

Estratégias Híbridas para Transmissão de Mídia Contínua Interativa com Compartilhamento de Fluxo

Marcus Rocha, Marcelo Maia, Ítalo Cunha, Jussara Almeida, Sérgio Campos

¹Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
Av. Antônio Carlos, 6627, Pampulha – Belo Horizonte, MG, Brasil, 31270-010

{mvrocha, mmaia, cunha, jussara, scampos}@dcc.ufmg.br

Abstract. *The impact of interactivity in media workloads has been previously studied. Buffer space, local to client media player, allows to avoid data retransmission and to reduce required server bandwidth. This work proposes a new set of strategies which use the period of time when the client is inactive, and thus is not consuming server bandwidth, to listen to and to store streams being delivered to other clients. This work also explores hybrid strategies, combining the new proposed optimizations with existing ones, in order to further reduce server bandwidth requirements over a wider range of request rates than any single optimization. Our optimizations reduce average server bandwidth in up to 66% over the original Bandwidth Skimming protocol.*

Resumo. *O impacto da interatividade na escalabilidade dos protocolos de transmissão de mídia contínua com fluxos compartilhados é alvo de estudos recentes. Áreas de armazenamento permitem evitar retransmissões e reduzir a banda média de servidor requerida. Este trabalho propõe estratégias que usam períodos inativos do cliente (períodos sem consumo de banda) para escutar e armazenar fluxos transmitidos a outros clientes. Também explora estratégias híbridas que combinam novas otimizações com estratégias anteriores e reduzem a banda média de servidor em uma faixa mais ampla de taxas de chegada que qualquer política individual. As otimizações propostas reduzem a banda média de servidor em até 66% sobre o protocolo Bandwidth Skimming original.*

1. Introdução

O uso de mídia contínua tem despertado interesse para uso em aplicações em áreas como educação, treinamento e entretenimento. A elevada demanda por banda de servidor e cliente, imposta por estas aplicações, ensejou o desenvolvimento de diversos protocolos para a transmissão escalável de mídia contínua [Bar-Noy and Ladner 2003, Bradshaw et al. 2003, Coffman et al. 2002, Eager and Vernon 1998, Eager et al. 2000, Eager et al. 2001, Gao and Towsley 1999, Hua et al. 1998, Hua and Sheu 1997, Ma et al. 2005].

Dois protocolos atraem especial interesse pois, além de propiciarem grande redução na largura de banda requerida por servidores, provêm serviço imediato ao cliente. Estes protocolos, *Patching* [Hua et al. 1998, Gao and Towsley 1999] e *Bandwidth Skimming* [Eager et al. 2000, Eager et al. 2001], já foram avaliados detalhadamente, especialmente para o *acesso seqüencial*, no qual toda a mídia é exibida sem interrupções.

Diversos estudos mostram que a *interatividade* – acesso a segmentos da mídia – está presente em diversas cargas reais [Almeida et al. 2001, Chesire et al. 2001, Costa et al. 2004, Padhye and Kurose 1998], e que a mesma tem impacto significativo sobre a escalabilidade de protocolos de compartilhamento de banda para mídia contínua [Jin and Bestravros 2002, Rocha et al. 2005a, Tan et al. 2002a]. Ainda que *Patching* e *Bandwidth Skimming* tenham a escalabilidade fortemente afetada pela interatividade, o presente trabalho tem seu foco no *Bandwidth Skimming*, que apresenta escalabilidade superior à do *Patching* tanto para acesso seqüencial [Eager et al. 2001, Gao and Towsley 1999], quanto para acesso com interatividade [Rocha et al. 2005b].

Otimizações e protocolos orientados para cargas interativas foram explorados em [Hua and Tran 2005, Liao and Li 1997, Ma et al. 2005, Rocha et al. 2005a, Rocha et al. 2005b, Rodrigues and Leão 2005]. Nestes trabalhos, as otimizações e estratégias mais bem sucedidas utilizam o armazenamento de conteúdo em área local ao programa cliente de exibição de vídeo, ou em área compartilhada, com a finalidade de evitar retransmissão de dados. Em geral este conteúdo é armazenado nos períodos de atividade dos clientes, durante os quais os mesmos estão requisitando mídia ao servidor e a estão exibindo para os usuários (e, portanto, consumindo banda do servidor).

Este trabalho explora um aspecto pouco tratado em trabalhos anteriores, os períodos de inatividade do cliente de exibição, durante os quais o cliente não requisita mídia para exibir para o usuário (e não está consumindo banda do servidor), ou durante os quais toda a mídia a exibir se encontra na área de armazenamento do cliente.

A seguir, apresenta estratégias híbridas, combinando otimizações de períodos ativo e inativo, e obtendo economia de banda em faixas mais amplas de taxas de chegada que as otimizações individuais, com economia de até 66% sobre o *Bandwidth Skimming*.

Estratégias de otimização baseadas no tempo inativo levam o cliente a escutar até dois fluxos nos períodos de tempo nos quais o cliente não estaria consumindo banda. Como resultado, tem-se impacto na largura de banda do cliente e na largura de banda de rede. Este trabalho faz uma avaliação preliminar da banda de rede considerando uma topologia de rede real, e aponta os principais compromissos envolvidos.

O foco do trabalho é a redução da banda média requerida para prover serviço imediato a usuários interativos. Em primeiro lugar, esta abordagem é adotada em trabalhos anteriores com os quais os presentes resultados podem ser comparados [Jin and Bestravros 2002, Tan et al. 2002a]. Em segundo lugar, [Tan et al. 2002b] mostrou que, para cargas seqüenciais, devido a flutuações estatísticas nas taxas de chegada para múltiplos arquivos e com chegadas *Poisson* [Almeida et al. 2001, Costa et al. 2004], a largura de banda requerida para que um servidor seja capaz de atender a uma certa carga é próxima à banda média requerida para atender a esta carga.

As principais contribuições deste trabalho incluem:

- Otimizações de período inativo com economia de banda de até 55% sobre o protocolo *Bandwidth Skimming* original;
- Estratégia híbrida que obtém economia de banda em faixas mais amplas de taxas de chegada do que as otimizações individuais;
- Estratégia híbrida que obtém economia de banda de até 66% sobre o protocolo *Bandwidth Skimming* original.

O restante deste trabalho está assim organizado. A seção 2 discute aspectos de cargas interativas com impacto sobre a escalabilidade de protocolos de compartilhamento de banda, apresenta as cargas empregadas e discute a síntese de cargas realistas. A seção 3 discute protocolos de compartilhamento de banda e o impacto da interatividade sobre a escalabilidade destes, e apresenta otimizações correntes para interatividade. A seção 4 apresenta e avalia novas otimizações para período inativo. A seção 5 apresenta e avalia estratégias híbridas de otimização. A seção 6 discute o impacto das novas estratégias na banda de rede. A seção 7 conclui e indica trabalhos futuros.

