

Virtualização com Hyper-V versus VMWARE ESXi, Virtual Server comparando desempenho sobre balanceamento de carga e tolerância à falhas.

Cristiano dos Santos Diedrich, Marcelo Henrique Euzebio Batista

Sistemas de informação – Faculdade Cenecista Nossa Senhora dos Anjos (Facensa)
Av. Dr. José Loureiro da Silva, 1991 - 94.010001 – Gravataí – RS – Brazil

csdiedrich@yahoo.com.br, marcelobatista@sinprors.org.br

Abstract. *This paper search submit a research on the differences between the technologies of virtualization and the use in cluster. Tests are presented in the structures that define what technologies would be better to an environment where you need high availability of virtual machines, for fault tolerance and performance of the same to the best use of hardware and a decrease in the allocation of these spaces with the decrease in spending on energy.*

Resumo. *Este artigo procura apresentar uma pesquisa sobre as diferenças entre as tecnologias de virtualização, e o uso delas em cluster. Serão apresentados testes realizados nas estruturas que definirão qual das tecnologias seria melhor para um ambiente onde se necessita de alta disponibilidade de máquinas virtuais, tolerância a falhas e desempenho das mesmas visando o melhor aproveitamento dos hardwares e diminuição nos espaços de alocação dos mesmos juntamente com a diminuição no gasto em energia.*

1. Introdução

Um dos maiores paradigmas no mundo de TI hoje é: Como garantir alta disponibilidade, escalabilidade, segurança e desempenho para a solução ou serviço prestado?

Para essa dúvida surgiram muitas respostas, dentre delas a clusterização, que nada mais é do que a divisão do processamento de um determinado serviço ou aplicação e virtualização de máquinas.

Cada uma dessas respostas soluciona apenas uma das dúvidas, por exemplo, a clusterização soluciona o problema de alta disponibilidade e desempenho, e a virtualização a de segurança e escalabilidade. O que se vem pesquisando hoje é a junção dessas duas soluções, para isso empresas como Microsoft, Sun, e IBM vem investindo cada ano mais recursos para que se encontre o melhor uso dessas duas tecnologias em um único ambiente.

Há grandes exemplos dessas estruturas, como por exemplo, o site de pesquisas Google, que utiliza, para aumentar a velocidade das pesquisas, cluster de servidores, isso faz com que todos os seus servidores tenham acesso às mesmas informações, fazendo com que não seja necessário o acesso a um único servidor para trazer as informações necessárias para uma pesquisa, ou seja, a pesquisa, quando realizada, é feita nos servidores mais próximos de onde a pessoa que efetuar a pesquisa está. Com este procedimento a pesquisa torna-se muito mais rápida.

No âmbito da virtualização o exemplo mais conhecido é da empresa Microsoft que têm 70% de seus servidores em uma estrutura de virtualização, essa estrutura traz economia e ganho de gerenciamento, pois, há mais máquinas virtuais em um ambiente de poucos servidores físicos, fazendo com que a economia em energia, refrigeração do ambiente e espaço sejam reduzidos e o gerenciamento desses servidores fique mais centralizado. (MICROSOFT, 2009).

2. Virtualização

Virtualização é um recurso que permite o uso do hardware físico para executar vários sistemas operacionais.

Esse recurso permite o uso de múltiplas máquinas virtuais, com sistemas operacionais heterogêneos executando isoladamente, lado a lado na mesma máquina física. Cada máquina virtual tem seu próprio hardware virtual em cima do qual um sistema operacional e aplicações são carregados. O sistema operacional enxerga o hardware como consistente, ou seja, um conjunto de hardware em relação aos componentes físicos de hardware atual. (CARMONA, 2008).

Máquinas virtuais são encapsuladas em arquivos, as tornando fácil de serem salvas, copiadas e de se prover uma máquina virtual.

Sistemas completos (com aplicações totalmente configuradas, sistema operacional, BIOS e hardware virtual) podem ser movidos, em questão de segundos, de um servidor físico a outro para manutenção em tempo mínimo e continuamente consolida a carga de trabalho. (CARMONA, 2008).

