



Département Informatique et Mathématiques Appliquées

Projet Long 2008

ADMINISTRATION AUTONOME DE SERVEURS SUR LA GRILLE AVEC UNE MACHINE VIRTUELLE

Spécifications

Responsable : Daniel Hagimont - Professeur INPT/ENSEEIHT - Daniel.Hagimont@enseeiht.fr

Co-encadrant : Laurent Broto – Etudiant en thèse à l'UPS - Laurent.Broto@irit.fr

Superviseur industriel : Emmanuel Murzeau - emmanuel.murzeau@airbus.com

Chef de projet : Ezequiel Geremia - ezequiel.geremia@etu.enseeiht.fr

Etudiants :

- Julien Louisy
- Julien Clariond
- Hery Randriamanamihaga
- Ezequiel Geremia
- Mathieu Giorgino

Sommaire

1 - Contexte général.....	3
2 - Analyse des besoins.....	5
2.1 - Objectif.....	5
2.2 - Besoins.....	5
2.2.1 - L'outil de paravirtualisation Xen.....	5
2.2.2 - Tune, gestionnaire de composants de type Fractal.....	5
2.2.3 - Mesure des performances.....	5
2.2.4 - Politique de distribution de machines virtuelles.....	6
3 - Exigences.....	6
3.1 - Exigences de réalisation.....	6
3.2 - Exigences fonctionnelles.....	9
3.2.1 - Propriétés relatives au mécanisme de migration.....	9
3.2.1.1 - Migration sans interruption d'activité.....	9
3.2.1.2 - Transparence.....	9
3.2.2 - Mesure des performances.....	10
3.2.3 - Intégration dans Tune.....	11
3.2.3.1 - Propriétés relatives au mécanisme de migration.....	11
3.2.3.2 - Administration autonome.....	12

1 - Contexte général

La virtualisation a pour but de permettre l'exécution de plusieurs systèmes d'exploitation sur une même machine physique. Le système d'exploitation installé sur la machine physique est appelé hôte. Celui-ci émule des machines qualifiées de virtuelles. Les machines virtuelles peuvent à leur tour accueillir des systèmes d'exploitation dans le but d'opérer des traitements particuliers.

Le cadre d'utilisation des outils de virtualisation est illustré dans l'exemple qui suit. Plus généralement, il s'agira de minimiser un certain critère (mesuré selon une métrique bien définie) grâce au mécanisme de migration. La migration consiste à déplacer un composant logiciel de la machine physique sur laquelle il s'exécute à une autre à travers un réseau (cf. *Figure 1.1*). Si le composant en question est un système d'exploitation, alors il est déplacé avec son contexte d'exécution, mais sans son système de fichiers.

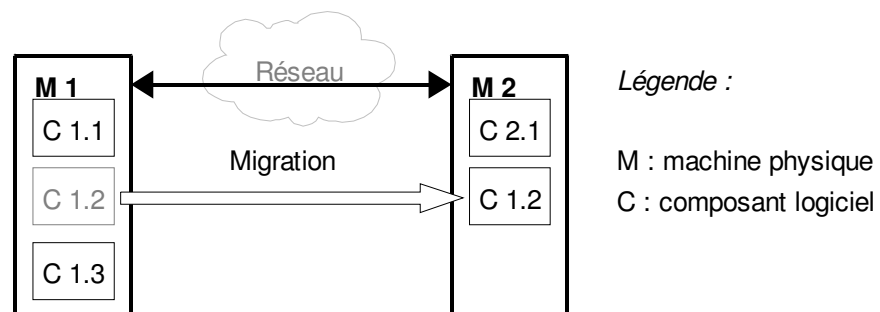


Figure 1.1 : Schéma illustrant le principe de la migration

Les grilles d'ordinateurs sont capables de traiter un volume important d'information correspondant à des calculs de simulation ou à des requêtes de clients vers des serveurs web. Certains hébergeurs proposent donc de louer des sous-ensembles de ces grilles. L'hébergeur doit donc faire face à la problématique suivante: comment maximiser le gain financier issu de la location de serveurs ?

Pour cela, il peut diminuer le coût d'exploitation de la grille, assimilé ici au coût de la consommation électrique. Cela se traduit notamment par l'extinction des machines physiques faiblement chargées, i.e. dont la charge processeur est sous un certain seuil minimum. Cependant, l'application qui s'exécute sur la machine à éteindre doit être déplacée avec son contexte d'exécution vers un autre serveur qui peut la supporter. Si l'ensemble des applications n'est plus supporté par les machines physiques allumées, alors d'autres machines physiques peuvent être mises sous tension afin d'accueillir les applications à forte consommation processeur.

