



**Fundação Educacional do Município de Assis
Instituto Municipal de Ensino Superior de Assis
Campus "José Santilli Sobrinho"**

FERNANDO GUAZELLI SOUTO DE OLIVEIRA

VIRTUALIZAÇÃO DO ARMAZENAMENTO DE DADOS

**Assis
2011**

FERNANDO GUAZELLI SOUTO DE OLIVEIRA

VIRTUALIZAÇÃO DO ARMAZENAMENTO DE DADOS

Projeto de pesquisa apresentado ao Curso de Bacharelado em Ciência da Computação do Instituto Municipal de Ensino Superior de Assis – IMESA e a Fundação Educacional do Município de Assis – FEMA, como requisito parcial à obtenção do Certificado de Conclusão.

Orientando: Fernando Guazelli Souto de Oliveira

Orientador: Rita de Cássia Cassiano

**Assis
2011**

FICHA CATALOGRÁFICA

OLIVEIRA, Fernando Guazelli

Virtualização do Armazenamento de Dados / Fernando Guazelli Souto de Oliveira . Fundação Educacional do Município de Assis - FEMA -- Assis, 2011.

59p.

Orientadora: Rita de Cássia Cassiano Lopes.

Trabalho de Conclusão de Curso – Instituto Municipal de Ensino Superior de Assis – IMESA.

1. O Início da Virtualização . 2. A Computação na Nuvem . 3. Virtualização do Armazenamento de Dados.

CDD:001.6

Biblioteca da FEMA

VIRTUALIZAÇÃO DO ARMAZENAMENTO DE DADOS

FERNANDO GUAZELLI SOUTO DE OLIVEIRA

**Trabalho de Conclusão de Curso apresentado
ao Instituto Municipal de Ensino Superior de
Assis, como requisito do Curso de Graduação,
analisado pela seguinte comissão analisadora:**

Orientador: _____

Analisador (1): _____

Assis

2011

RESUMO

Este trabalho constitui-se em uma pesquisa bibliográfica sobre a tecnologia de Virtualização do Armazenamento de Dados. A pesquisa conta com a definição de Virtualização, Computação na Nuvem e Virtualização do Armazenamento de Dados, em que Virtualização é a capacidade de se executar dois ou mais Sistemas Operacionais distintos, simultaneamente, sob o gerenciamento de um monitor de máquina virtual denominado *Hypervisor*. Já, Computação na Nuvem, é um novo conceito de disponibilização de serviços, que permite obter inúmeras possibilidades por meio da internet. A Virtualização do Armazenamento de Dados, como o próprio nome diz, utiliza o recurso da Virtualização somado à Computação na Nuvem para possibilitar que os dados, informações sejam alocados em unidades de discos virtuais, e com isto, fornecendo vantagens como flexibilidade, escalabilidade, restauração de backup com alto desempenho, heterogeneidade de equipamentos, console de gerenciamento única, e etc. Este trabalho permite concluir que a Virtualização do Armazenamento de Dados é uma tecnologia que se aplica em empresas de portes pequenos até grandes, oferecendo vantagens e benefícios concretos.

Abstract

This work is the content of a bibliographical research about the Storage Virtualization technology data. The research counts on the definition of Virtualization, Cloud Computing and Storage Virtualization Data in which Virtualization is the capacity of achieving two or more different Operational Systems simultaneously, under the monitor management of virtual machine called Hypervisor. As Cloud Computing is a new concept of disposal of services that permits to get numerous possibilities by internet. The Storage Virtualization Data as itself, uses the resource of Virtualization plus Cloud Computing to make it possible that the data and information be kept in virtual disk drive, and together offering advantages as flexibility, scalability, backup restoration with high performance, heterogeneity of equipment, single management console and etc. This work permits to conclude that the Storage Virtualization Data is a technology that applies on small and even big companies offering advantages and concrete benefits.

LISTA DE ILUSTRAÇÕES

Figura 1.1 – Abstração de Processo em um Sistema Operacional.....	15
Figura 1.2 – Níveis de Hierarquia.....	16
Figura 1.3 – Tipos de Máquina Virtual, e seus relacionamentos entre Hóspede e Hospedeiro.....	18
Figura 1.4 – Modelos de Serviços nas Nuvens.....	29
Figura 1.5 – <i>Storage Dell EqualLogic OS6000</i>	41
Figura 1.6 – Virtualização de Bloco.....	46

SUMÁRIO

1 INTRODUÇÃO.....	09
2 O INÍCIO DA VIRTUALIZAÇÃO.....	11
2.1 A EVOLUÇÃO DA VIRTUALIZAÇÃO.....	11
2.2 O CONCEITO.....	13
2.3 SIMULAÇÃO, EMULAÇÃO, VIRTUALIZAÇÃO TOTAL E PARAVIRTUALIZAÇÃO.....	20
3 COMPUTAÇÃO NA NUVEM.....	24
3.1 HISTÓRIA.....	24
3.2 ESTRUTURAS DA NUVEM.....	26
3.2.1 Modelos de Serviços na Nuvem.....	27
3.3 PROBLEMAS AO UTILIZAR TECNOLOGIAS NA NUVEM.....	29
3.3.1 A Segurança da Informação na Nuvem.....	32
3.4 POR QUE UTILIZAR COMPUTAÇÃO NA NUVEM?.....	37
4 VIRTUALIZAÇÃO DO ARMAZENAMENTO OU VIRTUALIZAÇÃO DE STORAGE.....	39
4.1 DEFINIÇÃO E CONCEITO.....	41
4.2 TIPOS DE VIRTUALIZAÇÃO DE <i>STORAGE</i>	42
4.3 POR QUE UTILIZAR A VIRTUALIZAÇÃO DE <i>STORAGE</i> ?.....	46
4.4 FERRAMENTAS DE VIRTUALIZAÇÃO DE <i>STORAGE</i>	50
5 CONSIDERAÇÕES FINAIS.....	54

7 REFERÊNCIAS.....	56
---------------------------	-----------

1. INTRODUÇÃO

A tecnologia quebra barreiras a todo instante, novas soluções para a área da Tecnologia da Informação são apresentadas em ritmo acelerado. Os objetivos da busca por novas tecnologias e novas soluções em TI são as mesmas: melhoram a flexibilidade, confiabilidade, segurança, estabilidade, escalabilidade, e é a partir desses objetivos que surgem duas vertentes no ramo da TI, a Virtualização, e a Computação em Nuvem. Juntas, essas tecnologias renovarão os conceitos de TI, abrindo caminho para diversas descobertas que derivarão da Virtualização e da Computação em Nuvem, uma delas é a Virtualização do Armazenamento ou Virtualização de *Storage*.

A citação Computação em Nuvem, originalmente do inglês *Cloud Computing*, é utilizada para designar um novo padrão de utilização de recursos. Que consiste em deslocar toda infraestrutura alocada fisicamente para a rede, o que permite a utilização de recursos de *software*, *hardware*, plataformas de desenvolvimento, altamente flexíveis em razão dos recursos virtualizados. Pode-se dizer que o componente chave da Computação em Nuvem é a virtualização, e mais do que isso, que a Nuvem nada mais é do que *hardware* e *software* virtualizados acompanhados de monitoramento e planejamento prévios. A essência da Computação em Nuvem segue três conceitos básicos, que são: 1º a virtualização, que é a capacidade de um S.O (Sistema Operacional) pai disponibilizar recursos de *hardware* para um segundo S.O virtual, de forma que este S.O virtual não perceba a existência do S.O pai, o que significa que o usuário final ou usuário leigo não terá conhecimento da plataforma computacional que estará trabalhando; 2º a escalabilidade, que seria a alta flexibilidade do ambiente em razão da virtualização, permitindo o aumento ou a diminuição do uso de recursos quando for necessário; 3º o modelo *pay-per-use*, onde os recursos utilizados na Nuvem são pagos pelo consumo dos mesmos, e não como um produto adquirido (VAQUERO et al, 2009).

Ainda com o objetivo de generalizar especificamente o conceito de Nuvem, Vaquero (2009, p.1) define Nuvem como: “Uma grande fonte de recursos virtualizados que podem ser dinamicamente modificados com a finalidade de se obter melhor aproveitamento dos mesmos”.

A virtualização é uma camada de *software* que possibilita simular uma máquina física através de um sistema operacional já em execução, possibilitando infinitas alternativas de uso para quantas máquinas virtuais o *hardware* suportar, além de permitir ambientes de testes, desenvolvimentos controlados, laboratórios de estudo sem a necessidade de várias máquinas físicas. No início da utilização do conceito de virtualização, a ideia era fazer uso de um Sistema Operacional virtual (máquina virtual) sobre um Sistema Operacional diferente, citando como exemplo usuários do S.O Macintosh que criaram uma máquina virtual executando Microsoft Windows, o que possibilitou a utilização de todos os *softwares* compatíveis com plataforma Windows ao mesmo tempo em que utilizavam *softwares* próprios de plataforma Macintosh, cada *software* em execução no Sistema Operacional o qual é compatível (Kace Systems Management, 2007).

Com essas duas tecnologias decolando no mundo da TI, este trabalho estará analisando os frutos da junção dessas duas tecnologias, Virtualização e Computação em Nuvem. Em específico, este trabalho abordará a Virtualização de Armazenamento de Dados, ou como é de fato conhecida, Virtualização de *Storage*.

Este trabalho será dividido em Capítulos, que seguem:

Introdução

Capítulo 1 – O início da Virtualização. (Serão esclarecidas as raízes da virtualização, a história de como foi desenvolvido o conceito).

Capítulo 2 – A Computação em Nuvem. (Abrange definição, aprofundamento técnico da Computação em Nuvem).

Capítulo 3 – Virtualização de *Storage*. (Definição, exemplos de utilização).

Considerações Finais

Referências

O INÍCIO DA VIRTUALIZAÇÃO

2.1 – A EVOLUÇÃO DA VIRTUALIZAÇÃO

Segundo a empresa VMWare, pioneira no fornecimento de software de virtualização para plataformas x86, a Virtualização se iniciou nos anos 60, quando os mainframes já com alto poder de processamento computacional começaram a ser desvantajosos, uma vez que o tempo de ociosidade era demasiado. Com isso, surgiram as primeiras ideias que culminaram na formação do conceito de Virtualização. As empresas perceberam que era necessário aproveitar ao máximo o poder de processamento desses grandes computadores e que para isso seria necessário que vários processos fossem executados “ao mesmo tempo”.

Em busca do melhor aproveitamento de tempo de processamento, surgiu o conceito de *TimeSharing*, ou seja, tempo compartilhado. Cada processo é executado durante um período curto de tempo, no entanto, com a alta velocidade dos computadores tal procedimento é transparente ao usuário. Por fim, todos os processos são executados um por vez, por um curto espaço de tempo e repetidamente em sequência até ser finalizado ou concluído [VMWARE, 2011].

O conceito de *TimeSharing* ainda em uso nos dias atuais, foi lançado na década de 70 pela IBM um mainframe capaz de executar simultaneamente sistemas operacionais diferentes, sendo gerenciados por um programa de controle denominado *Hypervisor*. Com o modelo 370, a IBM foi a pioneira no desenvolvimento de um computador comercial inteiramente projetado para virtualização. Utilizando o sistema operacional CP/CMS, era capaz de executar múltiplas instâncias simultaneamente. Posteriormente, o sistema IBMz/VM aproveitava a virtualização por hardware de forma mais eficaz, já que todas as

interfaces de hardware eram virtualizadas. O VM/CMS é o sistema de virtualização mais popular, largamente utilizado academicamente [VMWARE, 2011].

Alguns anos depois, com as aplicações sendo executadas em plataforma cliente/servidor e declínio da utilização dos mainframes com a ascensão da plataforma x86, a virtualização caiu no esquecimento. Mais tarde, a larga usabilidade dos S.O Windows e Linux contribuíram para a padronização da plataforma x86 pela indústria, segundo a VMware [MASINA et. al, 2008].

O alto custo na aquisição de um mainframe, fez com que empresas tendessem a buscar alternativas, e foi na plataforma x86 que se iniciou outro processo de padronização. Com a plataforma x86, as empresas passaram a investir em vários servidores menores para fazer o mesmo trabalho de um grande mainframe, processo este chamado de *end-low*. Esta estratégia permitiu que empresas comprassem servidores conforme a necessidade [VMWARE, 2011].