2. Cargas Interativas

Esta seção discute os aspectos da interatividade com impacto sobre a escalabilidade dos protocolos de compartilhamento de banda (seção 2.1), apresenta nossas cargas reais (seção 2.2) e o processo usado para geração de cargas sintéticas realistas (seção 2.3).

2.1. Interatividade e Cargas Interativas

Diferentemente do acesso seqüencial, onde toda a mídia é exibida do início ao fim sem qualquer interrupção, numa carga com *interatividade* um usuário inicia uma sessão de acessos a uma certa mídia e, ao longo desta sessão, emite uma ou mais requisições interativas para receber diferentes segmentos da mídia. Diversos trabalhos evidenciam a presença da interatividade em cargas reais [Almeida et al. 2001, Chesire et al. 2001, Costa et al. 2004, Padhye and Kurose 1998] e seu impacto sobre a escalabilidade de protocolos de compartilhamento de banda [Almeida et al. 2001, Chesire et al. 2001, Jin and Bestavros 2002, Tan et al. 2002a, Rocha et al. 2005a, Rocha et al. 2005b].

Para protocolos de compartilhamento de banda com cargas seqüenciais e chegadas *Poisson*, a banda média de servidor requerida para sustentar uma determinada carga depende apenas da taxa de chegada normalizada N , onde N indica o número de chegadas num período de tempo de duração igual à duração da mídia ($N = \lambda T$, sendo λ a taxa de chegadas e T a duração da mídia) [Eager et al. 2001, Gao and Towsley 1999].

A situação é bem mais complexa para cargas interativas. Além da taxa N , os seguintes aspectos da interatividade têm algum impacto sobre a escalabilidade de protocolos de compartilhamento de banda, como visto em [Rocha et al. 2005a, Rocha et al. 2005b]:

- *Posição inicial*: O ponto da mídia no qual a requisição se inicia;
- *Duração*: A duração do segmento de mídia requisitado;
- *Tipo de requisição interativa*: Pausa, saltos;
- *Distância de salto*: A distância, na mídia, entre o ponto em que uma requisição termina e o ponto em que a próxima começa. Distância 0 indica uma pausa, negativa indica salto para trás e positiva indica salto para a frente;
- O número de requisições por sessão.

Tais trabalhos identificam a *dispersão temporal* e a *dispersão espacial* como fatores presentes em cargas interativas que têm *impacto primário* sobre a escalabilidade de protocolos de compartilhamento de banda. Dispersão temporal é apenas um termo conveniente para a taxa de chegada normalizada N . Já a dispersão espacial permite representar o impacto da posição inicial e da duração das requisições. A figura 1 ilustra este impacto, apresentando três cargas com níveis crescentes de dispersão. No eixo x tem-se o tempo

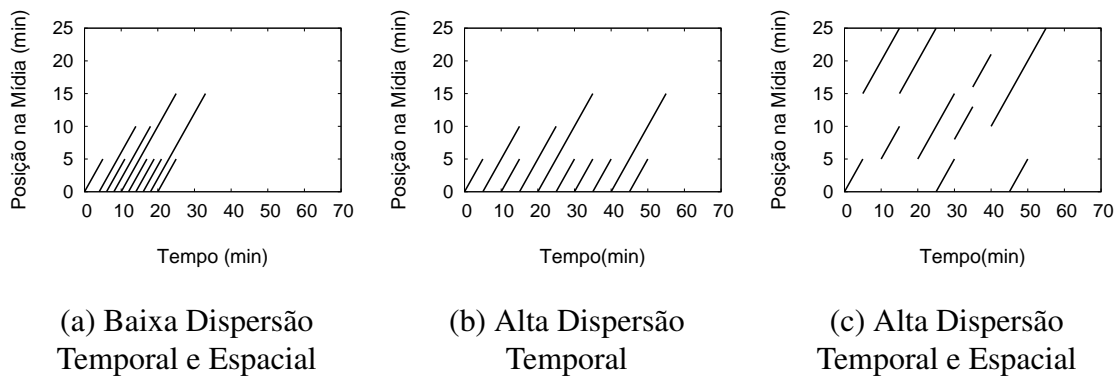


Figura 1. Cargas dispersas nas dimensões espacial e temporal

de chegada das requisições, e no eixo y as posições inicial e final de cada segmento requisitado. As três cargas incluem 10 requisições de mesma duração para a mesma mídia. Na figura 1(a) é vista a carga com menor dispersão. As requisições estão próximas no tempo, e a área de sobreposição entre as requisições, definida em termos de sua posição inicial e duração, é máxima. Nesta figura há melhores oportunidades de compartilhamento de banda. Na figura 1(b) a dispersão temporal aumenta, e as oportunidades de compartilhamento de banda diminuem. Na figura 1(c) a dispersão espacial aumenta, e as oportunidades de compartilhamento de banda se reduzem ainda mais.

O uso da dispersão permite classificar as cargas aqui usadas em grupos conforme o nível de interatividade presente nas mesmas [Rocha et al. 2005a, Rocha et al. 2005b]:

Interatividade Alta (IA): Cargas com alta dispersão espacial. A duração média das requisições está abaixo de 20% da duração da mídia, e a posição inicial média está entre 30% e 60% da duração da mídia. Tipicamente, menos de 30% das requisições começam na posição inicial da mídia, e as sessões têm pelo menos três requisições. Vídeos educacionais longos têm, tipicamente, este perfil.

Interatividade Média (IM): Cargas com nível intermediário de dispersão espacial. A duração média das requisições fica abaixo de 20% (assim como nas cargas IA), mas a posição inicial das requisições é mais concentrada (abaixo de 30% ou acima de 60% da duração da mídia). Em geral, as sessões têm menos de três requisições. Vídeos de entretenimento têm, tipicamente, este perfil.

Interatividade Baixa (IB): Cargas com pouca dispersão espacial. A duração média das requisições é mais longa (pelo menos 20% da mídia, em média), com a posição inicial fortemente concentrada no começo da mídia. Usualmente, temos cargas com menos de duas requisições por sessão. Cargas de áudio e de vídeos muito curtos (abaixo de 90 segundos) têm, tipicamente, este perfil.

2.2. Cargas Reais

Este trabalho faz uso um conjunto de cinco cargas reais oriundas de diversos domínios de aplicação: *vídeo educacional*, *vídeo de entretenimento*, *áudio de entretenimento*. Estas cargas são usadas em [Rocha et al. 2005a, Rocha et al. 2005b] como base para a geração

de um conjunto rico de cargas sintéticas realistas (ver seção 2.3). O sistema *eTeach*¹ contém vídeo educacional de uma grande universidade norte-americana, incluindo desde vídeos curtos com anúncios variados (abaixo de 5 minutos) até vídeos longos com aulas de 50 a 60 minutos. As cargas de entretenimento (uma de vídeo e duas de áudio), caracterizadas detalhadamente em [Costa et al. 2004], são oriundas de dois grandes provedores de conteúdo, um dos quais é UOL², e contém arquivos de áudio e vídeo curtos (tipicamente de até 10 minutos). As cargas contêm mídias com variados perfis de acesso e níveis de interatividade (seção 2.3), sendo a interatividade mais comum nas cargas de vídeo educacional. O sistema MANIC³, também de uma grande universidade norte-americana, contém vídeo educacional com características similares às das outras cargas [Rocha et al. 2005b].

