

Desempenho de servidores desktop WEB TV.

Rodrigo de Azevedo Carvalho, Atila Bohlke Vasconcelos

Sistemas de Informação – Centro Universitário Ritter dos Reis (Uniritter)
CEP 90840-440 – Porto Alegre – RS – Brasil

carioca_rs@yahoo.com.br, atila_vasconcelos@uniritter.edu.br

***Resumo.** Este artigo tem o objetivo de abordar os principais fatores que influenciam uma transmissão streaming através de redes (Local, MANs, WANs), com o intuito de obter amostras reais e supostas do desempenho de um servidor desktop. É demonstrado através das técnicas de medição de desempenho, o comportamento do servidor em situações normais, stress e suposições para uma maior abrangência em sua capacidade, assim como a utilização em arquiteturas variadas.*

1. Introdução

Atualmente está se tornando comum o fato de ouvir estações de rádio e assistir TV via Internet. Esta facilidade é proporcionada pelo contínuo aumento na largura de banda na Internet disponível para a população. Pessoas que possuam as ferramentas corretas podem servir e/ou usufruir desses serviços.

Os computadores pessoais (PCs-Desktop) modernos têm evoluído rapidamente, com maiores capacidades de memória, barramentos internos de alto desempenho e processadores com múltiplos núcleos. Essa evolução dá aos usuários finais a capacidade de prover e usufruir de serviços antes só implementados por empresas com computadores servidores dedicados. O computador residencial antes utilizado para ler e-mails e escrever textos, ganha novas funcionalidades, proporcionando uma divulgação multimídia pessoal através de redes e até da Internet.

Este artigo mostra diversos testes com ênfase na avaliação do desempenho e na difusão dos servidores de Web TV. Estes testes servem como forma momentos em que o servidor necessitará de mais desempenho para atender maior número de clientes conectados simultaneamente. O artigo pode e deve ser usado como forma de comparação com artigos e trabalhos anteriores, para um entendimento mais amplo sobre a evolução tecnológica nesta área.

2. Referencial Teórico

Atualmente, as redes de computadores vêm crescendo de forma considerável, empresas e residências estão se integrando cada vez mais nesta área para conduzir seus negócios e seus entretenimentos, de forma mais ágil e lucrativa. Muito desse crescimento pode ser creditado a Internet, o ponto em comum de todas as redes mundiais, sejam empresariais ou residenciais.

Mesmo com baixas velocidades foi analisado por empresas da área que o número de downloads de músicas através da Internet era crescente. Em 1995, a Progressive Networks começou a se popularizar com seu software Real Audio que

continha um decodificador, um servidor e um reproduutor de áudio (Kurose e Ross, 2003).

Existem controvérsias quanto às formas de transmissão de streaming; muitos profissionais e autores como Avila (2008) e outros, defendem que existem apenas duas formas:

- Sob demanda
- Em tempo real

Outros, como Forouzan (2006), Kurose e Ross (2003), defendem a existência de três formas:

- Sob demanda
- Tempo real
- Interativo

Streaming sob demanda ocorre quando o arquivo fica armazenado em um servidor, onde pode ser acessado diversas vezes. Ao acessá-lo o usuário faz uma ação que é muito próxima a um download, mas salvando em uma memória temporária (o Buffer), que terá seu conteúdo descartado após sua visualização. Exemplos dessa forma de distribuição são os web sites Youtube.com e Globo.com.

Em tempo real, como o próprio nome diz, consiste na recepção e exibição do fluxo de dados – o “streaming” – ao mesmo tempo em que este é enviado. O computador armazena um pequeno trecho de vídeo em sua memória temporária, o suficiente apenas para garantir a sincronização das imagens e sons. Deste modo, recursos como rever cenas, parar temporariamente o vídeo (“pause”) ou retroceder não são possíveis. Diversas rádios na Internet já transmitem sua programação desta forma, como por exemplo, a Rádio Globo, POP Rock e Jovem Pan.