A consequência do *End-Low* foi a ociosidade no processamento, já que as empresas superdimensionavam os servidores a fim de não terem problemas de dimensionamento de hardware. Desta maneira, segundo a *International Data Corporation*, o teto de uso de CPU dos servidores de plataforma x86, era estimado entre 10 a 15% da capacidade total. Por fim, estes servidores que portavam apenas uma aplicação, passaram a ter o mesmo problema dos mainframes, onde o alto poder de processamento era desperdiçado, uma vez que apenas 15% de sua capacidade total eram utilizados [VMWARE, 2011].

Em 1990, a VMWare Inc. introduziu o conceito de virtualização na plataforma x86, como uma possibilidade de melhor utilização do processamento dos servidores e até mesmo computadores domésticos [VMWARE, 2011].

Mais tarde, em 2005, as empresas de desenvolvimento de Hardware como a Intel e a AMD, lançaram tecnologias de melhor aproveitamento da virtualização, a Intel com a Intel-VT e a AMD-V. Estas tecnologias fornecem total interação dos sistemas virtualizados com o hardware, chamada (*full virtualization*) [VMWARE, 2011].

Já em 2007, a própria IBM trocou quatro mil servidores de pequeno porte por trinta mainframes Linux rodando máquinas virtuais. Com isto, a IBM economizou duzentos e

cinquenta mil dólares em manutenção e economia de energia [NEXT GENERATION CENTER, 2011].

Na década de 2010, a necessidade de sistemas mais leves, mais rápidos e que sustenta a demanda de processamento, fez com que a Virtualização retomasse o mercado tecnológico como uma solução inteligente, vantajosa, e que fornece confiabilidade e segurança, itens imprescindíveis no mercado de Servidores. Os pesquisadores juntamente com a indústria de tecnologia direcionam seus esforços na busca de tecnologias que forneçam segurança e facilidade no gerenciamento do ambiente, com a finalidade de fornecer acesso aos usuários em qualquer lugar, a qualquer hora [NEXT GENERATION CENTER, 2011].

2.2 – O CONCEITO

Para que esteja claro o conceito da virtualização, faz-se necessário o aprofundamento técnico desta tecnologia. Primeiramente, existe uma linguagem padrão ao se tratar de Virtualização: Máquina ou Servidor hospedeiro é o equipamento em que se aplica a virtualização, onde o sistema operacional é executado diretamente no hardware. Já o sistema Convidado não possui esta restrição, podendo ser executado sobre o hospedeiro como uma máquina virtual. Diferentemente do Hospedeiro, o sistema Convidado pode ser mais de um sendo executado sobre o mesmo hospedeiro, inclusive com sistemas operacionais distintos. Tais definições foram obtidas no portfólio on-line da Hewlett-Packard (HP).

Segundo Coelho, Calzavara e Di Lucia et al (2008), a virtualização nada mais é do que a emulação de ambientes isolados, podendo executar dois ou mais sistemas operacionais distintos em uma única máquina física, usufruindo toda capacidade de processamento

que o hardware dispõe. Além disso, a virtualização visa o proveito do máximo processamento, já que estaria diminuindo o tempo de ociosidade do hospedeiro.

É importante concretizar a independência do Hospedeiro com Convidado. Segundo a Hewlett-Packard (HP), uma das maiores empresas de TI, além do hardware, nada mais é comum entre os dois sistemas. Se tratando dos sistemas Convidados, não existe ligação lógica entre eles, possibilitando a utilização de diversos sistemas operacionais iguais, similares ou distintos, sendo apenas limitados pelo software que provê a virtualização. Normalmente, se um sistema operacional pode ser instalado em uma máquina física, ou hospedeiro, também poderá ser instalada em uma máquina virtual como um sistema convidado.

A Virtualização é uma forma de criar um ambiente de pequenos sistemas isolados, com facilidade de gerenciamento avançada, contribuindo para a segurança dos próprios sistemas. Não apenas isto, a Virtualização inclui todas as plataformas, desde aplicativos até sistemas operacionais, processadores, interconexões de plataforma, storage componentes [NEXT GENERATION CENTER, 2011].

Com diversas possibilidades de se aplicar o conceito de Virtualização, surgem vários cenários onde este conceito é utilizado com eficácia. Algumas das aplicações mais utilizadas pelas empresas são: Virtualização de Plataforma e Virtualização de Storage [NEXT GENERATION CENTER, 2011].

Para que os diferentes tipos de Virtualização sejam devidamente entendidos, alguns conceitos básicos relacionados ao Sistema Operacional serão esclarecidos, que seguem:

- Processos: Segundo [CARISSIMI, 2011], um processo nada mais é do que execução abstrata de um programa pela CPU (Central Processing Unit) ou Unidade Central de Processamento. Um processo possui quatro regiões distintas que são: *heap*, texto, dados e pilha, onde a região de *Heap* é responsável pela alocação dinâmica de porções de memória, a região de texto aloca as instruções a serem executadas pelo programa, já a região de dados mantém todas as variáveis globais, inicializadas ou não. A região de pilha é responsável por manter variáveis locais, além de alocar o endereço de retorno de chamadas de função e possibilitar a passagem de parâmetros. (Figura 1.1 – Processo-1)

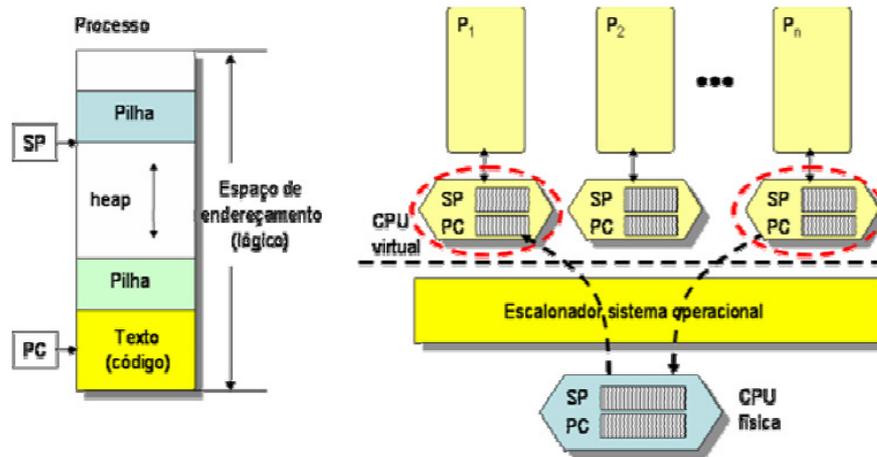


Figura 1.1 – Abstração de Processo em um Sistema Operacional.

- Threads: Segundo Rosely Scheffer, Threads ou processo leve, é uma entidade básica de utilização de CPU, que consiste em um apontador de instrução, conjunto de registradores e espaço de pilhas. As Threads não exigem alocação em memória de disco, pois estão totalmente carregadas em memória principal ou memória RAM (*Random Access Memory*) ou Memória de Acesso Randômico. Todo Sistema Operacional possui pelo menos uma Thread em execução, a *Main Thread* ou Thread Principal. Diferentemente dos Processos, as Threads não são independentes umas das outras, sendo possível invadir o espaço de outra a invalidando, não sendo protegidas umas das outras. A utilização de Threads possibilita o paralelismo, embora a CPU execute uma Thread por vez assim como um Processo, uma das diferenças está na utilização de Threads Múltiplas, onde uma Thread bloqueada não impede a execução de outra de uma mesma tarefa, ou várias Threads de uma mesma tarefa sendo executadas, fazendo com que o I/O (Input/Output) ou Entrada e Saída de dados seja maior, possibilitando maior desempenho.

- Estrutura em Camada Hierárquica: Com a evolução do Hardware acelerada e indiscutivelmente veloz, vários softwares também surgem, fazendo com que a complexidade no gerenciamento da relação entre hardware e software seja possível, eficaz, confiável e performática. Porém, isto não foi fator limitante das tecnologias em desenvolvimento, graças à estrutura em camada hierárquica dos Sistemas Operacionais. O uso de abstração e de interfaces permitiu que cada subsistema fosse visto como sendo independente, oferecendo serviços aos demais. No entanto, os detalhes de cada subsistema não precisam se expostos, sendo necessário apenas o conhecimento das interfaces e dos serviços oferecidos. Em suma, existem três tipos de interfaces, que são: instruções de máquina (privilegiadas), instruções de máquina (não-privilegiadas), chamadas de sistema e interface aplicativa de programação [CARISSIMI, 2011].

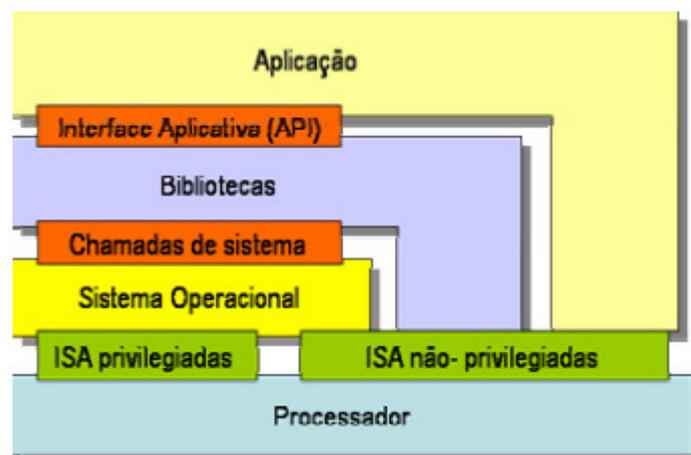


Figura 1.2 – Níveis de Hierarquia

Descrevendo a forma com que a primeira interface (Instruções de Máquinas com Nível Privilegiado) se comporta, o mesmo autor afirma que a unidade central de todo computador é o processador, este segue um conjunto de instruções de máquina ISA (*Instruction Set Architecture*), como por exemplo, o padrão utilizado por processadores Intel e AMD, denominado Intel IA-32 ou x86. Portanto, os softwares são compilados de forma que seus arquivos binários sejam compatíveis com o padrão ISA de instruções de máquina. Na verdade, as instruções de máquina ainda são divididas em dois grupos de níveis de acesso ao Hardware, o nível privilegiado, e o nível não-privilegiado. O nível privilegiado garante que sejam

executadas instruções especiais, como por exemplo, a de interrupção, utilizada pelos Sistemas Operacionais. Já o nível não-privilegiado é utilizado pelos processos de usuários, que não requerem a execução de instruções especiais [CARISSIMI, 2011].

A segunda interface como já dito, consiste nas instruções de máquinas não-privilegiadas e pelas chamadas de sistema. Essas instruções possibilitam que os processos não-privilegiados de usuários, executem instruções não-privilegiadas diretamente no processador, porém com restrição aos recursos de hardware (instruções privilegiadas. As chamadas de sistema é uma forma controlada e segura de fornecer aos processos de usuários recursos de hardware, fazendo com que as operações costumeiramente chamadas de E/S (entrada e saída) sejam autenticadas e validadas antes de ocorrerem [CARISSIMI, 2011].

Já a terceira interface consiste nas chamadas a funções de bibliotecas, que na realidade, são conhecidas informalmente por *API – Application Program Interface*, ou Interface Aplicativa de Programação [CARISSIMI, 2011].

Levando-se em conta os conceitos esclarecidos e principalmente as três interfaces abordadas enquanto se tratando de Níveis de Hierarquia, a implementação de Máquinas Virtuais podem ser feitas de duas maneiras. Primeiramente, é possível desenvolver um programa de aplicação que forneça um ambiente de execução para outras aplicações. Esse ambiente pode gerar instruções abstratas que permitem a execução de instruções não-privilegiadas, chamadas de funções e as *APIs* de bibliotecas que corresponderiam as ações abstratas desejadas. Este conceito é utilizado pela Máquina Virtual Java ou *Java Virtual Machine* (JVM). É possível também fazer um programa de aplicação, que emule chamadas de sistemas de outro sistema operacional, na prática isto se faz presente quando se executa um sistema operacional Linux em um sistema operacional Windows utilizando, por exemplo, *VMware Player* ou *Virtual Box*. Este tipo de virtualização pode ser demonstrado na figura 1.3 (a) [CARISSIMI, 2011].

