



Infrastructure pour le calcul de haute performance

Exercices pour le Module 11

INF8601 Systèmes informatiques parallèles

Michel Dagenais

École Polytechnique de Montréal
Département de génie informatique et génie logiciel

La virtualisation pour le Calcul de Haute Performance

Pourquoi la virtualisation est-elle parfois envisagée dans des environnements de calcul de haute performance?



La virtualisation pour le Calcul de Haute Performance

Pourquoi la virtualisation est-elle parfois envisagée dans des environnements de calcul de haute performance?

Un centre de calcul de haute performance, comme Calcul Québec, dessert de très nombreux usagers qui utilisent divers environnements de développement (Linux, Windows), un grand nombre de logiciels différents, et possiblement diverses versions de bibliothèques qui sont parfois incompatibles. Il est souvent difficile pour les usagers de s'adapter à l'environnement pré-installé sur la grappe. Avec la virtualisation, l'application vient sous la forme d'une image complète avec son propre environnement logiciel. Il y a cependant un coût en performance.



Virtualisation, paravirtualisation et conteneurs

Quels sont les avantages et inconvénients des conteneurs par rapport à la paravirtualisation et aux machines virtuelles?



Virtualisation, paravirtualisation et conteneurs

Quels sont les avantages et inconvénients des conteneurs par rapport à la paravirtualisation et aux machines virtuelles?

La virtualisation permet une flexibilité à peu près complète, mais occasionne un surcoût qui peut atteindre 20 ou 30%. La paravirtualisation requiert un système d'exploitation configuré de manière particulière (e.g. pilotes d'interface paravirtualisés), mais vient avec un surcoût réduit. Les conteneurs ont un surcoût presque nul et permettent toute flexibilité au niveau usager (programmes, bibliothèques), mais le système d'exploitation est imposé.



Tables de pages étendues

La technologie EPT (Extended Page Table) de Intel permet de diminuer le surcoût de la virtualisation en moyenne. Donnez un exemple où la technologie peut accélérer une application et une où l'application peut au contraire être ralentie.



Tables de pages étendues

La technologie EPT (Extended Page Table) de Intel permet de diminuer le surcoût de la virtualisation en moyenne. Donnez un exemple où la technologie peut accélérer une application et une où l'application peut au contraire être ralentie.

Sans la technologie EPT, il faut intercepter toutes les modifications aux tables de pages dans la machine virtuelle, pour les refléter dans les tables de pages sur l'hôte. Avec la technologie EPT, les tables de pages de la machine virtuelle seront connectées aux tables de pages de la machine hôte. Cependant, lors d'une faute de TLB, il faut traverser deux tables au lieu d'une seule pour faire la traduction. Ainsi, un programme qui fait beaucoup de modifications à sa table de pages (e.g. beaucoup d'appels mmap) sera plus rapide avec EPT. Un programme qui fait beaucoup de fautes de TLB (accède beaucoup de pages différentes), sera plus lent.

Gestionnaires de tâches

Un laboratoire comme ceux du département GIGL permettent l'accès à distance des noeuds par ssh. Pourquoi les grappes de calcul sont-elles dotées d'un gestionnaire de tâches comme SLURM, au lieu de simplement utiliser ssh?



Gestionnaires de tâches

Un laboratoire comme ceux du département GIGL permettent l'accès à distance des noeuds par ssh. Pourquoi les grappes de calcul sont-elles dotées d'un gestionnaire de tâches comme SLURM, au lieu de simplement utiliser ssh?

Si tous se connectent à distance sans coordination, on peut se retrouver avec trop d'utilisateurs en même temps, ou avec des noeuds comptant trop d'utilisateurs alors que d'autres sont vides. Un gestionnaire de tâches permet de répartir les tâches à exécuter entre les noeuds et dans le temps, tout en tenant compte des priorités et quotas. SLURM permet aussi de mettre en queue des tâches interactives, ce qui réserve des noeuds qui seront accessibles en ssh.

