

Virtualização de **Servidores**

Manoel Veras
Alexandre Carissimi



Virtualização **de Servidores**

Manoel Veras
Alexandre Carissimi

Revisão
Thiago Reis



Virtualização de Servidores

Manoel Veras

Alexandre Carissimi

Revisão

Thiago Reis

Rio de Janeiro

Escola Superior de Redes

2015

Copyright © 2015 – Rede Nacional de Ensino e Pesquisa – RNP
Rua Lauro Müller, 116 sala 1103
22290-906 Rio de Janeiro, RJ

Diretor Geral
Nelson Simões

Diretor de Serviços e Soluções
José Luiz Ribeiro Filho

Escola Superior de Redes

Coordenação
Luiz Coelho

Edição
Pedro Sangirardi

Revisão Técnica

Marcelo Castellan Braga

Coordenação Acadêmica de Administração de Sistemas
Sergio Alves de Souza

Equipe ESR (em ordem alfabética)
Celia Maciel, Cristiane Oliveira, Derlinéa Miranda, Edson Kowask, Elimária Barbosa, Lourdes Soncin, Luciana Batista, Luiz Carlos Lobato e Renato Duarte

Capa, projeto visual e diagramação
Tecnodesign

Versão
2.0.0-b

Este material didático foi elaborado com fins educacionais. Solicitamos que qualquer erro encontrado ou dúvida com relação ao material ou seu uso seja enviado para a equipe de elaboração de conteúdo da Escola Superior de Redes, no e-mail info@esr.rnp.br. A Rede Nacional de Ensino e Pesquisa e os autores não assumem qualquer responsabilidade por eventuais danos ou perdas, a pessoas ou bens, originados do uso deste material.
As marcas registradas mencionadas neste material pertencem aos respectivos titulares.

Distribuição
Escola Superior de Redes
Rua Lauro Müller, 116 – sala 1103
22290-906 Rio de Janeiro, RJ
http://esr.rnp.br
info@esr.rnp.br

Escola Superior de Redes

A Escola Superior de Redes (ESR) é a unidade da Rede Nacional de Ensino e Pesquisa (RNP) responsável pela disseminação do conhecimento em Tecnologias da Informação e Comunicação (TIC). A ESR nasce com a proposta de ser a formadora e disseminadora de competências em TIC para o corpo técnico-administrativo das universidades federais, escolas técnicas e unidades federais de pesquisa. Sua missão fundamental é realizar a capacitação técnica do corpo funcional das organizações usuárias da RNP, para o exercício de competências aplicáveis ao uso eficaz e eficiente das TIC.

A ESR oferece dezenas de cursos distribuídos nas áreas temáticas: Administração e Projeto de Redes, Administração de Sistemas, Segurança, Mídias de Suporte à Colaboração Digital e Governança de TI.

A ESR também participa de diversos projetos de interesse público, como a elaboração e execução de planos de capacitação para formação de multiplicadores para projetos educacionais como: formação no uso da conferência web para a Universidade Aberta do Brasil (UAB), formação do suporte técnico de laboratórios do Proinfo e criação de um conjunto de cartilhas sobre redes sem fio para o programa Um Computador por Aluno (UCA).

A metodologia da ESR

A filosofia pedagógica e a metodologia que orientam os cursos da ESR são baseadas na aprendizagem como construção do conhecimento por meio da resolução de problemas típicos da realidade do profissional em formação. Os resultados obtidos nos cursos de natureza teórico-prática são otimizados, pois o instrutor, auxiliado pelo material didático, atua não apenas como expositor de conceitos e informações, mas principalmente como orientador do aluno na execução de atividades contextualizadas nas situações do cotidiano profissional.

A aprendizagem é entendida como a resposta do aluno ao desafio de situações-problema semelhantes às encontradas na prática profissional, que são superadas por meio de análise, síntese, julgamento, pensamento crítico e construção de hipóteses para a resolução do problema, em abordagem orientada ao desenvolvimento de competências.

Dessa forma, o instrutor tem participação ativa e dialógica como orientador do aluno para as atividades em laboratório. Até mesmo a apresentação da teoria no início da sessão de aprendizagem não é considerada uma simples exposição de conceitos e informações. O instrutor busca incentivar a participação dos alunos continuamente.

As sessões de aprendizagem onde se dão a apresentação dos conteúdos e a realização das atividades práticas têm formato presencial e essencialmente prático, utilizando técnicas de estudo dirigido individual, trabalho em equipe e práticas orientadas para o contexto de atuação do futuro especialista que se pretende formar.

As sessões de aprendizagem desenvolvem-se em três etapas, com predominância de tempo para as atividades práticas, conforme descrição a seguir:

Primeira etapa: apresentação da teoria e esclarecimento de dúvidas (de 60 a 90 minutos).
O instrutor apresenta, de maneira sintética, os conceitos teóricos correspondentes ao tema da sessão de aprendizagem, com auxílio de slides em formato PowerPoint. O instrutor levanta questões sobre o conteúdo dos slides em vez de apenas apresentá-los, convidando a turma à reflexão e participação. Isso evita que as apresentações sejam monótonas e que o aluno se coloque em posição de passividade, o que reduziria a aprendizagem.

Segunda etapa: atividades práticas de aprendizagem (de 120 a 150 minutos).
Esta etapa é a essência dos cursos da ESR. A maioria das atividades dos cursos é assíncrona e realizada em duplas de alunos, que acompanham o ritmo do roteiro de atividades proposto no livro de apoio. Instrutor e monitor circulam entre as duplas para solucionar dúvidas e oferecer explicações complementares.

Terceira etapa: discussão das atividades realizadas (30 minutos).
O instrutor comenta cada atividade, apresentando uma das soluções possíveis para resolvê-la, devendo ater-se àquelas que geram maior dificuldade e polêmica. Os alunos são convidados a comentar as soluções encontradas e o instrutor retoma tópicos que tenham gerado dúvidas, estimulando a participação dos alunos. O instrutor sempre estimula os alunos a encontrarem soluções alternativas às sugeridas por ele e pelos colegas e, caso existam, a comentá-las.

Sumário

Escola Superior de Redes **v**

A metodologia da ESR *v*

Aspectos gerais da virtualização **1**

Introdução *1*

Tipos de virtualização *4*

Conceito e categorização *5*

Hipervisores *6*

Tipos de virtualização de servidores *7*

Tipos de virtualização 8

Virtualização completa 8

Paravirtualização 9

Virtualização assistida por hardware 10

Cenários da virtualização *11*

Tendências da virtualização *12*

Datacenter dinâmico 12

Computação em nuvem 13

Principais fornecedores de soluções de virtualização *14*

Licenciamento *15*

Gerenciamento e segurança com a virtualização *15*

Limitações da virtualização *15*

Desempenho e benchmarks *16*

Soluções de virtualização *16*

VMware ESX 16

Hyper-V 17

Xen 17

KVM 17

O KVM: origem *17*

Arquitetura 17

Gerenciamento de memória do Linux 18

Swapping 19

Balão de memória 19

Kernel Same-page Merging (KSM) 19

Live migration 19

Sistemas convidados suportados 20

Snapshots *20*

Projeto de virtualização **21**

Projeto de virtualização *21*
Educação e conscientização 22
Planejamento da capacidade 23
Coleta e levantamento 24
Análise 24
Prova de conceito 25
Avaliação de disponibilidade para virtualização 26
Ferramentas 26

Escolha do fornecedor *26*
Projeto 27
Implementação 27
Gerenciamento 28

Infraestrutura de hardware *28*
Arquitetura física do datacenter 29
Arquitetura virtual do datacenter 30

Hardware – descrição *31*
Servidores 31
Arquitetura do processador 32
Benchmark de servidores 33

Padrões da indústria *34*

Características dos servidores *35*
Disponibilidade 35
Desempenho 35
Gerenciamento 36
Escalabilidade 36

Dispositivos de armazenamento – storage *37*

Sistemas de armazenamento *37*
Servidores 37
Storage 38
Conectividade 38

Características do servidor de discos *38*

Tipos de storage *39*
JBOD 39
RAID 40
LUN 40
Unidade Lógica 40
Inteligente 41

Redes de storage *41*

Protocolo SCSI *43*

SAN *43*
SAN FC e FCoE 44
SAN IP 45

Storage e virtualização *46*

Infraestrutura de rede *46*

Virtualização da rede *47*

Benefícios da virtualização *48*
Redução do custo total de propriedade 48

Análise de TCO/ROI *49*
Total Cost of Ownership (TCO) 49
Return On Investment (ROI) 50

Alta disponibilidade *50*

Backup e restore *50*

Introdução ao Xen **53**

Histórico *53*

Versões *54*

Componentes *55*
Hipervisor Xen 56
Domínio convidado privilegiado 56
Domínio convidado não privilegiado 57

Virtualização no Xen *58*
Modos de operação do processador 58
Arquitetura dos drivers de dispositivo 59
Paravirtualização 60
Virtualização completa 61
Virtualização híbrida 61

XenMotion – migração de domínios convidados *62*

Outras formas de virtualização com o Xen *63*
XenDesktop 64
XenApp 64

Gerenciamento do Xen **67**

Infraestrutura de hardware para o Xen *67*
Processador e dispositivos de E/S 67
Memória RAM 68
Sistemas de armazenamento (storage) 69

Infraestrutura física de armazenamento 70
Logical Volume Management (LVM) 71
Migração e sistemas de armazenamento 72
Rede 73

Infraestrutura de software para o Xen *73*
Distribuições comerciais e livres 73
Hipervisor e sistemas operacionais para Dom0 74
Sistemas operacionais para DomU 74

Geração de máquinas virtuais paravirtualizadas *75*
Inicialização de sistemas Linux 75
Clonagem 76
Imagens preexistentes (appliances) 77
Ferramentas específicas às distribuições Linux 77

Qemu *78*

Gerenciamento de domínios convidados *78*
Comando de linha 78
Virt-manager 79
Xen Center 80

Arquitetura genérica de virtualização *81*

Introdução ao Hyper-V **85**

Características gerais *85*
Versões 86
Observa-se nos tópicos apresentados acima que apenas as versões Standard e Datacenter podem ser utilizados para infraestruturas de virtualização. 87
Benefícios 87
Licenciamento 88

Componentes da arquitetura Hyper-V *88*

Cenários de uso do Hyper-V *90*
Consolidação de servidores 90
Continuidade de negócios e recuperação de desastres 90
Desenvolvimento e teste 91
Datacenter dinâmico 91

Gerenciamento com Hyper-V *91*
Hyper-V Manager 91
System Center Virtual Machine Manager (SCVMM) 92
Cenários 92
Recursos do SCVMM 93

Snapshots com o Hyper-V *94*

Live Backup *95*

Avanços com o Windows Server 2012 R2 Hyper-V *96*
Armazenamento dinâmico de máquina virtual 96
Live Migration 96
Suporte de processador avançado 96
Suporte avançado de redes 97
Volumes compartilhados de cluster 97

Alta disponibilidade *98*
Guest Clustering 98
Host Clustering 99
Quick Migration 100
Live Migration 101
Gerenciamento do Live Migration 101
Live Migration *vs* Quick Migration 102

Gerenciamento do Hyper-V **103**

Infraestrutura de hardware *103*
Processador 103
Memória 104
Armazenamento 104
Rede 105
Outros componentes 105

Configurando máquinas virtuais *106*
Sistemas operacionais convidados 106
Discos suportados 106

Discos e armazenamento *108*
Opções de armazenamento em máquinas virtuais 108
Criação de discos rígidos virtuais 109
Discos físicos para máquinas virtuais 110

Aspectos de segurança *110*
Práticas recomendadas de segurança 111

Backup e Restore *113*
Backup on-line e off-line 114
Processo de restauração 114
Máquinas virtuais em cluster 115

Redes virtuais *115*
Noções básicas 116
Rede e máquinas virtuais 119
Configurando VLANs 120

Introdução ao VMware **121**

vSphere *122*

Availability 123
Security 124
Automation 124
Compute 125
Storage 125
Network 126

Disponibilidade *126*
Segurança 127
Escalabilidade 127
Vsphere hipervisor 128

Tipos de arquivos VMware *129*

Gerenciamento do VMware **133**

Componentes do gerenciamento *133*
VMotion 135
Alta Disponibilidade (HA) 136
Balanceamento dinâmico de recursos (DRS) 136
Balanceamento dinâmico de recursos de armazenamento (Storage DRS) 137
Storage VMotion 137
Site Recovery Manager (SRM) 138

Infraestrutura de TI *139*

Arquitetura do servidor *139*
Processador e memória 139

Armazenamento (storage) *140*
Storage Area Network (SAN) 141
Network-Attached Storage (NAS) 142

Rede Local – Local Area Network (LAN) *143*

Nuvem **145**

Computação em nuvem *145*

Principais benefícios *146*
Consolidação de um Datacenter 149

Modelos de computação em nuvem *149*

Serviços em nuvem *150*

Estratégia de TI em nuvem *152*

Consolidação do conhecimento **155**

Software *155*
Hipervisores 155
Microsoft Hyper-V 156

VMware ESX/ESXi 157
XenServer 160
Soluções de gerenciamento 161
Hardware *163*
Gerenciamento de memória 163
vCPU 164
Vlan tagging 164
Quantidades recomendadas 165
Sistemas de discos 165
Dell Virtualization Advisor Tool 167
Comparação entre soluções de virtualização 167
Bibliografia **169**

1

Aspectos gerais da virtualização

Introdução

- Time Sharing Processing (TSP)
- Multiprocessamento nos mainframes
- Hipervisor
- Da IBM à VMware
- Servidores x86
- A virtualização e seus benefícios

A virtualização é um assunto que tem despertado atenção, aparecendo como destaque no mundo da Tecnologia de Informação (TI), apesar de não ser uma novidade. Pode-se dizer que a ideia da virtualização nasceu com a publicação do artigo Time sharing processing in large fast computers, na Conferência Internacional de Processamento de Informação, realizada em Nova York em 1959. Escrito pelo cientista da computação Christopher Strachey, o texto tratou da multiprogramação em tempo compartilhado, estabelecendo um novo conceito de utilização de máquinas de grande porte. Esses grandes servidores poderiam, agora, aproveitar melhor os recursos de hardware. Baseado no trabalho inicial de Strachey, o MIT desenvolveu o padrão Compatible Time Sharing System (CTSS), que serviu como referência para diversos fabricantes.

Posteriormente, com base na evolução do padrão CTSS, a IBM introduziu o conceito de multiprocessamento nos mainframes, o que permitiu várias CPUs trabalharem como uma só, antecipando o conceito de virtualização. Esses mainframes também lançaram a concepção de memória virtual (virtual storage), como parte do Sistema Operacional (SO). Isso possibilitou a abstração e o mapeamento da memória real para memória virtual, além da especificação de partições (ou espaços de endereçamento), usadas por diferentes programas. Assim, surgiram as primeiras formas de virtualização.

Emulação
Capacidade de imitar o comportamento externo de um sistema, sem preocupação com estados e propriedades internas a ele.

Inicialmente, foi adotado o conceito de máquina virtual de processo – uma aplicação que executa sobre um SO A e emula o comportamento de um SO B. As aplicações desenvolvidas para B podem executar sobre A (figura 1.1). É importante salientar que essa técnica de implementação permite que binários de um processador sejam interpretados e substituídos por código equivalente de outro processador. Portanto, além de emular o Sistema Operacional, é possível emular processadores.

As desvantagens dessa técnica são basicamente duas: pior desempenho e desperdício de capacidades do hardware físico.

- O desempenho é sacrificado, já que há uma tradução de um sistema a outro, além da execução em modo de usuário;
- O desperdício da capacidade física do hardware vem do fato de que as máquinas virtuais de processo oferecem dispositivos de E/S genéricos e simples.

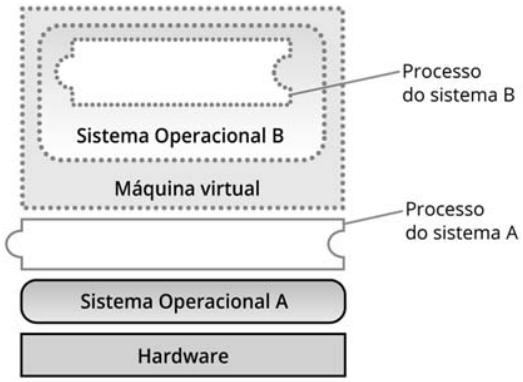


Figura 1.1 Máquina virtual de processo.

① Normalmente, há uma confusão no emprego dos termos *simulação* e *emulação*. A simulação envolve a modelagem matemática de um fenômeno, ou de um sistema, ressaltando as características-chave de seu comportamento, com vistas à predição e análise do objeto emulado. Com ela, é possível observar estados e propriedades como se fossem do sistema original real. Exemplos: simuladores de voo, de circuitos elétricos etc.

Os monitores de máquinas virtuais (Virtual Monitor Machine – VMM) surgiram para resolver desvantagens. Também conhecidos como hipervisores (hypervisors), são implementados como uma camada de software entre o hardware e o Sistema Operacional, oferecendo uma máquina virtual para o Sistema Operacional (figura 1.2). Dessa forma, eles conhecem e exploram eficientemente os dispositivos físicos de E/S. O desempenho tende a ser melhor, por não executarem em modo usuário, evitando chaveamentos de contexto. Existem duas técnicas usadas nos hipervisores: virtualização total e paravirtualização. A diferença essencial é a necessidade de o SO hóspede ser modificado (paravirtualização) ou não (virtualização total) para executar sob o hipervisor.

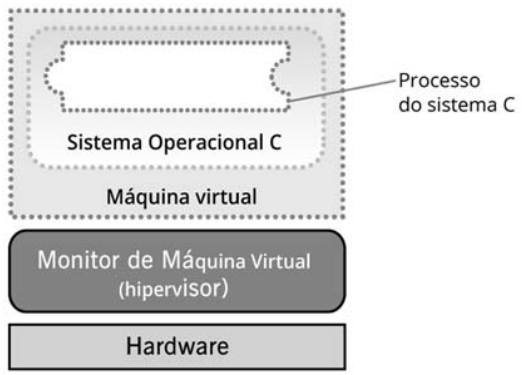


Figura 1.2 Monitor de máquina virtual (hipervisor).

O Sistema Operacional Virtual Machine (VM) da IBM surgiu baseado no conceito de hipervisor, programa que permite vários sistemas operacionais diferentes executarem de maneira isolada em um único hardware. O IBM VM/370, baseado na virtualização, foi bastante utilizado para a migração de um mainframe para outro ou de um Sistema Operacional para outro, permitindo que ambos agissem sob a supervisão do VM hypervisor.

No contexto dos anos 1970, quando a maioria dos mainframes – mesmo de um único fabricante – era caracterizada por ter seu próprio Sistema Operacional, o uso de máquinas virtuais permitiu que aplicações de sistemas legados fossem executadas nos novos sistemas. No entanto, na década de 1980, à medida que os computadores pessoais se popularizavam e a diversidade de sistemas operacionais foi reduzida para alguns poucos (Unix, Macintosh e Microsoft), a virtualização foi perdendo importância.

Vários fatores provocaram o ressurgimento da virtualização nos anos 1990, entre os quais podemos citar a máquina virtual Java – Java Virtual Machine (JVM); o aumento do poder computacional dos processadores e da capacidade de memória dos computadores pessoais; e a disseminação de sistemas distribuídos e da própria internet, como um modelo de serviços baseado em uma arquitetura cliente-servidor.

O marco inicial desta nova era foi o surgimento da VMware em 1998, criada por Diana Greene e Mendel Rosenblun. A VMware desenvolveu o primeiro hipervisor, que permitiu a virtualização de servidores em plataformas x86. Desde 1996, entretanto, outra empresa já tratava de virtualização, mas em ambiente Macintosh (Connectix, fundada em 1988). Em 2003, a Microsoft adquiriu a Connectix, e, no ano seguinte, a EMC comprou a VMware. Em seguida, lançou o Microsoft Virtual Server 2005, seu primeiro produto com foco na virtualização de servidores.

Um aspecto importante de qualquer projeto de virtualização é a possibilidade de redução do consumo de energia dos servidores e da refrigeração. O fato de otimizar o uso de recursos promove a economia, pois um servidor à plena carga e outro sem carga consomem praticamente a mesma quantidade de energia. A virtualização também é uma aliada importante dos projetos que envolvem a recuperação de desastres, por simplificar a construção de um ambiente com essa finalidade.

Da forma como é feita hoje, a virtualização é uma revolução na área de TI, o que pode ser confirmado por análises econômicas conduzidas pela Gartner Inc., pela criação e evolução de associações como a Enterprise Management Association (EMA) e pelo grande crescimento de empresas como a VMware. O International Data Corporation (IDC) prevê que o mercado e os investimentos em virtualização devem aumentar de 6,5 bilhões para 15 bilhões de dólares em 2011.

Tipos de virtualização

- Virtualização de servidores
- Virtualização de desktops
- Virtualização do armazenamento
- Virtualização das aplicações
- Virtualização de redes

A virtualização pode ser feita de diferentes formas, incluindo:

- **Virtualização de servidores**: a mais comum e fácil de ser justificada. Diferente da época dos mainframes, a virtualização agora é feita em servidores x86. É o tema central deste curso.
- **Virtualização de desktops**: trata da configuração dos desktops dos usuários finais em uma infraestrutura centralizada virtual. Isso significa que as aplicações de desktop também passam a executar em um datacenter, sob a forma de máquinas virtuais. Esse é o conceito de Virtual Desktop Infrastructure (VDI), que permite a montagem dinâmica de desktops, oferecendo maior confiabilidade e otimização do uso de espaço em disco com a consolidação do armazenamento e flexibilidade na escolha do Sistema Operacional. Existem limitações para o uso dessa técnica de forma generalizada. Normalmente, sua adoção é antecedida por um trabalho de levantamento da situação a ser considerada.
- **Virtualização do armazenamento (storage)**: a ideia é introduzir um componente (appliance) que permite às diversas unidades heterogêneas de armazenamento (discos físicos) serem vistas como um conjunto homogêneo de recursos. A virtualização do armazenamento não é tão popular quanto a de servidores.
- **Virtualização das aplicações**: trata do conceito de execução do programa por completo, em um repositório central, permitindo a configuração centralizada do aplicativo, o que melhora seu gerenciamento, por permitir que seja feita em um único lugar.
- **Virtualização de redes:** arquitetura que proporciona um ambiente de rede separado para cada grupo ou organização. Esses ambientes lógicos são criados sobre uma única infraestrutura compartilhada de rede. Cada rede lógica fornece ao grupo de usuários correspondente plenos serviços de rede, semelhantes aos utilizados por uma rede tradicional não virtualizada. A experiência da perspectiva do usuário final é a de ter acesso a uma rede própria, com recursos dedicados e políticas de segurança independentes. Assim, envolve a lógica segmentação da rede de transportes, dos dispositivos e de todos os serviços. Como as diversas redes lógicas compartilham uma infraestrutura comum, muitas vezes centralizada e com um conjunto de equipamentos e servidores, os grupos de usuários podem colaborar com maior flexibilidade e capacidade de gerenciamento. O que permite novos processos de negócio, que não seriam possíveis (e nem sequer imagináveis) por meio de uma rede tradicional.

Conceito e categorização

- Categorias de virtualização
- Consolidação de servidores
- Camada de abstração entre hardware e software
- Camada de virtualização, hipervisor e Virtual Machine Monitor

Os softwares de virtualização podem ser classificados em três categorias:

- **Nível do hardware**: a camada de virtualização é posta diretamente sobre a máquina física e a apresenta às camadas superiores, como hardware abstrato, similar ao original. Esse é o caso da maioria dos hipervisores (VMware ESX, Xen e Hyper-V).
- **Nível do Sistema Operacional**: mecanismo que permite a criação de partições lógicas em uma plataforma, de maneira que cada uma seja vista como máquina isolada, compartilhando o mesmo SO. A camada de virtualização está inserida entre o Sistema Operacional e as aplicações. Exemplos: Jails, OpenVZ, Solaris Zones, Containers, Linux-VServer, Parallels Virtuozzo, SandBox, KVM e Sun VirtualBox.
- **Nível da linguagem de programação**: a camada de virtualização é um programa de aplicação do Sistema Operacional da plataforma. Define uma máquina abstrata sobre a qual executa uma aplicação desenvolvida em uma aplicação de alto nível. A máquina virtual Java (JVM) é o exemplo mais marcante.

As categorias acima possuem objetivos diferentes, mas buscam aspectos comuns: (i) oferecer compatibilidade de software; (ii) permitir o isolamento entre máquinas virtuais, ou seja, um software em execução não deve ser afetado por outro; (iii) o encapsulamento, que possibilita a captura do estado da máquina virtual. A camada de virtualização deve ser projetada para não impactar o desempenho das aplicações.

No nível do hardware, a virtualização mais comum é a de servidores. Ela trata da consolidação de vários servidores físicos e subutilizados, em um servidor físico com alto grau de utilização, reduzindo a complexidade do gerenciamento, o espaço físico e os requisitos de energia e refrigeração.

Uma abordagem comum utilizada para conceituar a virtualização é defini-la como uma camada de abstração entre o hardware e o software, que protege o acesso direto do software aos recursos físicos do hardware. A virtualização permite que a camada de software (aplicações e Sistema Operacional) seja isolada da camada de hardware (figura 1.3). A forma pela qual essa camada de abstração é implementada dá origem às máquinas virtuais de processo e aos monitores de máquinas virtuais (hipervisores).

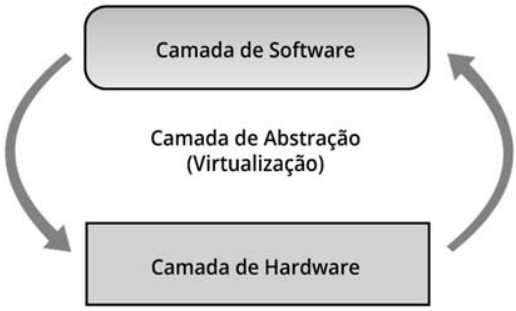


Figura 1.3 Papel da virtualização.

Hipervisores

- Características
- Tipos
  - Tipo I (bare metal)
  - Tipo II (hosted)

O hipervisor é a plataforma básica das máquinas virtuais como conhecemos. Suas principais funções consistem no escalonamento de tarefas, gerência da memória e manutenção do estado da máquina virtual. O desempenho e a escalabilidade do hipervisor definem a qualidade da virtualização. Podemos citar como características necessárias a um hipervisor: segurança sobre os recursos virtualizados e agilidade na reconfiguração de recursos computacionais, sem interromper as operações do servidor de máquinas virtuais.

Os hipervisores são classificados em dois tipos (figura 1.4):

- **Tipo I** (bare metal, nativo ou supervisor): executa diretamente no hardware do servidor. Controla o hardware e o acesso do Sistema Operacional convidado (guest OS). O papel do hipervisor nativo é compartilhar os recursos de hardware entre as máquinas virtuais, de forma que cada uma delas imagina ter recursos exclusivos. Exemplos: VMware ESX Server, Microsoft Hyper-V e Xen Server. Uma variação do tipo I é o "embedded hypervisor", que é instalado no firmware (como o VMware ESXi). Esse hipervisor é pequeno e tem um impacto mínimo nos recursos e no desempenho do servidor físico.
- **Tipo II** (hosted): aplicação que fornece um ambiente de execução para outras aplicações. Executa sob um Sistema Operacional nativo como se fosse um processo deste. A camada de virtualização é composta por um SO hóspede e um hardware virtual, que são criados sobre os recursos de hardware oferecidos por meio do SO nativo. Exemplos: VMware player, Virtualbox e Virtual PC.

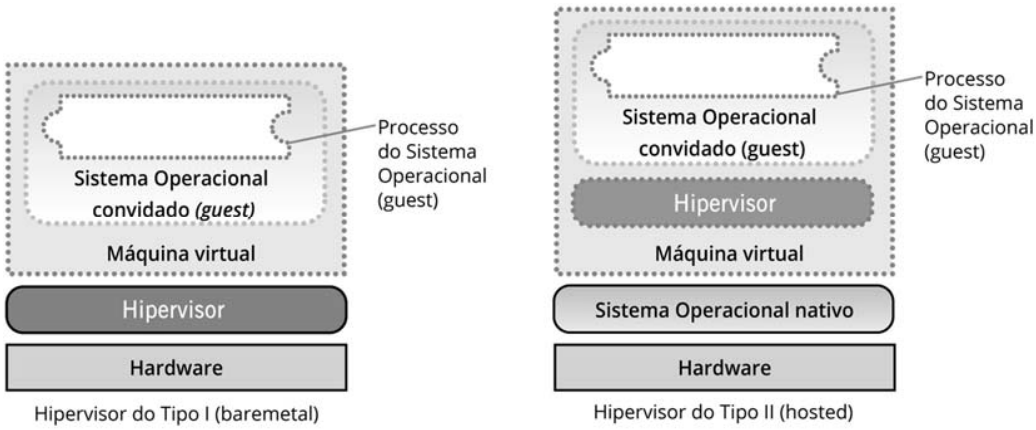


Figura 1.4 Tipos de hipervisores.

Tipos de virtualização de servidores

- Nível de privilégio da arquitetura x86
- Tipos
  - Virtualização completa
  - Paravirtualização
- Virtualização assistida por hardware
  - Intel VT
  - AMD V

A arquitetura x86 provê quatro modos de operação para o processador, identificados de 0 a 3, denominados anéis de proteção (rings) ou Current Privilege Level (CPL). Nos sistemas operacionais convencionais, como o Windows e o Linux, apenas os modos 0 e 3 são utilizados. O anel nível 0 detém os maiores privilégios de execução e é usado pelo Sistema Operacional. O anel nível 3, de menor privilégio, é empregado por processos do usuário. As instruções de máquina do processador são, dessa forma, divididas em instruções não privilegiadas que executam em modo usuário (anel 3) e instruções privilegiadas que atuam em modo protegido (anel 0). Se um processo de usuário tenta executar uma instrução privilegiada, ocorre uma exceção (trap) que deverá ser tratada adequadamente pelo Sistema Operacional. Entretanto, com a arquitetura x86, existem instruções que afetam o funcionamento do processador, como as privilegiadas, mas que são executadas em modo usuário sem gerar exceções (traps). Essas instruções são denominadas “sensíveis” e seu uso em sistemas operacionais nativos (não virtualizados) não causa problemas. Em ambientes virtualizados, as instruções sensíveis devem ser tratadas adequadamente, já que podem ser feitas por processo executado em um Sistema Operacional convidado (virtualizado) e, assim, alterar o comportamento do SO nativo ou de outro Sistema Operacional convidado.

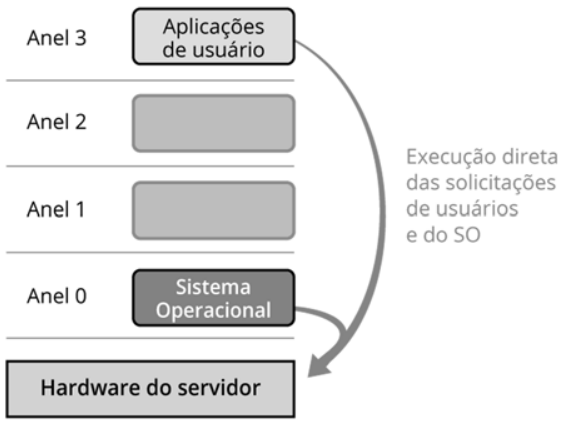


Figura 1.5 Nível de privilégio na arquitetura x86.

Tipos de virtualização

A virtualização pode ser realizada de diferentes maneiras, cada uma com seus prós e contras. Na prática, em arquiteturas x86, as opções de virtualização alteram o nível de privilégios (rings) default. As soluções baseadas em hipervisores incluem a virtualização completa e a paravirtualização.

Virtualização completa

A virtualização completa (full virtualization) realiza toda a abstração do sistema físico, sobre o qual o Sistema Operacional convidado é executado. Não é necessário fazer qualquer modificação no SO convidado ou em suas aplicações. Esse tipo de virtualização facilita a migração de máquinas virtuais entre servidores físicos, porque há total independência das aplicações e dos recursos físicos do servidor. Ainda, a segurança é facilitada pelo isolamento entre as máquinas virtuais, já que cada instância da máquina virtual é um processo do Sistema Operacional nativo.

Uma desvantagem da virtualização completa é o desempenho, pois o hipervisor verifica a execução de todas as instruções privilegiadas ou sensíveis feitas pelo Sistema Operacional convidado e as substitui por ações equivalentes controladas. Outro ponto desfavorável é a dificuldade de emular o funcionamento dos dispositivos de E/S, por conta da diversidade dos dispositivos existentes. A solução empregada consiste em implementar hardwares virtuais que emulam dispositivos genéricos, o que pode causar a subutilização dos dispositivos de E/S reais. A figura 1.6 ilustra a arquitetura de privilégio no x86.

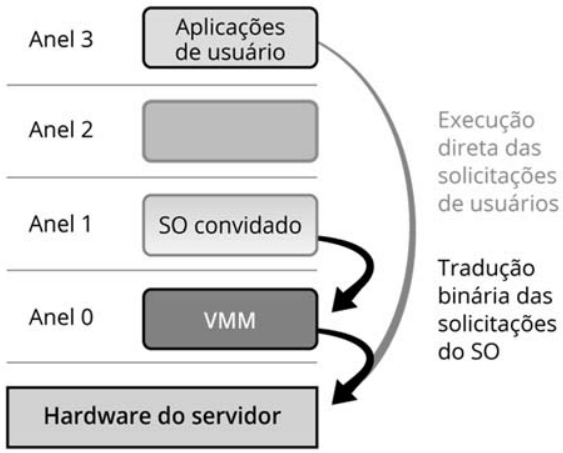


Figura 1.6 Virtualização completa no x86.

Paravirtualização

A paravirtualização é uma alternativa para contornar os problemas de desempenho e subutilização de recursos da virtualização total. Nesse caso, o Sistema Operacional convidado (ou hóspede) é alterado para chamar a máquina virtual (hypercalls) sempre que for executar uma instrução sensível (figura 1.7). As instruções não privilegiadas, aquelas realizadas pelos processos de usuários, podem ser executadas diretamente sobre o processador nativo. Em relação aos dispositivos de E/S, os hipervisores que empregam paravirtualização permitem que as máquinas virtuais usem os drivers do dispositivo físico real sob o controle do hipervisor, o que é interessante por otimizar o desempenho. A principal desvantagem da paravirtualização é a necessidade de modificação do Sistema Operacional hospedado ou convidado, o que pressupõe acesso ao código-fonte. O Xen Open Source é um exemplo de sistema baseado em paravirtualização, que virtualiza o processador e a memória, utilizando um núcleo Linux modificado e virtualizando o sistema de E/S com drivers de dispositivos customizados.

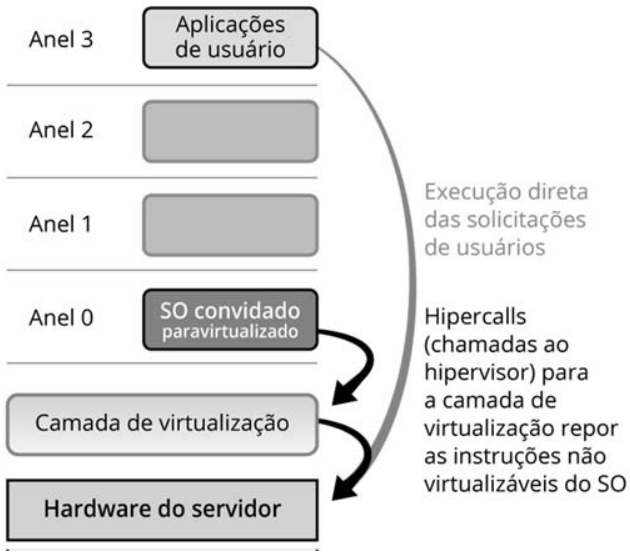


Figura 1.7 Paravirtualização no x86.

Virtualização assistida por hardware

Na prática, há uma relação custo/benefício entre a virtualização total e a paravirtualização. Enquanto a primeira permite o uso de um Sistema Operacional convidado sem modificações, a segunda precisa alterá-lo para substituir instruções privilegiadas e sensíveis por hypercalls, mas oferece um melhor desempenho. Sendo assim, os fabricantes Intel e AMD investiram em extensões na arquitetura x86, para suportar a virtualização e melhorar o desempenho da solução como um todo. Essas extensões são genericamente denominadas de virtualização assistida por hardware (Hardware Assisted Virtualization – HAV). O movimento da AMD e da Intel praticamente eliminou as vantagens de desempenho dos sistemas baseados em paravirtualização, que tinham o ônus de modificar o Sistema Operacional para funcionar. Por razões técnicas, apenas os novos sistemas x64 fazem uso do HAV.

Basicamente, tanto Intel (Intel VT) como AMD (AMD-V) alteraram o funcionamento dos anéis de proteção (rings) dos processadores x86. Como visto anteriormente, o anel nível 0 é utilizado pelo SO e executado em modo protegido, e o anel nível 3 (de menor privilégio) é empregado pelos processos realizados em modo usuário. Os anéis 1 e 2 não são usados pelos sistemas operacionais convencionais. A Intel e a AMD, de formas diferentes, mas com a mesma filosofia, criaram dois novos modos de operação para os anéis de proteção do processador: modo “root” e modo “non-root”. Os anéis de 0 a 3 executam no modo “non-root” e há um anel adicional – de maior prioridade – para o modo “root”. Dessa forma, um Sistema Operacional executa sem modificações nos níveis 0 a 3, e o hipervisor atua no anel adicional de maior prioridade (modo “root”). Portanto, o hipervisor passou a ter total prioridade sobre o SO. Nesse novo modo de operação, as instruções privilegiadas e sensíveis executadas pelo Sistema Operacional convidado causam um desvio (trap) para o hipervisor, que tem a responsabilidade de tratar adequadamente a ocorrência dessas ações.

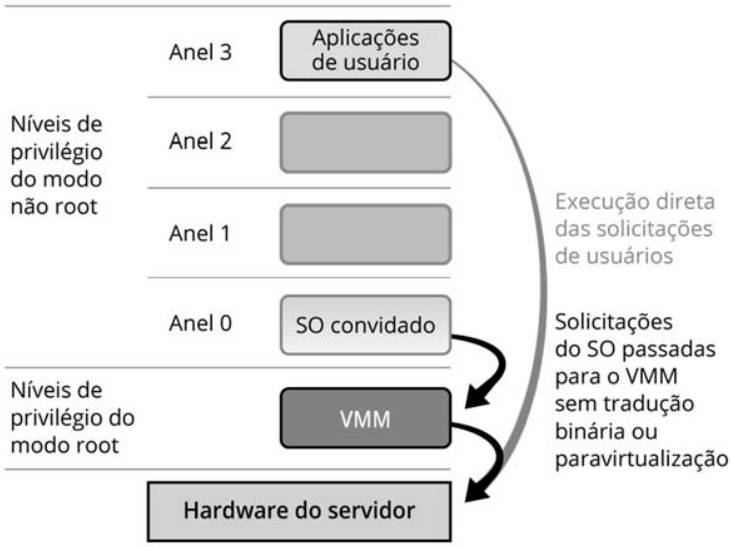


Figura 1.8 Virtualização assistida por hardware em arquiteturas x86.

Cenários da virtualização

- Consolidação de servidores
- Melhorar a continuidade dos negócios
- Criar um novo ambiente de testes e de desenvolvimento de software
- Proteger e gerenciar os desktops da empresa
- Hospedar aplicações legadas
- Datacenter dinâmico

A virtualização é o elemento central de um datacenter. É comum encontrarmos o emprego da filosofia "um servidor por serviço", por motivos que vão desde a segurança até o oferecimento de serviços de internet a máquinas-clientes. Nesse contexto, a carga de processamento não explora todo o poder computacional disponibilizado pelo processador do servidor, o que implica desperdício de ciclos de processamento. Pesquisas em Internet Data Centers (IDC) mostram que somente cerca de 15% da capacidade dos servidores é utilizada, com 85% de ociosidade. Além do desperdício, há todo o problema de consumo de energia, refrigeração, espaço físico, gerenciamento e manutenção dos vários servidores. A virtualização surge como solução para esses obstáculos. Ela possibilita a otimização dos recursos e torna a aplicação independente do hardware. A virtualização também está alinhada ao conceito de TI verde (green computing), já que permite uma economia significativa de energia no datacenter. A ideia básica é permitir que várias máquinas virtuais, cada uma responsável por um serviço, executem sobre uma única máquina física. Essa situação é denominada de consolidação de servidores e é especialmente importante no contexto de datacenters.

Outro ponto importante da virtualização é possibilitar a continuidade dos negócios a um custo adequado, utilizando recursos já incorporados nos produtos de virtualização, como a

alta disponibilidade (High Availability – HA) e a recuperação de desastres (Disaster Recovery – DR). Essencialmente, uma máquina virtual é uma imagem que pode ser armazenada e instanciada a qualquer momento. Isso permite que uma máquina virtual seja monitorada e migrada de um servidor físico para outro, facilitando a recuperação de falhas e o balanceamento de carga.

A flexibilidade e a portabilidade das máquinas virtuais também tornam interessante seu uso em ambientes de desktops. É possível imaginar, por exemplo, o desenvolvimento de produtos de software destinados a vários sistemas operacionais, sem a necessidade de uma plataforma física para criar e testar cada um deles. Nesse caso, as máquinas virtuais em desktops podem ser usadas para a definição de ambientes experimentais completos, sem interferir no Sistema Operacional original do equipamento.

Há ainda a virtualização de desktops, que permite a cada usuário estabelecer uma seção de trabalho dentro de um sistema centralizado, a partir de um cliente fino (thin client) ou de outro software-cliente. A diferença para soluções do tipo "Terminal Services" é que cada usuário pode empregar um Sistema Operacional diferente totalmente isolado dos demais. Como apenas o sistema centralizado possui as imagens das máquinas virtuais, as questões de configuração, gerenciamento e proteção ficam mais simples.

Por fim, a virtualização é uma ferramenta muito útil para hospedar e executar sistemas legados. Como uma máquina virtual é um ambiente que inclui um Sistema Operacional, bibliotecas e aplicações de forma totalmente independente e isolada de outra, é possível manter versões de antigos sistemas operacionais e bibliotecas exigidas por sistemas legados.

Por fim, o datacenter dinâmico utiliza os benefícios da virtualização para criar uma infraestrutura mais ágil, combinada com novos recursos de gerenciamento que permitem mover máquinas virtuais sem causar impacto sobre as atividades dos usuários.

Tendências da virtualização

- Datacenter dinâmico
- Computação em nuvem (cloud computing)
  - Estágios da computação em nuvem
  - Serviços de computação em nuvem

Datacenter dinâmico

A virtualização muda a maneira como um datacenter é gerenciado. O conceito de virtualização desfaz um modelo tradicional utilizado, que é o de associar cada aplicação ou serviço a uma máquina específica. Permite melhorar a utilização dos servidores que tradicionalmente rodam entre 10% a 15% da sua capacidade. A execução de várias máquinas virtuais em um servidor permite ocupar os ciclos ociosos do processador e eliminar a necessidade de outro servidor físico, reduzindo o consumo de energia e o espaço físico necessário. Com o conceito de datacenter dinâmico, é possível provisionar os recursos de forma imediata, mediante a demanda.

Computação em nuvem

A ideia central da computação em nuvem é possibilitar às aplicações que rodam em datacenters isolados rodarem na cloud (internet) em um ambiente de larga escala e com um uso "elástico" de recursos. A virtualização é o elemento central dessa visão, e a sua importância pode ser entendida facilmente. A computação em nuvem pode ser privada, pública ou uma combinação das duas (híbrida).

- **Cloud Privada**: acesso aos recursos por meio de infraestrutura da própria organização. Por pertencer a uma organização, o controle é mais rígido dos recursos, com segurança. Não há restrições quanto à banda, permitindo maior resiliência.
- **Cloud Pública**: acessa recursos por meio da internet, usando interface web. Alocação e pagamento por demanda (soluções elásticas). O fato de ser pública não significa que seja livre nem aberta.
- **Cloud Híbrida**: utiliza recursos de uma ou mais nuvens públicas e privadas.

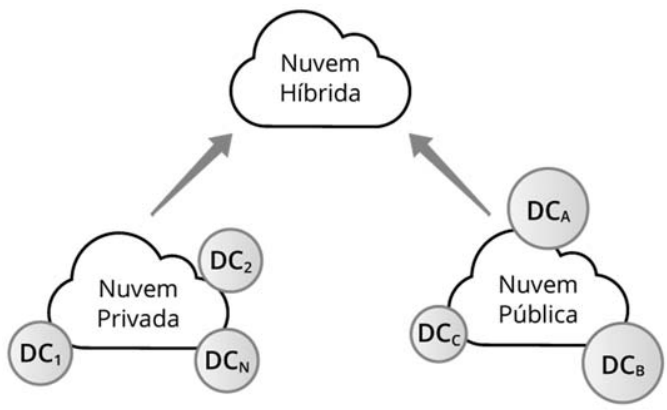


Figura 1.9 Computação em nuvem.

Em contexto mais simples, a virtualização permite que um único datacenter funcione como um conjunto de recursos virtuais. Num segundo momento, possibilita a otimização do uso dos recursos em uma configuração de datacenters redundantes. A diretriz da computação em nuvem é a utilização da virtualização para viabilizar o uso de um conjunto de datacenters em uma espécie de nuvem privada. Nesse estágio, a virtualização deixaria o uso de recursos privados e públicos em uma configuração híbrida de nuvem.

Mesmo aplicações que continuem a executar localmente podem empregar serviços de infraestrutura providos por uma nuvem pública, como, por exemplo, armazenamento de dados e acesso a serviços providos internamente e externamente a essa nuvem. O ambiente de computação em cloud é essencialmente distinto do ambiente tradicional de computação. É a aplicação que propicia disponibilidade e resiliência, e não o hardware redundante dentro dos nós individuais. Ainda, as aplicações de computação em nuvem são normalmente homogêneas e podem chegar a 1.500 nós utilizados.

Principais fornecedores de soluções de virtualização

- VMware
- Microsoft
- Citrix
- RedHat

Os principais fornecedores de software de virtualização para servidores são:

- **VMware**: VMware ESXi e VMware vSphere.
- **Microsoft**: Hyper-V, System Center Virtual Machine Manager (SCVMM).
- **Citrix**: Xen Server, Citrix Essentials for Hyper-V e Citrix Essentials for Xen Server.
- **RedHat**: KVM, Red Hat Enterprise Virtualization.
- **Huawei**: Data Center Virtualization Solution (Fusion Cloud suite — FusionSphere, FusionCube and FusionAccess).

Essas empresas comercializam os softwares de virtualização de diversas formas e com diferentes funcionalidades. Pesquisa com 365 profissionais de grandes organizações nos Estados Unidos – realizada pelo Enterprise Strategy Group em 2008 – indicou que todas as organizações consultadas já utilizam a virtualização de alguma maneira. A maioria das emprega o VMware como principal ferramenta, seguido da Microsoft e da Citrix. Algumas pesquisas feitas em 2014 mostram as movimentações das soluções com destaque para o posicionamento da Huawei. A figura 1.5 concentra o ranking de 2012 a 2014, conforme as publicações da empresa Gartner.



Figura 1.10 Posicionamento das soluções em 2012, 2013 e 2014.

Licenciamento

- Desafios
- Processadores de vários núcleos
- Modos de licenciamento

O licenciamento de plataformas virtualizadas varia de acordo com a estratégia de mercado de cada fabricante. Alguns preferem adotar a estratégia de licenciamento por soquete (CPU); outras preferem o licenciamento por núcleo.

O raciocínio é simples: é fácil mensurar os chips de processadores. Essa maneira de obter a licença mobiliza o pessoal de infraestrutura de TI para utilizar sempre os chips mais poderosos disponíveis. Obter o máximo de um software sempre exigiu hardware de alto desempenho; a diferença é que atualmente o desenvolvimento melhora com o aumento da quantidade de núcleos, em vez de crescer com a quantidade de megahertz (MHz). Essa estratégia da indústria está ligada ao crescimento por meio do aumento de núcleos, que é mais interessante, pois reduz o consumo de energia. Além disso, a linearidade que eventualmente poderia ser perdida com a utilização de muitos núcleos tem sido otimizada por fabricantes como Intel e AMD.

Para verificar o modelo utilizado por cada um dos fabricantes o modelo

Gerenciamento e segurança com a virtualização

- Gerenciamento
- Segurança

A virtualização introduz mais um elemento de gerenciamento que cuida das máquinas virtuais. Normalmente, é realizado por outro componente da solução de virtualização, que, tipicamente, necessita de uma infraestrutura específica.

A segurança da virtualização é um aspecto essencial a ser considerado durante a escolha do software. Ela é comumente garantida pelo hipervisor, por meio de alguns mecanismos de detecção de intrusão que auxiliam no controle do acesso aos discos e à rede.

Limitações da virtualização

- Limitações de carga excessiva
- Gerenciamento do licenciamento
- Falta de profissionais especializados

Principais limitações da virtualização:

- **Aplicativos de carga excessiva**: incluindo os sistemas gerenciadores de bancos de dados, podem ser um fator limitante. Considerando que sempre existe uma perda de desempenho introduzida pelo hipervisor, se uma aplicação ou um sistema gerenciador

de banco de dados já demanda boa parte dos recursos do servidor, qual seria a razão para virtualizar este servidor?

- **Gerenciamento do licenciamento**: as licenças podem ser um fator limitante quando é necessário saber exatamente a regra para cada aplicação; isso é feito de maneiras diferentes pelos diversos fabricantes. Em uma determinada situação de carga, o licenciamento é válido; já outro que utiliza uma configuração diferente de hardware pode não ser.
- **Falta de profissionais especializados**: como a virtualização é relativamente recente, ainda existem poucos profissionais experientes que dominem a técnica e as opções comerciais disponíveis.

Desempenho e benchmarks

- Desempenho é uma questão relevante
- Desenvolvimento de benchmarks específicos

Como a virtualização consiste basicamente em inserir uma camada de software adicional em um sistema computacional, a questão sobre quanto isso afeta o desempenho final é imediata. Estudos feitos pela VMware e pela XenSource apontam para uma queda de desempenho, em geral entre 2% e 10%, com algumas situações impondo perdas maiores. Cabe ressaltar que esses resultados foram obtidos por meio de benchmarks genéricos. Porém, atentos a esse aspecto, os fabricantes estão investindo esforços para reduzir a baixa de desempenho. Para a consolidação de uma forma padronizada e isenta de avaliação, um comitê específico (SPEC Virtualization Comitee) desenvolveu um benchmark para a virtualização, que pode ser encontrado em: www.spec.org

Soluções de virtualização

Principais soluções de virtualização no mercado:

- VMware ESX
- Hyper-V
- Xen
- KVM

VMware ESX

Possui poderosa infraestrutura e oferece muitos recursos avançados para gerenciamento e administração de um ambiente virtual, como VMotion (migração de máquinas virtuais), recursos de alta disponibilidade, tolerância a falhas, Storage Motion, entre outros. Mesmo sendo considerada uma solução de custo elevado, é o principal software usado por grandes empresas.

Hyper-V

É a solução oferecida pela Microsoft para sistemas de virtualização. Adotado principalmente em empresas de pequeno e médio porte, conta com recursos como Live Migration, Quick Migration e Dinamic Memory. É oferecido de forma gratuita, entretanto, é necessário ter a licença do Windows Server para usá-lo.

Xen

Baseado em Linux, é menos popular que o Hyper-V e o VMware ESX, mas constitui uma solução muito estável. Fornece interface de administração por linha de comando (CLI), além de uma alternativa gráfica (GUI): XenClient – essa é a melhor indicação para quem já utiliza produtos da Citrix.

KVM

O KVM aparece como alternativa completa para sistemas de virtualização baseados em Linux. Assim como o Xen, fornece administração por linha de comando e por interface gráfica; dispõe de recursos de migração on-line e off-line; possui o sistema de otimização do gerenciamento de memória Kernel Samepage Merging (KSM); utiliza o protocolo QEMU Monitor Protocol (QMP) – que é um protocolo Java Script Object Notation (JSON-based) –; é open source; e permite adicionar CPUs e dispositivos PCI em tempo real.

O KVM: origem

O Kernel-based Virtual Machine (KVM) foi inicialmente desenvolvido pela empresa Qumranet entre 2006 e 2007. Seu crescimento tornou-se mais acentuado em 2008, quando a empresa foi adquirida pela RedHat, que desde então incorporou o KVM às suas distribuições.