Outra maneira de se implementar uma Máquina Virtual, é fornecer uma camada de software entre o hardware e o sistema operacional, o qual terá como uma de suas funções, prover segurança no acesso da camada de software aos recursos físicos da

máquina. Esta camada oferece como interface ao sistema operacional um conjunto de instruções de máquina que podem ser idênticas às utilizadas pelo processador, ou outra. O diferencial, é que esta camada de software deverá estar disponível já na inicialização do sistema operacional *Host*, ou seja, deverá estar disponível sempre que o computador estiver ligado, assim como possibilite a utilização de vários programas diferentes ou iguais simultaneamente. Esta máquina virtual é comumente chamada de Monitor de Máquina Virtual, ou *VMM-Virtual Machine Manager*, ou ainda *Hypervisor*. Com isso, é possível executar vários programas (Sistemas Operacionais) diferentes em uma mesma plataforma. A figura 1.3 (b) demonstra este tipo de Máquina Virtual [CARISSIMI, 2011].

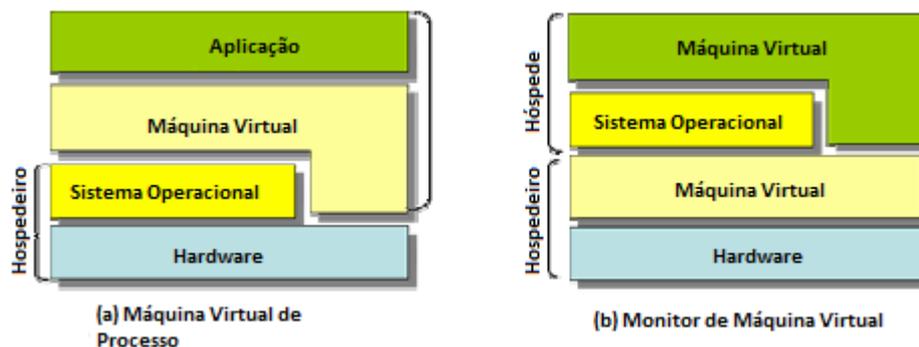


Figura 1.3 – Tipos de Máquina Virtual, e seus relacionamentos entre Hóspede e Hospedeiro.

Devemos destacar que um processo só existe enquanto um programa está em execução. Portanto, uma Máquina Virtual de Processo só existe enquanto estiver executando. Já o Monitor de Máquina Virtual está presente desde a inicialização do Sistema Operacional Hospedeiro [CARISSIMI, 2011].

Considerando que as Máquinas Virtuais dependem de acesso direto ao processador físico da máquina, o mesmo precisa oferecer suporte à Virtualização, provendo mecanismos básicos para que os sistemas virtualizados consigam utilizar os recursos de hardware da mesma maneira como o sistema operacional hospedeiro. Por isto, Popek e Goldberg (1974) citaram três propriedades necessárias para que sistemas computacionais oferecessem suporte eficiente à virtualização, que seguem:

Eficiência: todas instruções de máquina que não comprometem o funcionamento do sistema devem ser executadas diretamente no hardware sem intervenção da máquina virtual.

Controle de recursos: uma máquina virtual deve ter controle completo sobre os recursos virtualizados sendo estritamente proibido que um programa executando sobre a máquina virtual os acesse diretamente.

Equivalência: um programa executando sobre uma máquina virtual deve exibir um comportamento idêntico àquele apresentado caso a máquina virtual não existisse e o programa acessasse diretamente uma máquina física equivalente. Duas exceções são consideradas. Primeira, eventualmente, algumas instruções podem ter seu tempo de execução aumentado. Segunda, pode haver problemas de conflito de acesso a recursos, os quais devem ser resolvidos de forma apropriada.

Essas propriedades foram traduzidas no conjunto de instruções de máquina (*ISA*) em três grupos e dois teoremas. Os três grupos são: Instruções de máquina privilegiadas, que se executadas em modo usuário geram em consequência as *Traps* ou exceções. Sensíveis ao controle, que permitem alterações nos recursos do sistema. Sensíveis comportamentais, que dependem das configurações de registradores internos ou modos de execução do processador [CARISSIMI, 2011].

Antes de mencionar os dois teoremas, o entendimento básico da arquitetura x86 se faz necessário. Primeiramente, a arquitetura x86 estabelece quatro modos de operação do processador que são chamadas de anéis de proteção (*rings*) ou *CPL* (*Current Privilege Level*) que são identificados de 0 a 3. Em sistemas operacionais Windows e Linux (Unix), apenas dois modos são utilizados, *ring 0*, que possui privilégios avançados e é utilizado pelo sistema operacional, e o *ring 3* que detém privilégios restritos e é utilizado por processos de usuários. Caso um processo tente executar uma instrução em nível privilegiado, ocorrerá uma exceção (*trap*) que deverá ser tratada adequadamente.

Segundo [CARISSIMI, 2011] o primeiro teorema afirma que “um monitor de máquina virtual (VMM) poderá ser implementado se as instruções sensíveis de controle e comportamentais forem um subconjunto das instruções privilegiadas”, o que quer dizer

que qualquer instrução que afete o VMM, deverá passar por seu controle. Já o segundo teorema, cita a possibilidade de uma VMM executar uma instancia de si mesma.

No entanto, segundo o mesmo autor [CARISSIMI, 2011], o primeiro teorema é falho, pois é possível a implementação da virtualização em arquiteturas que não o respeitam em detrimento do desempenho. Duas técnicas são costumeiramente utilizadas por outras arquiteturas para prover a virtualização. A primeira técnica consiste em um desvio das instruções sensíveis para a VMM em tempo de execução, de forma que a VMM as trate adequadamente. A segunda técnica é aplicada de forma que o programa a ser executado na VMM é modificado para que as instruções sensíveis sejam substituídas para chamadas a VMM, técnica esta conhecida também como paravirtualização.

Os fabricantes de processadores AMD e Intel, no entanto, desenvolveram extensões para a arquitetura x86 suportar a virtualização. As duas tecnologias apresentadas pelas fabricantes foram a AMD-V (*AMD-Virtualizatio*) e Intel-VT (*Virtualization Technology*). Estas soluções foram desenvolvidas independentemente e são incompatíveis entre sí, embora sirvam para o mesmo fim. As diferenças são simples, enquanto a AMD implementa funções avançadas no processador cuja a execução é feita pelo *Hypervisor* e que podem controlar em seu nome. Já a Intel, introduziu mecanismos similares ao conceito de *Rings*, como já visto na arquitetura x86, onde existem dois níveis de segurança *root* e *não-root* que são utilizados pelo *Hypervisor* para definir as permissões de execução do hóspede sem que exista risco para o sistema hospedeiro [CARISSIMI, 2011].

2.3 – SIMULAÇÃO, EMULAÇÃO, VIRTUALIZAÇÃO TOTAL E PARAVIRTUALIZAÇÃO

Em razão da grande variedade de cenários ou ambientes onde a Virtualização é aplicada, há certa confusão sobre a diferença entre Simulação, Emulação, Virtualização,

Virtualização Total e Paravirtualização. Por isto, será feita uma breve descrição comparativa entre as duas tecnologias, que segue:

- Simulação: Do dicionário de língua portuguesa Houaiss, Simulação é o ato de representar com semelhanças, fingir ou imitar. Computacionalmente, isto significa que de forma abstrata ou física, um modelo matemático pode representar características necessárias para o entendimento de seu comportamento e, através disto, disponibilizar entrada de dados e/ou ações para que seja possível obter predição e análise de resultados em ambientes simulados. Os simuladores estão em uso em diversas áreas de estudo, como a química, física, dentre outros. A simulação interativa é largamente utilizada em simuladores de vôo e jogos de computador, onde um agente externo humano realiza a entrada de dados ou ações [CARISSIMI, 2011].
- Emulação: A Emulação consiste em reproduzir o comportamento exato e as funcionalidades de um programa, utilizando outro. Com isso, as entradas fornecidas ao programa emulado devem fornecer as mesmas saídas para o mesmo sistema em um ambiente não emulado, e com isso, validar que tal emulação reproduz totalmente as funcionalidades e processos do sistema quando não executado em cenário emulado. Um exemplo de software de emulação é o *Open Source* Basilisk II, que emula o Sistema Operacional Apple Macintosh. O mesmo pode ser executado em Sistema Operacional Windows, Linux, MAC OS X [CARISSIMI, 2011].
- Virtualização: É a tecnologia que possibilita a distribuição de recursos de hardware propriamente dito a sistemas operacionais (iguais ou diferentes) executando simultaneamente sobre o mesmo hardware através de um software de monitoramento denominado *Hypervisor*. Com esta tecnologia, vários Sistemas Operacionais podem ser executados simultaneamente sobre o mesmo hardware sendo que cada um consiga ter acesso ao hardware em si e desta forma, obter maior desempenho e confiabilidade. Além disto, esta técnica pode ser aplicada a

diversos ambientes, não somente com Sistemas Operacionais. Um exemplo é a Virtualização de *Storage*, que será analisada posteriormente neste trabalho [NEXT GENERATION CENTER, 2011].

- Virtualização Total: Segundo [CARISSIMI, 2011] esta técnica de virtualização é empregada na implementação de monitores de máquinas virtuais (VMM) ou máquinas virtuais de sistema e consiste em fornecer uma réplica virtual do hardware subjacente de forma que o sistema operacional e aplicações do sistema operacional hóspede possam executar diretamente sobre o hardware original. A vantagem desta técnica é que não existe necessidade de alteração do sistema operacional hóspede para ser executada sobre a VMM, no entanto, existem inconvenientes como a incompatibilidade, queda no desempenho.

As desvantagens surgem primeiro, em razão da tamanha diversidade de dispositivos existentes que compõe um computador, logo, é difícil desenvolver uma máquina virtual que imite o comportamento de cada tipo de dispositivo. Em segundo lugar, todas as instruções do sistema hóspede são testadas pelo VMM para que seja definida qual o nível de permissão necessário, se são sensíveis ou não, o que demanda processamento. Não menos importante, um dos grandes empecilhos da virtualização total é relacionada à alocação de memória, sistemas como Windows e Linux implementam o conceito de memória virtual através da paginação, o que também é necessário em sistemas hóspedes, gerando uma série de adaptações necessárias feitas pelo *Hypervisor*, gerando processamento adicional e queda de desempenho [CARISSIMI, 2011].

- Paravirtualização: O autor [CARISSIMI, 2011] explica que a paravirtualização é uma abordagem alternativa à virtualização total, com o objetivo de contornar as desvantagens da mesma. O termo *Hypercall*, ou chamada de *Hypervisor*, é definido pelo autor como sendo a “substituição da chamada de uma instrução sensível por um tratador de software (*trap*) com uma parametrização adequada de registradores”. Isto acontece toda vez em que o sistema hóspede executa uma

instrução ou ação considerada sensível, fazendo com que o teste feito pela VMM não seja mais necessário. Outra diferença considerável em relação à virtualização total, é que o hardware é acessado por *drivers* fornecidos pela própria VMM.

No entanto, se faz necessário esclarecer que para a utilização da virtualização total, o sistema operacional hóspede não necessita de qualquer alteração para que seja executada sobre uma VMM, ao passo que a paravirtualização exige modificações por parte do sistema hóspede para que “enxergue” a VMM e através disso utilize o hardware apropriadamente.

3 – COMPUTAÇÃO EM NUVEM

3.1 – HISTÓRIA

O conceito de Computação em Nuvem, embora pareça novo e recente, foi idealizado na década de 1960 pelo cientista de inteligência artificial John McCarthy, em uma oportunidade que o mesmo ministrava uma palestra, a qual já propunha o modelo de computação como um serviço distribuído semelhantemente a outros como água, energia, etc [RYDLEWSKI et al, 2009].

Seguindo a mesma linha de raciocínio de John McCarthy, um dos responsáveis pelo desenvolvimento da ARPANET (*Advanced Research Projects Agency Network*) Joseph Carl Robert Licklider, introduzia a idéia de uma rede de computadores global em que todos estariam conectados entre si, acessando programas, dados e serviços de qualquer lugar [MOHAMED et al, 2009].

Em 1999, com o surgimento da Salesforce.com, que passou a ser pioneira na disponibilização de aplicações empresariais através da internet, a Computação em Nuvem começou a ganhar mais espaço e alavancada pelas políticas de mercado, onde as concorrentes ao perceberem a forte adesão desta nova tendência, passaram a adotar a tecnologia para não cair na obsolescência. Grandes empresas como: Amazon, Google, IBM e a Microsoft também passaram a adotar a Computação em Nuvem [MOHAMED et al, 2009].

Já em 2002, a Amazon desenvolveu o Amazon Web Services, que proveu um leque de serviços na Nuvem como armazenamento ou *storage*, computação e inteligência artificial através da *Amazon Mechanical Turk* [MOHAMED et al, 2009].

