



# Virtualization avec Xen

*Daniel Veillard*

*veillard@redhat.com*



# Présentation Générale de Xen

- **Projet de Recherche, Cambridge UK**
- **Paravirtualization de Windows et Linux**
- **Licence GPL**
- **Présentation a OLS 2004**
- **Performances intéressantes**
- **Virtualization hardware**



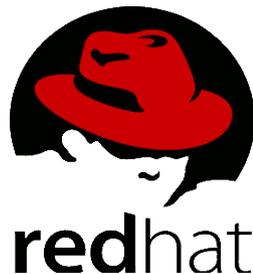
# Acteurs autour du projet Xen

- **Groupe de Recherche**
- **Xensource**
- **Communauté d'utilisateurs**
- **Distributeurs Linux (IBM, SuSE, Red Hat, ...)**
- **Constructeurs Intel/AMD, Dell, Fujitsu, ...**
- **Microsoft**



# La virtualisation : définition

- **Découpler les ressources informatique de leur implémentation**
- **Machine Virtuelle**
  - Processeur virtuels
  - Mémoire
  - Périphériques réels ou émulés (réseau, vidéo, disques)
- **Stockage Virtualisé**
  - Disque distants
  - Disques locaux émulés
- **Transition du multitâche au multi-OS**



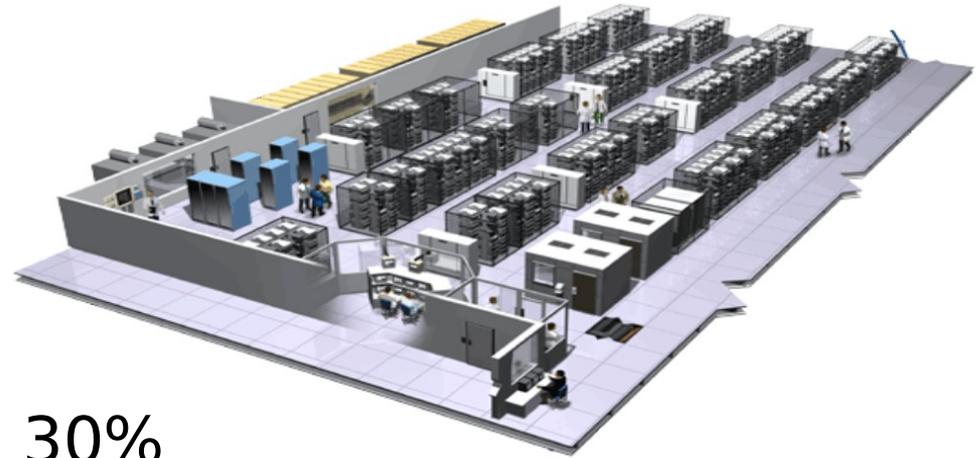
# Bénéfices de la virtualisation

- **Les bénéfices sont de deux ordres :**
- **Réduction des coûts**
  - Meilleure gestion de la capacité
  - Simplification de la maintenance
  - Mise à l'échelle
- **Réduction des risques**
  - Isolation et sécurité des services
  - Indépendance matérielle
  - Redondance aisée

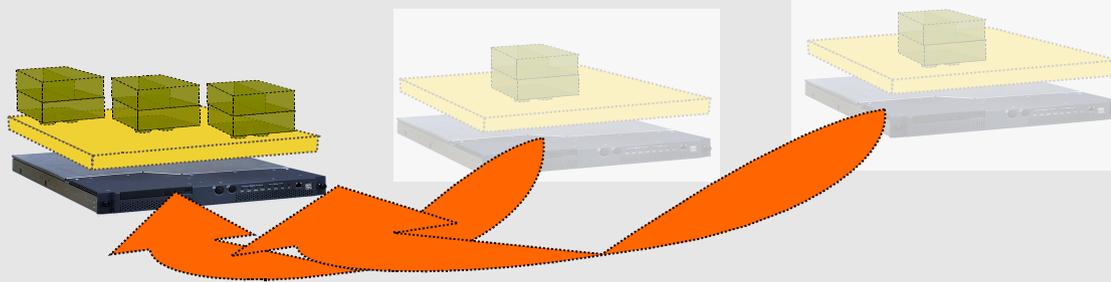


# Exemple: salle machine

- **Contraintes physique**
  - Taille, energie, clim.
- **Surdimentionés**
  - Utilisation moyenne de 10 à 30%
- **Un serveur par application/client**
  - Simplification de gestion, facturation
  - Limite les risques
- **Problèmes**
  - Coûts s'envolent
  - Administration du parc