É uma solução completa de virtualização para Linux com arquitetura x86 de 64 bits e explora as tecnologias de virtualização Intel VT e AMD-V dos processadores mais recentes. Distribuído como um componente integrado do kernel desde a versão 2.6.20, transforma o Linux em um hipervisor estável e de alto desempenho.

Arquitetura

Ao carregar o módulo do KVM no Linux, o kernel exporta um dispositivo chamado /dev/kvm, que fornece ao sistema um novo modo de execução, além dos modos já existentes kernel e user, chamado de guest mode (Figura 9.6). Todos os processos no espaço de usuário acessam de forma compartilhada os dispositivos do /dev, entretanto, o dispositivo /dev/kvm permite que uma máquina virtual tenha seu espaço de endereço independente do kernel ou de qualquer outra máquina virtual, permitindo o seu isolamento; ou seja, cada processo que acessa o dispositivo /dev/kvm recebe um mapa de endereços diferentes.

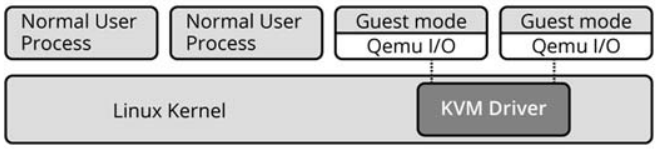


Figura 1.6 Arquitetura KVM.

A estrutura do KVM é dividida em dois componentes básicos:

- **KVM driver**: é um conjunto de módulos do kernel que fornece os drivers específicos para o tipo de processador e a infraestrutura necessária para a virtualização, atuando como interface para o user space.
- **User space**: roda uma versão modificada do qemu, que gerencia E/S e funciona como um espaço virtual para os sistemas hóspedes, além de fornecer mecanismos de controle e gerenciamento das máquinas virtuais.

As máquinas virtuais são tratadas como processos do Sistema Operacional e, por isso, é possível usar os comandos do sistema para manipulá-las. Junto com cada máquina, é instanciado um processo do qemu, levemente modificado, que intercepta qualquer solicitação de I/O feita pela máquina e a emula. Além das solicitações de I/O, o qemu é responsável por emular os dispositivos, como adaptadores gráficos, dispositivos de rede, discos, entre outros.

- i440FX PCI bridge e PIIX3 ISA bridge;
- Cirrus CLGD 5446 PCI VGA;
- Mouse e teclado PS/2;
- Drive de disquete;
- Adaptadores de rede PCI e ISA;
- Portas seriais;
- Placa de som Creative Sound Blaster 16;
- PCI UHCI USB controller e Hub USB virtual.

A utilização do KVM, junto com o qemu, permite que as máquinas virtuais rodem com baixo overhead, tornando-o uma alternativa competitiva às opções do mercado.

Gerenciamento de memória do Linux

O Linux é baseado no conceito de Memória Virtual e utiliza técnicas de paginação de registros, ou seja, quando um processo está rodando na máquina, não há alocação de toda a memória de uma única vez. Ela é feita de acordo com a necessidade, tornando a utilização da memória mais eficiente.

Quando a memória física é totalmente utilizada ou um processo fica determinado tempo sem ser utilizado, o Linux move a memória de um processo para o swap. Como o swap é feito em disco, torna-se significativamente mais lento que a RAM.

O KVM usa todos os benefícios do gerenciamento de memória do Linux. As máquinas virtuais são processos para o host, e o KVM pode alocar mais memória para os guests do que realmente existe no host. Não há um bloco de memória física dedicado; os guests podem ser manipulados com comandos do Sistema Operacional. Utilizando o swap, é possível diminuir a quantidade de memória real necessária para os guests.

SAIBA MAIS Quando o total de memória usada pelos guests é maior do que a memória física disponível, temos um overcommit. Existem alguns mecanismos específicos para tratar o overcommit em sistemas de virtualização.

Swapping

O host escolhe algumas páginas de memória do guest e envia para o disco. Se o guest necessitar daquela informação, o host traz de volta as páginas para a memória principal. É o método mais confiável do ponto de vista do host. Entretanto, para o guest pode não ser a melhor escolha, porque o host não tem controle sobre os processos que estão em execução dentro do guest.

Balão de memória

Mais eficiente do que o swapping do ponto de vista do guest, o balão de memória funciona de forma cooperada: o guest se comunica com o host e informa quais páginas de memória podem ser liberadas. Pelo fato de a escolha ser feita pelo guest, não tem tanta confiabilidade.

É possível obter informações sobre esse mecanismo por meio do console do qemu. Exemplo:

```
(qemu) info ballon
ballon: actual 1024
```

Kernel Same-page Merging (KSM)

Realiza uma unificação das páginas de memória que são iguais. Dessa forma, as páginas duplicadas são removidas, e seu espaço, liberado. Esse sistema oferece mais vantagem quando houver muitas máquinas virtuais homogêneas.

Para melhorar o processo de gerenciamento de memória, foi desenvolvido um daemon chamado ksmtuned, que gerencia dinamicamente os ajustes de memória do KSM, tomando como base a quantidade de memória livre disponível.

Live migration

Consistem em mover um guest de um host para outro. Dessa forma, libera a memória que estava sendo utilizada.

O KVM usa todas as técnicas para conseguir os melhores resultados no gerenciamento de memória. Preferencialmente, são utilizados os métodos de balão de memória e KSM; em seguida, o live migration; e, em último caso, o swap.

Sistemas convidados suportados

O KVM está em constante aperfeiçoamento e desenvolvimento. A lista dos sistemas operacionais convidados suportados está disponível na web.

Acesse a lista atualizada em http://linux-kvm.org

Snapshots

O termo "snapshot" teve origem no universo da fotografia, onde era utilizado para designar fotos instantâneas. Atualmente, é um termo muito usado em computação para denominar uma cópia de um sistema, filesystem, memória, banco de dados, entre outros, em determinado momento.

Diferente de um backup convencional, no qual todos os dados são copiados em um intervalo longo de tempo, um snapshot cria uma cópia do estado atual em alguns segundos. Isso ocorre porque os snapshots trabalham apenas com ponteiros e praticamente não fazem a movimentação dos dados.

O tamanho de um snapshot está relacionado à modificação dos ponteiros dos arquivos, e não a seu tamanho real.

O KVM permite a criação de snapshots dos discos das máquinas virtuais. Entretanto, apenas o formato qcow2 tem suporte a essa tecnologia.

2

Projeto de virtualização

A virtualização não deve ser utilizada indiscriminadamente. Cada organização tem características e necessidades próprias que precisam ser cuidadosamente analisadas e consideradas antes de se adotar a virtualização. Dois aspectos são fundamentais: o projeto de virtualização e a infraestrutura de hardware.

Primeiro, deve ser feito um levantamento da infraestrutura atual e os objetivos que se pretende atingir. Fazem parte dessa etapa tanto as análises de recursos financeiros e de seu retorno como as técnicas. Segundo, apesar de cada organização possuir características próprias, as plataformas existentes podem ser genericamente classificadas em três grandes grupos: servidores, dispositivos de armazenamento (storages) e infraestrutura de rede. O objetivo dessa seção é detalhar os dois aspectos, a começar pelas questões relacionadas ao projeto de virtualização.

Projeto de virtualização

Um projeto de virtualização merece cuidado especial. Definir corretamente a infraestrutura para suportar as aplicações no ambiente virtualizado é uma tarefa complexa, por envolver diversas variáveis. Também deve-se levar em consideração a necessidade de crescimento do ambiente e o consumo de recursos quando do uso de algumas funcionalidades. Se a infraestrutura já existe, compartilhá-la com a virtualização requer um estudo sobre as suas reais possibilidades.

Os principais passos para um projeto de virtualização estão listados abaixo resumidamente:

- Educação;
- Planejamento da capacidade (Capacity Planning);
- Escolha dos fornecedores;
- Projeto;
- Implementação (Deployment);
- Gerenciamento.

O projeto de virtualização começa com a educação e conscientização sobre os desafios e necessidades de sua organização. Em seguida, a etapa de planejamento da capacidade é necessária para atender às demandas da organização. Com base nas informações coletadas e nos cenários, pode ser feita a opção pelo fornecedor do software. As etapas de um projeto de virtualização estão ilustradas na figura seguinte.

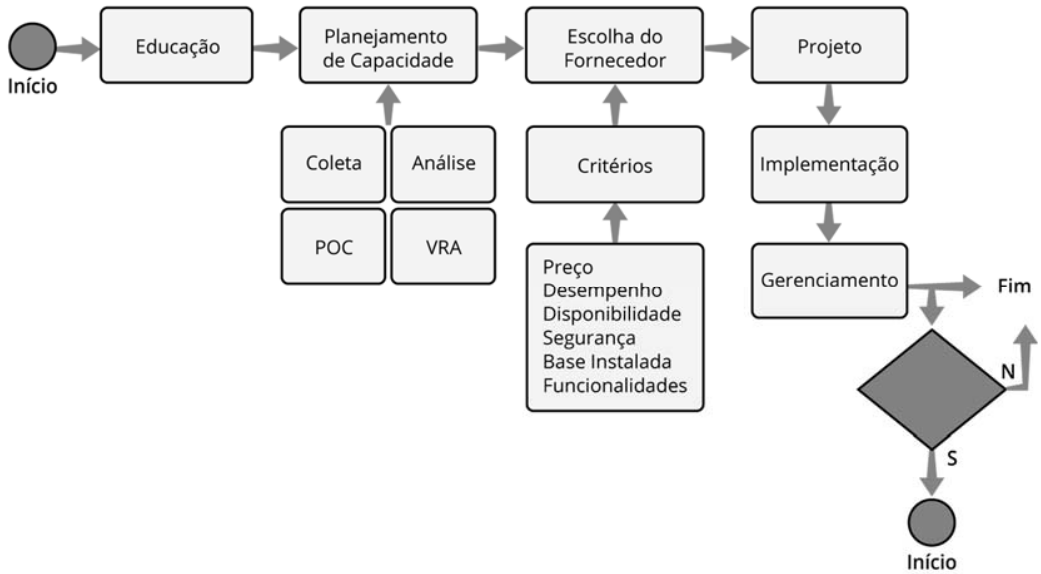


Figura 2.1 Etapas de um projeto de virtualização.

Educação e conscientização

A solução de virtualização de servidores pode incluir tecnologias avançadas, como tolerância a falhas, balanceamento de carga e redução do consumo de energia. Porém, todas têm um custo associado em termos de complexidade e recursos necessários – que aprimoram a operação da infraestrutura, reduzindo o tempo de interrupção e, consequentemente, melhoram os SLAs. No entanto, nem todos os projetos exigem sua aplicação.

Você deve procurar entender os objetivos do negócio, considerando que os diretores (reitores) e gerentes de infraestrutura possuem visões diferentes das necessidades, apesar de convergirem na necessidade de transformar a situação atual da infraestrutura de TI de suas organizações.

As justificativas para a elaboração de projetos que contemplem a virtualização incluem dificuldades atuais: provisão de novos servidores, operacionais, consumo de energia, mau uso dos recursos, falta de estratégia para alta disponibilidade e recuperação a desastres, aplicações legadas e novos hardwares.

Os desafios são diversos, como alterar a infraestrutura existente enquanto a organização continua suas atividades durante o período de ajustes. A meta é realizar a mudança para o novo ambiente com o mínimo de paralisações das operações e de limitação dos recursos humanos e físicos.

As organizações avançam na busca de soluções não proprietárias que simplifiquem o ambiente de infraestrutura de TI. Cerca de 90% dos servidores vendidos atualmente são baseados em plataforma x86 – uma tendência de mercado. A virtualização permite utilizar melhor os recursos de hardware, incluindo aperfeiçoar o uso de processador e a memória dos servidores. Além disso, é preciso reconhecer a importância da sua preparação antes de iniciar o projeto. Invista na sua educação.

Planejamento da capacidade

- Etapa fundamental para o planejamento da nova infraestrutura virtualizada
- Identificação dos servidores compatíveis
- Seleção dos serviços indicados para virtualização
- Coleta, análise, testes e elaboração do relatório

O planejamento da capacidade (Capacity Planning) em projetos de virtualização é um dos principais passos a serem realizados durante o projeto de virtualização. Planejar a infraestrutura a ser consolidada em um projeto requer cuidados, considerando a diversidade de servidores, os sistemas operacionais e cargas (workloads). Para a determinação da carga de trabalho, é fundamental a seleção de parâmetros de execução que tenham impacto no sistema. Em geral, o uso de recursos é muito baixo, o que permite aos projetos de virtualização uma alta taxa de consolidação. Podemos dividir o planejamento de capacidade nas fases descritas a seguir.

Coleta

Nesta primeira fase, devemos listar os hardwares e sistemas operacionais instalados, coletar o desempenho dos serviços, identificando o consumo de processamento, memória, disco e rede. É indicado realizar a coleta durante o período de pelo menos um mês, para uma cobertura completa dos eventos da organização. O resultado inclui o inventário dos servidores físicos na sua instalação e as métricas de desempenho de cada serviço.

Análise

Os dados coletados na fase anterior são comparados com os parâmetros predefinidos, para identificar os servidores, sistemas operacionais e serviços compatíveis com a virtualização. Esta etapa inclui uma análise do hardware necessário para suportar servidores virtualizados, análise do desempenho dos serviços e das necessidades de failover e expansão de curto prazo.

Testes

Com uma ferramenta de apoio ao planejamento de capacidade, o mais prudente é desenvolver uma prova de conceito (Proof Of Concept – POC) do ambiente a ser virtualizado, junto com uma análise do custo total do investimento. O resultado será uma solução que deve ser discutida com o cliente e, depois, refinada.

Elaboração do relatório

O relatório de planejamento de capacidade precisa incluir servidores compatíveis com a virtualização, os serviços indicados para o processo (número inicial de máquinas virtuais), a memória virtual necessária, o total de máquinas virtuais suportadas. Também deverá indicar as necessidades de treinamento da equipe. Esse tipo de relatório é conhecido como Avaliação de Disponibilidade para Virtualização (Virtualization Readiness Assessment – VRA).

Coleta e levantamento

Deve-se considerar o levantamento no caso de projeto para uma infraestrutura existente. As funcionalidades escolhidas para a virtualização definirão os recursos aproveitados da antiga infraestrutura. É preciso definir uma meta de utilização para os processadores, lembrando que o uso de determinados recursos, como a alta disponibilidade e o balanceamento dinâmico das cargas de trabalho, obrigam a nova infraestrutura a ter uma folga, tanto nos recursos de processamento quanto nos de memória. O ideal é trabalhar a infraestrutura nova para que ela funcione à média carga (em torno de 50% a 80%).

Informações importantes para levantamento sobre a infraestrutura existente:

- Número de servidores;
- Processamento total dos servidores;
- Memória total dos servidores;
- Utilização de banda por servidor;
- Espaço total de disco dos servidores;
- Performance dos processadores, memória e disco.

Análise

Este processo é fundamental porque o host tem uma quantidade limitada de recursos para serem compartilhados (processamento e memória), e as máquinas virtuais em execução devem realizar a ação sem comprometer o desempenho do sistema.

Com a introdução do hipervisor – entre a máquina virtual e o host – e do seu papel ativo no gerenciamento da memória, agendamento de processamento (schedule) e I/O, se faz necessária uma etapa de análise da performance desse serviço candidato à virtualização. Não podemos voltar aos primórdios do processo quando se ouvia que “os servidores Exchange não devem ser virtualizados” ou “você nunca deve virtualizar um servidor de SQL”. A análise não pode ser superficial. Um servidor de Exchange com um número muito grande de clientes – e uma enorme caixa postal – pode não ser um bom candidato. No entanto, o mesmo servidor com poucos clientes e caixas pequenas pode ser indicado para a virtualização. A análise deve focar no desempenho real do servidor; não na sua função. Durante as atividades práticas deste curso, optamos por omitir os tipos dos serviços providos por cada servidor, para deixar mais clara a sua independência na determinação da solução virtualizada.

Após a coleta dos dados e elaboração do inventário dos servidores físicos em seu ambiente, devemos partir para a fase de análise:

- Análise de desempenho quantitativo de cada servidor;
- Análise do hardware necessário para suportar servidores virtualizados;
- Análise das necessidades de tolerância a falhas.

Prova de conceito

A Prova de Conceito (Proof Of Concept – POC) foi concebida para prover as organizações com conhecimento das soluções de infraestrutura virtual e demonstrar as capacidades da tecnologia existente no software, além do modo como são aplicadas para resolver problemas reais.

Geralmente, a Prova de Conceito dura 30 dias e necessita do envolvimento de um engenheiro de sistemas da fornecedora de software ou de um parceiro, para auxiliar a:

- Instalar e configurar o software;
- Definir as melhores práticas para o desenho da solução, implementação e utilização;
- Testar as características e funcionalidades do software e como ele se aplica a diferentes cenários de utilização;
- Validar que a solução proposta esteja adequada às necessidades do negócio da organização.

Existem quatro estágios durante uma Prova de Conceito, concebidos para garantir que todos os requerimentos estejam definidos, com a identificação dos recursos e estabelecimento das expectativas. Isso para garantir que seja finalizada no tempo previsto e com sucesso.

- **Fase I – Levantamento**: definição dos requerimentos e planejamento do projeto; durante esta fase é necessário conduzir a descrição do escopo do projeto com os objetivos e requerimentos para a Prova de Conceito.
- **Fase II – Planejamento**: inclui o desenho da solução, o período de execução e o plano de testes; tipicamente trabalha-se junto ao cliente para definir o tempo de realização da Prova de Conceito, os cenários de utilização e os critérios de teste e avaliação para testar as características/funcionalidades da solução proposta.
- **Fase III – Construção**: instalação, configuração e teste, composta de dois estágios: a instalação ou criação da infraestrutura virtual e o teste/avaliação da solução. Após a fase de testes, segue a de validação, em que o cliente conduzirá novos testes para verificar se a solução atende às necessidades.
- **Fase IV – Gerenciamento**: contempla os objetivos alcançados, o refinamento da solução e os próximos passos; a fase final da Prova de Conceito é necessária para avaliar se a proposta atende aos requisitos de negócios do cliente. Nesta fase, o cliente já deve ter uma ideia clara das capacidades e usabilidade da solução, e também do impacto financeiro de sua adoção.

O Relatório de Impacto do POC é um documento elaborado pelo consultor técnico do parceiro/engenheiro de sistemas da fornecedora de software, que inclui informações técnicas referentes aos negócios e sobre o impacto da solução no ambiente do cliente.

Avaliação de disponibilidade para virtualização

O relatório de avaliação de disponibilidade para virtualização, ou Virtualization Readiness Assessment (VRA), descreve detalhadamente o impacto e os benefícios que serão alcançados após a consolidação dos servidores – aponta as melhorias proporcionadas na gestão dos serviços, na maximização da produtividade e na simplificação do processo de implantação de novos serviços. Deverá contemplar a estimativa financeira sobre o investimento aplicado na aquisição de softwares, hardwares e gastos com produtos para manter a infraestrutura, conhecida como Custo Total do Investimento ou Total Cost of Ownership (TCO). Também terá que contemplar o retorno previsto sobre o investimento, Return On Investment (ROI). O relatório gerado servirá como apoio para o gerente de TI na justificativa de investimento no projeto de virtualização.

Ferramentas

Ao entender a complexidade dessas tarefas, alguns fabricantes disponibilizaram ferramentas que permitem fazer o inventário do ambiente de TI – incluindo servidores, dispositivos de armazenamento de dados e rede, além de fornecer sugestões para o novo ambiente virtualizado, visando à otimização da nova infraestrutura. As ferramentas de planejamento da capacidade combinam o inventário com as informações de desempenho. Os fabricantes oferecem meios de permitir o correto planejamento da capacidade, com base na utilização da infraestrutura existente.

① Exemplos de ferramentas para o planejamento de capacidade:
VMware Capacity Planner Tool
Microsoft Assessment and Planning Toolkit
Novell PlateSpin Recon
CiRBA Data Center Intelligence
Lanamark

Escolha do fornecedor

- Preço
- Desempenho (densidade)
- Disponibilidade
- Segurança
- Base instalada
- Funcionalidades

Os critérios de escolha passam por desempenho do software de virtualização, disponibilidade, segurança, base atual instalada e funcionalidades ofertadas. A decisão quase sempre recai entre três opções de fornecedores: VMware, Microsoft e Citrix, cujas plataformas serão tratadas neste curso.

Escolhido o fornecedor, é necessário definir a infraestrutura de suporte ao novo projeto. A infraestrutura pode já existir, e, nesse caso, deverá passar por grandes ajustes. Se o projeto for novo, será preciso simular as cargas das aplicações de TI no ambiente virtualizado. Com a

infraestrutura existente, é mais comum fazer um levantamento da situação de carga encontrada para a definição da infraestrutura virtualizada. Fabricantes como a VMware e a Microsoft disponibilizam ferramentas que permitem avaliar a situação da infraestrutura existente em termos de processamento, memória, I/O e rede dos servidores.

Para a criação da infraestrutura, deve-se utilizar um software de simulação com uma situação real de carga. Essa opção, em geral, é complexa, e as ferramentas existentes são proprietárias. Outra possibilidade é simplificar o projeto, por meio de um cenário virtualizado baseado na experiência e em dados secundários existentes. Assim, foge-se da simulação no caso de uma nova infraestrutura e do levantamento no caso de infraestrutura existente. O Dell Virtualization Advisor é uma boa ferramenta para estimar a infraestrutura necessária. Os advisors podem ser utilizados para projetos de infraestrutura existente e para nova infraestrutura.

⚠ Recomenda-se deixar os servidores de firewall e servidores da DeMilitarized Zone (DMZ) separados dos servidores de virtualização usados na rede interna da organização.

Projeto

Como em qualquer projeto, é necessário utilizar uma ferramenta básica de gerenciamento que permita controlar o cumprimento dos prazos e os recursos para a realização das duas principais entregas descritas a seguir, além de estabelecer uma estratégia clara de comunicação.

- **Especificações do sistema**: devem estar incluídas as especificações detalhadas, os requisitos de licenciamento e de backup, implicações da virtualização na rede, nos dispositivos de armazenamento etc.
- **Plano de implementação**: precisa contemplar a verificação dos recursos, construção da infraestrutura, migrações e revisão.

O funcionamento da nova infraestrutura tem que ser cuidadosamente planejado, envolvendo testes das principais aplicações antes da entrada em produção, para verificar a operação no ambiente virtualizado. Por mais que os serviços de implementação tenham sido contratados, é preciso pensar no treinamento de pelo menos um profissional interno para a diminuição dos riscos com a nova operação. Essa opção deve fazer parte de um plano maior de estratégia para a redução de riscos.

Implementação

A implementação depende de detalhes específicos de cada infraestrutura e das plataformas disponíveis. As arquiteturas básicas serão apresentadas adiante de forma genérica e posteriormente detalhadas para cada software de virtualização.

Dicas importantes para servidores:

- Deverão ser utilizados servidores homologados para o software de virtualização;

- Definição das partições para o servidor virtualizado;
- Definição da forma de licenciamento do software de virtualização;
- Configurações suportadas para o software de gerenciamento;
- Migração: para converter um servidor físico em máquina virtual, podem ser usadas ferramentas dos próprios fabricantes. Esse procedimento é denominado Physical to Virtual (P2V);
- Projeto de arquitetura incluindo os servidores, armazenamento e backup.

Dicas importantes para dispositivos de armazenamento (storage):

- Em geral o próprio software de virtualização realiza o **failover** e o balanceamento de carga;
- Definição da necessidade de uso de zoneamento no storage para efeito de segurança;
- Definição da rede a ser utilizada e funcionalidades necessárias;
- Definição do nível de integração do backup.

Failover

Processo no qual uma máquina assume os serviços executados por outra, quando esta última apresenta algum tipo de falha. Pode ser feito de forma automática ou manual.

A utilização de recursos de alta disponibilidade exige que os servidores tenham processadores compatíveis para permitir que uma aplicação (ou sistema) execute em qualquer servidor do pool. É preciso, ainda, assegurar que o software de alta disponibilidade seja homologado para o storage empregado. A recuperação de desastres depende da integração do software de replicação aplicado no storage e da funcionalidade do software de virtualização.

Gerenciamento

Em geral, é utilizado um software de gerenciamento específico fornecido pelo fabricante do software de virtualização, que pode ser integrado a um framework de gerenciamento oferecido por alguns fabricantes.

Infraestrutura de hardware

- Arquitetura física de um datacenter
- Arquitetura virtual de um datacenter
- Hardware – descrição

Os datacenters são os principais beneficiados pela virtualização – caracterizados pelo grande conjunto de computadores (servidores, em especial) e por atender a necessidades específicas de vários clientes em relação a sistemas operacionais, aplicativos, questões de segurança e disponibilidade, entre outros. Sem o emprego da virtualização, é bastante comum o emprego da filosofia de “um serviço por servidor”, ou, na melhor das hipóteses, por clientes. Isso implica grande quantidade de servidores, elevando os custos de manutenção, consumo de energia e equipe de TI, entre outros. Além disso, não há carga de trabalho suficiente para explorar todo o poder de processamento disponível, ou seja, há um

desperdício de processamento. Com a virtualização, é possível instanciar múltiplas máquinas virtuais em um único servidor físico, reduzindo a quantidade de máquinas e, por consequência, o custo de manutenção da infraestrutura física.

Outro aspecto bastante comum em um datacenter é a necessidade de readequar dinamicamente a capacidade de armazenamento com o decorrer do tempo. Sem o emprego da virtualização, esse problema é resolvido com a aquisição de novos discos, com o seu particionamento, a instalação de um sistema de arquivos específico e a configuração desse novo disco em um Sistema Operacional existente. Novamente, os discos podem ser superdimensionados, gerando um desperdício de espaço de armazenamento, além de aumentar o custo de gerenciamento e manutenção. A virtualização permite que o espaço em disco também seja virtualizado por meio da criação de discos virtuais, que podem ser inseridos, removidos e configurados com maior facilidade em um sistema. A capacidade de armazenamento, física ou virtual, é o que se denomina storage. Ainda, tipicamente, os storages são acessíveis aos servidores por meio de redes de alto desempenho.

Os problemas do desperdício de capacidade de processamento e armazenamento, assim como a necessidade de redes de interligação, também existem em datacenters pertencentes a uma única organização. O mesmo ocorre em servidores corporativos que não configurem necessariamente um datacenter. A virtualização surge como uma solução comum aos dois casos, sendo genericamente chamada de virtualização de servidores.

Arquitetura física do datacenter

- Recursos de processamento e memória, incluindo servidores e clusters
- Recursos de storage
- Recursos de rede

O datacenter físico pode ser representado pelo diagrama mostrado na próxima figura. As aplicações executam de maneira isolada em diferentes servidores. O que acontece na prática é que determinada aplicação precisa de mais recursos de hardware, enquanto outra tem recursos sobrando. Com essa arquitetura não é possível fazer muita coisa. Os recursos de processamento e memória acabam sendo mal utilizados, implicando alto TCO para o datacenter. Além disso, a grande disponibilidade precisa ser pensada para cada aplicação, onerando o projeto como um todo.

Principais elementos da arquitetura do datacenter físico:

- Recursos de processamento e memória, incluindo hosts (servidores) e clusters;
- Recursos de storage;
- Recursos de rede.

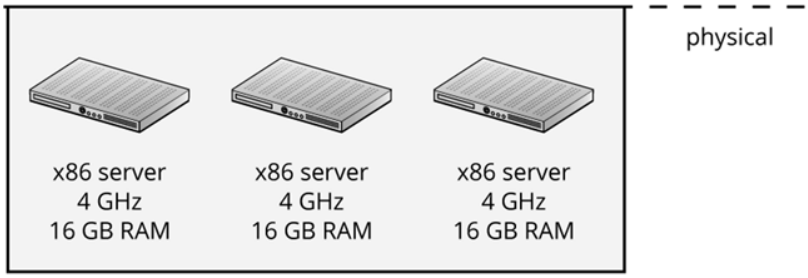


Figura 2.2 Datacenter físico.

Arquitetura virtual do datacenter

- Recursos de processamento e memória
- Recursos de storage e datastores
- Recursos de rede
- Máquinas virtuais (MVs)

A virtualização possibilita otimizar o uso da infraestrutura de TI, incluindo servidores (hosts), storage e os dispositivos de rede. Ela agrega os vários recursos e apresenta um conjunto simples e uniforme de elementos em ambiente virtual. O datacenter virtual pode, então, ser provisionado para o negócio com o conceito de infraestrutura compartilhada.

Principais elementos da arquitetura do datacenter virtual:

- Recursos de processamento e memória, incluindo servidores, clusters e Resource Polls (RPs) ou grupos de recursos;
- Recursos de storage e datastores;
- Recursos de rede;
- Máquinas virtuais propriamente ditas.

Os servidores (hosts) representam os recursos virtuais de processamento e memória de um servidor físico que esteja rodando o software de virtualização. O servidor é o local em que estão hospedadas as máquinas virtuais.

Já os clusters são um conjunto de servidores que possibilita agregar dinamicamente recursos de processamento e memória de forma coordenada, permitindo funcionalidades que não são obtidas com um host isolado. Os recursos de processamento e memória de servidores e clusters podem ser particionados em uma hierarquia de grupos de recursos.

- **Storage**: local onde ficam armazenados os dados, aplicações e informações de configuração.
- **Datastores**: representações virtuais de combinações de recursos físicos de storage.

As máquinas virtuais são associadas a um servidor particular, a um cluster, a um grupo de recursos e a um datastore. Provisionar máquinas virtuais é mais simples do que provisionar servidores físicos. Os recursos são provisionados para as máquinas virtuais com base nas

políticas definidas pelo administrador de sistemas, conferindo flexibilidade ao ambiente virtual. Os grupos de recursos podem ser reservados para uma máquina virtual específica, por exemplo. O resource pool é um recurso virtual que permite alocar recursos específicos para as aplicações com funcionamento dentro do cluster das máquinas virtuais.

A figura a seguir ilustra o datacenter virtual. Os recursos de hardware passam a fazer parte de verdadeiros pools de recursos utilizados de acordo com os níveis de serviço requeridos pelas aplicações.

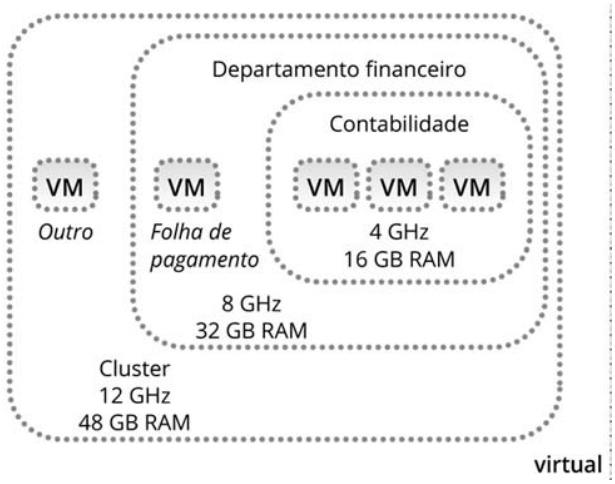


Figura 2.3 Datacenter virtual.

Hardware – descrição

Os principais blocos de construção da infraestrutura de hardware são:

- Servidores
- Storage
- Servidor de gerenciamento e backup
- Infraestrutura de redes – switches

Além disso, é necessário considerar no projeto os aspectos de alta disponibilidade, recuperação de desastres e segurança.

Servidores

- Arquitetura do processador
  - CISC vs RISC
  - Servidores x86
- Benchmarks
- Padrões da indústria
- Caraterísticas dos servidores

- Disponibilidade
- Desempenho
- Gerenciamento
- Escalabilidade

Arquitetura do processador

Os servidores utilizados nos datacenters variam desde mainframes – cujos processadores são baseados em instruções Reduced Instruction Set Computer (RISC) – até x86, que seguem o Complex Instruction Set Computer (CISC).

Os servidores do tipo CISC contêm um grande número de instruções. Já os modelos baseados em RISC deixam para o Sistema Operacional boa parte da execução das instruções e, portanto, têm um número reduzido delas. Os servidores em CISC – especificamente os modelos com a arquitetura x86 – são considerados um padrão da indústria e utilizados em larga escala.

Os servidores x86 são o foco atual da virtualização. Eles estão evoluindo rapidamente em termos de confiabilidade e poder de processamento, de modo seu hoje é dominante, nos mais diversos ambientes computacionais. A arquitetura desses servidores foi modificada recentemente com a introdução dos servidores de lâminas (blades).

A arquitetura de servidores baseados no processador x86 evoluiu ao longo do tempo. À medida que novos processadores com maior poder de processamento são lançados, memória e barramento local (E/S) naturalmente acompanham essa evolução, para garantir um sistema balanceado e sem gargalos.

A figura a seguir ilustra a arquitetura de um servidor padrão x86 baseado em processador Intel, com os principais componentes e suas interligações. Os componentes principais ilustrados são o processador, a memória e o barramento entre o processador e memória, conhecido como Front Side Bus (FSB), e o barramento de Entrada/Saída (E/S). As conexões de placas de rede (Network Interface Card – NIC), controladora de discos RAID e demais periféricos são normalmente realizadas pelo barramento de E/S.

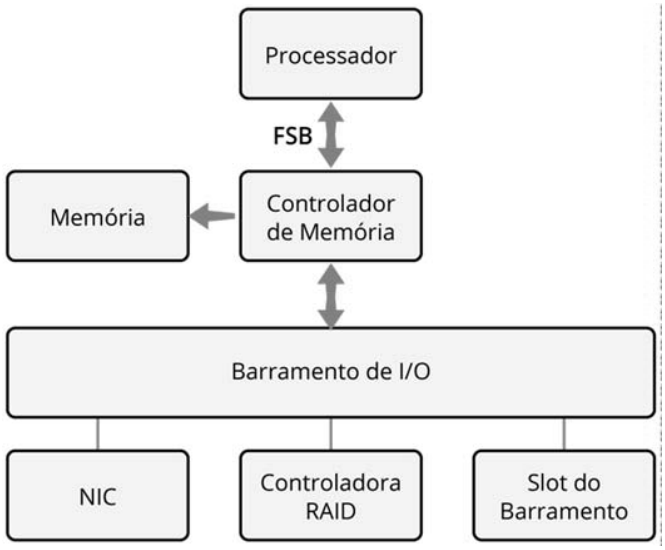


Figura 2.4 Arquitetura de servidor x86.

A arquitetura x86 continua a evoluir, mas mantém instruções derivadas da arquitetura 8086 de 16 bits. O que, talvez, seja o grande segredo da aceitação dessa arquitetura: a compatibilidade com as versões anteriores.

Existem duas principais arquiteturas x86: 32 bits e 64 bits, que a Intel convencionou chamar de (IA) 32 e (IA) 64. Por sua vez, a microarquitetura refere-se ao projeto físico de cada processador. Processadores com diferentes microarquiteturas podem utilizar um mesmo conjunto de instruções, ou seja, a mesma arquitetura, o que acontece com os processadores Intel e AMD. A microarquitetura Intel desenvolveu-se, passando da opção Netburst para núcleo e, agora, para Nehalen. Os processadores AMD Opteron são também largamente empregados na indústria de servidores e apresentam excelente desempenho.

Benchmark de servidores

- Spec CPU 2006
- SPECvirt_sc2010
- TPC

A ideia do benchmark é a de comparar uma mesma configuração e condições de carga para duas CPUs de fabricantes distintos ou os resultados obtidos por CPUs diferentes de um mesmo fabricante. Existem diversos benchmarks para diferentes funções executadas pelos servidores. O mais simples é o orientado para o throughput da CPU, como o SPEC CPU 2006, cujos resultados para as CPUs de mercado podem ser obtidos no site do Standard Performance Evaluation Corporation (SPEC): www.spec.org

Especificamente para o mundo da virtualização existe um benchmark desenvolvido pelo SPEC, voltado para a avaliação de servidores (que pode ser encontrado em www.spec.org/virt_sc2010/). Esse benchmark faz a medição fim a fim do desempenho de

todos os componentes (incluindo hardware e plataforma de virtualização) e do ambiente da máquina virtual (com Sistema Operacional e aplicações).

Há também os benchmarks para processamento de transações, que medem a habilidade de um sistema. Eles consistem em verificar o acesso a sistemas de banco de dados e atualizações. Na década de 1980, um grupo de engenheiros criou o Transaction Processing Performance Council (TPC - www.tpc.org). Trata-se de uma organização independente de fornecedores, com o intuito de criar benchmarks equilibrados e realistas para o processamento de transações on-line. Uma visita ao site é interessante para observar os resultados das diversas avaliações.

Padrões da indústria

- Servidores
  - Torre
  - Rack
  - Lâmina (blade)
- I/O
- Gerenciamento
- Chassis e racks

A utilização de servidores baseados na plataforma x86 ganhou impulso com o avanço da arquitetura cliente/servidor no início da década de 1990. Eles utilizam hardware altamente confiável e com poder de processamento elevado pela rápida evolução tecnológica, possibilitando seu uso em quase todo tipo de aplicação. Ainda existem limitações para determinadas cargas de trabalho, mas, de forma geral, os servidores x86 atendem às demandas do mundo corporativo.

Os servidores-padrão da indústria evoluíram ao longo dos anos. Os primeiros eram do tipo torre e atendiam aos propósitos básicos de entregar serviços de rede. Posteriormente, surgiram os servidores em rack, ideais para uso nos datacenters, por serem mais compactos e fáceis de gerenciar. Atualmente, os servidores em rack possuem diversos tamanhos e capacidades de processamento. Os racks são descritos em termos de unidades 1U (equivale a 44,45 mm). Os servidores de 1U e 2U de altura são os tamanhos mais comuns de servidores.

Mais recentemente, surgiram os servidores blades, que aperfeiçoam ainda mais o uso do espaço físico, simplificam o gerenciamento e, normalmente, consomem menos energia quando comparados aos servidores em rack. A altura do chassi típico é de 10 Us.

Os racks utilizados para acomodar os servidores e os chassis onde são instalados os blades são baseados em padrões ditados por normas. Suas dimensões de altura e largura são comuns entre os diversos fabricantes, com algumas diferenças na profundidade. Os racks costumam ter um KVM (Keyboard, Video e Mouse), que pode oferecer conectividade IP e a gerência dos servidores com um único console de gerenciamento. Há duas frentes para o gerenciamento das blades: o do próprio chassi e o das lâminas, que também pode e deve ser feito por meio de interfaces web.

As conexões de E/S são a parte crítica de qualquer projeto de blades, pois elas que permitem a conexão de servidores blades com o mundo externo. Em geral, esses servidores oferecem três tipos de conectividade: IP/Ethernet, Fibre Channel (FC) e Infiniband. A próxima figura exemplifica um conjunto de blades e suas conexões.

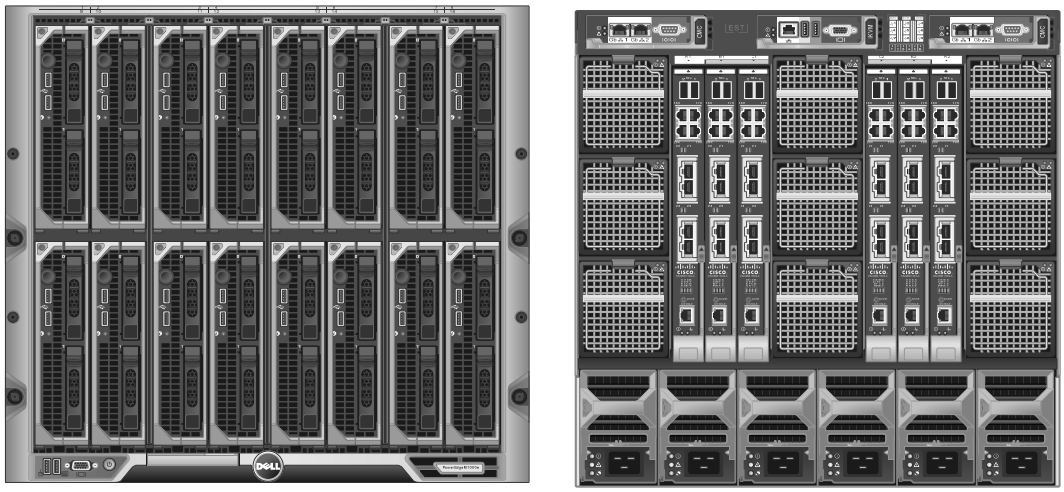


Figura 2.5 – Servidores blade e conexões de I/O.

Características dos servidores

- Disponibilidade
- Desempenho
- Gerenciamento
- Escalabilidade
  - Vertical (scale-out)
  - Horizontal (scale-up)

Disponibilidade

De forma geral, os servidores melhoraram a disponibilidade com o uso de fontes e ventiladores redundantes e discos em configuração de Redundant Storage of Inexpensive Drives (RAID). Diversas pesquisas realizadas por fabricantes demonstram que esses dois aspectos são críticos. Placas de rede duplicadas e a utilização de mais de um processador no servidor aumentam ainda mais o nível de disponibilidade. Hoje, as memórias também permitem o uso de configurações redundantes. A alta disponibilidade do servidor é obtida com a configuração em cluster de pelo menos dois servidores.

Desempenho

O desempenho de servidores x86 tem avançado rapidamente. Atualmente, existe também a possibilidade de utilizar processadores mais econômicos do ponto de vista energético, em

detrimento do desempenho. Uma comparação rápida de resultados obtidos com os benchmarks SPEC ou mesmo o TPC indicam claramente o ganho de desempenho conseguido a cada nova atualização dos processadores.

Gerenciamento

O gerenciamento do hardware do servidor é realizado por meio de um software de gerenciamento, normalmente fornecido pelo próprio fabricante do servidor, cujo console é baseado na web. O software costuma ser instalado em uma estação de gerenciamento e utiliza o protocolo Simple Network Management Protocol (SNMP). Também é possível gerenciar o servidor por uma placa de gerenciamento integrada, que permite o acesso remoto ao servidor.

Escalabilidade

Escalabilidade é a habilidade de um sistema computacional de lidar, de forma transparente, com um número crescente de usuários ao mesmo tempo. A escalabilidade em servidores é conseguida com a adição de mais processadores.

Os métodos tradicionais de escalabilidade em servidores são a vertical (scale-up) e a horizontal (scale-out).

- **Escalabilidade vertical**: utiliza múltiplos processadores e adiciona recursos em um único nó do sistema (mais memória ou mais processamento, por exemplo). Scale-up é uma arquitetura de hardware otimizada pelo software. O principal elemento que viabiliza esta modalidade é o multiprocessamento simétrico (SMP), que apresenta limites de linearidade em certas situações, dependendo da arquitetura da aplicação. Normalmente, exige maior investimento inicial e menor flexibilidade para a introdução de novas tecnologias.
- **Escalabilidade horizontal**: usa múltiplos servidores e adiciona mais nós ao sistema para crescimento, como um novo servidor a um sistema de banco de dados em cluster. Scale-out é uma arquitetura de software otimizada pelo hardware. A modalidade costuma apresentar menor investimento inicial e é mais flexível para a introdução de novas tecnologias.

Em sistemas x86, é mais comum a escalabilidade do tipo horizontal com um software de clustering. A escalabilidade do tipo vertical é padrão em sistemas RISC. O licenciamento da camada de software é sempre um aspecto importante a ser considerado na decisão por uma ou outra forma de crescimento do servidor.

As configurações de servidores para a virtualização variam entre as modalidades scale-out e scale-up. Ou seja, a virtualização pode ser adotada em pequenos servidores que crescem horizontalmente (aumento do número de servidores) ou em grandes servidores com crescimento vertical (aumento do número de processadores).

Os servidores respondem diretamente pela virtualização e podemos destacar os servidores do tipo rack e os do tipo blade, que são adequados para o processo. Os blades

especificamente mudam o perfil da densidade de energia dentro do datacenter e acarretam mudanças na estratégia de energizá-lo e refrigerá-lo.

Dispositivos de armazenamento – storage

- Histórico
- Situação atual
- Papel do storage no datacenter

O storage é hoje um componente fundamental da infraestrutura de TI, responsável direto pelo nível dos serviços de armazenamento fornecidos por um datacenter. A necessidade de suportar o crescimento da massa de dados digital e, ao mesmo tempo, de aumentar a confiabilidade das informações – devido a aspectos regulatórios e de operar cada vez mais em um sistema 24 horas por dia, 7 dias por semana (24 por 7) –, faz do storage o ponto focal de muitos projetos de virtualização da infraestrutura de TI. O storage responde pelo requisito de entrada/saída do sistema computacional.

O sistema de E/S sempre foi negligenciado na arquitetura de um sistema de computador e, até hoje, representa um gargalo. A evolução contínua dos processadores tem provocado o aumento da diferença de desempenho existente entre processadores e sistemas de armazenamento. Por exemplo, os discos mecânicos – de uso padrão na maioria das instalações de storage – impõem limites ao desempenho do sistema de E/S, devido às características mecânicas e rotacionais desse tipo de dispositivo quando em operação de leitura e escrita. Na prática, o processador fica esperando a operação de E/S e prejudica o desempenho das aplicações e do sistema como um todo. Recentemente, surgiu uma opção para substituir os discos mecânicos: os discos baseados em memória flash, com o nome de Solid State Disk (SSD). Ainda tem custo elevado, mas deve ser rapidamente adotada, porque melhora significativamente o desempenho do sistema de E/S.

Sistemas de armazenamento

- Servidores
- Storage
- Conectividade

O sistema de armazenamento é composto por diversos componentes, onde devemos considerar o servidor, o storage e a conectividade.

Servidores

Os usuários armazenam e recuperam dados por meio das aplicações, que rodam nos servidores. Estes são compostos por componentes físicos (hardware) e lógicos (software).

Componentes lógicos do servidor:

- Sistema Operacional: controla todos os aspectos do ambiente computacional;

- Device drivers: softwares especiais que permitem ao SO interagir com dispositivos específicos;
- Volume manager: software que roda no servidor e faz a interface entre o sistema de arquivos e o disco físico;
- Sistema de arquivos: relaciona-se à estrutura hierárquica de arquivos;
- Aplicação: interface entre o usuário e o servidor. O acesso aos dados pode ser feito por blocos ou por arquivos. O acesso por blocos é o mecanismo básico de acesso aos discos; por sua vez, o acesso por arquivos é uma abstração do acesso por blocos.

Storage

Componente principal do sistema de armazenamento, pode utilizar um meio magnético ou de estado sólido. Discos e fitas usam o meio magnético. Discos ópticos adotam um meio de estado sólido.

Conectividade

Refere-se à interconexão entre o servidor e o dispositivo de armazenamento (storage). A conectividade possui componentes físicos (hardware) e lógicos (protocolos).

Características do servidor de discos

- Como funciona
- Conexão direta
- Conexão via rede

Diferente de um servidor de arquivos, um storage pode ser visualizado como um servidor de discos. Quando conectado ao storage, um servidor só enxerga os discos no storage e utiliza o sistema de arquivos fornecido pelo próprio Sistema Operacional. A próxima figura ilustra a arquitetura básica de um storage com seus principais elementos (portas, cache, controladora e discos).

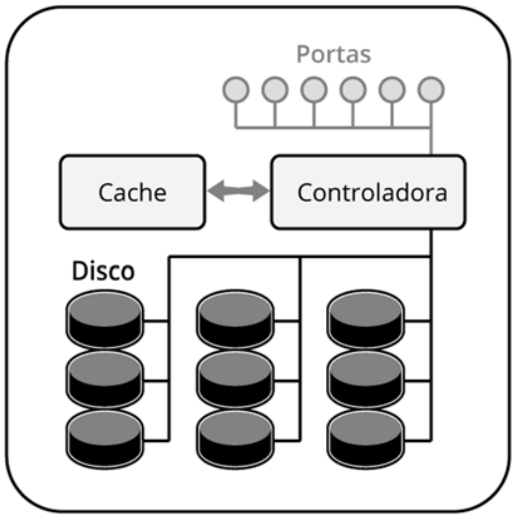


Figura 2.6 Arquitetura de um storage.

Os servidores são conectados ao storage localmente ou com o uso de tecnologias adequadas de comunicação em rede, tais como fibre channel ou iSCSI, permitindo acesso aos discos que compõem o storage. As portas de conexão são estendidas para os discos internamente, por meio de canais de E/S. Grandes subsistemas de discos podem ter diversas portas de conexão, controladoras redundantes e canais de E/S internos, com capacidade de armazenamento de vários terabytes de dados.

A conexão realizada via rede aumenta a quantidade de servidores a serem conectados ao storage e otimiza o uso dos discos. Outro aspecto importante do storage em rede é que o espaço em disco livre pode ser vinculado a qualquer servidor conectado ao storage, o que resolve o principal problema do emprego local de discos – no qual discos subutilizados não podem ser usados de maneira convencional por outros servidores.

Tipos de storage

- JOBD
- RAID
  - LUN
  - Unidade lógica
- Inteligente

O storage pode ser dividido em três tipos, como veremos a seguir.

JBOD

Se o subsistema de disco não possui controladora interna é considerado um Just a Bunch of Disks (JBOD). No caso de JBOD, as controladoras não têm a tecnologia RAID e fazem parte do servidor que estará conectado ao storage.

RAID

Redundant Array of Independent Disks (RAID) consiste em um conjunto de dois ou mais discos, com dois objetivos básicos: tornar mais rápido o acesso aos dados gravados em disco – para isso, é usada a técnica de divisão de dados (data stripping) –; e trazer mais segurança para os dados – nesse caso, adota-se a técnica de espelhamento (mirroring). Essas duas técnicas podem ser utilizadas isoladamente ou em conjunto. Os tipos de RAID mais empregados são: 0, 1, 2, 3, 4 e 5. Todos são basicamente a combinação dos RAID 0 e 1, acrescentando outras técnicas para melhorar o uso dos discos.

LUN

O Logical Unit Number (LUN) é o número usado em um storage para identificar uma unidade lógica – dispositivo endereçado pelo protocolo SCSI ou similares, como Fibre Channel ou iSCSI.

Unidade Lógica

A partir de um grupo de discos físicos, configurados em RAID, passamos a ter um disco único. Esse disco, então, pode ser dividido em vários "pedaços" de tamanho variável. Para cada um é associado um número, o Número da Unidade Lógica (LUN). A figura 2.7 mostra um grupamento RAID dividido em várias unidades lógicas e os respectivos LUNs.

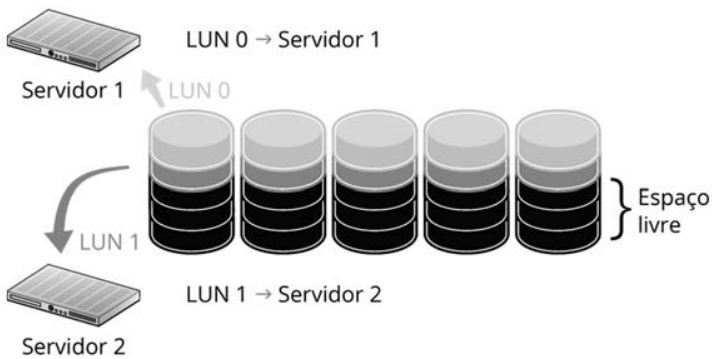


Figura 2.7 Partições/LUNs em um grupo RAID.

Essa mesma técnica aplica-se a um único disco, o que significa dividi-lo em unidades lógicas. A figura 2.8 ilustra a opção.

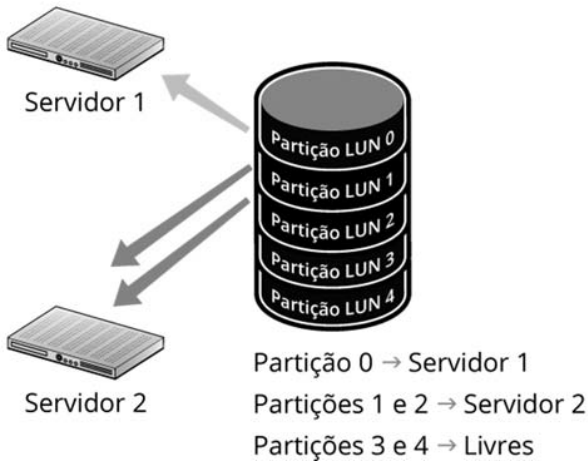


Figura 2.8 Partições/LUNs em um único disco.

Inteligente

Um sistema de armazenamento inteligente tem quatro componentes-chave: front-end, memória cache, back end e discos físicos.