Foram selecionados traços de acesso a 36 objetos escolhidos entre os mais populares nas cinco cargas. Destes, 11 se encaixam na categoria IA, 8 na categoria IM e 17 na categoria IB. A *localidade de referência* [Costa et al. 2004] está presente nestes traços: em média, 35% das interações são pausas e 45% são saltos para trás (a distância média de um salto é de 233 segundos). Os períodos de inatividade duram, em média, 232 segundos. Estes fatores motivam e são explorados pelas novas otimizações propostas.

2.3. Cargas Sintéticas Realistas

Este conjunto de cinco cargas reais é usado como base para gerar um amplo conjunto de cargas sintéticas realistas, as quais preservam as características chave da interatividade presentes nas cargas originais (seção 2.1). Estas cargas são usadas como entrada para um gerador de cargas validado em [Rocha et al. 2005b]. Este gerador reproduz, nas cargas sintéticas, os aspectos chave da interatividade presentes nas cargas reais.

Este gerador toma um traço de acessos a uma mídia numa carga real, e constrói um modelo de estados probabilístico correspondente, onde cada estado representa segmentos de 10 segundos da mídia, pausas, início de sessão e término de sessão. As probabilidades de transição entre os estados são computadas a partir do traço de acessos. As durações das pausas e dos segmentos exibidos parcialmente têm distribuição exponencial com média extraída do traço de acesso. Este modelo é chamado *perfil de carga*, e é usado para gerar sessões de acesso interativo. Dado um perfil, varia-se a taxa de chegada de sessões e a semente do gerador de números aleatórios para criar um conjunto extenso de cargas de perfil similar. Aqui, foram geradas algumas centenas de cargas de 36 diferentes perfis.

3. Protocolos

Esta seção discute protocolos escaláveis correntes de compartilhamento de banda (seção 3.1), apresenta o impacto da interatividade sobre a escalabilidade destes protocolos (seção 3.2) e discute otimizações para estes protocolos para cargas interativas (seção 3.3).

3.1. Protocolos Escaláveis de Compartilhamento de Banda

Protocolos escaláveis de compartilhamento de banda permitem obter redução de banda de servidor por meio da aplicação de um princípio simples comum: cada cliente recebe dois fluxos, um para ser exibido de imediato, e outro para armazenamento

¹eTeach - Learning on Demand - <http://eteach.cs.wisc.edu/index.html>

²Universo Online - <http://www.uol.com.br>

³RIPPLES/MANIC - <http://manic.cs.umass.edu>

e exibição futura. Quando a exibição do primeiro alcança o ponto em que o segundo começou a ser armazenado, o primeiro é desconectado e apenas o segundo permanece ativo [Bar-Noy and Ladner 2003, Coffman et al. 2002, Eager and Vernon 1998, Eager et al. 2000, Eager et al. 2001, Gao and Towsley 1999, Hua et al. 1998]. Dentre estes protocolos, são de maior interesse os que permitem serviço imediato ao cliente e, em especial, o *Bandwidth Skimming* [Eager et al. 2000, Eager et al. 2001], dada sua escalabilidade próxima ao ótimo, com cargas seqüenciais, para banda de servidor e rede [Eager et al. 2001, Zhao et al. 1999], e sua escalabilidade superior à do *Patching*, em banda de servidor, para cargas interativas [Rocha et al. 2005b].

Bandwidth Skimming baseia-se na *união hierárquica* de fluxos *multicast* para reduzir requisitos de largura de banda de servidor. A idéia é que, quando da chegada de uma nova requisição de um cliente, um novo fluxo *multicast* (IP ou ao nível de aplicação) é criado para entregar a mídia requerida. Na variação *Closest Target* deste protocolo, o cliente escuta este novo fluxo de *multicast* (garantindo, assim, serviço imediato), e escuta também o fluxo *multicast* já existente mais próximo da posição requisitada (este é o fluxo *alvo* do cliente). Quando o cliente completa a escuta, por meio do fluxo novo, de todo o trecho que ele perdeu com relação a seu alvo, este fluxo novo pode ser desconectado e dizemos que os clientes se uniram. A partir daí, os clientes do fluxo unido passam a escutar um novo alvo, criando-se assim uma árvore hierárquica de uniões de fluxos. A área de armazenamento, no cliente, onde o fluxo alvo é armazenado para fins de união de fluxos, é chamada *merge buffer*. Durante uma tentativa de união de fluxos, o fluxo alvo pode terminar antes que seja alcançado pelo novo fluxo. Neste caso, um novo fluxo alvo é selecionado usando-se o mesmo critério anterior. Todo o conteúdo recebido do fluxo alvo, e armazenado no *merge buffer* durante a tentativa de união não efetivada, é perdido.

Existem outras estratégias para seleção de fluxos alvos e construção da hierarquia de uniões [Bar-Noy and Ladner 2003, Coffman et al. 2002, Eager et al. 2000, Eager et al. 2001], mas o algoritmo *Closest Target* é simples e muito próximo do ótimo [Eager et al. 2001, Zhao et al. 1999]. Para o mesmo, a banda média de servidor requerida, para cargas seqüencias e chegadas *Poisson*, é logarítmica com relação à taxa de chegada normalizada N , enquanto que para *Patching*, esta banda cresce com a raiz quadrada. Neste trabalho, o termo *Bandwidth Skimming* é usado para referência ao protocolo *Closest Target* que exija uma banda de cliente igual ao dobro da taxa de exibição da mídia.

3.2. Impacto da Interatividade na Escalabilidade

A seção 2.1 discutiu a dispersão como fator de impacto primário da interatividade sobre a escalabilidade de protocolos de compartilhamento de banda. Este impacto foi avaliado em [Jin and Bestavros 2002, Tan et al. 2002a]. Em [Tan et al. 2002a] evidenciou-se que a banda média de servidor requerida por qualquer protocolo de compartilhamento de banda passa a crescer, sob interatividade, com a raiz quadrada de taxa de chegada normalizada N , em lugar do crescimento logarítmico exibido pelas cargas seqüenciais. Por fim, em [Rocha et al. 2005b], mostra-se que este impacto é mais forte nas cargas com nível alto ou médio de interatividade, o que faz destas cargas alvo de estratégias de otimização.

3.3. Otimizações para Cargas Interativas

Otimizações e protocolos prévios, orientados para cargas interativas, armazenam o conteúdo recebido do servidor em área local do cliente, ou em área comparti-

lhada, a fim de evitar retransmissão. Este conteúdo é obtido nos *períodos ativos do cliente*, durante os quais o cliente está requisitando dados do servidor para exibição [Hua and Tran 2005, Liao and Li 1997, Ma et al. 2005, Rocha et al. 2005a, Rocha et al. 2005b, Rodrigues and Leão 2005]. Tais trabalhos evidenciam a importância do uso de áreas de armazenamento para se reduzir a banda média de servidor requerida para cargas interativas.