No ano de 2006, a Amazon lançou a *Elastic Compute Cloud* ou Nuvem de Computação Elástica, a qual permitia que pequenas empresas alugassem computadores que executariam suas próprias aplicações, na Nuvem [MOHAMED et al, 2009].

Atualmente a Computação em Nuvem está definitivamente difundida, havendo inúmeras empresas oferecendo serviços e aplicações de diversos ramos de atividades, todas na Nuvem. Com a Web 2.0, empresas como Google passaram a oferecer serviços na Nuvem via browser, como é o caso do Google App, que consiste em uma série de ferramentas disponibilizadas via browser. Segundo Dan Germain, da empresa *Cobweb Solutions* que presta serviços no ramo de T.I (Tecnologia da Informação):

A contribuição mais importante da computação na nuvem tem sido a grande quantidade de aplicações sendo disponibilizadas pelas principais gigantes da tecnologia, como Google e Microsoft. Quando estas companhias entregam serviços confiáveis e fáceis de consumir, o efeito é uma generalização do aceite de serviços prestados online [MOHAMED et al, 2009].

Jamie Turner, pioneiro na Computação em Nuvem, revela que o amadurecimento da Virtualização, o desenvolvimento de interfaces universais de alta velocidade e padrões de interoperabilidade de *softwares*, foram os fatores chave para que a tecnologia de Computação em Nuvem evoluísse. Turner ainda acrescenta:

A Computação em Nuvem vai além de usuários leigos que utilizam Google Docs, nós podemos apenas começar a imaginar a abrangência e o alcance desta tecnologia. Praticamente, qualquer coisa pode ser entregue a partir da Nuvem [MOHAMED et al, 2009].

No ano de 2008, o site de notícias BBC informou que as gigantes da tecnologia HP, Yahoo e Intel firmaram uma parceria para promover com maior ênfase a Computação em Nuvem e seus serviços. Oferecendo armazenamento on-line e um leque de serviços para arquivos e equipamentos que poderiam ser utilizados “dentro da nuvem” [BBC, 2008].

3.2 – ESTRUTURAS DA NUVEM

A Computação em Nuvem é formada por *Data Centers*, que nada mais é do que grandes centros de processamento de dados, ou seja, um aglomerando de enormes quantidades de máquinas com alta capacidade de armazenamento e processamento trabalhando juntas [RYDLEWSKI et al, 2009].

Os *Data Centers* já estavam presentes antes da Computação em Nuvem, no entanto, a localização era necessariamente estratégica, uma vez que a taxa de transmissão de dados não era alta. Após a fibra ótica e a evolução da internet, a localização dos parques tecnológicos passou a ser irrelevante, sendo apenas uma questão de disponibilidade de energia elétrica e investimentos para custeá-los. Com a Virtualização, os *Data Centers* passam por uma reformulação, fazendo com o que a utilização de seus recursos sejam melhor aproveitados, fornecendo total elasticidade de demandas [RYDLEWSKI et al, 2009].

3.2.1 – MODELOS DE SERVIÇOS NA NUVEM

De acordo com Taurion et. al (2009) a Computação em Nuvem traz uma releitura de termos como rede, computação, internet, conectividade, acessibilidade e racionalização, que são vinculados com as propostas de virtualização e terceirização, podendo apresentar três formas:

- SaaS (*Software as a Service* – Software como Serviço): Esta forma de Nuvem pode disponibilizar vários tipos de aplicativos que podem ser pequenos e simples como planilhas de cálculos, até aplicativos comerciais ou edição de imagens. Um exemplo é o Google Docs, que oferece uma linha de aplicativos nomeados *offices* gratuitamente, sendo necessário apenas acesso à internet e um *browser* para usufruir das ferramentas [TAURION et al, 2009].

Outros exemplos de aplicativos disponibilizados pela Nuvem são: Web como Zoho (processamento de texto, planilha), Salesforce (CRM), SlideRocket (apresentações) ou serviços da Web como Google Search, Yahoo! Weather ou PayPal [IBM, 2011].

Segundo Martins (2011), com este modelo de prestação de serviços, o cliente não precisa se preocupar com qualquer tipo de inconveniente relacionado aos servidores que suportam as aplicações, à manutenção, atualização, enfim, todos os problemas, gerenciamentos, correções, são transparentes ao cliente.

O mesmo autor cita que segundo analistas de mercado, o crescimento pela adesão do modelo de SaaS e a alta pressão de reduzir custos de T.I são os maiores responsáveis pelo crescimento do SaaS, que automaticamente alavanca a evolução da Computação em Nuvem.

- IaaS (*Infrastructure as a Service* – Infraestrutura como Serviço): O autor Taurion et al (2009) afirma que IaaS consiste em abrigar grandes quantidades de máquinas em *Data Centers*, possibilitando oferecer ao cliente acesso aos servidores e, com isto, dispensar investimento em servidores locais, custos com refrigeração, energia,

manutenção, e com a vantagem de poder fazer upgrade em *real-time* ou tempo real, do hardware contratado.

Já Martins (2011) é mais sucinto, afirmando que o cliente tem total acesso às tecnologias adquiridas quando contrata serviços do modelo IaaS, como por exemplo, o próprio cliente provisiona o ambiente tal como, processamento, *storage*, rede e recursos computacionais fundamentais. Com isto, o cliente está apto a implantar e executar os softwares pertinentes, incluir sistemas operacionais e aplicações, sendo gerenciados por ele mesmo. No entanto, o mesmo não gerencia ou controla os recursos por trás dessa infraestrutura; contudo, tem possibilidade de controle ilimitado a componentes como *storage*, *firewall*, ativos de rede, por exemplo.

Alguns exemplos de serviços de IaaS são: *Amazon Web Services*, com o *Elastic Cloud (EC2)* para processamento, o *Simple Storage Service (S3)* para armazenamento e *Joyent* o qual provê uma infraestrutura sob demanda com possibilidade escalável [MARTINS et al, 2011].

Segundo Foster (2008), ao invés de vender infraestrutura de hardware, os provedores de IaaS oferecem infraestrutura virtualizada como serviço aos clientes, dessa forma, os recursos de hardware disponibilizados são abstraídos e encapsulados fornecendo à próxima camada (usuários finais), uma interface padronizada que unifica os recursos na forma de IaaS.

- PaaS (*Plataform as a Service – Plataforma como Serviço*): Permite que as empresas usufruam do poder de processamento e armazenamento da Computação em Nuvem de acordo com a demanda, personalizando suas necessidades. A PaaS geralmente são usados por desenvolvedores para escrever suas aplicações de acordo com as especificações de uma plataforma, sem se preocupar com a camada de infraestrutura de hardware (IaaS) [MARTINS et al, 2011].

Abaixo, uma figura que ilustra as relações entre as formas de serviços oferecidos pela Nuvem:

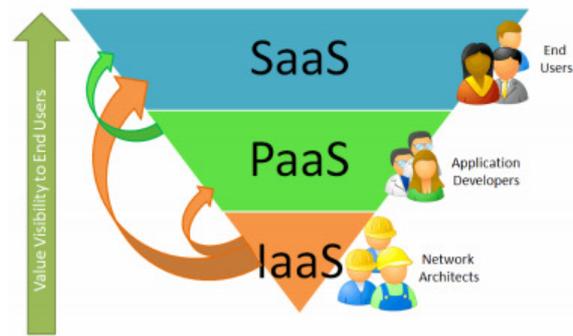


Figura 1.4 – Modelos de Serviços nas Nuvens.

3.3 – PROBLEMAS AO UTILIZAR TECNOLOGIAS NA NUVEM

A conexão com a rede mundial de computadores, a internet, já faz parte da vida cotidiana das pessoas, mas até hoje muito mais como uma opção do que como uma dependência substancial. No entanto, este cenário muda quando visto pelo ponto de vista empresarial. As empresas buscam formas de redundância, ou seja, maneiras alternativas para possibilitar acesso às informações ou serviços em caso de queda da conexão. Sem conexão, não existe Nuvem, em consequência, também não estarão disponíveis os serviços utilizados através da mesma. Com esta preocupação relacionada à conectividade, surgem custos duplicados para que a redundância e a garantia de conexão existam. Vale lembrar que não é só a desconexão que prejudica o dia-a-dia das empresas, mas também as oscilações e quedas nos serviços fornecidos pela Nuvem [TRAPANI, 2011].

A migração de ambientes para a Nuvem podem trazer diversos riscos, principalmente quando se tratando de empresas que não possuem estudo refinado de provisionamento, fazendo com que a chance de erro no dimensionamento da infraestrutura a ser utilizada na Nuvem seja alta. Além disto, empresas que estudam a possibilidade de migração de

serviços para a Nuvem devem maximizar gerenciamento e eficiência [MARTINS et al, 2011].

Uma questão importante se relaciona a velocidade de conexão, que é a capacidade de manter-se à frente do tempo e, conseqüentemente dos concorrentes, através de maior desempenho em transações. Os processos em andamento dentro das empresas são altamente dependentes de tempo de resposta, isso quer dizer que a velocidade de conexão influencia fortemente na resposta de transações e conseqüentemente se torna fator limitante ou decisivo para a empresa em questão manter-se à frente ou não, dos concorrentes. Levando em consideração esta variável (tempo de resposta), muitas empresas ainda não estão aptas a acompanhar e se readequar na velocidade com que a tecnologia evolui, e por este motivo, mantém a estrutura tradicional de aquisições de ativos de infraestrutura, trazendo consigo implicações negativas que comprometem a agilidade de provisionamento [MARTINS et al, 2011].

Segundo Callegari (2011), a segurança continua sendo o principal tema de discussão entre os especialistas em tendências de mercado, quando se tratando de Computação em Nuvem. Dados informados pela *Symantec* confirmam que a segurança é um fator a ser considerado pelas companhias para migrar para ambientes na Nuvem. Para reforçar isto, a *Symantec* intitulou em 2011 um estudo, que ficou conhecido como *State of Security*. O autor afirma que “uma área de investimento de T.I é a Nuvem”, e ressalta que “a nuvem privada ainda está à frente da pública e a proteção do ambiente é ponto crítico para as companhias” [CALLEGARI, 2011].

Este estudo, que foi encomendado pela *Symantec* à *Applied Research*, mostra que a adesão da Virtualização e da Computação em Nuvem têm gerado novos desafios. As empresas esperam uma série de benefícios quando adotada a virtualização e a computação na nuvem, como por exemplo, redução de custo, desempenho diferenciado, facilidade de recuperação em caso de desastres entre outros, no entanto, a pesquisa constatou que a tecnologia não consegue entregar o que é esperado. O gerente de engenharia de sistemas da *Symantec* Brasil explica:

Existe uma discrepância entre as expectativas e a realidade e isso tem haver com a maturidade das organizações, com o conhecimento que o cliente tem da tecnologia. Essa falta de conhecimento gera uma expectativa [aquém do que eles esperam], [CALLEGARI, 2011].

Seguindo o mesmo raciocínio, há uma terceira conclusão, de que as empresas querem uma garantia pelo serviço, ou seja, querem melhor desempenho com a solução de Computação em Nuvem e Virtualização. Quem migrou da solução *in house* (onde a infraestrutura de servidores está presente fisicamente na rede local) para solução virtual, tem que obter desempenho pelo menos igual [JEDRAS, 2011].

Um dos grandes empecilhos para a adoção em massa da Computação em Nuvem se dá ao investimento em hardware e software já feitos pelas empresas. Os fatores relacionados aos softwares que foram desenvolvidos para serem executados em infraestrutura local impede a migração para a nuvem. Para que isso se torne realidade para as empresas, esses softwares deveriam ser reescritos, o que não é interessante, como por exemplo, para bancos e grandes corporações [JÚNIOR et al, 2010].

No entanto, existe uma variável que está crescendo na contramão da ampla adoção da Computação em Nuvem, que é a diminuição considerável do custo de hardware nos últimos anos, tornando possível a utilização desta tecnologia em ambiente privado [JÚNIOR et al, 2010].

3.3.1 – A SEGURANÇA DA INFORMAÇÃO NA NUVEM

Como já dito anteriormente neste trabalho, o maior fator limitante à migração para ambientes na Nuvem, é a segurança. Neste ponto, surge um debate sobre até onde a privacidade é mantida quando os dados já não estão mais sobre gestão da T.I local. Shirey (2000) define da seguinte forma:

A privacidade pode ser definida como o direito de uma determinada entidade (normalmente um indivíduo), agindo em seu próprio nome, de determinar o grau de interação de suas informações com contexto onde se encontra inserida, incluindo o grau de comprometimento/disposição em divulgar essas informações para outras entidades.