- **Front-end**: interface entre o sistema de armazenamento e o servidor. Possui, normalmente, portas e controladoras de front-end. As portas executam os protocolos apropriados.
- **Memória cache**: nos subsistemas de disco, as caches são utilizadas para acelerar o acesso de leitura e escrita nos discos físicos. A cache existe no disco e nas controladoras. Todo disco possui uma pequena cache que serve para operações de escrita e leitura. A taxa de transferência do canal de I/O para a controladora é significativamente mais alta do que a velocidade que a controladora pode escrever ou ler do disco, por isso, também considera a ideia da cache.
- **Back-end**: interface entre a memória cache e os discos físicos. Formada por dois componentes: portas e controladoras de back-end.
- **Discos físicos**: os discos físicos são conectados ao back-end com interface SCSI ou Fibre Channel. O storage inteligente permite o uso de discos misturados como SCSI, FC ou SATA.

Redes de storage

- Necessidade
- DAS, NAS e SAN
- Protocolo SCSI

O crescimento da massa de dados organizacional e as novas demandas relativas à regulamentação trouxeram à tona a necessidade da utilização de unidades de storage

independentes do servidor. A maneira de conectar o storage evoluiu de soluções nas quais o storage era conectado diretamente ao servidor, Direct Attached Storage (DAS), para a criação de redes de storage independentes, que usam uma combinação de protocolos específicos e interfaces de discos, denominadas Storage Area Network (SAN) e Network Attached Storage (NAS).

A principal diferença entre SAN e NAS é: enquanto uma SAN oferece apenas um meio de armazenamento formado por blocos, sem criar um sistema de arquivos para eles, uma NAS fornece, além do meio físico de armazenamento, um sistema de arquivos. Apesar disso, SAN e NAS podem ser empregadas concomitantemente em uma única solução de storage, ou seja, uma parte do storage pode ser configurada para oferecer blocos para dados e outra, para blocos de dados e um sistema de arquivos.

O storage do tipo NAS é baseado em redes de storage IP e é primariamente utilizado para compartilhamento de arquivos. Quando comparado ao DAS é mais escalável, com melhor disponibilidade, além de mais fácil de gerenciar. Normalmente, o seu uso e gerenciamento requerem maior investimento inicial e conhecimento mais especializado.

NAS usa protocolos de rede e de compartilhamento de arquivos. Esses protocolos incluem o TCP/IP para a transferência de dados e CIFs e NFS para serviços de arquivo remoto. Usuários do Windows e do Unix podem compartilhar os mesmos dados armazenados em um servidor NAS, que é acessado por clientes e servidores em uma rede IP, e, muitas vezes, utiliza múltiplas interfaces de rede.

NAS tem seu próprio Sistema Operacional, para servir arquivos, ou seja, otimizar o I/O para diversos sistemas operacionais. Logicamente, o NAS pode servir a mais clientes do que um servidor de arquivos convencional.

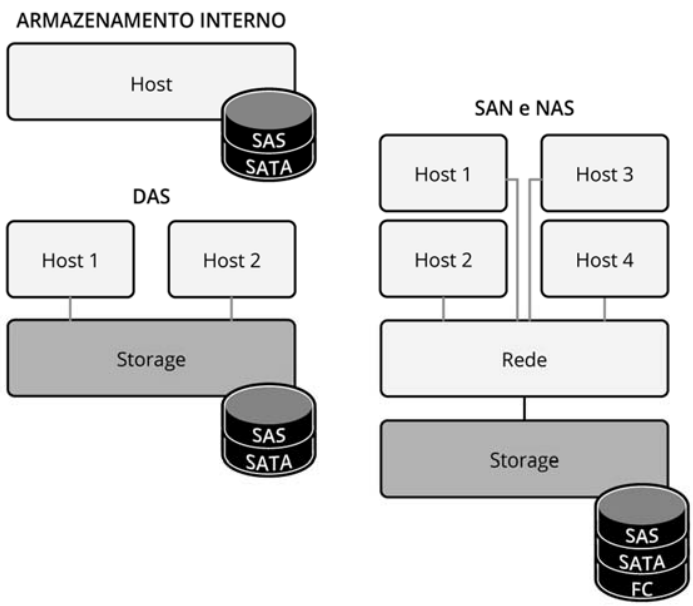


Figura 2.9 Técnicas de I/O.

Quando se compara uma rede do tipo SAN (ou NAS) a uma rede do tipo LAN é necessário entender algumas diferenças:

- O overhead de uma rede do tipo SAN é muito menor do que o de uma rede do tipo LAN;
- A proteção é uma característica muito mais necessária em uma rede LAN. O servidor acaba funcionando como um firewall para as redes SAN, o que simplifica a segurança;
- Um comportamento mais inteligente para o congestionamento na SAN é muito mais relevante do que na LAN, onde o TCP/IP simplesmente descarta o pacote. Com SAN, esse aspecto acaba sendo um problema.

A figura 2.9 esquematiza os principais tipos de técnicas de storage.

Protocolo SCSI

- Arquitetura do padrão SCSI-3
- Componentes da arquitetura
- Initiator
- Target
- Subsistema de entrega
- Serial Attached SCSI-SAS

O protocolo Small Computer System Interface (SCSI) foi desenvolvido para propiciar um mecanismo de transporte de dados eficiente entre os servidores e os periféricos, como discos e outros recursos. A arquitetura do padrão SCSI-3 (SCSI – SAM3) define um modelo cliente/servidor com um initiator (servidor), um target (disco) e um subsistema de transferência de dados, que pode ser um cabo paralelo, Fibre Channel ou iSCSI. A arquitetura do SCSI-3 envolve comandos específicos, os protocolos de transporte e a interconexão física que possibilita transferir os dados entre o initiator e o target.

Há também um tipo de storage baseado em protocolo SCSI, denominado Serial Attached SCSI (SAS). É um protocolo de gerenciamento e armazenamento de dados que passou a ser utilizado para fornecer storage a servidores.

SAN

- FC e FCoE
- IP

Nas redes SAN, a infraestrutura de rede pode ser Fibre Channel (FC) ou Gigabit Ethernet, enquanto os dados transportados são do tipo “bloco”. Nas redes NAS, a infraestrutura é quase sempre Gigabit Ethernet, e os dados armazenados, do tipo “arquivo”. O entendimento de quando utilizar uma ou outra infraestrutura é complexo e, muitas vezes, confuso, causando problemas quando soluções que deveriam ser baseadas em NAS são feitas em SAN e vice-versa.

O armazenamento do tipo Storage Area Network (SAN) é baseado em redes de armazenamento dedicadas e escaláveis, que conectam servidores e dispositivos de storage usualmente no nível de bloco (dados de aplicação).

ⓘ O protocolo SCSI continua sendo o padrão utilizado na comunicação entre o servidor e o storage. Na prática, protocolos como FC e iSCSI encapsulam os comandos SCSI dentro deles.

SAN FC e FCoE

- Conceito
- Protocolo FC
- Componentes da rede SAN FC

Os protocolos Fibre Channel (FC) foram rapidamente adotados como tecnologia viável para aplicações que tratam de blocos em nível de I/O. Fibre Channel simplifica as ligações entre servidores e dispositivos de storage, diminuindo a perda de sinal e aumentando as distâncias máximas permitidas quando comparado ao SCSI convencional. Importante ressaltar que o FC é um protocolo que pode usar fibra óptica ou cobre como meio de comunicação.

Inicialmente, a adoção das redes SAN baseadas no protocolo FC possibilitou a ligação dos discos SCSI aos servidores, aumentando a velocidade e o número de dispositivos permitidos. Também adicionou suporte para protocolos de várias camadas de alto nível, incluindo SCSI, Asynchronous Transfer Mode (ATM) e IP, sendo o SCSI o mais utilizado. Atualmente, com o surgimento de novas funcionalidades e dispositivos, as redes FC estão consolidadas.

As redes SAN FC possuem os seguintes componentes:

- **Portas dos nós** (nodes ports): nas redes FC, os dispositivos são chamados de nós. Cada nó é fonte ou destino da informação para um ou mais nós. Cada nó requer uma ou mais portas para se comunicar com outros nós.
- **Cabeamento**: SANs utilizam cabos de fibra óptica ou de cobre para pequenas distâncias.
- **Dispositivos de interconexão**: como hubs, switches e directors; hubs são pouco usados e compartilham a banda devido à transmissão de dados para todas as conexões. Switches são os componentes centrais de uma rede SAN e possuem função similar aos switches de LAN. Directors têm função similar aos switches, embora com um maior número de portas e mais robustez.
- **Unidades de armazenamento**: normalmente unidades inteligentes com diversas funcionalidades.
- **Software de gerenciamento da SAN**: gerencia a interface entre os servidores, dispositivos de interconexão e discos do storage.

Teoricamente, uma rede SAN baseada no protocolo Fibre Channel (FC) pode ter 15 milhões de nós. No caso do protocolo FC, os servidores utilizam placas do tipo Host Bus Adapter (HBA) para permitir a conexão à SAN.

O protocolo FC é o mais utilizado nas redes do tipo SAN e atende aos requisitos de desempenho e confiabilidade necessários, sendo implementado em cinco camadas. As implementações iniciais ofereciam throughput de 100MB/s (1062.5 Mb/s), que já era bem superior ao padrão Ultra SCSI de 20MB/s, usado em DAS até então.

O padrão FC pode operar em modo full-duplex, chegando, nesse caso, a 100 MB/s. Existe também o Fibre Channel over Ethernet, que nada mais é do que o encapsulamento dos quadros FC usando a rede Ethernet.

SAN IP

- Conceito
- Protocolo iSCSI
- Componentes da rede SAN IP

Recentemente, o padrão Internet SCSI (iSCSI), que adota o modelo Ethernet para comunicação na rede de storage, passou a ser uma opção confiável e mais barata. O protocolo iSCSI é um meio de transportar os pacotes SCSI pelo TCP/IP. Os principais componentes da rede iSCSI são servidores (initiators), targets e uma rede IP.

O protocolo iSCSI carrega os comandos SCSI do servidor (initiator) para os dispositivos target (storage). O iSCSI trabalha encapsulando os comandos SCSI dentro do TCP/IP e transportando-os por uma rede IP. O tráfego iSCSI pode utilizar dispositivos Ethernet, como switches e placas de rede (NICs) padrão, para escoar o tráfego de dados em blocos, como na SAN FC.

Em geral, o armazenamento em Internet SCSI (iSCSI) simplifica e reduz o custo de entrada em uma solução de storage. O mesmo ocorre com o TCO, quando comparado ao FC. A adoção do padrão Ethernet pelo protocolo iSCSI permite um ganho de aprendizado em relação ao protocolo FC. Outra vantagem do iSCSI é que ele pode ser incorporado a redes de armazenamento já existentes, baseadas em NAS e em SAN FC, não exigindo uma ruptura tecnológica, desde que o storage já possua portas iSCSI ou utilize um appliance. Com a chegada do padrão 10 Gb/s Ethernet, acredita-se que o iSCSI ganhará ainda mais espaço no mercado de redes de armazenamento. A rede iSCSI também trata de aplicações que usam blocos de dados como elementos padrão de I/O.

Storage e virtualização

- Interdependência
- SAN e NAS
- SAN FC e SAN IP
- Funcionalidades

Existe uma relação de interdependência entre o dispositivo de armazenamento e a virtualização. Boa parte das funcionalidades obtidas com a virtualização depende da consolidação do armazenamento em um único dispositivo. As redes de storage do tipo SAN e NAS são suportadas pelos softwares de virtualização, o que permite otimizar a utilização de recursos. Diversas funcionalidades oferecidas também dependem da plena interoperabilidade entre os dispositivos de storage e softwares de virtualização.

Infraestrutura de rede

- LAN
- Switch
- VLAN

Os dispositivos de rede como switches e roteadores são os elementos centrais da conectividade de qualquer rede. Os switches têm importância fundamental por serem os dispositivos normalmente empregados para interconectar os vários equipamentos (servidores, storages e mesmo os roteadores) de uma infraestrutura de rede local (LAN) de uma organização. Os switches estão em todo lugar, inclusive dentro das blades.

O funcionamento de um switch é baseado na construção de uma tabela de encaminhamento contendo os endereços MAC de todos os dispositivos conectados a cada porta do switch. Quando um quadro de dados chega a uma porta, o MAC de destino do dispositivo é verificado, e o quadro é enviado somente para a porta a qual foi destinado. Em um switch, o encaminhamento dos quadros é feito por um hardware específico conhecido como Application Specific Integrated Circuits (ASICS).

Um switch proporciona alta taxa de transferência com baixa latência.

A alta taxa de transferência em um switch é dada em bits por segundo ou pacotes por segundo (PPS), dependendo da tecnologia utilizada. Uma característica interessante dos switches é a capacidade de oferecer redes locais virtuais (Virtual Local Area Network – VLAN).

Uma Virtual LAN (VLAN) é um domínio de broadcast formado por portas de um ou mais switches. A VLAN é configurada no switch. Com o uso de VLANs, os administradores têm controle sobre portas e usuários. Dessa forma, é possível limitar o acesso à rede (por parte de usuários) a apenas algumas portas previamente configuradas. O administrador do datacenter controla cada porta e todos os recursos que serão permitidos.

Os grupos que formam as VLANs podem ser criados de acordo com os recursos de rede. Os switches têm a possibilidade de ser configurados para informar uma estação de

gerenciamento de rede a respeito de qualquer acesso não autorizado aos recursos. Se houver necessidade de comunicação entre VLANs, é necessário empregar roteadores.

Os switches são dispositivos da camada de enlace (camada 2 de acordo com o modelo de referência do Open System Interconnection – RM-OSI) e, por isso, não verificam o protocolo empregado na camada de rede. Essa característica faz com que o switch encaminhe todos os quadros de broadcast. Todavia, por meio de VLANs, domínios de broadcast são automaticamente criados. Broadcasts enviados de um nó na VLAN não serão encaminhados para as portas configuradas em outra VLAN. Ao associar portas de switch ou usuários para grupos de VLANs em um switch ou grupo de switches conectados, adquire-se flexibilidade para adicionar somente os usuários vinculados ao domínio de broadcast, independentemente de sua localização física. Isso pode, por exemplo, parar as tempestades de broadcasts causadas por uma falha em uma placa de rede (NIC) ou aplicativos que o estejam gerando.

Latência

Referente ao atraso que o dispositivo de rede introduz quando os quadros passam por ele. A latência é um aspecto importante para aplicações sensíveis ao tempo, como as transmissões de voz e vídeo.

As VLANs são tipicamente criadas pelo administrador do datacenter, o qual associa portas do switch a uma determinada VLAN. Essas são chamadas de VLANs estáticas. Se o administrador quiser desenvolver um trabalho pensando mais à frente e associar todos os endereços de hardware a um banco de dados, pode usar as VLANs dinâmicas. Os switches oferecem a possibilidade de ser configurados para associar VLANs dinamicamente. Quando uma VLAN se tornar grande demais, podem ser criadas mais VLANs, para que os broadcasts não consumam muita largura de banda.

Virtualização da rede

- Switches virtuais
- Interfaces de redes virtuais

É possível criar uma infraestrutura de rede virtual permitindo a interconexão das diferentes máquinas virtuais em execução em um ou vários servidores, como se elas estivessem em uma rede física. Os dois componentes de uma rede virtual são:

- **Switches virtuais**: é possível criar switches virtuais em um servidor. As funcionalidades de um switch virtual são as mesmas de um switch camada 2, suportando inclusive VLANs com controle pelas portas.
- **Interfaces de rede virtuais (Network Interface Controler – NIC)**: permitem conectar as máquinas virtuais entre si e com a rede externa. As interfaces de redes físicas funcionam como uplinks para as portas do switch virtual. O NIC Teaming possibilita conectar um switch virtual a múltiplos adaptadores Ethernet. As máquinas virtuais podem ser configuradas com um ou mais NICs virtuais, cada um com seu endereço IP e MAC.

Benefícios da virtualização

- Redução do Total Cost of Ownership (TCO) ou custo total de propriedade
- Redução do uso do espaço físico
- Redução do consumo de energia
- Isolamento de ambientes de testes, desenvolvimento e produção
- Flexibilidade na disponibilização de novos servidores
- Padronização das plataformas
- Gerenciamento centralizado
- Simplificação no uso de alta disponibilidade (HA) e recuperação de desastres
- Computação em nuvem e datacenter dinâmico

Novos negócios exigem uma infraestrutura de TI ágil e flexível para suportar a dinâmica dos seus processos. O uso da virtualização pode trazer grandes benefícios para a organização com a adequação da sua infraestrutura de TI ao negócio, mas requer planejamento e aquisição de novos recursos. Os prováveis benefícios obtidos com a virtualização são: redução do TCO e computação em nuvem.

Redução do custo total de propriedade

O custo total de propriedade pode ser reduzido com o uso da técnica de virtualização. Os fabricantes disponibilizam ferramentas que permitem o cálculo do Total Cost of Ownership (TCO), considerando a comparação de uma infraestrutura de TI com e sem virtualização. Em geral, é simples justificar um projeto de virtualização utilizando a abordagem de TCO, tanto para a atualização da infraestrutura física existente como para a construção de uma nova.

A redução do TCO é causada por:

- **Redução do uso do espaço físico**: a utilização da virtualização permite a redução do espaço físico à medida que considera menos servidores como solução. A consolidação das estruturas de armazenamento (storage) e backup, quase sempre contempladas num projeto de virtualização de servidores, também ajuda na redução do espaço como um todo.
- **Redução do consumo de energia**: quase sempre a consolidação física é acompanhada da redução do consumo de energia. Servidores são os responsáveis pelo maior consumo de energia entre os equipamentos de TI, e a consolidação acaba por diminui-lo.
- **Isolamento dos ambientes de teste, desenvolvimento e produção:** em muitas instalações, pode ser caro demais construir ambientes físicos diferentes para os ambientes de teste, desenvolvimento e produção. A utilização da virtualização otimiza o uso dos recursos, pois permite que os ambientes coexistam de maneira completamente isolada, mesmo executando nos mesmos servidores físicos.

- **Flexibilidade na criação de novos servidores**: as máquinas virtuais (servidores lógicos) podem ser criadas de forma automática em servidores físicos já existentes. Na prática, a demanda por um novo servidor físico que dependeria de aprovação, compra e entrega, pode ser atendida por uma máquina virtual pronta para rodar em qualquer servidor que tenha folga.
- **Padronização das plataformas**: à medida que o hipervisor passa a ser o elemento central do servidor virtualizado, todo o esforço de padronização da plataforma fica simplificado, pois a relação com o hardware ocorre por dele. Diferentes sistemas operacionais podem coexistir no mesmo servidor, sob o controle do hipervisor.
- **Gerenciamento centralizado**: o gerenciamento das máquinas virtuais fica centralizado em uma única ferramenta, facilitando o gerenciamento, reduzindo os custos operacionais e promovendo a simplificação do ambiente. **Facilita o uso de alta disponibilidade (HA) e recuperação de desastres.**
- **Implantação de técnicas de alta disponibilidade**: técnicas como clusters de servidores – e o uso de tecnologias de replicação para suportar a recuperação de desastres – podem ser simplificadas com o uso da virtualização. Ela permite a alta disponibilidade independentemente da adoção de técnicas de cluster, facilitando a criação do site secundário e otimizando os recursos alocados para o segundo site. Além disso, permite automatizar os processos de recuperação de desastres com a fácil integração, promovida por técnicas de replicação do storage.
- **Computação em nuvem e datacenter dinâmico**: a virtualização é o componente central do datacenter dinâmico, que, por sua vez, viabiliza a computação em nuvem. A computação em nuvem e o datacenter dinâmico se tornam viáveis à medida que as soluções de virtualização avançam nos aspectos referentes ao balanceamento de carga dinâmica, recuperação de falhas, segurança e interoperabilidade entre sistemas diferentes.

Análise de TCO/ROI

- Total Cost of Ownership (TCO)
- Return On Investment (ROI)
- Análises de TCO permitem realizar uma justificativa de projeto baseada em números
- Os principais fabricantes possuem ferramentas próprias de análise de TCO

Total Cost of Ownership (TCO)

A metodologia de cálculo do TCO considera quatro categorias de custo: hardware e software; operações; downtime; e administração. A metodologia do TCO inclui os custos diretos e indiretos, assim como os de administração, mais difíceis de serem medidos.

Return On Investment (ROI)

Medida usada para comparar o custo do projeto com os benefícios obtidos. A fórmula para calcular o ROI é dividir os benefícios resultantes de um período pelo montante investido no mesmo período.

ROI = Benefícios Quantificáveis/Custos Quantificáveis

Na avaliação de um projeto, as organizações calculam o ROI e o tempo requerido para que os benefícios sejam maiores que os custos. No caso de projetos de virtualização, a análise de TCO é um pré-requisito para o estudo de ROI, se os clientes não conseguirem determinar a economia gerada. O ROI para um projeto de virtualização é calculado pela comparação do custo existente com a redução de custo gerada pelo novo projeto.

Normalmente, os fabricantes de software possuem ferramentas exclusivas para cálculo de TCO. É possível importar os dados fornecidos por essas ferramentas para as ferramentas de ROI de cada fabricante. Isso significa que é possível readequar o projeto introduzindo dados reais de inventário e de uso na ferramenta de ROI, obtidos pela ferramenta de levantamento. Essas ferramentas são facilmente encontradas em pesquisa na internet.

Downtime

Período pelo qual um sistema computacional fica parado por problemas técnicos ou por manutenção preventiva.

Alta disponibilidade

- Alta disponibilidade (HA)
- Níveis de disponibilidade

A alta disponibilidade (HA) e a recuperação de desastres (DR) devem ser pensadas continuamente, já que acontecem em diversas camadas, e cada camada propicia os níveis de disponibilidade adequados para a camada superior.

- **Plataforma**: servidores com redundância de fonte e discos;
- **Dados**: redes SAN redundantes;
- **Aplicação**: cluster failover de aplicação;
- **Site**: replicação de site.

A alta disponibilidade e a recuperação de desastres não devem ser pensadas em um único nível. A ideia de fazer um cluster para a aplicação, por exemplo, não deve eliminar a redundância nas fontes de alimentação e o nível de RAID utilizado em um conjunto de disco dentro dos servidores. Da mesma forma, a replicação entre sites não deve eliminar a necessidade de se fazer o cluster da aplicação no local, por meio de técnicas de clusters de servidores.

Backup e restore

- Necessidades
- Aplicativos de backup
- Mídias

O backup é uma cópia dos dados de produção, criada e retida com o propósito de manter e fornecer segurança para os dados que a empresa considera importantes. O restore é o processo de recuperação desses dados. Os serviços de backup e restore em TI são parte de um processo mais amplo que visa a garantir a disponibilidade do datacenter e da infraestrutura de TI.

Necessidades

O backup/restore precisa ser realizado por várias razões, dentre as quais podemos destacar:

- Requisitos de negócio;
- Requisitos legais;
- Proteção contra falhas de hardware;
- Proteção contra falhas das aplicações;
- Proteção contra erros dos usuários;
- Recuperação de desastres;
- Alcance de níveis de serviço específicos.

O aspecto chave para definir a forma de realizar o backup é a natureza dos dados dele. Deve-se traçar também uma estratégia para a recuperação dos dados do backup e testá-la antes da sua necessidade real. É comum acontecerem surpresas quando se usa o restore, ou seja, quando há a tentativa de recuperação dos dados. A própria forma de realizar o restore dos dados deve ser devidamente entendida com o fabricante do software de backup. A sugestão é incluir o treinamento no restore do ambiente no momento da realização do treinamento específico da ferramenta de backup que será utilizada.

Nem todos os dados de uma organização precisam fazer parte de backup ou necessitam de backup em bases regulares. Os sistemas operacionais de servidores, por exemplo, a menos que sofram algum tipo de correção (patch), não exigem backup em bases regulares. Normalmente, pode-se fazer um questionário para o levantamento dos dados, com o objetivo de definir a estratégia de backup para cada necessidade específica. A partir dos dados levantados pelo questionário, será possível determinar o momento e a periodicidade com que o backup deve ser feito, os dados que precisam constar nele, por quanto tempo têm que ser mantidos e o local onde os backups devem ser feitos, dependendo da necessidade de velocidade para o restore (disco ou fita). Esse procedimento precisa fazer parte de um plano mais amplo de continuidade do negócio.

O processo de backup é ilustrado a seguir.

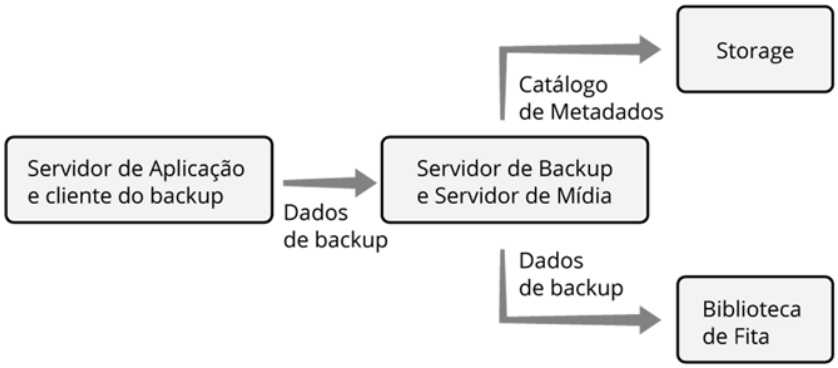


Figura 2.10 Processo de backup.

Aplicativos de backup

O aplicativo de backup é o software responsável por fazer a salvaguarda dos dados estratégicos a partir de políticas determinadas pelas necessidades da instituição. As políticas de backup determinam se, quando, onde e por quanto tempo o dado será salvo. O software deve possibilitar a realização de backups completos e periódicos (diário, semanal, mensal etc.). É importante testar os dados salvos por meio do restore, pois em muitos casos o backup é gravado em mídias defeituosas, o que impossibilita a sua recuperação.

Mídias

As unidades de fita são o dispositivo natural para a realização do backup, devido ao seu custo/benefício em comparação com outras mídias. Um aspecto-chave do backup baseado em fitas é a definição da política de rotação das fitas e o tempo de retenção delas. Parâmetros de tempo de retenção:

- On-site – 30 dias;
- Off-site – 180 dias (não considerando imposições legais).

Sobre a segurança do backup, técnicas de cifragem podem ser utilizadas, mas é importante lembrar que seu uso degrada o desempenho do backup e do restore.

3

Introdução ao Xen

- Histórico
- Versões

O hipervisor Xen é uma solução de virtualização baseada em monitor de máquina virtual, inicialmente concebido como um projeto de pesquisa no âmbito acadêmico. Graças ao seu desempenho e também ao fato de ter sido desenvolvido segundo uma filosofia de código aberto, o Xen rapidamente se tornou uma solução de virtualização interessante para o mundo dos sistemas operacionais de código aberto. O sucesso foi tão grande que seus criadores fundaram a empresa Xensource, posteriormente adquirida pela Citrix System Inc. Atualmente, o Xen é um produto de mercado denominado XenServer e possui uma versão livre e outras comerciais. A diferença entre elas está relacionada a alguns recursos avançados, como gerenciamento de energia, balanceamento de carga e ferramentas para gerenciar, monitorar e auditar ambientes virtualizados. O objetivo deste capítulo é apresentar os principais conceitos relacionados à concepção do Xen e suas características como produto.

Histórico

O Xen foi originalmente desenvolvido pelo Grupo de Pesquisa de Sistemas do Laboratório de Computação da Universidade de Cambridge, como parte do projeto XenoServer, coordenado por Ian Pratt e Keir Fraser. O nome Xen deriva do grego xenos, que significa estrangeiro ou desconhecido. O objetivo do projeto era proporcionar uma infraestrutura pública para computação distribuída em larga escala. O componente fundamental da arquitetura do XenoServer era um hipervisor denominado Xen, responsável por virtualizar e gerenciar os recursos físicos por meio de ambientes virtuais de execução.

A primeira versão do Xen foi liberada em 2003 e rapidamente se popularizou, por duas razões: ser um software livre e oferecer um desempenho muito bom se comparado com as demais soluções de virtualização de sua época. No entanto, por ser baseada em paravirtualização, sua grande desvantagem era a necessidade de modificação do código-fonte dos sistemas operacionais convidados. Isso limitava o seu emprego em sistemas operacionais de código aberto. Atualmente, o Xen tem capacidade para executar máquinas virtuais com sistemas operacionais não modificados (virtualização completa). É licenciado nos termos da GNU General Public Licence (GPLv2) e permite que vários sistemas operacionais hóspedes (convidados) sejam executados em um mesmo sistema hospedeiro.

O projeto XenoServer originou a XenSource Inc. Em outubro de 2007, a XenSource Inc. foi comprada pela Citrix System Inc., pelo valor de 500 milhões de dólares. Após a aquisição, a Citrix passou a suportar o desenvolvimento do projeto open source, mantendo a empresa

XenSource ativa, e começou a investir e a comercializar uma versão “empresarial” do Xen. Em decorrência do negócio, os produtos da XenSource Inc. foram rebatizados. O XenExpress se tornou XenServer Express Edition e XenServer OEM Edition – a última é uma versão embarcada do hipervisor Xen. O XenServer passou a ser denominado XenServer Standard Edition, e o produto XenEnterprise ganhou o nome de XenServer Entreprise Edition.

Em outubro de 2009, a Citrix transformou seus produtos comerciais baseados no XenServer em software livre e reorganizou a divisão de virtualização em duas linhas: infraestrutura; e gerenciamento e automação. Os produtos foram novamente renomeados e, agora, existem quatro edições: XenServer (livre), XenServer Advanced Edition, XenServer Enterprise e XenServer Platinum. Consulte o site da Citrix para conhecer as famílias de produtos XenServer.

Edição	Características
Free Edition	Plataforma de virtualização básica e sem custo.
Advanced Edition	Alta disponibilidade e ferramentas avançadas de gerenciamento.
Enterprise Edition	Balanceamento dinâmico de carga e de energia, com gerenciamento baseado em perfis.
Platinum Edition	Automação avançada e recursos de computação em nuvem.

Figura 3.1 Características e edições do XenServer.

Em 2012, os primeiros pacotes para o Xen Cloud Plataform foram liberados, facilitando assim a implantação da solução de computação em nuvem.

Os anos de 2013 e 2014 foram bem movimentados. O Xen tornou-se parte da Fundação Linux e passou a desenvolver projetos colaborativos. O site da comunidade deixou de ser o xen.org e utilizou outro domínio (XenProject.org). Para finalizar, a nova marca do projeto Xen foi lançada pela Fundação Linux.

O Xen pode ser obtido de forma livre a partir dos sites da XenSource (www.xenproject.org) ou da Citrix System Inc. (www.citrix.com).

Conheça a história do Xen em:
http://www.xenproject.org/about/history.html

Versões

A versão Xen 1.0 corresponde àquela de seu lançamento e, em um curto espaço de tempo, foi substituída pela versão 2.0, já desenvolvida pela XenSource Inc. Essa versão foi caracterizada por modificações que visavam a transformar o Xen em um produto competitivo para o mercado de virtualização. Como estratégia de negócios da XenSource Inc., o hipervisor Xen foi liberado como solução em software livre. Tanto a versão 1.0 como a 2.0 ofereciam apenas paravirtualização.

Em 2005, foi lançada a versão Xen 3.0 que, entre várias modificações, incluía um melhor suporte para arquiteturas SMP e de 64 bits, uma gama maior de dispositivos de hardware, além de estender a capacidade de endereçamento de memória de 4 GB para até 64 GB (suporte PAE36). Entretanto, a alteração mais significativa foi aproveitar a capacidade de virtualização por hardware dos processadores Intel-VT (Vanderpool) e AMD-V (Pacífica). Isso permitiu que o Xen suportasse uma virtualização completa, ou seja, a execução de sistemas operacionais hóspedes sem a necessidade de modificações.

Com o lançamento do Xen 4.0, otimizações relacionadas ao gerenciamento de memória e aspectos de segurança foram otimizadas. Assim, é possível fornecer maior escalabilidade para atender cargas de trabalho (workloads) intensivas, processamento e uso de rede, sendo apropriado para aplicações em nuvem (cloud computing). Tipicamente, a designação XenServer é empregada pela Citrix e o nome Xen é usado pela XenSource, mas ambas adotam o mesmo hipervisor como base.

Com uma plataforma completa e consolidada para virtualização de servidores, o XenServer encontra-se na versão 6.2 e está totalmente compatível e engajado nas soluções para computação em nuvem e redes do futuro.

Componentes

- Hipervisor
- Domínio convidado privilegiado
- Domínio convidado não privilegiado

O Xen é composto por três componentes fundamentais, como mostra a figura 3.2: hipervisor Xen, domínio convidado privilegiado e domínio convidado não privilegiado. O hipervisor Xen é a camada de base. Sobre ela estão um ou mais sistemas operacionais hóspedes, os quais o hipervisor escalona sobre os processadores disponíveis. Há sempre, no mínimo, um Sistema Operacional convidado na arquitetura Xen. Essa é uma máquina virtual automaticamente inicializada no boot do Xen e, na terminologia empregada, corresponde ao domínio convidado 0 ou, simplesmente, Dom0. Os demais sistemas convidados são denominados DomU, onde o “U” vem do inglês unprivileged.

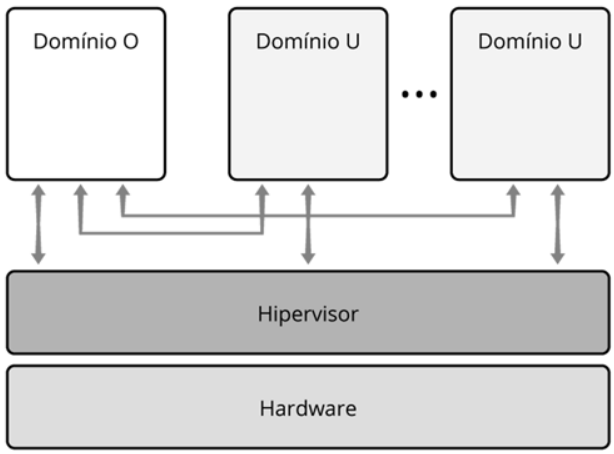


Figura 3.2 Componentes do Xen: hipervisor e domínios convidados.

Hipervisor Xen

O hipervisor Xen é um monitor de máquina virtual do tipo I (baremetal) que executa entre o hardware e um sistema operacional convidado (figura 3.2). Na prática, o hipervisor executa diretamente sobre o hardware e se torna um tipo de interface entre todas as requisições de hardware (processador, interrupções e disco) para os demais sistemas operacionais. Ao separar os convidados do hardware, o Xen está apto a executar múltiplos sistemas operacionais de forma segura e independente.

Uma característica interessante da solução Xen é que o hipervisor não possui nenhum tipo de driver de dispositivo. Isso significa, na prática, que o Xen não acessa efetivamente o hardware de E/S, mas apenas o gerencia. Os acessos aos dispositivos são feitos pelos drivers existentes no domínio convidado 0 (Dom0) de forma coordenada com o hipervisor. Devido a essa estrutura, o Dom0 é considerado um domínio privilegiado.

Domínio convidado privilegiado

Durante a inicialização, o domínio convidado privilegiado (referenciado como Dom0) é o primeiro Sistema Operacional convidado automaticamente a ser carregado pelo hipervisor Xen. O Dom0 é o único domínio convidado que tem privilégios especiais para acessar e administrar o hipervisor Xen, criar, controlar e parar outros domínios convidados.

- Só pode haver um Dom0

Dom0 pode ser um Sistema Operacional qualquer, desde que tenha a capacidade de interagir com o Xen. Isso se traduz no emprego de sistemas operacionais que seguem uma filosofia de código aberto (como GNU/Linux, FreeBSD, NetBSD e OpenSolaris), pois é necessário acessar o código-fonte antes de adaptá-lo ao hipervisor Xen. Os sistemas da família Microsoft Windows, por serem de código fechado, não podem ser usados como Dom0.

Já foi mencionado que o Xen não conta com nenhum tipo de driver de dispositivo e atua apenas controlando os recursos fundamentais do processador, como interrupções, gerência

de memória (MMU) e escalonamento. Cabe ao Dom0 acessar e controlar os dispositivos de E/S por meio de seus próprios drivers. As requisições de E/S provenientes dos demais domínios convidados são encaminhadas para o Dom0, como mostra a figura 3.3.

Por executar um Sistema Operacional, é possível que o usuário use um login no Dom0 de forma convencional. Entretanto, o Dom0 tem privilégios especiais sobre o hipervisor e demais domínios convidados. Um administrador de sistemas, ao efetuar o login no Dom0, tem a capacidade de gerenciar toda a máquina. Por isso, é importante que o Dom0 seja instalado e configurado para contemplar ao máximo o quesito segurança, evitando a execução de serviços desnecessários, entre outros problemas. Tudo isso deve ser feito para reduzir a chance de sofrer um ataque e acabar comprometido, o que prejudicaria também os demais sistemas convidados.

Domínio convidado não privilegiado

Os domínios convidados não privilegiados, denominados DomU, são carregados e inicializados a partir do Dom0. Em um sistema computacional, podem existir vários DomU e cada um deles executa de forma totalmente isolada e independente dos demais domínios existentes na máquina. Um domínio convidado opera de forma não privilegiada. Todos os acessos a recursos compartilhados (processador, memória, E/S) devem ser realizados por meio do Dom0 e do hipervisor Xen.

O Xen exigia, até a sua versão 2, que o Sistema Operacional de um domínio convidado fosse modificado justamente para substituir as chamadas de sistemas nativas por chamadas ao hipervisor (paravirtualização). Desde a versão 3, isso não é mais necessário, ou seja, é possível o emprego de sistemas operacionais não modificados, o que permite a execução dos sistemas da família Microsoft Windows (virtualização completa). Vale repetir que, no entanto, essa opção só é válida quando o processador conta com suporte por hardware para virtualização (Intel-VTx ou AMD-V).

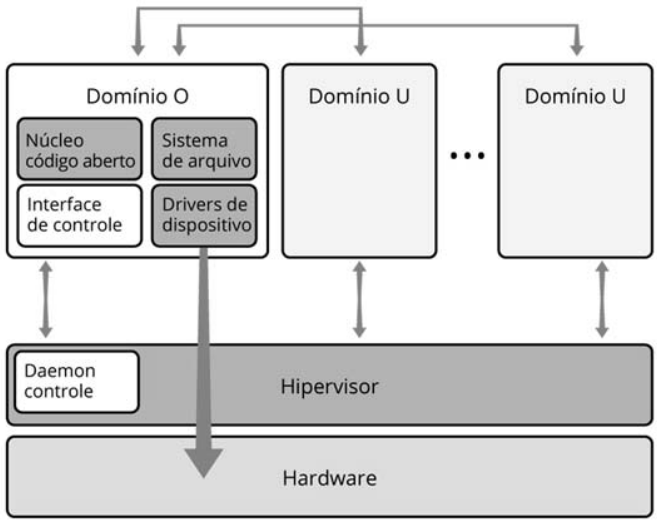


Figura 3.3 Interação entre Hipervisor, Domínio 0 e Domínios U.

Virtualização no Xen

- Modos de operação do processador
- Arquitetura de drivers de dispositivos
- Paravirtualização
- Virtualização completa
- Híbrida

Paravirtualização é a técnica de virtualização em que o Sistema Operacional convidado tem consciência de que está executando sobre um hipervisor e não sobre o hardware real da máquina. Nesse caso, o Sistema Operacional é modificado para não realizar chamadas de sistemas nativas (que são instruções privilegiadas do processador), e sim fazer chamadas ao hipervisor. Já na virtualização completa, o Sistema Operacional é iludido para acreditar que está executando diretamente sobre o hardware, quando, na verdade, o faz sobre um hipervisor. O uso dessas técnicas no Xen será detalhado a seguir.

Modos de operação do processador

Os sistemas operacionais definem dois modos de execução: espaço de usuário e espaço de núcleo. Quando opera em espaço de núcleo é possível usar qualquer instrução do conjunto assembly do processador. Já a execução em modo usuário permite o emprego apenas das instruções assembly não privilegiadas. A divisão entre instruções privilegiadas e não privilegiadas é feita pela configuração do processador em modos de operação ou proteção.

As arquiteturas x86 definem até quatro níveis de proteção (0,1, 2 e 3), denominados anéis (Current Privilege Level – CPL), mas os sistemas operacionais convencionais empregam apenas dois: 3 e 0. Esses níveis correspondem, respectivamente, aos espaços de usuário e núcleo. As aplicações de usuário executam no CPL 3, e o núcleo, no CPL 0. As chamadas de sistema são implementadas pela execução de uma interrupção de software, que recebe como

parâmetro, via registradores do processador, o código da função a ser realizada e seus argumentos. A operação da instrução de interrupção de software altera o anel de proteção para CPL 0. É assim que a execução é transferida do espaço de usuário para o de núcleo. A figura abaixo ilustra a estrutura dos anéis de proteção.

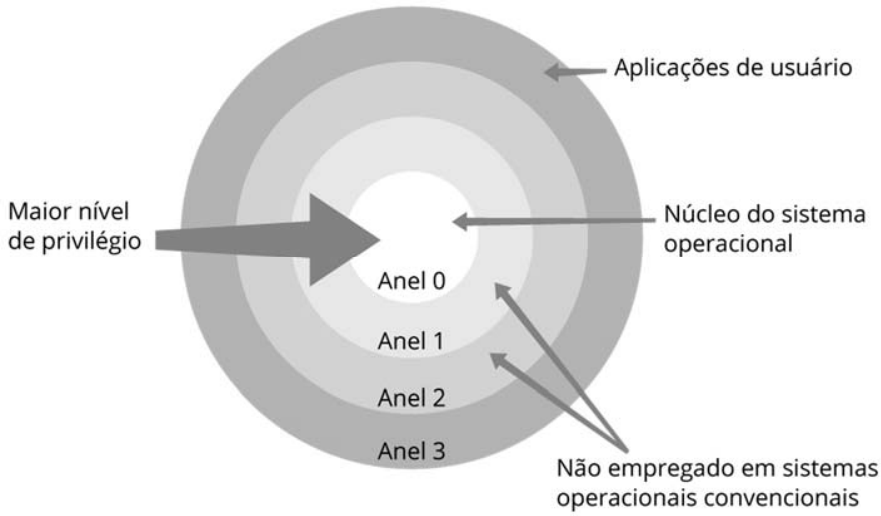


Figura 3.4 Anéis de proteção em arquiteturas x86.

Arquitetura dos drivers de dispositivo

O domínio 0 é o único que acessa diretamente os dispositivos de E/S por meio dos respectivos drivers instalados no núcleo do Dom0. Na arquitetura Xen, os domínios convidados não privilegiados (DomU) não têm autorização para acessar os dispositivos de E/S. No entanto, é claro que as aplicações que executam nos DomU efetuam operações E/S. Surge a questão: como os DomU fazem acessos a E/S, se eles não são autorizados?

A solução para esse dilema é fazer com que os DomU solicitem ao Dom0 o acesso aos dispositivos de E/S em nome deles. Para isso, o Dom0 tem uma camada de software adicional denominada backend, que disponibiliza drivers genéricos virtuais para dispositivos de E/S orientados a bloco e rede. Os DomU têm acesso a esses drivers genéricos por meio de outra camada de software, o frontend. Dessa forma, um DomU realiza requisições de E/S para um driver, o frontend, da mesma forma que faria para um núcleo Linux. Entretanto, o frontend é apenas uma interface virtual, sem acesso ao hardware real, que encaminha essas requisições ao backend, delegando a tarefa de E/S para o Dom0. Do backend, elas são reencaminhadas para o driver real, conforme ilustra a figura a seguir.

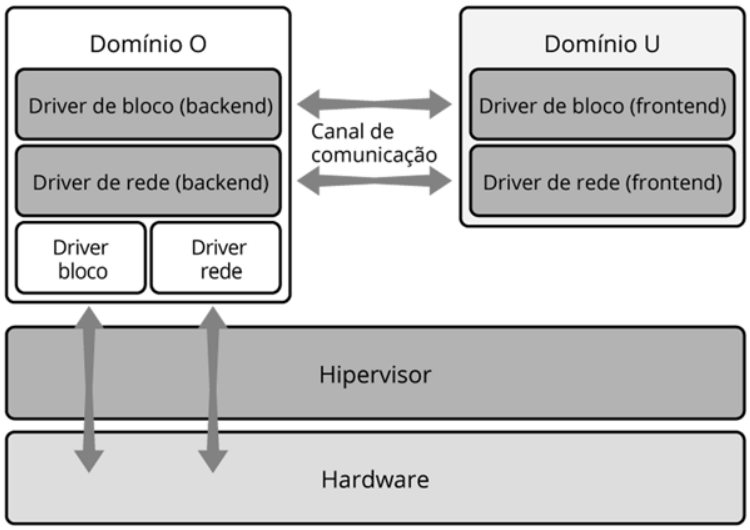


Figura 3.5 Mecanismo de backend e frontend.

Sempre que um DomU é instanciado, cria-se um canal de comunicação entre ele e o Dom0. Esse canal é implementado por uma área de memória compartilhada, de onde mensagens de requisições e de dados são enviadas do DomU para o Dom0 e vice-versa. Existe uma área de memória compartilhada exclusiva entre cada DomU e o Dom0.

Na realidade, devido a questões de desempenho, essa arquitetura modular de backend e frontend não é a única disponível no Xen. Na versão mais recente do hipervisor, há um mecanismo denominado Xen PCI Passtru, no qual é possível dar privilégios especiais para um DomU acessar diretamente dispositivos de E/S sem passar pelo Dom0.

Paravirtualização

Os sistemas operacionais convencionais, como o Linux e a família Microsoft Windows, nas arquiteturas x86, implementam as chamadas de sistema pela interrupção de software int 0x80. Lembre-se de que as aplicações executam no CPL 3, e o núcleo, no CPL 0. A paravirtualização no Xen consiste em fazer o hipervisor operar no anel CPL 0, e o Sistema Operacional, no CPL 1. As aplicações de usuários continuam a ser executadas no CPL 3.

O hipervisor Xen introduz o conceito de hypercall, que é uma interrupção de software (INT 0x82), mas altera o anel de proteção para CPL1 em vez de CPL 0. Dessa forma, os sistemas operacionais convidados são modificados para executar no anel 1 e, no lugar de realizarem chamadas de sistema (int 0x80), fazem hypercalls (int 0x82). Cada hypercall verifica a função a ser executada, cuida da consistência dos argumentos de execução, garante a não interferência de um núcleo em outro e realiza acessos controlados ao hipervisor no CPL 0. O hipervisor, então, efetua a operação desejada.

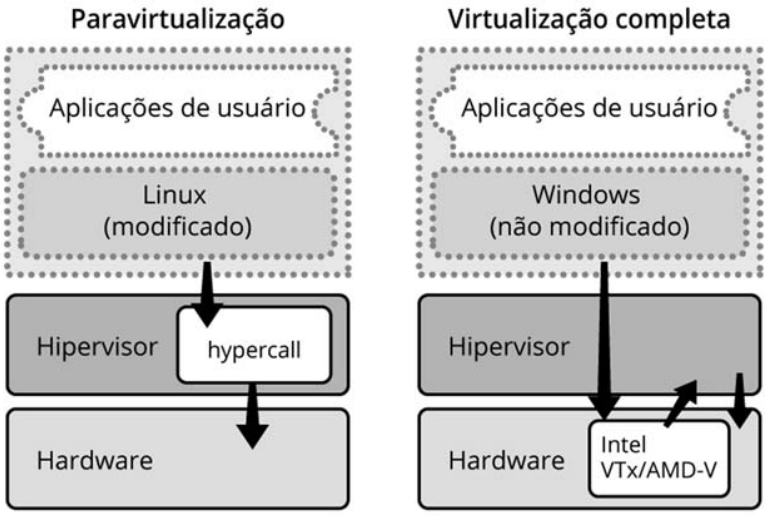


Figura 3.6 Implementação de paravirtualização e virtualização completa no Xen.

Virtualização completa

Técnica de virtualização que permite aos sistemas operacionais convidados executarem sem modificações. Com a virtualização completa, o Sistema Operacional não distingue se está operando diretamente sobre o hardware real ou sobre um hipervisor. Entretanto, para que isso seja possível, é necessário que o processador ofereça suporte à virtualização por hardware (Hardware Assisted Virtualization – HAV).

No caso do Xen, os processadores Intel-VTx e AMD-V fornecem esse suporte por meio de soluções próprias. Para tratar a diferença entre essas implementações, o Xen oferece uma camada de abstração chamada de Hardware Virtual Machine (HVM). O HVM concede uma interface única para o ambiente virtualizado e possui implementações específicas para uma ou outra família de processadores.

Outro detalhe importante é que um Sistema Operacional não modificado, ao inicializar, espera encontrar uma BIOS. Nesse caso, o Xen oferece uma BIOS virtual para dar a ilusão de que o Sistema Operacional está carregando em uma máquina real. Normalmente, a BIOS virtual é implementada pelo emulador Qemu.

Em decorrência das técnicas de virtualização empregadas por Xen, é bastante comum encontrar referências à terminologia “convidado-PV”, para indicar o uso de paravirtualização, e “convidado HVM”, no caso de ser usada a virtualização completa. A figura 3.6 mostra os dois tipos.

Virtualização híbrida

Existe ainda uma terceira possibilidade de virtualização denominada híbrida, mas que, atualmente, está em estágio de pesquisa e desenvolvimento. O princípio básico é lembrar que são os drivers do dispositivo que realmente executam operações em dispositivos de E/S.

Normalmente, os drivers são um módulo à parte, separados do núcleo, que por questões de desempenho são ligados junto ao núcleo, formando um código monolítico. O núcleo em si é responsável apenas pelo recebimento e identificação de interrupções, pela chamada da rotina de tratamento de interrupções adequada, pela gerência de memória e pelo escalonamento.

A ideia da virtualização híbrida é explicitar essa separação entre os drivers de dispositivos e o núcleo. Dessa forma, o núcleo pode executar sem modificações, porque não acessará nenhum dos dispositivos de E/S. Nesse caso, os drivers de dispositivos passam a ser paravirtualizados, ou seja, eles são alterados para funcionar sob a coordenação de um hipervisor e não mais sob a supervisão do núcleo do Sistema Operacional. Até o período da elaboração deste material (out. 2010), não foi criada nenhuma ferramenta de virtualização que ofereça esse tipo de suporte na forma de um produto; há apenas trabalhos em fase de desenvolvimento.

XenMotion – migração de domínios convidados

- Regular ou freeze-and-copy
- Direta ou live migration

Uma capacidade bastante interessante do Xen é a migração de máquinas virtuais (domínios convidados), ou seja, passar um domínio convidado que executa em uma determinada máquina física para outra. Essa capacidade é fornecida pela funcionalidade XenMotion.

Existem duas formas de realizar a migração. A primeira, chamada de regular ou freeze-and-copy, consiste em parar a máquina a virtual, fazer um cópia de seu contexto de execução, transferir essa cópia para a máquina-destino e reiniciá-la. No entanto, o segundo tipo, denominado migração direta ou live migration, é mais poderoso, pois permite que a máquina virtual seja migrada sem ter a necessidade de suspender a execução da máquina virtual.

Algumas condições devem ser satisfeitas para que a migração aconteça no Xen. Inicialmente, é necessário que tanto a máquina física-origem quanto a de destino estejam executando o daemon de controle xend. ou seja, executem o hipervisor Xen. Segundo, a máquina física de destino deve possuir recursos físicos suficientes para atender às demandas do domínio que está sendo migrado. Por fim, o Xen não emprega nenhum mecanismo ou método para efetuar a migração do sistema de arquivos usado pelo domínio convidado que está sofrendo a migração. Portanto, considera que está sendo adotado algum tipo de sistema de arquivos compartilhado, como, por exemplo, o NFS. Assim, o domínio convidado pode migrar de uma máquina física a outra, já que o sistema de arquivos está localizado fisicamente em uma terceira máquina, o servidor de arquivos.

Tecnicamente, para que a migração direta (live migration) ocorra, é necessário que a memória do domínio convidado seja copiada para a máquina de destino. Para evitar a parada, essa cópia é feita de forma iterativa entre a máquina de origem e destino, de modo a permitir que a execução continue na origem até que a cópia seja concluída, e o domínio, reinicializado no destino. A migração direta começa com uma reserva prévia de recursos de memória na máquina de destino, seguida de uma transferência de dados iterativa e por demanda, na qual são passados apenas os dados modificados entre uma iteração e outra.

