

Inserindo a virtualização nos serviços e processos de TI

Rogério de Oliveira ¹

1) Universidade Presbiteriana Mackenzie, São Paulo, Brasil

roger.oliveira@mackenzie.br

Resumo

A virtualização vem sendo largamente empregada para a redução de custos nos *data centers*, principalmente no que se refere a redução de espaço, energia e custos de gerenciamento de infraestrutura. Ela reduz o custo dos serviços de Tecnologia da Informação (TI) e desse modo viabiliza novas facilidades para o desenvolvimento de aplicações como mobilidade, recuperação de desastre, disponibilidade e alocação flexível de recursos, antes limitados. Apesar desses benefícios as tecnologias de virtualização requerem cuidados em sua aplicação e trazem impactos significativos na forma de operação e organização dos *data centers*. Este trabalho apresenta os principais conceitos de virtualização, seus benefícios e impactos, com ênfase em tecnologias de virtualização de servidores. Propõe, então, que a inserção das tecnologias de virtualização nos *data centers* seja feita com foco em um *processo de virtualização*, de forma a minimizar riscos e garantir os benefícios da virtualização no longo prazo.

Palavras chave: virtualização, paravirtualização, monitores de máquina virtual, VMWare, Xen, Virtual Server, Data Centers

1 Introdução

A grande proliferação de máquinas em anos recentes, em geral, com baixa utilização, altos custos de espaço, energia e gerenciamento, e maior fragilidade e exposição dos ambientes criou um ambiente propício ao ressurgimento das alternativas de virtualização [Figueiredo et al. 2005],[Rosenblum e Garfinkel 2005]. Pesquisas apontam para a virtualização como uma direção estratégica de fornecedores de *hardware*, *software* e empresas usuárias [MacLaughlin 2008],[Ferguson 2007a],[Ferguson 2007b],[Lai, 2008],[Donston 2008],[Gillet e Schreck 2006]. Seus números diferem de acordo com o mercado e as metodologias empregadas mas em geral pode se concluir que:

- a) ***A maior parte das grandes empresas emprega alguma tecnologia de virtualização.*** Ela é mais presente nas grandes empresas, mas também está presente nos segmentos menores. É também crescente o interesse por uso de tecnologias de virtualização.
- b) ***Tecnologias de virtualização de servidores são as mais empregadas.*** A virtualização está comumente associada à virtualização de servidores (máquinas) e é a mais empregada. Em segundo lugar aparece o emprego da virtualização de *storage*. A VMWare¹ aparece como o fornecedor líder de virtualização de servidores, mas é crescente o interesse por soluções abertas como o Xen ou soluções Microsoft.
- c) ***Cresce o interesse por tecnologias de virtualização para computação pessoal.*** Embora ainda pouco explorada é crescente o interesse por tecnologias que reduzam o custo e o esforço de gerenciamento de um grande número de *desktops*.

¹ A Tabela 1 indica o *site* de referência dos produtos virtualização citados.

- d) **A redução de custos ainda é a principal motivo para adoção da virtualização.** As reduções de custo de energia e espaço, que podem variar de 20-80%, são a principal causa para o investimento em tecnologias de virtualização e têm sido alcançadas satisfatoriamente na maior parte das empresas [MacLaughlin 2008].
- e) **Cresce o interesse por outros benefícios da virtualização.** Além da redução de custos a virtualização é uma alternativa para problemas como o provisionamento flexível de recursos para aplicações, soluções de segurança e o suporte a planos de recuperação de desastres e *backup*.

Recentemente, um reforço adicional ao emprego de técnicas de virtualização vem sendo dado pelo interesse em iniciativas do uso sustentável de energia, *green* e *utility computing* [Computerworld 2008].

A crescente consolidação das tecnologias de virtualização aos *data centers* traz mudanças significativas e torna importante a avaliação dos impactos da virtualização e os benefícios que ela traz.

Neste trabalho abordam-se os benefícios da virtualização no desenvolvimento de aplicações e seus impactos. Com foco maior na virtualização de servidores, e com base na experiência prática de projetos de virtualização, apontam-se riscos e impactos dessa abordagem indicando-se, em cada caso, possíveis ações de mitigação.

2 Virtualização e seus conceitos

Virtualização é uma camada de abstração entre os dispositivos físicos (*hardware*) e as aplicações usuárias desses recursos. Ela quebra o paradigma da arquitetura de sistemas em que, a um sistema operacional (SO), encontram-se associados recursos físicos dedicados como processadores, memória, discos e dispositivos de rede. Mais associada a virtualização de servidores, a virtualização é um conceito que aplica-se a servidores, dispositivos de armazenamento (*storage*), de rede e de computação pessoal assim como aplicações. Entre as diferentes formas de virtualização empregadas hoje em TI destacam-se:

- a) **Virtualização de servidores.** É o modelo dominante e possui diversas formas de implementação para diferentes arquiteturas de processadores. Basicamente ela permite que uma máquina física apareça como múltiplas máquinas. Essas máquinas virtuais, dentre outras propriedades, oferecem um ambiente de processamento de programas idêntico a máquina física. Esse modelo é tratado em detalhe a seguir.
- b) **Virtualização de armazenamento.** Existem diferentes níveis de virtualização de dispositivos de armazenamento (*storage*). Pode-se entender o RAID como o mecanismo mais elementar de virtualização uma vez que ele implementa discos lógicos com uma diferente estrutura de *hardware*. Mas o termo é mais adequado para tecnologias que oferecem uma capacidade de armazenamento dinâmica e independente de sua localização física nos discos ou no tipo de mídia. Em geral ela está associada a tecnologias de *arrays* de discos acessados via rede de dados (*network attached storage*, NAS) ou uma rede dedicada de armazenamento (*storage area network*, SAN).
- c) **Virtualização de redes.** Virtual LANs (VLAN), virtual IPs (VIP) e Virtual Private Networks (VPN), embora nem sempre reconhecidas como formas de virtualização [Dittner e Rule 2007], constituem-se nos principais mecanismos para tornar o acesso aos recursos de rede flexíveis e independentes dos recursos físicos utilizados. *Switches* virtuais ainda estão associados a máquinas virtuais em algumas tecnologias [IBM 2004].