A privacidade é composta por três elementos: o sigilo, o anonimato e o isolamento, de modo que o sigilo é um problema ligado à confiabilidade, o anonimato à identidade do sujeito e o isolamento é o direito de manter-se invisível perante outros indivíduos [Fischer-Hübner, 2001].

O direito à privacidade é consolidado em áreas como a médica, jurídico e a fiscal. No caso da área médica, a privacidade é a limitação de acesso às informações de um indivíduo, em outras palavras, é o direito que cada indivíduo tem de não ter sua vida ou seus dados observados sem autorização. Em outros contextos/cenários, privacidade é o direito de ficar sozinho, ou e realizar ações sem que algum sistema o identifique [JÚNIOR et al, 2010].

O primeiro trabalho com a finalidade de discutir questões relacionadas à privacidade em um ambiente computacional foi [Ware, 1973], onde foram definidos vários princípios

relacionados à privacidade de dados pessoais alocados em um sistema. A partir deste trabalho e de trabalhos recentes como [Fischer-Hübner, 2001; Hinde, 2003; Rezgui et al, 2003; Yee e Korba 2009], pode-se fundamentar alguns princípios relacionados à privacidade em um contexto mais amplo, não somente computacional explanados por JÚNIOR et al. (2010) como:

1. Responsabilidade: Uma organização é responsável pelas informações pessoais sob o seu controle e deve designar indivíduos responsáveis pela organização e conformidade dessas com a legislação e políticas internas. Devem existir termos sobre responsabilidade de uso, com sanções levais em caso de mau uso.
2. Identificação de objetivos: Os objetivos para os quais as informações pessoais são coletadas devem ser identificados pela organização previamente ou enquanto a informação é coletada.
3. Consentimento: O consentimento dos indivíduos é necessário para a coleta, uso e/ou divulgação de informações pessoais. As informações não podem ser transferidas para outras entidades, salvo se autorizado e com um nível de proteção adequado.
4. Limite de coleta: A coleta de informação pessoal será limitada ao necessário para os fins identificados pela organização. As informações devem ser coletadas por meios conhecidos e com amparo legal.
5. Limite de uso, divulgação e retenção: As informações pessoais não devem ser utilizadas ou divulgadas para outros fins que não aqueles para os quais foram coletadas, exceto com o consentimento do indivíduo ou se exigido por lei. Além disso, as informações pessoais serão mantidas apenas o tempo necessário para o cumprimento desses propósitos.
6. Precisão: As informações pessoais deverão ser tão precisas, completas e atualizadas quanto for necessário para os propósitos definidos.
7. Salvaguardas: Medidas de segurança adequadas à sensibilidade das informações devem ser usadas para protegê-las, tanto em relação à sua confidencialidade quanto à sua integridade.
8. Transparência: A organização deve tornar disponíveis aos indivíduos informações específicas sobre suas políticas e práticas relativas à gestão das informações privadas.

9. Acesso individual: A pedido, todo indivíduo deve ser informado da existência, uso e divulgação de suas informações pessoais e deve ser permitido o acesso a essa informação.
10. Crítica à conformidade: Um indivíduo deve ser capaz de criticar a precisão e integridade de suas informações e modifica-los se necessário. Também deve ser facultado ao indivíduo questionar os princípios anteriores ou sujeitos responsáveis a respeito da política de privacidade adotada.

A privacidade é atributo imprescindível para a Computação em Nuvem, pois influi diretamente tanto na confiança do cliente quanto nas conformidades legais. Na medida em que a Computação em Nuvem evolui, novos questionamentos sobre a privacidade são feitos [Mather et al, 2009]:

1. Acesso: os donos dos dados têm o direito de saber quais informações são mantidas e, em alguns casos, de solicitar a remoção dessas informações. Se um usuário exerce o seu direito de solicitar ao provedor a eliminação dos seus dados, será possível garantir que todas as suas informações foram eliminadas da nuvem?
2. Aderência: quais são os requisitos de conformidade à privacidade neste ambiente? Quais leis, regulamentos, normas ou compromissos contratuais regem o uso das informações, e quem são os respectivos responsáveis? Um ambiente de nuvem pode atravessar várias jurisdições, no caso de dados armazenados em vários países. Qual o foro competente para regulamentar esse ambiente ou as informações armazenadas nele?
3. Armazenamento: em qual parte da nuvem as informações estão armazenadas? Estão em um centro de dados de outro país? A legislação sobre a privacidade pode limitar a capacidade dos provedores em transferir alguns tipos de informações para outros países. Todavia, quando os dados são armazenados na nuvem, essa transferência pode ocorrer sem o conhecimento dos mesmos.
4. Retenção: por quanto tempo a informação pode ser retida na nuvem? Quais as políticas de retenção e descarte dessas informações? Quem rege essas políticas, o

consumidor que armazenou as informações ou o provedor de computação em nuvem? Quais as exceções à política?

5. Destruição: como deve ocorrer a destruição das informações que identificam o consumidor? Como garantir que não foram conservadas cópias destas? A disponibilidade através de replicação pode ser um problema no momento da destruição de informações. Será que as réplicas da informação foram destruídas ou apenas tornaram-se inacessíveis?
6. Auditoria: como as organizações consumidoras podem monitorar e verificar se seus fornecedores estão cumprindo os requisitos de privacidade?
7. Violação da privacidade: em casos confirmados de violação, quais os responsáveis pela notificação, processos (e custos associados)? Como determinar a culpa de cada entidade envolvida?

Cada cenário possui sua característica específica, logo, ameaças quanto à privacidade diferem a cada ambiente onde a Computação em Nuvem está presente. Segundo [Pearson et al, 2009], vários cenários de riscos podem afetar a privacidade: os usuários de serviços poderiam ser forçados a aceitar o monitoramento de suas informações, ou induzidos a isto, ou até mesmo fornecer dados pessoais contra a própria vontade; os provedores de serviços poderiam utilizar as informações para outros fins dos quais o cliente não estaria ciente antes da contratação. Por outro lado, as prestadoras de serviços também poderiam sofrer consequências legais pela exposição de informações sigilosas, em caso de problemas, perdendo credibilidade e prejuízos financeiros.

Algumas normas foram criadas com o intuito de facilitar a gestão da T.I como um todo, além de agregar informações relevantes para a implementação de segurança da informação. Conjunto de normas como a ITIL (*IT Infrastructure Library*) [ITIL, 2010] ou até mesmo o padrão ISSO 27001:2005 [ISSO, 2005], não abordam a virtualização ou o armazenamento virtual de dados e serviços na nuvem, não foram criadas com este

propósito. Referenciando este contexto, [Doelitzscher et al, 2010] caracteriza alguns problemas relacionados à Computação em Nuvem:

- Cumprimento das leis e políticas: os consumidores são responsáveis pela segurança e integridade de seus dados, inclusive nos casos em que estes estão sob a custódia de terceiros.
- Privilégios de controle de acesso: o armazenamento ou tratamento de informações sensíveis na nuvem gera um risco adicional: os serviços de computação em nuvem são controlados por terceiros. É necessário um controle de acesso restrito sobre os administradores para evitar acesso indevido às informações dos consumidores.
- Segmentação das Informações: as informações de vários consumidores da nuvem podem ser armazenadas no mesmo disco rígido físico, separadas pelo uso da virtualização. Para a proteção das informações contra acesso não autorizado, os provedores de serviços na nuvem precisam prover criptografia dos dados. Além disso, é necessário manter registros de auditoria, pois falhas durante o processamento das informações podem acarretar em perda destas.
- Incidentes de segurança: os provedores de serviços na nuvem oferecem poucos recursos no caso de incidentes de segurança. Em um ambiente onde máquinas virtuais são continuamente iniciadas e encerradas por vários consumidores, a atualização das informações dos usuários e dos registros de atividade é uma tarefa desafiadora. Normalmente, apenas o registro de atividades de um usuário na nuvem é oferecido em casos de investigação de incidentes.

Com todos os problemas e questionamentos quanto a segurança das informações contidas na Nuvem, esta tecnologia continua a crescer no cenário corporativo. Embora os riscos sejam uma realidade, as grandes empresas como Google, Microsoft, Amazon dentre várias outras, investem em alta disponibilidade, poder de processamento, redundância, tecnologia de ponta para garantir a integridade dos dados. Esse zelo é revertido em confiabilidade, os questionamentos continuam, fazem parte do pacote de

problemas que profissionais qualificados resolvem a cada dia, no entanto, um dado confirma que a Computação em Nuvem veio para revolucionar e manter-se no cenário de T.I moderna.

3.4 – MOTIVOS PARA UTILIZAR COMPUTAÇÃO EM NUVEM

Embora existam diversos fatores relacionados à Computação em Nuvem que de certa forma tornam esta tecnologia insegura para os profissionais mais conservadores, por outro lado a mesma está em constante desenvolvimento e aprimoramento. Novos problemas surgem a cada dia, assim como são solucionados pelos profissionais. Uma prova disso é a onda de contratações de profissionais da área de TI por parte não só da Google, mas também da HP e de outras empresas [ARRINGTON, 2011].

Segundo executivos da IBM, Fujitsu, Intel e Google, a Computação em Nuvem é uma tecnologia evolutiva, no entanto, ainda não está pronta para armazenar 100% dos dados “nas nuvens” [FRAGA, 2011].

Sébastien Marotte, vice-presidente de empreendimentos do Google, sugere que as empresas utilizem mais as vantagens que a Computação em Nuvem oferece, e complementou: [FRAGA, 2011]

Não podemos por tudo em nuvem agora, mas meu conselho é que coloquem tudo o que puderem na Nuvem, pois os benefícios são altos. São benefícios e melhorias em escala, manutenção, acessibilidade e evolução.

Segundo [FERRER, 2011], a Computação em Nuvem atingirá a marca de US\$ 1,45 bi em 2015 referentes a movimentações e transações financeiras. Empresas prestadoras de serviços de armazenamento de dados tiveram um aumento significativo de 56% em sua receita no ano de 2011.

VIRTUALIZAÇÃO DO ARMAZENAMENTO OU VIRTUALIZAÇÃO DE STORAGE

A gestão do armazenamento já foi simples, com poucas unidades de disco, quase nenhuma demanda do provisionamento, afinal, caso houvesse necessidade de maior espaço em disco, bastava trocar a unidade física (*Hard Disk Drive*). No entanto, esta demanda cresceu e várias unidades físicas foram constantemente sendo adicionadas, como consequência a gestão dos dados ficou mais complexa tornando a demanda de novas soluções uma realidade urgente [KAY, 2011].

Para contornar o problema em evidência, foi criada o que é chamado de R.A.I.D (*Redundant Array of Independent Disks*). Essa tecnologia permite agrupar várias unidades físicas de armazenamento (H.D.D), facilitando a gestão dos dados neles inseridos e até mesmo fornecendo maior performance e maior segurança. No entanto, com o avanço da demanda de armazenamento, o R.A.I.D também passou a não portar tamanha necessidade de gestão, ainda mais com a heterogeneidade de equipamentos (*Storages*) [KAY, 2011].

A mais nova resposta da tecnologia de armazenamento de dados à demanda de mercado é a chamada Virtualização de Armazenamento, ou Virtualização de *Storage*. Esta tecnologia implementa uma camada de software ou hardware entre o sistema de armazenamento e os servidores, fazendo com que os aplicativos não tenham conhecimento de quais sub-sistemas estão sendo utilizados, ou onde os dados estão de fato armazenados. Através dessa mudança de paradigma na maneira como os dados são tratados, os administradores têm em mão uma interface única de gerenciamento, consolidando todos os diferentes tipos de necessidades no manuseio dos dados. A flexibilidade na expansão do tamanho de espaço utilizável também é uma vantagem da Virtualização de *Storage*, uma vez que a alteração é feita em tempo real, diminuindo o tempo de inatividade dos serviços dependentes sem prejudicar desempenho [KAY, 2011].