Enquanto essa cópia vai sendo realizada, o mecanismo de migração direta busca identificar momentos em que seja possível suspender a execução do domínio convidado na máquina de origem e reativá-lo na de destino – de forma que resulte em um tempo mínimo de indisponibilidade. Se, por ventura, tal condição não for identificada, ou caso o número de iterações de atualização for muito alto, o domínio convidado é suspenso na origem para que a migração seja concluída. Como a cópia de memória já estava em curso, o tempo de interrupção tende a ser menor do que se tivesse sido feito de forma regular.

Cabe ressaltar que a migração de domínios é interessante para atender a dois aspectos fundamentais em uma infraestrutura de serviços: manutenção e balanceamento de carga. Tipicamente, realizar manutenção em servidores provoca uma descontinuidade no oferecimento dos serviços que ele provê. A virtualização, associada à migração, é o instrumento que permite reduzir o tempo de indisponibilidade dos serviços (downtime). Assim, se os serviços são providos por um domínio convidado, basta migrá-lo para outro servidor enquanto é feita a manutenção no servidor original. Se o recurso de livre migration for empregado, o fornecimento do serviço praticamente não sofre interrupção.

O balanceamento de carga é apropriado quando um ou mais servidores estão ociosos ou com uma baixa taxa de processamento, enquanto outros servidores estão sobrecarregados. Nesse caso, os domínios convidados que executam nos servidores sobrecarregados podem ser transferidos para os ociosos. O objetivo é tentar manter todas as máquinas físicas com uma carga de processamento similar, para atender melhor todos os usuários dos serviços providos. A situação mais comum é que o sistema ofereça serviços automáticos para a detecção de servidores disponíveis e a realização da migração. Isso passa por um monitoramento dos recursos dos servidores físicos.

É comum que as capacidades de migração e de balanceamento de carga façam parte de um "pacote empresarial" de uma solução de virtualização. No caso específico do Xen, a capacidade de migração (XenMotion) é disponível na forma de software livre; já o balanceamento de carga é feito por ferramentas comerciais específicas. A Citrix oferece balanceamentos de carga para a família XenServer apenas nas versões Enterprise e Platinum (ver tabela 3.1).

Outras formas de virtualização com o Xen

- XenDesktop
- XenApp
- Computação em nuvem

O enfoque deste curso é a virtualização de servidores. Mas lembre-se de que ela não é a única forma de virtualização existente, pela qual uma organização pode tirar proveito para reduzir custos e, ao mesmo tempo, flexibilizar sua infraestrutura de TI e torná-la mais ágil. Em um contexto empresarial, tão importante quanto a virtualização de servidores é a virtualização de desktops, de aplicações e também a computação em nuvem. Dada a importância dessas atividades, o Xen fornece soluções por meio da XenSource e da Citrix.

XenDesktop

O mesmo tipo de tecnologia e de princípio utilizado para virtualizar servidores também pode ser empregado em máquinas de usuários finais, os desktops. A ideia fundamental por trás da virtualização dos desktops é centralizar a geração de imagens e a administração de sistemas operacionais empregados pelos usuários finais. A partir de um sistema computacional com recursos modestos ou mesmo sem muitos cuidados de administração, um usuário final pode acessar recursos de uma infraestrutura corporativa de forma segura e funcional. Entretanto, para realizar esse acesso é necessário que a máquina do usuário final execute um programa-cliente de conexão remota, como o Remote Desktop Connection (RDC).

A princípio, o conceito de virtualização de desktops parece similar ao de terminal services da Microsoft, mas não é. A diferença fundamental entre eles é que as soluções baseadas no tipo terminal services permite aos usuários dividir um ambiente de desktop com outros. Na virtualização de desktop, o usuário tem acesso a um ambiente próprio de desktop em um servidor central – é como se cada um estivesse sozinho usando seu próprio desktop. As principais vantagens da virtualização de desktop advêm dessa característica. Primeiro, se cada usuário corresponde a uma máquina virtual instanciada em um servidor central (e remoto), esses equipamentos estão compartilhando recursos de rede, de processamento, memória e armazenamento, que provavelmente não estariam disponíveis em um desktop físico. Segundo, a imagem do Sistema Operacional é mantida de forma centralizada, por uma equipe de TI dedicada; no físico, não haveria a possibilidade de um usuário final ser o administrador do desktop virtual. Por fim, eventuais problemas em um desktop não interferem na seção de trabalho de outro usuário em outro desktop, ou seja, explora-se a capacidade de encapsulamento das máquinas virtuais, confinando cada usuário em um sistema próprio.

As soluções para virtualização de desktop somam uma série de componentes, incluindo o lado que executa na máquina do usuário final, um protocolo de conexão remota, o hipervisor e as máquinas virtuais no servidor central. Existe ainda uma série de outras facilidades com vistas ao gerenciamento das imagens, tolerância a falhas etc. Em função disso, é bastante comum encontrar a referência ao acrônimo Virtual Desktop Infrastructure (VDI).

O XenDesktop é o produto disponibilizado pela Citrix para a virtualização de desktops. Há cinco versões do produto: express, standard, advanced, enterprise e platinum. As diferenças principais entre elas estão relacionadas à quantidade de usuários finais suportados; capacidade de failover; acesso remoto; gerenciamento remoto e de imagens; suporte a aplicações por demanda; entre outros. Para mais detalhes, consulte o site do fabricante (www.citrix.com).

XenApp

A virtualização de aplicações (AppV) é o nível mais alto de abstração da virtualização. Essencialmente, é a capacidade de usar um software ou serviço de forma completamente isolada do Sistema Operacional nativo da máquina local. Uma forma pragmática de enxergar a virtualização da aplicação é vê-la como um executável único que possui todos os recursos de softwares (módulos, bibliotecas, chaves de registro etc) necessários à sua operação, sem

usar os do Sistema Operacional. No estágio atual, a virtualização de aplicações está restrita a sistemas operacionais da família Microsoft Windows.

Existem duas grandes vantagens para se empregar a virtualização de aplicações. Primeiro, é bastante comum que durante a instalação de um aplicativo as configurações de um Sistema Operacional sejam afetadas de alguma forma. Por isso também, a instalação de aplicativos deve ser feita com direitos administrativos (root do sistema). Eventualmente, após várias instalações, o Sistema Operacional pode ficar instável ou lento. Ao virtualizar uma aplicação, ela encapsula todos os elementos de que precisa sem ter que alterar, configurar ou incluir módulos no SO nativo. Segundo, uma aplicação virtualizada pode executar em qualquer versão de Windows. Isso porque a camada de virtualização de aplicação gerencia todas as interações entre a aplicação e o Sistema Operacional da máquina.

Existem vários fornecedores de soluções para a virtualização de aplicação. A Microsoft, por meio da ferramenta denominada Microsoft Application Virtualization (MAV), antigamente chamada de SoftGrid, possibilita a criação de aplicações virtualizadas e oferece sua instalação por demanda – um usuário que não tenha o software instalado em sua máquina, pode tê-lo de forma automática. A VMware oferece o ThinApp. A Citrix Inc. tem o XenApp, originalmente conhecido como Citrix Presentation Server. Essas ferramentas são similares ao MAV.

4

Gerenciamento do Xen

Infraestrutura de hardware para o Xen

- Processador e dispositivos de E/S
- Memória RAM
- Sistema de armazenamento (storage)
- Rede

O objetivo deste capítulo é fornecer informações adicionais e práticas consideradas no desenvolvimento de uma solução de virtualização baseada no Xen para projetos de infraestrutura básica. Inicialmente, serão abordados aspectos relacionados a processador, memória, subsistema de armazenamento e rede. Esse entendimento é importante para ajudar o dimensionamento dos recursos empregados na solução de virtualização, de acordo com as demandas específicas de cada organização. Serão abordadas ainda a geração, instalação e utilização das imagens rodadas nos domínios convidados (máquinas virtuais).

Processador e dispositivos de E/S

- XenServer
- DomU
- Hipervisor Xen
- Dom0

O XenServer executa em uma vasta gama de processadores da família Intel e AMD. Tanto os processadores quanto o próprio Xen estão constantemente em evolução, o que torna bem difícil informar uma tabela de compatibilidade sem correr o risco de rápida desatualização. Como regra geral, pode-se dizer que os processadores derivados da família x86 são aqueles com melhores garantias de funcionamento e estabilidade do Xen. É recomendável o emprego de um processador de 64 bits com frequência de operação superior a 1.5 GHz e, de preferência, multicore. O Xen pode suportar 32 processadores ou núcleos.

É possível analisar o XenServer por seus três grandes componentes: Domínio 0 (Dom0), Domínio U (DomU) e o hipervisor Xen. O Dom0 é um sistema operacional GNU/Linux com o núcleo modificado para interagir com os DomU e o hipervisor. Os requisitos de processador e de hardware são os de uma instalação Linux convencional, ou seja, não há grandes restrições e exigências para executar o Dom0.

Para o Domínio U é necessário considerar o emprego da paravirtualização ou da virtualização completa. Assim como o Dom0, cada DomU precisa respeitar as mesmas

necessidades de processador, memória e hardware recomendados para o Sistema Operacional que utiliza. Entretanto, para empregar máquinas virtuais com sistemas operacionais não modificados (virtualização completa) no DomU, como é o caso da família Microsoft Windows, é preciso que o processador ofereça suporte à virtualização por hardware (Hardware Assisted Virtualization – HAV). Atualmente, esse suporte é disponível nos processadores Intel-VTx e AMD-V (SVM) e costuma ser habilitado por meio de opções da BIOS.

Apesar das soluções oferecidas pela Intel e pela AMD seguirem um mesmo princípio de funcionamento, elas são implementadas de formas distintas, o que torna um código dependente do processador. Para mascarar essas diferenças, o Xen define uma camada de compatibilidade chamada de Hardware Virtual Machine (HVM).

Na prática, o hipervisor do Xen é uma camada de software com tarefas bastante específicas e básicas – como configuração de controladores dos dispositivos de hardware e da unidade de gerência de memória (MMU); escalonamento; troca de contexto; e tratamento de interrupções de hardware. O hipervisor em si não impõe fortes restrições quanto ao processador. A prova disso é que o hipervisor Xen é capaz de executar em um Pentium 4 com 512 MB de memória RAM, embora seja relativamente óbvio que um ambiente real de produção exige recursos de processamento e de memória bem mais robustos.

Como visto no capítulo anterior, para acessar os dispositivos de entrada e saída, o Xen emprega os drivers do Sistema Operacional que executa no Dom0. Portanto, a compatibilidade do Xen quanto ao suporte a dispositivos é exatamente a oferecida pelo SO do Dom0. A opção mais comum para o Dom0 é o Sistema Operacional GNU/Linux, com suporte para uma variedade de dispositivos de E/S.

Memória RAM

- Capacidade de endereçamento
- Aplicações
- DomU
- Dom0
- Hipervisor

Como qualquer processo, o Xen tem uma capacidade de endereçar memória limitada pelo barramento de endereços do processador. Em sistemas com 32 bits, é possível acessar apenas 4 GB de memória (232). Uma das evoluções do Xen 3.0 foi oferecer para as arquiteturas de 32 bits suporte ao mecanismo de Physical Address Extension (PAE), que permite acessar até 64 GB de memória. Já para sistemas com arquiteturas de 64 bits, o limite é de 1 TB de memória RAM. A quantidade de memória desejável é aquela que possibilita executar as máquinas virtuais e as aplicações de maneira satisfatória ao usuário final. A questão é dimensionar adequadamente essa quantidade de memória, o que envolve custo.

Inicialmente, como em qualquer sistema computacional, a quantidade de memória a ser empregada depende muito do que se espera executar. No caso do Xen, serão considerados a quantidade de máquinas virtuais que se deseja instanciar, o Sistema Operacional escolhido

para o Dom0 e para cada DomU e as aplicações que essas máquinas virtuais vão executar. Em relação às aplicações e serviços empregados nas máquinas virtuais (DomU), o dimensionamento de memória deve ser o usado em ambientes não virtualizados.

Um ponto de partida bastante lógico para os domínios U é indicado no wiki do Xen: determinar o número de máquinas virtuais que serão usadas simultaneamente e somar a capacidade mínima de memória recomendada pelo fabricante de cada Sistema Operacional, empregado em cada máquina virtual (DomU). Grosso modo, isso é equivalente a multiplicar por 1,5 a quantidade de máquinas virtuais instanciadas. O resultado é a quantidade de memória em GB necessária para os domínios U. É preciso ainda considerar a memória necessária ao Dom0 e ao hipervisor Xen.

Estudos e análises de desempenho com o Xen demonstraram que o Dom0 é capaz de executar com estado mínimo de funcionamento em apenas 64 MB de RAM. Certamente, essas condições não são apropriadas para ambientes de produção. A memória consumida pelo Dom0 depende fortemente dos recursos empregados pelo Sistema Operacional que nele executa. É importante salientar que o papel desempenhado pelo Dom0 é praticamente o de gerenciar o Xen e oferecer os drivers de dispositivos de E/S do sistema. Portanto, seja por questões de redução de espaço de memória como de segurança, é desejável que o Dom0 opere com o mínimo possível de opções de configuração do Sistema Operacional. Um valor tipicamente considerado para o Dom0 é 1 GB de RAM. Quanto ao hipervisor, o consumo de memória tem cerca de 128 MB de memória RAM.

Sistemas de armazenamento (storage)

- Infraestrutura física de armazenamento
- Logical Volume Management (LVM)
- Migração e sistemas de armazenamento

O sistema de armazenamento empregado pelo Xen, novamente, depende da quantidade de máquinas virtuais instanciadas e das aplicações executadas. Entretanto, há algumas diretrizes gerais que devem ser consideradas no dimensionamento da capacidade de disco e no tipo de sistema de armazenamento a ser usado.

Cada máquina virtual tem seu próprio sistema de arquivos e, por consequência, deve haver espaço em disco suficiente para cada uma delas. O local exato onde os arquivos serão armazenados, assim como o tipo de sistema de arquivos a ser empregado em cada uma das partições, depende de uma série de fatores que precisam ser analisados caso a caso. Neste capítulo, serão discutidos os principais fatores. Em todo caso, é possível afirmar que os sistemas de arquivos e dispositivos de armazenamento suportados pelo Xen são os oferecidos pelo Dom0 – variam desde sistemas de arquivos locais a sistemas remotos pelo emprego de Network Attached Storage (NAS) e Storage Area Network (SAN). Lembre-se de que NAS possui um sistema de arquivos associado, usualmente acessado via Network File System (NFS) ou Samba, enquanto um SAN disponibiliza apenas blocos brutos de dados, com acesso por protocolos específicos como iSCSI ou Fiber Channel.

Infraestrutura física de armazenamento

Genericamente, as opções de meios físicos de armazenamento de instalação Xen recaem sobre a escolha entre usar um dispositivo local ou remoto. Assim como em sistemas operacionais convencionais, a latência e a banda passante do disco são importantes para o desempenho global. Entretanto, é importante considerar que no caso do Xen vários domínios U podem tentar executar acessos quase simultaneamente, gerando um volume de dados relevante. Por isso, é preciso projetar o sistema de armazenamento para evitar que se torne um gargalo.

Sobre o armazenamento em discos locais, é necessário considerar soluções que melhorem a latência e a banda passante. O primeiro ponto é a própria tecnologia do disco a ser empregado. A opção mais indicada para servidores são os discos SCSI, que tipicamente oferecem uma maior taxa de rotação – medidos em RPM, o que reduz o tempo de latência rotacional – e um baixo tempo de posicionamento de cabeçotes de leitura e escrita (seek time). Dessa forma, se comparado com outras tecnologias, como discos IDE, ATA e SATA, um disco SCSI apresenta baixa latência de acesso e alta taxa de transferência de dados (banda passante). É possível ainda supor a presença de vários discos físicos, com suas respectivas controladoras, de forma a distribuir entre eles os acessos físicos aos discos.

Outra possibilidade é adotar um sistema de armazenamento remoto. Dois aspectos são particularmente importantes. O primeiro diz respeito à forma pela qual os dados são lidos e escritos no dispositivo remoto. Nessa transferência, podem ser empregados protocolos orientados a blocos, como os usados pelo Fiber Channel (FC) ou pelo iSCSI, ou protocolos orientados a sistemas de arquivos (NFS, Samba e CIFS). Os protocolos orientados a blocos favorecem o desempenho. O segundo aspecto é relacionado à taxa de transferência da rede de interconexão. Quanto maior a velocidade e a taxa de transferência, melhor é o desempenho. É aconselhável ainda que a rede empregada pelo sistema de armazenamento não dispute banda passante com outras aplicações e atividades de gerência do sistema. Recomenda-se, portanto, o uso de uma rede de interconexão exclusiva entre os servidores que executam Xen e o servidor de arquivos.

⓵ Um servidor de arquivos não precisa ser um servidor Xen.

Outra questão a ser observada é a tolerância a falhas. É interessante considerar uma configuração RAID apropriada tanto para discos locais como para discos remotos. A escolha correta do nível de RAID permite uma redundância física de informação no disco, de tal forma que a falha física de um disco pertencente ao RAID pode ser recuperada a partir dos outros discos do RAID. Pela distribuição dos dados nos discos do RAID, os acessos podem ser feitos de forma concorrente, aumentando a banda passante total disponibilizada pelo sistema de armazenamento. O uso de NAS ou SAN também contribui com a tolerância a falhas, pois a falha de um servidor não compromete o acesso ao sistema de armazenamento.

Logical Volume Management (LVM)

Logical Volume Management (LVM) é uma forma de agrupar espaços em disco em porções denominadas volumes lógicos, que podem ser redimensionados ou removidos. Com o emprego do LVM, é possível redimensionar partições de um sistema de armazenamento sem refazer o particionamento físico do disco. Essencialmente, o LVM tem cinco componentes:

- **Volume físico** (Physical Volume – PV): empregado para designar um dispositivo físico de armazenamento, como um disco rígido, um sistema RAID ou uma partição.
- **Unidades físicas** (Physical Extents – PE): são unidades físicas de alocação de dados, similares aos blocos de dados. Um volume físico (PV) é dividido com base nessa unidade.
- **Volume lógico** (Logical Volume – LV): equivalente a uma partição tradicional de disco e, por conseguinte, pode ser formatado logicamente para conter um sistema de arquivos.
- **Unidade Lógica** (Logical Extents – LE): unidades lógicas de alocação de dados, similares aos blocos de dados e empregadas no volume lógico. As unidades lógicas (LE) possuem o mesmo tamanho das unidades físicas (PE).
- **Grupo de volume** (Volume Group – VG): corresponde ao nível mais alto de abstração do LVM. Reúne uma coleção de volumes lógicos (LV) em uma mesma unidade de gerenciamento. Em analogia com discos tradicionais, é possível considerar um grupo de volume, como um disco que contém várias partições (LVs), que são compostas por blocos de dados (LEs).

Dessa forma, um disco físico (PV) é dividido em unidades físicas (PEs) que correspondem às unidades lógicas (LEs). Um conjunto de LEs forma um volume lógico (LV); por sua vez, um conjunto de LVs forma um grupo de volume (VG). É importante salientar que os LEs empregados para compor um volume lógico podem ser provenientes de discos físicos diferentes. Esse tipo de estrutura permite que, para aumentar a capacidade de armazenamento de uma instalação já existente, seja suficiente adquirir um novo disco físico, formatá-lo em unidades físicas (PEs) do LVM, associá-las a unidades lógicas (LEs) e fazer a distribuição entre os diferentes volumes lógicos (LVs) conforme a necessidade. É possível ainda retirar unidades lógicas de um volume lógico e atribuí-las a outro. A figura a seguir ilustra a relação entre esses elementos.

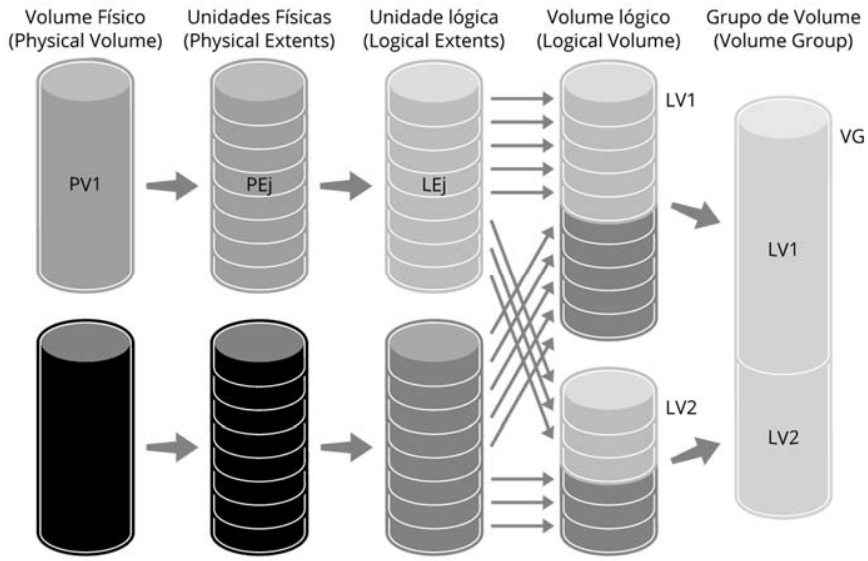


Figura 4.1 Logical Volume Management (PVs, PEs, LEs, LVs e VG).

O emprego de partições LVM é recomendável para o redimensionamento do espaço de disco usado pelas máquinas virtuais, de acordo com a necessidade.

Migração e sistemas de armazenamento

A migração é uma das grandes facilidades oferecidas pelo Xen. É possível migrar domínios de um servidor físico a outro, mas com algumas considerações relativas aos domínios 0 e U. Existem duas modalidades de migração: normal e "a quente" (live). Na migração normal, o domínio é parado e congelado, formando um arquivo-imagem com contexto de execução, que é transferido para outro servidor é lá é reinicializado. Na migração live, o domínio virtual é transferido para outro servidor sem a necessidade de interromper a execução.

Entretanto, a migração de domínios entre servidores físicos só é possível se o tipo de armazenamento configurado no domínio a ser migrado for do tipo remoto, ou seja, se todas as partições montadas pelo domínio não forem dispositivos físicos acessados diretamente do Dom0 local. Em uma instalação típica, o Dom0 não sofre migração e, embora possa empregar sistemas de arquivos remotos via NFS ou mesmo Samba, ele é instalado diretamente em um disco local com um sistema de arquivos qualquer provido pelo Linux. Já os DomU podem ser migrados "a quente".

Vale observar que, em uma migração, apenas as informações relacionadas ao contexto de execução do processador e de memória são transferidas de um servidor físico a outro. Todos os arquivos pertencentes a um domínio permanecem em seu lugar de origem no sistema de arquivos remoto. Por motivo de configuração de rede e funcionamento dos sistemas de arquivos remotos, o servidor físico de destino deve estar na mesma sub-rede que o servidor físico de origem e o servidor de arquivos remoto. Portanto, a migração é um fator importante a ser considerado no projeto de sistema de armazenamento que será utilizado e na própria infraestrutura de rede.

Rede

De maneira similar ao que acontece com os discos físicos, os requisitos de rede de uma instalação Xen dependem do volume de tráfego de dados gerado pelas aplicações e do número de máquinas virtuais instanciadas em cada servidor físico. Tipicamente, uma instalação Xen emprega uma rede do tipo Ethernet. Existe grande variedade de placas de rede Ethernet com seus respectivos drivers para o Linux. Isso não constitui um problema.

Como um único servidor físico é capaz de executar múltiplas máquinas virtuais – que efetuam tarefas envolvendo acesso à rede –, pode haver uma forte competição pelo recurso de interface física de rede e, assim, comprometer o desempenho da solução de virtualização. Em função disso, recomenda-se que sejam empregadas interfaces de rede de pelo menos 1 Gbit/s, além de usar mais de uma interface física de rede por servidor, para separar os diferentes tráfegos existentes:

- Tráfego normal, gerado pelas aplicações;
- Tráfego de storage e gerenciamento.

Uma característica interessante do Xen, relacionada à rede, é sua capacidade de oferecer e configurar uma infraestrutura de rede virtual. É possível definir interfaces e criar uma topologia de rede permitindo a comunicação entre os domínios convidados, além de efetuar NAT, bridging e aplicar regras de firewall, tudo completamente virtualizado. A infraestrutura virtual é tão flexível e poderosa quanto àquela que pode ser obtida com os comandos Linux brctl e iptables. O brctl permite criar, gerenciar e monitorar bridges ethernet, enquanto o iptables possibilita a definição de regras de firewall.

Infraestrutura de software para o Xen

- Distribuições comerciais e livres
- Hipervisor e sistemas operacionais para Dom0
- Sistemas Operacionais para DomU

O primeiro passo para instalar o Xen é, obviamente, obtê-lo. Uma instalação Xen é formada pelo hipervisor, pelo domínio 0 e por domínios U. O hipervisor e o Dom0 se confundem. Já o DomU é composto por uma imagem do sistema de arquivos raiz, um núcleo modificado para o Xen (no caso de paravirtualização) e um arquivo de configuração que descreve os recursos, como memória, processador e dispositivos de E/S. A seguir, serão discutidos alguns detalhes sobre a obtenção e instalação desses componentes.

Distribuições comerciais e livres

Desenvolvido dentro da filosofia de software livre, o Xen é disponibilizado de duas formas: distribuições comerciais e livres. A diferença fundamental entre elas é o suporte à ferramentas específicas (ou melhoradas) para instalação, configuração e monitoramento do Xen, assim como suporte, treinamentos e manuais específicos. Além disso, atividades relacionadas à migração, alta disponibilidade e recuperação a falhas normalmente são

oferecidas nas distribuições comerciais. As distribuições livres não fornecem essas facilidades ou as disponibilizam em versões de funcionalidades limitadas. As principais distribuições comerciais do Xen são: Citrix XenServer Product Group (XenServer Standard e XenServer Enterprise); Virtual Iron Software Inc. (Virtual Iron Virtualization Manager and Service); SUSE Linux Enterprise Server 10; e Red Hat Enterprise Linux 5 (RHEL 5).

Na opção por uma distribuição livre do Xen, basta empregar uma distribuição Linux convencional que tenha suporte ao hipervisor Xen. As distribuições Linux mais conhecidas são Ubuntu, Fedora, Debian, OpenSuse e Gentoo – além do XenServer, que é derivado da distribuição CentOS. É possível ainda usar o Xen com distribuições que não sejam Linux, como, por exemplo, OpenSolaris, FreeBSD, OpenBSD e NetBSD. Assim como vários outros pacotes de software, o Xen também é disponibilizado de forma binária e pode ser instalado via gerenciadores de pacotes da distribuição (apt-get, YaST, yum), ou por meio de seus arquivos-fonte em tarball, que precisam ser configurados, compilados e instalados. A escolha entre as distribuições disponíveis, como pelas opções de binários pré-compilados ou tarballs, é feita mais por familiaridade que por critério ou requisito técnico específico.

Hipervisor e sistemas operacionais para Dom0

O hipervisor Xen, na verdade, é uma camada de software que age como se fosse uma espécie de micronúcleo. Ele é usado para realizar algumas tarefas fundamentais, como escalonamento e gerência de memória, além de tratar interrupções e gerir as chamadas de sistemas do Xen, denominadas hypercalls. O hipervisor, sozinho, não tem valia e sua instalação é sempre confundida com a do Dom0. Ao instalar o Dom0, o hipervisor é instalado junto. Para o Dom0, pode ser empregada qualquer uma das distribuições comentadas.

Sistemas operacionais para DomU

Os sistemas operacionais das máquinas virtuais hóspedes dependem do tipo de virtualização empregada. Como foi discutido, o Xen oferece duas formas de virtualização: completa e paravirtualização. No caso da virtualização completa, a instalação do Sistema Operacional para o DomU segue exatamente os mesmos procedimentos e condições de um sistema convencional não virtualizado. Isso significa que é necessário ter as mídias de instalação e, principalmente, respeitar as condições de licenciamento do SO a ser instalado. Só é possível executar no DomU uma máquina virtual com Sistema Operacional não modificado, como é o caso da família Microsoft Windows, se o processador oferecer suporte em hardware para a virtualização (HAV).

Sem a possibilidade de suporte, é necessário empregar a paravirtualização. Nesse caso, é preciso ter um Sistema Operacional convidado que tenha sido modificado para interagir com o Xen. Isso restringe o DomU a empregar sistemas operacionais que seguem uma filosofia open source, como o Linux, o OpenSolaris, o FreeBSD e o NetBSD. Lembre-se de que, conforme vimos no capítulo anterior, existe outra alternativa para o uso de núcleos não modificados: a virtualização híbrida (ainda em fase de desenvolvimento).

Geração de máquinas virtuais paravirtualizadas

- Inicialização de sistemas Linux
- Clonagem
- Imagens preexistentes
- Ferramentas específicas às distribuições Linux
- Qemu

Um domínio 0 paravirtualizado é formado por um núcleo com suporte ao Xen, um sistema de arquivos e um arquivo de configuração. Dessa forma, para instanciar uma máquina virtual para DomU é preciso fornecer um núcleo e definir o sistema de arquivos. O núcleo pode ser obtido de três formas:

- Utilizar a própria imagem do núcleo do domínio 0;
- Usar uma imagem de um núcleo já pronta;
- Compilar um núcleo.

A vantagem das duas primeiras é que empregam um núcleo pronto e funcional, embora ele possa não ter todos os módulos e funcionalidades desejados para um DomU. Esse inconveniente é contornado quando se compila o núcleo, pois, dos eventuais problemas que possam ocorrer durante o processo, é possível customizá-lo para atender aos requisitos do DomU. Um ponto importante a ressaltar é que, independentemente da forma como se obtém o núcleo, sua imagem é armazenada fora do DomU.

Há quatro formas para configurar um DomU paravirtualizado: clonagem, imagens preexistentes, uso de ferramentas específicas de uma distribuição Linux e via qemu. Porém, antes de apresentar mais detalhes sobre os modos de instalação da paravirtualização, é interessante salientar alguns aspectos da inicialização e carga de um Sistema Operacional GNU/Linux convencional, para uma melhor compreensão desses métodos.

Inicialização de sistemas Linux

Na inicialização de um Sistema Operacional, estão envolvidos três componentes fundamentais: carga e inicialização (bootstrap); imagem do núcleo; e o sistema de arquivos raiz (Root File System – RFS). O procedimento de carga e inicialização é responsável por identificar o dispositivo usado para armazenar a imagem do núcleo do SO e de um Root File System inicial e temporário. Na maior parte dos ambientes Linux, a carga é separada em duas etapas. Na primeira, é posto em memória um pequeno código – responsável por transferir para ela o programa que efetivamente realizará a carga do núcleo do Sistema Operacional (lilo ou grub). O carregador do núcleo copia para a memória a imagem do núcleo e, na sequência, transfere para ele o controle. A imagem do núcleo é um arquivo compactado, que normalmente tem um nome como vmlinuz.

Em seguida, é preciso carregar os drivers de dispositivos, montar o sistema de arquivos e, então, executar os scripts de inicialização (aqueles que estão em /etc). Mas, nisso, surge um

problema: para ler o disco e inicializar o sistema é necessário que os drivers já tenham sido carregados, e o sistema de arquivos já esteja montado. A solução para o impasse é empregar um sistema de arquivos raiz (temporário) inicial, com o mínimo de funcionalidades; armazenado em RAM. Esse primeiro sistema de arquivos faz parte do Initial RAM Disk (initrd). O initrd possui um pseudo-driver de dispositivo para emular em RAM um disco onde está armazenado o sistema de arquivos raiz inicial. O initrd pode ser um arquivo separado da imagem do núcleo (vmlinuz) ou estar embutido nela. No caso de arquivo separado, ele deve ser carregado em memória antes do controle e da inicialização serem passados ao núcleo.

Uma vez que todos os drivers necessários estejam carregados, e tenham sido montados os sistemas de arquivos definitivos, o sistema de arquivos raiz inicial é substituído pelo sistema de arquivos raiz real, armazenado em disco. Importante salientar que o núcleo (kernel) e o sistema de arquivos raiz (Root FS) são desvinculados entre si, assim como aqueles montados depois. Por isso, é possível atualizar o núcleo de uma instalação Linux sem a necessidade de reinstalar todo o sistema.

Root File System

É importante não confundir o uso da palavra inglesa *root*. Root File System indica a raiz inicial do sistema de arquivos, que é representado e iniciado em /, e não o sistema de arquivos (diretórios e arquivos) pertencente ao usuário root (administrador do sistema).

Clonagem

Este método parte de uma instalação já existente com um núcleo modificado para interagir com o Xen. O primeiro passo é criar um arquivo inicialmente vazio, que serve como uma espécie de partição virtual de disco, onde os diretórios e arquivos do root FS serão armazenados. Além dele, é possível criar outro arquivo para servir como partição de swap para o DomU. Tipicamente, esses arquivos são criados com o auxílio do comando Linux dd e formatados logicamente com o comando mkfs. O resultado final é análogo a ter (vitualmente) duas partições de disco, onde o sistema convidado será instalado.

Na sequência, é preciso popular o root FS do sistema convidado (DomU) com os arquivos necessários. Esses arquivos são copiados do sistema de arquivos do Sistema Operacional convencional usado como base para gerar o convidado. A parte mais complicada dessa cópia é analisar, no root FS do SO base, o que pode ou deve ser copiado para o root FS do DomU, que está sendo gerado.

Com as imagens de um root FS e uma área de swap, basta definir um arquivo de configuração Xen para mapear e montar esses arquivos como dispositivos de E/S no Sistema Operacional convidado (DomU). Esse arquivo de configuração, além da imagem de um núcleo (vmlinuz) e de um RFS inicial (initrd), define outros parâmetros, como nome da máquina, memória a ser usada pela máquina virtual e endereço IP. Para executar a máquina virtual, é preciso apenas empregar o comando Xen xm com a opção create, passando como parâmetro o nome do arquivo de configuração.

Imagens preexistentes (appliances)

Criar as imagens necessárias para o DomU é uma tarefa que exige algum conhecimento e paciência; mesmo com os devidos cuidados, ainda está sujeita a erros. Uma forma de facilitar a tarefa de obter um DomU é empregar imagens preexistentes, denominadas appliances. A desvantagem desse método é que uma appliance pode conter softwares desnecessários e, eventualmente, usar módulos que não estão necessariamente instalados na máquina-alvo. Diversos sites disponibilizam appliances, sendo os mais conhecidos:

- jailtime.org
- virtualappliances.net
- jumpbox.com
- rpath.com

As appliances costumam ser disponibilizadas em três formatos de arquivos: imagem de disco, imagem de partição e imagem comprimida. Na prática, esses modelos são empregados de modo semelhante: deve-se recuperar a imagem, instalá-la de forma adequada em um sistema de arquivos e, eventualmente, adaptar os arquivos de configuração. Os detalhes para instanciar um DomU com uma appliance estão normalmente nos próprios sites que as oferecem.

Ferramentas específicas às distribuições Linux

As principais distribuições Linux concedem ferramentas específicas para gerenciar máquinas virtuais e, em especial, auxiliar na criação do DomU. Entre as comumente empregadas, citamos:

- **YaST Virtual Machine Management** (OpenSUSE): é um módulo que pode ser instalado junto com o Yast, que permite, de forma gráfica, criar e gerenciar DomU. Nessa ferramenta, é possível escolher entre instalar um Sistema Operacional de forma convencional – via mídia eletrônica (cd/DVD) – ou empregar como modelo uma instalação já existente.
- **Virt-manager** (CentOS/Fedora): parte integrante de um conjunto de ferramentas de gerenciamento e monitoramento de máquinas virtuais denominado virt-tools (xen-tools). O virt-manager oferece uma interface gráfica e costuma ser disponibilizado no menu “Aplicações”, submenu “Ferramentas de sistema”. Seu uso é bastante intuitivo. Tipicamente, os passos consistem em definir um hardware virtual (nome do sistema, disco, memória) e, na sequência, escolher o tipo de instalação a ser feita, entre as opções baseadas em uma appliance ou instalação via mídia eletrônica (cd/DVD).
- **debootstrap** (Debian/Ubuntu): pacote que pode ser instalado no sistema por meio de apt-get. Essencialmente, o deboostrap permite fazer a instalação de um sistema em um diretório qualquer. O sistema de arquivos resultante desse procedimento corresponde ao root FS do DomU que está sendo gerado.

O procedimento geral executado por qualquer uma dessas ferramentas é o mesmo. As diferenças fundamentais entre elas estão na facilidade de uso. Nesse quesito, há certa preferência geral por virt-manager. Observe que é possível gerar imagem para DomU com uma ferramenta qualquer e, posteriormente, usar a imagem gerada, customizada, como se fosse uma appliance de terceiros. Novamente, um DomU nada mais é que uma imagem de núcleo, um sistema de arquivos e um arquivo de configuração.

Qemu

- Drivers virtuais
- Software livre
- Emulador de processadores

O Qemu é um software livre (licença GNU LGPL) que emula uma série de processadores do mercado e também oferece um conjunto vasto de dispositivos de E/S virtuais, ou seja, na prática o Qemu é o hardware de uma máquina virtual. Uma forma simples de gerar imagem para DomU é instalar o Qemu e, a partir dele, fazer a instalação do Sistema Operacional desejado para o DomU. Uma vez concluída a instalação, é possível copiar o sistema de arquivos gerado nessa instalação para ser usado como sistema de arquivos do DomU. A imagem do núcleo a ser empregada no DomU pode ser a do Dom0 ou de outro núcleo compatível com o Xen. Por fim, como nos casos anteriores, é necessário criar um arquivo de configuração que defina os recursos da máquina virtual (memória RAM, disco e rede) e descreva onde estão os arquivos-imagem do núcleo e do sistema de arquivos.

A vantagem é que o usuário faz a instalação de um Sistema Operacional no Qemu de forma tradicional, como se fosse uma máquina física. Dessa forma, o nível de detalhamento, complexidade e conhecimento é reduzido; basta copiar as imagens do núcleo e do sistema de arquivos e fazer o arquivo de configuração.

Gerenciamento de domínios convidados

- Comandos de linha
- Virt-manager
- XenCenter

O Xen possui uma série de ferramentas que permitem a criação, o gerenciamento e o monitoramento dos domínios convidados. A base de funcionamento é o processo xend, que, sendo um daemon, executa em background no domínio convidado 0 com privilégios de usuário root.

Comando de linha

O comando de linha xm (xen management) é o mecanismo fundamental para se trabalhar com um sistema composto por um domínio 0 e vários domínios U. Sua interface é feita por um console no domínio 0, que pode ser obtido de forma remota (SSH, por exemplo). A partir do comando xm é possível criar, examinar e terminar domínios, obter informações sobre

domínios e adicionar e remover dinamicamente dispositivos de bloco e de rede. Fazem parte dos comandos de linha xenstore, xenmon, xenperf, xentrace e xentop.

A desvantagem dos comandos de linha é justamente o fato de se empregar uma interface de texto por meio de uma Command Line Interface (CLI). Para suprir essa necessidade, existem vários projetos em código aberto que criam interfaces gráficas de gerenciamento. Essas interfaces fornecem acesso às funcionalidades do Xen de forma simplificada e intuitiva, se comparada com a linha de comando. Entre as principais iniciativas estão XenMan, Enomalism, Manage Large Networks (MLN) e OpenRQM. Também existem ferramentas específicas para distribuições, produtos comerciais em licenças freeware ou proprietárias.

Virt-manager

RedHat e Fedora disponibilizam ferramentas customizadas para a criação e gerenciamento de máquinas virtuais Xen. Nessas distribuições, existem as opções de linhas de comando (virsh e virt-install) e uma interface gráfica, o Virtual Machine Manager (virt-manager). O virt-manager é uma interface gráfica que pode ser executada em um desktop e oferece facilidades para criar, monitorar o desempenho e visualizar recursos e estatísticas dos domínios convidados. A criação de domínios é feita de forma simplificada por um "assistente de instalação" (wizard), que permite a definição de recursos do hardware virtual, guiando a instalação do sistema. O virt-manager possui embutido um cliente VNC, que possibilita o acesso a uma console gráfica e a um domínio convidado. A distribuição SUSE também oferece, pelo YaST2, uma ferramenta gráfica intuitiva para trabalhar com domínios convidados. Essas ferramentas são as mais empregadas na linha de distribuições Linux; entre as distribuições mais difundidas em servidores virtualizados Xen está a CentOS (derivada da RedHat), e as da Novell, que utilizam o SUSE.

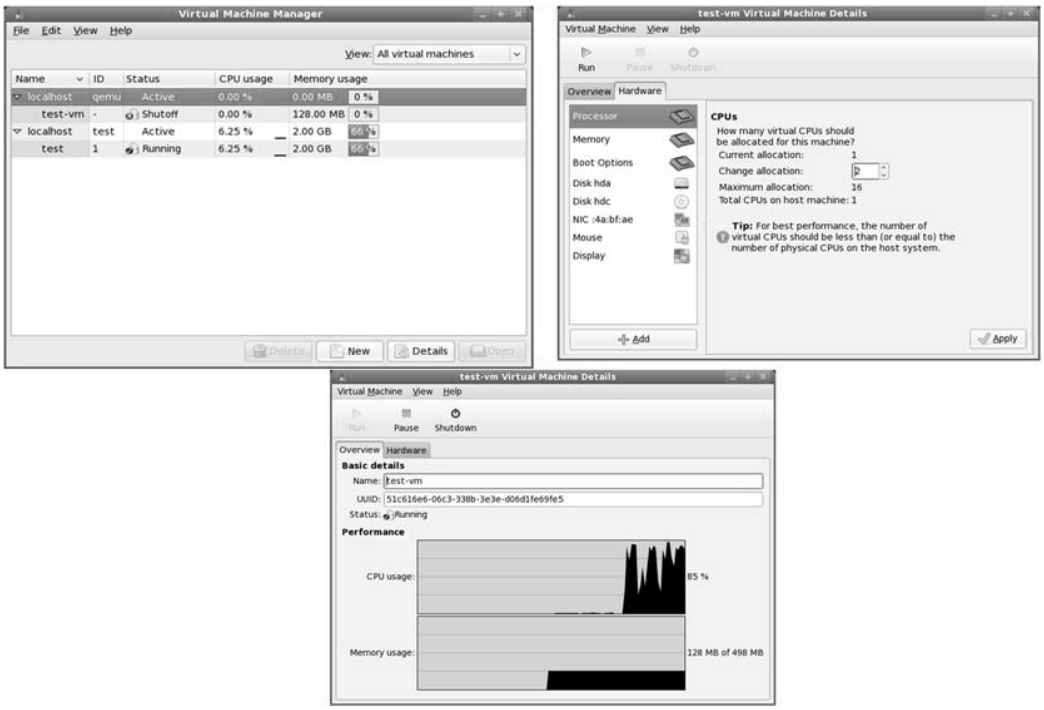


Figura 4.2 Telas de apresentação do virt-manager.

Xen Center

Citrix XenCenter é uma aplicação Windows que oferece uma interface gráfica para gerenciar o Citrix XenServer, cujas principais funcionalidades são a criação, o gerenciamento e o monitoramento de máquinas virtuais. De forma similar às interfaces gráficas das distribuições Linux, o XenCenter permite a conexão a domínios convidados, por meio de uma interface gráfica. O grande diferencial do XenCenter em relação às demais ferramentas é que permite realizar migrações de máquinas virtuais e controlar o balanceamento de carga (de forma gráfica). É possível fazer o download do Citrix XenCenter na página oficial da Citrix Inc. A figura a seguir mostra telas do Xen Center.

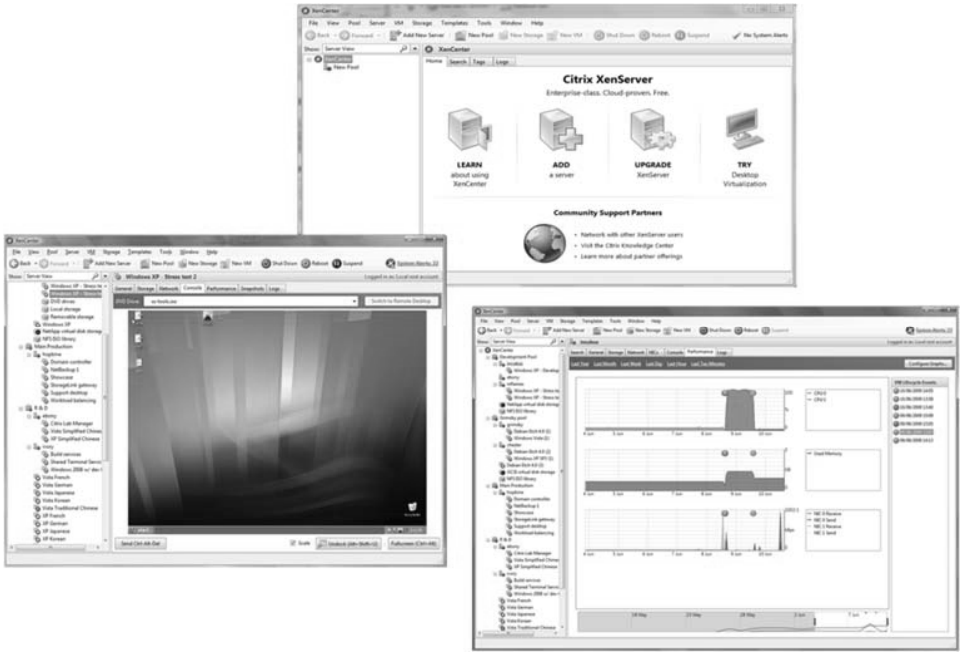


Figura 4.3 Telas de apresentação do XenCenter.

Arquitetura genérica de virtualização

- Migração e alta disponibilidade (HA)
- Servidor de licenças
- Repositório de imagens

Em uma solução corporativa ou de missão crítica, há aspectos inter-relacionados que precisam ser considerados e levam a uma arquitetura genérica para a solução de virtualização. Os pontos fundamentais são a capacidade de migração e de alta disponibilidade (High Availability). A migração consiste em mover uma máquina virtual de um servidor físico para outro. As razões para isso vão desde paradas programadas para manutenção a balanceamento de carga. No entanto, para que seja viável, o sistema de arquivos não pode ser local. De forma similar, quando uma máquina física falha e a recuperação consiste em instanciar as máquinas virtuais em outro servidor funcional, é preciso que o sistema de arquivos das máquinas virtuais seja acessível a partir desse novo servidor. A solução imediata é empregar um sistema de arquivos remoto, o que justifica o uso de sistemas de armazenamento (storage) com tecnologias NAS ou SAN. A garantia de tolerância a falhas do storage é mantida com investimentos em soluções redundantes e de alta confiabilidade, como, por exemplo, sistemas baseados em RAID.

Como o uso de storage gera um volume de acessos (tráfego), é desejável que o storage empregue um mecanismo de interconexão isolado dos demais. Tal interconexão pode ser proprietária e específica de uma tecnologia de disco. Outra opção é adotar uma rede ethernet e, nesse caso, é importante que ela seja destinada ao tráfego de dados entre as máquinas

virtuais e o storage – para que não haja concorrência com o tráfego normal de dados – e possua uma taxa de transferência (1 ou 10 Gbit/s).

Uma consequência da migração e da alta disponibilidade é o fato de um sistema hóspede, e mesmo o hipervisor, não terem uma licença vinculada a uma máquina física. A solução consiste em ter um conjunto de licenças compartilhado pelas máquinas e hipervisores. Um sistema pode recuperar a licença liberada em função da falha de outro. Para que esse mecanismo funcione, surge a ideia de um servidor de licenças. No caso específico do Xen, o hipervisor possui licenças para o uso das versões Advanced, Enterprise e Platinum, além de licenças de sistemas operacionais proprietários e de aplicações. O servidor de licenças deve ser uma máquina Windows Server.

Por fim, as imagens ISO que serão utilizadas nos domínios U precisam estar armazenadas em algum local. A solução é colocá-las em um lugar acessível por qualquer máquina. Surge, então, a noção de um repositório de imagens ISO, que usualmente emprega um sistema de armazenamento remoto.

⓵ Detalhe importante: no caso do Xen, o repositório de imagens é separado daquele utilizado para armazenar os sistemas de arquivos das máquinas virtuais.

Xen Cloud Platform

O Xen Cloud Platform (XCP) é um open source de virtualização de servidores enterprise-ready e uma plataforma de computação em nuvem. Entrega o hipervisor Xen com suporte para uma variedade de sistemas operacionais convidados, incluindo Windows, Linux e suporte de rede, armazenamento e ferramentas de gerenciamento.

XCP atende às necessidades dos provedores de computação em nuvem, combinando o isolamento e a multialocação de recursos do hipervisor Xen – como segurança avançada, armazenamento e tecnologias de virtualização de rede –, para oferecer um rico conjunto de serviços em nuvem na infraestrutura virtual. A plataforma também aborda as necessidades dos utilizadores em matéria de segurança, desempenho, disponibilidade e isolamento entre nuvens privadas e públicas.

Consolida as cargas de trabalho do servidor, permite a economia de energia, refrigeração e gerenciamento de custos – contribuindo para a computação ecologicamente sustentável –, uma maior capacidade de adaptação à constante mudança nos ambientes de TI, o uso otimizado do hardware existente e a melhoria do nível de confiabilidade da área de TI.

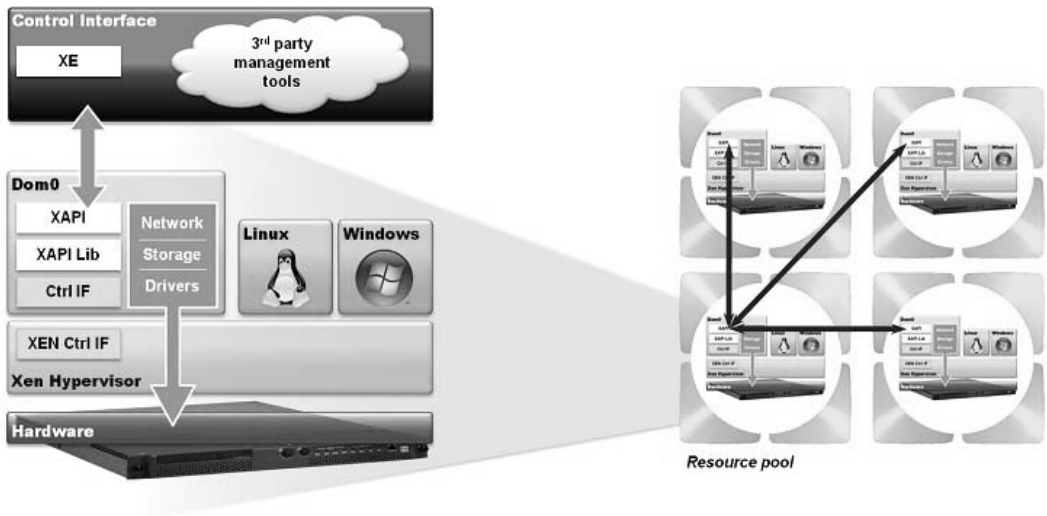


Figura 4.4 Arquitetura XCP. Fonte: Citrix.

O XCP foi originalmente derivado do Citrix XenServer. Hoje, o código dele está sob a Licença Pública Geral GNU (GPL2), disponível sem custo, tanto de origem quanto de formato binário. Ele é e sempre será de código aberto, unindo a indústria e o ecossistema Xen para acelerar a adoção de tecnologias de virtualização e computação em nuvem. Trabalha ativamente com open source e padrões abertos para ajudar a superar os desafios da mobilidade em nuvem.

5

Introdução ao Hyper-V

- Características gerais
- Versões
  - Windows Server 2012 com Hyper-V
  - Hyper-V Server 2008
- Benefícios
- Licenciamento

No Windows Server 2012, o suporte à virtualização de servidores está disponível como um recurso do Sistema Operacional (perfil) denominado Hyper-V. Ele pode ser considerado uma função do SO. A Microsoft recomenda que o Hyper-V seja o único perfil habilitado em servidores que empregam soluções de virtualização.

Segundo a fabricante, o Windows Server 2012 com Hyper-V é uma tecnologia de virtualização baseada em hipervisor, que permite a melhor aplicação dos investimentos em hardware de servidor por meio da consolidação de múltiplos servidores em máquinas virtuais separadas, executadas em um único computador físico. Com o Hyper-V, é possível rodar vários sistemas operacionais – como, por exemplo, versões diferentes de Windows, GNU/Linux, entre outros –, em um único servidor físico e aproveitar o poder de processamento das arquiteturas de 64 bits.