Nos anos 60 foi introduzido a virtualização, permitindo o particionamento de

grandes mainframes, um raro e caro recurso. Após um tempo, microcomputadores forneceram mais eficiência e uma melhor forma de distribuir o poder de processamento das máquinas, então, perto dos anos 80, virtualização não era mais empregada extensamente. (BOSILCA; FEDAK, 2001).

Nos anos 90, pesquisadores começaram a enxergar que a virtualização poderia resolver alguns dos problemas associados à proliferação de hardwares mais baratos, incluindo subutilização, custos de gerenciamento e vulnerabilidade.

Recentemente as maiores fabricantes de processadores do mercado, Intel e AMD, desenvolveram otimizações no hardware, com novas instruções de CPU, para uso específico de virtualização de sistemas operacionais. A tecnologia da Intel se chama Vanderpool e a da AMD de Pacifica, e ambas permitem que um processador funcione como se fossem vários processadores trabalhando em paralelo de modo a permitir que vários sistemas operacionais sejam executados ao mesmo tempo em uma mesma máquina. (CROSBY, 2009; TORRES, 2008; UHLIG, 2009).

Existe uma confusão relacionada à virtualização e sistemas multitarefa. Nos sistemas operacionais multitarefa existe um único sistema operacional e vários programas trabalhando em paralelo. Na virtualização você pode ter vários sistemas operacionais trabalhando em paralelo, cada um com vários programas distintos em execução. Cada sistema operacional roda em um processador virtual, ou seja, em uma máquina virtual. (CROSBY; BROWN, 2009).

3. Clusterização

A ideia inicial que conduz ao cluster foi desenvolvida na década de 1960 pela IBM como uma forma de interligar grandes mainframes, visando obter uma solução comercialmente viável de paralelismo. Cada mainframe era um cluster, entretanto, o cluster ganhou força até que três tendências convergiram nos anos 1980: microprocessadores de alto desempenho, redes de alta velocidade, e ferramentas padronizadas para computação distribuída de alto desempenho. Uma quarta tendência possível é a crescente necessidade de poder de processamento para aplicações científicas e comerciais unida ao alto custo e a baixa acessibilidade dos tradicionais supercomputadores. (DOWD, SEVERANCE, 1998).

No final de 1993, Donald Becker e Thomas Sterling iniciaram um esboço de um sistema de processamento distribuído construído a partir de hardware convencional como uma medida de combate aos custos dos supercomputadores. No início de 1994, com o patrocínio do projecto HTPCC/ESS, criaram o primeiro cluster desse tipo, o projecto Beowulf.

Tal acontecimento levou a um desenvolvimento de software interfaces que tornavam mais fácil a coordenação, debug ou submissão de grandes programas (ou um grande número de programas independentes). Os mais rápidos (e maiores) supercomputadores são clusters. (SPECTOR, 2000; VRENIOS, 2002).

3.1 Tipos de Cluster

Alta Disponibilidade (High Availability (HA) and Failover): Esse tipo de cluster é utilizado para garantir que o serviço ou aplicação não saia do ar, quando um dos nós do cluster falhar, outro assume a tarefa no mesmo instante. Neste sistema temos geralmente

uma máquina inicial que corre as aplicações necessárias, e uma segunda que substitui a primeira no caso desta falhar. No caso de aplicação em base de dados, um primeiro servidor é dado como principal e corre a aplicação normalmente, existindo um segundo que vai mantendo uma cópia da informação e assume o papel principal caso haja uma falha de funcionamento no primeiro. Isto permite manter todo o sistema a funcionar sem prejuízo para utilizadores. Esta é uma configuração simples e existem numerosas formas de dividir os dados, podendo os computadores estar em salas, edifícios, países ou até continentes diferentes, de modo a evitar catástrofes naturais em larga escala ou acidentes físicos locais. (WEYGANT, 2001).

Balanceamento de carga (Load Balancing): Os sistemas de clusters orientados para o Load Balancing destinam-se a distribuir requisições para os diferentes nós do cluster de modo a aumentar o desempenho do sistema. Sempre que um nó falhar, as suas requisições serão redistribuídas para os nós disponíveis. Este tipo é muito utilizado na Internet, em servidores de e-mail, comércio eletrónico e em sistemas de lojas. (VRENIOS, 2002).