Se tratando de uma camada intermediária entre os sistemas de armazenamento e os servidores, a Virtualização de *Storage* torna-se a interface principal de gerenciamento em que os servidores “enxergam” todos os dispositivos de armazenamento como apenas um sistema de armazenamento, e o sistema de armazenamento “enxerga” todos os servidores como apenas um servidor através dessa interface. Com isso, a proteção dos dados é melhorada, afinal os servidores não interagem diretamente com os dispositivos e sim com uma interface de gerenciamento, isso é gradativamente benéfico uma vez que diferentes dispositivos são tratados da mesma forma e com total compatibilidade, até mesmo em casos onde há a necessidade de restauração de backup, onde a interface de gerenciamento pode restaurar um backup de uma fita para um dispositivo mais lento, de menor performance, para não prejudicar o desempenho dos dispositivos de maior performance que são utilizados para tarefas que dependem de desempenho avantajado. Como resultado, temos uma total flexibilidade gerenciada por apenas uma interface, que permite mover recursos de armazenamento livremente entre os dispositivos com transparência para o usuário final, e com a devida segurança [KAY, 2011].

Segundo Kay (2011), existem três tipos de implementações de Virtualização de *Storage*:

- Baseado em *Host*: Neste tipo de implementação, todo o I/O de dados é gerenciado por um único driver genérico. Este driver recebe todas as requisições de I/O por parte dos dispositivos de armazenamento e redireciona para os Servidores [KAY, 2011].
- Baseado em Dispositivo de Armazenamento: Já neste caso, a virtualização é implementada dentro do *Storage*, e toda interoperabilidade com outros equipamentos que sejam acoplados ou adicionados ao grupo de *Storage* passa a fazer parte da gestão deste dispositivo primário. Isso é possível, inclusive, pela maneira como os softwares que implementam R.A.I.D funcionam, agrupando novos recursos físicos [KAY, 2011].
- Baseado em Rede: Neste cenário, o *Storage* Virtualizado é visto como um dispositivo de rede, que geralmente utiliza de interfaces SAN através de Fibre Channel [KAY, 2011].

4.1 – DEFINIÇÃO E CONCEITO

Anterior ao conceito de Virtualização de *Storage* se faz necessário o entendimento do equipamento denominado *Storage*. Como podemos observar na figura a seguir:



Figura 1.5 – Storage Dell EqualLogic OS6000

O *Storage* tradicional, como pode se observar na Figura 1.5, nada mais é do que uma CPU (*Central Processor Unit*) com os demais componentes, memória RAM, memória física para portar o Sistema Operacional e é claro, os discos de armazenamento H.D.D. Esses discos podem variar entre o padrão SATA ou padrão SCSI, possuem fonte de alimentação de energia redundante para garantir a integridade do equipamento e, dos dados contidos nele em caso de queima de um dos dispositivos. É claro que este

equipamento é customizável, existem diversos fabricantes de *Storages* como é o caso de IBM, EMC, HP, etc [BUNN, 2011].

Uma das possíveis definições para Virtualização de *Storage* pode ser dita como “abstração do detalhe”. Esta seria a essência da virtualização. A virtualização provê uma interface simples e flexível para sistemas e interfaces complexas. Uma breve comparação, por exemplo, enquanto dirigimos um carro, não sabemos necessariamente o que ocorre internamente no motor, não sabemos os processos de combustão que ocorrem, o pedal do acelerador “virtualiza” essa operação, ou seja, esconde, simplifica todo processo interno do motor em apenas um meio que de fato interessa ao usuário, que é o próprio pedal de aceleração [BUNN, 2011].

Outras duas definições foram criadas pela SNIA – *Storage Networking Industry Association*, que segue:

- Ação de abstrair, esconder ou isolar funções internas do *storage* ou serviço de aplicações, *hosts*, ou recursos de rede, com o propósito de permitir a independência no controle do armazenamento de dados, controle de aplicações e recursos de rede [BUNN, 2011].
- A virtualização de *Storage* ou aplicação ou ainda de dispositivos tem o propósito de agregar funcionalidades escondendo a complexidade real das operações ou reduzindo a necessidade dos recursos do *Storage* [BUNN, 2011].

4.2 – TIPOS DE VIRTUALIZAÇÃO DE STORAGE

Segundo a SNIA, (citado por BUNN, 2011) existem cinco diferentes tipos de Virtualização de *Storage*: bloco, disco, fita, sistema de arquivo e virtualização de arquivo.

- Virtualização de Disco: É uma das Virtualizações de *Storages* mais antigas, e é feita diretamente no *firmware* do disco.

Em nível de hardware, a localização magnética do disco é definida pelos cilindros, trilhas, setores, no entanto, cada disco possui uma quantidade de cilindros, trilhas e setores diferentes, é por isso que a capacidade muda de disco para disco, de fabricante a fabricante. Com tamanha heterogeneidade de discos é impossível que o sistema operacional “saiba” exatamente todos os espaços de endereçamento dos mesmos [BUNN, 2011].

Para contornar este problema, é utilizada a virtualização de disco, onde o firmware transforma os endereçamentos físicos em blocos lógicos enumerados ou ainda endereçamento lógico de blocos, para serem utilizados pelos sistemas operacionais e aplicações. Essa solução conhecida como LBA (*Logical Block Addressing*) revolucionou a maneira como os sistemas operacionais se comunicam e se “entendem” com os discos. A partir dessa solução, o tamanho do disco é definido pela quantidade de blocos lógicos [BUNN, 2011].

- Virtualização de *Storage* em Fita: Este tipo de virtualização de *storage* é feita por diversas *Libraries*, equipamentos estes que são utilizados para backup, em que várias fitas denominadas LTO (*Linear Open Tape*) portam o backup da empresa em questão. No entanto, existem duas subdivisões dentro desta virtualização, que é a virtualização de fita LTO, e virtualização do *drive* de fita ou cabeça de gravação de fita [BUNN, 2011].

- Virtualização de Fita LTO: Neste caso, a Virtualização de Fita utiliza discos como cache para emular a leitura e escrita antes de gravar diretamente na fita física. Fazendo uso desta tecnologia, é possível melhorar tanto o desempenho de backup quanto o tempo de vida útil da cabeça de gravação de fita. O desempenho é melhorado em razão do buffer, que é utilizado sobre os discos, reduzindo as variações de velocidade causadas pela rede ou por *hosts* ocupados. A maneira como os dados são gerenciados antes de serem devidamente gravados na fita física melhora não só o desempenho, mas também o nível de compressão dos dados, e como consequência, o “aumento” do espaço disponível. Isso é possível, pois a virtualização de fita emula um ambiente com um número alto de pequenas fitas, desta forma, os

dados são salvos fora da fita em alta velocidade tornando possível a alta compressão antes de salvar os dados na fita física [BUNN, 2011].

- Virtualização de Cabeça Leitura e Escrita de Fita: Com a tecnologia de SAN e Fibre Channel foi possível disponibilizar através da rede um único *drive* (cabeça de leitura e escrita de fita) para diversos servidores. Em consequência a esse compartilhamento, um problema se torna real, o corrompimento dos dados a partir do momento em que diferentes servidores ou aplicações fazem acesso ao *drive*. A virtualização de *drive* soluciona este problema de integridade de dados, promovendo uma estabilização das fitas garantindo a segurança das informações [BUNN, 2011].

Um único *drive* pode ser mostrado como vários *drives* virtuais para servidores ou aplicações independentes, logo, quando algum destes servidores ou aplicações faz a requisição de utilização do *drive* virtual, esta requisição é enviada ao *drive* físico que reserva uma fita e a disponibiliza ao servidor pelo *drive* virtual [BUNN, 2011].

Quando a operação é dada como completa, o controlador devolve o *drive* de fita física para ser disponibilizado novamente, dessa forma, podendo ser utilizado por outro *host*. A inteligência inclusa na virtualização de *drive* faz com que aplicativos e servidores não conflitem ao acessar uma determinada fita, seja para ler ou gravar [BUNN, 2011].

- Virtualização do Sistema de Arquivos: Este tipo de virtualização tem um conceito muito simples, que pode ser exemplificado através de um servidor NAS (*Network Attached Storage*), em que um servidor de arquivos é disponibilizado através da rede utilizando sistemas de arquivos NFS ou CIFS. Neste tipo de virtualização, um servidor dedicado gerencia os recursos de alocação de arquivos que são compartilhados na rede. Independente de qual sistema operacional os arquivos são alocados, ou o local onde são acessados, as aplicações e os usuários utilizam de forma transparente os dados neste servidor remoto. Esta abstração da localização da informação é um exemplo claro de uma das importantes vantagens da virtualização de *storage*: transparência na localização [BUNN, 2011].

Em um ambiente de banco de dados, a sobrecarga no sistema de arquivos degrada o desempenho. Por isto, alguns administradores utilizam partições RAW

para estes cenários em detrimento da facilidade de controle, já outros continuam a utilizar sistemas de arquivos para facilitar a visualização dos dados e com isso ter melhor controle sobre os mesmos. No entanto, a virtualização de *storage* consegue unir os benefícios de ambas as soluções, a facilidade de gerenciamento e controle do sistema de arquivo, com o desempenho das partições RAW. Esta junção fornecida pela virtualização permite que o sistema de arquivo seja visível ao administrador, permitindo uma experiência melhorada no gerenciamento das entidades de banco de dados. No entanto as entidades estão alocadas fisicamente em discos RAW, o sistema de arquivos permanece escondido e dessa maneira o *buffer* de I/O é ignorado resultando máximo desempenho [BUNN et al, 2003].

Virtualização de Arquivo: Neste caso, a virtualização de arquivos é utilizada geralmente para a migração de arquivos que são raramente acessados para *storages* de baixo custo, ou *storages* secundários como fitas, discos padrão *Serial ATA (SATA)*. Toda a transação de migração é transparente ao usuário e à aplicação em execução, sendo que os dados continuam a serem acessados no *storage* principal. Mais uma vez, a virtualização de *storage* baseia-se na transparência da localização dos dados [BUNN, 2011].

- Virtualização de Bloco: A virtualização de bloco é o “passo mais atual” da virtualização de *storage*. Anteriormente, este trabalho salientou sobre o conceito de LBA, onde a virtualização provê a criação de blocos lógicos, com o objetivo de gerenciar adequadamente os diferentes tipos de discos magnéticos e suas capacidades. Logo, a idéia de virtualização de bloco é simples, constituindo em agrupar vários blocos lógicos, o chamado *block aggregation*. No entanto, outros serviços ou aplicações podem ser inclusas na virtualização de bloco, e esta é de fato a “faceta” deste tipo de virtualização [BUNN, 2011].

- **Physical disks** (Discos Físicos)

- Fixed size (Tamanho Fixo)
- Bounded performance (Desempenho Limitado)
- Do break (occasionally) (Quebra, ocasionalmente)



- **Virtual disks** (Discos Virtuais)

- As big, small, or numerous as users need (Grande ou pequeno, como o usuário precisa)
- As fast as users need (Tão rápido quanto o usuário precisa)
- Can be “very reliable” or not (Pode ser muito confiável, ou não)
- Can grow(!), shrink, or morph (Pode ser incrementado, decrementado)

Figura 1.6 – Virtualização de Bloco

Os administradores de *Storages* querem utilizar uma ferramenta que possibilite controlar vários tipos de servidores e sistemas de armazenamentos (*storages*) diferentes, com o objetivo de simplificar o gerenciamento. A virtualização de *storage* é uma ferramenta avançada para os administradores de *storages*, que pode ajuda-los a encontrarem estes objetivos [BUNN, 2011].

O objetivo da virtualização de bloco é controlar o *storage* físico de modo que possibilite aos volumes lógicos oferecer capacidade, desempenho e confiança sem perder o nível de detalhes sobre o mesmo. Além disso, para quem está utilizando diretamente a virtualização de bloco, não percebe o envolvimento dos volumes lógicos com os volumes físicos, apenas percebe que a capacidade atende às demandas da aplicação. A camada de virtualização é responsável por mapear o *Input/Output* requisitado pelo volume lógico para o *storage* físico. Simplificando, a virtualização de bloco cria *storages* virtuais utilizando *storages* físicos, oferecendo capacidade, confiabilidade e desempenho que os administradores precisam. Citando alguns casos práticos, temos: [BUNN, 2011]

- Se for necessário adicionar um novo disco para aumentar a capacidade de armazenamento, um novo volume é criado ou o disco já em utilização é expandido. É claro que o reverso disto também é válido. Qualquer alteração dessa magnitude é transparente ao usuário [BUNN, 2011].
- Caso a necessidade seja aumento de desempenho, existe a possibilidade de adição de novo (os) discos para que o RAID 0 seja mais efetivo. Lembrando que não é apenas de *Stripe* que o desempenho pode ser melhorado, existem outras maneiras além desta [BUNN, 2011].
- Já se a necessidade for confiabilidade, também existem diversas possibilidades para garantir a integridade dos dados. Um exemplo é a clusterização, assim como a utilização de *mirroring* ou espelhamento [BUNN, 2011].