- d) **Virtualização de computação pessoal.** Diferentemente das tecnologias anteriores a virtualização de computação virtual (*client* ou *end-user virtualization*) não está associada a concentração e centralização de recursos. Os propósitos desse modelo são em geral o de simplificar o gerenciamento de *software* de um grande número de clientes, prover mobilidade ou ainda isolar aplicações com objetivos de segurança. São exemplos desse tipo de tecnologia o Microsoft Terminal Server e o VMWare Virtual Desktop.
- e) **Virtualização de aplicações.** A aplicação dos conceitos de virtualização sobre *software* leva a um conceito controverso que engloba diferentes abordagens. Elas trazem em comum a transparência de localização dos serviços ou dados de uma aplicação, em geral dispersos, provida por uma camada de *software* adicional. Sob essa abrangente definição se encaixam tecnologias como grade de dados, computação em grade [Meier 2008], *web services* e outras arquiteturas orientadas a serviço [Vogels 2008].

Embora muitas considerações sobre virtualização possam ser aqui aplicadas a tecnologias de virtualização em geral, nosso foco será na virtualização de servidores. A caracterização e os principais conceitos dessa forma de virtualização encontram-se a seguir.

2.1 Virtualização de servidores

Os requerimentos básicos e a formalização da virtualização foram definidos em [Popek e Goldberg 1974]. Esses princípios basicamente não sofreram alterações e podem ser diretamente aplicados, muito embora as tecnologias de virtualização tenham evoluído, por exemplo, incorporando assistência de *hardware* e recursos de rede à virtualização [Figueiredo et al. 2005]. Um modelo de camadas, Figura 1, normalmente é empregado para representar a virtualização como uma interface do sistema para a execução de programas.

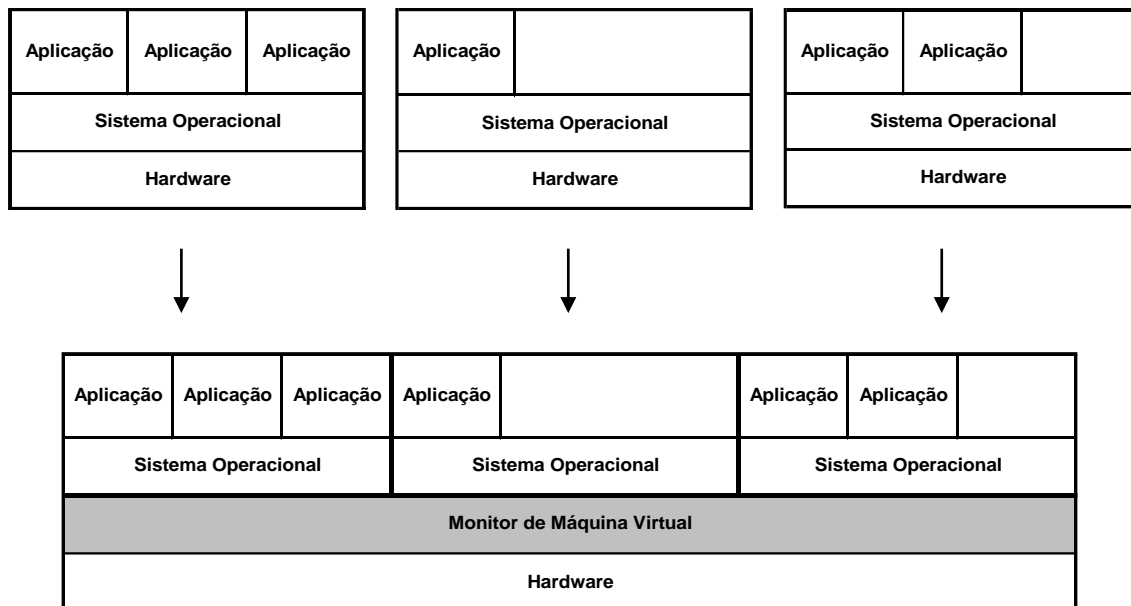


Figura 1 – O Monitor de Máquina Virtual fornece uma camada de abstração suficientemente próxima do *hardware* para a execução de quaisquer programas escritos para o uso daquele *hardware*. O processo de consolidação permite, então, centralizar várias aplicações em um único *hardware*.

Uma *máquina virtual* é duplicata eficiente e isolada de uma máquina física. As máquinas virtuais são providas por um *monitor de máquina virtual* (MMV) com as propriedades [Popek e Goldberg 1974]: i) *Compatibilidade* - provê um ambiente para os programas que é essencialmente idêntico a máquina original; ii) *Desempenho* - o programas que executam nesse ambiente têm um desempenho semelhante ao desempenho na máquina nativo; iii) *Controle* - o MMV tem completo controle dos recursos do sistema. Esses recursos, como processadores, memória e dispositivos de E/S, são *virtualizados*, isto é, suas interfaces e todos os recursos visíveis dessas interfaces, são mapeadas para interfaces do MMV, enquanto os recursos do sistema real oferece a implementação dessas interfaces [Smith e Nair 2005a].

Essas três propriedades são determinantes no funcionamento dos MMV. Por exemplo, o fato de as máquinas virtuais proporcionarem um ambiente para a execução de programas com um desempenho próximo da execução nativa, implica que grande parte das instruções deva ser executada diretamente no processador real, sem intervenção do MMV (ver Figura 2). Eliminam-se assim, da categoria de virtualização, programas como emuladores e interpreters como o Bochs e Microsoft Virtual PC.

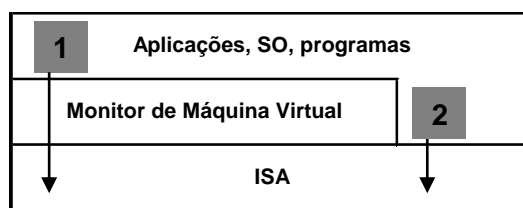


Figura 2 – Diferentemente dos emuladores a virtualização permite que instruções não privilegiadas (2) sejam executadas diretamente (Instruction Set Architecture) provendo aos programas um desempenho mais próximo da execução no *hardware* nativo. A forma como o Monitor de Máquina Virtual permite a execução das instruções privilegiadas (1) define diferentes tipos de máquinas virtuais. Na solução geral a máquina virtual oferece algum *overhead* na execução desse tipo de instruções.