L'hébergeur peut également maximiser le taux d'occupation de la grille, i.e. maximiser la charge processeur sur l'ensemble des machines physiques. Il s'agit, d'une part, de concentrer les applications à faible consommation CPU sur un nombre réduit de machines physiques et d'autre part d'allouer dynamiquement des machines physiques à des applications fortement consommatrices en CPU. Ceci permet d'augmenter le nombre de machines physiques disponibles et, de ce fait, d'augmenter la capacité d'accueil de nouvelles applications. Il en résulte qu'il est possible de louer plus de machines physiques qu'il

n'en existe sur la grille, tout en assurant un fonctionnement global optimal. En effet, la période de pointe, en terme de volume d'information à traiter, pour certaines applications, coïncident avec la période creuse pour d'autres et vice versa.

Enfin, si la grille est répartie sur des sites spatialement distants, alors le déplacement d'applications qui communiquent entre-elles sur des points plus proches du réseau est une solution qui diminue le temps de communication inter-applicatif.

Ces trois politiques, résumées dans la *Figure 1.2*, peuvent être mises en oeuvre grâce aux outils de virtualisation. Il s'agit de mettre en place une architecture de base capable de fournir des fonctionnalités élémentaires.

Contexte	parc de serveurs payants	parc de serveurs à haute consommation électrique	parc de serveurs distribué sur des sites distants
Métrique	prix du temps d'occupation par serveur et par seconde	ampérage électrique de l'alimentation par serveur	distance de la route séparant deux serveurs et charge réseau
But	minimiser le coût = concentrer les tâches à faible charge processeur	minimiser la consommation électrique = éteindre les machines physiques à faible charge processeur (sans perte de tâche)	minimiser le temps de communication = regrouper les tâches sur un réseau local

Figure 7. Tableau récapitulatif de la politique à mettre en oeuvre en fonction du contexte

2 - Analyse des besoins

2.1 - Objectif

L'objectif est triple : la maîtrise de la technologie de paravirtualisation Xen, son intégration dans le gestionnaire de composants Tune en vue d'une administration autonome et l'implantation d'une politique de distribution des machines virtuelles en fonction de la charge processeur.

2.2 - Besoins

2.2.1 - L'outil de paravirtualisation Xen

L'outil de paravirtualisation Xen est un outil de virtualisation dont les machines virtuelles sont des systèmes d'exploitation portés pour Xen. Ceci permet d'offrir des performances supérieures aux virtualiseurs qui ne font qu'émuler les machines virtuelles.

B1	<i>Énoncé du besoin</i> L'outil de virtualisation Xen sera déployé sur un système d'exploitation GNU/Linux.
-----------	---

2.2.2 - Tune, gestionnaire de composants de type Fractal

Tune permet l'administration de haut niveau de composants de type Fractal. Afin d'intégrer l'architecture associée au projet à Tune, chaque système d'exploitation s'exécutant sur les machines virtuelles devra être encapsulé dans un composant Fractal.

B2	<i>Énoncé du besoin</i> Le gestionnaire de composants est Tune.
-----------	---

2.2.3 - Mesure des performances

Un outil de mesure des performances devra être développé pour démontrer la fiabilité des méthodes d'installation, de configuration et d'utilisation des différentes technologies. Il offrira une mesure de la charge processeur des machines physiques et du temps d'exécution des applications s'exécutant sur les machines virtuelles.

B3	<i>Énoncé du besoin</i> Un outil de mesure des performances sera associé à l'architecture.
-----------	--

2.2.4 - Politique de distribution de machines virtuelles

L'architecture intégrée devra implanter une politique de distribution des machines virtuelles en fonction de leur charge processeur. Cette politique permettra d'implanter une gestion autonome des machines virtuelles dans l'environnement Tune.