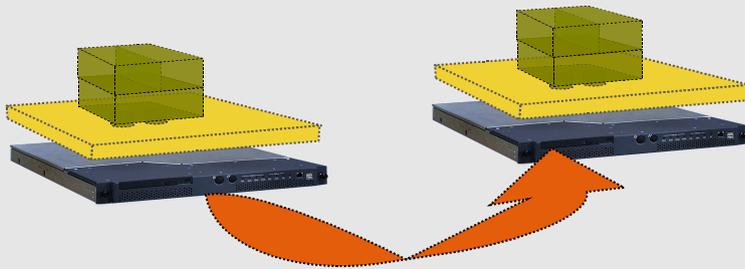
# La virtualisation en action



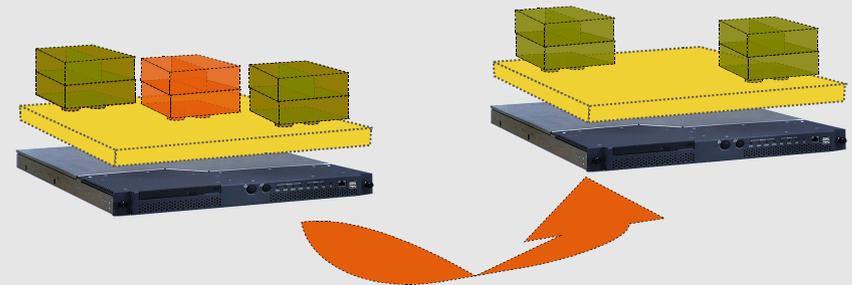
Consolidation de serveurs sous-utilisés



Politique de  
Sécurisation



Pas d'interruption de service,  
migration



Répartition dynamique de  
charge

# Techniques de virtualisation

- **Problème**
  - Le jeu d'instruction du i386 n'est pas virtualisable
- **Émulation complète (Qemu)**
  - Contrôle total, mais lent
- **Réécriture du code au vol (VMWare)**
  - Plus rapide mais très complexe
- **Paravirtualisation**
  - Recompilation pour la plateforme virtuelle, idéal
- **Virtualisation matérielle**
  - Nouveaux CPUs, reste complexe



# Architecture Générale

- **Hyperviseur**
- **Domaine 0, OS hôte et pilotes de périphérique**
- **Domaines U, OS invité, pilotes paravirtuels**
- **Bus entre domaines**
- **Appels hyperviseurs**
- **XenStore base de données**
- **Xend démon**



