

Universidade Federal do Paraná

Departamento de Informática

Patrícia Sereda & Roberto A Hexsel

Estimativa de Banda de Arquivos de Vídeo
Através da Utilização de Filtros Passa-baixa

Relatório Técnico
RT-DINF 004/2003

Curitiba, PR
2003

Estimativa de Banda de Arquivos de Vídeo Através da Utilização de Filtros Passa-baixa

Patrícia Sereda & Roberto A Hexsel

Departamento de Informática
Universidade Federal do Paraná
Centro Politécnico, C Postal 19081 – 81531-990 Curitiba, PR
{patricia,roberto}@inf.ufpr.br

Abstract A video server must control which fraction of network capacity is used by the video streams being transmitted in order to avoid network congestion and to service as many concurrent requests as possible. This paper presents a method for deterministic bandwidth estimation of video files that improves network utilization by concurrent video streams, without losses in quality at the clients. Experiments show that network utilization increases between 40-55%, which in turn allows for 50-60% more requests being granted by the video server.

Resumo Este artigo demonstra a utilização de filtros passa-baixa para melhorar a estimativa de largura de banda de arquivos de vídeo que serão transmitidos por servidores através de uma rede local. A estimativa de banda de rede é realizada por um algoritmo determinístico para controlar a admissão de novas requisições, evitando que ocorram sobrecargas na interface de rede e possíveis atrasos na entrega de fluxos de vídeo a clientes conectados no servidor. Nos testes realizados, a utilização dos filtros possibilitou ganhos entre 50 e 60% no número de requisições aceitas, e melhora na utilização da capacidade da rede entre 40 e 55%.

Introdução

Um *servidor de vídeo* é um programa que disponibiliza fluxos de vídeo a clientes que reproduzem o vídeo sem a necessidade de armazenamento local. As requisições são recebidas a qualquer momento e referem-se a algum vídeo armazenado em disco, ou codificado ao vivo. Este tipo de serviço é conhecido como Vídeo sob Demanda [16].

Todos os recursos utilizados pelo serviço de distribuição de vídeo devem ser controlados com cuidado para que os servidores possam disponibilizar vários fluxos de vídeo simultaneamente. Dentre estes recursos, a banda de rede merece atenção especial. O método de controle de banda de rede implementado nos servidores depende do tipo de arquivo de vídeo disponibilizado. Arquivos de vídeo podem ter a banda de rede variável ou constante. Arquivos de vídeo de banda constante são pouco utilizados por causa de sua baixa qualidade, relativamente a arquivos de banda variável.

O controle de banda de rede em um servidor que disponibiliza arquivos com banda variável é dificultado pela própria variabilidade na banda necessária para o envio dos fluxos de vídeo. Podem ocorrer intervalos durante a transmissão em que a banda do vídeo é suficientemente alta a ponto de sobrecarregar a interface de rede e ocasionar atraso na entrega de fluxos de vídeo aos clientes.

Para evitar atrasos na entrega de fluxos de vídeo os servidores empregam algoritmos determinísticos de controle de banda de rede, que garantem que a interface de rede não seja sobrecarregada. Estes algoritmos estimam a utilização e reservam banda de rede para cada arquivo de vídeo com base na taxa máxima transmitida durante o intervalo de transmissão de todo o vídeo [7]. Geralmente, as estimativas superam o valor médio efetivamente transmitido, resultando na sub-utilização da rede e do servidor de vídeo propriamente dito.

Este artigo descreve um método para estimar a banda de vídeos baseado em filtros passa-baixa que produz estimativas mais próximas dos valores médios, permitindo assim melhor utilização da banda de rede e aproveitamento do servidor [13]. Os testes aqui descritos mostram ganhos na utilização da rede da ordem de 56%, e ganhos no número de requisições aceitas pelo servidor da ordem de 63%.

O método de estimativa de banda com filtro passa-baixa foi implementado no *Stored Video File server (SVFserver)* [13] para controlar a utilização da banda de rede. O servidor *SVFserver* transmite apenas a exata quantidade de dados para a reprodução dos quadros no cliente, sem enviar dados adiantadamente, o que elimina a necessidade de armazenagem temporária dos quadros nos clientes. O *SVFserver* somente transmite fluxos de vídeo a partir de arquivos armazenados em disco, e opera com maior eficiência em redes locais.

O texto está organizado em 5 seções. A seção 1 contém conceitos referentes a vídeo digital e a servidores de vídeo. A seção 2 analisa o método determinístico para controle de banda de rede. A seção 3 descreve os filtros passa baixa propostos. A seção 4 mostra os resultados da aplicação do filtro em um servidor de vídeo. A seção 5 apresenta as conclusões.

1 Servidores de Vídeo

Vídeo digital é a composição de uma seqüência de imagens estáticas, sincronizadas a um sinal de áudio. Cada imagem estática é chamada de *quadro*. Para se obter a aparência de movimento das imagens, vários quadros com pequenas diferenças entre si são reproduzidos em seqüência a uma taxa maior ou igual a 24 quadros por segundo para se obter o efeito de persistência visual [5].