Contudo, o uso de área local de armazenamento não garante redução, e pode até mesmo aumentar, a banda média requerida [Rocha et al. 2005a, Rocha et al. 2005b]. Ao combinar protocolos de compartilhamento de banda com armazenamento no cliente, entram em jogo duas forças opostas. Por um lado, a área de armazenamento evita retransmissão (já que parte das requisições é servida localmente). Por outro, apenas fragmentos das requisições são enviadas ao servidor. A carga por ele percebida se torna mais dispersa, reduzindo-se as oportunidades para compartilhamento de banda.

Dentre as otimizações previamente propostas, três são de especial interesse [Rocha et al. 2005b], porque são particularmente úteis para a composição de estratégias híbridas. Em primeiro lugar, porque obtêm grande redução de banda de servidor requerida. Em segundo, porque exploram características distintas dos protocolos e cargas reais interativas, com resultados distintos para diferentes cargas e taxas de chegada (embora compartilhem o princípio comum de armazenar conteúdo numa área local ao cliente).

KEEP MERGE BUFFER (KMB) armazena em área local todos os segmentos de mídia que seriam descartados do *merge buffer* quando uma tentativa de união falha (seção 3.1). O cliente tem pouco controle sobre o conteúdo armazenado, o qual depende dos alvos não atingidos, resultando em grande fragmentação da área de armazenamento.

LOCALITY (LOC) armazena todo o conteúdo exibido para o usuário em área local ao cliente, explorando a localidade de referência presente em cargas interativas reais [Almeida et al. 2001, Costa et al. 2004]. Como o conteúdo armazenado ao longo de uma requisição somente poderá ser usado na requisição seguinte, LOC também explora o número alto de requisições por sessão frequente nas cargas mais interativas (seção 2.1). Por fim, diferentemente de KMB, o conteúdo armazenado por LOC depende apenas dos segmentos requisitados pelo usuário em uma sessão, e não da taxa de chegada N .

SILENT PREFETCH (SP) oferece uma forma simples e ágil para a escolha e escuta de fluxos no período inativo. O cliente escuta o fluxo mais próximo do último segmento de mídia exibido para o usuário. Também escuta o alvo (no sentido de *Bandwidth Skimming*) deste fluxo, quando houver. Se o fluxo escutado terminar antes do período de inatividade, o cliente escuta o fluxo seguinte mais próximo e também seu eventual alvo. A figura 2 ilustra a escolha de fluxos, apresentando o conteúdo da área de armazenamento local do cliente para uma mídia de duração T num dado instante. Os fluxos disponíveis para escuta estão transmitindo os pontos A , B , C e D . Os trechos marcados em preto indicam área já preenchida, e P indica o último ponto da mídia exibido. O fluxo C será escutado porque é o mais próximo do último ponto exibido P . Seu alvo D também será escutado. SP explora a localidade de referência presente em cargas interativas [Costa et al. 2004], pois o primeiro fluxo escutado no início do período inativo é o mais próximo do último segmento de mídia exibido. A escolha é limitada porque, ao longo deste período, é restrita a fluxos sucessivos. Isto não impede a escuta de conteúdo já armazenado, mas reduz a fragmentação

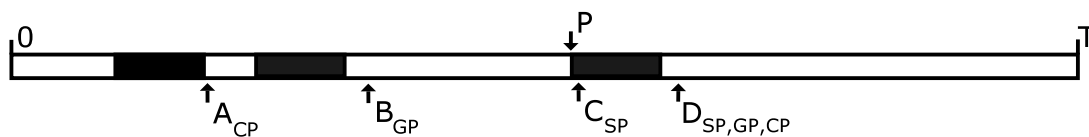


Figura 2. Escolha de fluxos de período inativo

da área de armazenamento. Por fim, esta escolha se justifica porque operações de pausa e saltos para a frente são frequentes em cargas interativas [Costa et al. 2004].

As seções seguintes apresentam e avaliam detalhadamente as novas otimizações para período inativo. A seguir, estratégias híbridas são usadas para combinar as otimizações de período ativo e inativo, e, por fim, é feita uma avaliação preliminar do impacto destas estratégias na banda de rede para uma topologia real, que mostra os principais compromissos envolvidos.

4. Novas Estratégias de Otimização para Período Inativo

Esta seção introduz novas otimizações para o período de inatividade do cliente. A seção 4.1 apresenta as otimizações, a seção 4.2 avalia os resultados obtidos, e a seção 4.3 discute aspectos de implementação.

4.1. Estratégias de Otimização

A existência de períodos de inatividade relativamente longos entre as requisições de um usuário a uma mesma mídia (232 segundos em média nos 36 perfis de carga usados) motiva o projeto de estratégias de otimização que exploram tais períodos para receber conteúdo antecipadamente e colocá-lo em área de armazenamento local ao cliente. O cliente é dito inativo quando o usuário não está exibindo mídia, ou quando a exibe a partir da área de armazenamento. As novas estratégias propostas usam estes períodos de inatividade para receber fluxos criados para atender a outros clientes, colocando os dados recebidos na área local de armazenamento. Elas escutam um máximo de dois fluxos, seguindo uma configuração típica para o *Bandwidth Skimming* original (seção 3.1). Dessa forma, as requisições do usuário são atendidas diretamente dessa área sempre que possível, evitando retransmissão de conteúdo. O número de requisições por sessão tem impacto nestas otimizações, já que o conteúdo armazenado em um período inativo somente será usado em períodos ativos subsequentes. Estas otimizações, assim como qualquer otimização para protocolos de compartilhamento de banda que use área de armazenamento local ao cliente, estão sujeitas à fragmentação da carga imposta ao servidor (seção 3.3).

GREEDY PREFETCH (GP) evita a escuta de conteúdo já armazenado, buscando povoar ao máximo a área de armazenamento local. Dois fluxos são escolhidos independentemente (um não precisa ser alvo do outro). Os fluxos escutados são os mais próximos do último ponto exibido da mídia (para frente ou para trás, explorando a localidade de referência) e que transmitam dados inéditos. Na figura 2, os fluxos escutados são D e B , nesta ordem (C não é escolhido porque transmite conteúdo já armazenado). GP não previne a fragmentação da área de armazenamento, pois um fluxo pode preencher parcialmente uma área vazia e criar novos fragmentos. É o que acontecerá na figura 2 se o

fluxo B terminar antes de alcançar o ponto P . Se não há fluxo disponível com conteúdo inédito, escolhe-se o que transmite o mínimo de conteúdo já armazenado.