# Architecture de Xen

## Domaine0

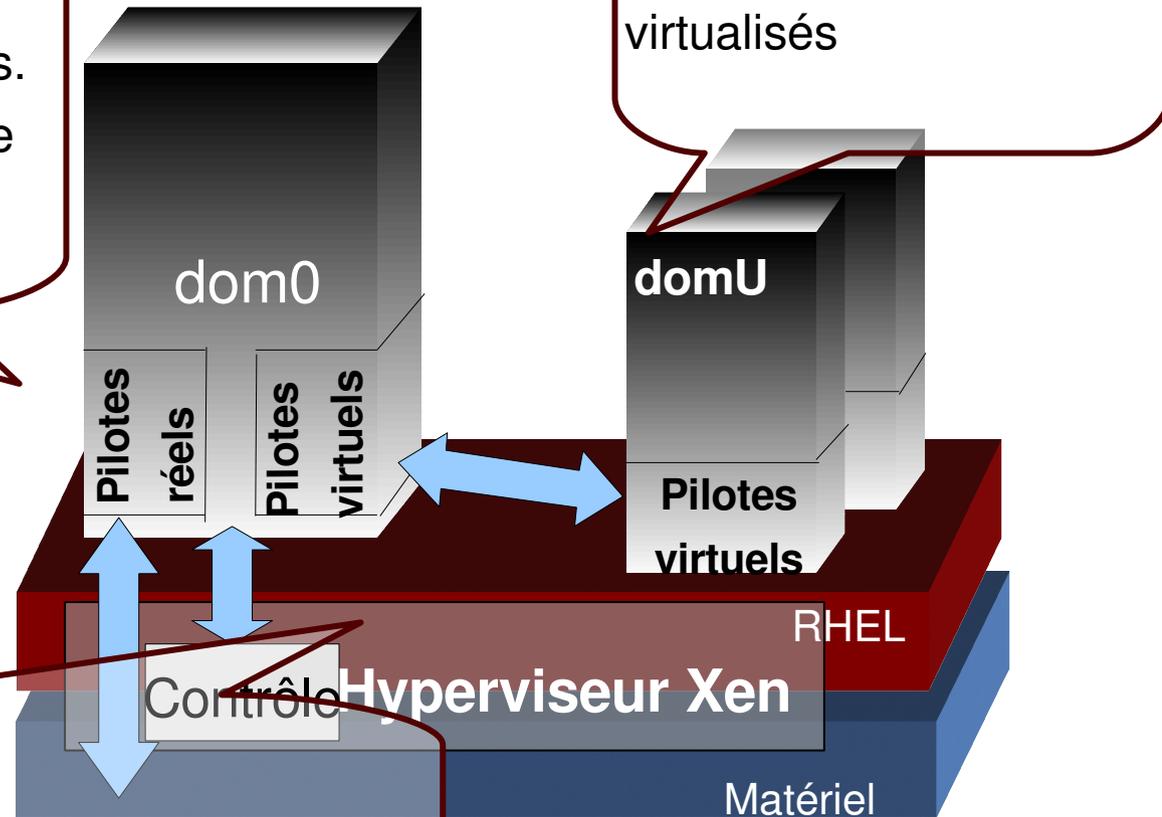
Privilégié, gère les périphériques.  
Fournit les interfaces de contrôle  
et outils de gestion

## Domaine non privilégié:

Machine Virtuelle  
CPU, Memoire, Stockage  
virtualisés

## Hyperviseur Xen

routage IRQ, séquenceur , and communications entre  
domaines. Implémente le partage transparent des ressources  
avec les pilotes réels et limitations d'usage



# Architecture logicielle

- **Hyperviseur privilégié Ring 0**
- **Domaine 0 noyau Ring 1/3**
- **Domaines U noyaux Ring 3**
- **Patches pour le noyau linux (0 et U)**
- **Problème: taille des patches, intrusifs**



# Appels Hyperviseurs

- **Très bas niveau**
- **Noyaux Dom0/DomU trappe logicielles**
- **Usermode Dom0 via ioctl()**
- **libxc GPL'ed**
- **Encapsulation python**
- **Problème: instabilité des appels, très spécifiques**
- **Upstream: paravirt\_ops**



# Le Bus Xen

- **Communication entre Dom0, domU via hyperviseur**
- **Sécurité**
- **Prévention des famines**



# Le XenStore

- **Base de donnée**
- **accès concurrent Dom0, DomU et Hyperviseur**
- **Informations partagées ou privées**
- **Démon dans Domaine 0**
- **snooping**