Um quadro de vídeo digital é composto por elementos de imagem chamados *pixels*. Em uma imagem digital cada *pixel* é composto por três tipos de informações. No padrão *RGB*, cada *pixel* é composto de informações correspondentes às três crominâncias *Cr-Red*, *Cg-Green* e *Cb-Blue*, sendo cada informação geralmente representada por 8 bits. No padrão *YCrCb*, cada *pixel* é composto de informações correspondentes à luminância *Y* e às crominâncias *Cr-Red* e *Cb-Blue*, sendo a luminância representada por 16 bits e cada crominância representada por 8 bits.

Cada quadro é representado por uma grande quantidade de informações, e portanto um vídeo usa muito espaço para seu armazenamento, e seu transporte necessita de banda relativamente larga. Para reduzir o espaço e a banda, e tornar viável a utilização de vídeos digitais, foi desenvolvido um padrão de compressão para vídeo digital denominado *Motion Picture Experts Group (MPEG)* [4]. O padrão *MPEG* reduz drasticamente a quantidade de dados de um vídeo através da eliminação de redundâncias. As redundâncias eliminadas são de 3 tipos: redundância espacial, temporal e de áudio.

A *redundância espacial* corresponde aos dados de elementos de imagem que preenchem áreas de uma imagem com informações repetidas. A *redundância temporal* corresponde às informações repetidas em quadros consecutivos ao longo da reprodução de um vídeo. A eliminação da redundância temporal produz três tipos de quadros: quadro *Intracoded (I)*, possui informações do quadro original eliminando somente a redundância espacial; quadro *Predictive (P)*, possui informações que diferenciam o quadro atual com relação a um quadro *I* ou *P* anterior; e quadro *Bidirectional (B)*, possui informações referente a diferenças do quadro atual com os quadros anterior e posterior *I* ou *P*, ao mesmo tempo. Desta forma, um arquivo de vídeo é formado por seqüências de quadros *I*, *P* e *B*.

As redundâncias espacial e temporal são variáveis a cada quadro fazendo com que o fluxo de dados resultante do processo de compressão possua largura de banda variável no tempo. Quando o vídeo possui largura de banda variável no tempo ele é denominado *Variable Bit Rate (VBR)*. No caso inverso, quando a compressão força a criação de quadros de mesmo tamanho, excluindo informações de quadros com muito detalhes e incluindo zeros para preencher quadros com pouca informação, o vídeo resultante possui largura de banda constante no tempo e é denominado *Constant Bit Rate (CBR)* [15]. O envio de fluxos de vídeo de arquivos *CBR* facilita a administração dos recursos por parte do servidor, pois a utilização de cada recurso é previamente conhecida e constante, mas estes arquivos são pouco utilizados devido a sua baixa qualidade quando comparados com arquivos *VBR* [12].

Arquivos *VBR* possuem o inconveniente de produzir rajadas de dados nos momentos em que são transmitidos pacotes com imagens ou som com baixo grau de compressão. Estas rajadas dificultam o controle sobre a banda de rede, pois o valor estimado para utilização da banda de rede de um arquivo de vídeo é a maior quantia de dados enviada, por intervalo, durante todo o período de reprodução do vídeo, sendo este valor muito superior ao valor médio transmitido. O valor exagerado provoca a sub-utilização da banda de rede, e força o servidor a aceitar um número pequeno de conexões.

O gráfico da Figura 1 mostra a variabilidade dos tamanhos dos quadros de um filme comprimido com *MPEG*. Observa-se que este vídeo possui quadros com cerca de 100.000 bytes, mas a média do tamanho dos quadros é em torno de 8.000 bytes. Devido à escala horizontal do gráfico, não se percebe claramente a seqüência de quadros, que é geralmente consiste de um quadro *I* grande, seguido de alguns quadros *B* e *P* com tamanho médio inferior a 5000 bytes.

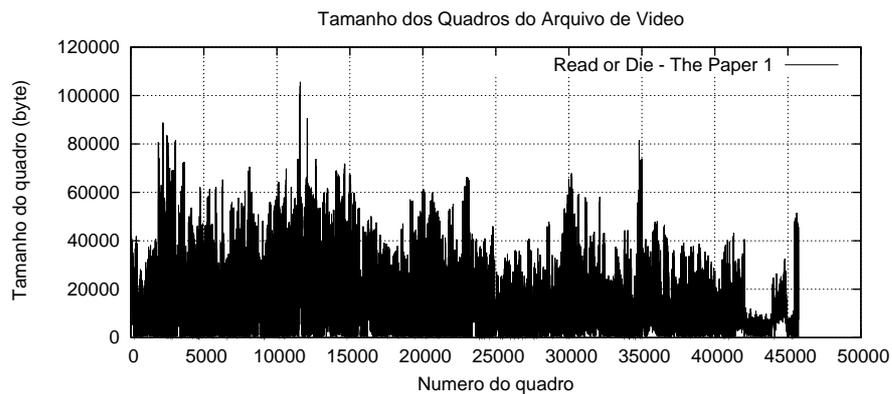


Figura 1: Tamanho dos quadros do desenho *Read or Die – The Paper 1*.