B4	<i>Énoncé du besoin</i>	Une politique de distribution de charge sera implantée dans le cadre de l'administration autonome de machines virtuelles avec Tune.
-----------	-------------------------	---

3 - Exigences

3.1 - Exigences de réalisation

ER1	<i>Énoncé de l'exigence</i>	Le système d'exploitation utilisé sur les machines physiques devra être de type GNU/Linux.
	<i>Rationnel</i>	C'est une exigence du client. La stabilité est une propriété essentielle des systèmes d'exploitation hébergeant des machines virtuelles.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B1

ER2	<i>Énoncé de l'exigence</i>	Xen sera l'outil de virtualisation utilisé.
	<i>Rationnel</i>	C'est une exigence du client. Xen est un outil de paravirtualisation. Il offre, à ce titre, des performances supérieures aux outils d'émulation et des fonctionnalités de haut niveau, notamment pour effectuer des migrations.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B1

ER3	<i>Énoncé de l'exigence</i>	Tune sera le gestionnaire de composants.
	<i>Rationnel</i>	C'est une exigence du client. Tune offre des fonctionnalités de haut niveau pour l'administration autonome de composants.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B2

ER4	<i>Énoncé de l'exigence</i>	Les machines virtuelles devront être encapsulées dans des composants de type Fractal.
	<i>Rationnel</i>	La spécification Fractal fournit la possibilité d'interfacer des composants logiciels de manière modulaire. Tune permet l'administration des composants de type Fractal.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B2

ER5	<i>Énoncé de l'exigence</i>	Un diagramme de reconfiguration, propre à l'environnement, contrôlera la migration de machines virtuelles.
	<i>Rationnel</i>	Les diagrammes de reconfiguration seront utilisés pour instancier les politiques de distribution de charge processeur choisies, en vue d'une administration au sein de Tune.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B2

ER7	<i>Énoncé de l'exigence</i>	Les machines physiques seront connectées par un réseau
	<i>Rationnel</i>	Les machines physiques devront être capables de communiquer afin de permettre la migration des machines virtuelles qu'elles hébergent.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B4

ER7	<i>Énoncé de l'exigence</i>	Le protocole NFS (Network File System) sera utilisé pour l'échange de fichiers sur le réseau.
	<i>Rationnel</i>	Les machines virtuelles sont migrées sans leur système de fichiers, l'accès à des fichiers distants est donc inévitables.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B4

ER8	<i>Énoncé de l'exigence</i>	Dans l'environnement Tune, une sonde sera créée pour prélever la charge processeur des machines physiques sur lesquelles s'exécutent les systèmes d'exploitation hôtes.
	<i>Rationnel</i>	L'observation des paramètres d'évolution des composants est indispensable dans le cadre d'administrateurs de composants autonomes car ceux-ci doivent réagir, selon une certaine politique à ces événements.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2

ER9	<i>Énoncé de l'exigence</i>	La valeur élémentaire de la charge processeur sera évaluée au moins toutes les 100 ms.
	<i>Rationnel</i>	La charge processeur évolue rapidement, il faut donc en rendre compte par des mesures régulières sans l'affecter.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2

ER10	<i>Énoncé de l'exigence</i>	La valeur moyenne de la charge processeur sera calculée sur la base de 10 valeur élémentaires.
	<i>Rationnel</i>	La valeur moyenne permet de régulariser la mesure et supprimer le bruit dû à des pics d'activité. Dans le cadre de la migration de machines virtuelles, l'interprétation de ces mesures est d'autant plus pertinente qu'elle rend compte de tendances étalées dans le temps.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2

ER11	<i>Énoncé de l'exigence</i>	La migration d'une machine virtuelle sera déclenchée lorsque la valeur moyenne de la charge de la machine physique, sur laquelle s'exécute le système d'exploitation hôte, dépassera 70%.
	<i>Rationnel</i>	La valeur seuil de la charge processeur moyenne doit être suffisamment grande pour justifier la migration d'une machine virtuelle mais doit rester raisonnable pour garantir une migration rapide.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2

ER12	<i>Énoncé de l'exigence</i>	La migration des machines virtuelles les plus consommatrices en temps processeur seront déclenchées en premier.
	<i>Rationnel</i>	La politique de distribution consiste minimiser le nombre de migration vers d'autres systèmes hôtes afin d'en garder un maximum sur la machine d'origine.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2

3.2 - Exigences fonctionnelles

3.2.1 - Propriétés relatives au mécanisme de migration

3.2.1.1 - Migration sans interruption d'activité

EF1	<i>Énoncé de l'exigence</i>	La commande xm migrate --live <nom de VM> <machine destination> sera utilisée pour effectuer des migrations sans interruption d'activité.
	<i>Rationnel</i>	La commande est fournie par le système d'exploitation hôte utilisant l'outil de paravirtualisation Xen.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B1, B4
	<i>Référence du test</i>	T1