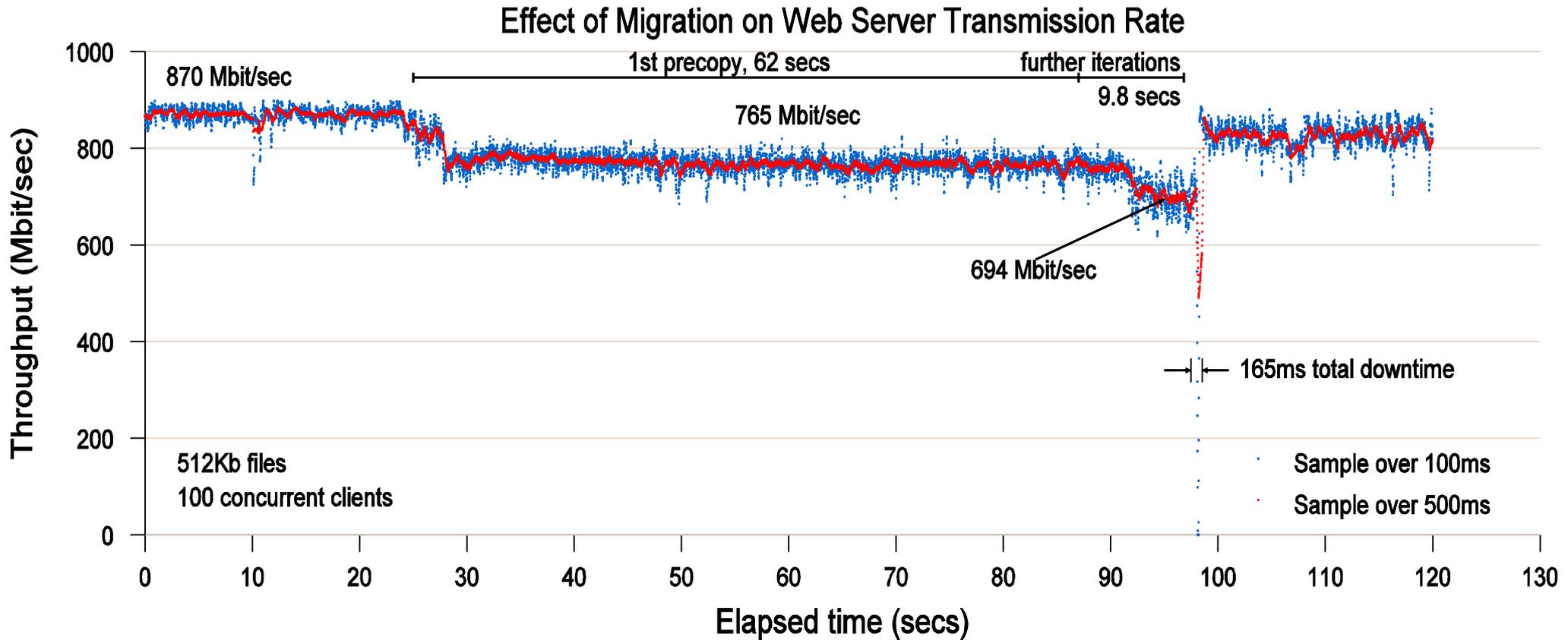


# Ca fonctionne !

- **Cycle de vie classique de la machine virtuelle**
- **Sauvegarde/restauration de l'image mémoire**
- **Migration a froid et a chaud**
- **Redimensionnement memoire et CPUs a chaud**
- **Pilotes de périphériques complets via Linux**
- **Bon support i386/x86\_64 , ia64 et ppc64 en progrès**
- **Robuste, beaucoup de tests**
- **Isolation correcte mais pas absolue**



# Migration a chaud d'un serveur web



Source: présentation Xen à OLS 2006



# Performances

- **Très bonnes en general**
- **Classiquement 0-10% de dégradation/machine physique**
- **Mais uniquement en paravirtualization !**



# Virtualization complète

- **Virtualisation matérielle Intel et AMD**
- **émulation matérielle basée sur QEmu**
- **Pas de sauvegarde, pas de migration**
- **Récent, moins optimisé**
- **Support actif de AMD et Intel**
- **Ca fonctionne, en général**



# Problèmes !

## ■ **Architecturaux**

- code neuf
- complexité
- HVM incomplet
- isolation imparfaite

## ■ **Logiciel**

- integration upstream, API changeantes
- tres gros patch, 2.6.16
- implementation xend et xenstore
- conditions limite
- processeur PPC64



# Les petits nouveaux

- **KVM et lguest**
- **Virtualization complète dans le noyau Linux 2.6.20**
- **Linux est l'hyperviseur**
- **KVM utilise aussi Qemu**
- **moins de code, mais très récent**



# Perspectives

- **Séparation Domaine 0 (Hurd, OSF-1/MK, ...)**
- **Amélioration du support HVM**
- **Intégration 'upstream' pour Xen et VMWare**
- **Amélioration des pilotes et de l'émulation**
- **Place centrale de Qemu**
- **Nombreux modèles de virtualisation**
- **Besoins variés**



# Libvirt



- **API stable pour la virtualisation**

- Projet Open Source <http://libvirt.org/>
- API générique (Xen, QEmu, KVM)
- Librairie LGPL

- **Opérations de base**

- Créer/détruire/suspendre/sauver
- Description XML des machines virtuelles

- **Caractéristiques**

- Gestion d'erreurs
- Accès lecture seule possible pour l'utilisateur
- Gestion des anciennes versions Xen
- Écrite en C, bindings Python et Perl