Em sua forma básica um MMV executa em modo privilegiado tendo acesso a todos os recursos do sistema. As máquinas virtuais, entretanto devem executar em modo não privilegiado tanto as instruções privilegiadas (*kernel* do SO) como as não privilegiadas. Quando uma máquina virtual tenta executar um código privilegiado a CPU passa o controle (*traps into*) ao MMV que emula a instrução privilegiada para o estado da máquina virtual [Rosenblum e Garfinkel 2005]. O desafio da virtualização está em garantir nesse processo as propriedades de compatibilidade, desempenho e controle, uma vez que esse mecanismo não corresponde ao comportamento de todas as instruções nas arquiteturas atuais.

Por exemplo, em arquitetura de processadores ISA-32 (x86), 17 instruções com diferentes níveis de privilégio, com acesso a recursos de sistema e comumente executadas por SOs, podem causar falhas no sistema ou dar falso retorno quando não executadas em modo privilegiado, não passar o controle ao MMV e, portanto, entrar em conflito com o MMV [Dittner e Rule 2007]. A instrução POPF (*pop CPU flags from stack*) é um popular exemplo desse tipo de instrução. Dentre as técnicas empregadas na solução desse tipo de problema prevalecem a *execução direta*, combinada com a tradução binária de código, e a *paravirtualização* [Rosenblum e Garfinkel 2005].

Xen e XenServer são exemplos de sistemas que empregam a técnica de paravirtualização (ver Tabela 1). Nesses sistemas a interface das instruções que apresentam conflito para virtualização direta é substituída por um conjunto de instruções virtualizável equivalente. Embora não traga impacto para a execução da maior parte das aplicações, a paravirtualização requer a modificação

dos sistemas operacionais que executam sob as máquinas virtuais (*OS guest*). A mudança não é significativa e por isso essa desvantagem vem sendo superada com a adesão dos fornecedores de SO.

VMWare e Virtual PC (ver Tabela 1) empregam a técnica de execução direta combinada com a tradução binária de código, e oferecem assim uma virtualização *completa* (*full virtualization*). Em uma máquina VMWare as instruções normais executam sem mudança enquanto, instruções que requerem um tratamento diferenciado, como POPF, são traduzidas para um bloco de instruções substituto equivalente, mas que permite a execução direta [Adams e Agesen 2006]. O bloco substituto é *cached* de modo que a tradução é executada uma única vez. Aplicações ou sistemas operacionais executam assim sem qualquer mudança. A tradução binária substitui as instruções durante a primeira execução do código, enquanto a paravirtualização substitui as instruções no código fonte dos sistemas operacionais ou programas [Rosenblum e Garfinkel 2005]. A virtualização completa requer uma combinação entre o *hardware* e o *software* de virtualização empregado, o que dificulta a garantia de compatibilidade de versões superiores de *software* com versões anteriores (*backward compatibility*). Ela também traz uma pequena penalidade de desempenho pela tradução binária. Essas desvantagens, entretanto, parecem não ser consideráveis para o mercado de aplicações.

Existem diversas outras tecnologias de virtualização, tanto para processadores x86 como para outras arquiteturas. Em particular IBM, Sun Microsystems e Hewlett-Packard (HP) oferecem tecnologias maduras de virtualização para outras arquiteturas no modelo de *partições lógicas* (*LPARs*, *Solaris Containers* e *Virtual Server Environment* respectivamente) e fornecedores de *hardware* vem promovendo alterações nas arquiteturas IA-32 (Intel e AMD) e PowerPC (IBM) para fornecer suporte mais nativo a virtualização [Neiger et al. 2006]. Para uma discussão detalhada e uma taxonomia atual dos diferentes modelos de virtualização consultar [Smith e Nair 2005a]. Para maiores detalhes do funcionamento da virtualização completa dos sistemas VMWare ver [Adams e Agesen 2006]. Para um comparativo das técnicas de virtualização XenServer e VMWare consultar [VMWare 2005].

Esses conceitos permitem discutir a seguir os principais benefícios, impactos e recomendações quando do uso de tecnologias de virtualização.

Nome	Fornecedor/Site (Documentação)	CPU						SO						Tipo de Virtualização
		Host			Guest			Host			Guest			
		x86	PowerPC	IA-64	x86	PowerPC	IA-64	Windows	Linux	Outros	Windows	Linux	Outros	
Virtual PC	www.microsoft.com	X	X		X			X		X	X	X	X	Completa
Virtual Server	www.microsoft.com	X			X			X			X			Completa
WMware WS	www.vmware.com	X			X			X	X		X	X	X	Completa
WMware Server	www.vmware.com	X			X			X	X		X	X	X	Completa
WMware ESX	www.vmware.com	X			X					None	X	X	X	Completa
Xen	www.cl.cam.ac.uk	X	X	X	X				X	X	X	X	X	Paravirtualização
XenServer	www.xensource.com	X			X					None	X	X		Paravirtualização
Bochs	bochs.sourceforge.net	X	X		X			X	X	X	X	X	X	Emulação
KVM	sourceforge.net/projects/kvm	X			X				X		X	X		Outro
QEMU	fabrice.bellard.free.fr/qemu	X	X	X	X	X			X	X	X	X	X	Completa
Virtual Iron	www.virtualiron.com	X			X					None	X	X	X	Outro

Host none = bare metal (execução direta sobre o *hardware*)

Tabela 1 – Principais MMV para processadores x86. A tabela sumariza os sistemas suportados (cpu/SO, Host/Guest) e o tipo de virtualização a partir das informações dos sites de origem.

3 Benefícios comuns da virtualização

Do ponto de vista da organização de serviços de TI a virtualização é, em sua essência, uma atividade de custos [Vogels 2008]. Ela permite reduzir diretamente os custos de um serviço de TI ou viabiliza serviços, como mobilidade e capacidade de recuperação, a uma fração do custo de outras tecnologias.

Seus benefícios são largamente encontrados na literatura de virtualização [Figueiredo et al. 2005],[Rosenblum e Garfinkel 2005],[IBM 2004],[Meier 2008],[Smith e Nair, 2005a] incluindo relatórios de institutos de pesquisa e direção estratégica como o IDC e Forrester. A tabela 2 sumariza alguns dos problemas comuns dos serviços de TI e as abordagens da virtualização como uma alternativa de custo menor a solução desses problemas.