Arquitetura de um Servidor de Vídeo

Um *Servidor de vídeo* é um programa que disponibiliza fluxos de vídeo a clientes que os reproduzem. A Figura 2 mostra a arquitetura de um servidor de vídeo. A figura mostra os dois tipos de serviços que podem ser disponibilizados num servidor, a transmissão de vídeos ao vivo, e o acesso a arquivos de vídeo armazenados em disco.

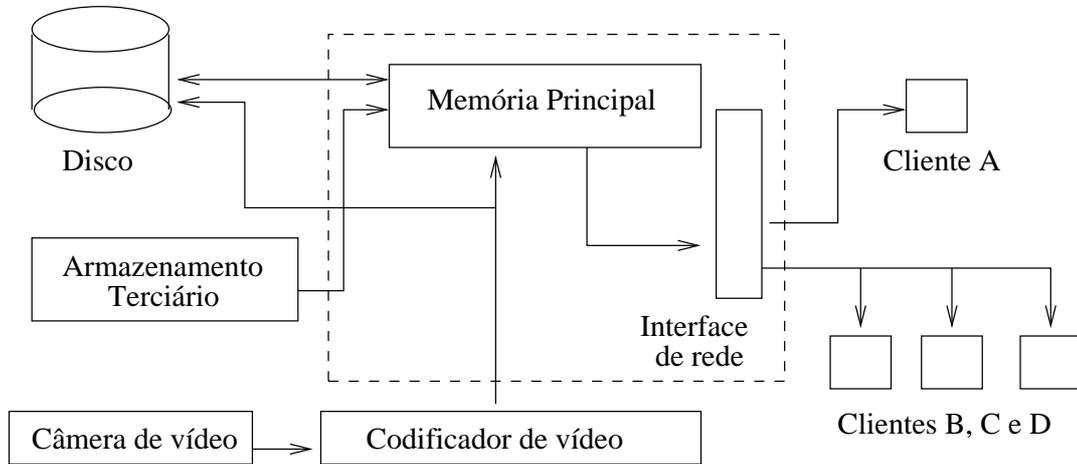


Figura 2: Arquitetura de um servidor de vídeo.

Na transmissão ao vivo, a imagem e o som são capturados e codificados como fluxos de vídeo, e então enviados ao servidor que repassa-os aos clientes. Na transmissão de arquivos de vídeo, estes podem estar armazenados diretamente em disco, ou em armazenamento terciário como fitas ou *compact discs*. Neste caso o servidor copia o arquivo do disco (ou armazenamento terciário) para a memória principal antes de transmiti-lo ao cliente através da interface de rede.

Para que um cliente conectado a um servidor de vídeo possa reproduzir o vídeo sem perda de qualidade, o servidor deve garantir que os fluxos de áudio e de vídeo sejam enviados sem atraso. O servidor deve limitar o número de requisições aceitas para controlar a utilização da rede. Quando o servidor está atendendo a muitas requisições, a aceitação de novas requisições poderia causar sobrecarga em algum recurso, prejudicando assim o envio dos fluxos de vídeo das requisições correntes. Um *algoritmo de controle de admissão* é empregado para decidir se novas requisições podem ser atendidas. Se a aceitação da nova requisição puder provocar sobrecarga na interface de rede, a requisição é negada.

O servidor, ao ser inicializado, analisa cada arquivo de vídeo disponível em uma pasta específica de arquivos de vídeo, e obtém informações que serão utilizados nos algoritmos que controlam recursos do servidor. Após análise dos arquivos, o servidor fica aguardando requisições. Ao receber uma requisição, o servidor utiliza o algoritmo de controle de admissão para verificar se existem recursos suficientes para acrescentar o fluxo de vídeo requisitado aos fluxos das requisições previamente aceitas. Este algoritmo analisa a banda necessária para o arquivo requisitado e a compara com a banda ainda disponível. Se a requisição não pode ser aceita, o servidor envia uma mensagem ao cliente negando o pedido. Se a requisição pode ser aceita, o servidor informa ao cliente da aceitação, e reserva a banda de rede e outros recursos necessários para satisfazer a esta nova requisição.

Um servidor de vídeo envia uma determinada quantidade de dados a seus clientes em intervalos de tempo fixo. A quantidade de dados enviados em cada período deve ser suficiente para que os quadros sejam reproduzidos sem que ocorra falta de dados e a conseqüente interrupção na reprodução do vídeo no cliente. Este período entre a transmissão de dados é denominado *ciclo do servidor*. Os ciclos do servidor determinam a quantidade de dados que é enviada para todos os clientes conectados ao servidor. Por exemplo, o ciclo de um servidor pode ser definido em 1 segundo, e neste período devem ser transmitidos os fluxos de vídeo correspondentes aos quadros que serão reproduzidos em todos os clientes naquele período.

O gráfico da Figura 3 mostra a quantidade de dados que deve ser enviada pelo servidor na transmissão do arquivo de vídeo da Figura 1, com ciclo do servidor de 1 segundo. Observa-se neste gráfico a grande variabilidade na taxa de dados que devem ser enviados em cada ciclo, que pode variar de 2.000 a 940.000 bytes. Como o consumo dos dados pelo cliente depende do tamanho dos quadros que estão sendo reproduzidos, o servidor deve garantir que os fluxos de vídeo serão entregues a tempo, para que não ocorram falhas na reprodução.