COOPERATIVE PREFETCH (CP) tem como objetivo evitar a fragmentação permitida por GP. O primeiro fluxo é escolhido seguindo os mesmos critérios de GP. Já o segundo é escolhido de forma a minimizar a fragmentação, e para tanto, nunca se escolhem fluxos cujo conteúdo ocupe o meio de uma área vazia (como o fluxo B na figura 2). Em lugar disto, escolhe-se o fluxo mais próximo do último ponto exibido e que esteja ou no início de uma área vazia ou no meio de uma área já ocupada. Dessa forma, caso o fluxo termine antes de preencher toda a área, garante-se que ela não será dividida e, sim, apenas reduzida. No exemplo da figura 2, o primeiro fluxo escolhido será D e o segundo será A . Os dois fluxos cooperam entre si, um preenchendo a área de armazenamento, o outro removendo a fragmentação. Mas, como o segundo fluxo não previne receber conteúdo já armazenado, o preenchimento da área de armazenamento é menos agressivo que em GP.

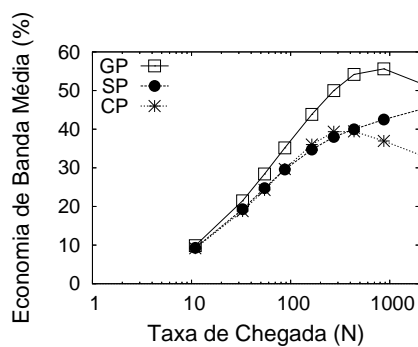
4.2. Avaliação

Esta seção apresenta uma avaliação quantitativa das novas otimizações GP e CP e as compara com a otimização prévia de período inativo SP. Para tanto, um simulador para *Bandwidth Skimming* foi construído e validado para cargas seqüenciais. A banda média de servidor requerida diferiu, para uma ampla faixa de taxas de chegada, em apenas 11% do resultado analítico correspondente [Eager et al. 2001]. Este simulador foi, então, adaptado para implementar as novas otimizações. Os resultados aqui apresentados são a média de 5 execuções, com desvio padrão que não excede a 2%.

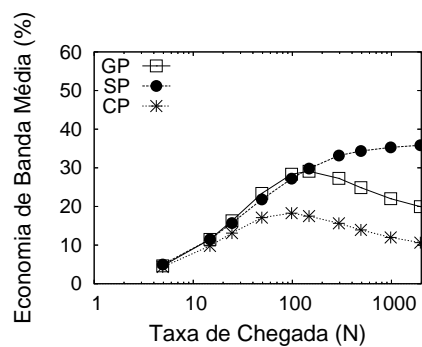
A medida usada para avaliar a escalabilidade das estratégias de otimização propostas é a economia de banda do servidor em relação ao protocolo *Bandwidth Skimming* original. A figura 3 resume os principais resultados obtidos para as três otimizações e para diferentes níveis de interatividade. O eixo x apresenta os valores crescentes da taxa de chegada de requisições normalizada N . O eixo y mostra a economia de banda obtida se comparada ao protocolo original. Valores acima de 0 indicam economia de banda. Resultados para outras cargas da mesma classe são qualitativamente similares aos apresentados.

As curvas para as três estratégias de otimização apresentam um comportamento geral similar. Para valores mais baixos de taxa de chegada de requisições N , a dispersão temporal é alta e conseqüentemente há poucas oportunidades para união de fluxos. O pouco conteúdo armazenado pelo cliente já permite redução na retransmissão, e leva a alguma economia de banda de servidor. À medida que a taxa de chegada cresce, mais conteúdo é armazenado. Como resultado, o impacto da redução de retransmissão domina, e a economia obtida cresce. A partir de certo valor de N , a dispersão espacial introduzida na carga entregue ao servidor domina, e a economia de banda começa a decrescer (seção 3.2). O ponto em que o pico ocorre depende da carga e da estratégia de otimização.

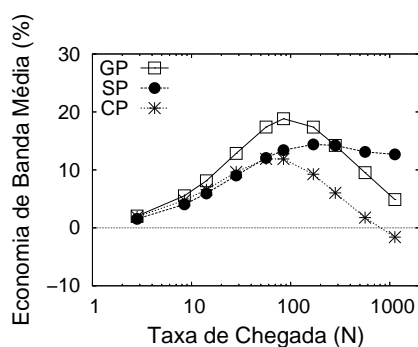
Um aspecto notável da otimização GP é verificado quando se comparam as figuras 3(a,b). Na primeira, GP obtém economia significativamente superior à de SP, enquanto que a situação se inverte na segunda. Esta diferença se deve ao preenchimento agressivo da área de armazenamento presente nas otimizações gulosas. Na carga da figura 3(b) a duração média das requisições é superior em 25% ao tempo médio de inatividade. Com requisições relativamente longas, GP consegue ocupar 90% do tempo disponível com escuta de conteúdo, mesmo com taxas tão baixas como $N = 48$ (100% do tempo é ocupado



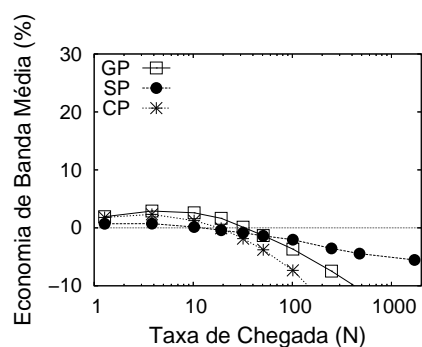
(a) Interatividade Alta, GP supera SP



(b) Interatividade Alta, SP supera GP



(c) Interatividade Média



(d) Interatividade Baixa

Figura 3. Economia de banda de servidor para diferentes níveis de interatividade

com $N = 291$). Com isto, a capacidade de GP de armazenar conteúdo, e portanto de evitar retransmissão e economizar banda, se esgota a taxas de chegada relativamente baixas. Já SP, que é mais restrita na escolha de fluxos, somente ocupa 90% do tempo de inatividade com escuta de conteúdo com $N = 486$. Portanto, SP consegue economia de banda crescente mesmo para taxas de chegada mais altas. Essa situação é similar em todas as cargas nas quais a duração média das requisições excede o tempo médio de inatividade.

SP não sofre fortemente o efeito degenerativo da fragmentação e, portanto, sua economia de banda somente decresce para taxas de chegada mais altas (figura 3(b)). É a estratégia que permite economia para faixa mais ampla de taxas de chegada. GP é mais agressiva ao preencher a área de armazenamento, e obtém economia superior a qualquer outra otimização (figuras 3(a,c)). Como GP está sujeita ao efeito degenerativo da fragmentação, a economia de banda cai abruptamente. CP não se mostrou efetiva para reduzir a fragmentação. A economia obtida com o preenchimento agressivo é tão elevada, que GP sempre supera aquela obtida com CP (figura 3).

A figura 4 ilustra o preenchimento agressivo da área de armazenamento de GP, mostrando o percentual médio de preenchimento da área de armazenamento local dos clientes para diversas taxas de chegada, para a carga da figura 3(a). A capacidade de GP economizar banda depende justamente deste preenchimento agressivo. Enquanto o percentual de preenchimento cresce rapidamente com a taxa de chegada, a economia de

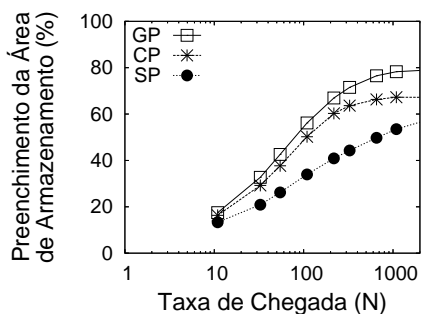


Figura 4. Preenchimento da área de armazenamento (Carga IA).

