9ª MOSTRA ACADÊMICA UNIMEP
08 a 10/11 de 2011
Tema:
"Ambiente e Sustentabilidade"

9º Congresso de Pesquisa

O IMPACTO DO SERVIÇO DE NAT E FIREWALL NO ATENDIMENTO DE REQUISIÇÕES WEB

Autor(es)

JOSE LUIS ZEM

1. Introdução

Atualmente é impensável o dia-a-dia sem o uso de computadores, principalmente computadores interligados através de redes locais ou de abrangência global. Se por um lado essa interligação favorece a troca de informações entre os computadores e pessoas que os operam por outro lado trazem uma preocupação adicional com tais informações compartilhadas, como a segurança no acesso e a manipulação das mesmas.
Assim, mecanismos destinados a prover essa segurança são muito comuns nestas redes e dois deles que se destacam, no caso o firewall e o serviço de tradução de endereços (NAT). Em função da grande utilização destes mecanismos torna-se interessante, e importante, conhecer quais as influências trazidas por essas tecnologias sobre as redes de computadores e serviços fornecidos por ela.

2. Objetivos

Desta maneira, o principal objetivo deste artigo é o de apresentar o relato sobre a realização de um experimento onde se buscou mensurar o impacto ocasionado no desempenho de um servidor web ao se utilizar de um servidor de NAT e de um firewall, tanto em ambientes com computadores reais e também virtuais.

3. Desenvolvimento

Em termos gerais, firewalls são barreiras interpostas entre a rede privada da organização e a rede externa, podendo ser baseados na combinação de hardware e software ou somente em software [Tanenbaum, 2009] [Tanenbaum, 2003]. Eles analisam o tráfego entre a rede interna e a rede externa, em tempo real, permitindo ou bloqueando o tráfego de acordo com as regras definidas previamente.
São utilizados como um dos principais instrumentos de defesa de uma rede corporativa, controlando o acesso aos sistemas e computadores da organização através do estabelecimento de regras e a filtragem de tráfego entre duas redes.
Podem ser implementados para a filtragem de serviços, controle de acesso, registros e estatísticas de utilização da rede como um todo, além de impor uma política de uso e de acesso à rede de computadores.
Conforme a Internet e seu uso cresceram ficou evidente que haveria problemas com o sistema de endereçamento utilizado, principalmente no tocante a disponibilidade de endereços válidos para todos os usuários.
Algumas alternativas foram disponibilizadas, como por exemplo a utilização de subredes e a implementação de um mecanismo que permita que múltiplos computadores de uma mesma rede possa operar ao mesmo tempo com apenas um endereço globalmente válido [Comer, 2007]. A este mecanismo dá-se o nome de Tradutor de Endereços de Rede ou NAT (Network Address Translator).
Essa solução, que também pode ser viabilizada por hardware ou software, torna possível que uma rede conectada à Internet tenha um endereço IP único e válido e o compartilhe com múltiplos computadores, sem nenhum conflito. Para evitar tais conflitos de endereço o serviço NAT atribui para cada computador um endereço local único e para evitar o consumo de endereços IP, os endereços locais são privados, o que significa que não são válidos na Internet. O serviço traduz o endereço em todos os pacotes, de forma que os computadores na Internet nunca vejam os endereços privados [Comer, 2007].
Para a realização foram preparados seis cenários, variando-se de dois a três computadores presentes em cada situação analisada. Os