3.2.1.2 - Transparence

EF2	<i>Énoncé de l'exigence</i>	Les connections TCPdes machines virtuelles devront être maintenues lors d'une migration.
	<i>Rationnel</i>	Dans le cas de l'exécution de serveurs web sur les machines virtuelles, les connections des clients ne doivent pas être perdues.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B4
	<i>Référence du test</i>	T2

EF3	<i>Énoncé de l'exigence</i>	La migration d'une machine virtuelle n'affectera pas le résultat des applications qui s'y exécutent.
	<i>Rationnel</i>	La migration doit conserver l'état de la machine virtuelle et le contexte d'exécution des applications qu'elle héberge.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B4
	<i>Référence du test</i>	T3

EF4	<i>Énoncé de l'exigence</i>	La migration de machine virtuelle sera effectuée quelque soit la charge processeur de la machine physique contenant le système d'exploitation hôte.
	<i>Rationnel</i>	La migration de machines virtuelles est effectuée pour diminuer la charge processeur de la machine physique contenant le système d'exploitation hôte.
	<i>Priorité</i>	P0
	<i>Référence du besoin</i>	B4
	<i>Référence du test</i>	T4

3.2.2 - Mesure des performances

EF5	<i>Énoncé de l'exigence</i>	La durée de la migration d'une machine virtuelle devra être quantifiée.
	<i>Rationnel</i>	Une étude quantitative des performances permettra de mettre en place une configuration optimale.
	<i>Priorité</i>	P1
	<i>Référence du besoin</i>	B3
	<i>Référence du test</i>	T5

EF6	<i>Énoncé de l'exigence</i>	Le délai introduit par la migration devra être quantifiée.
	<i>Rationnel</i>	Une étude quantitative des performances permettra de mettre en place une configuration optimale.
	<i>Priorité</i>	P1
	<i>Référence du besoin</i>	B3
	<i>Référence du test</i>	T6

EF7	<i>Énoncé de l'exigence</i>	La durée de l'interruption de service lors de la migration d'une machine virtuelle devra être quantifiée.
	<i>Rationnel</i>	Une étude quantitative des performances permettra de mettre en place une configuration optimale.
	<i>Priorité</i>	P1
	<i>Référence du besoin</i>	B3
	<i>Référence du test</i>	T7

3.2.3 - Intégration dans Tune

3.2.3.1 - Propriétés relatives au mécanisme de migration

EF8	<i>Énoncé de l'exigence</i>	Toute migration dans l'environnement Tune devra satisfaire à l'exigence EF2.
	<i>Rationnel</i>	L'environnement d'administration de composants Tune ne doit pas dégrader les propriétés de la migration de machines virtuelles.
	<i>Priorité</i>	P1
	<i>Référence du besoin</i>	B2
	<i>Référence du test</i>	T8

EF9	<i>Énoncé de l'exigence</i>	Toute migration dans l'environnement Tune devra satisfaire à l'exigence EF3.
	<i>Rationnel</i>	L'environnement d'administration de composants Tune ne doit pas dégrader les propriétés de la migration de machines virtuelles.
	<i>Priorité</i>	P1
	<i>Référence du besoin</i>	B2
	<i>Référence du test</i>	T9

EF10	<i>Énoncé de l'exigence</i>	Dans l'environnement Tune, toute migration devra satisfaire à l'exigence EF4.
	<i>Rationnel</i>	L'environnement d'administration de composants Tune ne doit pas dégrader les propriétés de la migration de machines virtuelles.
	<i>Priorité</i>	P1
	<i>Référence du besoin</i>	B2
	<i>Référence du test</i>	T10

EF11	<i>Énoncé de l'exigence</i>	Dans l'environnement Tune, le déclenchement de la migration devra être automatique.
	<i>Rationnel</i>	Le déclenchement automatique de la migration de machines virtuelles joue un rôle important dans le caractère autonome de l'administration dans l'environnement Tune.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2
	<i>Référence du test</i>	T11

3.2.3.2 - Administration autonome

EF12	<i>Énoncé de l'exigence</i>	La migration ne s'effectuera que vers un système hôte ayant une capacité processeur suffisante pour accueillir une nouvelle machine virtuelle.
	<i>Rationnel</i>	Dans le cas où la charge processeur de toutes les machines physiques dépasse 70% alors aucune migration ne sera déclenchée.
	<i>Priorité</i>	P2
	<i>Référence du besoin</i>	B2
	<i>Référence du test</i>	T12