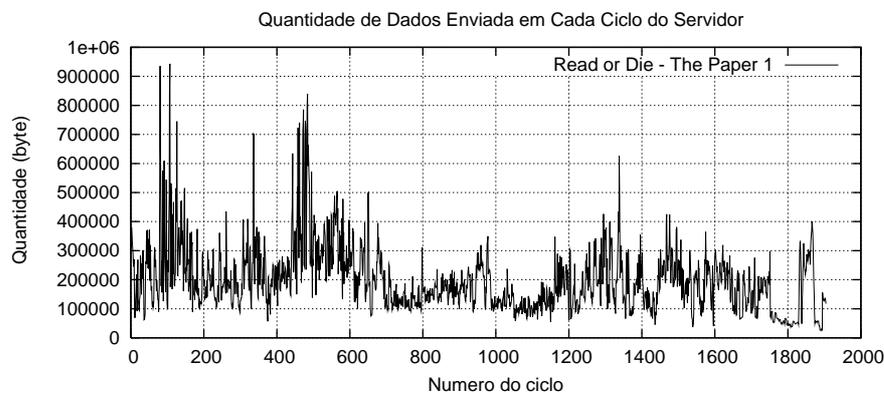


Figura 3: Quantidade de dados enviada a cada ciclo do servidor.

2 Controle de Banda de Rede

O controle de banda de rede em um servidor de vídeo pode ser efetuado por métodos determinísticos ou por multiplexação estatística. A multiplexação estatística emprega algoritmos estatísticos na determinação da banda utilizada. Este método não assegura o controle da banda de rede pois permite que o servidor aceite um número elevado de conexões, diminuindo a qualidade do serviço oferecido porque podem ocorrer atrasos na entrega dos fluxos de vídeo.

A implementação de algoritmos determinísticos considera o pior caso na estimativa da banda de rede utilizada, assegurando que não ocorrerão atrasos na entrega dos fluxos de vídeo. Este método pode sub-utilizar a banda de rede, pois os valores considerados no cômputo da utilização deste recurso são geralmente muito superiores ao valor médio utilizado.

O controle da banda de rede é realizado de forma diferenciada para cada tipo de codificação de vídeo. Em [10, 12] são propostos algoritmos determinísticos que controlam a banda da rede no servidor utilizando arquivos de vídeo no formato *CBR*. Estes trabalhos mostram a facilidade em administrar e calcular os recursos quando um servidor de vídeo trabalha com arquivos neste formato, mas aqueles são raramente usados devido à sua baixa qualidade.

Para arquivos com codificação *VBR* foram propostos vários métodos de estimativa de banda de rede, como os descritos em [11, 6, 1, 2, 3, 9, 8]. Nestes trabalhos são propostos algoritmos que utilizam cálculos determinísticos e multiplexação estatística para estimar a utilização de banda da rede. Em cada um destes trabalhos, o algoritmo implementado depende do programa cliente que reproduz o vídeo, limitando portanto seu escopo de aplicação.

O trabalho de Knightly e Zhang [7] propõe a utilização de um método determinístico para o controle da banda de rede. No que segue, este método é denominado *método K-Z*. O método K-Z considera a diferença entre uma estimativa de banda baseada no tamanho do maior quadro de um arquivo de vídeo, e outra estimativa baseada na maior soma dos tamanhos de um conjunto de quadros consecutivos. A segunda forma é razoável por causa da seqüência de quadros típica num arquivo de vídeo, na qual um quadro grande de tipo *I* é seguido de alguns quadros pequenos de tipo *P* e *B*.

Considere, como exemplo um arquivo de vídeo, com 25 quadros por segundo, e no qual o maior quadro-I tem 50.000 bytes, os quadros que o seguem são três quadros-B de 1.100, 1.200 e 1.000 bytes, e um quadro-P de 4.500 bytes. O tamanho destes quadros, somado ao tamanho do quadro-I é a maior soma de uma seqüência de cinco quadros deste arquivo de vídeo, perfazendo 57.800 bytes. Para este arquivo, a estimativa de banda de rede baseada somente no tamanho do maior quadro é $50.000 \times 25 = 1.250.000$ bytes/s. Por outro lado, considerando-se a maior seqüência de cinco quadros consecutivos, a estimativa é $57.800 \times 5 = 289.000$ bytes/s. Se a banda for reservada com base no tamanho dos quadros-I, haverá sub-utilização porque estes correspondem a apenas 10 a 20% dos quadros em arquivos de vídeo. Desta forma, a reserva baseada na soma de alguns quadros obtém uma melhor utilização da banda de rede.

É possível portanto estimar a banda de rede de arquivos de vídeo através da soma dos tamanhos dos quadros que são reproduzidos em um segundo. Neste caso o algoritmo de controle considera a maior quantidade de dados enviados durante 1 segundo para estimar a largura de banda utilizada em cada arquivo de vídeo disponibilizado. Para o arquivo de vídeo da Figura 3 a banda estimada corresponde ao maior pico do gráfico, que é 943.000 bytes, embora a média seja bem menor que este valor. Com base na estimativa de largura de banda de cada arquivo, o algoritmo de controle de banda de rede adiciona os valores estimados de todos os arquivos que estão sendo transmitidos e determina se a interface de rede poderá sofrer sobrecarga em algum momento.