O Windows Server 2012 com Hyper-V fornece uma plataforma de virtualização dinâmica, confiável e escalável. Somada a um conjunto de ferramentas de gerenciamento integradas, tanto para os recursos físicos como para os recursos virtuais, permite a criação de um datacenter dinâmico.

Características gerais

Principais características do Hyper-V:

- **Arquitetura de 64 bits**: o Hyper-V foi concebido para explorar as capacidades das arquiteturas de 64 bits, possibilitando, assim, um bom desempenho.
- **Suporte a Sistemas Operacionais**: capacidade para execução simultânea de diferentes tipos de sistemas operacionais convidados, incluindo sistemas de 32 bits e 64 bits, como Windows, Linux e outros.
- **Suporte a Multiprocessadores Simétricos (SMP)**: suporte para até quatro processadores no ambiente de uma máquina virtual, o que permite o aproveitamento dos aplicativos multithreaded.

- **Arquitetura de compartilhamento de hardware**: estrutura criada para oferecer acesso de forma eficiente e segura aos drivers de dispositivos de E/S. Baseada em um modelo provedor-consumidor de serviço (**VSP/VSC**) que será detalhado adiante.
- **Snapshot de máquina virtual**: o Hyper-V oferece a capacidade de gerar imagens congeladas (snapshots) de uma máquina virtual em execução. Isso permite que um usuário possa retornar facilmente a um estado anterior e aprimorar a solução de backup e recuperação de falhas.
- **Escalabilidade**: por oferecer suporte a múltiplos processadores (e núcleos) e acesso avançado à memória nas máquinas virtuais, pode-se fazer a escalabilidade vertical do ambiente de virtualização. O objetivo é oferecer suporte a uma grande quantidade de equipamentos virtuais, dentro de um determinado servidor e, posteriormente, continuar a migração rápida para a escalabilidade horizontal, por meio de múltiplos servidores.
- **Extensível**: as interfaces e APIs do Hyper-V seguem o padrão Word Manufactorers Identification (WMI), permitindo que desenvolvedores e fabricantes de software independentes produzam com rapidez ferramentas personalizadas, como utilitários e aprimoramentos na plataforma de virtualização.

VSP/VSC

A arquitetura de virtualização da Microsoft (Windows Server Virtualization – WSV) é inspirada em um modelo cliente-servidor, no qual o Virtualization Service Provider (VSP) executa na partição pai, sendo o único componente autorizado a acessar os drivers de dispositivos reais. As partições filho executam o Virtualization Service Client (VSC), que oferece drivers virtuais de dispositivos. As requisições feitas a um driver virtual são recebidas pelo VSC, que as reencaminha para o VSP.

Versões

Existem basicamente duas versões de Hyper-V:

- Windows Server 2012 Hyper-V;
- Microsoft Hyper-V Server.

A tabela 5.1 compara as versões do Windows Server 2012, fazendo uma relação entre os recursos de cada edição e a forma de licenciamento.

Edição	Comparação de recursos	Modelo de licenciamento
Foundation	1 processador Recursos limitados	Por Servidor Limite de 15 usuários
Essentials	2 processadores Recursos limitados	Por Servidor Limite de 25 usuários
Standard	Duas instâncias virtuais Todos os recursos	Processador CAL**
Datacenter	Instâncias virtuais ilimitadas Todos os recursos	Processador CAL**

Tabela 5.1 Versões do Hyper-V.

** CALs (client access licenses – licenças de acesso para o cliente): licença para cada usuário ou dispositivo que acessará o servidor.

Como pode ser visto na tabela 5.1, o licenciamento Microsoft foi revisto e simplificado, se comparado com as versões anteriores. No novo modelo, o usuário pode definir qual versão melhor o atende, considerando alguns aspectos.

- **Foundation Edition:** para pequenas empresas com até 15 usuários e servidores com um único processador.
- **Essentials Edition:** para pequenas empresas com até 25 usuários e servidores com até dois processadores.
- **Standard Edition:** para ambientes não virtualizados ou levemente virtualizados.
- **Datacenter Edition:** para ambientes de nuvem privada altamente virtualizados.

Observa-se nos tópicos apresentados acima que apenas as versões Standard e Datacenter podem ser utilizados para infraestruturas de virtualização.

O produto Microsoft Hyper-V Server é uma solução gratuita e mais simples. Oferece apenas a opção de consolidação de servidores e a definição de máquinas virtuais com diferentes sistemas operacionais. A última característica permite que sejam criados ambientes de teste e desenvolvimento de software para outras plataformas.

Benefícios

O Windows Server 2012 Hyper-V simplifica a interação entre o hardware, o Sistema Operacional e as máquinas virtuais, enquanto fortalece os componentes centrais da virtualização. O Windows Server 2012 Hyper-V concede confiabilidade e escalabilidade com uma arquitetura de hipervisor construída sobre um micronúcleo (micro-kernel), que não

inclui qualquer device driver de terceiros. Os device drivers utilizados são aqueles desenvolvidos para o Sistema Operacional Windows nativo.

Para a Microsoft, a virtualização representa uma oportunidade de aumentar o nível de segurança. As funcionalidades providas pelo Hyper-V incluem proteção de rede como NAT, firewall e via políticas de acesso. Isso permite que as máquinas virtuais gerem regras de segurança (para evitar a exposição dos servidores compartilhados) e, assim, reduzam o risco de ataques.

O desempenho é um aspecto fundamental da virtualização. Para atender a esse importante requisito, o Hyper-V inclui:

- Suporte a processadores multicore, em que cada máquina virtual pode ter acesso a até quatro processadores lógicos;
- Desempenho otimizado, devido ao baixo custo computacional oferecido pela arquitetura do Hyper-V;
- Suporte a máquinas virtuais que executam sistemas operacionais de 64 bits;
- Acesso a até 64 GB de memória por máquina virtual.

Licenciamento

Os direitos de virtualização por licenciamento de servidor em relação ao uso do Hyper-V precisam ser considerados, variando de acordo com a versão do Windows Server. A tabela abaixo ilustra os requisitos de licença por servidor.

Microsoft Hyper-V Server 2008	Windows Server 2012 Standard	Windows Server 2012 Datacenter
Cada máquina virtual (VM) guest requer uma licença de servidor	1 computador físico + 2 VM	1 computador físico + número ilimitado de VMs

ⓘ O Windows Server 2012 pode ser licenciado sem o Hyper-V.

Componentes da arquitetura Hyper-V

- Hipervisor
- Partições pai e filho
- Máquinas virtuais e sistemas operacionais convidados
- Dispositivos sintéticos e emulados
- Serviços integrados

O hipervisor é o componente central do Hyper-V, responsável por criar e gerenciar ambientes isolados de execução (chamados de partições). A arquitetura do hipervisor é aquela de um monitor de máquina virtual, mas que segue uma filosofia denominada micronúcleo. Nessa arquitetura, em oposição à arquitetura chamada de monolítica, o hipervisor possui apenas as funções essenciais para gerenciar os recursos de hardware –

como atendimento a interrupções, escalonamento do processador (ou núcleos) e gerência de memória. Os drivers de dispositivos são externos ao hipervisor. A próxima figura ilustra as diferenças entre a arquitetura monolítica e a arquitetura baseada em micronúcleo.

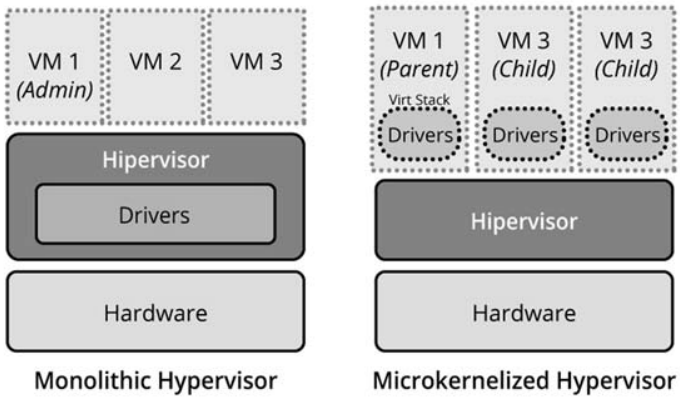


Figura 5.2 Arquitetura monolítica *versus* micronúcleo.

Quando o Hyper-V é carregado pela primeira vez, cria-se uma partição chamada de partição raiz ou partição pai. Ela hospeda a instância do Windows Server 2012 que estava executando antes da habilitação do perfil Hyper-V. A partição pai controla os dispositivos de hardware e responde pela alocação de memória em si, além de solicitar ao hipervisor a criação das partições filho, que não têm acesso direto ao hardware.

As requisições de E/S de uma máquina virtual são encaminhadas para os adaptadores físicos do sistema por meio da partição pai. O modelo indireto de E/S utilizado pelo Hyper-V permite que os equipamentos virtuais sejam independentes de tipos específicos de dispositivos de hardware no servidor físico. Os dispositivos virtuais que uma máquina expõe para o Sistema Operacional convidado podem ser de dois tipos:

- Dispositivos virtuais emulados;
- Dispositivos virtuais sintéticos.

VMBUS
Canal de comunicação entre os VSCs das partições filho e o VSP da partição pai.

Dispositivos virtuais emulados são uma implementação feita por softwares de um dispositivo PCI tradicional. Para um Sistema Operacional convidado, o dispositivo emulado aparece como um PCI físico. Dispositivos virtuais sintéticos são executados por softwares e baseados na arquitetura VSP-VSC do Hyper-V. Esses dispositivos utilizam um canal de comunicação **VMBUS** entre as diferentes partições. Os dispositivos sintéticos têm menor custo computacional. A próxima figura ilustra os componentes.

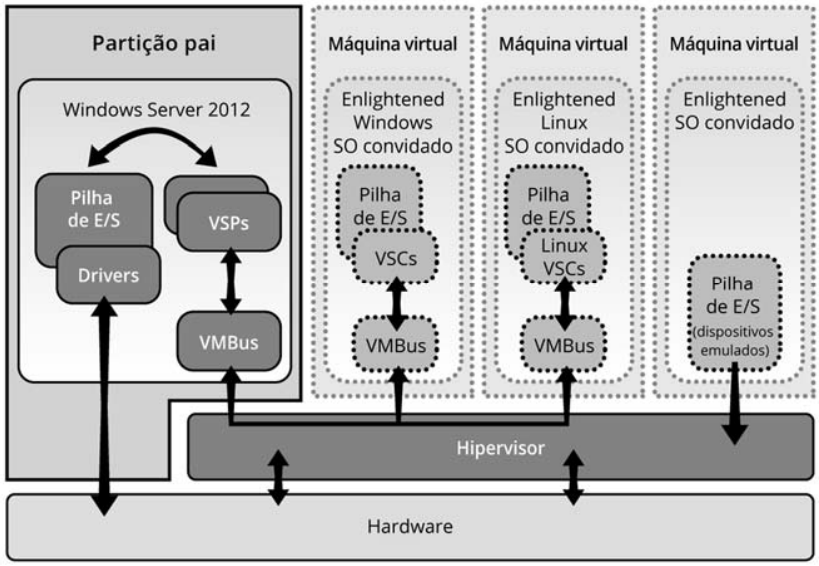


Figura 5.3 Arquitetura do hipervisor.

Cenários de uso do Hyper-V

- Consolidação de servidores
- Continuidade de negócios e recuperação de desastres
- Desenvolvimento e teste
- Datacenter dinâmico

Consolidação de servidores

As empresas são pressionadas a simplificar o gerenciamento e reduzir os custos, ao mesmo tempo em que precisam manter e aprimorar suas vantagens competitivas – como flexibilidade, confiabilidade, escalabilidade e segurança. O uso do Hyper-V para a consolidação de vários servidores em um único sistema físico, mantendo o isolamento, ajuda a suprir tais necessidades. Um dos principais benefícios da consolidação de servidores é a redução do custo total de propriedade (TCO), não apenas por diminuir os requisitos de hardware, mas também os custos com energia, refrigeração e gerenciamento.

As empresas podem se beneficiar ainda da virtualização com a otimização da infraestrutura, tanto do ponto de vista da utilização dos ativos quanto da capacidade de balancear as cargas de trabalho por diferentes recursos. Outro benefício é a capacidade de integrar livremente cargas de trabalho de 32 bits e 64 bits no mesmo ambiente.

Continuidade de negócios e recuperação de desastres

A continuidade de negócios é a capacidade de minimizar o tempo de inatividade programado e também o não programado. Isso inclui o tempo gasto em funções de rotina, como manutenção, backup e interrupções de energia não previstas. O Hyper-V contém recursos

para prover a continuidade de negócios, como backup em tempo real e migração rápida, permitindo que as empresas atendam aos parâmetros de tempo de ativação e resposta.

A recuperação de desastres é o ponto central da continuidade de negócios. Desastres naturais, ataques mal-intencionados e mesmo os problemas simples de configuração, como conflitos de software, podem degradar os aplicativos e os serviços – até que os administradores resolvam os problemas e restaurem os dados de backup. Com o Windows Server 2008, o Hyper-V oferece suporte à recuperação de desastres (Disaster Recovery – DR) dentro dos ambientes dos datacenters, utilizando as capacidades de agrupamento geograficamente dispersas. A recuperação rápida e confiável de desastres ajuda a garantir o funcionamento das capacidades de gerenciamento remoto e a perda mínima de dados.

Desenvolvimento e teste

Desenvolvimento e teste geralmente são as tarefas que podem usufruir muito da tecnologia Hyper-V. Com o uso de máquinas virtuais, as equipes de desenvolvimento são capazes de criar e testar uma ampla variedade de cenários, em um ambiente seguro e independente, que aproxima perfeitamente a operação dos servidores físicos e clientes. O Hyper-V maximiza a utilização dos hardwares de teste, com redução de custos, melhora do gerenciamento do ciclo de vida e aprimoramento da cobertura dos testes. Com recursos abrangentes de suporte aos sistemas operacionais convidados e pontos de verificação, o Hyper-V é uma ótima plataforma para ambientes de desenvolvimento e testes.

Datacenter dinâmico

O Hyper-V, em conjunto com as soluções de gerenciamento de sistema já existentes, como o Microsoft System Center, pode ajudar a criar um datacenter dinâmico, com sistemas de autogerenciamento e agilidade operacional. Por meio de mecanismos como a reconfiguração automatizada de máquinas virtuais, o controle flexível de recursos e a migração rápida, é possível criar um ambiente de TI dinâmico, que utiliza o Hyper-V não somente para resolver problemas, mas também para antecipar o aumento de demanda.

Gerenciamento com Hyper-V

- Hyper-V Manager
- System Center Virtual Machine Manager (SCVMM)

O Hyper-V pode ser gerenciado por duas principais ferramentas: a própria do Hyper-V, chamada de Hyper-V Manager, e uma externa denominada Systems Center Virtual Machine Manager (SCVMM).

Hyper-V Manager

- Interface MMC
- Funcionalidades

O Hyper-V utiliza para gerenciamento a interface padrão Microsoft Management Console (MMC). As máquinas virtuais e os servidores são configurados com o MMC, que permite aplicar plugins de terceiros e melhorias por meio de comandos criados pelos usuários do tipo Windows PowerShell. Uma única instância do MMC pode conectar múltiplos hospedeiros Hyper-V. Entretanto, cada hospedeiro – e as máquinas virtuais que possui – é gerenciado de modo independente. Para administrar o ambiente de Hyper-V em cluster, é necessário usar o console Failover Cluster.

Funcionalidade do MMC:

- Criar e gerenciar máquinas virtuais, redes virtuais e discos virtuais;
- Exportar/importar máquinas virtuais;
- Fazer snapshots de máquinas virtuais.

System Center Virtual Machine Manager (SCVMM)

- Conceito
- Propriedades
- Cenários
- Recursos

O SCVMM é uma solução de gerenciamento heterogênea e abrangente para datacenters virtualizados. Facilita a administração de máquinas virtuais Windows e também de ambientes físicos. Oferece melhor utilização do servidor físico e permite:

- A consolidação simples e rápida de infraestrutura virtual, com a identificação integrada de candidato à consolidação;
- P2V rápida, gerenciamento centralizado da infraestrutura de máquinas virtuais, rápido fornecimento de novos equipamentos virtuais pelo administrador e usuários finais;
- Distribuição inteligente da carga de trabalho, com base no conhecimento de desempenho, e diretivas comerciais definidas pelo usuário.

Cenários

- Consolidação de servidor de produção
- Aumento da agilidade operacional
- Gerenciamento integrado

Consolidação de servidor de produção

À medida que as organizações buscam consolidar seus servidores de produção, o SCVMM oferece uma maneira de transferir o conhecimento sobre o sistema e o ambiente pelo processo de virtualização, que ajuda a manter a continuidade do conhecimento. Pela consolidação de vários servidores de produção com o Virtual Server 2005 R2 ou a virtualização do Windows Server, as empresas reduzem o custo total de propriedade geral e ainda mantêm um framework unificado de gerenciamento em seus ambientes físico e virtual.

Aumento da agilidade operacional

Companhias de todos os segmentos procuram maneiras de aumentar a eficiência por meio de seus ambientes de TI e, com isso, aumentar a agilidade operacional. O SCVMM oferece um mecanismo para permitir funcionalidades como os rápidos aprovisionamento de servidor e recuperação, além da capacidade de migração escalável, para tornar toda a infraestrutura virtual robusta e fácil de gerenciar.

Gerenciamento integrado

O SCVMM ajuda a criar uma infraestrutura de gerenciamento centralizado de máquina virtual em múltiplas máquinas que empregam o Virtual Server 2005 R2 e a virtualização do Windows Server. As organizações, em geral, estão adotando a virtualização nas áreas de produção, teste e desenvolvimento. A sofisticação dos recursos de gerenciamento ajuda os administradores a implantar e gerenciar ambientes virtuais e físicos, com uma abordagem integrada.

Recursos do SCVMM

- Identificação de candidatos à consolidação
- Disposição inteligente
- Aprovisionamento de servidores
- Biblioteca central
- Aprovisionamento de autoatendimento

O SCVMM se concentra em requisitos únicos de máquinas virtuais; projetado para permitir a maior utilização possível dos servidores físicos, gerenciamento centralizado de infraestrutura de equipamento virtual e rápido aprovisionamento de novas máquinas.

Recursos disponíveis no SCVMM:

Identificação de candidato à consolidação

O primeiro passo na migração de um centro de dados físico com um modelo de carga de trabalho por servidor é identificar as cargas de trabalho físicas apropriadas para a consolidação no hardware virtual. Os fatores de decisão para determinar os candidatos adequados se baseiam em vários pontos, como desempenho histórico, características de pico de carga e padrões de acesso. O SCVMM utiliza os históricos de desempenho existentes no banco de dados do System Center Operations Manager para listar os candidatos à consolidação em ordem de classificação.

Disposição inteligente

É a designação e ativação da carga de trabalho virtual de um servidor físico, maximizando o emprego de ativos físicos. O System Center Virtual Machine Manager faz uma abordagem profunda e holística da disposição, combinando o conhecimento de dados históricos de desempenho da carga de trabalho com informações sobre as máquinas virtuais. Regras

comerciais e modelos associados também são utilizados pelo SCVMM para determinar as opções de disposição.

Aprovisionamento de servidores

O SCVMM identifica as máquinas virtuais na empresa pela integração com o Active Directory. Isso ajuda as organizações a escalar facilmente o gerenciamento de equipamentos físicos e virtuais no datacenter e escritórios remotos.

Biblioteca central

O SCVMM oferece um repositório central a todos os blocos de construção para uma máquina virtual, como Virtual Hard Disk (VHDs), máquinas virtuais off-line, modelos e até mesmo imagens ISO. Cada item da biblioteca possui modelos ou metadados que permitem um gerenciamento mais controlado dos objetos. O modelo é um novo objeto que possibilita ao administrador criar configurações aprovadas de equipamento virtual, que servem como um padrão para subsequentes implantações de máquinas virtuais.

Aprovisionamento de autoatendimento

A infraestrutura virtual é comumente usada em ambientes de teste e desenvolvimento em que há aprovisionamento coerente e desmontagem de máquinas virtuais para a finalidade de teste. Com o SCVMM, os administradores podem estender seletivamente os recursos de autoatendimento a grupos de usuários e também definir cotas. A ferramenta de aprovisionamento automático gerencia as máquinas virtuais por seus ciclos de vida, incluindo as desmontagens.

Snapshots com o Hyper-V

- Instantâneo da máquina virtual
- Backup do estado da máquina virtual
- Testes e desenvolvimento
- Reduz o desempenho da máquina virtual em produção

Os snapshots de uma máquina virtual são arquivos baseados no estado, nos dados do disco e na configuração do equipamento virtual num determinado ponto no tempo. É possível tirar vários snapshots de uma máquina virtual, mesmo quando ela está ativa, e, em seguida, revertê-la a qualquer um dos estados anteriores, aplicando o snapshot apropriado. Há possibilidade de usar tanto o Hyper-V Manager como o Virtual Machine Connection para fazer um snapshot. Todas as tarefas passíveis de serem realizadas com snapshots estão disponíveis pelo Hyper-V Manager, como aplicar ou excluir snapshots e exibir a lista de todos os snapshots existentes para uma máquina virtual específica.

Os snapshots podem ajudar a aumentar a eficiência em muitas configurações, nas quais pode ser preciso recriar diferentes ambientes e reproduzir condições como as de desenvolvimento e teste, além de tarefas relacionadas a suporte técnico. Entretanto, o mesmo poder e flexibilidade que torna os snapshots úteis e eficazes em determinadas configurações podem causar consequências indesejadas e potencialmente graves em outras configurações. Isso

inclui os riscos inerentes à perda involuntária de dados, se os snapshots não forem gerenciados apropriadamente.

Uma das situações mais favoráveis para o uso dos snapshots refere-se ao desenvolvimento e teste de atividades, incluindo a utilização da máquina virtual como um servidor de teste para as atualizações e hotfixes, antes de implantá-los nos servidores de produção. Não é recomendável empregar snapshots em equipamentos virtuais que prestam serviços sensíveis no tempo, tais como serviços do Active Directory (AD), ou quando o desempenho ou a disponibilidade de espaço de armazenamento é crítica.

Além disso, antes de utilizar os snapshots é preciso considerar:

- Fazer um snapshot reduz o desempenho da máquina virtual, enquanto ele é criado. Não se deve usar snapshot em equipamentos virtuais que prestam serviços em um ambiente de produção;
- Não é recomendável o uso de snapshots em máquinas virtuais configuradas com discos rígidos virtuais fixos, porque reduzem os benefícios de desempenho que podem ser obtidos com a utilização de discos rígidos virtuais fixos;
- Os snapshots exigem espaço de armazenamento adequado, como arquivos AVHD no mesmo local no disco rígido virtual. Vários snapshots podem rapidamente consumir uma grande quantidade de espaço. Quando o Hyper-V Manager é usado para excluir um snapshot, ele é removido da árvore de snapshots, mas o arquivo AVHD não é excluído até que a máquina virtual seja desligada.

Live Backup

- Conceito
- Métodos básicos

É possível efetuar backup automático das máquinas virtuais e seus dados sem qualquer downtime, caso o Sistema Operacional convidado tenha suporte ao Volume Shadow Copy Service (VSS). Para implementar os cenários de backup e recuperação, é preciso usar um aplicativo de backup compatível com o gravador VSS do Hyper-V. O backup do Windows Server exige a adição de uma chave de registro ao gravador VSS do Hyper-V.

Há dois métodos básicos para realizar um backup:

Servidor que executa o Hyper-V

Recomendado para realizar um backup completo do servidor. Se o aplicativo de backup for compatível com o Hyper-V e o gravador VSS do Hyper-V, é possível salvar todos os dados necessários para restaurar totalmente o servidor, exceto as redes virtuais. Os dados acrescidos no backup incluem a configuração de máquinas virtuais, instantâneos associados a elas e discos rígidos virtuais. Como resultado, a recuperação do servidor pode ser mais fácil, se isso for preciso, pois não existe a necessidade de criar novamente os equipamentos virtuais ou reinstalar o Hyper-V. Entretanto, as redes virtuais não estão incluídas em um backup completo de servidor. Será preciso reconfigurar a rede, criando novamente as redes

virtuais e reconectando os adaptadores virtuais de cada máquina à rede virtual apropriada. Parte do planejamento de backup, há necessidade de documentar a configuração e todas as definições relevantes de rede virtual, caso seja preciso criá-la novamente.

Sistema Operacional convidado de uma máquina virtual

Método usado quando for preciso fazer o backup dos dados de um repositório que não possui suporte do gravador VSS do Hyper-V. Nesse caso, o aplicativo de backup é executado no Sistema Operacional convidado da máquina virtual.

O processo de backup em fita aproveita a vantagem da funcionalidade no backup virtual do Hyper-V. Por exemplo, se um servidor incorpora um script que automatiza o backup dos dados para a fita, esse processo continua sendo utilizado se o servidor é convertido para máquina virtual.

Avanços com o Windows Server 2012 R2 Hyper-V

- Armazenamento dinâmico de máquina virtual
- Live Migration
- Suporte de Processador Avançado
- Suporte avançado de redes
- Volumes compartilhados de cluster

Armazenamento dinâmico de máquina virtual

O Windows Server 2012 R2 Hyper-V suporta a inserção e a remoção “a quente” de armazenamento. Com a adição ou remoção de arquivos Virtual Hard Drive (VHD) e de discos passthrough enquanto a máquina virtual está sendo executada, o Windows Server 2012 R2 Hyper-V possibilita a rápida configuração dos equipamentos virtuais para atender aos requisitos de mudança. O recurso permite tanto a adição como a remoção de arquivos VHD ou de discos passthrough aos controladores SCSI existentes nas máquinas virtuais.

A inserção ou remoção “a quente” de armazenamento requer os serviços de integração do Hyper-V, fornecidos pelo Windows Server 2012 R2 para instalação no Sistema Operacional convidado.

Live Migration

Live Migration é um dos mais novos recursos do Windows Server 2012 R2 Hyper-V. Ainda neste capítulo, descreveremos o recurso em detalhes. Daremos informações sobre como ele move as máquinas virtuais em execução e descreve cenários nos quais é particularmente útil, além dos seus requisitos de implementação.

Suporte de processador avançado

O Windows Server 2012 R2 Hyper-V suporta até 32 núcleos de processadores lógicos. O suporte avançado aos processadores torna possível executar cargas de trabalho ainda mais

exigentes em um único computador físico ou consolidar mais cargas de trabalho em um único computador físico. O Windows Server 2012 R2 Hyper-V também aguenta os recursos Second-Level Address Translation (SLAT) e o core parking de CPU.

A SLAT usa a funcionalidade de CPU disponível nos processadores Intel – que suportam as tabelas de página estendida – e nos processadores AMD, que possuem a Indexação Rápida de Virtualização (Rapid Virtualization Indexing) para reduzir a carga de tradução de endereços físicos convidados para endereços físicos reais. Isso diminui significativamente o tempo de CPU do hipervisor e economiza memória para cada máquina virtual, permitindo ao computador físico trabalhar mais, com menos recursos do sistema. O core parking de CPU possibilita a economia de energia; agenda a execução do equipamento virtual em apenas alguns dos núcleos de CPU do servidor e coloca o resto em estado "sleep".

Suporte avançado de redes

No Windows Server 2012 R2 há três novos recursos de rede que melhoram o desempenho da rede no ambiente de virtualização. O suporte a jumbo frames, antes disponível em ambientes não virtuais, tem sido ampliado para ser aplicado nas máquinas virtuais. Ele permite que os equipamentos usem jumbo frames de até 9014 bytes, se a rede física subjacente aguentar. O suporte a jumbo frames reduz a carga da pilha de rede e aumenta o processamento. Além disso, há também uma significante diminuição de utilização da CPU, devido ao menor número de chamadas da pilha de rede para o driver de rede.

O TCP Chimney, que transfere a subcarga de processamento TCP/IP para o hardware de rede, também tem sido estendido para o trabalho no mundo virtual. O TCP Chimney melhora o desempenho da máquina virtual, permitindo a redução do processamento de rede para o hardware, especialmente em redes de mais de 1 Gigabit/s. Esse recurso é especialmente benéfico para as funções que envolvem grande quantidade de transferência de dados, tal como a função de servidor de arquivo. O Consulta de Máquina Virtual (Virtual Machine Queue – VMQ) possibilita que os adaptadores de rede dos computadores físicos usem DMA para colocar os conteúdos dos pacotes diretamente na memória do equipamento virtual, aumentando o desempenho de E/S.

Volumes compartilhados de cluster

Com o Windows Server 2012 R2, o Hyper-V é capaz de usar o armazenamento em Cluster Shared Volume (CSV) para simplificar e aumentar o armazenamento compartilhado. O CSV permite que múltiplos servidores Windows acessem o armazenamento SAN com um único espaço de nome consistente para todos os volumes, em todos os servidores. Múltiplos servidores podem acessar o mesmo Número de Unidade Lógica (LUN) no armazenamento de SAN.

O CSV possibilita migrações instantâneas mais rápidas e facilita o gerenciamento de armazenamento para o Hyper-V, quando usado em uma configuração de cluster. Os volumes compartilhados de cluster estão disponíveis como parte do recurso de Clustering de Failover do Windows Server 2012 R2.

Alta disponibilidade

- Guest clustering
- Host clustering
- Quick Migration
- Live Migration
- Live Migration vs Quick Migration

O Windows Server 2008 com Hyper-V oferece duas soluções para Alta Disponibilidade (HA): Hyper-V Guest Clustering e Hyper-V Host Clustering. No primeiro, o serviço de clusters é executado no Sistema Operacional convidado, dentro da máquina virtual. Nessa configuração, as aplicações é que estão com alta disponibilidade. A figura 7.5 ilustra o cluster em uso por duas máquinas virtuais hospedadas em um mesmo servidor físico. Se o servidor falhar ou houver problema na partição pai do Windows, os dois equipamentos virtuais falharão juntos.

Guest Clustering

A segunda solução de uso do Hyper-V Guest Clustering, com mais de um servidor, é apresentada na figura 7.6. O cluster é realizado entre duas máquinas virtuais instanciadas em servidores físicos distintos. Nesse caso, se um servidor físico (ou mesmo a partição pai) falhar, a aplicação em cluster continuará disponível. O detalhe desse cluster é que a conexão com o sistema de armazenamento (storage) tem que ser feita via iSCSI – os iSCSI initiators executam dentro do Sistema Operacional convidado –, pois só assim a máquina virtual pode se comunicar com o sistema de armazenamento sem precisar da partição pai.

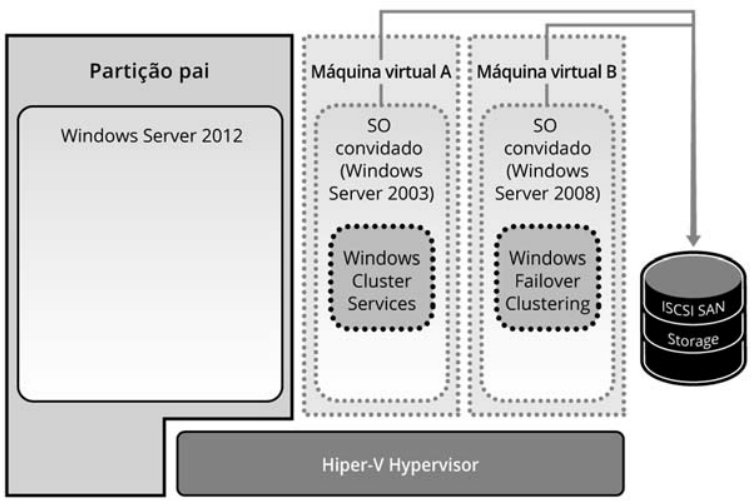


Figura 5.4 Guest Clustering (mesmo servidor físico).

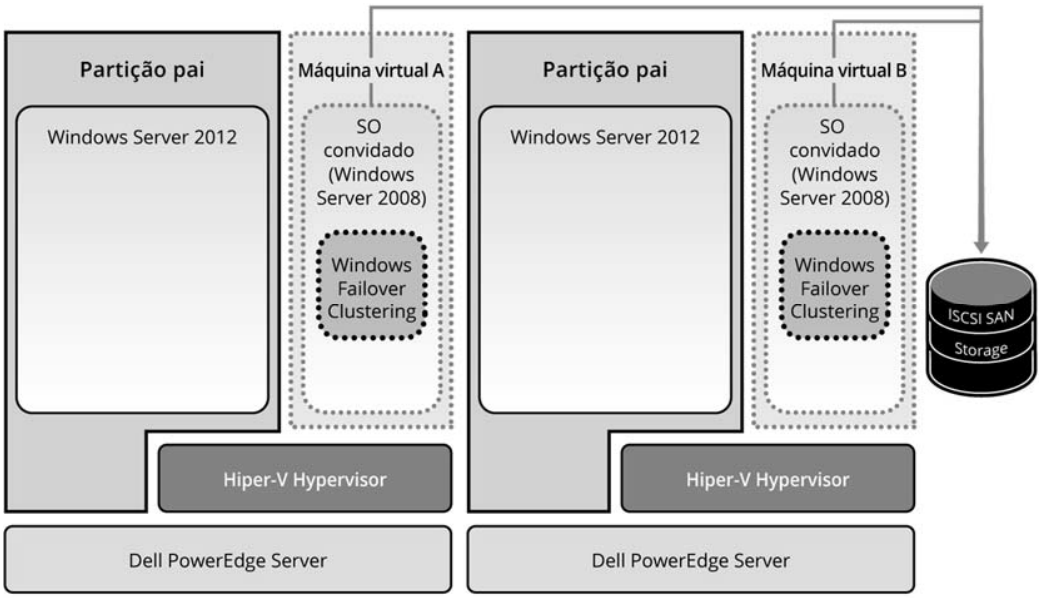


Figura 5.5 Guest Clustering (diferentes servidores físicos).

Principais características do Hyper-V Guest Clustering:

- Aplicações que executam nos sistemas operacionais convidados são as entidades do cluster;
- A alta disponibilidade é da aplicação e não da máquina virtual;
- O Sistema Operacional convidado (guest) precisa ser suportado pelo Hyper-V;
- Requer que o Sistema Operacional convidado tenha capacidade de cluster;
- Requer que o software de cluster tenha suporte a iSCSI;
- Cluster tem de ser configurado em toda máquina virtual;
- As cargas de trabalho (aplicações) devem ser cluster-aware.

Host Clustering

A segunda possibilidade para a alta disponibilidade é o Hyper-V Host Clustering. Nesse caso, o serviço de cluster é o próprio do Windows Server 2008 64 bits, o failover clustering, executado na partição pai. Pode ser implementado em sistemas de armazenamento (storage) do tipo DAS, iSCSI e FC SAN. As máquinas virtuais são como recursos de cluster e não precisam de serviços de cluster nos sistemas operacionais convidados. O requisito é que os arquivos de equipamentos virtuais no storage sejam acessados por todos os servidores físicos que fazem parte do cluster. Esse é o tipo de cluster mais recomendado, por permitir uma manutenção planejada e a melhor distribuição das cargas das máquinas virtuais. Um cluster Hyper-V Host pode ter até 16 servidores físicos e ser implantado nas versões de 64 bits Enterprise e datacenter. A figura 7.7 ilustra o uso do Hyper-V Host Clustering.

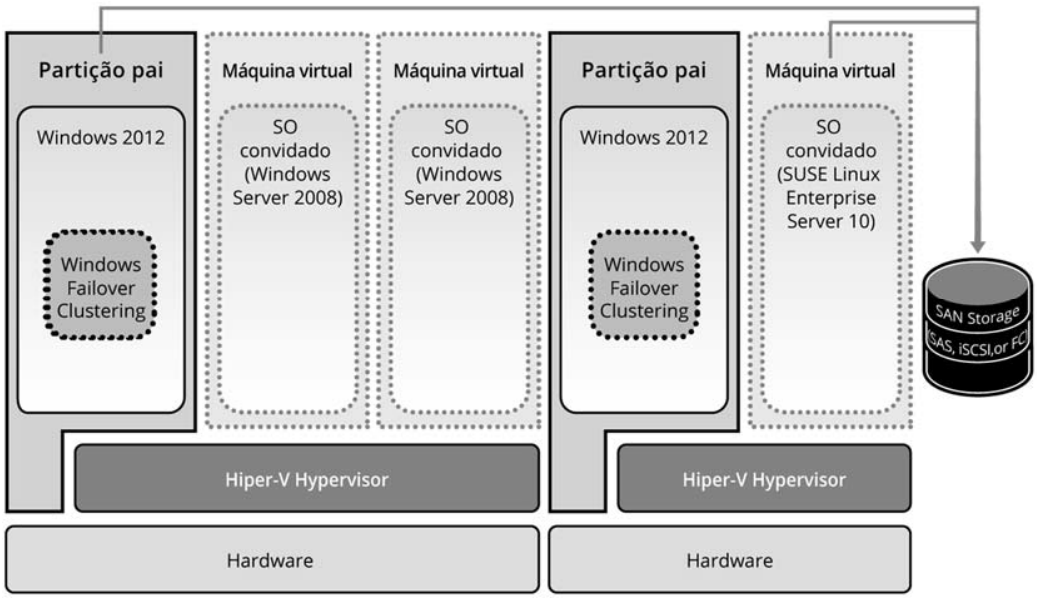


Figura 5.6 Hyper-V Host Clustering.

Principais características do Hyper-V Host Clustering:

- As máquinas virtuais são as entidades do cluster;
- A alta disponibilidade é no nível da máquina virtual;
- Independe do Sistema Operacional convidado, pois o software de cluster executa na partição pai;
- O cluster é configurado só na partição pai de cada servidor;
- A implementação é independente do número de equipamentos virtuais;
- As cargas de trabalho (aplicações) dentro das máquinas virtuais não precisam ser cluster-aware.

O Hyper-V Host clustering permite tanto o failover planejado, conhecido como Quick Migration/Live Migration, quanto o failover não planejado. O Hyper-V possibilita a Quick Migration (migração rápida) de uma máquina virtual em execução, de um sistema físico hospedeiro para outro, com um tempo de inatividade mínimo. Assim, aproveita as capacidades já conhecidas de alta disponibilidade das ferramentas de gerenciamento do Windows Server e do SCVMM.

Quick Migration

O Quick Migration combina a virtualização baseada em hipervisor do Windows com o cluster failover disponibilizado nas versões Windows Server 2012 Enterprise e Windows Server 2012 datacenter. O Quick Migration não apresenta **Recovery Point Objective** (RPO) zero, pois tanto o Sistema Operacional convidado quanto a carga de trabalho da máquina ficam indisponíveis por algum tempo, enquanto são executados os procedimentos de cluster. Esse intervalo costuma variar entre 15 segundos e alguns minutos. Devido às características de

cluster, e pela forma que interage com o processador, é recomendado que os processadores dos servidores envolvidos sejam idênticos. No caso de failover não planejado, o serviço de cluster reinicializa automaticamente todas as máquinas virtuais em outro nó.

Live Migration

O Live Migration é integrado com o Windows Server 2012 R2 Hyper-V e com o Microsoft Hyper-V Server 2008 R2. Com o Live Migration, pode-se mover as máquinas virtuais em execução de um servidor físico Hyper-V para outro, sem qualquer interrupção de serviço ou tempo de indisponibilidade perceptível. Dessa forma, o Live Migration oferece uma boa agilidade, pois os datacenters com múltiplos servidores físicos Hyper-V serão capazes de mover os equipamentos virtuais (em execução) para um computador físico melhor, a fim de obter maior desempenho e escalabilidade ou otimizar a consolidação sem impacto para os usuários. Ainda, o Live Migration torna viável manter as máquinas virtuais on-line, mesmo durante a manutenção, aumentando a produtividade dos usuários e dos administradores de servidor. Os datacenters também serão capazes de reduzir o consumo de energia, elevando dinamicamente as taxas de consolidação e desligando os servidores físicos não utilizados durante os horários de menor demanda.

Recovery Point Objective

Processo que estima o quanto será necessário retroceder no tempo para obter uma cópia consistente para a reinicialização correta.

Etapas básicas realizadas pelo Livre Migration:

1. Todas as páginas de memória da máquina virtual são transferidas do computador físico Hyper-V fonte para o computador Hyper-V de destino. Enquanto isso, é feito o rastreamento de qualquer modificação do equipamento virtual para suas páginas de memória.
2. As páginas que foram modificadas durante a etapa anterior são transferidas para o computador físico de destino.
3. O controle de armazenamento dos arquivos VHD das máquinas virtuais são movidos para o computador físico de destino.
4. A máquina virtual de destino é trazida para o modo on-line no servidor Hyper-V de destino.

Dessa forma, o Live Migration reduz significativamente o tempo de indisponibilidade da máquina virtual que está sendo migrada. Isso o torna o tipo de migração indicada quando os usuários devem ter acesso ininterrupto ao equipamento. Como o Live Migration é projetado para finalizar suas atividades em um tempo menor que o do timeout das conexões TCP, os usuários não terão nenhuma quebra de conexão na migração de máquina virtual durante as etapas 3 e 4 do processo.

Gerenciamento do Live Migration

O SCVMM, usado em conjunto com o Live Migration, pode aumentar a capacidade de uma organização de responder às mudanças de níveis de uso e de requisitos de sua infraestrutura de TI. O SCVMM também é útil no gerenciamento de computadores físicos Hyper-V locais e remotos. Quando o SCVMM gerencia um computador Hyper-V, que foi configurado para alta

disponibilidade, ele é capaz de iniciar o Quick Migration (ou o Live Migration) a partir de seu console de gerenciamento. Isso fornece uma ferramenta única para todas as tarefas de gerenciamento das máquinas virtuais, incluindo o Live Migration.

O Windows PowerShell é a base de todo o acesso ao servidor SCVMM. Como o SCVMM Admin Console baseia-se inteiramente no Windows PowerShell, é possível criar scripts Windows PowerShell para serem executados pelo SCVMM. Ao empregar o SCVMM para migrar uma máquina virtual, pode-se gerar o script PowerShell associado. Assim, há possibilidade de repetir essa ação no futuro via script ou alterá-lo para realizar a migração de uma máquina virtual diferente, e, ainda, para alterar os computadores físicos de origem e de destino da migração.

O SCVMM oferece relatório sobre virtualização de servidor físico e colocação de máquina virtual. Esses relatórios podem ser usados em processos de tomada de decisão para a colocação de novas máquinas virtuais ou para a migração das existentes. Especialmente em ambientes muito densos, como vários datacenters, ou muitos dispersos, como locais remotos, a informação correta sobre o desempenho da virtualização pode ser vital para atender aos requisitos de uptime e disponibilidade. O SCVMM facilmente fornece as informações necessárias para o gerenciamento eficiente de múltiplos computadores físicos Hyper-V ou de máquinas virtuais. Como o Live Migration do Hyper-V facilita a mobilidade dos equipamentos virtuais de um computador físico para outro, a obtenção de informações corretas sobre as máquinas Hyper-V no ambiente é especialmente importante.

Live Migration *vs* Quick Migration

O Quick Migration é um recurso do Windows Server 2012 Hyper-V e do Windows Server 2012 R2 Hyper-V. Tanto o Live Migration quanto o Quick Migration movem as máquinas virtuais em execução de um computador físico Hyper-V para outro. A diferença é que o Quick Migration primeiro salva e depois move a máquina virtual; só então a restaura, resultando em um tempo de indisponibilidade. O processo do Live Migration usa um mecanismo diferente para mover o equipamento virtual em execução para um novo computador físico.

O Windows Server 2012 Hyper-V suporta o Quick Migration. O Windows Server 2012 R2 Hyper-V suporta tanto o Quick Migration quanto o Live Migration. O Live Migration foi disponibilizado pela versão Windows 2008 Server R2 e é integrado ao Windows Server 2012 R2 Hyper-V e ao Microsoft Hyper-V Server 2012 R2.

6

Gerenciamento do Hyper-V

Infraestrutura de hardware

- Processador
- Memória
- Armazenamento
- Rede
- Outros componentes

O Hyper-V está disponível em edições de 64 bits do Windows Server 2012, mais especificamente, nas edições de 64 bits do Windows Server 2012 Standard e Windows Server 2012 Datacenter. Não há no mercado Hyper-V para edições de 32 bits (x86), ou Windows Server 2012 para sistemas baseados em Itanium. Já as ferramentas de gerenciamento do Hyper-V podem ser encontradas para edições de 32 bits.

Para planejar e implantar o Hyper-V com eficiência, é preciso entender os requisitos e as configurações máximas do hardware físico e do ambiente virtual que compreenderá o local de servidor virtualizado. O objetivo dessa seção é fornecer elementos para um melhor dimensionamento de uma solução de virtualização com o emprego do Hyper-V.

Processador

O Hyper-V precisa de processadores com características específicas, como arquitetura de 64 bits, suporte de hardware à virtualização (Hardware Assisted Virtualization – HAV) e o modo Hardware Enforced Data Execution Prevention (DEP) – aplicado por hardware disponível e habilitado. Os processadores Intel-VTx e AMD-V contemplam os três requisitos.

É possível buscar os servidores que oferecem suporte à arquitetura x64 e ao Hyper-V no catálogo do Windows Server. As configurações de virtualização assistida por hardware (HAV) e DEP aplicado por hardware estão disponíveis na BIOS. Entretanto, os nomes das configurações podem diferir dos que constam no catálogo. Para saber se um determinado modelo dispõe de suporte ao Hyper-V, deve-se verificar com o fabricante do servidor.

A versão de lançamento do Hyper-V conta com suporte em computadores físicos com até 320 processadores lógicos; enquanto as versões anteriores, com apenas 24. É possível configurar o total de 64 processadores virtuais em uma máquina virtual. Observe que o número de processadores com suporte a um Sistema Operacional convidado pode ser menor.

Memória

Quando trata-se de infraestruturas de virtualização, talvez não haja recurso de hardware tão importante para o desempenho geral como a memória física. É essencial alocá-la de maneira que garanta a cada máquina virtual (VM) ter o espaço necessário para manter a infraestrutura funcional. Atualmente, o Windows Server 2012 suporta até 4 TB de memória RAM, e as VMs podem ser configuradas com até 1 TB de memória.

Armazenamento

O Hyper-V oferece suporte a uma variedade de opções de armazenamento. Os seguintes tipos de armazenamento físico podem ser utilizados com um servidor que executa o Hyper-V:

- **Disk Array Subsystem (DAS)**: Serial Advanced Technology Attachment (SATA), Parallel Advanced Technology Attachment (PATA), Serial Attached SCSI (SAS), SCSI, USB e Firewire.
- **Storage Area Network (SAN)**: tecnologias Internet SCSI (iSCSI), Fibre Channel e SAS. A Microsoft não oferece suporte a NAS (armazenamento conectado à rede) para o Hyper-V.

É possível configurar uma máquina virtual para usar os seguintes tipos de armazenamento:

- **Dispositivos virtuais IDE**: cada máquina virtual fornece suporte a até 4 dispositivos IDE. O disco de inicialização, também denominado disco de boot, deve estar conectado a um dos dispositivos IDE. O disco de inicialização pode ser rígido virtual ou físico. Embora um equipamento virtual deva empregar um dispositivo IDE virtual, como o disco de inicialização do Sistema Operacional convidado, há várias opções ao selecionar o dispositivo físico que fornecerá o armazenamento para o IDE. Podemos citar qualquer um dos tipos mencionados de armazenamento físico.
- **Dispositivos virtuais SCSI**: cada máquina virtual oferece suporte a até 4 controladores SCSI virtuais, e cada controlador, a até 64 discos. Isso significa que cada equipamento pode ser configurado com até 256 discos virtuais SCSI. O uso de dispositivos SCSI virtuais requer que os serviços de integração estejam instalados no Sistema Operacional convidado. Para obter uma lista de SOs, para os quais os serviços de integração estão disponíveis, consulte os itens máquinas virtuais e sistemas operacionais convidados.
- **Discos rígidos virtuais de até 2 TB**: há possibilidade de empregar discos rígidos virtuais fixos, o que expande dinamicamente os discos virtuais e os diferencia.
- **Discos físicos**: os discos físicos conectados diretamente a uma máquina virtual não têm limitação de tamanho, a não ser a do Sistema Operacional convidado.
- **Capacidade de armazenamento da máquina virtual**: com discos rígidos virtuais, cada máquina fornece suporte a até 512 TB de armazenamento. Para discos físicos, esse número é ainda maior, dependendo do que suporta o SO convidado.

Embora o desempenho de E/S dos dispositivos SCSI e IDE físicos possa diferir bastante, o mesmo não ocorre com os dispositivos SCSI e IDE virtualizados no Hyper-V. Eles oferecem

alto desempenho de E/S, quando os serviços de integração são instalados no Sistema Operacional convidado.

Rede

O Hyper-V concede uma variedade de configurações e opções para atender a diferentes requisitos de rede. A rede do Hyper-V contém o seguinte suporte:

- Cada máquina virtual pode ser configurada com até 12 adaptadores virtuais de rede: 8 podem ser do tipo “adaptador de rede” e 4, do tipo “adaptador de rede herdado”. O primeiro oferece melhor desempenho e requer um driver para máquina virtual incluso nos pacotes de serviço de integração.
- Cada adaptador de rede virtual tem a alternativa de ser configurado com um endereço MAC estático ou dinâmico.
- Cada adaptador de rede virtual oferece suporte a até 4094 VLANs.
- Pode-se ter um número ilimitado de redes virtuais, com até 512 máquinas virtuais por rede.

Não é possível conectar uma rede virtual a um adaptador de rede sem fio. Como resultado, não é possível fornecer recursos de rede sem fio a máquinas virtuais.

Outros componentes

O Hyper-V oferece suporte para a criação de dispositivos de E/S virtuais comuns às máquinas físicas, como:

- **Unidade de DVD**: uma máquina virtual tem uma unidade de DVD virtual por padrão quando é criada. Elas podem ser configuradas com até 3 unidades de DVD conectadas a um controlador IDE. Os equipamentos fornecem suporte a até 4 dispositivos IDE, mas um dispositivo deve ser o disco de inicialização. Uma unidade de DVD virtual pode acessar CDs e DVDs, arquivos .ISO ou mídia física. Entretanto, apenas uma máquina virtual tem possibilidade de ser configurada para acessar uma unidade de CD/DVD física por vez.
- **Porta serial (COM)**: cada máquina virtual é configurada com duas portas (COM) seriais virtuais que podem ser conectadas a um canal (pipe) nomeado para comunicação com um computador físico local ou remoto. Nenhum acesso a uma porta COM física é disponibilizado por um equipamento virtual.
- **Unidade de disquete**: cada máquina virtual está configurada com uma unidade de disquete virtual, que pode acessar os arquivos contidos no Virtual Floppy Disk (VFD). Nenhum acesso a uma unidade de disquete física é disponibilizado para equipamentos virtuais.

Configurando máquinas virtuais

- Sistemas Operacionais convidados
- Discos suportados

É possível utilizar o Hyper-V para configurar e usar várias máquinas virtuais ao mesmo tempo. O número específico depende de dois fatores. Um é a disponibilidade dos recursos físicos no servidor que executa o Hyper-V. Outro corresponde à capacidade máxima do Hyper-V. Há a opção de configurar até 512 máquinas virtuais em um servidor que executa o Hyper-V. Com os recursos físicos apropriados, a versão de lançamento do Hyper-V oferece suporte a até 128 equipamentos em execução ao mesmo tempo. Um hotfix (KB956710) está disponível para aumentar para 192 o número máximo de máquinas virtuais em operação.

Sistemas operacionais convidados

O Hyper-V oferece suporte para sistemas operacionais convidados da família Microsoft Windows e GNU/Linux, sendo capaz de executar SOs de 32 bits e 64 bits ao mesmo tempo. No caso da família Windows, é possível usar as seguintes edições de 32 e 64 bits do Windows Server 2012, além de versões anteriores – como Windows 2000, 2003, Vista e Seven – como um sistema operacional convidado. O que inclui suporte em uma máquina virtual configurada com 1, 2 ou 4 processadores virtuais:

- Windows Server 2012 Standard e Windows Server 2012 Standard sem Hyper-V;
- Windows Server 2012 Datacenter e Windows Server 2012 Datacenter sem Hyper-V;
- Windows Web Server 2012;
- Windows Server 2012 HPC Edition.

No caso do Sistema Operacional GNU/Linux, um servidor Hyper-V permite a execução de várias distribuições em uma máquina virtual configurada com um processador virtual:

- Suse Linux Enterprise;
- Open SUSE;
- Red Hat Linux;
- Ubuntu;
- CentOS.

