

Computação em nuvem (Mais) Virtualização

Gabriel V C Candido

Instituto Federal do Paraná - Pinhais

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

Definição formal

Duplicata eficiente e isolada de uma máquina real

Definição formal

Duplicata eficiente e isolada de uma máquina real

Equivalência

- ▶ ambiente de execução quase idêntico ao real

Definição formal

Duplicata eficiente e isolada de uma máquina real

Equivalência

- ▶ ambiente de execução quase idêntico ao real

Controle de recursos

- ▶ hipervisor tem acesso aos recursos do sistema real
- ▶ sistema virtual só acessa o que o hipervisor alocar para ele
- ▶ hipervisor intermedeia os acessos aos recursos reais



Definição formal

Duplicata eficiente e isolada de uma máquina real

Equivalência

- ▶ ambiente de execução quase idêntico ao real

Controle de recursos

- ▶ hipervisor tem acesso aos recursos do sistema real
- ▶ sistema virtual só acessa o que o hipervisor alocar para ele
- ▶ hipervisor intermedeia os acessos aos recursos reais

Eficiência

- ▶ queremos um bom desempenho
- ▶ executar a maior parte das coisas diretamente pelo processador real
- ▶ o que não puder ser executado diretamente deve ser interpretado pelo hipervisor e traduzido

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

O que precisamos emular?

- ▶ registradores
- ▶ memória (estado)
- ▶ tabela de páginas; segmentos
- ▶ interrupções
- ▶ relógio
- ▶ barramentos
- ▶ BIOS
- ▶ disco
- ▶ interfaces de rede
- ▶ ...

Suporte de hardware

Intel Pentium IV (x86 antigos) não era virtualizável

Suporte de hardware

Intel Pentium IV (x86 antigos) não era virtualizável

- ▶ instruções sensíveis podiam ser executadas em modo usuário sem gerar exceções

Suporte de hardware

Intel Pentium IV (x86 antigos) não era virtualizável

- ▶ instruções sensíveis podiam ser executadas em modo usuário sem gerar exceções

Soluções:

- ▶ tradução dinâmica
- ▶ alterar o SO
- ▶ tornar todas as instruções sensíveis privilegiadas (suporte em hardware)

Anéis de proteção



Figura: Níveis de proteção em processadores AMD/Intel. Fonte: CM 33

Anéis de proteção



Figura: Níveis de proteção em processadores AMD/Intel. Fonte: CM 33

Precisa alterar o sistema convidado

Anéis de proteção



Figura: Níveis de proteção em processadores AMD/Intel. Fonte: CM 33

Precisa alterar o sistema convidado: solução atual consiste em criar um modo de virtualização que tem os 4 níveis

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

Técnicas

Emulação completa

- ▶ instruções
- ▶ memória
- ▶ periféricos
- ▶ *alto custo; alta flexibilidade*
- ▶ *QEMU*

Emulação completa

- ▶ instruções
- ▶ memória
- ▶ periféricos
- ▶ *alto custo; alta flexibilidade*
- ▶ *QEMU*

Virtualização da interface de sistema

- ▶ somente instruções privilegiadas
- ▶ dispositivos
- ▶ *sistema convidado enxerga o processador real*
- ▶ *baixo custo; baixa flexibilidade*
- ▶ *VMware Workstation, VirtualBox, KVM*

Técnicas

Tradução dinâmica

- ▶ desmontar, reorganizar e traduzir o binário
- ▶ otimização
- ▶ código intermediário
- ▶ *Java*

Tradução dinâmica

- ▶ desmontar, reorganizar e traduzir o binário
- ▶ otimização
- ▶ código intermediário
- ▶ *Java*

Paravirtualização

- ▶ hardware similar, mas não idêntico ao real
- ▶ convidado acessa parte do hardware diretamente
- ▶ *baixo custo; adaptação do sistema convidado*
- ▶ *Xen*

Paravirtualização

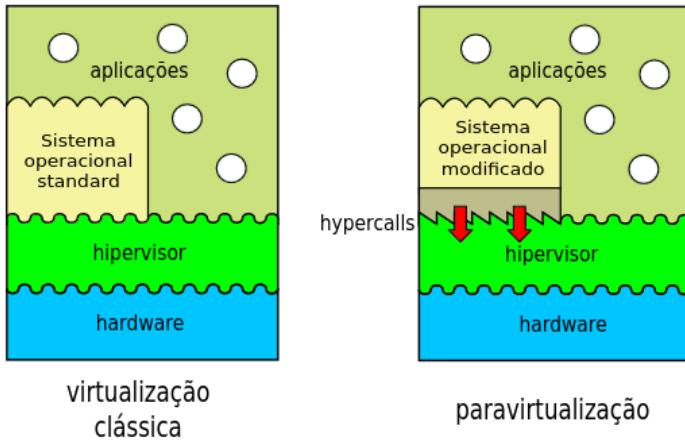


Figura: Paravirtualização. Fonte: CM 33

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

Desempenho

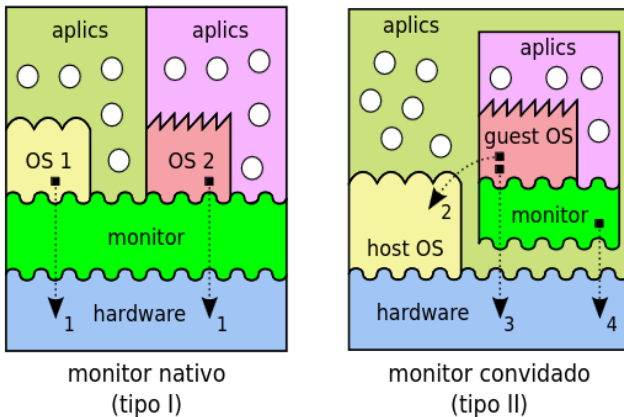


Figura: Otimizações em máquinas virtuais. Fonte: CM 33

Desempenho

Acessos diretos, que “pulam” camadas/etapas:

1. SO convidado acessa o hardware nativo: precisa de suporte do SO convidado e no hipervisor
 - ▶ gerência de memória do Xen
2. SO convidado acessa o SO nativo: precisa de suporte do hipervisor
 - ▶ VMware usa o sistema de arquivos nativo
3. SO convidado acessa o hardware nativo: precisa de suporte do SO nativo e do hipervisor
 - ▶ VMware usa essa estratégia para drive de CD, interface de rede e acesso à placa de vídeo
4. hipervisor acessa o hardware nativo: precisa de suporte do SO nativo
 - ▶ VMmware

Desempenho

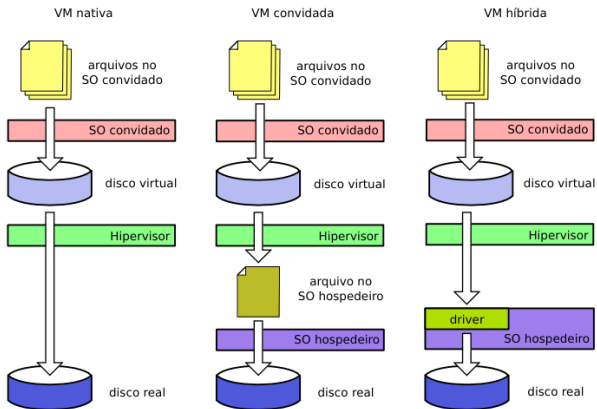


Figura: Desempenho de hipervisores. Fonte: CM 33

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

Migração

Transferir uma máquina virtual de um hipervisor para outro através da rede

Migração

Transferir uma máquina virtual de um hipervisor para outro através da rede

- ▶ *Stop-and-copy*: pausa a máquina, faz a cópia e inicia novamente no novo *host* (pode demorar minutos)

Migração

Transferir uma máquina virtual de um hipervisor para outro através da rede

- ▶ *Stop-and-copy*: pausa a máquina, faz a cópia e inicia novamente no novo *host* (pode demorar minutos)
- ▶ *Demand-migration*: migra inicialmente só o núcleo convidado e depois migra o que é acessado, sob demanda (demora segundos, mas a migração continua ocorrendo por muito tempo até migrar tudo)

Migração

Transferir uma máquina virtual de um hipervisor para outro através da rede

- ▶ *Stop-and-copy*: pausa a máquina, faz a cópia e inicia novamente no novo *host* (pode demorar minutos)
- ▶ *Demand-migration*: migra inicialmente só o núcleo convidado e depois migra o que é acessado, sob demanda (demora segundos, mas a migração continua ocorrendo por muito tempo até migrar tudo)
- ▶ *Pre-copy*: faz uma cópia de tudo sem parar a máquina virtual; quando terminar, pausa a máquina e copia de novo somente o que foi modificado (é o meio do caminho das duas anteriores)

Sumário

Definição formal

Suporte de hardware

Técnicas de virtualização

Desempenho

Migração

Virtualização na prática

Para quê serve?

- ▶ binários portáteis
- ▶ compartilhamento de hardware
- ▶ suporte a aplicações legadas
- ▶ experimentos em redes
- ▶ ensino
- ▶ segurança
- ▶ ...

- ▶ Usa tradução dinâmica

Dois modos de emulação:

- ▶ emulação total do sistema; ou
 - ▶ emulação no modo de usuário (Linux)
-
- ▶ VirtualBox é construído sobre o QEMU

- ▶ Embutido no kernel Linux
- ▶ Processadores virtuais e recursos paravirtualizados
- ▶ Gerenciamento de máquinas virtuais: LibVirt
- ▶ Sistemas convidados usam a biblioteca *VirtIO* para acessar os recursos virtualizados

Docker

- ▶ Framework para construção de contêineres
- ▶ Uma aplicação pode ser empacotada com todas as suas dependências em um contêiner
- ▶ Distribuição e implantação são agilizados
- ▶ No Linux usa o LXC ou *systemd*, *OverlayFS* e *cgroups*



- ▶ Depuração de memória através de tradução dinâmica

- ▶ Depuração de memória através de tradução dinâmica

Passo-a-passo:

1. ao carregar o programa, traduz o binário para uma representação intermediária (IR)
2. instrumenta o código em IR para verificar as operações de memória
3. traduz para um novo binário (para um processador virtual)

- ▶ Depuração de memória através de tradução dinâmica

Passo-a-passo:

1. ao carregar o programa, traduz o binário para uma representação intermediária (IR)
 2. instrumenta o código em IR para verificar as operações de memória
 3. traduz para um novo binário (para um processador virtual)
- ▶ Pode ser até 50 vezes mais lento!

- ▶ Depuração de memória através de tradução dinâmica

Passo-a-passo:

1. ao carregar o programa, traduz o binário para uma representação intermediária (IR)
 2. instrumenta o código em IR para verificar as operações de memória
 3. traduz para um novo binário (para um processador virtual)
-
- ▶ Pode ser até 50 vezes mais lento! No geral não estamos preocupados com isso quando estamos depurando programas (não usamos em produção)

User-mode Linux

- ▶ Cria um espaço de usuário Linux como um processo do próprio Linux nativo
- ▶ Isolamento
- ▶ Acesso aos recursos que o sistema nativo oferece
- ▶ Simples → baixo custo