Testes Realizados com o Método K-Z

Durante o desenvolvimento do *SVFserver* foram realizados testes para avaliar o método K-Z de controle de banda de rede [13]. O algoritmo de controle de banda de rede mantém uma estimativa da banda utilizada pelos vídeos que estão sendo transmitidos. Quando é recebida uma nova requisição, o algoritmo de controle de admissão do servidor decide, com base nos resultados do algoritmo de controle de banda de rede, se uma nova requisição pode ser atendida. Se o novo valor total não ultrapassar a 80% da largura de banda disponível, a requisição é aceita.

Os testes realizados no servidor foram executados em uma rede *Ethernet* 100Mbps, com a servidora e três clientes. As quatro máquinas empregadas nos testes são modelo Presario da Compac, e contém processador Pentium III de 933 MHz, com memória RAM de 384 Mbytes, disco modelo Maxtor 36147H8 de 60 Gbytes com 512 kbytes de cache SDRAM, 5.400 rpm e tempo de busca abaixo de 9,5 ms, e uma interface de rede do modelo Accton SMC2-1211TX.

Foram realizados dois experimentos para comprovar a eficácia do método K-Z. O primeiro consiste de uma série de requisições ao servidor para o mesmo arquivo de vídeo, (Figura 3), e o segundo experimento consiste de uma série de requisições a quatro arquivos distintos. A Tabela 1 mostra as características dos filmes utilizados nos testes. A tabela mostra o valor da banda de rede estimada para cada arquivo de vídeo, e este valor é a máxima banda utilizada durante todo o período de transmissão do vídeo.

Nome	Tipo	Qps	Resolução	Duração[s]	Banda
<i>Read or Die - The Paper 1</i>	desenho	24	640x480	1905	943.320
<i>Read or Die - The Paper 2</i>	desenho	24	640x480	1850	722.550
O Senhor dos Anéis 1 †	filme	24	576x240	2163	494.593
<i>From the Hell</i> ‡	filme	24	595x240	1755	491.600

† Somente os 36 minutos iniciais; ‡ somente os 30 minutos iniciais.

Tabela 1: Características dos filmes utilizados nos testes.

As requisições foram geradas por um programa cliente codificado especificamente para estes testes, que efetua requisições aos arquivos de vídeo e verifica a quantidade de dados recebidos [13]. Este cliente possui a opção de armazenar os dados recebidos em um arquivo no disco, permitindo verificar se o arquivo recebido é idêntico ao arquivo transmitido pelo servidor.

No experimento com somente um filme, uma nova requisição é realizada a cada segundo, forçando assim a coincidência de intervalos em que são transmitidos os ciclos com maior quantidade de dados. No experimento com quatro filmes, as requisições são realizadas em intervalos de 1 até 5 segundos. Cada um dos testes foi repetido 5 vezes.

Após o início dos experimentos, é observado se em algum momento o tráfego na interface de rede ultrapassa 80% da banda disponível. Como os testes foram realizados em uma rede do tipo *Ethernet* de 100Mbps, o valor de banda disponível às requisições corresponde a 80Mbps, ou cerca de 10MB/s. A medida de tráfego da rede é realizada com o programa *Xnetload* [14]. Este programa informa, ao final do teste, a taxa máxima transmitida através da interface de rede, e permitem verificar se o método utilizado para estimar a largura de banda é válido.

A Tabela 2 mostra os resultados dos experimentos. A primeira e segunda colunas mostram o número de requisições aceitas pelo servidor e a banda máxima medida na interface de rede para o experimento com um filme. A quinta e sexta colunas da tabela mostram o número de requisições aceitas pelo servidor e a banda máxima medida na interface de rede para o segundo experimento. No segundo experimento, a diferença no número requisições aceitas em cada teste é devida a aleatoriedade nas requisições.

Observa-se nos resultados mostrados na Tabela 2 que, apesar de o servidor aceitar um número de conexões que supostamente causaria saturação na rede, pois 80% da banda disponível de rede estaria sendo utilizada, o tráfego máximo medido é de 6.6 Mbytes/s (médio de 6.4 Mbytes/s) para 10 requisições no primeiro experimento, enquanto que no segundo o tráfego máximo é de 5.2 Mbytes/s para 16 requisições e a média de 4.7 Mbytes/s para 14,8 requisições. Estes valores indicam que a largura de banda é sub-utilizada por conta do exagero na estimativa de banda dos arquivos, ficando abaixo de 70% da capacidade.

No segundo experimento a banda utilizada é menor que no primeiro. Isto acontece porque os ciclos que enviam as maiores quantidades de dados em cada um dos filmes são mais bem distribuídos no tempo, o que não ocorre com requisições feitas a somente um arquivo. Assim, quando são enviados fluxos de diversos arquivos, geralmente os picos dos arquivos não coincidem, como ocorre nas requisições realizadas as mesmo arquivo de vídeo.