computadores reais empregados tinham como configuração básica:
•Processador Intel Core 2 Duo e7400 @ 2.80 GHz.
•Memória RAM de 4 GB.
•Disco rígido de 160 GB.
•Sistema operacional LINUX e MS-Windows XP.
Como infraestrutura da rede de comunicação que interligou os computadores foi utilizado um chaveador (switch) com portas de 10 Mbps e de 100 Mbps, com suporte a frames Ethernet e Fast Ethernet. O computador que gerou as requisições foi conectado à porta de 100 Mbps e os demais computadores foram conectados às portas de 10 Mbps.
Para implementar a máquina virtual, utilizada em algumas situações, utilizou-se o ambiente de virtualização (hypervisor) VirtualBox [Oracle, 2010], sendo que o computador virtual, criado para alguns dos cenários de teste, tinha como configuração básica:
•Processador com núcleo único.
•Memória RAM de 256 MB.
•Disco rígido de 1GB.
•Sistema operacional Linux.
Nos equipamentos Estação (emissor das requisições web), Servidor web e Servidor NAT utilizou-se o sistema operacional Linux, em especial a distribuição Slax [Slax, 2010]. Já o equipamento onde foi instalado o hypervisor utilizou-se do sistema operacional MS-Windows XP.
O primeiro cenário construído e apresentado na Figura 01A (Cenário A) consistiu em dois computadores pertencentes a uma mesma rede de comunicação, ou seja, tendo apenas o chaveador como intermediador da comunicação entre eles. Neste cenário não havia a presença do serviço de firewall.
O segundo cenário, representado na Figura 01B (Cenário B), difere pouco do Cenário A, ou seja, dois computadores interligados na mesma rede de comunicação, porém o equipamento que atua como servidor web também executa um serviço de firewall.
O terceiro cenário (Cenário C, Figura 01C) foi uma derivação do Cenário A, ou seja, um computador atuando como emissor de requisições web, um equipamento atuando como servidor web (sem o serviço de firewall) e, agora, um terceiro equipamento, atuando como um servidor de NAT. Todas as requisições que eram enviadas ou recebidas pelo equipamento Estação passavam, obrigatoriamente, pelo Servidor de NAT.
O quarto cenário (Cenário D, Figura 01D) tem o mesmo arranjo e características do Cenário C, porém adicionando-se o serviço de firewall junto ao equipamento que atua como servidor web.
O quinto cenário (Figura 01E, Cenário E) consistiu em implementar o Servidor web em uma máquina virtual, sem a presença do serviço de NAT e no sexto cenário (Figura 01F, Cenário F) acrescentou-se o serviço de firewall na máquina virtual.
Ao longo deste artigo faz-se referência apenas às identificações dos cenários (A,B,C,D,E,F) e não mais às figuras. Onde ser referenciar (x,y) por exemplo, deseja-se mencionar a comparação entre o cenário x contra o cenário y.

4. Resultado e Discussão

Os testes consistiram, basicamente, em enviar, a partir de um computador (Estação), uma quantidade de requisições web a um outro computador (Servidor web) e, em seguida, coletar as estatísticas sobre o tempo demandado para executar a bateria total de testes ou para cada lote de requisições, o uso da rede e a quantidade de erros. Eventualmente, ativou-se ou não o serviço de firewall ou então adicionou-se, ou não, um equipamento intermediário para interceptar tais requisições.
A intenção destes testes foi a de confrontar os ambientes com e sem o serviço de firewall, com e sem o serviço de NAT, a utilização do computador real e do virtual e detectar a variação de desempenho nas referidas situações.
Para se gerar as requisições, em quantidade e velocidade suficientes, foi empregada uma ferramenta de stress-test para ambientes web chamada httperf [Httperf, 2010] e para se produzir os lotes de requisições construiu-se um script específico para tal função.
Para cada cenário foram executadas três baterias de testes, sendo que os valores finais usados para as análises comparativas foram obtidos através da média aritmética das referidas baterias, para o cenário analisado.
Para se obter os tempos totais consumidos pelas baterias de testes empregou-se o comando time [Time, 2010] do Linux. Esse comando emite como resultado uma representação do tempo total consumido entre o início e o encerramento da execução de um comando que lhe é informado como parâmetro. A representação de todos os tempos totais gastos para a execução dos testes, em todos os cenários (A,B,C,D,E,F), pode ser encontrada na Figura 02A.
Ao analisar os valores apresentados pela Figura 02A é possível perceber que, entre os cenários (A,B) a diferença dos valores obtidos é mínima, sendo que os testes tiveram sua execução mais rápida no cenário B que no cenário A. Isto também foi presenciado quando se compara os cenários C com o D. Porém, na comparação dos cenários E e F observa-se que a diferença no tempo a favor do ambiente que não possuía o serviço de firewall.
Quando se comparam os tempos gastos nos cenários sem a presença do serviço de firewall (Figura 02B) nota-se uma diferença de tempo de execução considerável entre os cenários (A,C) e (A,E).
E quando se comparam os tempos totais gastos nos cenários onde o serviço de firewall estava ativado (Figura 02C) observa-se novamente, uma diferença de tempo de execução considerável entre os cenários (B,D) e (B,F).