Toda essa revolução pôde ser realizada por conta de aumentos constantes e significativos nas velocidades de acesso a Internet. Hoje no Brasil é possível atingir taxas que vão até 20 Mbps, temos total condição de colocar em prática projetos de streaming vídeo com significativo nível de qualidade. A prova dessas alternativas é a popularidade que os vídeos sob demanda têm atingido em sites como Globo.com, Youtube.com entre outros. Essa velocidade vem mostrar que quaisquer usuários, além de usufruir de serviços, também podem prestar tais serviços de modo a transmitir eventos familiares, festas, aulas, propagandas e qualquer outra informação que queiram compartilhar com outras pessoas na rede.

3. Avaliação e desempenho do sistema

A análise de desempenho de sistemas computacionais pode ser entendida como uma combinação de medidas e interpretações sobre diversas características do sistema, como desempenho, velocidade de comunicação, tamanho dos dados, etc. (Jain, 1991). Os objetivos da análise de um sistema computacional dependem da situação específica, dos interesses e da habilidade do analista (Lilja, 2000). Para uma correta análise de desempenho computacional é preciso optar por quais componentes serão abordados, e quais as variantes que podem influenciar nesse processo. Essas opções limitam a análise

a um foco apropriado de interesse, dando ênfase aos componentes de maior influencia no sistema a ser testado.

O sistema de um computador é muito complexo e pode conter oscilações que possivelmente acarretarão erros na precisão das métricas adotadas para cada componente. Por exemplo, uso do processador, quantidade de memória alocada, etc. Para uma análise efetiva, é necessário desenvolver técnicas que perturbem o mínimo possível o sistema a ser analisado (Vasconcelos, 2006). Assim é preciso estabelecer objetivos de forma a organizar a análise de maneira mais adequada.

De acordo com David J. Lilja (2000) existem alguns objetivos comuns, são eles:

- Comparação das alternativas: Ao estabelecer qual será o hardware necessário para o uso de determinada aplicação é possível cair em muitos impasses tais como, quantidade e velocidade de memória RAM, clock, arquitetura e quantidade de núcleos do processador, velocidade de placas de rede, dentre outros aspectos. Isso leva a ter de optar qual a melhor alternativa para o tipo de aplicação a ser usada e como deixar os requisitos prontos para um futuro upgrade.
- Determinação de impacto de um componente: Com a evolução tecnológica agindo de maneira tão acelerada, é prudente ter em vista o impacto que a adição, remoção ou substituição de um componente pode causar em seu sistema.
- Ajuste do Sistema: Este item se refere à análise do sistema, de forma a ajustar valores de parâmetros para tentar atingir o melhor desempenho.
- Identificação de desempenho relativo: Através da base de dados obtidos anteriormente, identificar o desempenho relativo do sistema.
- Análise de problemas de desempenho: O bom funcionamento do programa é um pré-requisito fundamental, pois é imprescindível verificar problemas e desperdícios e saná-los antes que interfiram na produtividade do serviço. Uma vez o sistema estando instável provavelmente suas medidas não demonstrarão com exatidão o que deve ser apresentado.
- Estabelecer expectativas: Toda pessoa que lida com computadores, normalmente tem uma idéia de como será a próxima geração de computadores. A questão pertinente neste caso é ajustar apropriadamente o sistema para a tarefa que ele deve fazer.

Quando alguém se confronta com o desafio de medir hardware, existem três técnicas fundamentais que podem ser utilizadas para encontrar as soluções desejadas. Essas são:

- Medida em um sistema existente
- Simulação do sistema
- Modelagem analítica

Normalmente medidas em um sistema existente, fornecem os melhores resultados desde que usadas as ferramentas necessárias para a função, não simplesmente supor o que precisa ser feito. Essas características também fazem os resultados baseados

nas medições serem mais confiáveis quando são apresentados. Realizar medidas em sistemas reais não é muito flexível, uma vez que estas medidas proporcionam informação apenas sobre sistema medido, impossibilitando variações na configuração. Por exemplo, em muitos sistemas não é possível alterar as velocidades de *clock* do processador ou da memória. (Lilja, 2000)