Alternativa da virtualização	Problemas comuns dos serviços de TI
Otimização e centralização de recursos	Baixa utilização dos recursos de <i>hardware</i>
Isolamento de aplicações	Segurança
Flexibilidade na alocação de recursos	Disponibilidade e confiabilidade dos serviços
Simplificação da infraestrutura	Processos laboriosos de operação

Tabela 2 – Problemas comuns dos serviços de TI encontram na virtualização uma alternativa.

A figura 3 mostra uma típica utilização da capacidade de processamento e armazenamento em *data centers*. Estima-se que o uso da capacidade de processamento na maior parte das empresas seja da ordem de 15% a 20% [Vogels 2008]. A virtualização permite otimizar o emprego desses

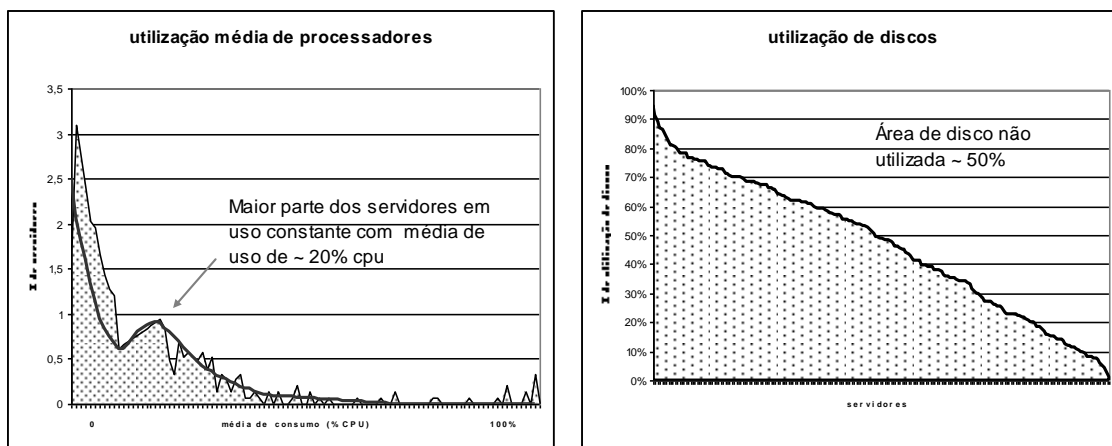


Figura 3 – Uso médio de processadores e percentual de uso de discos externos em um ambiente comercial típico sem virtualização. A maior parte dos servidores avaliados (cerca 2000) e que têm uso constante apresentam um consumo médio próximo a 20% de cpu. O grande percentual de servidores com consumos inferiores a 10 e 5% indica servidores com uso pouco frequente. No gráfico de uso de disco (120TB de discos externos distribuídos por cerca de 200 servidores) os servidores encontram-se ordenados da mais alta para a mais baixa taxa de utilização. A área do gráfico representa o *storage* alocado (área pontilhada) e a área superior o espaço disponível. Cerca 50% das áreas encontram-se disponíveis (~60TB). O espaço não utilizado seria ainda maior se considerados os discos não alocados mas disponíveis nas controladoras de armazenamento externo (reservas técnicas de storage).

intensificando a utilização de recursos por máquina física integrando várias máquinas virtuais no mesmo *hardware* através da consolidação (ver Figura 1). Os benefícios da consolidação são

imediatos: redução do número de máquinas físicas e correspondente redução de espaço e consumo de energia. Numa abordagem prática, a redução de máquinas pode chegar a reduções da ordem de 50% de espaço e consumo de energia.

A redução de máquinas e uso de máquinas virtuais traz uma simplificação significativa dos ambientes de *data centers*. Atividades laboriosas como instalação física dos equipamentos e cabeamento são substituídas pela alocação dinâmica de máquinas virtuais. Simplificam-se também processos manuais intensivos de controle de ativos.

A virtualização provê um ambiente isolado para aplicações idêntico ao de uma máquina física. Traz assim, alternativas para isolar máquinas sujeitas a falhas de segurança e ataques, protegendo os demais serviços. A um custo reduzido, viabiliza-se um maior número de instâncias de sistemas com diferentes e mais apropriadas configurações de segurança.

Depois da redução de custo a flexibilidade na alocação de recursos é o que traz maiores mudanças na forma de como as organizações de TI operam. Alguns recursos e usos comuns da das tecnologias de virtualização neste sentido são:

- a) *Encapsulamento e live migration* – Embora não faça dos requisitos de virtualização, uma funcionalidade adicional dos sistemas recentes de virtualização, como o VMWare ESX, é a capacidade do MMV encapsular por completo o estado de uma máquina virtual em execução, incluído seus dados, SO, aplicações e processos [Rosenblum e Garfinkel 2005]. Isso permite ao MMV suspender a execução de uma máquina virtual e reiniciá-la no mesmo ponto em outro sistema físico. Esse recurso aumenta a disponibilidade dos sistemas oferecendo um *downtime* mínimo para as aplicações em diversas situações como contorno para paradas programadas e a migração de uma máquina virtual para outra com recursos mais adequados para as aplicações. Esse recurso, combinado com outras ferramentas, ainda possibilita a execução de backups instantâneos das aplicações (*snapshot images*).
- b) *Gerenciamento flexível da carga de trabalho dos sistemas* – Máquinas virtuais podem ser migradas facilmente para *hardwares* diferentes e permite aos administradores de sistema proteger o nível de serviço em geral das aplicações (tempo de resposta, número de usuários, etc.) contra a falta de recursos por não provisionamento.
- c) *Serviços de alta disponibilidade e de recuperação de desastres* – A virtualização não oferece o mesmo nível de tolerância a falha e de recuperação de desastres que as tecnologias que empregam redundância de *hardware* (*clusters* e soluções de replicação). Entretanto, ela permite fornecer serviços de alta disponibilidade e de recuperação de desastre, com um nível de serviço ligeiramente menor mas sem os altos custos da redundância, viabilizando esses serviços por uma fração do custo onde antes não havia qualquer solução disponível.