Discos suportados

Os tipos de disco utilizados pelo Hyper-V recaem em dois grandes grupos: Virtual Hard Disk (VHD) e pass-through. Um Virtual Hard Disk (VHD) é um arquivo que reside em uma partição formatada a que a partição pai tem acesso. Esse arquivo aparece para o Sistema Operacional convidado como um disco com um conjunto de blocos em estado bruto. A próxima figura ilustra o uso de VHD.

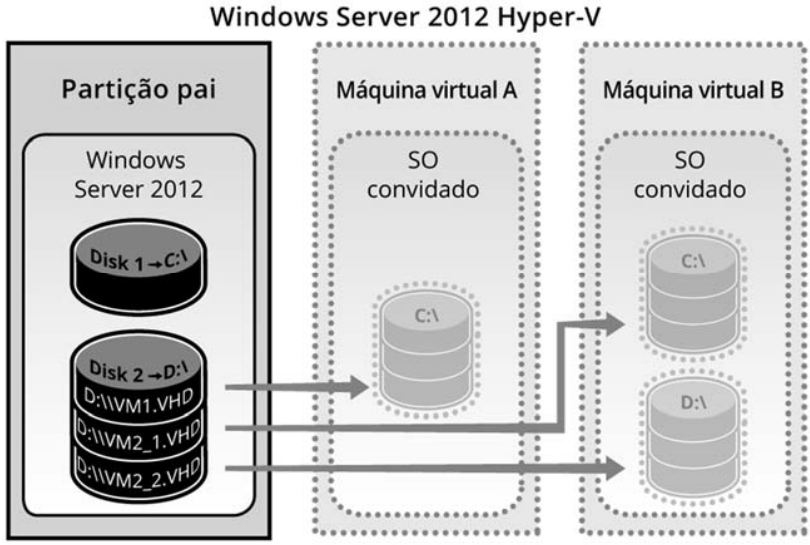


Figura 6.1 Máquina virtual com Virtual Hard Disk (VHD).

O pass-through é um disco físico mapeado diretamente em uma máquina virtual. Para a partição pai, o disco está off-line, e o acesso direto para escrita e leitura, indisponível. Os requisitos de E/S para a máquina virtual são passados por meio da partição pai para o disco. Com a liberação da partição pai, sobram ciclos adicionais de CPU para processar as máquinas virtuais. Quando mapeados para um disco pass-through, o SO convidado tem acesso direto aos blocos brutos do storage físico. A figura a seguir ilustra o uso de pass-through.

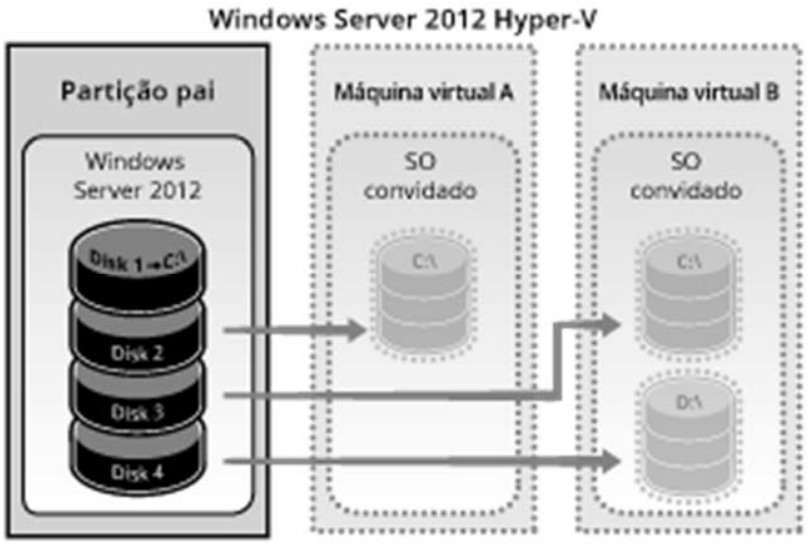


Figura 6.2 Máquina virtual usando disco pass-through.

Discos e armazenamento

- Opções de armazenamento em máquinas virtuais
- Criação de discos rígidos virtuais
- Discos físicos para máquinas virtuais

Opções de armazenamento em máquinas virtuais

Embora o desempenho de E/S dos dispositivos SCSI e IDE físicos possa variar bastante, isso não é válido para os dispositivos SCSI e IDE virtualizados no Hyper-V. Os dispositivos IDE e SCSI do Hyper-V oferecem alto desempenho de E/S quando serviços de integração são instalados no Sistema Operacional convidado. A tabela a seguir descreve as opções de armazenamento disponíveis com dispositivos SCSI.

Cenário	HD Virtual SCSI Local	SCSI Local diretamente conectado	HD Virtual SCSI Remoto	SCSI Remoto diretamente conectado
Tipo de armazenamento	Armazenamento diretamente conectado	Armazenamento diretamente conectado	SAN, Fibre Channel/iSCSI	SAN, Fibre Channel/iSCSI
Tipo de disco exposto ao Sistema Operacional de gerenciamento	Disco rígido virtual em NTFS	Disco físico diretamente conectado a uma máquina virtual	Disco rígido virtual em NTFS	Disco físico diretamente conectado a uma máquina virtual
Tamanho máximo de disco com suporte em máquina virtual	2 TB	Nenhum limite de tamanho a não ser o oferecido pelo Sistema Operacional convidado	2 TB	Nenhum limite de tamanho a não ser o oferecido pelo SO convidado
Instantâneos de disco rígido virtual têm suporte	Sim	Não	Sim	Não
Disco rígido virtual de expansão dinâmica	Sim	Não	Sim	Não

Disco rígido virtual de diferenciação	Sim	Não	Sim	Não
Capacidade de acesso dinâmico (adição automática) a qualquer disco das máquinas virtuais	Não	Não	Não	Não

Tabela 6.3 Armazenamento SCSI.

Criação de discos rígidos virtuais

O Sistema Operacional de gerenciamento permite a seleção do uso de discos rígidos virtuais ou discos físicos diretamente conectados a uma máquina virtual. Os discos rígidos virtuais podem ter uma capacidade de até 2 TB e incluir os seguintes tipos:

- **Fixo**: ocupa o espaço em disco físico no SO de gerenciamento igual ao tamanho máximo do disco, independentemente de uma máquina virtual exigir o espaço em disco. Leva mais tempo para ser criado que outros tipos de disco, porque o tamanho alocado do arquivo .VHD é determinado durante a sua criação. Esse tipo de disco rígido virtual proporciona melhor desempenho em comparação a outros tipos, porque os discos rígidos virtuais fixos são armazenados em um bloco contíguo no Sistema Operacional de gerenciamento.
- **Expansão dinâmica:** é aquele em que o tamanho do arquivo .VHD cresce à medida que os dados são gravados no disco. Esse tipo permite o uso mais eficiente do espaço em disco. É preciso monitorar o espaço em disco disponível para evitar ficar sem espaço no SO de gerenciamento.
- **Diferenciação**: armazena as diferenças do disco rígido virtual no Sistema Operacional de gerenciamento. Isso possibilita isolar as alterações em uma máquina virtual e manter um disco rígido virtual em estado inalterado. O disco de diferenciação no SO de gerenciamento pode ser compartilhado com máquinas virtuais e, como prática recomendada, deve permanecer como "somente leitura". Caso contrário, o disco rígido virtual da máquina virtual será invalidado.

Com discos rígidos virtuais, cada máquina virtual oferece suporte a até 512 TB de armazenamento. Os discos físicos diretamente conectados a um equipamento virtual não têm limite de tamanho, a não ser a do Sistema Operacional convidado.

É possível utilizar discos rígidos virtuais como uma opção de armazenamento no SO de gerenciamento e disponibilizar o armazenamento para as máquinas virtuais. A criação de um novo disco virtual é feita por meio do Assistente para Criação de Discos Rígidos Virtuais ou

Assistente de Nova Máquina Virtual. Se for um disco de expansão dinâmica, o assistente de Nova Máquina Virtual mostra uma forma de criar armazenamento para o novo equipamento virtual, sem executar o Assistente para Criação de Discos Rígidos Virtuais. Isso pode ser útil para instalar um Sistema Operacional convidado em uma máquina virtual logo após a sua criação.

Quando um novo disco rígido virtual é criado, são necessários um nome e um local de armazenamento. Os discos são armazenados como arquivos .VHD, o que os torna portáteis, mas também representa uma ameaça de segurança potencial. É preciso eliminar esse risco tomando algumas precauções, como armazenar os arquivos .VHD em um local seguro. Não se deve criar o disco rígido virtual em uma pasta marcada para criptografia. O Hyper-V não oferece suporte ao uso de mídia de armazenamento se o Encrypting File System tiver sido usado para o arquivo .VHD. Entretanto, é possível usar os arquivos armazenados em um volume que use Criptografia da Unidade de Disco BitLocker do Windows.

Discos físicos para máquinas virtuais

É possível usar discos físicos diretamente conectados a uma máquina virtual como opção de armazenamento no Sistema Operacional de gerenciamento. Isso permite que os equipamentos acessem o armazenamento mapeado diretamente para o servidor que executa o Hyper-V, sem primeiro configurar o volume. O armazenamento pode ser um disco físico interno do servidor ou um LUN (número de unidade lógica) em um SAN mapeado para o servidor. A máquina virtual deve ter acesso exclusivo ao armazenamento, de modo que ele possa ser definido em um estado off-line em gerenciamento de disco. O armazenamento não tem limite de tamanho, portanto, um LUN pode ser ilimitado, ou seja, limitado ao tamanho máximo suportado pelo Sistema Operacional.

Ao usar discos físicos diretamente conectados a uma máquina virtual, deve-se ter ciência de três limitações: esse tipo de disco não pode ser expandido dinamicamente; não é viável usar discos de diferenciação com eles; nem gerar snapshots do disco rígido.

Aspectos de segurança

- Práticas recomendadas

O servidor de virtualização pode ser protegido com as mesmas medidas usadas para qualquer servidor que execute o Windows Server 2012. Além disso, é preciso empregar algumas ações extras para ajudar na segurança das máquinas virtuais, dos arquivos de configuração e dos dados. Proteger as máquinas virtuais em execução no servidor de virtualização significa agir de acordo com os procedimentos desse tipo de servidor ou carga de trabalho. Por exemplo, se as diretivas e os procedimentos exigirem a aplicação de um antivírus, ele deve ser executado no equipamento. Se houver um requisito de diretiva para segmentar o servidor físico para uma rede particular, siga também a diretiva para a máquina virtual.

Práticas recomendadas de segurança

Usar uma instalação Server Core do Windows Server 2012 para o Sistema Operacional de gerenciamento.

Uma instalação Server Core oferece a menor superfície de ataque e reduz o número necessário de correções, atualizações e reinicializações para manutenção.

Não executar nenhum aplicativo no Sistema Operacional de gerenciamento.

Todos os aplicativos devem ser executados em máquinas virtuais. Ao manter o Sistema Operacional de gerenciamento sem aplicativos e executando uma instalação principal do Windows Server 2012, menos atualizações são obtidas para o SO de gerenciamento. A justificativa é que não serão necessárias atualizações de software, apenas do Server Core, dos componentes de serviço do Hyper-V e do hipervisor.

Usar o nível de segurança das máquinas virtuais para determinar o nível de segurança do Sistema Operacional de gerenciamento.

Implantar as máquinas virtuais nos servidores de virtualização com requisitos de segurança similares. Por exemplo, considerar que o nível de risco e o esforço para proteger os servidores estão classificados em três categorias: "proteção", "mais proteção" e "proteção máxima". É preciso empregar mais procedimentos de controle e esforço de conformidade nos servidores com proteção máxima do que nos servidores com proteção. Isso será verdadeiro se o servidor for físico ou estiver em execução em uma máquina virtual. Se forem implantadas máquinas virtuais com proteção máxima no Sistema Operacional de gerenciamento, será necessário proteger o servidor de virtualização como um servidor de "proteção máxima". A implantação de máquinas virtuais com níveis de segurança similares em um servidor de virtualização pode facilitar o gerenciamento e a mobilidade das máquinas virtuais.

Não conceder permissões aos administradores de máquina virtual no Sistema Operacional de gerenciamento.

De acordo com o princípio de privilégio mínimo, concede-se aos administradores de uma máquina virtual (às vezes chamados de administradores de departamento ou administradores delegados) as permissões mínimas necessárias. O gerenciamento delas em todos os objetos associados ao equipamento pode ser complexo e causar problemas de segurança potenciais se não for manipulado corretamente. O controle de acesso baseado em função permite especificar o controle de acesso em termos de estrutura organizacional de uma empresa, por meio da criação de um novo objeto denominado "função". É possível atribuir uma determinada função a um usuário quando ele exerce um cargo. O Hyper-V usa as diretivas do Gerenciador de Autorização para controle de acesso baseado em função.

Verificar se as máquinas virtuais estão totalmente atualizadas antes de serem implantadas em um ambiente de produção.

Como as máquinas virtuais possuem mais mobilidade e são mais rápidas de implantar do que os computadores físicos, há risco maior de que uma máquina virtual não totalmente

atualizada ou com todas as correções aplicadas seja implantada. Para gerenciar isso com eficiência, devem ser usados os mesmos métodos e procedimentos de atualização de máquinas virtuais empregados na atualização de servidores físicos. Por exemplo, se for permitido o emprego de atualizações automáticas com o Windows Update, Microsoft System Center Configuration Manager ou outro método de distribuição de software, é necessário verificar se as máquinas virtuais estão atualizadas e/ou com as correções aplicadas antes de serem implantadas.

Usar um adaptador de rede dedicado para o Sistema Operacional de gerenciamento do servidor de virtualização.

Por padrão, nenhuma rede virtual está configurada para o Sistema Operacional de gerenciamento. É recomendado o uso de um adaptador de rede dedicado para gerenciar o servidor que executa o Hyper-V e não o expor ao tráfego de rede não confiável. Não se deve permitir que máquinas virtuais utilizem esse adaptador de rede. É recomendado que sejam usados um ou mais adaptadores de rede dedicados diferentes para a rede da máquina virtual. Isso permite aplicar níveis diferentes de configuração e diretiva de segurança de rede para elas. Por exemplo, é possível configurar a rede para que as máquinas virtuais tenham acesso de rede diferente em relação ao Sistema Operacional de gerenciamento, incluindo o uso de VLANs, IPsec, Proteção de Acesso à Rede (NAP) e Microsoft Forefront Threat Management Gateway.

Usar a Criptografia de Unidade de Disco BitLocker para proteger os recursos.

Funciona com os recursos de hardware e firmware de servidor para oferecer criptografia de unidade de disco e inicialização de sistemas operacionais seguros, mesmo quando o servidor não estiver ativo. Isso ajuda a proteger os dados se um disco for roubado e instalado em outro computador para mineração de dados. A Criptografia de Unidade de Disco BitLocker também ajuda na segurança dos dados se um invasor usar um Sistema Operacional diferente ou executar uma ferramenta de ataque a software para acessar um disco.

A perda de um disco físico representa risco maior em cenários de pequenas e médias empresas e em escritórios remotos, no qual a segurança física do servidor pode não ser tão rigorosa como em um datacenter corporativo. Entretanto, o uso da Criptografia de Unidade de Disco BitLocker faz sentido para todos os computadores. Use a Criptografia de Unidade de Disco BitLocker em todos os volumes que armazenam arquivos de máquina virtual, incluindo discos rígidos virtuais, arquivos de configuração, snapshots e qualquer recurso de máquina virtual, como imagens ISO e disquetes virtuais. Para um nível de segurança mais alto que contenha inicialização protegida, a Criptografia de Unidade de Disco BitLocker requer o hardware Trusted Platform Module (TPM).

É recomendado o uso da Criptografia de Unidade de Disco BitLocker no Sistema Operacional de gerenciamento do Hyper-V, além de proteger os volumes que contêm arquivos de configuração, discos rígidos virtuais e instantâneos. Não se deve executar a Criptografia de Unidade de Disco BitLocker em uma máquina virtual, pois ela não oferece suporte para isso.

Desabilitar as configurações de BIOS de virtualização quando desnecessárias.

Quando um servidor empregado para virtualização não for mais usado, por exemplo em um cenário de teste ou desenvolvimento, é recomendada a desativação das configurações de BIOS de virtualização assistidas por hardware e necessárias ao Hyper-V.

Backup e Restore

- Backup on-line e off-line
- Processo de restauração (restore)
- Considerações sobre máquinas virtuais em cluster

Quando se planeja uma estratégia de backup e recuperação para um ambiente de servidor virtualizado, há diversos fatores a considerar. É necessário avaliar os vários tipos de backups que podem ser feitos, o estado da máquina virtual e o tipo de armazenamento em uso. Abordaremos as vantagens e desvantagens desses fatores.

Como visto na seção anterior, o Hyper-V oferece duas possibilidades diferentes de execução do backup: a partir do servidor que o executa e a partir do Sistema Operacional convidado. No entanto, é preciso planejar a estratégia de backup em função da compatibilidade entre as soluções de backup e armazenamento. Alguns pontos a considerar:

- **Discos rígidos virtuais:** oferecem a melhor compatibilidade e podem ser armazenados em muitos tipos de mídia física.
- **Discos físicos diretamente conectados a uma máquina virtual:** o gravador Volume Shadow copy Service (VSS) do Hyper-V não pode fazer o backup desses discos. Consequentemente, esse tipo de disco não será incluído em nenhum backup executado por um programa de backup que use o gravador VSS do Hyper-V. A situação exige o emprego de outro processo para fazer o backup do disco físico, como rodar um aplicativo de backup no Sistema Operacional convidado.
- **Armazenamento baseado em iSCSI:** armazenamento com suporte para backup do gravador VSS do Hyper-V quando o armazenamento estiver conectado pelo SO de gerenciamento e for usado pelos discos rígidos virtuais.
- **Armazenamento acessado a partir de uma máquina virtual com iniciador iSCSI no Sistema Operacional convidado:** esse armazenamento não é incluso em um backup do computador físico. Por isso, é preciso usar outro processo para fazer o backup dos dados do armazenamento baseado em Internet SCSI (iSCSI) antes da realização de um backup completo de servidor. Por exemplo, é possível fazer um backup dos dados no armazenamento iSCSI de um aplicativo de backup em execução no SO convidado.

Backup on-line e off-line

A execução de backup em estado on-line ou off-line depende da possibilidade de ser feito sem tempo de inatividade. É possível fazer um backup on-line sem nenhum tempo de inatividade em uma máquina virtual em execução, quando as seguintes condições forem atendidas:

- Os serviços de integração estão instalados e o serviço de integração de backup não foi desabilitado;
- Todos os discos que estiverem sendo usados pela máquina virtual são configurados no Sistema Operacional convidado como discos básicos formatados em NTFS. As máquinas virtuais que usam o armazenamento em que as partições físicas foram formatadas como discos dinâmicos (ou o sistema de arquivos FAT32) impedem a realização de um backup on-line. Isso não é o mesmo que expandir dinamicamente os discos rígidos virtuais, que são totalmente suportados pelas operações de backup e restauração;
- O Serviço de Cópia de Sombra de Volume deve ser habilitado em todos os volumes usados pela máquina virtual com uma configuração específica. Cada volume também deve servir como local de armazenamento de cópias de sombra do volume. Por exemplo, o armazenamento de cópia de sombra do volume **C:** deve estar localizado em **C:**.

Quando não for possível realizar um backup on-line, será feito um backup off-line. Esse tipo resulta em algum tempo de inatividade. Vários fatores podem afetar o tempo necessário para a realização de um backup off-line. Se a máquina virtual estiver em execução ou pausada, ela será colocada em um estado “Saved” como parte do processo de backup off-line. Após a conclusão do backup, a máquina virtual retornará ao seu estado atual.

Processo de restauração

O processo de restauração (restore) é direto, desde que as recomendações descritas anteriormente sejam seguidas quando os backups forem criados. Isso inclui passar pelas etapas recomendadas para assegurar que os dados não inclusos em um backup de servidor completo possam ser recuperados ou criados novamente.

Para executar uma restauração quando todos os componentes do conjunto de backups contam com suporte do gravador VSS do Hyper-V, basta usar as mídias disponíveis e fazer uma restauração de todo o sistema ou de uma máquina virtual específica, dependendo das circunstâncias. O gravador VSS trata o Hyper-V como um aplicativo cujo backup pode ser feito. Isso significa que permite recuperar máquinas virtuais individuais. Porém, não é possível usar esse método para recuperar apenas uma porção de uma máquina virtual.

Se o conjunto de backups contiver mídia sem suporte para gravador VSS, deve-se seguir uma etapa adicional. Primeiro, é necessário restaurar todo o sistema ou uma máquina virtual, dependendo das circunstâncias. Na sequência, é feita a restauração da mídia sem suporte ao VSS no Sistema Operacional convidado. Ao tentar restaurar uma máquina virtual enquanto estiver em execução, ela será desativada e excluída antes da restauração do backup.

Ao iniciar uma máquina virtual restaurada a partir de um backup on-line, o sistema enviará uma mensagem informando que o Sistema Operacional não foi desligado corretamente, mensagem que pode ser ignorada de forma segura.

Máquinas virtuais em cluster

Se a solução de virtualização considerar o armazenamento das máquinas virtuais em cluster, há fatores adicionais que precisam ser considerados no planejamento do backup e na restauração dessas máquinas virtuais. Antes de tentar restaurar as máquinas virtuais clusterizadas, ou fazer o seu backup, é necessário cuidar de dois pontos. O primeiro diz respeito à aplicação de um hotfix para impedir possíveis falhas de um backup completo em um servidor, quando a máquina virtual usar um volume montado por meio do mecanismo de Globally Unique IDentifier (GUID). Quando o hotfix for aplicado, um caminho de diretório que não puder ser determinado impedirá o backup apenas da máquina virtual que usa o caminho de diretório. Entretanto, quando o hotfix não for aplicado, um volume montado com uma GUID poderá causar a falha de toda a operação. O segundo ponto é avaliar a necessidade e as consequências de pôr a máquina virtual off-line antes de executar um backup ou restaurar uma máquina virtual.

Redes virtuais

- Noções básicas
- Rede e máquinas virtuais
- Configurando VLANs – redes locais virtuais

É possível criar redes virtuais no servidor Hyper-V para definir várias topologias entre as máquinas virtuais e o servidor de virtualização. Usando o Gerenciador de Rede Virtual (acessado pelo Gerenciador Hyper-V), há três tipos diferentes de redes virtuais.

- **Redes virtuais externas:** permitem que as máquinas virtuais se comuniquem com servidores localizados externamente e com o Sistema Operacional de gerenciamento (partição pai). Esse tipo ainda possibilita que máquinas virtuais no mesmo servidor físico se comuniquem.
- **Redes virtuais internas**: possibilitam a comunicação entre máquinas virtuais no mesmo servidor físico e o SO de gerenciamento. A rede virtual interna é uma rede virtual não associada a um adaptador de rede físico. Esse tipo é geralmente usado para criar um ambiente de teste, onde é preciso conectar as máquinas virtuais a partir do SO de gerenciamento.
- **Redes virtuais privadas:** deixam a comunicação apenas entre máquinas virtuais no mesmo servidor físico. A rede virtual privada é uma rede virtual sem um adaptador virtual no Sistema Operacional de gerenciamento. As redes virtuais privadas geralmente são usadas quando é preciso isolar as máquinas virtuais do tráfego de rede no SO de gerenciamento e das redes externas.

Noções básicas

Enquanto o Hyper-V permite configurar ambientes complexos de rede virtual, o conceito básico de rede virtual é direto. Para uma configuração simples, é recomendado ter pelo menos dois adaptadores de rede no servidor que executa o Hyper-V: um adaptador de rede dedicado ao computador físico para gerenciamento remoto e um ou mais adaptadores de rede dedicados às máquinas virtuais. Se estiver executando um iniciador iSCSI para armazenamento em disco rígido virtual, é recomendado o uso de adaptadores de rede adicionais no Sistema Operacional de gerenciamento. Esse sistema é uma partição que chama o hipervisor do Windows e solicita que novas partições sejam criadas. Só pode haver um SO de gerenciamento.

Ao adicionar a função Hyper-V durante uma instalação completa do Windows Server 2012, há a opção de configurar uma ou mais redes virtuais externas. Essa possibilidade não fica disponível ao se executar a instalação Server Core do Windows Server 2012. Os adaptadores de rede virtuais poderão ser renomeados para reflexão se forem atribuídos ao computador físico ou às máquinas virtuais.

Quando é instalado o Hyper-V e criada uma rede virtual externa, o Sistema Operacional de gerenciamento usa um novo adaptador de rede virtual para conexão à rede física. As conexões de rede são compostas pelo adaptador de rede original e pelo novo adaptador de rede virtual. O adaptador de rede físico original não tem nada associado a ele; já o adaptador de rede virtual tem todos os protocolos padrão e os serviços.

O Hyper-V associa o protocolo de serviço de rede virtual a um adaptador de rede físico quando uma rede virtual externa é criada. Deve-se estar ciente de que a conectividade de rede externa será interrompida temporariamente quando uma rede virtual externa for criada ou excluída.

Uma vez criada, a rede virtual funciona como uma rede física, exceto pelo fato de que o comutador é baseado em software, e as portas podem ser adicionadas ou removidas dinamicamente conforme necessário.

Quando uma rede virtual externa for configurada, todo o tráfego de rede será encaminhado pelo comutador virtual. Por esse motivo, recomendamos usar pelo menos um adaptador de rede físico adicional para o gerenciamento do tráfego de rede. O comutador virtual funciona como um comutador físico e encaminha o tráfego de rede pela rede virtual ao seu destino. A figura a seguir é um exemplo de rede virtual externa.

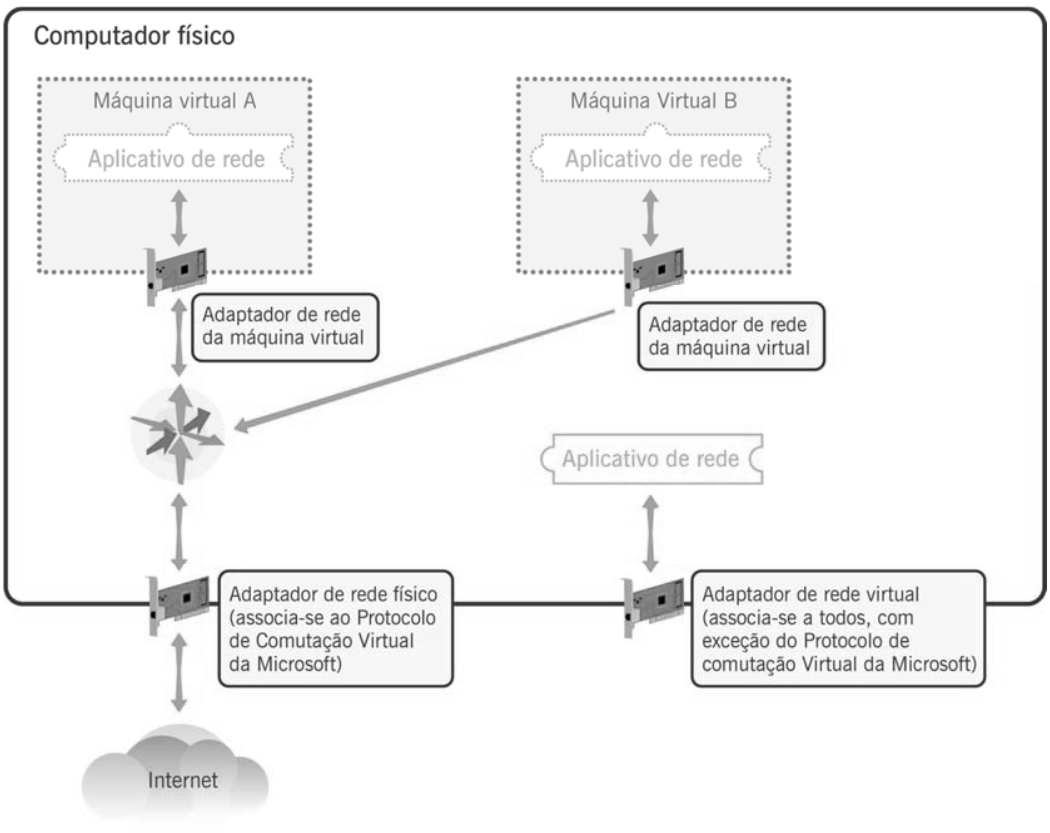


Figura 6.4 Rede virtual externa.

Em redes virtuais internas, só é permitida a comunicação entre máquinas virtuais no mesmo servidor físico e entre elas e o Sistema Operacional de gerenciamento. A próxima figura ilustra um exemplo de rede virtual interna.

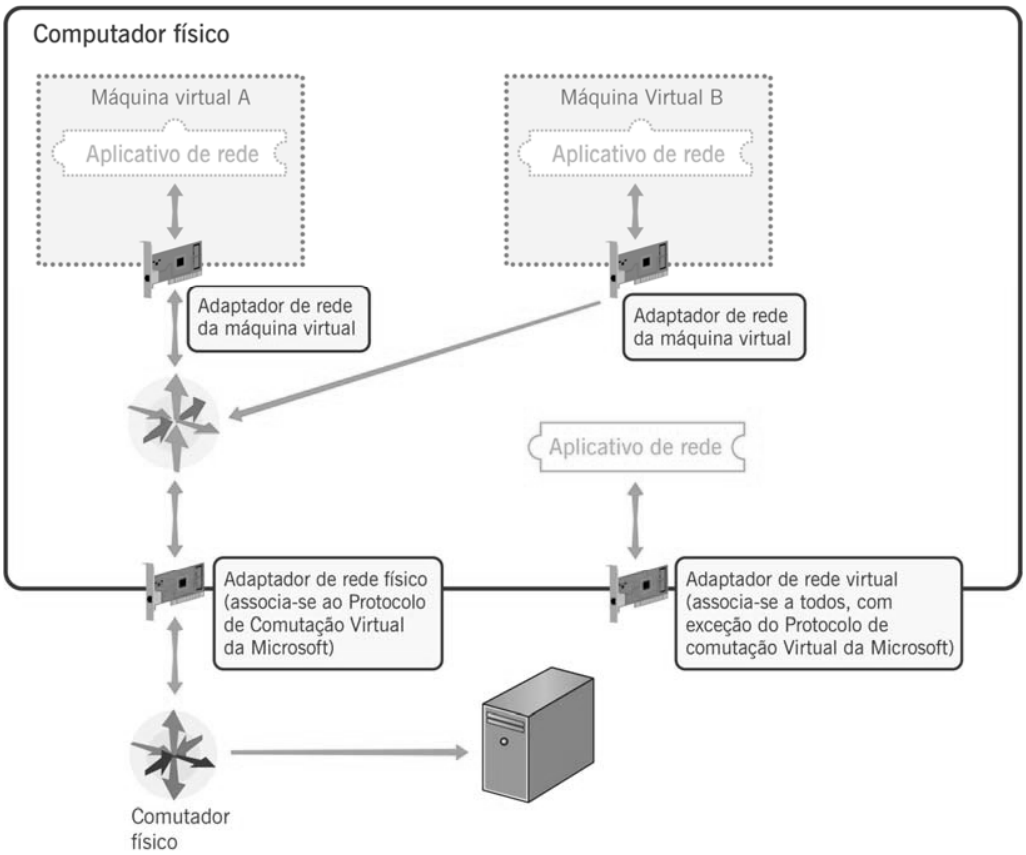


Figura 6.5 Rede virtual interna.

Use uma rede virtual privada quando quiser permitir a comunicação apenas entre máquinas virtuais no mesmo servidor físico. A figura seguinte mostra um exemplo de rede virtual privada.

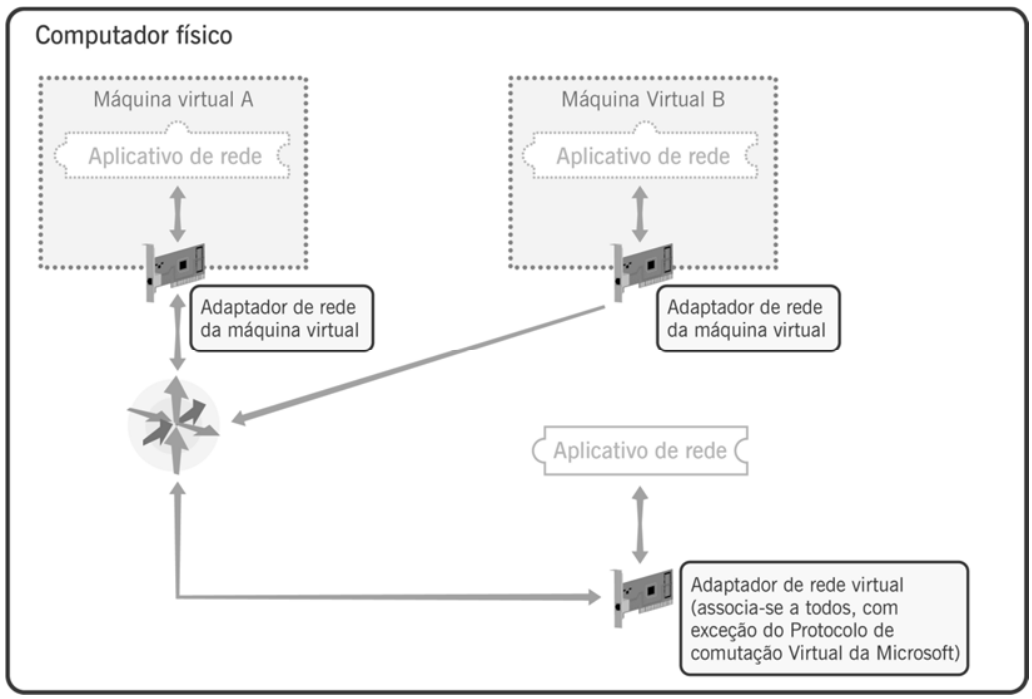


Figura 6.6 Rede virtual privada.

Rede e máquinas virtuais

No Hyper-V, quando uma máquina virtual é criada e vinculada a uma rede virtual, a conexão se estabelece por meio de um adaptador de rede virtual. Há dois tipos de adaptadores de rede disponíveis para o Hyper-V: o de rede e o de rede legado. Para que o adaptador de rede funcione, os serviços de integração devem estar instalados, o que faz parte da instalação do Hyper-V. Se os serviços de integração não puderem ser instalados por causa da versão do Sistema Operacional, o adaptador de rede não poderá ser usado. Em vez disso, será preciso adicionar um adaptador de rede legado que emule um PCI Fast Ethernet baseado em Intel 21140 e funcione sem a instalação de um driver de máquina virtual. Um adaptador de rede legado também oferece suporte a instalações baseadas em rede, porque ele pode inicializar no Pre-Boot Execution Environment (PXE). Ele será necessário também se uma máquina virtual precisar ser inicializada a partir de uma rede, para desabilitar o adaptador de rede, após a inicialização no PXE.

A máquina virtual está logicamente conectada a uma porta na rede virtual. Para que um aplicativo de rede da máquina virtual se conecte externamente, primeiro ele é encaminhado por meio do adaptador a uma porta lógica na rede virtual externa, a qual a VM está vinculada. O pacote de rede é direcionado ao adaptador de rede físico e para uma rede física externa.

Para que a máquina virtual se comunique com o Sistema Operacional de gerenciamento, há duas opções. Uma é encaminhar o pacote de rede pelo adaptador de rede físico e para a rede física, que retorna o pacote ao servidor Hyper-V usando o segundo adaptador de rede físico.

Outra mais eficiente é encaminhar o pacote de rede pela rede virtual. A opção selecionada é determinada pela rede virtual, que contém um algoritmo de aprendizado; ele, por sua vez, determina a porta mais eficiente para direcionar o tráfego; enviar o pacote de rede a essa porta. Até que essa determinação seja feita pela rede virtual, os pacotes de rede serão enviados a todas as portas virtuais.

Configurando VLANs

O Hyper-V oferece suporte a Redes Locais Virtuais (VLANs) e, como a configuração de uma VLAN é baseada em software, os computadores podem ser movidos facilmente e ainda manterem suas configurações de rede. Para cada adaptador de rede virtual conectado a uma máquina virtual, é possível configurar um ID de VLAN para a máquina virtual. Será necessário o seguinte para configurar VLANs:

- Um adaptador de rede físico que ofereça suporte a VLANs;
- Um adaptador de rede físico que ofereça suporte a pacotes de rede com IDs da VLAN já aplicados.

No Sistema Operacional de gerenciamento, a rede virtual deve ser configurada para permitir o tráfego de rede na porta física para os IDs da VLAN, que se deseja usar internamente com máquinas virtuais. Em seguida, é preciso configurar a máquina virtual para especificar a LAN virtual que ela usará para todas as comunicações de rede.

Existem dois modos de configuração de uma VLAN: modo de acesso e modo de tronco. No modo de acesso, a porta externa da rede virtual fica restrita a um único ID da VLAN na interface do usuário. É possível ter várias VLANs usando Windows Management Interface (WMI). Deve-se utilizar o modo de acesso quando o adaptador de rede físico estiver conectado a uma porta no comutador de rede físico, que também está no modo de acesso. Para dar acesso externo a uma VM na rede virtual que está no modo de acesso, será preciso configurar a VM para usar o mesmo ID da VLAN configurado no modo de acesso da rede virtual.

O modo de tronco permite que vários IDs da VLAN compartilhem a conexão entre o adaptador de rede físico e a rede física. Para fornecer acesso externo às máquinas virtuais na rede virtual em várias VLANs, é preciso configurar a porta na rede física que deve estar no modo de tronco. Também é necessário conhecer as VLANs específicas e todos os IDs da VLAN usados pelas máquinas virtuais às quais a rede virtual oferece suporte.

7

Introdução ao VMware

- Produtos
  - VMware vSphere
  - VMware vCenter
  - VMware hipervisor

A VMware é a empresa que introduziu o conceito de virtualização na arquitetura x86 de computadores. Fundada por Diane Greene e seu marido, Dr. Mendel Rosenblum; dois estudantes da Universidade de Stanford; e um colega de trabalho em 1999 –, a empresa lançou seu primeiro produto de virtualização baseado na plataforma x86, hoje conhecido como VMware Workstation.

Em 2000, a VMware criou o GSX Server 1.0, o primeiro software desenvolvido para servidores. Para a ferramenta entrar em produção, era necessário instalar um Sistema Operacional (Windows ou Linux). No ano seguinte, a VMware disponibilizou a ferramenta ESX 1.0 e, com ela, o modelo de virtualização baseado em hipervisor (bare-metal) foi introduzido. Dessa forma, passou a ser uma solução mais robusta, estável e com melhor performance.

De 2002 até agora, a VMware aprimorou os recursos existentes e acrescentou diversas funcionalidades, principalmente voltadas para recuperação de desastres e otimização dos recursos computacionais.

Atualmente, a VMware deixou de ser uma empresa que comercializa apenas software para a virtualização de servidores. A empresa visa a entregar todos os componentes existentes no datacenter, como software, trazendo mais uma nova solução: os datacenters definidos por software. Uma plataforma automatizada e gerenciada que abrange aplicativos com a possibilidade de implantação rápida de datacenters, nuvens e dispositivos móveis.

Os produtos VMware disponíveis para a consolidação de servidores estão divididos em dois grandes pacotes:

- **VMware vSphere**: permite transformar o datacenter em uma infraestrutura em nuvem, oferecendo como principais recursos os serviços de infraestrutura e de aplicativos.
- **VMware vCenter**: tem como recursos principais o gerenciamento centralizado da infraestrutura e a operação da plataforma virtualizada; administração das aplicações e do negócio de TI. O gerenciamento da plataforma é essencial para administrar os serviços disponibilizados pelo vSphere. Serviços de infraestrutura oferecidos:
  - **vCompute**: composto pelo hipervisor (ESXi) e pelo VMware Distributed Resource Scheduler (DRS).

- **vStorage**: permite acesso mais eficiente aos sistemas de armazenamento em ambientes virtualizados, abstraindo a complexidade desses sistemas.
- **vNetwork**: os serviços do vNetwork objetivam facilitar o gerenciamento de redes em ambientes virtualizados. Os serviços de aplicativos são focados na melhor utilização possível da plataforma de virtualização. São divididos em serviço para tempo de inatividade planejado (VMware vMotion e Storage vMotion); serviço para tempo de inatividade não planejado (VMware High Availability – HA); Fault Tolerance (FT) e VMware Data Recovery; e serviços de segurança (VMware vShield Zones e VMsafe).

Na parte de infraestrutura, o vCenter é composto por diversos produtos: vCenter Server, Heartbeat, Operations, Orchestrator, CapacityIQ, Site Recovery Manager (SRM), Configuration Manager, Lab Manager e Converter. Já em aplicações: vCenter Application Discovery Manager, AppSpeed e VMware Studio. Para o gerenciamento de negócios, há o vCenter Chargeback e o Service Manager.

- **VMware Hipervisor**: hipervisor bare-metal da arquitetura VMware, funciona diretamente sobre o servidor físico, sem a necessidade de um Sistema Operacional. Permite isolar o SO/aplicativo do hardware, possibilitando carregar vários aplicativos em um único servidor físico, sem a chance do mau funcionamento de um aplicativo interferir no outro.

vSphere

- Conceito
- Componentes da solução

O VMware vSphere foi o primeiro Sistema Operacional para computação em nuvem (cloud computing). Ele utiliza os recursos da virtualização para transformar datacenters em infraestruturas de computação em nuvem, consideravelmente simplificadas, e permite que as organizações de TI forneçam a próxima geração de serviços flexíveis e confiáveis – por meio de recursos internos e externos, com segurança e baixo risco. Baseado nos recursos da plataforma VMware Infrastructure, o vSphere reduz os custos operacionais e de capital, além de aumentar o controle sobre o fornecimento de serviços de TI, enquanto preserva a flexibilidade de escolha entre qualquer tipo de Sistema Operacional, aplicativo e hardware hospedados internamente ou que usam recursos externos.

O VMware vSphere é formado por seis componentes, conforme pode ser visto na figura 7.1.

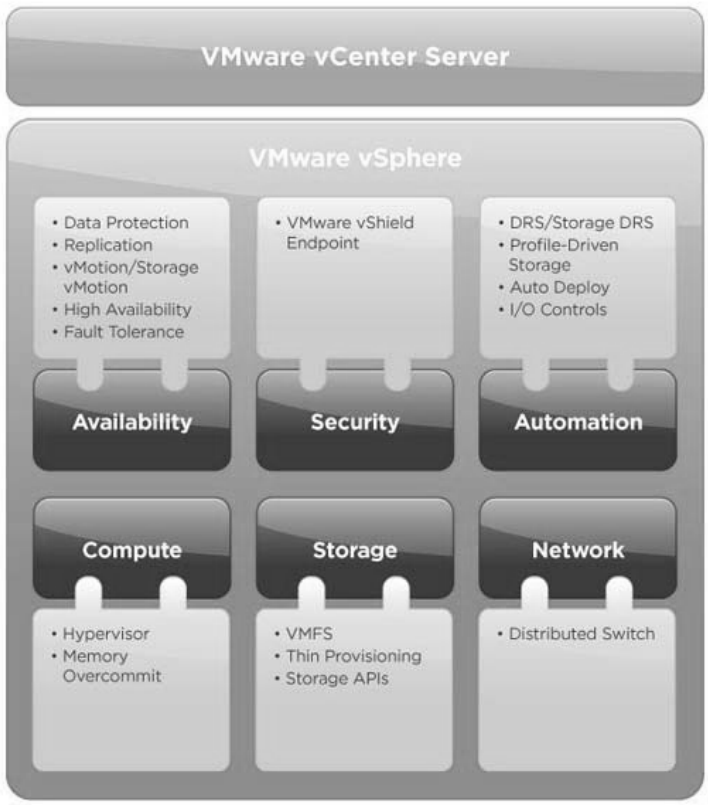


Figura 7.1 VMware vSphere.

Availability

Os recursos de disponibilidade permitem que o setor de TI forneça aplicativos com níveis variados de alta disponibilidade, de acordo com a prioridade e a necessidade, sem precisar de hardwares complexos redundantes, nem software de cluster.

- **Data Protection**: permite backup e recuperação simples e econômica de máquinas virtuais, sem agentes externos ou de terceiros.
- **Replication**: mecanismo de replicação proprietário da VMware. Copia apenas blocos alterados no site de recuperação. Essa abordagem reduz a utilização de largura de banda e possibilita RPOs (recovery point objectives, objetivos de ponto de recuperação) mais agressivos do que a replicação de máquina virtual manual completa do sistema.
- **vMotion/Storage vMotion**: elimina a necessidade de programar o tempo de inatividade de aplicativos em virtude da manutenção de servidor planejada ou durante as migrações de servidores/armazenamento, por meio da migração "a quente" de máquinas virtuais/discos, sem interrupção das atividades dos usuários nem perda de serviço.

- **High Availability (HA):** oferece a inicialização automática e econômica em minutos para todos os aplicativos, em caso de falhas de hardware ou do Sistema Operacional.
- **Fault Tolerance (FT):** oferece disponibilidade contínua para qualquer aplicativo sem perda de dados nem tempo de inatividade, em caso de falhas de hardware.

Security

Os recursos de segurança objetivam garantir uma maior proteção para as máquinas virtuais e suas aplicações.

- **vShield Endpoint**: permite gerenciar políticas de antivírus e antimalware para ambientes virtualizados, com as mesmas interfaces de gerenciamento usadas para proteger a infraestrutura física. O vShield Endpoint reforça a segurança da virtualização, com proteção aprimorada dos endpoints, ao descarregar o processamento de AV em um appliance virtual fornecido por parceiros da VMware vShield Endpoint.

Automation

Com a automação de tarefas, é possível fornecer soluções mais precisas, consistentes e repetíveis, que economizam tempo e esforços dos administradores de TI, além de minimizarem os erros provenientes de rotinas manuais. O vSphere oferece algumas funcionalidades que permitem automatizar tarefas administrativas da infraestrutura virtual.

- **DRS/Storage DRS**: balanceia a capacidade de processamento/armazenamento por cluster a fim de oferecer um desempenho otimizado para hosts, storage e máquinas virtuais.
- **Profile-Driven Storage**: permite visibilizar pools de armazenamento, otimizar e automatizar o aprovisionamento do armazenamento. Possibilita que os administradores do VMware vSphere superem os desafios do aprovisionamento antecipado de armazenamento, como planejamento de capacidade, níveis de serviço diferenciados e gerenciamento do espaço de capacidade.
- **Auto-Deploy**: facilita a implantação de servidor e o aprovisionamento rápido de hosts, aproveitando os recursos de inicialização de rede dos servidores x86. Os hosts do vSphere são inicializados em rede de um servidor central do Auto Deploy, no qual o software do ESXi é instalado diretamente na memória do servidor.
- **I/O Controls**: possibilita priorizar E/S para máquinas virtuais em execução num grupo de hosts do VMware vSphere, que têm acesso a um pool de armazenamento compartilhado.

Compute

O VMware Compute oferece os serviços de infraestrutura que virtualizam recursos de servidor com eficiência e os incorpora a conjuntos lógicos (pools) – que podem ser alocados com precisão em aplicativos.

- **Hipervisor:** a solução VMware Hipervisor é baseado no servidor ESXi; um hipervisor independente que é executado diretamente no hardware do hospedeiro. Ou seja, é um monitor de máquina virtual do tipo I (bare metal). O ESXi ocupa apenas 32 MB de memória e, inclusive, pode ser integrado diretamente no firmware do host. Ao considerar que os servidores em geral estão subutilizados, a virtualização com o ESXi possibilita otimizar o uso dos recursos e, ao mesmo tempo, reduzir o consumo de energia, além de conceder nível maior de disponibilidade.
  - **VMware ESXi**: oferece uma camada de virtualização robusta e de bom desempenho, que abstrai os recursos de hardware do servidor e permite o compartilhamento desses recursos entre várias máquinas virtuais. Os recursos de gerenciamento de memória e programação do VMware ESX e ESXi fornecem taxas de consolidação e bom desempenho de aplicativo.
- **Memory Overcommit:** é uma funcionalidade do hipervisor que permite a uma máquina virtual utilizar mais espaço de memória que o disponível no host. Por exemplo, um servidor de virtualização como o VMware ESXi, com 2 GB de memória física, pode executar quatro máquinas virtuais, cada uma com 1 GB de espaço de memória alocada. A ideia pode parecer perigosa, porque um computador irá apresentar falhas se a memória física esgotar. Porém, na prática, o comprometimento excessivo de recursos de computação de servidor é inofensivo, considerando que apenas uma pequena parte da memória física de fato é alocada para uma máquina virtual. Por exemplo: uma máquina virtual com 1 GB de memória física alocada para ele pode precisar apenas de 400 MB, deixando 600 MB de espaço alocado não utilizado. Com quatro máquinas, ao usar 400 MB, o servidor físico estará na prática consumindo 1,6 GB dos 2 GB existentes. Caso todas VM's ainda assim estejam utilizando simultaneamente toda a quantidade de memória alocada para ela, será possível adotar outras técnicas de gerenciamento de recursos para evitar que a infraestrutura física seja comprometida. Dentre elas, podemos citar: Memory Ballooning e Transparent Page Shareing (TPS).

Storage

O Storage é composto pelos serviços de infraestrutura, que abstraem os recursos de armazenamento (storage) da complexidade dos sistemas subjacentes de hardware. Tem o objetivo de proporcionar o uso mais eficiente da capacidade de armazenamento em ambientes virtualizados.

- **VMware vStorage Virtual Machine File System (VMFS):** sistema de arquivos de cluster de alto desempenho, que possibilita compartilhamento eficiente e controla o acesso simultâneo de servidores virtualizados aos subsistemas de armazenamento.

- **VMware vStorage Thin Provisioning**: oferece alocação dinâmica de espaço em disco. Isso permite a utilização mais eficiente do espaço de armazenamento, postergando os investimentos na aquisição de novos recursos de armazenamento (discos), até que se tornem realmente necessários. Reduz os gastos com armazenamento em até 50%.
- **Storage APIs**: os fornecedores de armazenamento podem usar a vSphere API para fornecer ao vSphere informações sobre arrays de disco específicos, a fim de aumentar a integração entre o armazenamento e a infraestrutura virtual.

Network

O VMware Network proporciona mecanismos para administrar e gerenciar redes em ambientes virtuais. O VMware vNetwork Distributed Switch simplifica e melhora o provisionamento, a administração e o controle da rede de máquinas virtuais em ambientes VMware vSphere. Além disso, permite que switches virtuais distribuídos por terceiros (como o Cisco Nexus 1000v) sejam usados em ambientes VMware vSphere – oferecendo aos administradores de rede interfaces familiares para controle da qualidade do serviço no nível da máquina virtual.

Disponibilidade

- VMware Storage vSphere
- VMware Hight Availability (HA)
- VMware Fault Tolerance
- VMware Data Protection

Os serviços de disponibilidade permitem que o setor de TI forneça aplicativos com níveis variados de alta disponibilidade, de acordo com a prioridade e a necessidade, sem precisar de hardwares complexos redundantes nem software de cluster.

- **VMware Storage vSphere**: elimina a necessidade de programar o tempo de inatividade de aplicativos – em virtude da manutenção de servidor planejada ou durante as migrações de armazenamento por meio da migração “a quente” de discos de máquinas virtuais –, sem interrupção das atividades dos usuários nem perda de serviço.
- **VMware High Availability (HA):** oferece a inicialização automática e econômica em minutos para todos os aplicativos, em caso de falhas de hardware ou do Sistema Operacional.
- **VMware Fault Tolerance**: concede disponibilidade contínua para qualquer aplicativo, sem perda de dados nem tempo de inatividade, em caso de falhas de hardware.
- **VMware Data Protection**: fornece proteção de dados mais eficiente, dimensionável e simples para máquinas virtuais.

Segurança

- **VMware vShield Zones**
- **VMware VMsafe**

Os serviços de segurança permitem que a infraestrutura de TI forneça aplicativos com o nível apropriado de políticas de segurança de forma operacionalmente eficiente.

- **VMware vShield Zones**: simplifica a segurança de aplicativos ao utilizar políticas de segurança corporativa, no nível do aplicativo em um ambiente compartilhado. Mantém, ao mesmo tempo, a confiabilidade e a segmentação de rede de usuários e dados confidenciais.
- **VMware VMsafe**: permite o uso de produtos de segurança que funcionam em conjunto com a camada de virtualização. Isso para oferecer às máquinas virtuais níveis mais altos de segurança do que os dos servidores físicos.