Ao se analisar tais valores e diferenças percebe-se que o impacto causado pela presença ou não do serviço de firewall não é significativo, ou melhor, nos cenários onde ele estava presente houve uma melhoria (redução) nos tempos totais de execução dos testes. Porém, o impacto causado pela presença do NAT ou pela implementação do servidor em máquina virtual foi significativo nas medições dos tempos totais, ou seja, um ocasionou um aumento nos tempos totais gastos.

O aumento nos tempos totais de execução para os cenários que não contavam com a presença do firewall foi de, aproximadamente, 89,33% entre os cenários (A,C) e de 58,48% entre os cenários (A,E). Já para os cenários que contavam com a presença do firewall aumentou-se, aproximadamente, em 89,38% entre os cenários (B,D) e em 60,68% entre os cenários (B,F).

Para melhor entender a razão desse aumento nos tempos totais de execução dos testes decidiu-se analisar o comportamento dos cenários a partir dos lotes de requisições individuais que foram enviados aos servidores nos respectivos cenários.

Para extrair os dados dos relatórios gerados pelo httperf e observar o uso dos recursos (tempo de processamento e rede, bem como a quantidade de erros ocorrida) consumidos em cada lote de requisições foi necessário construir um novo script.

Ao se extrair os dados individuais dos lotes, para todos os cenários, e confrontá-los foi possível ilustrar o comportamento dos mesmos na Figura 03A. Nela observa-se uma linearidade nos resultados entre os cenários (A,B) e entre os cenários (C,D), também é possível observar a diferença detectada nas comparações entre os cenários A e C e entre os cenários B e E. O que chamou a atenção na referida figura foi a alternância no comportamento dos cenários (E,F).

As diferenças nos tempos totais gastos, apresentadas nas Figuras 02B e 02C também podem ser verificadas nas Figuras 03B e 03C, exceto pela oscilação dos cenários E e F. Assim, o comportamento observado nas representações dos tempos totais gastos coincidem com os tempos gastos individualmente em cada lote de requisições e a mesma consideração sobre o impacto, praticamente nulo, do firewall nos cenários estava presente, além do aumento ocasionado pelo serviço de NAT e pelo uso da máquina virtual.

Uma vez que o impacto do firewall foi muito pequeno e a inserção do NAT produziu um impacto considerável decidiu-se verificar se o uso da rede de comunicação e a manipulação dos pacotes pelo NAT poderiam ser as razões desse acréscimo nos tempos totais gastos para as execuções. Os dados extraídos sobre a utilização da rede de comunicação são apresentados na Figura 03D.

É possível perceber na Figura 03D que os cenários (A,B) tiveram um melhor aproveitamento no uso da rede de comunicação. Já os cenários que utilizaram do NAT, cenários (C,D) no caso, obtiveram uma taxa de transmissão menor em relação à capacidade de comunicação dos cenários (A,B).

Quanto ao uso da rede de comunicação nos cenários (E,F) detectou-se no início uma utilização superior àquela dos demais cenários, porém, a partir do lote com 1750 requisições, percebe-se uma oscilação para níveis muito baixos em sua capacidade de uso.

Com base nestes comportamentos percebe-se que o uso do serviço de NAT pode produzir um impacto (negativo) considerável no provimento de serviços em uma rede de computadores, mas com respeito ao uso de máquinas virtuais, os valores não permitem afirmações conclusivas.

Para tentar esclarecer o ocorrido com os cenários (E,F) quanto ao uso da rede de comunicação buscou-se recuperar os dados referentes à quantidade de erros ocorrida durante os testes e verificar se a mesma tinha alguma relação com a oscilação.