O alcance dos benefícios acima, entretanto, não pode ser alcançado sem custos ou riscos. A seção seguinte discute potenciais impactos negativos da virtualização e aponta recomendações para a redução desses impactos e mitigação dos riscos envolvidos na adoção de tecnologias de virtualização.

4 Implementando a virtualização

Ao emprego de tecnologias de virtualização, vale lembrar, aplicam-se as recomendações e boas práticas gerais de projetos e mudanças em TI como a definição de padrões e estratégias próprias, e a revisão de políticas, processos e controles face ao emprego da nova tecnologia. Ainda como boa prática, a infraestrutura de virtualização deve ser construída em torno do atendimento dos serviços de TI e não da tecnologia. Observado isso podem-se tratar impactos e recomendações mais específicas aplicáveis a virtualização.

4.1 A virtualização como um processo

Embora diversas empresas apontem alcançar com sucesso seus objetivos na adoção de tecnologias de virtualização [MacLaughlin 2008], no médio e longo prazos a opção por tecnologias de virtualização deve estar amparada em um amplo *processo de virtualização* no *data center*. A virtualização traz benefícios muito imediatos, e a simples consolidação de máquinas proporciona reduções de espaço, energia que garantem um rápido retorno dos investimentos (ROI). Essa abordagem, além de não extrair todos os benefícios da virtualização, tem seu retorno rapidamente degradado à medida que o ambiente torna-se mais complexo: proliferam-se o número de imagens, alocam-se recursos temporariamente mas sem a desalocação e migram-se máquinas virtuais sem um controle apropriado. A ausência de controle acurado de uso dos recursos e rapidamente degrada a eficiência de uso das máquinas e o tempo para entrega de recursos para as aplicações. A Figura 4 ilustra a evolução dos custos na implementação de tecnologias de virtualização para a abordagem baseada em processos e a abordagem não estruturada. De fato, sendo relativamente recente a adoção de tecnologias de virtualização nos *datacenters*, muitas empresas podem estar se beneficiando dos retornos mais imediatos da virtualização e o real sucesso de suas estratégias deverá ser revisto nos próximos anos.

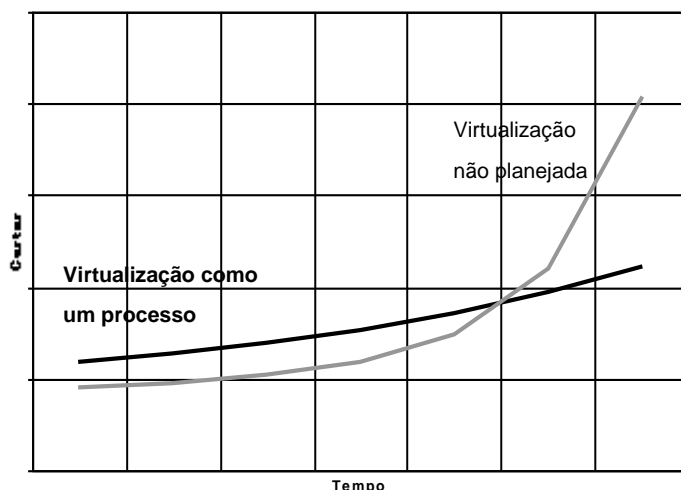


Figura 4 – A adoção da virtualização com foco unicamente na redução de máquinas físicas (virtualização não planejada) apresenta um custo inicial menor que a criação de um processo permanente com vistas a virtualização. Este último requer um investimento inicial maior em processos, ferramentas e automação.

Entretanto, a medida que o ambiente torna-se mais complexo, prolifera-se o número de imagens e alocações dinâmicas, torna-se mais difícil o controle do uso de recursos. Aumenta-se a necessidade de provisionamento de recursos ociosos e o tempo de alocação de recursos para as aplicações.

4.2 Iniciando o processo

Podemos buscar na gerência de projetos as etapas básicas para a implementação de um processo de virtualização: estabelecer um objetivo; definir uma estratégia; definir as soluções para essa estratégia; elaborar um plano; realizar um piloto; e implementar a solução.

O objetivo da virtualização deve ser claro, tangível e mensurável. O objetivo pode estar relacionado a custos ou melhora nos serviços. Esse objetivo pode definir estratégias diferentes de virtualização e levar a diferentes tipos de tecnologias. Por exemplo, diferentemente de servidores genéricos e de aplicações, servidores de arquivos podem ser virtualizados em dispositivos de *storage* do tipo NAS [Meier 2008],[IBM 2004] no lugar de máquinas virtuais x86. Ou pode-se optar por uma consolidação apenas de espaço dos servidores, com uma

sensível redução do consumo de energia, com a adoção de equipamentos do tipo *Blade Servers* [IBM 2004].

Focando-se na virtualização de servidores x86, deve-se ter em mente um cenário final do emprego da virtualização do no *datacenter*. Contrinui para a definição dessa estratégia uma lista de *máquinas candidatas a virtualização* (existentes ou futuras). Nem todas as máquinas ou aplicações são elegíveis para a virtualização. Em particular nas arquiteturas x86 a camada de virtualização sempre irá proporcionar um custo de desempenho das aplicações. Aplicações que são intensivas em E/S (entrada/saída) ou acesso a memória, como grandes bancos de dados, podem não ser elegíveis para virtualização. Incluem-se aqui também servidores de comunicação, como ftp, email servers e máquinas para a virtualização de *desktops*. O uso intensivo de recursos de rede e picos de utilização por esse tipo de aplicações pode penalizar outras aplicações compartilham a mesma banda de rede de uma mesma máquina física.

A definição da solução deve contemplar não só a escolha da MMV, como *hardwares* adequados e produtos adicionais que suportem todas as disciplinas de um processo de virtualização completo como monitoramento, produtos de migração (*live migration*) e gerenciamento de *hardware* e máquinas virtuais. O tamanho da solução, e das máquinas virtuais, deve ser tal que todo o *workload* dos sistemas origem possa ser comportado incluindo-se serviços complementares de redundância, recuperação de desastre e alta disponibilidade.

Por fim é uma boa prática a adoção de implementações por domínios específicos onde *workloads* conhecidos são mais ou menos semelhantes.