Teste	um único filme		quatro filmes distintos	
	requis.	banda [kB/s]	requis.	banda [kB/s]
1	10	6.305	16	4.229
2	10	6.281	15	4.410
3	10	6.305	13	4.434
4	10	6.319	16	5.147
5	10	6.594	14	5.182

Tabela 2: Quantidade de dados enviada pelo servidor com um e com quatro filmes.

3 Estimativa de Banda com Filtros Passa-baixa

Os testes aqui descritos, com o algoritmo baseado no trabalho de Knightly e Zhang, comprovam que este impede a ocorrência de sobrecarga na rede, mas pode induzir a sua sub-utilização. Para aumentar a utilização da banda de rede pode ser empregado um método que utiliza um ‘filtro’ para determinar o valor da banda de acordo com a quantidade de dados que deve ser transmitida durante um intervalo de vários ciclos contíguos.

A filtragem é necessária porque em alguns arquivos de vídeo a banda estimada como no método K-Z é determinada por picos isolados. A Figura 4 mostra a quantidade de dados enviada em cada ciclo do servidor para o filme *From the Hell*, durante os 30 minutos iniciais. Há um pico isolado de 491 kbytes próximo ao ciclo 900, e portanto 491 kbytes é o valor considerado na estimativa da banda, embora esta quantidade seja muito superior ao valor médio é transmitido nos outros ciclos, ocasionando assim sub-utilização de banda pelo servidor.

O *filtro de estimativa de banda* é aplicado para calcular a média ponderada da quantidade de dados a transmitir num ciclo com relação aos ciclos vizinhos. Desta forma, o valor estimado é reduzido, evitando que picos isolados distorçam a estimativa, mas possibilitando que os intervalos de muitos ciclos nos quais são transmitidas grandes quantidades de dados determinem a estimativa de banda utilizada. Por exemplo, na região entre os ciclos 670 e 710 da Figura 4 devem ser transmitidos vários ciclos de 200 a 400 kbytes. Para este arquivo a estimativa de largura de banda pelo método K-Z é de 491 kbytes/s, enquanto que com o filtro passa-baixa da Equação 1, a banda estimada é de 320 kbytes/s. Este valor é mais próximo da média e permite ao servidor aceitar mais requisições e utilizar melhor a banda de rede.

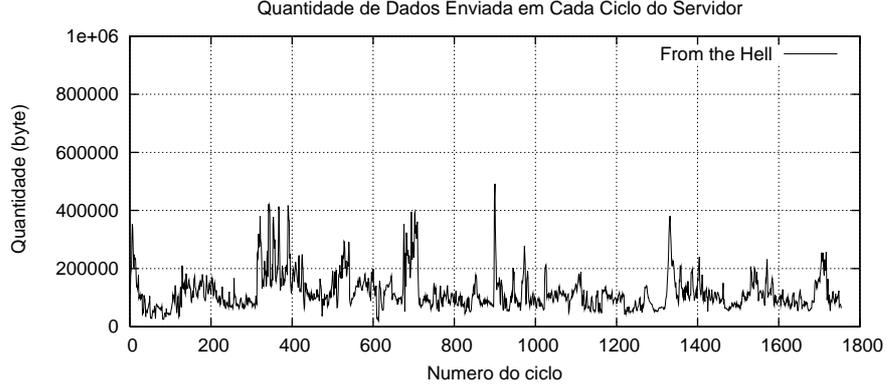


Figura 4: Quantidade de dados enviada a cada ciclo do servidor, filme *From the Hell*.

O filtro de estimativa de banda atua como um filtro passa-baixa ao reduzir a influência dos picos isolados numa seqüência de quadros, de maneira análoga a eliminação do ruído de um sinal. Os filtros \mathcal{B}_1 e \mathcal{B}_2 , definidos pelas Equações 1 e 2 respectivamente, são utilizados para estimar os valores de largura de banda dos arquivos de vídeo. Até onde sabemos, a utilização de métodos similares aos filtros propostos não é mencionada na bibliografia sobre algoritmos determinísticos de controle de banda de rede.

$$\mathcal{B}_1 = \sum_{-6 \leq n \leq -4} 0.05 \cdot ciclo_n + \sum_{-3 \leq n \leq +3} 0.1 \cdot ciclo_n + \sum_{+4 \leq n \leq +6} 0.05 \cdot ciclo_n \quad (1)$$

$$\mathcal{B}_2 = 0.025 \cdot ciclo_{n-10} + \sum_{-9 \leq n \leq +9} 0.05 \cdot ciclo_n + 0.025 \cdot ciclo_{n+10} \quad (2)$$

O filtro \mathcal{B}_1 calcula o valor da banda considerando também as quantias enviadas nos 3 ciclos anteriores e posteriores. O filtro \mathcal{B}_2 estima a largura de banda considerando as quantidades enviadas nos 10 ciclos adjacentes, anteriores e posteriores. Outros parâmetros de filtragem foram testados, com variações na largura de intervalo e nos pesos dos coeficientes, e verificou-se que ao alargar o intervalo, a banda estimada diminui, tendendo ao valor da média da quantidade de dados transmitida ao longo de todo o filme.