4.3 – A VIRTUALIZAÇÃO DE STORAGE

Com diversas interfaces de gerenciamento, *consoles* de manutenção, *softwares* de controle, servidores, sem contar sistemas operacionais e sistemas de armazenamento (*storage*), os ambientes de T.I são extremamente diversificados e heterogêneos. Para unificar e resolver o problema da dificuldade de gerenciamento, a Virtualização de *Storage* surge como solução. Segundo [BUNN, 2011] Áreas de problemas para soluções de Virtualização.

Ainda segundo o mesmo autor, algumas áreas que apresentam problemas em acesso à informação podem ser resolvidas pela virtualização:

- Falhas de comunicação entre o *Storage* e *switches* SAN são comuns em ambientes que não possuem solução de virtualização, podendo ainda ser

necessário alto investimento para restauração ou resolução do problema [BUNN, 2011].

- Definir QoS em *Storage* não é uma tarefa simples, no entanto, a virtualização pode não só prover QoS como possibilitar o ganho de performance e a sua gerencia em tempo real [BUNN, 2011].
- Dados alocados em tecnologias SAN são críticos na operação diária de uma empresa e a perda de acesso quanto a perda dos dados propriamente ditos, são seguidos de consequências incalculáveis. No entanto, a virtualização de *storage* possibilita maiores facilidades em casos de *disaster recover* ou *data archiving* para clientes que não tem acesso a tecnologias de ponta [BUNN, 2011].

Neste caso de avaliação da tecnologia, é necessário comparar as tecnologias virtualizadas e não virtualizadas.

Segundo [BUNN, 2011], arquiteturas de *storage* não virtualizado são controlados e gerenciados e desenvolvidos como discos físicos. Isto implica na necessidade de configurações de acesso, conexões, estudo de capacidade futura ou mais conhecido como *provisioning*, o que claramente expõe um modelo que não permite escalabilidade.

Comparativamente, a solução com virtualização oferece inúmeros benefícios e vantagens, como por exemplo, a redução no tempo de manutenção em razão da facilidade apresentada pela interface de gerenciamento provida pela virtualização. Basicamente, é muito mais fácil gerenciar recursos físicos do que recursos virtuais, além é claro de toca a capacidade de escalabilidade, flexibilidade que também são diferenciais fornecidos pela virtualização. Não obstante a isso, as empresas que compravam *storage* de um fornecedor específico e que posteriormente optou por mudar a marca do equipamento em uma nova compra, pode fazer isto tranquilamente, pois a camada de virtualização sem encarrega de abstrair o recurso físico, não havendo problemas de compatibilidade ou escalabilidade mesmo entre *hardwares* diferentes. O autor [BUNN, 2011] chega a dizer que a virtualização é a chave para o controle do fornecimento previsível de QoS (*Quality of Service*) de serviço na SAN.

Abaixo estão três motivos pelo qual uma empresa deve optar por virtualização de *storage*:

- Melhora no controle de sistemas de *storage* heterogêneos.

- Melhora na confiabilidade e tempo de resposta com controle automático.
- Melhor utilização do *storage*.

Um exemplo prático e amplo da utilização de virtualização de *storage* como solução, pode ser analisado com o caso da empresa *Champion Technologies*. Roland Etcheverry ao entrar para esta empresa percebeu a necessidade de mudanças no gerenciamento de dados, e ao mesmo tempo, já sabia que a virtualização de *storage* era necessária. Roland já havia implementado estas soluções, da mesma maneira que já havia trabalhado com o conceito de SAN (*Storage Area Network*), e sabia o quanto a utilização destes recursos poderia ajuda-lo com a reformulação do ambiente de *storage* [GRUMAN, 2011].

Para Etcheverry, a virtualização de *storage* permite que os funcionários gastem menos tempo gerenciando um ambiente de *storage* quando o mesmo apresenta apenas um único espaço, embora demasiadamente grande, possibilitando a centralização do controle. Como estudo de caso, Roland ainda cita algumas vantagens, com o uso eficaz do *storage*, pois os dados são alocados de forma que todos os discos “participem”, ao contrário de utilizar apenas alguns discos enquanto outros ficam subutilizados. O conceito de SAN agregado com a virtualização de *storage* permite ainda a adição ou a remoção de discos sem impactos diretos aos usuários, pois o *software* de virtualização se encarrega disso. Além disso, restaurar um *backup* também é mais performático, pois os dados não precisam ser copiados em sua totalidade, e sim, apenas o que foi alterado é restaurado [GRUMAN, 2011].

Em estudos para futuras ampliações do ambiente, a virtualização de *storage* facilita o planejamento refinado ou *thin provisioning*, pois torna possível o gerenciamento centralizado de diversos fornecedores, permitindo sempre optar pelo melhor equipamento independente de qualquer franquia já em produção [GRUMAN, 2011].

Segundo Etcheverry, “Muitas pessoas não pensam adequadamente sobre *storage*, não fazem um plano para reter custos”. Nik Simpson, um analista de *storage* da empresa *Burton Grup*, afirma que a virtualização de *storage* é uma forma muito diferente de gerenciar a informação, e com isto, as pessoas perdem muitas das vantagens que esta solução possibilita, como por exemplo, a redução de custos e as melhorias na produtividade da empresa onde será implantada [GRUMAN, 2011].

Embora a virtualização de *storage* possua diversas vantagens, uma das principais esta relacionada aos *backups*. Como já citado acima, a virtualização possibilita uma rápida restauração de *backup*, no entanto, mais do que isso, é possível mover os dados entre diversos equipamentos diferentes e com isso, em casos de *disaster recovery* ou restauração de desastre, ter maior confiança na integridade dos dados uma vez que a virtualização provê replicação em ambientes em que existe um *cluster* [GRUMAN, 2011].

Segundo a *Network1* (2011) algumas outras vantagens são:

- Redução de tempo de *backup*.
- Tempo de parada de ambiente reduzido.
- Tempo de recuperação de arquivos mais rápidos via VLS (*Virtual Library System*).
- Melhoria na performance de recuperação.
- Redução nos custos de mídia.
- Eliminação da necessidade de gerenciamento de *Disk Array*.
- Diminuição das janelas de *Backup*.
- Segurança e redundância das informações.

4.4 – FERRAMENTAS DE VIRTUALIZAÇÃO DE STORAGE

Neste tópico, este trabalho vai citar algumas ferramentas de virtualização de *storage* que estão em comercialização. A primeira delas é a ferramenta da IBM (*International Business Machines*) denominada IBM *TotalStorage®*:

- IBM *TotalStorage®*: A IBM disponibilizou ferramentas para a virtualização de *storage* subdivida em três tipos: IBM *TotalStorage® SAN Volume Controller*, IBM

TotalStorage® SAN Integration Server e *IBM TotalStorage® SAN File System*. No entanto, este trabalho focará apenas duas soluções:

- *IBM TotalStorage® SAN Volume Controller*: Esta ferramenta foi projetada para diminuir a complexidade dos ambientes de *storage* e reduzir os custos de gerenciamento de redes neste cenário. Para alcançar este objetivo, esta ferramenta possibilita que os administradores façam alterações como incremento de discos físicos e lógicos, sem que os aplicativos sejam interrompidos. Segundo a IBM, esta ferramenta foi projetada para:
 - Oferecer um ponto centralizado de gerenciamento de uma SAN heterogênea a partir de um único console.
 - Ajudar a tomada de decisão utilizando a virtualização existente, e centralizando o gerenciamento
 - Redução do tempo de inatividade durante pausas planejada e não planejada, manutenção e *backups*.
 - Aumentar a utilização da capacidade de armazenamento, o tempo de execução, a produtividade e eficiência do administrador [IBM, 2011].
 - Possibilitar serviços avançados de cópia e *backup* para vários dispositivos de armazenamento [IBM, 2011].

- *IBM TotalStorage® SAN File System*: Esta ferramenta foi projetada para permitir o compartilhamento de informações em ambientes heterogêneos, e com isso, possibilitar o melhor e mais adequado controle do ambiente. Como objetivo maior, esta solução pretende reduzir custos de gerenciamento e aprimorar a produtividade, oferecendo gerenciamento centralizado, maior utilização do armazenamento e disponibilidade de aplicativo. Esta ferramenta foi projetada para:
 - Melhorar a produtividade e reduzir a complexidade do armazenamento, utilizando de uma central de gerência [IBM, 2011].
 - Melhorar a utilização do *storage* fazendo melhor proveito de sua real capacidade.
 - Reduzir o tempo de *downtime* de aplicativos.

- Prover o acesso de dados compartilhados mais facilmente [IBM, 2011].

A segunda ferramenta de virtualização de *storage* a ser brevemente apresentada, é a solução da *Hitachi Data Systems*, denominada Hitachi Universal Storage Platform®:

- Hitachi Universal Storage Platform®: Esta ferramenta foi desenvolvida para atender não somente a grandes empresas, mas também as de pequeno e médio porte. No entanto, a solução é robusta o suficiente para suportar aplicativos de grandes empresas. Segundo [HITACHI, 2011], pela primeira vez as pequenas empresas poderão desfrutar de replicação de dados, dentre outros benefícios em razão da virtualização.

Ainda segundo o mesmo autor, a ferramenta USP possui a mais alta confiabilidade se tratando de solução de armazenamento. Juntamente com o armazenamento de dados e serviços que incluem tanto o *software Hitachi Dynamic Provisioning*, aplicativo este que funciona como uma central de gerenciamento entre sistemas de *storage* heterogêneos [HITACHI, 2011].

Algumas funcionalidades desta ferramenta podem ser analisadas abaixo: [HITACHI, 2011]

- Alto desempenho com 1.2 milhões por segundo.
- Escalabilidade suportando 96 pentabytes de total de *storage*.
- *Flashdrives* para desempenho excepcional
- *Storage* baseado em *failovercluster*
- Consolidação de vários protocolos (Fibre Channel, IBM FICON, SCON e NAS)

Também são relacionados alguns benefícios: [HITACHI, 2011]

- Redução dos riscos para a empresa
- Incremento na agilidade de transações empresariais
- Diminuição do custo total de aquisição

A terceira ferramenta a ser apresentada é a da produtora de *Storage* EMC. Esta empresa tem como parceira a VMware, e oferece alguns benefícios quando é utilizada em conjunto com soluções da pioneira da virtualização de plataforma x86, que segue:

- Reduz custos em razão da eficiência e automação.
- A EMC garante que o armazenamento unificado que oferece, é 20% mais eficiente do que outro *array* unificado.
- Melhor tempo de resposta do VMware View em mais de 80%
- Redução no consumo do armazenamento utilizando até 50% de compactação avançada.

Além desses benefícios, a EMC ainda oferece compatibilidade aprimorada com VMware:

- Provisionamento do VMware vCenter Server em alguns cliques.
- Aumente o desempenho do servidor VMware em até dez vezes.
- Obtenha implementações, clones, snapshots até dez vezes mais rápidos.

Embora todas as tecnologias e ferramentas de virtualização de *storage* aparentam “rodear” os mesmos benefícios, e prover as mesmas funcionalidades, todas também possuem pontos de enfoque, como por exemplo, a EMC que foca totalmente a solução em compatibilidade com a VMware.

5. CONSIDERAÇÕES FINAIS

O mercado de tecnologia segue em ritmo acelerado, novas ideias se transformam em tecnologias que mudam a vida das pessoas. A cada ano, soluções para os problemas do momento são descobertas, novos problemas surgem, outras soluções são providenciadas, e assim a evolução da tecnologia acontece desenfreadamente.

De maneira mais específica, a TI evolui da mesma maneira, no entanto, um objetivo fica claro neste contexto definido, a busca pela melhor solução. Este objetivo impulsiona as pesquisas tecnológicas. A cada ano milhares de profissionais de engenharia de computação, cientistas da computação, entram para o mercado de trabalho e contribuem para a evolução da TI com suas pesquisas.

Seguindo este raciocínio, a virtualização surge como uma descoberta da década de sessenta, e que em constante desenvolvimento e aprimoramento, chegou ao patamar de tecnologia transformadora, não somente da parte física e lógica de um *data center*, mas também dos profissionais de TI.