Na seção seguinte são apresentados alguns dos principais problemas e questões enfrentados na criação de processo de virtualização e algumas alternativas de solução.

4.3 Problemas, riscos e recomendações

Os riscos e problemas mais comuns enfrentados na virtualização dos *data centers* podem ser reduzidos e solucionados através ações específicas. A Tabela 3 sumariza essas questões e indica algumas abordagens possíveis. Esses problemas e riscos, e as abordagens menos óbvias, são discutidas a seguir.

Custos de entrada na tecnologia – Uma máquina física pode comportar dezenas de máquinas virtuais. A máquina física comporta a soma do *workload* de todas as máquinas virtuais e exige uma infraestrutura de rede e de armazenamento compatível. Custos adicionais de rede, *storage* e licenciamento de *software*, incluindo não só os MMV mas ferramentas periféricas, podem ser facilmente ignorados quando a virtualização tem foco somente na redução de máquinas e levar a um cálculo errôneo do custo da solução. Suponha a virtualização de 40 máquinas e considere 36GB de área de SAN para cada servidor a ser virtualizado (hoje os discos internos são maiores mas raramente são completamente ocupados), serão necessários 1.4TB de *storage* externo, o que, certamente requer planejamento e recursos. Soluções completas de virtualização, para que tenham retorno do investimento, exigem uma certa escala de máquinas elegíveis a virtualização que venha a compensar o investimento inicial.

Ausência de dados e métricas para recursos e serviços compartilhados – A determinação dos custos de infraestrutura de uma aplicação ou serviço é fundamental para se determinar o retorno efetivo de uma solução. O benefício deve justificar os custos e, portanto, precisa ser determinado. Embora a virtualização, com a otimização do uso de recursos, tenda a proporcionar custos de aquisição menores, a ausência de dados e métricas para recursos e serviços compartilhados não permite ainda determinar o custo do uso de recursos por aplicação ou seus usuários. A implementação de soluções de virtualização devem adotar padrões, de arquitetura, recursos, nomes etc. que facilitem a medida do consumo de recurso e o rateio entre as aplicações e usuários.

Problema/Risco comum	Abordagem/Mitigação
Custos de entrada na tecnologia	<ul style="list-style-type: none"> - Elaborar um caso de negócio com todos os custos envolvidos: servidores, dispositivos de armazenamento externo e rede, licenças de software etc. - Verificar que há um número de equipamentos suficiente para virtualização
Ausência de dados e métricas para recursos e serviços compartilhados	<ul style="list-style-type: none"> - Estabelecer padrões, como nomes de recursos, significativos para o negócio e usuários - Automatizar a medição do uso de recursos - Reportar periodicamente o uso de recursos aos usuários e envolvidos - Quando viável, segregar fisicamente ambientes de testes, QA e produção - Implementar métricas acuradas para os recursos alocados
Licenças de software	<ul style="list-style-type: none"> - Exigir dos fornecedores modalidades de licenciamento próprias para virtualização - Operar com modalidades de licença flexíveis que permitam proteger o custo presente mas permitam uma renegociação valores inferiores no futuro
Recursos de rede	<ul style="list-style-type: none"> - Garantir recursos de rede e banda suficientes para o tráfego de dados centralizado - Implementar tecnologias que permitam o acesso de máquinas virtuais a diferentes redes (<i>switches virtuais e multi port NICs</i>) em um mesmo <i>hardware</i>
Gerenciamento	<ul style="list-style-type: none"> - Implementar ferramentas de monitoração e gerenciamento dos MMV - Implementar soluções de gerenciamento da(s) máquina(s) física(s) como <i>RSA's</i> - Automatizar a monitoração das máquinas virtuais ou certificar que as soluções de monitoramento existentes não conflitam com a virtualização - Implementar processos para o gerenciamento de imagens ativas e desativas
Impactos humanos	<ul style="list-style-type: none"> - Fornecer treinamento, em nível apropriado, a técnicos, desenvolvedores, projetistas e gestores de TI em tecnologias de virtualização - Compartilhar da informação e dos benefícios da Virtualização na Organização
Suporte de fornecedores	<ul style="list-style-type: none"> - Assegurar que hardwares e softwares atendam a mesma matriz de compatibilidade - Em particular assegurar que <i>SO's</i>, <i>softwares</i> de <i>backup</i>, monitoramento e gerência de <i>storage</i> são compatíveis e podem operar com a virtualização sem restrições
Aumento da complexidade do ambiente	<ul style="list-style-type: none"> - Estabelecer padrões, como nomes de recursos, que atendam a operação técnica do ambiente - Quando viável, segregar fisicamente ambientes de testes, QA e produção
Questões de segurança	<ul style="list-style-type: none"> - Quando viável, segregar fisicamente ambientes de testes, QA e produção - Criar políticas e práticas de segurança específicas para o ambiente virtualizado
Compartilhamento de ambientes com diferentes níveis de serviço	<ul style="list-style-type: none"> - Quando viável, segregar fisicamente ambientes de testes, QA e produção - Elaborar um plano de consolidação baseado em dados de <i>workload</i> dos sistemas
Desempenho	<ul style="list-style-type: none"> - Empregar máquinas de processadores com novas tecnologias de virtualização (Intel, AMD, IBM) (para MMV de virtualização completa) - Estabelecer procedimentos de testes de aplicações para virtualização
Definição de <i>hardware</i> e capacidade	<ul style="list-style-type: none"> - Considerar o uso de máquinas tipo <i>Blade</i> ou máquinas de médio porte (até 8 processadores) - Considerar a criação de <i>farms</i> de várias máquinas físicas para virtualização - Definir uma arquitetura de <i>hardware</i> e um capacidade em conformidade com estratégias de emprego da virtualização para políticas de recuperação de desastre e alta disponibilidade - Realizar também um planejamento de capacidade de rede e <i>storage</i>

Tabela 3 – Questões e riscos envolvidos no emprego da virtualização e possíveis abordagens e ações.

Licenças de software – O custo de licenças de *software* apresentar incertezas ao longo do tempo. Muitos fabricantes ainda não tem uma política definida de preços que atendam a virtualização. No pior caso as licenças seriam cobradas por máquina virtual, mas isso, sendo as máquinas alocadas dinamicamente traria um custo excessivo e desnecessário a virtualização. Políticas por uso ou alocação ainda não são claramente presentes no mercado. Isso pode trazer custos inesperados para projetos ou aplicações com a mudança de política por parte de fornecedores.