A Tabela 3 mostra, para os quatro filmes, a estimativa de banda pelo método K-Z, e os valores calculados com os filtros \mathcal{B}_1 e \mathcal{B}_2 . Na tabela também são mostrados o percentual de redução com relação à estimativa da banda pelo método K-Z. Observa-se que a diferença entre os valores calculados por \mathcal{B}_1 e \mathcal{B}_2 é da ordem de 6%, enquanto que a diferença entre o resultado obtido com o filtro \mathcal{B}_1 e o valor original é cerca de 35%.

Arquivo	Banda		Redução		Redução	
	K-Z	\mathcal{B}_1	(%)	\mathcal{B}_2	(%)	
<i>Read or Die - The Paper 1</i>	943,3	606,7	35,7	570,0	39,6	
<i>Read os Die - The Paper 2</i>	722,5	483,6	33,1	449,7	37,8	
O Senhor dos Anéis 1 (36 m)	494,6	372,8	24,6	344,8	30,3	
<i>From the Hell</i> (30 m)	491,6	322,8	34,3	291,8	40,6	

Tabela 3: Banda de rede ([kbytes/s]) estimada com os filtros \mathcal{B}_1 e \mathcal{B}_2 .

4 Controle de Admissão com Filtros Passa-baixa

Os testes descritos na seção anterior foram repetidos com as novas estimativas de largura de banda, calculados pelos filtros \mathcal{B}_1 e \mathcal{B}_2 .

Os resultados dos cinco testes realizados com o desenho *Read or Die - The Paper 1* com os dois filtros encontram-se na Tabela 4. Nestes testes o algoritmo de controle de admissão aceitou 16 requisições e utilizou uma média de 9.698 kbytes/s para o filtro \mathcal{B}_1 , e aceitou 17 requisições e utilizou uma média de 10.078 kbytes/s para o filtro \mathcal{B}_2 , obtendo um ganho de 60% e 70% no número de requisições aceitas pelo servidor, e um melhora na utilização da banda de rede da ordem de 50% e 58%, com relação aos testes que utilizaram o método K-Z para a estimativa da banda.

Teste	\mathcal{B}_1		\mathcal{B}_2	
	req.	Banda	req.	Banda
1	16	9.927	17	9.997
2	16	10.023	17	10.561
3	16	8.840	17	10.053
4	16	10.089	17	9.578
5	16	9.109	17	10.299

Read or Die - The Paper 1

Tabela 4: Banda utilizada ([kbytes/s]) com estimativas pelos os filtros \mathcal{B}_1 e \mathcal{B}_2 .

Observa-se na Tabela 4 que em alguns testes a banda medida na interface de rede pelo programa *Xnetload* ultrapassou o máximo alocado pelo algoritmo de controle de banda de rede, que é 10 Mbytes/s, mas em nenhum dos testes ocorreram atrasos na entrega dos fluxos de vídeo aos clientes.

No segundo conjunto de testes, os valores utilizados no cálculo da estimativa da banda são os da Tabela 3, e o resultado destes testes é mostrado na Tabela 5. Nos testes realizados com o filtro \mathcal{B}_1 , o algoritmo de controle de admissão aceitou até 26 requisições, com média de 22,4 requisições, atingindo utilização média de banda de 6.541 kbytes/s. Com o filtro \mathcal{B}_2 , o servidor aceitou até 26 requisições com média de 24,2 requisições, atingiu utilização média da banda de 7.324 kbytes/s. O número médio de requisições aceitas aumentou em 51% e 63%, e

a utilização da rede obteve um ganho de 39% e 56% para os filtros \mathcal{B}_1 e \mathcal{B}_2 respectivamente, com relação aos testes com o método K-Z.

Teste	\mathcal{B}_1		\mathcal{B}_2	
	req.	Banda	req.	Banda
1	22	5.969	24	6.598
2	24	6.600	26	7.043
3	19	6.851	22	8.358
4	26	6.630	23	7.759
5	21	6.653	26	6.864

Requisições aos quatro filmes

Tabela 5: Banda utilizada ([kbytes/s]) com estimativas pelos filtros \mathcal{B}_1 e \mathcal{B}_2 .

Observa-se nestes testes que a banda máxima chegou a 6.851 kbytes/s com o primeiro filtro, e 8.358 kbytes/s com o segundo. O maior valor de banda medido corresponde ao teste com menor número de requisições aceitas, porque neste teste ocorreu um número desproporcional de requisições ao arquivo de vídeo com maior largura de banda. Para o filtro \mathcal{B}_1 , a utilização da interface de rede não atingiu 70% da banda reservada, enquanto que para o filtro \mathcal{B}_2 , a utilização ficou em torno de 83% dos 10 Mbytes disponibilizados.