Aplica-se a um cluster quando há uma máquina que fornece um determinado serviço e, por alguma razão, esta deixa de conseguir (ou nunca conseguiu) dar resposta (satisfazer todos os pedidos). Montam-se então várias máquinas para fornecer esse mesmo serviço e dividem-se os pedidos entre elas. Um exemplo: uma máquina que executa um Web Server pode não conseguir satisfazer todos os pedidos de quem visita as páginas alojadas em disco. Neste caso existe a hipótese de montar outra máquina com o mesmo Web Server e uma cópia exata do disco do servidor inicial. (As máquinas encontram-se interligadas e uma delas (slave) irá modificar o seu conteúdo em disco sempre que algo for mudado no master.) Finalmente pode-se direcionar os visitantes para um ou outro servidor conforme o acesso de cada computador, evitando a sobrecarga de cada um dos servidores. (WEYGANT, 2001; DOWD, SEVERANCE, 1998).

4. Trabalhos relacionados

A Case for High Performance Computing with Virtual Machines

Escrito por Wei Huang, Jiuxing Liu, Bulent Abali e Dhabaleswar K. Panda o artigo apresenta os benefícios da virtualização em cluster utilizando o Xen como software de virtualização. Nesta pesquisa é demonstrado uma breve comparação entre os diferentes ambientes onde pode-se utilizar a virtualização como meio para reduzir custos e maximizar o gerenciamento das máquinas virtuais. Infelizmente não é abordado mais a fundo como uma ambiente de máquinas virtuais compartilha-se em cluster quando há alguma indisponibilidade de máquinas físicas. (HUANG, 2009).

Neste artigo foi proposto o seguinte teste: teste de latência, ou seja, foram configurados dois servidores com o software de virtualização xen, foi criada uma máquina virtual em um dos servidores, e foi realizada a migração da máquina virtual para o outro servidor, foi utilizado o serviço de ping para verificar quantos pacotes de dados foram perdidos entre a migração de um servidor para o outro.

Task-aware Virtual Machine Scheduling for I/O Performance

Escrita por Hwanju Kim, Hyeontaek Lim, Jinkyu Jeong, Heeseung Jo e Joowon Lee a pesquisa traz bem detalhadamente os procedimentos a serem tomados para garantir o

melhor desempenho de máquinas virtuais utilizando o Xen como software de virtualização. Esta pesquisa traz um breve resumo do funcionamento desse software, e como o mesmo funciona para virtualização de máquinas. (KIM, 2009).

Depois de repassado todo o funcionamento desse software, é demonstrado o procedimento que deve-se seguir para assegurar que o Xen utilize a maior capacidade de processamento, acesso a disco e rede possível na máquina física onde o mesmo está sendo executado. É bastante enfatizada neste trabalho a forma de como o Xen trabalha com o acesso a disco e rede, e do ganho de desempenho utilizando-o.

No artigo citado, é proposto testes de desempenho utilizando transferência de arquivos entre máquinas virtuais com isso é abordado o desempenho de I/O (input/output) de disco das máquinas virtuais, e desempenho de rede. (KIM, 2009).

VirtualCluster : Customizing the Cluster Environment through Virtual Machines

Escrito por Kang Chen, Jun Xin, Weimin Zheng, este artigo aborda a utilização de Virtual Cluster e virtualização, demonstrando como implementar uma estrutura que utilize ambas as tecnologias. Trás uma breve pesquisa sobre os softwares utilizados atualmente para criar cluster. Em seguida são abordados vários pontos sobre a clusterização, como por exemplo a interface de administração do cluster. (CHEN, 2009).

O artigo propõe uma estrutura de virtualização em cluster utilizando um storage onde são armazenados os discos virtuais das máquinas virtuais, com isso é proposta uma análise do tempo gasto na transferência de uma máquina virtual para outro nó do cluster, ou seja é verificado quanto tempo uma máquina física demora para assumir uma máquina virtual, estando seu disco virtual em um storage, onde todos os nós do cluster tem acesso. (CHEN, 2009).