É mostrada na Figura 03E que essa influência realmente ocorre pois a oscilação presente nas Figuras 03D e 03E estão relacionadas, ou seja, quando da ocorrência do erro existe uma utilização menor da rede de comunicação por parte dos cenários (E,F). Já com os cenários (A,B) e (C,D) não foram detectados erros.

A questão referente à quantidade de erros, representada na Figura 03E, está relacionada também com os tempos totais gastos, ou seja, em função da quantidade de erros houve uma menor utilização da rede, gerando muitas vezes atrasos no retorno das requisições e isso interferiu diretamente nos tempos finais das execuções.

5. Considerações Finais

Com base nos testes realizados e nos resultados obtidos conclui-se que o uso de firewalls em ambientes web não causa impacto no desempenho do serviço, muito pelo contrário, pois nos testes os ambientes que faziam uso desse recurso apresentaram resultados melhores do que aqueles que não o utilizaram.

O que chamou a atenção foi o comportamento do NAT e o seu impacto perante a rede de comunicação. Nos ambientes onde ele estava presente, presenciou-se um aumento no tempo de retorno das requisições, concluindo assim que o uso deste recurso causa um impacto considerável no desempenho do servidor web.

E finalmente a adoção de máquina virtual no lugar da máquina real não se mostrou tão atraente, pois o desempenho ficou em um nível abaixo daquele obtido pelo computador real.

Como continuidade desse estudo seria interessante que fosse ampliado os testes descritos nesse artigo com uma quantidade maior de requisições para se verificar se o padrão observado nos experimentos se mantém. Também seria interessante utilizar outras velocidades e tecnologias na rede de comunicação para se verificar se em algum momento ocorre alguma inversão daquilo que se obteve nos testes realizados.

Além disso, seria interessante realizar os experimentos com outros serviços como servidores de banco de dados, de transferência de arquivos, processamento entre outros para certificar se os resultados aqui apresentados se restringem apenas ao serviço de web ou não.

Referências Bibliográficas

Comer, Douglas E. (2007) “Redes de Computadores e Internet”, 4ª Edição, Porto Alegre:Editora Bookman, Brasil.
Httperf (2010) “Welcome to the httperf home page”,http://www.hpl.hp.com/research/linux/httperf, Junho.
Oracle (2010) “Virtualbox”, http://www.virtualbox.org, Junho.
Slax (2010) “Slax: your pocket operating system”, http://www.slax.org, Junho.
Tanenbaum, Andrew S (2003) “Redes de Computadores”, 4ª Edição, Rio de Janeiro:Editora Campus, Brasil.
Tanenbaum, Andrew S (2009) “Sistemas Operacionais Modernos”, 3ª Edição, São Paulo:Editora Pearson Prentice Hall, Brasil.
Time (2010) “Linux/Unix Command: time”, http://linux.about.com/library/cmd/blcmdll_time.htm, Junho.

Anexos

Figura 03 – Comparação entre Tempos

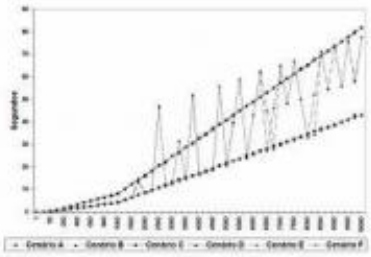


Figura 03A. Tempos gastos em cada lote e obtidos em todos os cenários.

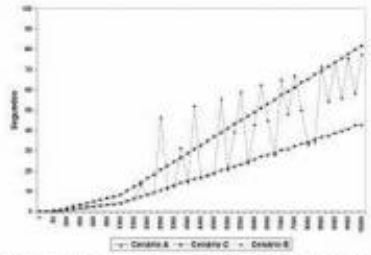


Figura 03B. Comparação entre os tempos gastos em cada lote nos cenários sem firewall.