Com base nos resultados obtidos nos testes, deve-se levar em consideração duas situações para a aplicação de filtros. Na primeira, podem ocorrer várias requisições concomitantes a um arquivo de vídeo, ocasionando a sobreposição dos ciclos que enviam os maiores quadros. Neste caso é preferível superestimar a largura de banda para garantir que o algoritmo de controle de banda não cause sobrecarga na rede. A segunda situação é o caso mais geral em que ocorrem requisições a diversos arquivos bem distribuídos no tempo. A estimativa de largura de banda pode ser mais otimista, possibilitando ao servidor atender um número maior de requisições e assim melhor aproveitar os recursos na rede.

5 Conclusões

Um servidor de vídeo sob demanda necessita controlar a fração da capacidade da rede que está sendo utilizada pelos fluxos de vídeo correntes, para decidir se novas requisições podem ser aceitas sem prejuízo a qualidade de reprodução nos clientes. O controle da capacidade utilizada é efetuado pelo algoritmo de controle de banda, enquanto que o algoritmo de controle de admissão decide se uma nova requisição pode causar sobrecarga na rede.

Este artigo descreve um método determinístico para estimar a largura de banda de arquivos de vídeo. Os métodos descritos na literatura tendem a superestimar a banda dos vídeos, impedindo a plena utilização da capacidade da rede e limitando artificialmente o número de requisições que poderiam ser atendidas pelo servidor. Em particular, o método determinístico de estimativa de banda proposto por Knightly e Zhang pode subutilizar a banda de rede. Este método é eficaz quando é utilizado conjuntamente com um filtro passa-baixa que melhora substancialmente a estimativa da banda de rede dos arquivos de vídeo.

Considerando os resultados dos testes aqui relatados, conclui-se que a utilização dos filtros \mathcal{B}_1 e \mathcal{B}_2 para a estimativa de largura de banda de vídeos permite aumentar a utilização da rede e obter um melhor aproveitamento do servidor sem interferir negativamente no desempenho dos clientes. Os testes realizados indicam ganhos na utilização da rede da ordem de 58% e acréscimos no número de requisições aceitas pelo servidor da ordem de 60% para o filtro \mathcal{B}_1 , e ganhos na utilização da rede da ordem de 58% e no número de requisições aceitas pelo servidor da ordem de 70% para o filtro \mathcal{B}_2 . O filtro \mathcal{B}_1 produz estimativas mais conservadoras que o filtro \mathcal{B}_2 e é portanto indicado quando um controle mais estrito deve ser exercido sobre a utilização da rede.

Referências

- [1] G. Neufeld D. Makaroff and N. Hutchi. Network bandwidth allocation and admission control for a continuous media file server. *Interactive Distributed Multimedia Systems and Telecommunication Services*, pages 337 – 350, 1999.
- [2] W. Feng and S. Sechrest. Critical bandwidth allocation for delivery of compressed video. *Computer Communications*, 18:709 – 717, October 1995.
- [3] W. Feng and S. Sechrest. Smoothing and buffering for delivery of prerecorded compressed video. *Proceedings of the IS&T/SPIE Symposium on Multimedia Computing and Networking*, pages 234 – 244, February 1995.
- [4] D. Gall. MPEG: A video compression standart for multimedia applications. *Communications of ACM*, 34(4):47–58, April 1991.
- [5] B. Grob. *Televisão e Sistema de Vídeo*. Editora Guanabara S.A., 1989.
- [6] E. Knightly and H. Zhang. Traffic characterization and switch utilization using a deterministic bounding interval dependent traffic model. *Proc. of IEEE INFOCOM*, pages 1137 – 1145, 1995.
- [7] E. Knightly and H. Zhang. D-BIND: An accurate traffic model for providing QoS guarantees to VBR traffic. *IEEE/ACM Transactions on Networking*, 5(?):219 – 231, April 1997.
- [8] S. Keshav M. Grossglauser and D. Tse. RCBR: A simple and efficient service for multiple time-scale traffic. *IEEE/ACM Transactions on Networking*, 5(6):741 – 755, 1997.
- [9] J. McManus and K. Ross. Prerecorded VBR sources in ATM networks: Piecewise constant-rate transmission and transport. Technical report, University of Pennsylvania, September 1995.
- [10] J. McManus and K. Ross. Video on demand over ATM: Constante rate transmission and transport. *Proceedings of IEEE INFOCOM*, pages 1357 – 1362, March 1996.
- [11] K. K. Ramakrishnan N. G. Duffield and A. R. Reibman. SAVE: An algorithm for smoothed adaptive video over explicit rate networks. *IEEE/ACM Transactions on Networking*, 6(6):717 – 728, 1998.
- [12] B. Khasnabish S. Gringeri, A. Lewis and B. Basch. Traffic shaping, bandwidth allocation, and quality assessment for MPEG video distribution over broadband networks. *IEEE Networking*, 12(6):94 – 107, November/December 1998.
- [13] Patrícia Sereda. Servidor de vídeo SVFserver. Dissertação de Mestrado, Depto. de Informática, Universidade Federal do Paraná, Agosto 2003.
- [14] Roland Smith. Roland’s Software Page, 2003. <http://www.xs4all.nl/~rsmith/software/>.
- [15] A. S. Tanenbaum. *Redes de Computadores*. Editora Campus, 1997.
- [16] M. Vernick. The design, implementation and evolution of the Stony Brook video server. Technical report, State University of New York, NY, December 1996.