A simulação é um instrumento para modelar as características importantes de um sistema. Assim um simulador pode ser descrito, como sendo nada menos que uma forma de teste que pode ser modificada para estudar o efeito causado por este impacto no sistema como um todo. O custo da simulação depende da complexidade do sistema que está sendo observado e do grau de detalhamento que está sendo dado ao simular tais funções. A primeira limitação na construção de um simulador é a impossibilidade de se modelar todos os pequenos detalhes do sistema que está sendo estudado. Consequentemente simplificações deverão ser supostas para que seja possível a adoção do simulador. Essas simplificações causarão limitações quanto à credibilidade do sistema avaliado, tornando-o suscetível a dados com leves desvios. (Lilja, 2000)

A modelagem analítica é uma descrição matemática do sistema. Comparada com as outras técnicas, medidas em um sistema existente e simulação de sistema, os resultados da modelagem analítica tendem a ser menos confiáveis e menos precisos. Entretanto as informações obtidas através desta técnica podem ajudar a obter maiores detalhes ao utilizar as outras técnicas. (Lilja, 2000)

Antes de começar a entender os aspectos de um sistema, é preciso determinar quais são realmente os itens importantes a serem medidos. As características básicas necessárias para fazer medições em computadores são:

- Contagem de repetições de um evento.
- A duração de algum intervalo de tempo
- Noções de alguns parâmetros.

Por exemplo, se for preciso medir quantas vezes um processador faz uma requisição de entrada/saída e se for de interesse medir quanto tempo essas requisições demoram isso certamente determinará o número de bits que podem ser transmitidos e guardados (salvos). Assim como nesse exemplo, os métodos podem ser aplicados para a maioria dos componentes de um computador em diversas situações como largura de banda, velocidade de memória e etc. Com essas formas de medir os valores, poderá ser obtido o valor desejado, que será utilizado para descrever o desempenho do sistema. Esses valores são chamados de métricas. (Lilja, 2000)

Se estiver interessado especificamente em contabilizar o tempo, contagem ou no tamanho do valor medido, o valor pode ser utilizado diretamente como a métrica de desempenho. Entretanto, se estiver interessado em normalizar a contagem dos eventos, é comum prover uma métrica de velocidade como em operações executadas por segundo. Esta forma de métrica é chamada de “rate metric” (métrica de taxa) ou “throughput” (tempo de execução da aplicação) e é normalmente calculada dividindo a contagem de eventos que ocorreram em um intervalo, pelo tempo de intervalo que eles ocorreram. Esses processos são usados para comparar diferentes medidas vindas de diferentes intervalos de tempo. (Lilja, 2000)

4. Resultados Parciais

Este artigo referencia um trabalho de conclusão em andamento, por este motivo os resultados demonstrados abaixo são provisórios podendo sofrer alterações.

Foram testados diversos aspectos do servidor, dentre eles banda de rede local (LANs) para exibições corporativas ou educacionais, banda de Internet (WANs e MANs) para apresentações a longas distâncias, desempenho da CPU e alocação de memória. Para as medidas parciais foi utilizado o sistema operacional Windows XP sp2 como servidor, conectado a três tecnologias de rede distintas sendo elas LAN (ethernet cabos CAT5), LAN (Wireless 802.11g) e WAN (Cable Modem Virtua), o acesso a todas essas tecnologias foi estabelecido por um roteador Dlink DIR-300. Os componentes dos testes foram escolhidos tendo em vista tecnologias de acesso popular, sendo possível atingir melhores resultados com equipamentos mais robustos (que fogem ao foco deste trabalho).

Para esses testes foi utilizado um computador baseado na tecnologia desktop de duplo núcleo, sendo sua configuração: Athlon X2 6000+ (clock 3 GHz por núcleo), 2 GB RAM (DDR2 800 MHz), conexão com rede cabeada e sem fio (cabeada a 100 Mbps e sem fio a um máximo de 54 Mbps), conexão com a Internet (cable modem com banda de 3 Mbps download e 512 Kbps upload).