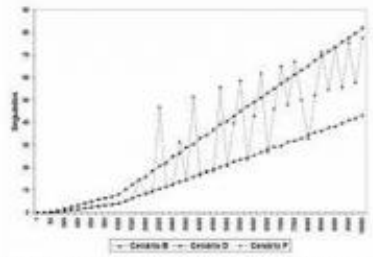


Figura 03C. Comparação entre os tempos gastos em cada lote nos cenários com firewall.

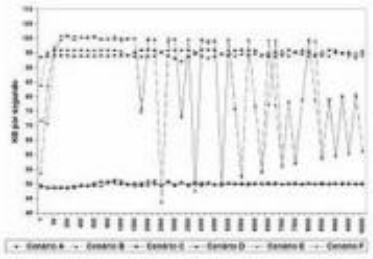


Figura 03D. Uso da rede em todos os cenários.

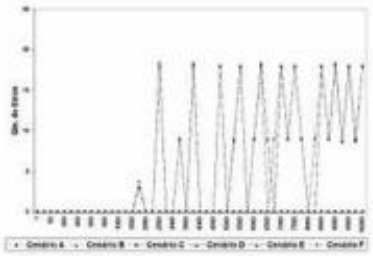


Figura 03E. Ocorrência de erros em todos os cenários.

Figura 01 – Cenários

(A) - Estação e Servidor *web* sem *firewall*, na mesma rede de comunicação.

(B) - Estação e Servidor web com *firewall*, na mesma rede de comunicação.

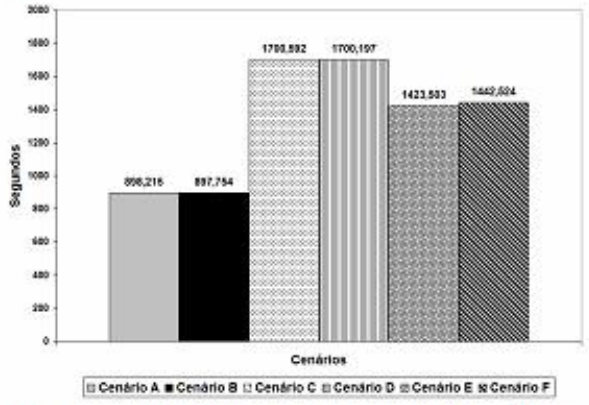
(C) - Estação e Servidor Web sem *firewall*, em redes de comunicação diferentes.

(D) - Estação e Servidor Web com *firewall*, em redes de comunicação diferentes.

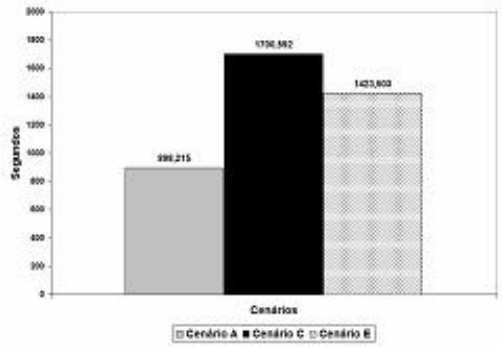
(E) - Estação (máquina real) e Servidor *web* sem *firewall* (máquina virtual), na mesma rede de comunicação.

(F) - Estação (máquina real) e Servidor *web* com *firewall* (máquina virtual), na mesma rede de comunicação.

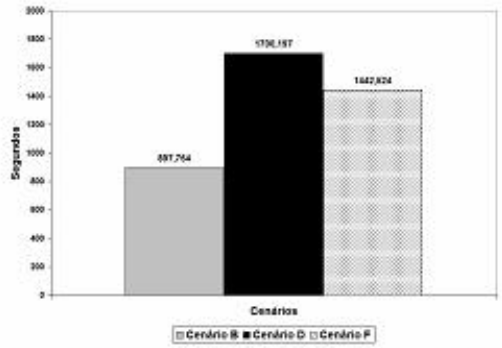
Figura 02 – Tempos Totais



(A) Tempos totais gastos obtidos em todos os cenários.



(B) Comparação entre os tempos totais gastos nos cenários sem firewall.



(C) Comparação entre os tempos totais gastos nos cenários com firewall.