Escalabilidade

- VMware DRS
- VMware Storage DRS
- Adição dinâmica
- Conexão dinâmica
- Extensão dinâmica de discos virtuais

Os serviços de escalabilidade fornecem o volume adequado de recursos a cada aplicativo, com base nas necessidades, sem interrupções.

- **VMware DRS**: balanceia dinamicamente a carga de recursos do servidor para oferecer a quantidade certa de recursos a cada aplicativo, com base na prioridade dos negócios; permite que o consumo de recursos por um aplicativo cresça ou diminua, conforme a necessidade.
- **VMware Storage DRS**: equilibra continuamente a utilização de espaço e a carga de entrada/saída do armazenamento, ao mesmo tempo em que evita o gargalo de recursos para atender aos níveis de serviço dos aplicativos.
- **Adição dinâmica**: permite a adição de CPU e memória a máquinas virtuais quando necessário, sem interrupções nem tempo de inatividade.
- **Conexão dinâmica**: concede a adição ou a remoção de armazenamento virtual e dispositivos de rede em máquinas virtuais, sem interrupção ou tempo de inatividade.
- **Extensão dinâmica de discos virtuais**: oferece a adição de armazenamento virtual para a execução de máquinas virtuais, sem interrupções nem tempo de inatividade.

Vsphere hipervisor

- Hipervisor nativo
- Permite executar diversos sistemas operacionais em um único servidor
- Isolamento entre as máquinas virtuais

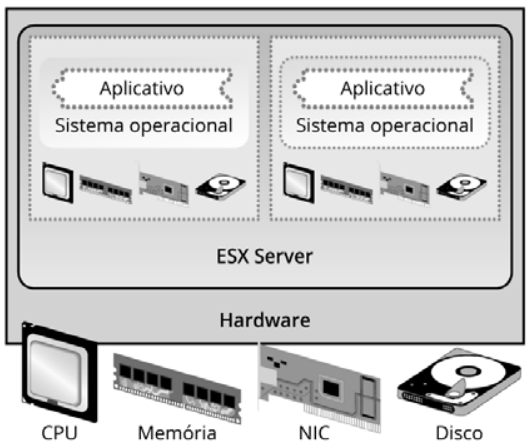


Figura 7.2 Servidor VMware vSphere (ESXi).

O ESXi é gratuito e com um licenciamento simples, pelo site da VMware. Suporta o gerenciamento centralizado com o vCenter e as demais capacidades oferecidas pelo vSphere – como a movimentação de máquinas virtuais, o balanceamento de carga e a alta disponibilidade. O VMware ESXi é composto por um Sistema Operacional denominado VMkernel (figura 5.3), em que os processos são executados, incluindo máquinas virtuais, aplicações de gerenciamento e agentes.

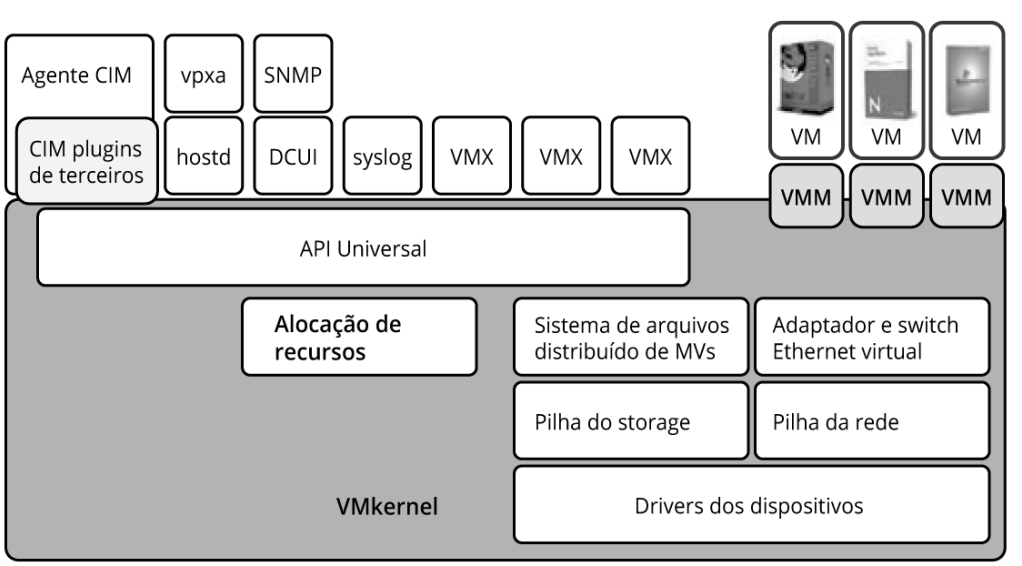


Figura 7.3 Arquitetura do hipervisor.

Principais componentes da arquitetura vSphere Hipervisor, com suas respectivas particularidades:

- **Boot from SAN:** o VMware EXSi suporta inicialização pela rede a partir da versão 4.1.
- **Simple Network Management Protocol (SNMP):** o VMware ESXI só suporta SNMP quando licenciado para VI Foundation, VI Standard ou VI Enterprise.
- **Integração com Active Directory (AD):** o VMware ESXi utiliza uma licença VI em conjunto com o Virtual Center, que permite a autenticação dos usuários por meio do AD. Nessa configuração, os usuários podem se logar diretamente no hospedeiro ESXi, com nome e senha locais na autenticação.
- **Atualizações e correções de software:** no VMware ESXi e no firmware, as correções e atualizações ocorrem da mesma forma. Qualquer versão de correção ou atualização é all-inclusive, ou seja, uma instalação de determinada correção inclui as correções das versões anteriores.

Tipos de arquivos VMware

A solução VMware vSphere utiliza diversos tipos de arquivos para armazenar as máquinas virtuais, configuração e snapshots. As extensões são:

- **<**Nome_da_MV**>.VMX** (Configuração): arquivo que contém todas as informações de configuração e definições do hardware da máquina virtual.
- **<**Nome_da_MV**>.NVRAM** (BIOS): arquivo que contém o BIOS Phoenix da máquina virtual, usado durante o boot.
- **VMDK**: quatro diferentes tipos de arquivos “Virtual Disk Data” que podem ser utilizados pelas máquinas virtuais.
- **<**Nome_da_MV**>.FLAT.VNDK**: arquivo que contém os dados do disco virtual, gerado quando um HD virtual é criado. O tamanho do arquivo dependerá da forma de definição, que pode ser thick disk ou thin disk.
- **<**Nome_da_MV-nº_sequencial**>.DELTA.VMDK**: tipo de arquivo criado quando um **snapshot** é gerado e o arquivo FLAT.VMDK se torna read-only. Qualquer alteração no HD virtual é gravada neste arquivo (DELTA.VMDK). É criado com 16 MB e incrementado de 16 em 16 MB, chegando ao tamanho máximo do arquivo flat.
- **<**Nome_da_MV**>-RDM.VMDK**: arquivo de mapeamento para o **RDM (Raw Device Mapping),** que gerencia o mapeamento de dados para os dispositivos RDW. A camada de virtualização de armazenamento passa o dispositivo mapeado como um dispositivo SCSI virtual para a máquina virtual. Um arquivo desse tipo é criado para cada RDM gerado para a máquina virtual.

- **CTK.VMDK**: arquivo utilizado pelo recurso **Changed Block Tracking (CBT)** do vSphere**.** Seu tamanho é determinado pelo do HD virtual, sendo de aproximadamente 5 MB para cada 10 GB. Um arquivo desse tipo vai existir para cada HD virtual que tenha o CBT ativado.

- **<**Nome_da_MV**>.VMSN** (estado do snapshot): tipo de arquivo usado com **snapshots** para armazenar o estado de uma máquina virtual. Um arquivo é criado para cada snapshot da máquina virtual e destruído automaticamente quando o snapshot é eliminado. O tamanho do arquivo vai variar dependendo da inclusão ou não do estado da memória da máquina virtual no snapshot.

- **<**Nome_da_MV**>.VMSD** (metadados do snapshot): tipo de arquivo usado com snapshots para armazenar os metadados e outras informações sobre cada snapshot ativo em uma máquina virtual. É iniciado com tamanho zero, sendo atualizado a cada snapshot criado ou destruído. Independentemente da quantidade de snapshots gerados, existe apenas um arquivo desse tipo para cada máquina virtual.

- **<**Nome_da_MV**>.VSWP** (memória swap): tipo de arquivo criado quando uma máquina virtual é ligada (power on). O arquivo swap é usado quando falta memória física no host e para permitir overcommitement. Seu tamanho é determinado pela quantidade de memória definida para a máquina virtual menos qualquer quantidade de memória eventualmente reservada pelo equipamento. Só é usada se não houver mais memória real disponível, podendo ocupar grandes espaços em disco. Lembre-se de que a máquina virtual não ligará se não houver espaço suficiente para o arquivo swap. É destruído quando a máquina virtual é suspensa ou desligada. Pode ser armazenado em **datastore** separado.

- **<**Nome_da_MV**>.VMSS** (Suspend): tipo de arquivo usado quando uma máquina virtual é suspensa. Serve para manter o conteúdo da memória no momento da suspensão e permitir que, no retorno, volte no mesmo estado em que parou. Seu tamanho será muito próximo ao da memória RAM disponível para a máquina virtual. Destruído quando a máquina virtual é desligada (power off).

- **vmware-nº_sequencial.LOG**: arquivos que registram a atividade de uma máquina virtual. Como todo arquivo de log, tem como principal utilidade auxiliar na depuração de eventuais erros.

- **<**Nome_da_MV**>.VMXF**: arquivo suplementar de configuração que não é usado pelo ESX, mas mantido por motivos de compatibilidade com o VMware Workstation.

Snapshot de uma máquina virtual: instantâneo tirado em um determinado momento. Serve, principalmente, para testes e depuração de possíveis erros. Como exemplo, antes de instalar qualquer software que possa interferir no funcionamento da máquina virtual, deve-se tirar um snapshot. Caso dê erro no software, basta retornar ao estado anterior com o uso do snapshot.

Memory overcommit (overcommitment): técnica utilizada pelo hipervisor (ESX/ESXi), que permite a uma máquina virtual usar mais memória do que o host tem para disponibilizar.

Datastore: local, em disco, onde são armazenados os arquivos de um determinado objeto. Por exemplo, o datastore das máquinas virtuais, dos arquivos ISO, entre outros.

thick disk/thin disk: são os HDs virtuais que, para as máquinas virtuais, parecem possuir um tamanho, mas ocupam apenas um HD virtual que foi definido com 10 GB – e usam metade. Ocupam, no host, apenas 5 GB. Para a máquina virtual, aparentam ter 10 GB. Os thick disks utilizam todo o espaço definido.

RDM (Raw Device Mapping): arquivo especial existente em volumes VMFS (Virtual Machine File System), que agem como proxy para os dispositivos raw. Permite acesso aos dispositivos de armazenamento, como discos, diretamente; sem passar pelos cachês e buffers do Sistema Operacional.

Dispositivos raw: em sistemas Unix-like, é um tipo especial de dispositivo que usa bloco como unidade e permite acesso direto ao dispositivo de armazenamento, disco, CD/DVD e regiões de memória.

Changed Block Tracking (CBT): o recurso mantém uma relação de blocos em disco que sofreram alteração, a partir de um determinado momento. Permite um ganho no desempenho aos softwares de backup incremental.

ⓘ A solução VMware Workstation, que opera no Linux ou Windows, utiliza as seguintes extensões:
LOG – VMware Workstation log file
NVRAM – VMware virtual machine BIOS
VMDK – VMware virtual disk file
VMSD – VMware snapshot metadata
VMX – VMware virtual configuration
VMXF – VMware team member
VMEM – VMEN snapshot file
VMSN – VMware virtual machine snapshot

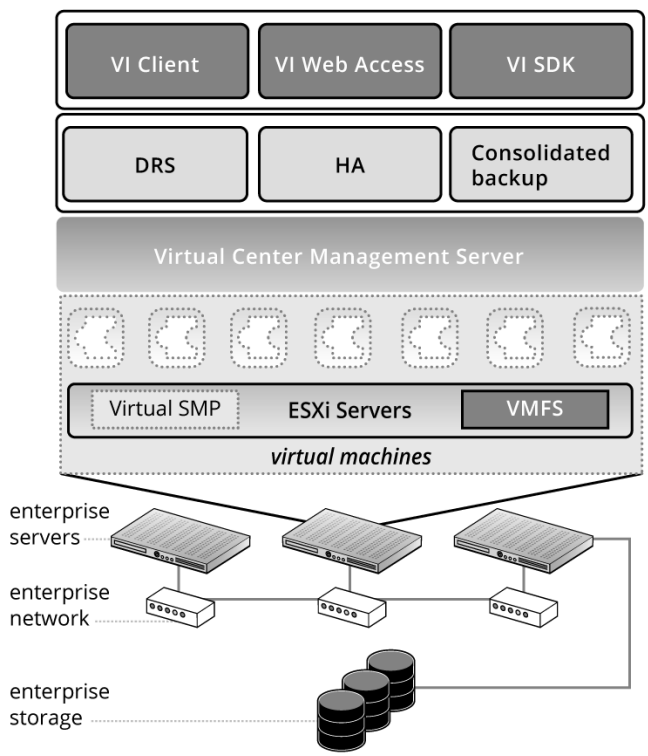
8

Gerenciamento do VMware

Componentes do gerenciamento

- Componentes da arquitetura
- VMotion
- Alta Disponibilidade (HA)
- Balanceamento dinâmico de recursos (DRS)
- Balanceamento dinâmico de recursos de armazenamento (Storage DRS)
- Storage VMotion
- Site Recovery Manager (SRM)

O VMware VSphere possui os componentes principais mostrados na figura 8.1.



Figuras 8.1 Componentes da arquitetura VMware.

Componentes da arquitetura VMware:

- **Servidor VMware ESXi**: camada de virtualização que executa em servidores físicos e possibilita que os principais recursos do servidor sejam provisionados para múltiplas máquinas virtuais.
- **VMware Virtual Center Management Server (vCenter Server)**: permite o gerenciamento de todo o ambiente virtual, automação operacional e otimização de recursos. É executado em um servidor Windows. A partir da versão 2.01 (patch 2) pode ser utilizado em cluster.
- **VMware Update Manager**: automatiza o gerenciamento de atualizações (upgrades) e correções (patches) em servidores ESXi e máquinas virtuais que executam nesses servidores. Importante conferir na documentação os sistemas suportados. É um plugin do vCenter, e o cliente deve ser instalado em máquina Windows. Permite criar regras de atualização e correção, e a sua instalação é recomendada em hardware de 64 bits.
- **VMware VMFS (Virtual Machine File System)**: sistema de arquivos para máquinas virtuais. Possibilita que vários servidores ESXi acessem simultaneamente o sistema de armazenamento da máquina virtual. Cada máquina virtual é encapsulada em um pequeno conjunto de arquivos, e o VMFS é o sistema de arquivos. É possível conectar até 32 servidores ESXi em um volume VMFS. A figura a seguir ilustra o funcionamento do VMFS.

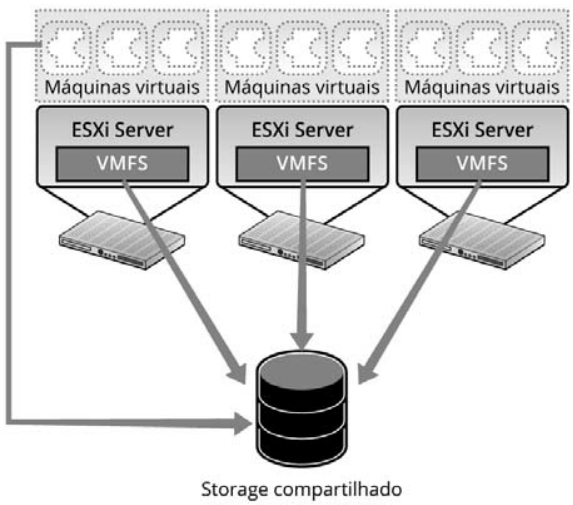


Figura 8.2 Virtual Machine File System (VMFS).

- **Virtual Symetric Multi-Processing (VMware SMP)**: permite que uma única máquina virtual utilize múltiplos processadores simultaneamente.
- **VM**: instância de máquina virtual, ou seja, cada máquina virtual possui seu próprio hardware virtual e Sistema Operacional.
- **VI Client**: interface que possibilita aos usuários e administradores conectarem-se ao servidor.
- **VI Web Access**: interface web para o gerenciamento das máquinas virtuais e console de acesso remoto.

- **VMware Infrastructure SDK**: propicia uma interface padrão para o VMware, permitindo que terceiros desenvolvam soluções para a infraestrutura VMware.
- **Recurso de Balanceamento de Carga (Distributed Resource Scheduler – DRS)**
- **Alta Disponibilidade (High Availability – HA)**

Os três últimos itens serão detalhados ainda neste capítulo.

VMotion

- Movimentação de máquinas virtuais
- Sem impacto no funcionamento da máquina virtual
- Utilizada para alta disponibilidade e balanceamento de recursos

A figura a seguir ilustra o funcionamento do VMotion, um recurso básico para o funcionamento da alta disponibilidade e do balanceamento dinâmico de recursos das soluções de virtualização da VMware. O VMotion permite a migração de máquinas virtuais em tempo real. Exige a utilização de um storage compartilhado por vários servidores, onde a máquina virtual é encapsulada por um conjunto de arquivos armazenados nesse storage. A migração realizada com o VMotion possibilita, por exemplo, que máquinas virtuais sejam transferidas de servidores muito carregados para outros ociosos.

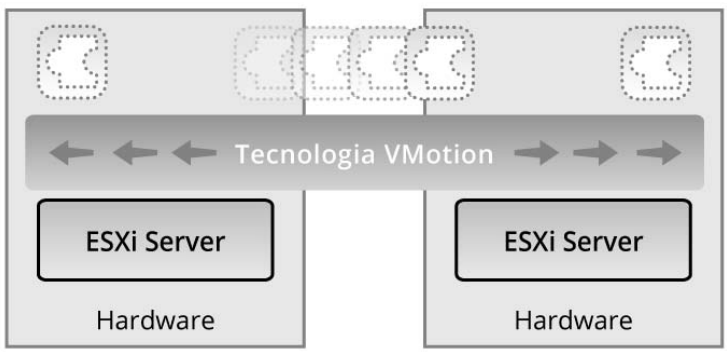


Figura 8.3 VMware VMotion.

Principais características do VMotion:

- Permitir múltiplas e simultâneas migrações concorrentes;
- Agendar migrações;
- Priorizar migrações "a quente" (live) para a utilização de recursos como alta disponibilidade;
- Manter uma trilha de auditoria com detalhes da migração.

Requisitos necessários para a utilização do VMotion:

- Sistema de arquivos VMFS;

- Storage compartilhado em redes do tipo SAN FC, SAN IP ou NAS;
- Rede dedicada gigabit ethernet;
- Servidores de origem e destino com CPUs compatíveis, o que implica utilizar CPUs de um mesmo fabricante e de gerações próximas.

Alta Disponibilidade (HA)

- Disponibilidade do serviço mesmo em caso de falha do host
- Processo automático

A Alta Disponibilidade (High Availability – HA) está disponível para qualquer máquina virtual, independentemente do hardware ou Sistema Operacional utilizado. Se um servidor de produção falha, as máquinas virtuais que estão rodando são automaticamente reiniciadas em outro servidor de produção. A figura a seguir ilustra o uso da alta disponibilidade.

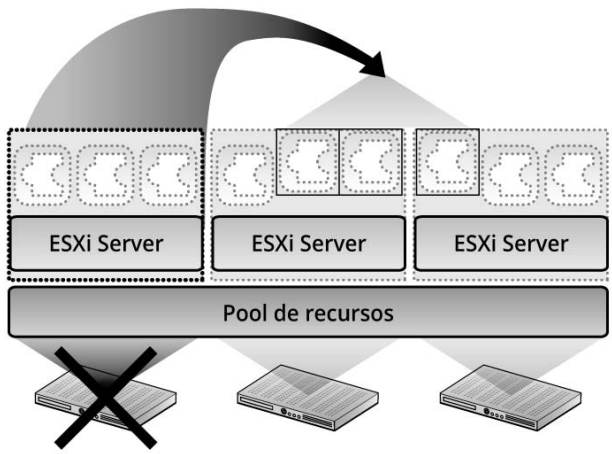


Figura 8.4 Alta Disponibilidade.

Balanceamento dinâmico de recursos (DRS)

- Distribuição automática de MVs com base na carga de trabalho
- Visa a melhorar o desempenho dos serviços

O balanceamento dinâmico de recursos (Distributed Resource Scheduler – DRS) do VMware permite que as máquinas virtuais possam ser redistribuídas para outros servidores mediante algumas regras previamente estabelecidas. Elas visam, normalmente, à melhoria do desempenho.

O DRS monitora dinamicamente a carga de trabalho das máquinas virtuais que estão em execução e a utilização dos recursos dos servidores físicos. Ele também verifica o uso dos recursos considerando políticas predefinidas pelo administrador. E, se for o caso, aloca e distribui dinamicamente no hardware a carga de trabalho das máquinas virtuais.

Com foco na otimização dos recursos de processamento e memória, é prática comum utilizar o DRS em conjunto com recursos de storage – como o Navisphere Quality of Service Manager (NQM), do EMC CLARiiON, que otimiza as operações de E/S. A próxima figura ilustra o uso de DRS.

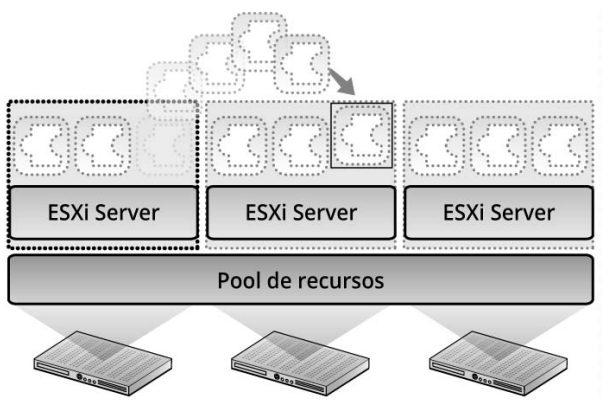


Figura 8.5 Distributed Resource Scheduler (DRS).

Balanceamento dinâmico de recursos de armazenamento (Storage DRS)

- Conceito
- Características
- Requisitos

O VMware vSphere Storage DRS equilibra continuamente a utilização de espaço e a carga de entrada/saída do armazenamento, ao mesmo tempo em que evita o gargalo de recursos para atender aos níveis de serviço dos aplicativos. O uso do Storage DRS permite:

- Implantar e integrar capacidade de armazenamento adicional em um pool de armazenamento;
- Aproveitar a nova capacidade;
- Melhorar os níveis de serviço de todos os aplicativos;
- Aumentar a produtividade do administrador do VMware vSphere, permitindo que monitore e gerencie infraestruturas adicionais.

Storage VMotion

- Conceito
- Características
- Requisitos

O Storage VMotion do VMware possibilita a migração “a quente” de arquivos de discos das máquinas virtuais entre diferentes unidades de storage, sem a necessidade de interromper

os sistemas envolvidos. O objetivo da migração é permitir um melhor desempenho dos acessos aos arquivos – quer seja por movê-los para a máquina onde os processos realizam maior número de acessos ou para reduzir tráfego e latência de rede. A migração pode ser feita empregando os protocolos FC ou iSCSI. Para utilizar o VMware Storage Motion é preciso instalar o VMware vCenter Server e o vCenter Agent.

Site Recovery Manager (SRM)

- Conceito
- Funcionalidades

Site Recovery Manager (SRM) é a solução de automação e gerenciamento de recuperação de desastres integrada ao VMware Virtual Infrastructure, VMware vCenter e ao software de replicação do storage. A próxima figura ilustra o SRM, que permite:

- Gerenciar planos de recuperação de desastres;
- Realizar testes de recuperação de desastres sem interrupção;
- Automatizar o failover e a recuperação;
- Simplificar e automatizar as cargas de trabalho envolvidas na recuperação de desastres;
- Configurar testes de failover e failback;
- Fazer o gerenciamento central de planos de recuperação, integrado ao VirtualCenter;
- Transformar os processos manuais de recuperação em planos de recuperação automatizados;
- Simplificar a integração com a replicação de storage de terceiros.

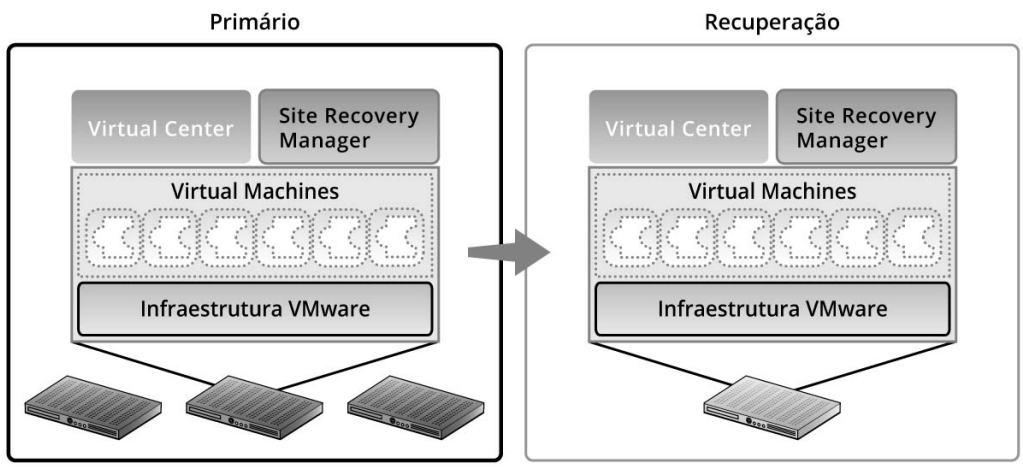


Figura 8.6 VMware Site Recovery Manager.

Infraestrutura de TI

Os principais blocos da infraestrutura de TI para o VMware são: servidores (hosts), redes de storage (SAN ou NAS) e redes locais (LAN). Esses blocos já foram apresentados de forma genérica no capítulo 2. Neste, serão tratados apenas aspectos específicos de infraestrutura relacionados ao VMware.

Arquitetura do servidor

Os servidores são responsáveis pelo processamento no datacenter. Com a virtualização, os fabricantes Intel e AMD fizeram um grande esforço para compensar no hardware (processador) a perda de desempenho ocasionada pelo custo computacional exigido pela virtualização. O resultado disso é que os novos processadores oferecem um bom desempenho, mesmo com o emprego de camadas de virtualização.

Além dos servidores que implementam a virtualização, é preciso considerar a necessidade de servidores ou máquinas adicionais para as tarefas de gerenciamento da infraestrutura de virtualização. Por exemplo, o vCenter Server, necessário para a execução das atividades de gerência do VMware, executa apenas sobre o Sistema Operacional Windows. Portanto, deve-se prever uma máquina com hardware e software apropriados.

> ⓘ Aspecto importante em um projeto de virtualização é que o hardware dos fabricantes precisa estar homologado para as soluções de virtualização disponíveis no mercado.

Processador e memória

O VMware é intensivo em utilização de CPU, portanto:

- Deve sempre ser considerada a aquisição de processadores de maior desempenho, avaliando a opção quadcore;
- No VMware ESXi, as necessidades de processamento das máquinas virtuais são distribuídas pelos diversos núcleos;
- O VMware é intensivo quanto ao consumo de memória:
  - Para o cálculo do tamanho da memória, deve ser considerado o número de máquinas virtuais, a quantidade de memória necessária para cada máquina virtual e uma capacidade adicional para migração;
  - A memória consumida pelo hipervisor varia de acordo com o número de máquinas virtuais e com a memória alocada para cada uma;
  - A memória empregada varia de 384 MB a 32 GB por máquina virtual. Em um cenário “pesado”, o hipervisor pode precisar de 1 GB.

Fabricantes de servidores como Dell, HP e IBM possuem diversas configurações que suportam o hipervisor VMware vSphere ESXi.

Armazenamento (storage)

Uma rede SAN típica utilizada com o VMware é ilustrada a seguir.

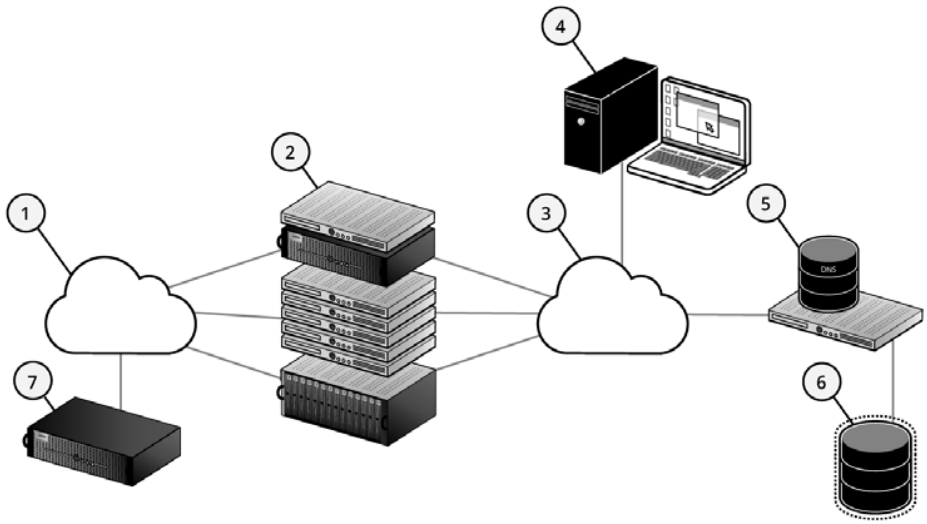


Figura 8.7 Rede SAN com VMware.

Os componentes da SAN podem ser agrupados em:

1. Rede de interconexão da SAN – Fibre Chanel (FC) ou iSCSI;
2. Hospedeiros do hipervisor ESXi;
3. Rede de interconexão da infraestrutura VMware (Ethernet);
4. VI vClient;
5. Servidor vCenter e servidor de licenças;
6. Banco de dados do vCenter;
7. Storage para as máquinas virtuais.

Sistemas de arquivos suportados pelo VMware:

- **Virtual Machine File System (VMFS)**: o VMware ESXi pode utilizar o VMFS em discos locais, volumes iSCSI ou volumes Fibre Channel (FC), criando um diretório para cada máquina virtual. O VMFS é um sistema de arquivos que pode ser usado simultaneamente por vários servidores ESXi.
- **Raw Device Mapping (RDM)**: mecanismo que oferece suporte a vários sistemas de arquivos existentes em um volume. Com ele, as máquinas virtuais têm acesso direto ao disco sem usar o sistema de arquivos VMFS.
- **Network File System (NFS)**: o VMware ESXi permite o emprego do sistema de arquivos NFS convencional. Nesse caso, o VMware ESXi monta um volume NFS, criando um diretório para cada máquina virtual.

A arquitetura de armazenamento do VMware possibilita que discos SCSI sejam vistos por servidores ESXi conectados a um barramento virtual. O disco SCSI virtual é provisionado pelo **datastore**.

Datastore
Container lógico formatado que simplifica a complexidade da rede de storage, podendo ser baseado em VMFS ou NFS.

Um disco virtual dentro de uma máquina virtual é localizado em um ou mais volumes no storage físico e tratado como um volume VMFS ou RDM.

O raw device é um volume lógico usado pela máquina virtual, mas que não está formatado com VMFS ou NFS. O RDM é um arquivo especial que atua como um tipo de proxy para o raw device, mapeando o volume lógico diretamente para a máquina virtual.

A máquina virtual acessa o datastore por meio de comandos SCSI, que permitem acessar o disco virtual. Esses comandos são empacotados de várias formas diferentes, dependendo do protocolo que o servidor ESXi utiliza para se conectar ao dispositivo de storage. O VMFS foi projetado para gerenciar múltiplos acessos e pode coordenar os acessos aos arquivos do disco virtual e aos arquivos internos de informação (metadados). O controle permite limitar o acesso dos servidores ESXi a determinados volumes.

Storage Area Network (SAN)

- Protocolo FC
- Protocolo iSCSI
  - Funcionalidades e limitações do hardware iSCSI
  - Funcionalidades e limitações do software iSCSI

Existem dois padrões dominantes de protocolos utilizados em Storage Area Networks (SAN): Fibre Channel e iSCSI. O princípio de operação desses dois protocolos já foi abordado. Por isso, agora serão comentados apenas alguns aspectos relacionados às limitações de hardware e de software de seu emprego, em especial do iSCSI. São pontos importantes a considerar:

- Capacidade e tempo de resposta para a definição do tipo de E/S. Essa capacidade é calculada a partir do nível de RAID empregado e do volume interno necessário às máquinas virtuais e aplicações;
- O número de hospedeiros e o tipo do equipamento físico de interconexão adotados devem ser considerados na decisão sobre o storage;
- O desempenho do disco e da banda fornecida necessários para E/S.

Principais funcionalidades e limitações de um hardware iSCSI:

- A inicialização do servidor ESXi em SANs iSCSI só é possível com initiator por hardware;

Nem todas as interfaces iSCSI disponíveis no mercado permitem realizar configurações de balanceamento de carga com múltiplas interfaces de rede (NIC teaming). Por exemplo, os modelos QLA4010 só possibilitam operação em **failover**.

- Oferece suporte para os mecanismos VMotion, VMware HA e VMware DRS, mas não para o VMware Consolidated Backup nem para o Microsoft Cluster Server;
- Possui suporte para RDM.

Failover

Ação de uma máquina assumir os serviços executados por outra.

Principais funcionalidades e limitações de um software iSCSI:

- Não oferece suporte para inicialização do servidor ESXi;
- O initiator por software suporta somente uma única interface de rede (denominada vmhba40);
- Não é possível executar balanceamento de carga por software com configurações de múltiplas interfaces de rede (NIC teaming). Essa configuração permite apenas failover;
- Suporte para os mecanismos VMotion, VMware HA e VMware DRS; não para Microsoft Cluster Server;
- Possui suporte para RDM.

O sistema de armazenamento baseado em Internet SCSI (iSCSI) simplifica e reduz o custo inicial de uma solução de storage, e mesmo o TCO, quando comparado à alternativa fundamentada em Fibre Channel (FC). Com a chegada do padrão 10 Gbit/s Ethernet, acredita-se que o padrão iSCSI ganhará ainda mais espaço no mercado de redes de armazenamento. A rede iSCSI também trata de aplicações que utilizam blocos de dados como elementos padrão de transferência de E/S.

A decisão de empregar uma rede SAN baseada no protocolo FC ou iSCSI com a camada de virtualização deve considerar a homologação dos produtos e as funcionalidades requeridas. Existe uma interdependência entre o dispositivo de armazenamento e a virtualização – boa parte das funcionalidades obtidas com a virtualização depende da consolidação do armazenamento em um único dispositivo. As redes de storage do tipo SAN e NAS são suportadas pelos softwares de virtualização, permitindo a otimização de recursos.

A decisão entre utilizar um initiator em hardware ou software é essencialmente um compromisso de custo *versus* desempenho. Ao usar a ferramenta I/O Meter como gerador de carga de trabalho, a VMware realizou um estudo para avaliar a diferença de desempenho entre essas duas opções. A mesma carga de trabalho foi colocada em servidores idênticos com recursos isolados de storage. O initiator por hardware disponibilizou 150% a mais de vazão (troughput) e exigiu apenas 25% dos recursos de processamento usados pelo initiator por software. A diferença de preço entre um adaptador Gigabit e um adaptador iSCSI é da ordem de US$ 400.

Network-Attached Storage (NAS)

Servidor de arquivos conectado a uma rede de computadores que fornece acesso aos dados para usuários de várias plataformas (Windows, Linux). Seu acesso é feito no nível de arquivo e não em bloco, como em SAN. O acesso aos arquivos é feito por meio de protocolos como Network File System (NFS), Server Message Block/Common Internet File System (SMB/CIFS) ou Apple Filing Protocol (AFP).

Rede Local – Local Area Network (LAN)

- Opções
- Switches virtuais
- NICs virtuais

As opções de rede são providas pelo ESXi e gerenciadas pelo VMware Virtual Center. Com a rede virtual, é possível criar redes virtuais dentro de um servidor ESXi ou por meio de múltiplos servidores. Os dois componentes de uma rede virtual são:

- **Switches virtuais**: é possível criar até 248 switches virtuais em um servidor ESXi. As funcionalidades de um switch virtual são as mesmas de um de switch convencional de camada 2, suportando inclusive VLANs com controle por portas. O protocolo spanning tree não é necessário, pois a topologia de rede permitida é de apenas um nível.
- **Interfaces virtuais de rede**: as Network Interconnection Cards (NICs) permitem conectar as máquinas virtuais entre si ao console e à rede externa, como ilustra a próxima figura**.** As placas de redes físicas funcionam como uplinks das portas do switch virtual para o switch físico. Com o uso da configuração NIC Teaming, é possível conectar um switch virtual a múltiplos adaptadores Ethernet. As máquinas virtuais podem ser configuradas com uma ou mais interfaces virtuais de rede, cada uma com seu endereço IP e MAC.

É importante salientar que as portas empregadas por interfaces de redes são determinantes para o desempenho global da solução. Recomenda-se a utilização de: uma interface de rede exclusiva para o gerenciamento; pelo menos duas para a implementação de uma solução com alta disponibilidade; interfaces dedicadas para as atividades de migração (uso do VMotion); duas ou mais interfaces redundantes para iSCSI; e, finalmente, duas ou mais interfaces redundantes para as máquinas virtuais.

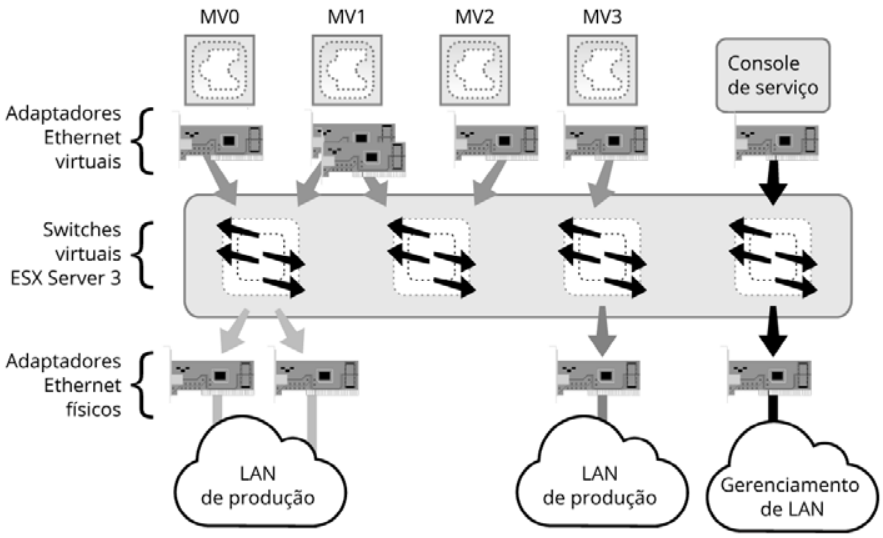


Figura 8.8 Virtualização da rede com o VMware.

9

Nuvem

- Computação em nuvem
- Principais benefícios
- Modelos de computação em nuvem
- Disponibilidade de Serviços em nuvem
- Estratégia de TI em nuvem

Computação em nuvem

- A computação em nuvem é:
  - A entrega de um serviço ou
  - A entrega de um recurso na rede
- Promove disponibilidade de recursos e é composto por cinco características:
  - Self-service
  - Acesso à rede de forma ampla
  - Pool de recursos
  - Elasticidade
  - Serviço medido

A computação em nuvem é a entrega de um serviço ou um recurso na rede. Mais especificamente, refere-se aos recursos de TI e serviços que são abstraídos da infraestrutura e fornecidos "sobre demanda" e "em escala" para um ambiente de compartilhamento com vários usuários. O National Institute for Standards and Technology (NIST) descreve computação em nuvem como um modelo que permite o acesso à rede por meio de um conjunto compartilhado de recursos computacionais configuráveis (por exemplo, redes, servidores, armazenamento, aplicativos e serviços), que podem ser rapidamente provisionados e ofertados com o mínimo esforço de gestão e interação com o provedor do serviço. Esse modelo de nuvem promove disponibilidade de recursos e é composto por cinco características essenciais:

- **Self-service:** o cliente pode, unilateralmente, provisionar as capacidades de computação, tais como capacidade do servidor e armazenamento de dados, automaticamente conforme a necessidade e sem precisar de interação com o provedor de cada serviço.
- **Acesso à rede de forma ampla:** os recursos estão disponíveis por meio da rede e são acessados por meio de mecanismos-padrão, que promovem o uso de plataformas heterogêneas (por exemplo, telefones celulares, laptops e PDAs).

- **Pool de recursos:** os recursos do provedor do serviço são reunidos para atender vários usuários por um modelo “multi-inquilino”, com diferentes recursos físicos e virtuais, atribuídos dinamicamente e novamente de acordo com a demanda do cliente. Há um senso de independência local em que o usuário geralmente não tem controle ou conhecimento sobre a localização exata dos recursos disponibilizados. Exemplos de recursos incluem o armazenamento, processamento, memória, banda de rede e máquinas virtuais.
- **Elasticidade:** as capacidades podem ser rápida e elasticamente provisionadas e automaticamente liberadas ao usuário. Para ele, as capacidades disponíveis ao armazenamento de dados, muitas vezes, parecem ser ilimitadas e podem ser adquiridas em qualquer quantidade e momento.
- **Serviço medido:** o sistema controla e otimiza automaticamente o uso dos recursos, alavancando a capacidade de medição em algum nível de abstração apropriado para cada tipo de serviço. O uso de recursos pode ser monitorado, controlado e comunicado; proporcionando transparência para o provedor e usuário do serviço utilizado.

A computação em nuvem é o resultado de avanços tecnológicos diversos, incluindo:

- Redes de alta velocidade;
- Capacidades de virtualização;
- Comodização do hardware dos servidores;
- Software de fonte aberta (como Linux, Apache e Hadoop), que reduziu o custo do software para datacenters;
- Adoção de padrões abertos da Web 2.0, o que tornou o desenvolvimento de aplicações na nuvem muito mais fáceis e rápidas.

O Gartner Hype Cycle publica a cada ano como estão os aspectos da computação em nuvem. Também identifica as aplicações/tecnologias com adoção significativa e as que são razoavelmente maduras.

Principais benefícios

- A computação em nuvem pode ser definida como o uso de clusters de recursos computacionais trazendo os benefícios:
  - Escalabilidade
  - Eficiência
  - Contenção de custos
  - Flexibilidade
  - Disponibilidade
  - Resiliência

A computação em nuvem muda fundamentalmente a maneira como os serviços de TI são entregues e como os dados são armazenados e protegidos. As organizações podem utilizar a

computação em nuvem para satisfazer as suas necessidades de TI adotando um modelo flexível, sob demanda e rapidamente escalável, que não requer nenhum fornecimento de recursos dedicados. A computação em nuvem, nesse contexto, é definida como o uso de clusters de recursos computacionais padronizados e gerenciados de uma forma automatizada para atender a uma determinada necessidade ou negócio, pois:

- Não possui uma infraestrutura dedicada, reduzindo custos com investimento;
- Consome recursos como serviços, pagando-se apenas pelos recursos que são utilizados, como um utilitário ou preço da assinatura;
- Não exige qualquer conhecimento dos detalhes sobre a prestação do serviço;
- Permite atender às necessidades de recursos dinâmicos, escaláveis ou potencialmente imprevisíveis.

Ao considerar que o grau do benefício vai depender do modelo de nuvem adotado, a transição para os serviços em nuvem pode oferecer os seguintes benefícios aos membros associados:

Benefícios	Detalhes
Escalabilidade	Maior agilidade de resposta pode beneficiar a prestação de serviços e atender às necessidades dos cidadãos, universidades, institutos e fornecedores. Opção de escalabilidade é oferecida sem os compromissos financeiros necessários para a aquisição de infraestrutura e manutenção. Provisionamento e implementação são realizados sob demanda, permitindo picos de utilização e reduzindo o tempo para implementar novos serviços. Entidades, no entanto, precisam estar cientes de que, a transição de sistemas legados, migração de dados e gestão da mudança podem retardar o processo de adoção da computação em nuvem.

Eficiência	A realocação de atividades das áreas operacionais de TI oferece oportunidade para as instituições se concentrarem em: Pesquisa e desenvolvimento, incluindo novas e inovadoras aplicações que favorecem os negócios e crescimento do produto e melhora da prestação de serviços; Criação de novas soluções que não eram tecnicamente e/ou economicamente viáveis sem o uso de serviços em nuvem; Prototipagem e validação de mercado com novas abordagens de forma muito mais rápida e menos dispendiosa; Racionalização de sistemas legados.
Contenção de custos	Alterações do modelo de custo das instituições podem ser modificados da seguinte forma: **Transferência de custos** De CAPEX para OPEX. Os serviços e armazenamento tornam-se disponíveis sob demanda, não havendo assim necessidade de reserva de recursos financeiros para a compra de infraestrutura e/ou manutenção, pois esses serão pagos como serviços; **Redução dos custos operacionais** Redução do consumo de energia e das despesas com a gestão de sistemas de TI; Diminuição do custo e complexidade na execução das tarefas operacionais e na solução de problemas que exigem grande capacidade de processamento; Redução de custos de suporte e manutenção, considerando a transição de sistemas legados para novos sistemas.
Flexibilidade	Menor tempo de set-up, pois a computação em nuvem se torna funcional muito mais rápida do que outros sistemas; A possibilidade de acessar uma tecnologia mais recente, que é atualizada constantemente e automaticamente.
Disponibilidade	Maior flexibilidade e disponibilidade de informação, pois é possível realizar o acesso à informação de qualquer lugar do mundo, sendo somente necessária uma conexão com a internet.
Resiliência	Potencial de falha em um ambiente de computação altamente reduzido. A falha de um nó de um sistema em um ambiente de nuvem não terá impacto na disponibilidade da informação e reduzirá o risco de inatividade perceptível.

Consolidação de um Datacenter

- Consumo reduzido de manutenção de hardware, diminuição das despesas com servidores por meio de uma infraestrutura mais eficiente e melhor aproveitada;
- Menores custos para investimento por meio da padronização e compartilhamento de ativos;
- Redução do investimento em uma infraestrutura de computação com aluguel baseado em computação utilitária e tempo de processamento;
- Diminuição dos custos na aquisição de equipamentos, em função da adoção do processo de virtualização de servidores, gerando maior taxa de utilização;
- Menores custos de recuperação de dados, considerando uma maior facilidade para a recuperação de desastres.

Modelos de computação em nuvem

- Nuvem pública
- Nuvem privada ou interna
- Nuvem de comunidade
- Nuvem híbrida
- Colaboração em nuvem

Existem cinco modelos de implantação de nuvem: pública, privada, comunidade, híbrida e colaborativa.

- Nuvem pública

Está disponível para todas as pessoas que acessam a internet. As nuvens públicas são abertas para todos os usuários e beneficiam uma base de clientes situados em ampla área geográfica.

- Nuvem privada ou interna

Os dados armazenados são acessados somente por meio de uma conta autorizada ou unidade federada. As nuvens privadas podem servir a uma única instituição ou a um grupo restrito de instituições.

- Nuvem de comunidade

Nesse caso, os dados armazenados podem ser acessados por um grupo específico, que tem autoridade para partilhar e usar a informação. Apenas os seus membros têm acesso aos dados e/ou para realizar a transferência de dados entre eles.

- Nuvem híbrida

É a combinação de nuvem pública e nuvem privada. Qualquer usuário pode ver apenas uma quantidade específica de dados; somente pessoas autorizadas têm acesso a todos os dados

disponíveis na nuvem. Como, a exemplo de um portal, no qual um funcionário vê todas as informações, e um cliente, apenas uma informação generalista.

- Colaboração em nuvem

É uma forma emergente de partilha eco-autoria de arquivos pelo uso de computação em nuvem, na qual os documentos são enviados para "a nuvem central" com objetivo de armazenamento e podem ser acessados por qualquer usuário ou instituição. Novas tecnologias de colaboração em nuvem têm permitido aos usuários carregar, comentar e colaborar em documentos e até mesmo alterá-los.

Serviços em nuvem

- Abordagens de tipos de serviços:
  - Armazenamento como Serviço – Storage as a Service (SaaS)
  - Infraestrutura como Serviço – Infrastructure as a Service (IaaS)
  - Plataforma como Serviço – Platform as a Service (PaaS)
  - Software como Serviço – Software as a Service (SaaS)
  - Tudo como Serviço – Everything as a Service (EaaS)

O serviço de nuvem pode ser usado para fornecimento com várias abordagens:

- **Armazenamento como Serviço – Storage as a Service (SaaS)**

É um modelo no qual o provedor de serviço aluga espaço em sua infraestrutura de armazenamento para as instituições ou usuários. Os fornecedores de SaaS estão disponibilizando aplicações de armazenamento secundário como uma maneira conveniente para gerenciamento de backup. A principal vantagem do serviço SaaS para uma instituição está na economia de custos – em pessoal, hardware e espaço de armazenamento físico. Por exemplo, no lugar de se manter uma biblioteca de fitas e organizar o arquivo de fitas externo, um administrador de rede pode utilizar o serviço para realizar o backup dos dados, definindo quais devem ser copiados.

A instituição firmará um contrato com o provedor de serviços, no qual será identificado o nível de serviço (SLA). Em caso do registro de corrompimento ou perda de dados, o administrador da rede deverá contatar o provedor de serviços SaaS e solicitar uma cópia dos dados.

Geralmente visto como uma boa alternativa para entidades que não têm orçamento suficiente e/ou pessoal técnico para implementar e manter sua infraestrutura própria. O SaaS também está sendo promovido como uma forma de mitigar os riscos na recuperação de desastres para todas as empresas. Fornece retenção a longo prazo para os dados e melhora tanto a continuidade do negócio quanto sua disponibilidade.

Infraestrutura como Serviço – Infrastructure as a Service (IaaS)

Nesse modelo, a instituição terceiriza o equipamento usado para apoiar as operações, incluindo armazenamento, hardware, servidores e componentes de rede. O prestador de serviços possui o equipamento e é responsável pela habitação, execução e manutenção dele. O pagamento dos serviços pela instituição é feito com base na utilização dos recursos. O IaaS é também conhecido como Hardware as a Service (Haas).

Plataforma como Serviço – Platform as a Service (PaaS)

É uma maneira de alugar hardware, sistemas operacionais, armazenamento e capacidade de rede pela internet. O modelo de prestação de serviço permitirá ao cliente alugar servidores virtualizados e serviços associados para a execução de aplicativos existentes ou desenvolver e testar novos. A PaaS é um desdobramento do Software como Serviço (SaaS), um modelo de distribuição de software em que as aplicações hospedadas são disponibilizadas aos usuários pela internet. O modelo de serviço PaaS apresenta várias vantagens, principalmente para os desenvolvedores. Com o PaaS, recursos do Sistema Operacional podem ser alterados e atualizados com frequência. Equipes geograficamente dispersas podem trabalhar juntas em projetos de desenvolvimento de software. Os serviços também são obtidos de diversas fontes que cruzam fronteiras internacionais. Diminui os custos iniciais e em regime pelo uso de serviços de infraestrutura de um único fornecedor, em vez de manter múltiplas instalações que, muitas vezes, desempenham funções duplicadas ou sofrem com problemas de incompatibilidade. As despesas gerais também podem ser minimizadas com a unificação dos esforços de desenvolvimento de programação.

Software como Serviço – Software as a Service (SaaS)

O SaaS é um modelo de distribuição de software em que os aplicativos são hospedados em um provedor ou fornecedor de serviço e disponibilizados aos usuários por meio de uma rede, geralmente a internet. O SaaS está se tornando um modelo de entrega predominante, assim como as tecnologias subjacentes que suportam serviços web e arquitetura orientada a serviços (SOA). Novas abordagens de desenvolvimento, tais como Ajax, tornam-se cada vez mais populares. Enquanto isso, os serviços de banda larga ficam mais disponíveis e apoiam o acesso de usuários em todo o mundo. Vantagens do modelo SaaS:

- Facilita a administração;
- Permite realizar atualizações automáticas e gerenciamento de correções (patches);
- Assegura a compatibilidade, permitindo que todos os usuários tenham a mesma versão do software instalado, o que favorece a colaboração;
- Oferece acessibilidade global.