A compatibilidade com os sistemas operacionais Linux e Windows é fundamental para o processo, por conta da abrangência que esses sistemas atingem atualmente. Foi escolhido inicialmente o protocolo de fluxo contínuo mms (Microsoft media service). A adoção do mms se deve ao software escolhido, o Windows Media Encoder, o adotar como padrão. A compatibilidade do mms com ambos os sistemas operacionais possibilitou testes eficientes tanto no servidor como no cliente.

O vídeo foi codificado e transmitido utilizando o codec wmv (Windows Media Vídeo, padrão do encoder), a 23 quadros por segundo e uma resolução de 320 x 240 pixels. O áudio foi enviado utilizando 2 canais a uma taxa de 32 kbps. O arquivo originalmente encontrava-se a uma resolução de 640 x 336 pixels; com 29 quadros por segundo o áudio codificado em MP3 a 128 kbps em 2 canais.

Inicialmente foram feitos testes referentes a variação de banda de rede, para isso foram disponibilizados de um a cinco clientes formando a bateria de testes para a LAN cabeada. Abaixo são apresentados todos os resultados obtidos até o momento.

Para apenas um cliente foi obtida a variação demonstrada no gráfico abaixo:



Figura 1 – Utilização de Largura de Banda para um cliente

Os testes para WAN obtiveram os mesmos resultados dos demonstrados para um cliente e ao serem feitos com dois a banda que é exigida foi superior do que a disponível impossibilitando os testes. É provável que sejam obtidos resultados muito semelhantes em relação à utilização de banda de rede, a variação neste caso se daria em torno da latência.

Para dois clientes foi obtida a variação demonstrada no gráfico abaixo:



Figura 2 – Utilização de Largura de Banda para dois clientes

Para três clientes foi obtida a variação demonstrada no gráfico abaixo:



Figura 3 – Utilização de Largura de Banda para três clientes

Para quatro clientes foi obtida a variação demonstrada no gráfico abaixo:



Figura 4 – Utilização de Largura de Banda para quatro clientes

Para cinco clientes foi obtida a variação demonstrada no gráfico abaixo:



Figura 5 – Utilização de Largura de Banda para cinco clientes

A média de todas as medidas mostra um crescimento aproximadamente linear da largura de banda utilizada com o número de clientes, como é apresentado na figura 6:

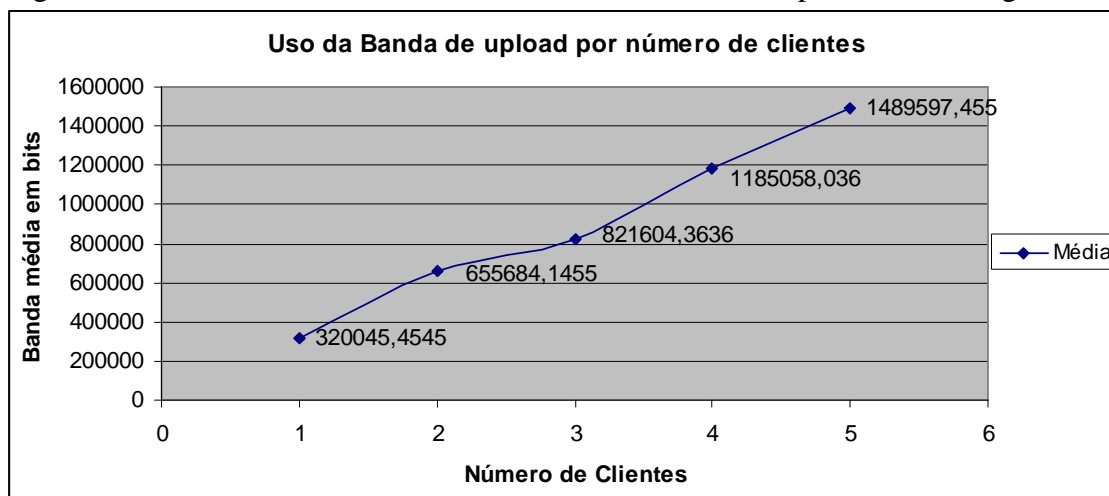


Figura 6 – Largura de Banda Utilizada versus Número de Clientes

Os testes com rede sem fio (Wireless), não puderam ser feitos, pois mesmo com taxas melhores na transmissão a tecnologia não permitiu uma visualização, por parte do cliente, sem cortes ou interrupções impossibilitando uma boa transmissão.