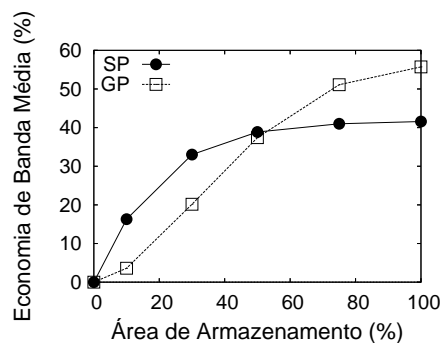


Figura 5. Impacto da restrição na área de armazenamento (Carga IA, N=652).

banda cresce. Para taxas mais altas, o preenchimento cresce mais lentamente. A economia de banda atinge então seu máximo e começa a decrescer. CP não previne a escuta de conteúdo já armazenado, o que limita sua capacidade de armazenamento. A situação é agravada em SP que, além desta limitação, tem escolha de fluxos mais restrita.

Para cargas IM os resultados são similares mas como estas cargas são menos dispersas, o impacto da fragmentação é mais forte e a economia de banda mais reduzida. Para cargas IB, ainda menos dispersas que as cargas IM, o impacto é ainda mais significativo.

Outras estratégias foram exploradas no intuito de reduzir a fragmentação, porém com economia muito modesta. Exigir que os fluxos a escutar não causem fragmentação limita seriamente o repertório de fluxos disponíveis e, portanto, a quantidade de dados armazenados. A redução de retransmissão é pequena e a economia obtida também o é.

Até aqui, este trabalho considerou que o cliente dispunha de área de armazenamento suficiente para comportar toda a mídia. O impacto da restrição desta área é avaliado a seguir. A figura 5 mostra o impacto da restrição para as duas estratégias mais eficientes ($N = 652$, carga da figura 3(a)). No eixo x temos o tamanho da área como um percentual da duração da mídia e no eixo y temos a economia de banda sobre o *Bandwidth Skimming* original. Com área restrita a 40% da duração da mídia, SP apresenta economia de banda próxima àquela obtida quando não há restrições. Já GP, que preenche a área mais agressivamente, se beneficia do aumento da área disponível. Em síntese, estratégias gulosas, como GP, exigem que o cliente disponha de área suficiente para armazenar toda a mídia ou seus benefícios serão muito limitados. Já SP permite reduzir a área disponível no cliente, sem com isto reduzir significativamente a economia de banda obtida.

4.3. Aspectos de Implementação

O processo de seleção de fluxos usado por GP faz necessário enviar, do servidor ao cliente, informação sobre fluxos disponíveis para escuta. Mas, como o número de fluxos ativos é limitado (seção 3.2), torna-se viável enviar mensagem do servidor ao cliente com a lista de fluxos disponíveis, de forma que o próprio cliente selecione os mais adequados. Embora este trabalho se refira a área de armazenamento local ao cliente, ela poderia se situar num agente, de forma que o cliente dispensasse qualquer modificação. Por fim, muitas destas otimizações poderiam ser aplicadas, sem grande esforço, a outros protocolos de compartilhamento de banda.

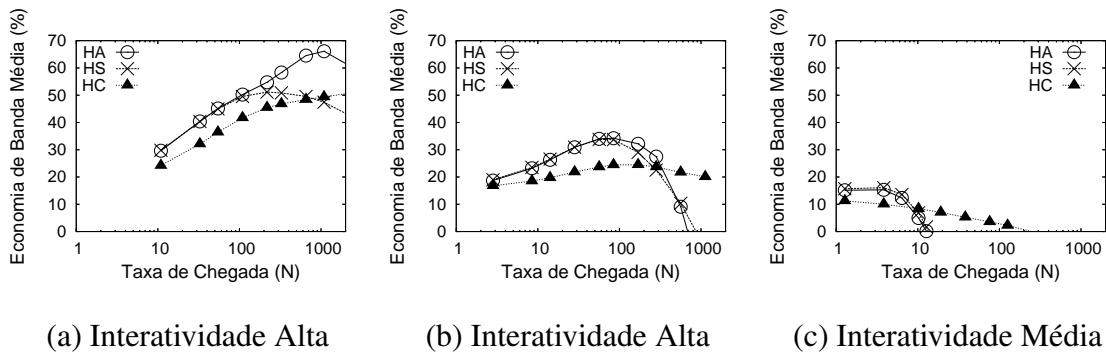


Figura 6. Economia de banda de servidor para estratégias híbridas

5. Estratégias Híbridas

Esta seção introduz novas estratégias híbridas. A seção 5.1 apresenta as estratégias, a seção 5.2 avalia seus resultados, e a seção 5.3 discute aspectos de implementação.

5.1. Estratégias

Estratégias híbridas combinam diferentes otimizações com o intuito de explorar as diferentes características de cada uma delas. De um lado, buscam-se estratégias que atuem sobre uma faixa mais ampla de taxas de chegada. De outro, estratégias com economia de banda mais expressiva, ainda que em cenários mais restritos.

HYBRID SYNERGIC (HS) é uma estratégia que combina as otimizações mais bem sucedidas no período ativo, LOC e KMB, com a otimização de período inativo menos sujeita à fragmentação, SP. Aqui, o objetivo é usar SP para minimizar o efeito degenerativo da fragmentação introduzido por KMB (seção 3.3).

HYBRID CONSERVATIVE (HC) combina apenas estratégias menos sujeitas ao efeito degenerativo da fragmentação, SP e LOC, a fim de obter economia de banda em faixas amplas de taxas de chegada. Em contrapartida, limita-se a economia de banda.

HYBRID AGGRESSIVE (HA) objetiva o preenchimento agressivo da área de armazenamento, como forma de maximizar a economia de banda por redução de retransmissão. Para tanto, combina LOC, que armazena todo o conteúdo exibido, KMB, que evita o descarte do conteúdo do *merge buffer*, e GP, que dá preferência ao armazenamento de conteúdo inédito. Em contrapartida, potencializa-se o efeito da fragmentação.