Recursos de rede – Gigabit Ethernet NICs (*Network Interface Cards*) são essências na infraestrutura de virtualização de *data centers*[VMWare 2003]. Embora não faça uso completo da banda de rede cada servidor físico está normalmente associado a um *port* dedicado de rede de 10-100 ou 1000 Megabits. Suponha a consolidação de 40 servidores para os quais deseja-se garantir uma banda dedicada média de 100 Megabits. A máquina *host* necessitará uma conexão de 4 Gb. Essa necessidade é ainda maior quando diversas máquinas *host* são agrupadas em um único *farm* de virtualização e requerem que as interfaces de rede suportem múltiplos *ports* (*multiport* NICs). Boas práticas de arquitetura de virtualização ainda recomendam *ports* dedicados para monitoramento e migração de máquinas virtuais. Isso pode requerer um planejamento de capacidade adequado da rede envolvendo banda, *switch ports* e cabeamento.

Gerenciamento – Gerenciamento, monitoramento e controle do ambiente virtualizado são fatores críticos de sucesso. Falhas no *hardware*, no SO hospedeiro ou no MMV tornam-se críticas devido a concentração de máquinas virtuais. Ferramentas para gerenciamento e monitoramento automatizado de dezenas ou centenas de máquinas virtuais parece essencial. A maior parte dos fornecedores de virtualização vem desenvolvendo e ofertando soluções nessa

área. O monitoramento das máquinas virtuais pode, em geral, ser feito pelas soluções que gerenciam outros sistemas não virtualizados mas é necessário certificar-se que essas soluções são compatíveis com a virtualização. Soluções como o IBM RSA (*Remote Supervisor Adapter*) [IBM 2003] permitem ainda o controle remoto de múltiplos *hardwares*.

Impactos humanos – [Mann 2008] Essa é uma questão frequentemente negligenciada mas que é crucial para o sucesso de um processo de virtualização. Usuários e desenvolvedores encontram-se habituados a ter recursos bem definidos. Embora limitados esses recursos contam com uma capacidade ociosa conhecida e da qual normalmente, usuários e desenvolvedores, consideram-se proprietários e estão acostumados a considerar o seu uso de forma independente. A virtualização muda dramaticamente esse cenário. Seu objetivo é reduzir essa capacidade ociosa (sendo a média de utilização de cerca de 20% muitos casos encontram-se abaixo disso) distribuindo esse recurso a outras aplicações. Desse modo o uso de capacidades adicionais, embora possível, traz um custo adicional ao custo de operação normal, embora menor que no primeiro caso, e que precisa ser eventualmente negociado com outras áreas e usuários. Além disso os usuários e desenvolvedores precisam estar habituados a pensar na virtualização como uma alternativa sendo criativos para poderem tirar o máximo benefício dessa tecnologia para suas aplicações e processos. Por último processos de virtualização como a consolidação precisam frequentemente do apoio de usuários e desenvolvedores, para testes e verificações por parte dos aplicativos, o que requer que os benefícios da virtualização sejam bem claros para os envolvidos.

Suporte de fornecedores – É necessário ter certeza de que *hardware* e *softwares* suportam e são compatíveis com a MMV e a virtualização. Na virtualização completa a compatibilidade de *hardware* é mais óbvia enquanto para a paravirtualização é óbvia a necessidade de compatibilidade de *software* (SO). A incompatibilidade pode trazer resultados imprevistos. Por exemplo o uso de Windows NT em uma máquina virtual VMWare, embora pareça executar normalmente para um grande número de aplicações pode trazer uma degradação de desempenho que inviabiliza o uso do aplicativo. Ainda há questões do suporte dos fornecedores quando de incidentes. Para alguns produtos existe a perda de suporte quando o produto é empregado em máquinas virtuais. Isso pode ser significativo em especial para aplicações de missão crítica. Em particular deve-se assegurar que os SOs elegíveis a virtualização, *softwares* de backup, monitoramento e gerência de *storage* (incluindo-se *drives* e *placas*) são compatíveis e podem operar com a virtualização sem restrições.

Aumento da complexidade do ambiente – Uma camada adicional de *software*, o compartilhamento e a alocação dinâmica de recursos tornam o ambiente virtuais mais complexos que ambientes isolados e estáticos da não virtualização. É óbvia necessidade de gerenciamento, manutenção e monitoramento dos ambientes virtuais. Isso por si só já traz um novo desafio para as organizações. Menos óbvio, por exemplo, é a necessidade do gerenciamento de imagens, as quais podem requerer atualizações e processos novos de distribuição. Uma camada adicional também torna mais difícil o diagnóstico e determinação de problemas, sobretudo problemas relacionados a recursos compartilhados por diversas máquinas virtuais como *storage* e *network*.

Questões de segurança – Maior segurança é um dos benefícios a que se propõe a virtualização. Entretanto, uma vez que aumentamos a densidade de aplicações por equipamento, a segurança se torna um item de criticidade maior. Uma camada adicional de *software* e a manipulação de imagens traz novos pontos potenciais de falha. Uma falha ou ataque ao MMV pode significar danos a dezenas de máquinas virtuais.

Compartilhamento de ambientes com diferentes níveis de serviço – Uma vez que as máquinas virtuais com diferentes workloads compartilham alguns recursos comuns, como a interface de rede e *storage*, existe o potencial de interferência de uma aplicação sobre a outra, como um consumo excessivo de banda ou I/O por uma aplicação. Isso não deve ocorrer para recursos como processador e memória uma vez que estes estão limitados pela MMV.

Desempenho – A compatibilidade [Popek e Goldberg 1974] garante que as máquinas virtuais fornecem um ambiente para a execução de programas idêntico a máquina física. Essa compatibilidade, entretanto, não garante acesso a mesma quantidade de recursos da máquina física, em geral compartilhados, ou o desempenho. Os desenvolvedores devem ter conhecimento que o desempenho das aplicações pode ser afetado pela virtualização [Drepper 2008] ou mesmo apresentarem incompatibilidade [Lai 2008]. É portanto importante estabelecer-se procedimentos que validem aplicações para ambientes virtualizados, incluindo testes funcionais e carga. Atenção também deve ser dada a adoção de tecnologias de *hardware* que melhorem o desempenho de tecnologias de virtualização como o suporte de processadores Intel, AMD, IBM para virtualização [Rosenblum e Garfinkel 2005].