O modelo tradicional de distribuição de software, no qual o ele é comprado e instalado em computadores pessoais, é por vezes referido como o “software como um produto”.

- **Tudo como Serviço – Everything as a Service (EaaS)**

Tudo como um serviço (XaaS) visa a tornar vários recursos disponíveis para utilização. Isso representa mais um passo lógico em computação em nuvem após a aquisição do software, dos ambientes de execução, do hardware e do trabalho humano como um serviço.

Estratégia de TI em nuvem

- Necessidade estratégica:
  - Melhorar a prestação de serviços
  - Melhorar o acesso aos serviços
  - Aumentar a eficiência da prestação de serviços

A necessidade de uma estratégia de TI que suporte os objetivos principais dos setores público e privado, proporcionando excelência administrativa, demanda:

• Melhorar a prestação de serviços público/privado;

• Melhorar o acesso aos serviços público/privado;

• Aumentar a eficiência da prestação de serviços.

É possível utilizar várias estratégicas de forma global para o TI, como:

- **Rede do setor público e privado:** racionalizar e padronizar para criar uma "rede de redes", permitindo comunicações seguras e móveis para uma maior capacidade a um preço inferior.
- **Racionalizar a propriedade de TI**: usar a computação em nuvem para aumentar a capacidade e a segurança, reduzir custos e acelerar a velocidade de implantação.
- **Datacenters:** racionalização dos datacenters para diminuir custos e aumentar a resiliência e a capacidade.
- **Serviços compartilhados:** continuamente mover sistemas para a nuvem e usá-los em atividades comuns de serviços compartilhados fornecidos pela nuvem.
- **Arquitetura e padrões:** criar um ambiente que permita a muitos fornecedores trabalharem juntos, cooperarem e interagirem de maneira segura, transparente e eficiente.
- **Software livre, padrões abertos e estratégia de reutilização:** nivelamento do jogo para a aquisição, com maior reutilização das ferramentas existentes. Menor aquisição e maior inovação, tudo por um custo menor.
- **TI verde:** cumprindo metas de sustentabilidade, os departamentos de TI se posicionam de forma mais eficiente e com preço inferior.
- **Segurança da Informação:** proteger os dados (cidadãos e empresas) de quaisquer danos, seja acidental ou mal-intencionado.
- **Profissionalização da TI:** melhorar as capacidades, conhecimentos, habilidades e experiência dos envolvidos no domínio das TIs.

- **Entrega de projeto confiável:** usar gestão de carteiras e de benefícios de ativos para garantir que o governo comprometa-se com os projetos certos e da maneira correta.
- **Fornecimento de gestão:** trabalhar em conjunto para obter o máximo de valor dos fornecedores, tanto para as organizações individuais quanto para o setor público.
- **Alinhamento e coordenação internacional:** assegurar que os tratados e diretrizes internacionais reflitam nos requisitos nacionais.

10

Consolidação do conhecimento

- Revisão das soluções

Neste capítulo, faremos uma consolidação do conhecimento adquirido durante o curso. O objetivo é mostrar as diferenças e similaridades entre os diversos softwares apresentados, tanto os hipervisores quanto os de gerenciamento, além do hardware envolvido em ambientes virtualizados.

Como vimos, os fabricantes de hardware estão bastante focados em facilitar o uso da virtualização, desenvolvendo desde chips de CPU até servidores de discos (storage) voltados especificamente para a utilização em ambientes virtualizados de datacenter. Vamos também sugerir procedimentos para a implementação de melhores práticas, visando a otimizar o desempenho dos ambientes virtualizados.

Software

- Hipervisores
  - Microsoft Hyper-V
  - VMware vSphere
  - Citrix XenServer
- Gerenciamento
  - vCenter
  - XenCenter
  - Hyper-V SCVMM

Hipervisores

Quando comparamos os diversos hipervisores disponíveis atualmente, percebemos que todos têm facilidades semelhantes. Afinal de contas, as funções de um hipervisor são basicamente disparar máquinas virtuais e alocar recursos para cada uma delas. É importante notar que, apesar disso, cada um tem as suas particularidades, seja na arquitetura, nos moldes de licenciamento, no uso das facilidades, nos preços etc.

Microsoft Hyper-V

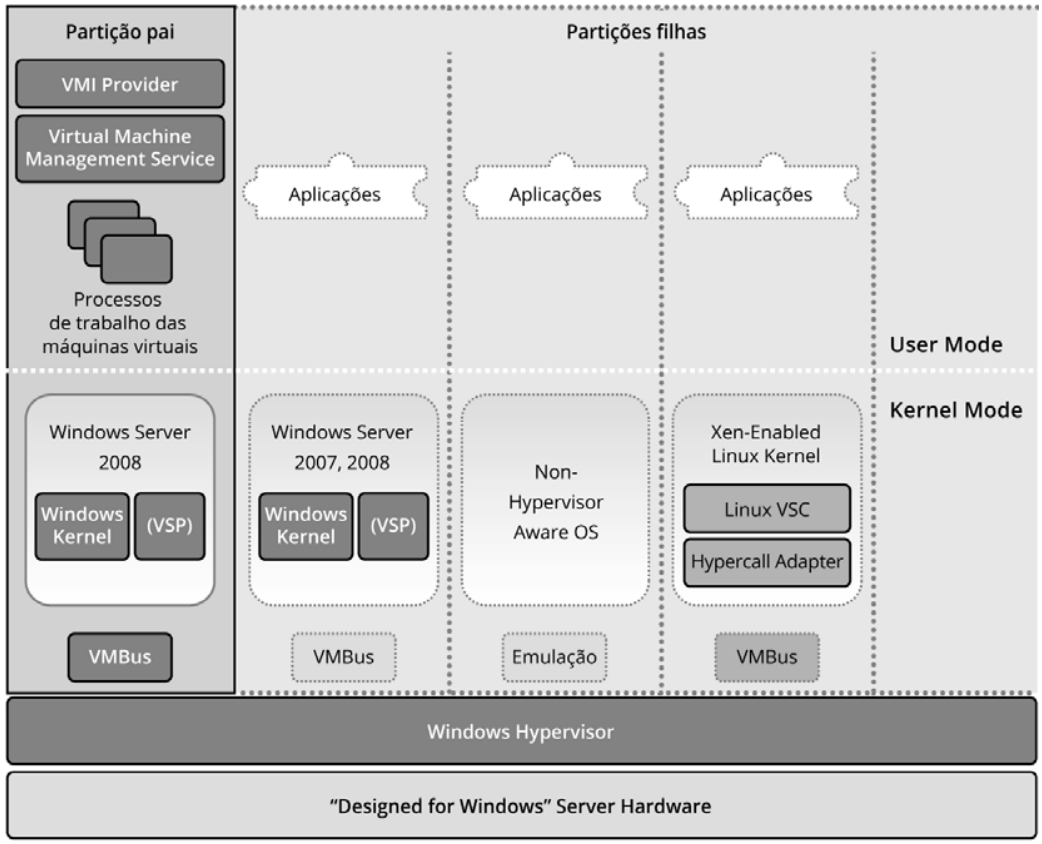


Figura 10.1 Arquitetura do Hyper-V.

Como vimos anteriormente, a variada oferta de edições (Standalone, Standard, Enterprise ou Datacenter) permite que cada organização possa escolher a que melhor resolve o tripé necessidade/preço/desempenho. Para o melhor aproveitamento dos recursos, a Microsoft recomenda para as máquinas virtuais o uso de versões Windows preparadas para virtualização, a saber: Windows 7, Vista e Server 2008.

Dispositivos sintéticos

São os novos dispositivos de alto desempenho suportados pelo Hyper-V.

No ambiente de rede, sempre que possível, deve-se escolher para as máquinas virtuais os **dispositivos sintéticos** em lugar dos emulados. Na interface de configuração de rede para a máquina virtual, essas denominações não aparecem, mas as associações são:

- Legacy Network Adapter – emuladas;
- Network Adapter – sintéticas.

O Hyper-V precisa de um número significativo de placas de rede (Network Interface Card – NIC). Como regra geral, sugere-se usar duas placas quad-port para complementar as duas interfaces de rede que vêm normalmente nos racks. Existem duas opções de discos para as máquinas virtuais: dinâmicos e fixos.

- **Discos dinâmicos**: permitem um melhor uso do espaço disponível no storage. Se for criado um disco dinâmico de 20 GB, a utilização efetiva pode ser de apenas alguns MBs, e o crescimento se dá na medida da necessidade.
- **Discos fixos**: alocam todo o tamanho especificado.

  Nesse caso, o ganho de espaço no storage pelo uso de discos dinâmicos acarreta uma queda no desempenho. A sugestão é usar discos fixos para máquinas de produção e discos dinâmicos para testes e desenvolvimento.

Alguns administradores recomendam o uso do formato pass-thru (PTD), que dá ao Hyper-V a capacidade de apresentar LUNs/volumes diretamente às máquinas virtuais.

Outro ponto a se considerar é o uso de programas antivírus que podem causar degradação do desempenho. Caso as regras de segurança obriguem o uso do antivírus, deve-se então eliminar alguns diretórios de configuração de máquinas virtuais, como:

- C:*ProgramData**Microsoft**Windows**Hyper-V*;
- C:*Users**Public**Documents**Hyper-V**VirtualHard Disks*;
- Diretórios de snapshots e executáveis como Virtual Machine Management Service Process (Vmms.EXE) e Virtual Machine Worker process (Vmwp.EXE).

⚠ O Hyper-V em produção não deve ser usado como uma workstation. Também não deve permitir o uso da rede de gerenciamento para acessar a internet.

VMware ESX/ESXi

- Melhorar a confiabilidade e a segurança
- Reduzir o número de correções
- Simplificar a distribuição e a configuração
- Reduzir o overhead de gerenciamento

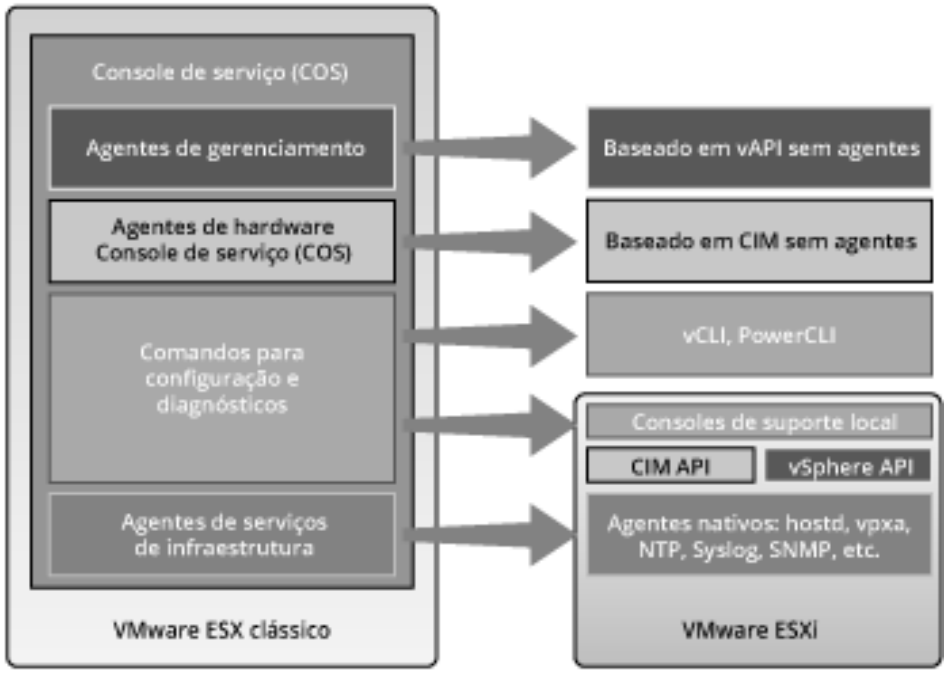


Figura 10.2 VMware: ESX *vs* ESXi.

A principal diferença de arquitetura entre ESX e ESXi está no console de serviço. Os dois hipervisores mudam também no tamanho, que permite ao ESXi ser iniciado a partir de um pendrive. Essa diferença é da ordem de MB para o ESXi, enquanto o ESX chega a GB. O ESX Classic (mostrado na figura anterior) inclui um console de serviço no vmkernel (virtualization kernel), fazendo com que ele fique muito maior que o ESXi. Tradicionalmente, a VMware distribui seus agentes de gerenciamento nesse console junto com agentes de infraestrutura (serviço de nomes, de tempo, de log etc.). Também permite que administradores usem o console para configurar, diagnosticar problemas e rodar scripts. A próxima figura ilustra a arquitetura em um servidor ESX. Apresenta os mesmos níveis de desempenho e escalabilidade do servidor ESX, segundo a VMware. Tanto a versão ESXi como a versão ESX Server são construídas a partir de um núcleo Linux. Na versão ESXi, o Linux Service Console foi excluído, o que impossibilita se conectar diretamente no console da máquina que executa o hipervisor. A perda de funcionalidade ocorrida pela retirada do Service Console é compensada pelo acréscimo de comandos remotos de interface de linha.

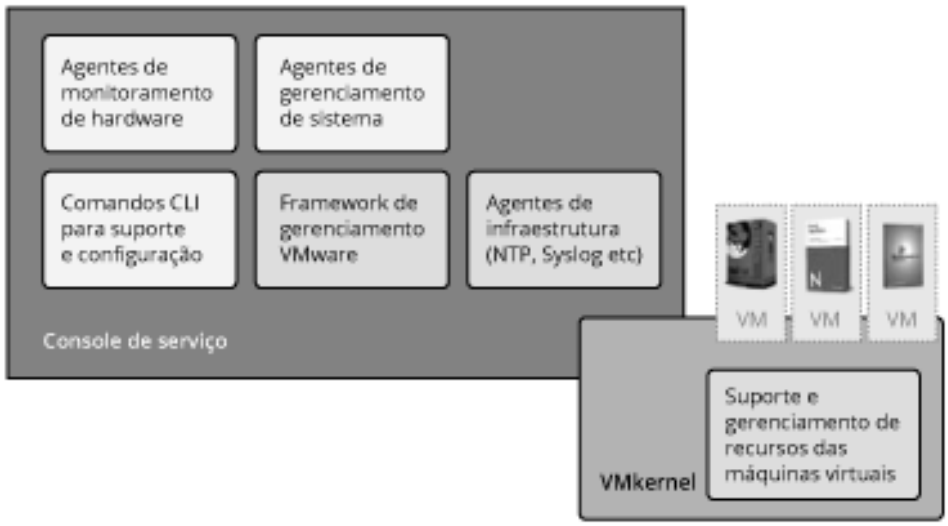


Figura 10.3 Arquitetura ESX Classic.

Vejamos agora a arquitetura ESXi (figura 9.4), na qual o console foi substituído por um Shell Posix (baseado em **Busybox**), trazendo com isso uma atuação mais "Linux like", porém com menos funcionalidades e, possivelmente, menos segurança. Nessa arquitetura, os agentes rodam diretamente no vmkernel. Os módulos de terceiros, como drivers e monitores de hardware, só são permitidos quando assinados digitalmente pela VMware.

Busybox

Combina pequenas versões de vários utilitários Unix em um único e compacto executável. É considerado o "canivete suíço" dos Linux embarcados.

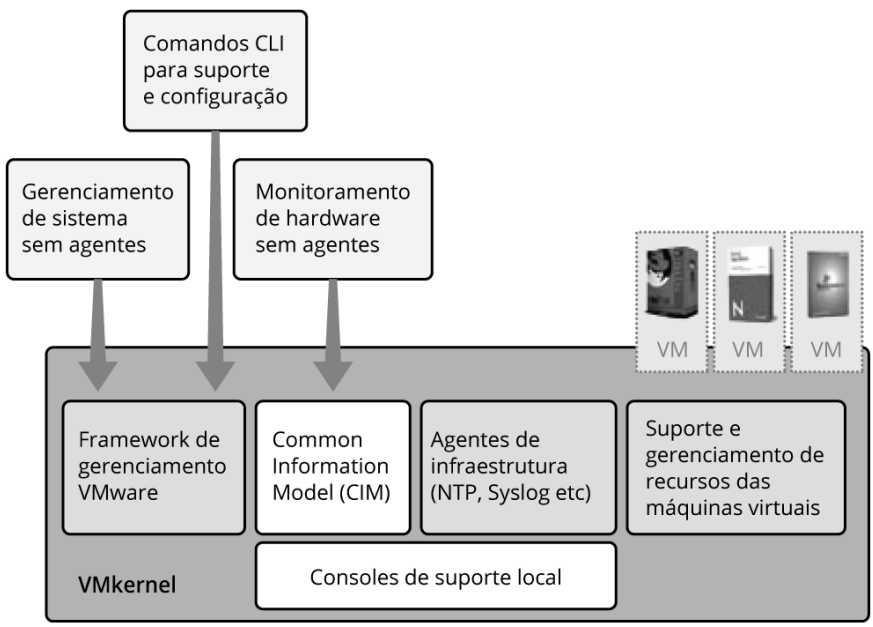


Figura 10.4 Arquitetura ESXi.

XenServer

A Citrix Systems tem uma história de desenvolvimento de sistemas voltados para distribuição de software em ambientes cliente/servidor. Por isso, existe uma preocupação com os desktops virtuais, sendo disponibilizado um conjunto de mecanismos para ajuste e otimização, específico para esse workload.

Um dos problemas conhecidos da virtualização de desktops é o consumo de recursos mesmo quando não há nenhum usuário conectado. No XenServer 5.6, foi introduzido o recurso chamado de Dynamic Memory Control (DMC), que tem por finalidade eliminar a necessidade de reservar memória física antes mesmo de ligar (power on) o desktop virtual. Ele pode resolver, em parte, os problemas de falta de memória, desempenho e número de máquinas virtuais; ou aumentamos o número de máquinas virtuais com perda de desempenho. O problema é a dificuldade de saber quando uma máquina virtual precisa de mais memória, a quantidade de memória de que necessita e por quanto tempo.

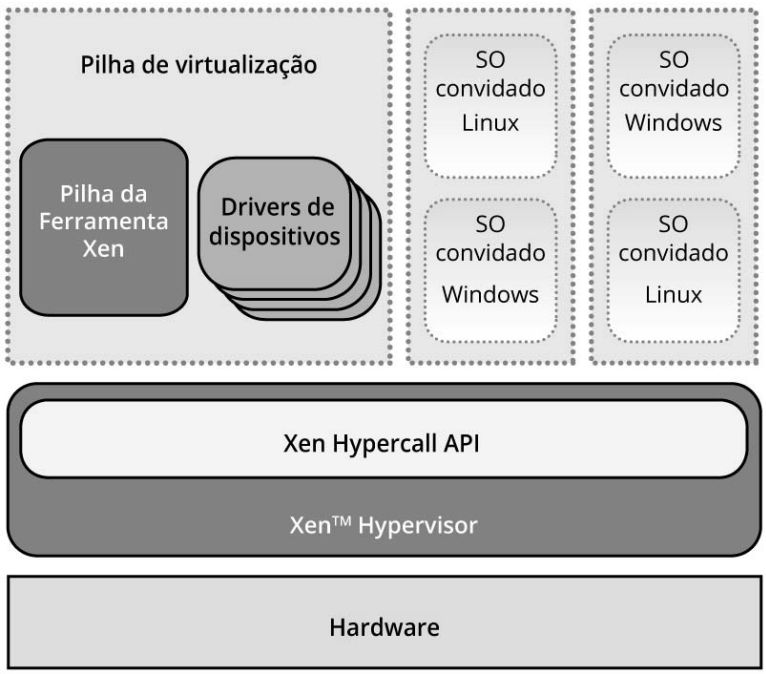


Figura 10.5 Arquitetura XenServer.

Como vimos, o Dom0 tem a função de fazer a comunicação entre as máquinas virtuais (DomU) e o mundo real. Como toda comunicação, obrigatoriamente, tem que passar pelo Dom0, um gargalo pode ser gerado caso o Dom0 não esteja configurado corretamente. Cada DomU usa 6 MB de RAM para ligar, o que não é muito, mas se levarmos em conta que, tipicamente, o Dom0 recebe 752 MB e utiliza 400 MB, sobram apenas 352 MB e, portanto, quando chegar a 60 máquinas virtuais, o pool de memória estará esgotado. Essa é uma situação comum em hipervisores. Como o ESX reserva memória para o gerenciamento, naturalmente o Hyper-V também precisa de uma boa quantidade de memória. A alocação de memória no XenServer é mais específica, e um planejamento malfeito pode até fazer com que o XenServer pare de funcionar. Se o problema de desempenho for de E/S, é recomendado

monitorar a rede e a atividade de disco, por meio de comandos como *iostat, hdharm, tcpdump, netstat* e *ifconfig*. Shell scripts podem ser construídos para facilitar o levantamento.

Esses dados foram tirados de *Hypervisor management and optimization*, de Mike Laverick, publicado no site SearchServerVirtualization.com.

Soluções de gerenciamento

Os sistemas de gerenciamento de plataformas virtualizadas são ferramentas indispensáveis para a administração desses ambientes. No decorrer do curso, foram apresentadas as ferramentas de gerenciamento, quando tivemos a oportunidade de instalá-las e usá-las. Os recursos de todas são semelhantes e, como cada uma é desenvolvida para um determinado hipervisor, não é possível testar todas em ambientes diferentes. A exceção é o XenCenter, disponibilizado para Xen e para Hyper-V.

É importante lembrar que diversas empresas oferecem de forma gratuita ferramentas de gerenciamento para as várias plataformas. O trabalho, portanto, é todo de pesquisa e testes para a escolha das ferramentas. A tendência é a automação de todo o processo para os softwares de gerenciamento, no intuito de manter as plataformas virtualizadas funcionais na maior parte do tempo – tomando conta dos recursos disponíveis e distribuindo as máquinas virtuais, de servidores mais carregados para aqueles com mais disponibilidade; fazendo do balanceamento dinâmico de recursos o carro-chefe do ambiente.

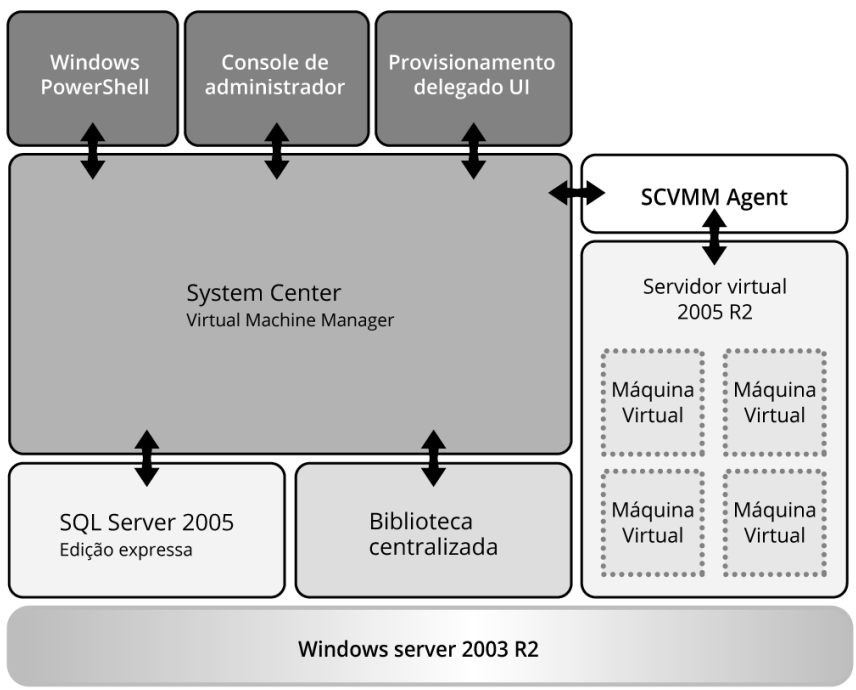


Figura 10.6 Solução de gerenciamento Hyper-V.

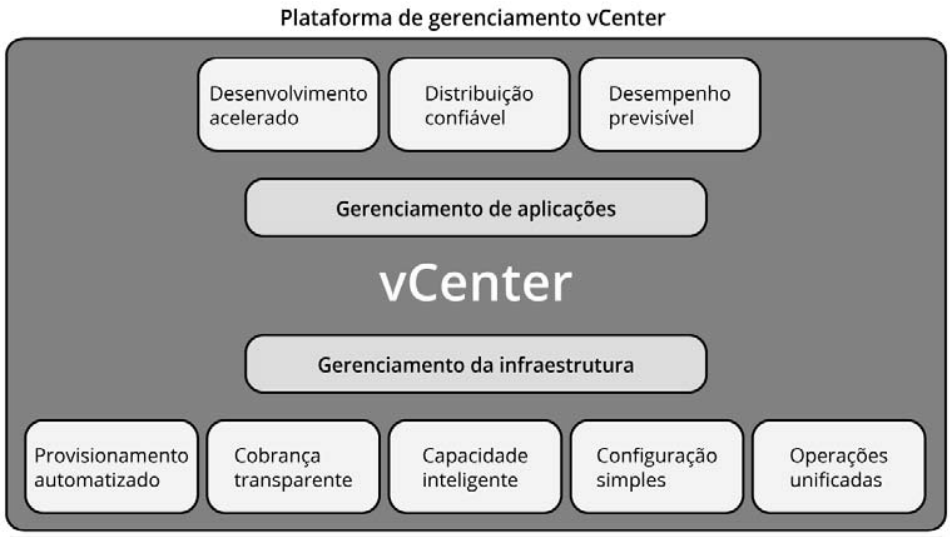


Figura 10.7 Solução de gerenciamento VMware vCenter.

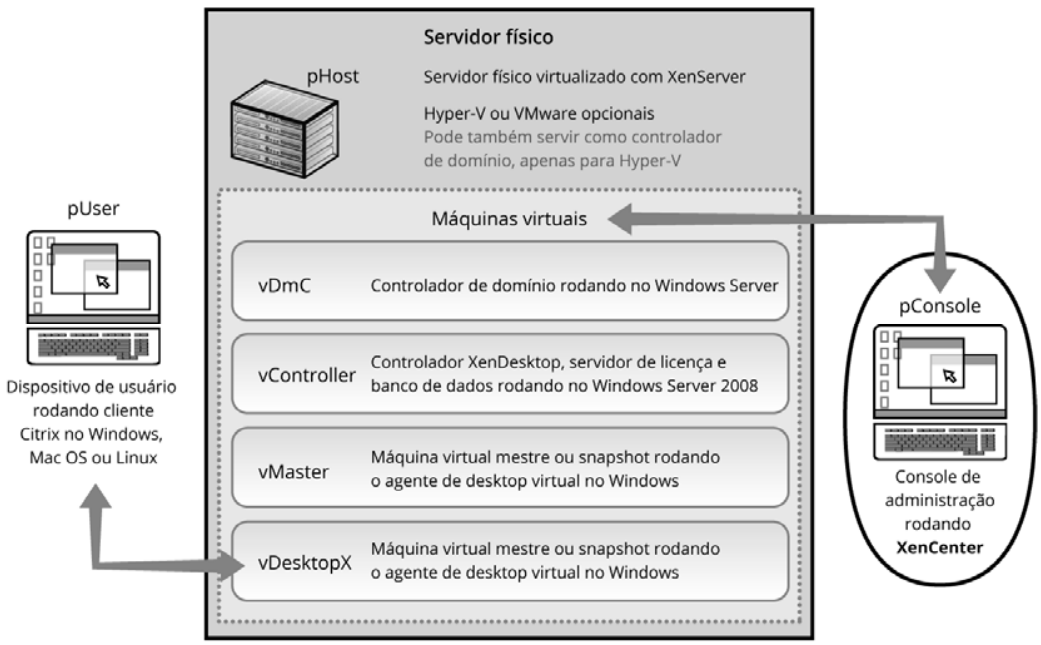


Figura 10.8 Solução de gerenciamento Citrix XenCenter.

"A virtualização sem gerenciamento é mais perigosa do que não usar a virtualização", Tom Bittman, vice-presidente do Gartner, Inc.

Hardware

Para facilitar o trabalho das equipes de suporte, alguns procedimentos devem ser utilizados para simplificar e otimizar o gerenciamento dos três principais hipervisores do mercado: Citrix XenServer, Microsoft Hyper-V e VMware ESX.

Não importa o hipervisor escolhido para seu ambiente virtualizado; todas as três soluções têm praticamente as mesmas demandas e desafios a resolver: memória, CPU, disco (storage) e recursos de rede.

Gerenciamento de memória

- Memory Overcommit
- Ballooning
- Transparent Page Sharing (TPS)
- Swap

A memória deve ser quantificada antes da escolha do hipervisor, por ser o primeiro recurso a se esgotar. É muito comum sistemas virtualizados terem desempenho prejudicado por falta de memória, bem antes de esgotarem a capacidade de processamento da CPU. Portanto, é muito importante no projeto de virtualização o dimensionamento correto da quantidade de memória necessária para evitar perda de rendimento. O ideal é configurar o máximo que for possível dentro do orçamento disponível. A longo prazo, a utilização de um número menor de servidores com grande quantidade de memória se mostra mais econômica do que o uso de um número maior de servidores pequenos, em consideração a fatores como consumo de energia, refrigeração e licenciamento.

A técnica conhecida como Memory Overcommit permite ao hipervisor oferecer às máquinas virtuais mais memória do que a memória física existente no host. Como podemos imaginar, essa técnica pode ser perigosa, pois o estouro de memória é capaz de parar um computador. A razão para o emprego dela é a constatação de que a maioria das máquinas virtuais utiliza menos memória física do que possui. É comum uma máquina virtual ter 1 GB de memória alocada e só usar de 300 MB a 400 MB, deixando liberados até 700 MB. Essa porção é, então, disponibilizada para outras máquinas virtuais. O ESX utiliza essa técnica por meio de três mecanismos:

- **Ballooning**: quando um host precisa recuperar parte da memória da máquina virtual, informa ao driver de balloon a quantidade de memória necessária. Em seguida, o driver de balloon começa a "inflar" dentro do Sistema Operacional, fazendo com que ele use suas próprias técnicas de gerenciamento de memória para oferecer a memória necessária ao driver de balloon.
- **Transparent Page Sharing (TPS)**: técnica que possibilita às máquinas virtuais com páginas de memória idênticas compartilharem-nas. Isso permite que o ESX armazene apenas uma cópia dessa página em sua memória física (ex: máquinas virtuais que usam o mesmo Sistema Operacional).

- **Swap**: durante a inicialização de uma máquina virtual, é criado um arquivo .VSWP no disco, utilizado durante o processo de swapping. Quando o host sente a necessidade de fazer o swapping, transfere páginas da memória física da MV para esse arquivo, liberando essas páginas na memória física do host para serem utilizadas por outras máquinas virtuais. O Hyper-V promete incorporar essa técnica a partir do Service Pack 1, com a denominação de memória dinâmica. Já o XenServer usa ballooning, mas demanda a instalação de driver específico (balloon driver).

vCPU

Apesar da proliferação de CPUs multicore, no caso de ambientes virtualizados, quanto maior for a quantidade de máquinas virtuais criadas com mais de uma vCPU, maior será o trabalho do hipervisor para escalonar em cada CPU as vCPUs dos sistemas convidados. Uma dica é ter cuidado com o uso excessivo de vCPUs em uma máquina virtual, pois o desempenho pode ser abaixo do esperado. O emprego de múltiplas vCPUs é indicado quando a aplicação for desenvolvida para multiprocessamento. Além disso, o excesso de vCPUs pode causar degradação do sistema com o aumento de **colisões** no processo de escalonamento das vCPUs.

Colisões

Mais de uma vCPU executando no mesmo núcleo ao mesmo tempo.

Normalmente, a utilização da mesma CPU por várias máquinas virtuais não configura um problema; pelo contrário, esse mecanismo é o alicerce da virtualização. O problema é o uso excessivo, fazendo com que o servidor físico fique sobrecarregado com a respectiva queda de desempenho dos sistemas convidados. Importante lembrar que os recursos computacionais são finitos e há casos em que aumentá-los pode ficar caro, tanto em relação ao tempo quanto ao custo financeiro.

Outra consideração importante é sobre o uso de sistemas de arquitetura de 32 bits rodando nos hipervisores de 64 bits. Os sistemas operacionais de 64 bits precisam de mais memória apenas para "ligar" uma máquina virtual. Lembre-se que, se o seu hipervisor não for inteligente o suficiente para distribuir de forma balanceada as cargas de trabalho (workloads), você mesmo terá que fazê-lo. Não tem sentido colocar duas máquinas virtuais, grandes consumidoras de recursos, no mesmo host. Em máquinas virtuais que fazem uso intensivo de CPU (CPU bound), o tamanho da memória cache pode ser o fator mais importante para seu desempenho.

Vlan tagging

É o processo em que o hipervisor acrescenta um identificador aos frames ethernet. Esses pacotes são identificados para permitir que, ao chegarem em um switch físico, possam ser direcionados à sub-rede apropriada. A recomendação para a definição do número de placas de rede é usar uma placa para cada tipo de tráfego e isolar cada tráfego em um domínio de broadcast. Para fazer esse isolamento, o ideal é o uso de VLANs e de **VLAN Tagging** para delimitar as sub-redes. O uso de VLAN Tagging pode reduzir significativamente o número de interfaces de rede.

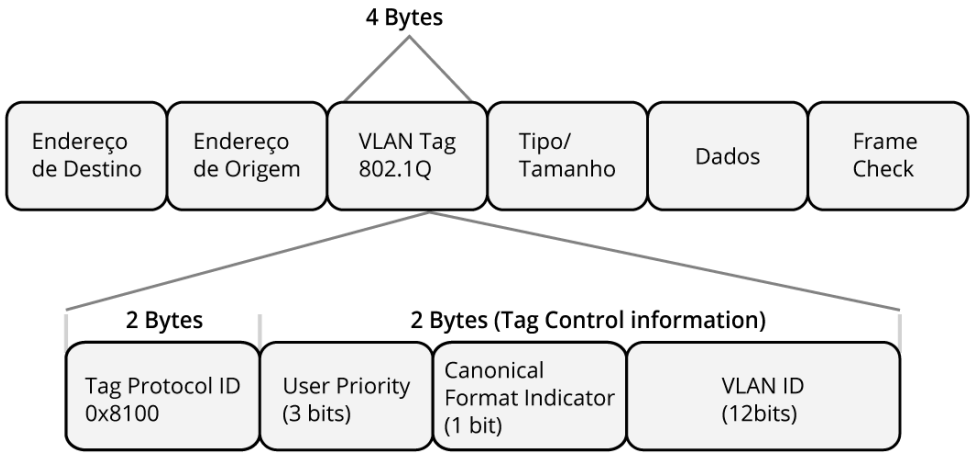


Figura 10.9 VLAN Tagging.

Quantidades recomendadas

1. Pelo menos uma interface para o tráfego de gerenciamento.
2. Pelo menos duas interfaces para o tráfego das máquinas virtuais.
3. Pelo menos duas interfaces para o tráfego de dados de storage IP (NAS ou iSCSI).
4. Pelo menos uma interface para live migration.

Em muitos datacenters, é comum usar a interface de gerenciamento para fazer "live migration", visto que o tráfego de gerenciamento pode ser controlado. O custo financeiro diminui, porém a segurança da instalação pode ser comprometida.

Sistemas de discos

Um projeto de datacenter é bem-sucedido quando consegue unir servidores poderosos com uma infraestrutura de storage igualmente capaz de suportar por anos as necessidades das aplicações e dos usuários.

A escolha do tipo de storage para seu ambiente virtual não é simples, pois a oferta engloba Network-Attached Storage (NAS) e também outras arquiteturas de Storage Area Network (SAN) – como Fibre Channel (FC), Internet SCSI (iSCSI) e Fibre Channel over ethernet (FCoE).

Máquinas virtuais nada mais são do que arquivos levados para um servidor e, por meio do hipervisor, executadas. As máquinas virtuais também precisam carregar e descarregar dados (arquivos). Como esses arquivos estão todos disponíveis em um servidor de discos, percebe-se a importância da escolha da forma de acesso a esse servidor. O uso de discos locais nos servidores impossibilita alguns recursos avançados de virtualização, como "live migration", e softwares de backup, como o VMware Consolidated Backup (VCB). Portanto, a utilização de storage compartilhado é praticamente mandatória em um ambiente virtualizado. Então, na

hora de projetar o storage compartilhado, é importante considerar os custos de manutenção e atualização.

Outro ponto de suma importância é o gerenciamento da plataforma. Os administradores terão que conviver com a criação e a manutenção de um enorme número de Logical Unit Number (LUNs), que, dependendo do tamanho da organização, pode ser de centenas ou até milhares. O aumento da largura de banda (bandwidth) não necessariamente aumentará o desempenho do sistema de discos. O maior desempenho pode ser alcançado com medidas que diminuam o acesso ao sistema, como a eliminação dos arquivos de swap do Windows.

Também pode ser uma boa estratégia configurar ambientes heterogêneos: empregar tecnologia SAN para aplicações de missão crítica e outras que sejam orientadas a bloco, e usar NAS para as aplicações secundárias, por ser simples e mais barata.

Fibre Channel (FC)

Principal tecnologia para arquitetura SAN, atualmente com velocidade de 10 Gbps, vem perdendo espaço para as novas tecnologias, principalmente pelo alto custo dos componentes da solução: Host Bus Adapters (HBA), switches, cabeamento etc. É claro que se a organização já possui uma base FC, os custos podem cair drasticamente.

Internet SCSI (iSCSI)

Representa uma mudança drástica para ambientes SAN. Mais fácil de instalar, configurar e gerenciar, com desempenho similar ao do FC – hoje já pode ser usado com velocidade de até 10 GbE (Gigabit Ethernet) e custos menores. Estudos como o do Enterprise Strategy Group (ESG) demonstram que o custo com iSCSI chega a ser 30% menor em comparação ao de FC. Outras vantagens da tecnologia iSCSI estão relacionadas à sua escalabilidade e segurança, por poder usar a encriptação do IPsec.

Fibre Channel over Ethernet (FCoE)

É o encapsulamento dos quadros FC em redes ethernet. Permite a utilização de FC em redes 10 Gb Ethernet ou maior, enquanto preserva o protocolo de Fibre Channel. Com FCoE, o FC passa a ser mais um protocolo correndo junto com o tráfego IP. FCoE opera diretamente acima do ethernet na pilha de protocolos de rede, diferentemente do iSCSI, que roda no topo do TCP/IP. Como consequência, FCoE não é roteável pela camada IP e, por isso, não funciona por meio de redes IP roteadas. Com relação a preço, iSCSI continua sendo mais barato e também roda em redes 10 GbE, o que anula o principal argumento em prol das opções FC/FCoE.

Dell Virtualization Advisor Tool

Em adição à ideia de utilizar uma arquitetura de referência, a Dell desenvolveu um advisor tool que visa a atender a necessidades específicas. O Dell Virtualization Advisor (mostrado na próxima figura) recomenda uma configuração de hardware baseada no fornecimento de dois tipos de informações:

- Descrição do ambiente existente;
- Características das cargas de trabalho (workloads) previstas para a infraestrutura virtual.

Em ambos os casos, a ferramenta propicia flexibilidade para a escolha das funcionalidades desejadas, e a entrada de dados é utilizada para determinar os requisitos de storage, servidores e rede. A saída é uma lista de hardware suportada pela Dell que atende aos requisitos iniciais.

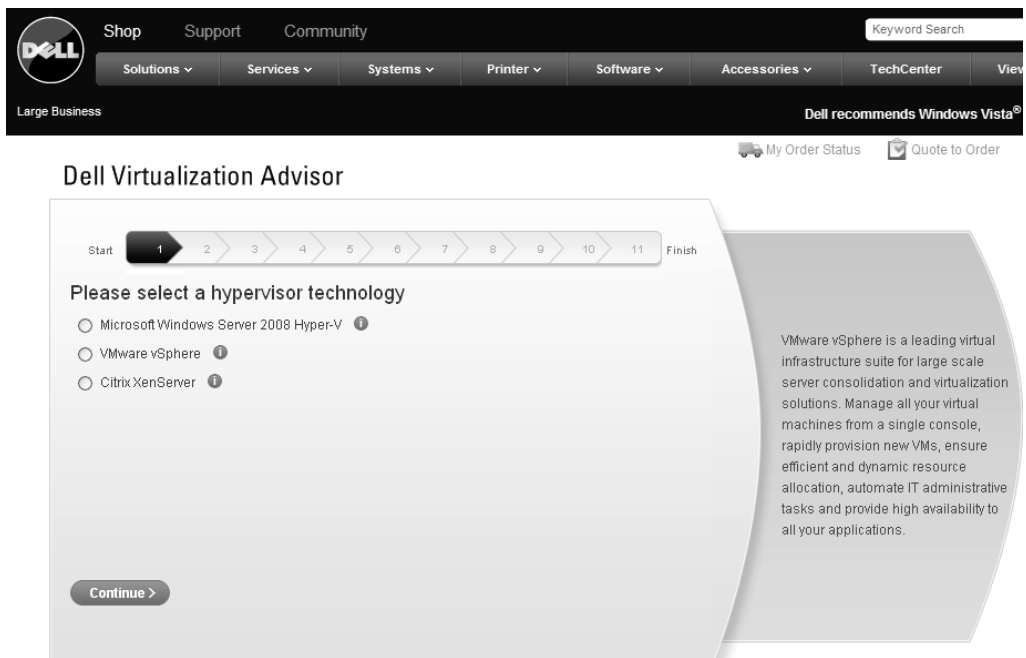


Figura 6.3 Dell Virtualization Advisor.

Comparação entre soluções de virtualização

Durante o curso, foram apresentados os subsídios para identificar as ferramentas mais adequadas às necessidades da sua organização. Sabemos que as empresas da área de virtualização estão constantemente envolvidas no processo de inovação de novas ferramentas, além de fusões e incorporações que têm permitido o rápido avanço dos últimos anos. Para se manter atualizado, é importante usar as ferramentas que fornecem informações sobre o assunto, acompanhar blogs, receber alertas do Google e assinar distribuições de RSS. A quantidade de informação atualmente disponível é muito grande. Por isso, é necessário bom senso para filtrar as notícias pertinentes dentre as descartáveis. A

tendência dos fabricantes é sempre mostrar as melhores partes de seu produto, destacando suas vantagens em relação aos concorrentes. Por isso, muito cuidado com os comparativos. Procure opiniões de consultores idôneos e independentes.

A principal recomendação é fazer a avaliação dos produtos disponibilizados para "test drive" e tirar suas próprias conclusões quanto à adequação da solução frente às necessidades da sua empresa. Sabemos que nem sempre há tempo para uma avaliação criteriosa. Por esse motivo, o planejamento é a alma do negócio. O orçamento é um limitador na escolha da plataforma de virtualização (hardware e software); portanto, seja criterioso e deixe sempre margem para atualizações que não demandem a parada dos sistemas que estão sob sua responsabilidade.

Bibliografia

- BARHAM, Paul *et al*. Xen and the art of virtualization. In: **Symposium on Operating Systems 19**. Proceedings. Bolton Landing, NY: AMC, 2003. Disponível em: http://www.cl.cam.ac.uk/research/srg/netos/papers/2003-xensosp.pdf
- BARHAM, P.; DRAGOVIC, B.; FRASER, K.; HAND, S.; HARRIS, T.; HO, A.; NEUGEBAUER, R.; PRATT, I. WARFILED, A. **Xen and the Art of Virtualization**. SOSP, 2003.
- BLADE.ORG. **Blades platforms and network convergence**. Piscataway, NJ: Blade.org, 2008. Disponível em: http://www.bladenetwork.net/userfiles/file/PDFs/Blade_WP03_net_convergence.pdf
- CARISSIMI, Alexandre S. Virtualização: Princípios básicos e aplicações. In: DORNELES, Ricardo V.; STEIN, Benhur; HOLBIG, Carlos; NAVAUX, Philippe O. A. (Org.). **ERAD 2009 – 9ª Escola Regional de Alto Desempenho**. Porto Alegre, 2009, v. 1, p. 39-69.
- CISCO SYSTEMS. **Data Center Blade Server Integration Guide**. San Jose, CA: Cisco, c2006. Disponível em: http://docstore.mik.ua/univercd/cc/td/doc/solution/bladserv.pdf
- CITRIX PRODUCT DEVELOPMENT TEAM. XenServer Administrator's Guide 5.5.0. 1.0 ed. [s.l.]: Citrix, 2009. Disponível em: http://download.alyseo.com:81/contrib/Xensource/ISOs/5.5.0/XenServer-5.5.0-installation.pdf
- COMPUTERWORLD. **Virtualização: eficiência sob medida**. Rio de Janeiro; São Paulo, 2009. Disponível em: http://www.openinnovatio.org/wp-content/uploads/2009/09/EB_CW_Virtualizacao.pdf
- CLARK, Christopher *et al*. Live migration of virtual machines. In: **Symposium on networked system Design & Implementation**. Proceedings. Berkeley, CA: USENIX association, 2005. Disponível em: http://www.cl.cam.ac.uk/research/srg/netos/papers/2005-migration-nsdi-pre.pdf
- CLARK, Tom. **Designing Storage Area Network: a practical reference for implementing Fibre Channel and IP SANs**. 2. ed. Boston, MA: Addison-Wesley, 2003.
- _____. **IP SANs: a guide to iSCSI, iFCP and FCIP protocolos for Storage Area Networks**. Boston, MA : Addison-Wesley, c2002.
- DELL. **Citrix XenServer 5.5 Dell edition: solution guide**. Round Rock, TX: Dell, 2009. Disponível em: http://support.dell.com/support /edocs/software/citrix/
- _____. **Citrix XenServer Dell Edition Reference Architecture: virtualization solutions engineering**. Round Rock, TX: Dell, 2008. Disponível em: http://i.dell.com/sites/content/business/solutions/ engineering-docs/en/Documents/dell_xenServer_reference_architecture_v1.1.pdf

- _____. **Dell VMware ESX/ESXi 3.5 Virtualization Reference Architecture**. Round Rock, TX: Dell, c2008. Disponível em: http://i.dell.com/sites/content/business/solutions/whitepapers/en/Documents/virtualization_ref_architecture_v1_1.pdf
- _____. **VMware ESX3.5 Update: Deployment Guide**. Round Rock, TX: Dell, c2007-2009.
- DORNAN, Andy. **Virtualização e múltiplos núcleos: como licenciar?** Parte I. InformationWeek, 2007. Disponível em: http://www.itweb.com.br/noticias/index.asp?cod=44388
- _____. **Virtualização e múltiplos núcleos: como licenciar?** Parte II. InformationWeek, 2007. Disponível em: http://www.itweb.com.br /noticias/index.asp?cod=44389
- ESG. **Research Report: The Impact of Server Virtualization on Storage**. Milford, MA: Enterprise Strategy Group, 2007.
- HALETKY, Edward L. **VMware ESX server in the enterprise: planning and securing virtualization servers**. Upper Saddle River, NJ: Prentice Hall, c2008.
- HAND, Steven *et al*. Controlling the XenoServer Open Platform. In: **IEEE Conference on Open Architectures and Network Programming**. Proceedings. San Francisco, CA: Institute of Electrical and Electronics Engineers, 2003. Disponível em: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.9501&rep=rep1&type=pdf
- HEWLETT-PACKARD. **The AMD processor roadmap for industry standard-servers**. Palo Alto, CA: HP, 2008. Disponível em: http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00428708/c00428708.pdf
- _____. **The Intel processor roadmap for industry standard servers**. Palo Alto, CA: HP, 2007. Disponível em: http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00164255/c00164255.pdf
- LOFFINK, John. **Dell PowerEdge M1000e Modular Enclosure Architecture**. Round Rock, Tx: Dell, 2008. Disponível em: http://www.dell.com/downloads/global/products/pedge/en/pedge_m1000e_white_paper.pdf
- LOWERY, J. Craig. **10 Gigabit Ethernet Unifying Fabric: Foundation for the Scalable Enterprise**. Dell Power Solutions, 2007. Disponível em: http://www.dell.com/downloads/global/vectors/2007_gigabit.pdf
- MARSHALL, David; REYNOLDS, Wade A.; McCrory, Dave. **Advanced server virtualization: VMware and Microsoft platforms in the virtual data center**. Boca Raton: Auerbach Publications, 2006.
- MATTHEWS Jeanna *et al*. **Running Xen: a hands-on guide to the art of virtualization**. Upper Saddle River, NJ: Prentice Hall, c2008.

- MICROSOFT. **Leveraging Microsoft Optimization to create your dynamic IT Roadmap**. Redmond, WA: Microsoft, 2009. Disponível em: http://microsoftio.partnersalesresources.com/content/overview/fy09_optimization_datasheet.pdf
- ROSEN, Rami. **Virtualization in Xen 3.0**. Linux Journal, [s. l.], 02 mar. 2006. Disponível em: http://www.linuxjournal.com /node/8909/print
- RUEST, Danielle; RUEST, Nelson. **Virtualization: a beginner's guide**. New York: McGraw Hill, c2009.
- SOMASUNDARAM, G.; SHRIVASTAVA, ALOK (Eds.). **Information Storage and Management: Storing, Managing and protecting Digital Information**. Indianapolis: Wiley Publishing, 2009.
- TAKEMURA, Chris; CRAWFORD, Luke. **The book of Xen: a practical guide for the system administrator**. San Francisco: No Starch Press, c2010.
- VERAS, Manoel. **Datacenter: componente central da infraestrutura de TI**. Rio de Janeiro: Brasport, 2009.
- VMware ESXi. Product datasheet. VMware, 2008.
- VMWARE. **Building the Virtualized Enterprise with VMwere Infrastructure**. Palo Alto, CA: VMware, 2006. Disponível em: http://www.vmware.com/pdf/vmware _infrastructure_wp.pdf
- _____. **iSCSI Design Considerations and Deployment Guide**. Palo Alto, CA: VMware, c2007. Disponível em: http://www.vmware.com/files/pdf/iSCSI_design_deploy.pdf
- _____. **Reducing Server Total Cost of Ownership with VMware Virtualization Software**. Palo Alto, CA: VMware, 2006. Disponível em: http://www.vmware.com/pdf/TCO.pdf
- _____. **SAN systems design and deployment guide**. 2 ed. Palo Alto, CA: VMware, 2007. Disponível em: http://www.vmware.com/ pdf/vi3_san_design_deploy.pdf
- _____. **Security Design of the VMware Infrastructure 3 Architecture**. Palo Alto, CA: VMware, c2007. Disponível em: http://www.vmware.com/pdf/vi3_security_architecture_wp.pdf
- _____. **The architecture of VMware ESXi**. Palo Alto, CA: VMware, c2007. Disponível em: http://www.vmware.com/files/pdf/vmware_esxi_ architecture_wp.pdf
- _____. **Understanding Full Virtualization, Paravirtualization and Hardware Assist**. Palo Alto, CA: VMware, 2007. Disponível em: http://www.vmware.com/files/pdf/VMware_ paravirtualization.pdf
- _____. **Virtualization Overview**. Palo Alto, CA: VMware, 2007. Disponível em: http://www.vmware.com/pdf/virtualization.pdf

- _____. **VMware and VMotion and CPU Compatibility**. Palo Alto, CA: VMware, c2008. Disponível em: http://www.vmware.com/files/pdf/vmotion_info_guide.pdf
- _____. **VMware Consolidated Backup: best practices and deployment considerations for SAN environments**. Palo Alto, CA: VMware, c2007. (Best practices). Disponível em: http://www.vmware.com/files/pdf/vcb_best_practices.pdf
- _____. **VMware Consolidated Backup: Improvements in version 3.5**. Palo Alto, CA: VMware, c2008. Disponível em: http://www.vmware.com/files/pdf/vcb_35_new.pdf
- _____. **VMware ESX Server using EMC Clariion Storage Systems**. Palo Alto, CA: EMC Corporation, c2007. Disponível em: http://www.vmware.com/vmtn/resources/540
- _____. **VMware virtual networking concepts**. Palo Alto, CA: VMware, c2007. Disponível em: http://www.vmware.com/files/pdf/virtual_networking_concepts.pdf
- VON HAGEN, William. **Professional Xen virtualization**. Indianapolis, IN: Wiley Pub., c2008.
- PFEIFER MACEDO, Leandro; LAWRENCE WEBSTER, Ian. **Projeto CCD "Centro Compartilhado de Dados"** RNP Brasil, 2012.

esr.rnp.br