5.2. Avaliação

O mesmo simulador usado na seção 4 foi estendido para permitir a combinação de estratégias. A figura 6 apresenta a economia de banda de servidor para as estratégias híbridas para cargas IA, IM e IB típicas. HC tem sucesso ao combinar otimizações que não apresentam efeito degenerativo acentuado. A economia de banda obtida excede a economia das estratégias individuais, e a faixa de taxas de chegada onde se obtém economia significativa é ampliada. HA obtém melhor economia de banda que HC, especialmente para taxas de chegada mais baixas, já que armazena conteúdo de forma agressiva, e tem mais sucesso ao evitar retransmissões. HS falha na tentativa de suprimir o efeito degenerativo da fragmentação herdado de KMB. Este efeito degenerativo, presente em HA e HS,

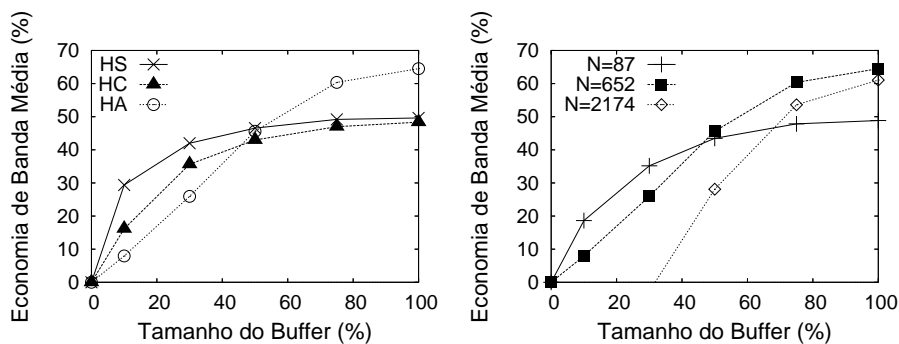
dá ensejo a um futuro desenvolvimento de protocolo adaptativo, que monitore a carga e decida pelo preenchimento mais ou menos agressivo da área de armazenamento.

O impacto da restrição de espaço na área de armazenamento é avaliado a seguir. Assim como na seção 4.2, a localidade de referência guia a gerência da área de armazenamento do cliente. A figura 7 mostra o impacto da restrição do tamanho da área de armazenamento para as estratégias híbridas. No eixo x temos o tamanho da área como um percentual da duração da mídia e no eixo y temos a economia de banda sobre o *Bandwidth Skimming* original. Para cargas IA e $N = 652$ (figura 7(a)), HA preenche a área agressivamente e sempre se beneficia do aumento da área disponível para armazenamento. HC e HS não preenchem a área de forma tão agressiva e obtêm, com área de armazenamento restrita a 50% da duração da mídia, economia de banda próxima à obtida sem restrições nesta área. A figura 7(b) mostra, para a estratégia HA e uma carga IA, o impacto da restrição de área em diversas taxas de chegada. HA, que preenche a área de armazenamento agressivamente, é fortemente afetada pela restrição da área de armazenamento. Para taxas de chegada mais altas, mesmo uma pequena redução na área de armazenamento tem forte impacto na economia de banda obtida pela estratégia.

Em síntese, similarmente ao que ocorre para otimizações individuais (seção 4), estratégias agressivas no preenchimento da área de armazenamento (HA) obtêm economias de banda superiores, mas exigem que o cliente disponha de área suficiente para armazenar toda a mídia ou seus benefícios serão muito limitados. Já estratégias menos agressivas (HS, HC) permitem ao cliente operar com área de armazenamento reduzida a 50% da duração da mídia e, ainda assim, obter economia de banda expressiva.

5.3. Aspectos de Implementação

A implementação de uma política híbrida não é significativamente mais complexa que a das políticas individuais que a integram, desde que se exclua a combinação de estratégias fortemente conflitantes na forma de escolha de fluxos (como GP e CP). Isto ocorre porque os eventuais conflitos entre conteúdo obtido por diferentes otimizações são prontamente resolvidos pelo cliente dando-se preferência para armazenamento ao conteúdo mais próximo do último segmento de mídia exibido.



(a) Interatividade Alta, $N=652$

(b) Interatividade Alta, HA

Figura 7. Impacto da restrição na área de armazenamento

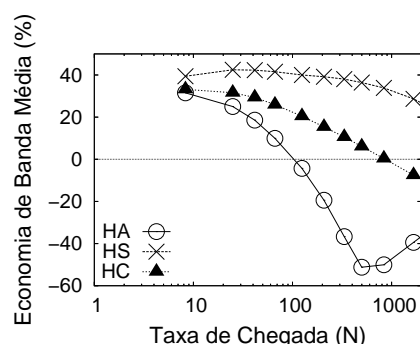


Figura 8. Economia de banda de rede para nível de interatividade alto

6. Análise de Banda de Rede

As novas estratégias de otimização usam períodos de inatividade do cliente para escutar fluxos sendo entregues a outros clientes. Se os fluxos estiverem sendo transmitidos para outra região da rede que não a do cliente, novos fluxos de rede serão criados para atendê-lo. Esta seção faz uma avaliação preliminar do impacto das estratégias híbridas sobre a banda de rede, apontando os principais compromissos envolvidos.

A topologia de rede usada foi levantada por meio de sondagens *traceroute*⁴, acionadas a partir de 53 servidores na Internet. De abril a setembro de 2005 foram feitas 6 coletas diárias das rotas entre estes 53 servidores. Removidos erros e rotas incompletas, restaram pelo menos 300 rotas entre cada par de servidores, num total de 4915 nodos.

Um dos 53 nodos de sondagem foi escolhido para sediar o servidor. Os clientes foram distribuídos uniformemente dentre os demais nodos, podendo existir mais de um cliente por nodo. A escolha do servidor é arbitrária porque os clientes são uniformemente distribuídos e a distância de qualquer nodo de sondagem aos outros tem distribuição normal. Os fluxos percorrem o menor caminho do servidor ao cliente. A banda de rede é calculada como o somatório da duração dos fluxos nos enlaces que conectam os nodos. A banda é afetada pelo número de fluxos de servidor que atravessam a rede, e pelo número de fluxos de rede criados para compartilhar os fluxos de servidor por entre múltiplos nodos. Não são necessários fluxos de rede para compartilhar fluxos entre clientes de um mesmo nodo. Este cenário representa o pior caso de distribuição de clientes, já que quanto mais tendenciosa é a distribuição, menos fluxos de rede são necessários para atendê-los.

A figura 8 apresenta a economia de banda média de rede para as estratégias HS, HA e HC para uma carga IA típica. HS é a estratégia com a maior economia de banda de rede (até 40% sobre o protocolo *Bandwidth Skimming*) porque, para cargas altamente interativas, a grande economia de banda do servidor sobrepuja o aumento de banda causado pela escuta de fluxos em períodos inativos. HA, que incorpora a otimização gulosa GP, ocupa rapidamente todo o tempo disponível para a escuta de fluxos e é, portanto, pouco efetiva em economia de banda de rede. HC é menos efetiva que HS porque esta incorpora KMB, a qual fragmenta fortemente a área de armazenamento. Com a carga mais fragmentada, a duração dos períodos de inatividade é reduzida, o que reduz o impacto sobre a banda de rede.