5. Resultados obtidos

A a implantação realizada foi possível obter uma pesquisa detalhada sobre as vantagens da utilização do Hyper-v em comparação aos mais utilizados no mercado, como o VMWARE ESXi e Virtual Server. A escolha pelo hyper-v vem de uma necessidade da empresa RedeHost, para a qual o autor trabalha, em obter mais informações sobre essa ferramenta. A escolha pelo VMWARE ESXi foi pelo fato de ser um dos softwares de virtualização mais conhecidos hoje em dia, e que é utilizado por grandes empresas para ambiente de cloud computing. O Virtual Server foi escolhido devido ao fato da empresa já vir trabalhando com essa ferramenta. Através dos resultados obtidos foi possível definir as vantagens na utilização de uma ou outra ferramenta.

Tendo como base os trabalhos descritos no capítulo anterior, foi possível analisar quais os testes poderiam trazer mais informações úteis para o ambiente proposto, onde necessita-se de alta disponibilidade de informações e tolerância a falhas.

Para que possa ficar mais claro, a máquina virtual que estará no nó do cluster será um webserver, ou seja, um servidor de páginas web, com isso será possível verificar quanto tempo um serviço crítico, como é um site, por exemplo, ficará indisponível. A figura 1 ilustra como será o cenário.

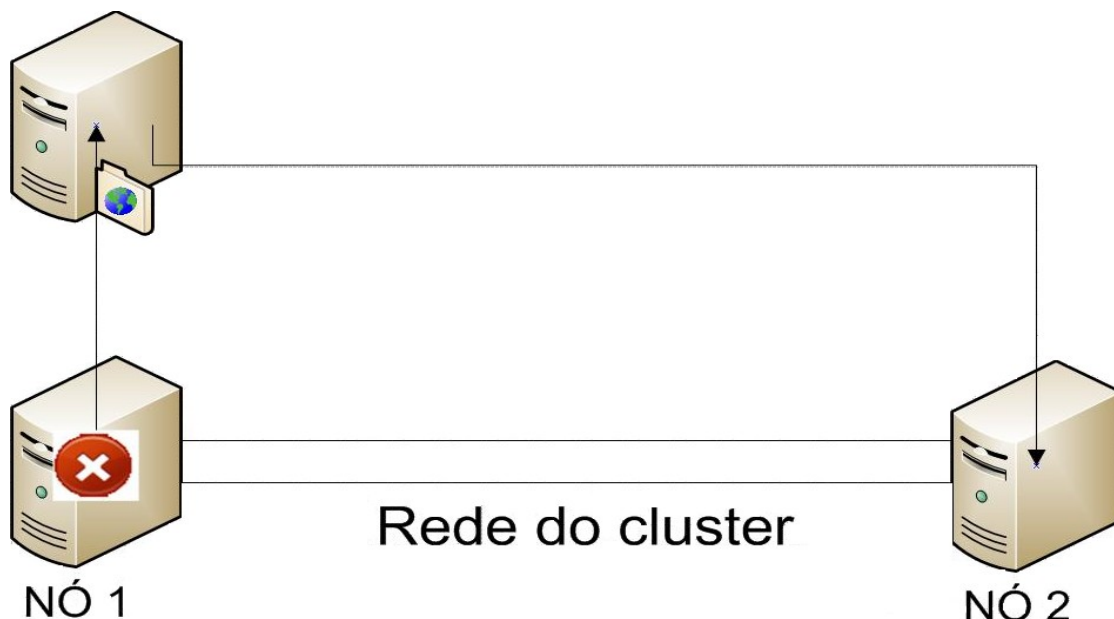


Figura 1: Cenário de teste.

O objetivo do trabalho foi de mesclar soluções já testadas e ferramentas que ainda não são tão conhecidas pela maioria dos profissionais, como é o caso do hyper-v, e trazer com a maior exatidão possível os resultados obtidos através dos testes:

- Testes de balanceamento de carga: Simular situações onde uma máquina física ficará sobre carregada com alto consumo de memória, e o processamento das máquinas virtuais que estarão sendo executadas na máquina física será dividido entre os demais nós do cluster. O Nível de processamento será medido em %, e através das ferramentas de virtualização será possível definir até quanto de processamento uma máquina física poderá ter antes do processamento da mesma ser distribuído entre os demais nós do cluster. O percentual adotado para esta pesquisa será de 50% de processamento na máquina física.