Não foi preciso demonstrar o uso de CPU de acordo com o aumento do número de clientes, pois durante os testes foi observado que não houve nenhuma variação com esse aumento:



Figura 7 – Utilização de CPU

Seguindo os testes pelo acompanhamento de alocação de memória durante todo o processo de transmissão, que é demonstrado abaixo:

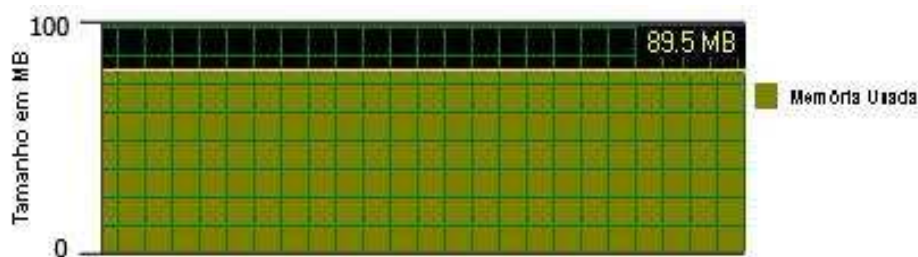


Figura 8 – Utilização de Memória

Para memória também não foi necessária a demonstração de alocação de acordo com o aumento do número de clientes, pois durante os testes foi observado que não houve nenhuma variação.

5. Considerações finais

Conforme os estudos feitos neste artigo, são apresentados os princípios da manipulação de som e imagem, assim como, formas de transmissão dos mesmos através de redes de computadores. Isto, com o objetivo de dar ênfase nos itens e situações que influenciam no desempenho de um servidor streaming. Visando esse desempenho, foram exibidas formas para a análise de um conjunto de hardwares do sistema, com o intuito de obter métricas comparativas para um estudo aprofundado de seu comportamento. São introduzidos conceitos básicos para uma análise completa de um servidor baseado em tecnologia desktop, voltado a transmissões de web tv em redes de abrangência, local (LAN), metropolitana (MAN) e redes de grande abrangência (WAN).

Também foram analisadas taxas e formas de compressão, através dos processos mais comuns, tais informações influenciam diretamente no processo de transmissão da imagem e som. Observaram-se imperfeições que podem ter sido ocasionadas por uma degradação durante o processo de transmissão ou pela forma em que a imagem e o som vêm a se comportar após o processo de codificação.

Referências

- AVILA, Renato Nogueira Perez (2008). **Aprenda a Criar e Instalar sua Radio**. Editora: Ciência Moderna, 1a Edição.
- Kurose, James F. e Ross, Keith W (2003). **Redes de Computadores e a Internet uma nova abordagem** Editora: Addison Wesley, 1a Edição.
- Forouzan, Behrouz A (2006). **Comunicação de Dados e Redes de Computadores** Editora: Bookman, 3a Edição.
- Patterson, David e Henessy, John. **Computer Organization and Design: The Hardware/Software Interface**. Editora: Morgan Kaufmann Publishers.
- Lilja, David J. (2000). **Measuring Computer Desempenho: A Practitioner's Guide** Editora: Cambridge University Press.
- Vasconcelos, Atila Bohlke (2006). **Modelo de Desempenho para Código com Desvios de Execução em Hardware Gráfico**. Universidade Federal do Rio Grande do Sul.
- Yang, Fan (2004). **Bit Allocation for Scalable Video Streaming over Mobile Wireless Internet**. IEEE InfoCom.
- IETF. **IETF Documents**. Disponível em: < <http://tools.ietf.org/html/>>. Acesso em 10 de Jun de 2008.