Definição de *hardware* e capacidade – Máquinas tipo *Blade* [Meier 2008] são associadas a um *chassis* que concentra vários equipamentos e fornece energia e cabeamento de modo a economizar espaço e consumo de energia sem entretanto reduzir o número de máquinas físicas. Esses equipamentos podem ser empregados em conjunto com a virtualização tornando ainda maior a concentração de máquinas. Entretanto uma maior densidade de máquinas, virtuais ou não, irá requerer uma concentração ainda maior de interfaces de rede de dados e de *storage*. Essa opção deve, portanto, ser ponderada evitando-se uma densidade crítica de recursos em um mesmo *rack*. O mesmo critério leva a opção de hospedar a virtualização em máquinas médias, de até 8 processadores, uma vez que máquinas maiores (16-32 processadores) podem levar a uma densidade de máquinas virtuais crítica tanto para recursos de *hardware* (interfaces e cabeamento) como *software* (limitações do MMV).

Embora comuns, todos esses fatores são cruciais para a implementação de um processo de virtualização consistente, robusto e de longo prazo.

5 Conclusão

As tecnologias de virtualização estão mudando dramaticamente a forma de trabalho das organizações de TI e em particular dos *data centers*. Ela traz reduções de custo significativas mas também traz novas oportunidades para a oferta de melhores serviços de TI. Não obstante seus benefícios a implementação de tecnologias de virtualização requer cuidados que reduzam potenciais impactos negativos da tecnologia e a adequação organizacional, técnica e de custos da empresa para fazer frente a dinâmica da virtualização e obter o máximo de seus resultados. Este trabalho apresenta os principais conceitos de virtualização, com foco maior na virtualização de máquinas x86, e propõe que as que, no médio e longo prazos, a opção por tecnologias de virtualização seja amparada por amplo *processo de virtualização* no *data centers*. Esse processo busca adequar a organização e inserir a tecnologia de virtualização nos serviços de *data center* os reduzindo os riscos dessa mudança e se constitui na principal contribuição deste trabalho. Futuramente esses processos serão explorados para outras modalidades de virtualização como a virtualização de aplicações e de computação pessoal que poderão trazer impactos ainda maiores as organizações de TI.

6 Referências

- Adams, K., Agesen, O., “A Comparison of *Software* and *Hardware* Techniques for x86 Virtualization”, *ACM ASPLOS* Outubro, 2006.
- Computerworld, “*Executive Briefings: Getting to Green*”, Strategic Insights from the Editors of Computerworld, 2008.
- Dittner, R., Rule, D., *The Best Damn Server Virtualization Book Period*, Syngress Publishing e Elsevier, 2007.

- Donston, D., "Virtualization: Survey says ...", *eWeek*, Vol. 25 Issue 5, 11/2/2008, 43-43.
- Drepper, U., "The Cost of Virtualization", *ACM QUEUE* January/February (2008), 30-33.
- Ferguson, S., "Adoption of Virtualization Continues to Grow: Report", <http://www.eweek.com/c/a/Infrastructure/Adoption-of-Virtualization-Continues-to-Grow-Report/>, (5 de Fevereiro de 2008), 12/2/2007a.
- Ferguson, S., "Companies embrace virtual technologies", *eWeek*, 22/10/2007b, Vol. 24 Issue 33, 28-30.
- Figueiredo, R., Dinda, P.A., Fortes, J., "Introduction: Resource Virtualization Renaissance", *Computer*, Volume 38, Issue 5 (2005), 28-31.
- Gillett, F.E., Schreck, G., "Server Virtualization Goes Mainstream", Forrester Research, 22/2/2006.
- IBM, *Virtualization and the On Demand Business*, IBM Red Paper, REDP9115, 2004.
- Lai, E., Thibodeau, P., "App Support Limitations Could Impede Virtualization", *Computerworld*, Vol. 42 Issue 2, 7/1/2008, 12-12.
- MacLaughlin, L. "Virtualization in the Enterprise Survey: Your Virtualized State in 2008", *CIO – Business Technology Leadership*, <http://www.cio.com/article/print/168401>, (1 de Abril de 2008), 2/1/2008.
- Mann, A. "The Pros and Cons of Virtualization", *Business Trends Quartely*, <http://http://www.btquarterly.com/?mc=pros-cons-virtualization&page=virt-viewresearch>, (23 de Abril de 2008), 2008.
- Meier, S., *IBM Systems Virtualization: Servers, Storage, and Software*, IBM Red Paper, REDP4396, 2008.
- Neiger, G., et al. "Intel Virtualization Technology: Hardware Support for Efficient Processor Virtualization", *Intel Technology Journal*, Volume 10, Issue 3, Agosto (2006), 167-178.
- Popek G.J, Goldberg R.P., "Formal Requirements for Virtualizable Third-Generation Architectures", *Communications of the ACM*, Julho (1974), 412-421.
- Rosenblum, M., Garfinkel, T., "Virtual machine monitors: current technology and future trends", *Computer*, Volume 38, Issue 5 (2005), 39-47.
- Smith, J.E., Nair, R. *Virtual Machines: Versatile Platforms for Systems and Processes*, Morgan Kaufmann, 2005a.
- Smith, J.E., Nair, R. "The Architecture of Virtual Machines", *Computer*, Volume 38, Issue 5 (2005b), 32-38.
- VMWare, "Virtual Infrastructure: Multiport NICs Are Critical When Consolidating Servers and Improving IT Management Through Virtualization", *VMWare White Paper*, http://http://www.vmware.com/pdf/vi_intel_vmware.pdf, (7 de Fevereiro de 2008), 2003.
- VMWare, "Virtualization: Arquitectural Considerations And Other Evaluation Criteria", *VMWare White Paper*, http://www.vmware.com/pdf/virtualization_considerations.pdf, (7 de Fevereiro de 2008), 2005.
- Vogels, W., "Beyond Server Consolidation", *ACM QUEUE* January/February (2008), 20-26.