- Testes de Failover: Simular situações onde uma máquina física ficará indisponível, e a máquina virtual que estiver executando nessa máquina migrará para o outro nó do cluster. Para verificar qual das estruturas terá menor tempo de indisponibilidade de serviços, será medido, através do software XNote Stopwatch (que é um software shareware, ou seja, pode ser usado durante 30 dias gratuitamente, para cronometrar o tempo em milissegundo, segundos e minutos), em segundos (s) o tempo que levará para um nó do cluster assumir os serviços do nó que ficará indisponível.

- Teste de perda de pacotes: Para verificar quantos pacotes de dados foram perdidos na migração entre os nós do cluster, será disparada uma instrução de ping pelo segundo nó do cluster para a máquina virtual que esta sendo migrada, com isso será possível saber quantos pacotes foram perdidos na migração de uma máquina virtual para outro nó do cluster.

O ambiente utilizado para testes contava com os seguintes recursos:

- 2 Servidores cada um com 2 processadores Xeon QuadCore de 2.66GHz e 16GB de memória RAM.

Todos os testes propostos foram realizados em um laboratório criado dentro da

empresa RedeHost, e todo o procedimento foi documentado para incorporar essas informações na base de conhecimento da empresa. A tabela 1 ilustra os resultados obtidos nos testes.

<i>Software</i>	<i>Teste realizado</i>		
	<i>Teste de balanceamento de carga (Sucesso/Insucesso)</i>	<i>Teste de Failover (Tempo em segundos (s))</i>	<i>Teste de perda de pacotes (numero de pacotes perdidos)</i>
<i>Hyper-V</i>	<i>Sucesso</i>	<i>13</i>	<i>6</i>
<i>VMWARE ESXi</i>	<i>Sucesso</i>	<i>17</i>	<i>10</i>
<i>Virtual Server</i>	<i>Sucesso</i>	<i>15</i>	<i>12</i>

Tabela 1: Comparação entre os resultados

Como pode ser visto na tabela 1, ambos os softwares realizaram aquilo que foi proposto, que foi a questão do balanceamento de carga entre os dois servidores. Já nos outros dois testes o hyper-v saiu-se melhor, tendo um melhor desempenho no tempo de migração das máquinas virtuais entre os servidores físicos, e ativando a interface de rede mais rapidamente, garantindo assim um menor número de perda de pacotes.

6. Conclusão

Este trabalho possibilitará que profissionais da área de infra-estrutura de TI tenham maiores detalhes de como os softwares Hyper-V, Virtual Server e VMWare ESXi funcionam. Procurar através dos testes fornecer informações confiáveis sobre a virtualização de ambos os softwares, isso fará com que esses profissionais trabalhem com maior certeza do que estão implantando e terão dados dos testes feitos para descobrir qual dos softwares garante um ambiente com pouco tempo de indisponibilidade de serviços e sistemas operacionais.

O trabalho trás também um relato sobre como é a integração entre as tecnologias de virtualização e clusterização nos softwares escolhidos, essa integração é bastante difundida, mas há poucos casos onde a mesma é utilizada.

Como sugestão para uma futura implementação, fica a realização dos mesmos testes em outros softwares de virtualização, alguns cogitados é o Xen da Citrix, o que proporcionará maiores informações sobre o uso da tecnologia de virtualização e clusterização, e possibilitará, a quem desejar, ter uma fonte mais abrangente de informação sobre esses softwares.

Este trabalho teve como foco verificar o comportamento dos softwares escolhidos em um ambiente de alta disponibilidade e tolerância a falhas, e certamente servirá como documento de análise para futuras implementações comerciais dos mesmos.

7. Referências

Este capítulo tem como objetivos trazer as fontes consultadas para a realização dessa pesquisa, assim como demais informações sobre as mesmas.

Referências