Em paralelo à Virtualização em ascensão, uma maneira inovadora de acesso à informação e disponibilização de serviços de todos os tipos, embalada pela facilidade de acesso à banda larga, surge para fazer uso da Virtualização como forma primordial de evolução, a chamada “Computação em Nuvem”. Este novo conceito de TI quebra paradigmas e altera radicalmente o rumo com que empresas prestadoras de serviços trabalham.

Este trabalho permite concluir que são infinitas as possibilidades de serviços de TI que podem ser disponibilizados somando a virtualização à computação na nuvem. Desta forma, alguns serviços como a Virtualização de *Storage*, IaaS, PaaS, se destacam pelas vantagens quando em comparação com soluções fora do contexto da Nuvem.

Com os estudos realizados, é possível afirmar que a Virtualização de *Storage* não é uma tendência, e sim uma realidade, empresas como IBM, HP, Hitachi, Google, utilizam desta

tecnologia como carro chefe, afinal, sem a Virtualização de *Storage* não seria possível oferecer diversos outros serviços, como por exemplo, o *SkyDrive* da Microsoft, o *GoogleDocs* da Google. Não somente isto, empresas de pequeno, médio e grande porte podem e devem fazer uso dessa tecnologia, pois a mesma impacta fortemente na redução de custos, e até mesmo na maior lucratividade da empresa, uma vez que o tempo de *downtime* é reduzido drasticamente.

Do ponto de vista do profissional que vai trabalhar com um cenário de Virtualização de *Storage*, é fenomenal, pois já não é necessária a preocupação com equipamentos de diferentes fabricantes, desde que possua suporte ao mesmo protocolo, todos interagem com uma única interface de gerenciamento, a flexibilidade e a rapidez na execução de reparos, melhoras, ou manutenções básicas, é muito maior.

É claro que toda tecnologia possui um ponto fraco, e na Virtualização de *Storage* aplica-se da mesma maneira. Pode-se afirmar, em razão da pesquisa realizada neste trabalho que a Computação em Nuvem está resolvendo seus problemas com o tempo, é o caso do grande entrave, a segurança. Os empresários ainda estão duvidosos quando se trata de deixar informações “em mãos” que não conhece, o conceito de Nuvem é abstrato demais e a sensações de fragilidade e insegurança tomam conta durante o processo de decisão. Com toda certeza, existe um trabalho muito árduo a ser executado até que soluções bem definidas, claras e simples quanto à segurança da informação, sejam providenciadas.

A conclusão que se pode tirar de todo escopo abordado neste trabalho, é que de fato a Virtualização de *Storage* está cada vez mais perto de pequenas, médias e grandes empresas, pois é visível que as vantagens que esta tecnologia oferece são imensamente maiores do que as prováveis desvantagens, o que no fim das contas, se torna atrativo não somente para o empresário que busca o melhor custo x benefício, mas também para o profissional que gerencia o ambiente. Embora algumas questões sobre a privacidade das informações sejam de fato fundamentadas, este trabalho permite concluir que a boa gerência do ambiente garante a segurança da informação, tornando uma solução que utiliza a Virtualização de *Storage* viável, atrativa, segura, e muito confiável.

REFERÊNCIAS

ARRINGTON, Michael. **Google, Facebook Battle for Computer Science Grads. Salaries Soar.** Local disponível em: <http://techcrunch.com/2008/01/30/stanford-computer-science-grads-getting-95k-offers-from-google/>. Acessado em: 11/10/2011

BBC. **HP, Intel, Yahoo in cloud tie-up.** Disponível em: <http://news.bbc.co.uk/2/hi/technology/7531352.stm>. Acesso em: 20/09/2011.

BUNN, Frank. **Storage Virtualization.** Disponível em: www.snia.org/sites/default/files/snivirt.pdf. Acesso em: 26/09/2011.

CALLEGARI, Lucas. **Empresas ainda estão estudando a nuvem.** Local disponível em: <http://computerworld.uol.com.br/tecnologia/2011/06/21/empresas-ainda-estao-estudando-a-nuvem-diz-pesquisa/>. Acessado em: 21/09/2011.

CARISSIMI, Alexandre. **Virtualização: da teoria a soluções.** Universidade Federal do Rio Grande do Sul. Disponível em: www.gta.ufrj.br/ensino/CPE758/artigos-basicos/cap4-v2.pdf. Acesso em: 10/09/2011.

COELHO, Fabio de Azeredo, CALZAVARA, Gustavo Santos, DI LUCIA, Ricardo. **Virtualização – VMWare e Xen.** Local disponível em: http://www.gta.ufrj.br/grad/09_1/versao-final/virtualizacao/index.html. Acesso em: 19/06/2011.

DIKAIKOS, M. D.; PALLIS, G.; KATSAROS, D.; MEHRA, P.; VAKALI, A. **Cloud Computing – Distributed Internet Computing for IT and Scientific Research.** IEEE Internet Computing, 13(5): 10-13, setembro/outubro 2009.

Doelitzscher, F., Reich, C., e Sulistio, A. (2010). Designing cloud services adhering to government privacy laws. Em International Symposium on Trust, Security and Privacy for Emerging Applications.

EMC². **A melhor opção de armazenamento voltado à virtualização.** Local disponível em: <http://brazil.emc.com/campaign/global/virtualization/index.htm>. Acesso em: 20/10/2011.

FERRER, Rafael. **Computação em nuvem poderá movimentar US\$ 1,45 bi em 2015.** Disponível em: <http://exame.abril.com.br/tecnologia/noticias/computacao-em-nuvem-podera-movimentar-us-1-45-bi-em-2015>. Acesso em: 05/10/2011.

Fischer-Hübner, S. (2001). IT-Security and Privacy: Design and use of privacy-enhancing security mechanisms. Em Goos, G., Hartmanis, J., e van Leeuwen, J., editores, Lecture Notes in Computer Science, volume 1958. Springer-Verlang.

FOSTER, I. What is the Grid? A Three Point Checklist. Argonne National Laboratory & University of Chicago, julho 2002.

FRAGA, Renê. **Google: computação em nuvem é uma tecnologia evolutiva.** Disponível em: <http://googlediscovery.com/2011/03/09/google-computacao-em-nuvem-e-uma-tecnologia-evolutiva/>. Acesso em: 21/09/2011.

Goldberg, R.P. .Survey of virtual machine research.. IEEE Computer, pp. 34-35, june 1974. Popek, G.; Goldberg, R. .Formal requirements for virtualizable 3rd generation architectures.. Communications of the ACM, v.17, n.7, pp. 412-421, 1974.

GRUMAN, Galen. **How to do Storage Virtualization Right.** Local disponível em: http://www.cio.com/article/136600/How_To_Do_Storage_Virtualization_Right. Acessado em: 20/10/2011.

Hinde, S. (2003). Privacy legislation: a comparison of the US and european approaches.

Computers & Security, 22(5):378–387.

HITACHI. **Hitachi Universal Storage Platform VM**. Local disponível em: <http://www.hds.com/products/storage-systems/universal-storage-platform-vm.html>. Acesso em: 20/10/2011

HP, **O que é Virtualização e o que ela pode fazer pela minha empresa? Local disponível em:** http://www.hp.com/latam/br/pyme/solucoes/apr_solucoes_01.html. Acesso em: 15/06/2011.

IBM, **Realidades da Computação em Nuvem de Software Livre, Parte 1: Nem Todas as Nuvens São Iguais**. Local disponível em: <http://www.ibm.com/developerworks/br/library/os-cloud-realities1/>. Acessado em: 19/09/2011

IBM. **SAN Volume Controller**. Local disponível em: http://www.ibm.com/br/systems/storage/san_volume_ct.phtml. Acesso em: 20/10/2011.

ISO (2005). ISO/IEC 27001 - Information technology - Security techniques - Information security management systems - Requirements. International Organization for Standardization.

ITIL (2010). IT Infrastructure Library. Office of Governance Commerce, UK.

JEDRAS, Jeff. **Segurança em Cloud começa no chão**. Local disponível em: <http://computerworld.uol.com.br/tecnologia/2011/09/13/seguranca-em-cloud-comeca-no-chao-nao-na-nuvem/>. Acessado em: 21/09/2011.

JÚNIOR, Arlindo Marcon da. Aspectos de Segurança e Privacidade em Ambientes de Computação em Nuvem. Programa de Pós-Graduação em Informática - PPGIa, Dezembro de 2007.

KACE SYSTEMS MANAGEMENT, INC. **Bests Practices in Virtualization: Integrated Management**. Local Disponível em: <https://www.kace.com/resources/Best-Practices-in-Virtualization-Integrated-Management>. Acesso em: 20/03/2011.

KAY, Russell. **QuickStudy: Storage virtualization**. Local disponível em: http://www.computerworld.com/s/article/325633/Storage_Virtualization. Acessado em: 13/10/2011.

MARTINS, Andre. **Fundamentos de Computação Nuvem para Governos**. Disponível em: <http://www.serpro.gov.br/wcge/artigos/ArtigoFundamentos%20de%20Computacao%20NUvem%20para%20Governos.pdf>. Acessado em: 20/09/2011.

MASINA, Daniel. **PERCEPÇÃO EM RELAÇÃO AO PROCESSO DE VIRTUALIZAÇÃO NO DATA CENTER DA PROCERGS: UM ESTUDO DE CAS**. 2008. 79p. Trabalho de Conclusão de Curso – Instituto de Ciências Exatas e Tecnológicas - Centro Universitário Feevale, Rio Grande do Sul, Novo Hamburgo, 2008.

Mather, T., Kumaraswamy, S., e Latif, S. (2009). *Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance*. O'Reilly Media.

MOHAMED, A. A history of Cloud Computing. ComputerWeekly.com, março 2009.

NETWORK1. **Benefícios da Virtualização de Storages**. Local disponível em <http://www.network1.com.br/mkt/virtualizacao/index.asp>. Acesso em: 20/10/2011.

NEXT GENERATION CENTER INTEL. **Virtualização**. Local disponível em <http://www.nextgenerationcenter.com/detallecurso/Virtualiza%C3%A7%C3%A3o.aspx?PagelD=1>, ao clicar no link “Baixar todo conteúdo do curso” será iniciado o download da apostila em que se encontra o material referenciado. Acesso em: 20/06/2011.

Pearson, S., Shen, Y., e Mowbray, M. (2009). A privacy manager for cloud computing. Em Jaatun, M., Zhao, G., e Rong, C., editores, Cloud Computing, volume 5931 of LNCS, páginas 90–106. Springer. 10.1007/978-3-642-10665-1-9.

Rezgui, A., Bouguettaya, A., e Eltoweissy, M. Y. (2003). Privacy on the web: Facts, challenges, and solutions. IEEE Security and Privacy, 1(6):40–49.

RYDLEWSKI, C. (2009) Computação Sem Fronteiras. Revista Veja, ed. 2125, São Paulo, ago 2009.

Shirey, R. (2000). RFC 2828 - Internet Security Glossary. The Internet Society.

SILVA, Rodrigo Ferreira. Virtualização de Sistemas Operacionais. Laboratório Nacional de Computação Científica - LTCC, 2010

TAURION, C. (2009) Cloud computing: Computação em Nuvem: transformando o mundo da tecnologia da informação. Rio de Janeiro: Brasport.

TRAPANI, Gina. **Os riscos ocultos da Computação em Nuvem**. Local disponível em: <http://www.gizmodo.com.br/conteudo/lifehacker-os-riscos-ocultos-da-computacao-em-nuvem>. Acessado em: 19/09/2011

Turn, R. e Ware, W. H. (1975). Privacy and security in computer systems. Technical Report P5361, Rand Corporation.

VAQUERO, L. M.; MERINO-RODERO, L.; CACERES, J.; LINDNER, M. A Break in the Clouds: Towards a Cloud Definition. ACM SIGCOMM Computer Communication Review, 39(1): 50-55, janeiro 2009.

VMWARE, **Conceitos básicos da virtualização**. Local disponível em <http://www.vmware.com/br/virtualization/virtualization-basics/history.html>. Acesso em: 10/06/2011.

VOAS, J.; ZHANG, J. Cloud Computing: New Wine or Just a New Bottle? *IT Professional*, 11(2): 15-17, março/abril 2009.

Yee, G. e Korba, L. (2009). Personal privacy policies. Em Vacca, J., editor, *Computer and Information Security Handbook*, páginas 487–505. Morgan Kaufmann.