⁴Traceroute Tool - <ftp://ftp.ee.lbl.gov/traceroute.tar.gz>

7. Conclusão

Este trabalho aborda novas otimizações para protocolos de compartilhamento de banda para período inativo e novas estratégias híbridas, que combinam seletivamente diversas otimizações. As novas estratégias híbridas se mostraram muito efetivas, e especialmente para as cargas mais interativas, justamente aquelas nas quais o protocolo *Bandwidth Skimming* tem sua escalabilidade mais afetada pela interatividade. As estratégias híbridas agressivas se mostram capazes de obter maior economia de banda de servidor que qualquer estratégia individual. Já as não agressivas, embora com economia mais limitada, se mostram capazes de obter economia de banda em faixas mais amplas de taxas de chegada.

Estratégias híbridas não agressivas permitem obter economia significativa de banda de servidor, mesmo com a área de armazenamento restrita a 50% da duração da mídia. O mesmo não ocorre com as estratégias agressivas, que necessitam de uma área de armazenamento capaz de conter toda a mídia, a fim de permitirem economia de banda próxima ao máximo que pode ser obtido por uma dada estratégia.

Estratégias híbridas não agressivas permitem obter grande economia de banda de rede para cargas de nível alto interatividade, e podem se aproximar da banda de rede para o *Bandwidth Skimming* original, em virtude da grande economia de banda de servidor. Já as estratégias mais agressivas são pouco econômicas em banda de rede, apesar de eficazes com relação à banda de servidor, por usar agressivamente o período inativo para coletar conteúdo. A análise detalhada do uso da banda de rede é deixada como trabalho futuro.

Considerando banda de servidor, as estratégias mais competitivas são:

- HYBRID CONSERVATIVE (HC), que permite economia significativa (até 51%) em ampla faixa de taxas de chegada, e para ampla variedade de cargas,
- HYBRID AGGRESSIVE (HA), que permite maior economia (até 66%).

Como a interatividade tem impacto significativo sobre o desempenho das estratégias híbridas, é deixado como trabalho futuro o projeto de um novo protocolo híbrido adaptativo que monitore a carga, identifique aspectos relevantes (como o nível de interatividade, a duração dos períodos inativos), e decida quais otimizações a ativar.

Por fim, registra-se o agradecimento ao doutorando Alex Borges e ao mestrando Eric Guimarães (UFMG) que cederam os dados da topologia de rede real.

Referências

- Almeida, J., Krueger, J., Eager, D., and Vernon, M. (2001). Analysis of Educational Media Server Workloads. In *Proc. NOSSDAV*, Port Jefferson, NY.
- Bar-Noy, A. and Ladner, R. (2003). Competitive On-Line Stream Merging Algorithms for Media-on-Demand. *Journal of Algorithms*, 48(1):59–90.
- Bradshaw, M., Wang, B., Sen, S., Gao, L., Kurose, J., Shenoy, P., and Towsley, D. (2003). Periodic Broadcast and Patching Services- Implementation, Measurement and Analysis in an Internet Streaming Video Testbed. *Multimedia Systems*, 9(1):78–93.
- Cheshire, M., Wolman, A., Voelker, G., and Levy, H. (2001). Measurement and Analysis of a Streaming Media Workload. In *Proc. USENIX-USITS*, San Francisco, CA.
- Coffman, E., Momcilovic, P., and Jelenkovic, P. (2002). The Dyadic Stream Merging Algorithm. *Journal of Algorithms*, 43(1):120–137.

- Costa, C., Cunha, I., Borges, A., Ramos, C., Rocha, M., Almeida, J., and Ribeiro-Neto, B. (2004). Analyzing Client Interactive Behavior on Streaming Media Servers. In *Proc. 13th World Wide Web Conference*, New York, NY.
- Eager, D. and Vernon, M. (1998). Dynamic Skyscraper Broadcasts for Video-on-Demand. In *Proc. Int'l Workshop on Multimedia Information Systems*, Istanbul, Turkey.
- Eager, D., Vernon, M., and Zahorjan, J. (2000). Bandwidth Skimming: A Technique for Cost-Effective Video on Demand. In *Proc. MMCN*, San Jose, CA.
- Eager, D., Vernon, M., and Zahorjan, J. (2001). Minimizing Bandwidth Requirements for On-Demand Data Delivery. *IEEE Trans. on Knowledge and Data Engineering*, 13(5):742–757.
- Gao, L. and Towsley, D. (1999). Supplying Instantaneous Video-on-Demand Services Using Controlled Multicast. In *Proc. IEEE MMCN*, Florence, Italy.
- Hua, K., Cai, Y., and Sheu, S. (1998). Patching: A Multicast Technique for True Video-on-Demand Services. In *Proc. ACM MULTIMEDIA*, Bristol, UK.
- Hua, K. A. and Sheu, S. (1997). Skyscraper Broadcasting: A New Broadcasting Scheme for Metropolitan Video-on-Demand Systems. In *Proc. ACM SIGCOMM*, Cannes, France.
- Hua, K. A. and Tran, D. A. (2005). Range Multicast for Video on Demand. *Multimedia Tools and Applications*, 27(3):367–391.
- Jin, S. and Bestavros, A. (2002). Scalability of Multicast Delivery for Non-sequential Streaming Access. In *Proc. ACM SIGMETRICS*, Marina Del Rey, CA.
- Liao, W. and Li, V. O. K. (1997). The Split and Merge Protocol for Interactive Video-on-demand. *IEEE MultiMedia*, 4(4):51–62.
- Ma, H., Shin, K., and Wu, W. (2005). Best-effort Patching for Multicast True VoD Service. *Multimedia Tools Applications*, 26(1):101–122.
- Padhye, J. and Kurose, J. (1998). An Empirical Study of Client Interactions with a Continuous-Media Courseware Server. In *Proc. NOSSDAV*, Cambridge, UK.
- Rocha, M., Maia, M., Almeida, J., and Campos, S. (2005a). Escalabilidade de Protocolos com Compartilhamento de Banda para Cargas de Mídia Contínua Realistas. In *Proc. SBC WPerformance*, São Leopoldo, RS, Brasil.
- Rocha, M., Maia, M., Ítalo Cunha, Almeida, J., and Campos, S. (2005b). Scalable Media Streaming to Interactive Users. In *Proc. ACM MULTIMEDIA*, Singapore, Singapore.
- Rodrigues, C. and Leão, R. (2005). Novas Técnicas de Compartilhamento de Banda para Servidores de Vídeo Sob Demanda Com Interatividade. In *Proc. Simpósio Brasileiro de Redes de Computadores*, Fortaleza, CE, Brasil.
- Tan, H., Eager, D., and Vernon, M. (2002a). Delimiting the Range of Effectiveness of Scalable On-Demand Streaming. In *Proc. Performance*, Rome, Italy.
- Tan, H., Eager, D., Vernon, M., and Guo, H. (2002b). Quality of Service Evaluations of Multicast Streaming Protocols. In *Proc. ACM SIGMETRICS*, Marina Del Rey, CA.
- Zhao, Y., Eager, D., and Vernon, M. (1999). Network Bandwidth Requirements for Scalable On-Demand Streaming. In *Proc. IEEE INFOCOM*, New York, NY.