

Agregação Dinâmica de Fluxos com Duração Conhecida em Redes WDM através de Balanceamento de Tráfego

Juliana de Santi¹, André C. Drummond¹, Nelson L. S. da Fonseca¹

¹Instituto de Computação – Universidade Estadual de Campinas (UNICAMP)
Caixa Postal 6167 – 13084-971 – Campinas – SP – Brasil

{santi, andred, nfonseca}@ic.unicamp.br

Abstract. *In this paper, we propose a new dynamic traffic grooming approach to known duration connection, which exploit a cost function to induce to traffic balancing in WDM mesh optical networks. The main contribution of the proposed approach is to allow a better traffic distribution on the network to reducing the demands rejection rate. We compared the performance of our algorithm to original HTA and results show significant improvements in blocking probability of connections, which highlights the benefits of the proposed approach.*

Resumo. *Neste trabalho, propõe-se uma nova abordagem para agregação dinâmica de tráfego com duração conhecida que utiliza uma função de custo que induz ao balanceamento de tráfego em redes ópticas em malha WDM. A principal contribuição da abordagem proposta é permitir uma melhor distribuição da carga na rede reduzindo a taxa de rejeição das demandas. Para avaliar o desempenho do algoritmo proposto foram realizadas comparações com a política de HTA original. Os resultados demonstram melhorias significativas na probabilidade de bloqueio das conexões, o que evidencia os benefícios da abordagem proposta.*

1. Introdução

Ao multiplexar vários comprimentos de onda em uma mesma fibra óptica, a tecnologia *Wavelength Division Multiplexing* (WDM) aumentou extraordinariamente a quantidade de banda disponível. Embora as aplicações com requisitos de banda passante elevada estejam aumentando de maneira vertiginosa, estes requisitos estão distantes de conseguir utilizar completamente a capacidade dos canais ópticos. O custo do provisionamento de um caminho óptico é elevado e independente da quantidade de recursos que serão utilizados. Logo, o custo de utilização de um caminho óptico é aumentado mediante a utilização parcial dos recursos provisionados. Além disso, a evolução das aplicações de Internet impulsionou o aumento na quantidade de dados gerados, para se ter idéia, estima-se que para 2012 o tráfego gerado por fluxos IP atinja 44 Exabytes (2^{60}) por mês [Systems 2008]. Desta forma, a preocupação com a utilização dos recursos das redes ópticas WDM é pertinente e necessária e o grande desafio é superar a grande disparidade entre a capacidade de um comprimento de onda, o qual pode ser da ordem de Gigabits por segundo (ex: OC-192, OC-768), e a baixa granularidade dos fluxos IP, a qual varia de poucos Megabits por segundo. A agregação de tráfego em redes WDM lida com esta disparidade, ou seja, tenta resolver o problema de agregar demandas de baixa velocidade (*subwavelengths*) em circuitos ópticos de alta velocidade para maximizar a utilização dos comprimentos de onda e minimizar custo.

O problema de agregação de tráfego pode considerar um cenário estático, no qual todas as chamadas são conhecidas a priori, o que pode ser formulado como um problema de programação linear. No entanto, com a expansão do uso da tecnologia WDM para as redes locais e metropolitanas, as demandas de tráfego passam a apresentar maior dinamicidade e precisam ser modeladas como uma função de tempo, dando, assim, origem a técnicas de agregação dinâmica de tráfego.

Com aumento de banda para os usuários finais, elevou-se as expectativas destes com relação a forma como as suas solicitações são atendidas. O usuário, atualmente, espera que suas solicitações sejam atendidas, levando-se em consideração um conjunto de requisitos de qualidade de serviço (*Quality-of-Service* - QoS), os quais são especificados em contratos (*Service Level Agreements* - SLAs) com o provedor dos serviços. As especificações dos SLAs possibilitam o conhecimento da informação sobre o tempo de duração (*holding-time*) das conexões, o que é também denominado na literatura de previsibilidade de tráfego.

Aplicações tais como servidores de *backup* ou vídeo conferência são uma nova tendência [Liu and Chen 2007] e a previsibilidade inerente a este tipo de tráfego tem emergido como uma estratégia para aumentar a eficiência da utilização dos recursos ópticos. Em [Doumith and Gagnaire 2007], investiga-se o impacto da previsibilidade de tráfego no desempenho da rede. Os autores do trabalho [Cavdar et al. 2009], levam em consideração o tempo de duração das demandas para o escalonamento de conexões sujeitas a proteção que admitem certa tolerância a atrasos. No trabalho [Tornatore et al. 2008], introduz-se uma abordagem para explorar o conhecimento do tempo de duração das conexões para a agregação dinâmica de demandas em redes ópticas WDM em malha. O conhecimento do tempo de duração das conexões é utilizado, também, para o provisionamento de demandas *multicast* em [Munir et al. 2009].

A abordagem *Holding-Time-Aware* (HTA) utilizada para a solução do problema de agregação dinâmica com conhecimento da duração das conexões reduz as taxas de bloqueio, quando comparada com algoritmos tradicionais para a realização de agregação em uma topologia restrita [Tornatore et al. 2008]. No entanto, este algoritmo tende a minimizar o número de caminhos ópticos utilizados e não se preocupa com o balanceamento da carga na rede.

A presente proposta introduz uma nova política de agregação dinâmica de tráfego para redes ópticas WDM em malha, a qual emprega uma função de custo baseada tanto no tempo de duração das conexões quanto na distribuição de tráfego na rede, para balancear o tráfego e, conseqüentemente, minimizar a taxa de rejeição das demandas solicitadas.

Foram realizadas simulações utilizando quatro topologias clássicas derivadas de cenários realistas compostas por OXCs com capacidade restrita de agregação. Os resultados obtidos mostram que o algoritmo proposto é capaz de manter a distribuição da taxa de bloqueio estável para todos os pares de comunicação, além de obter taxas de bloqueio significativamente inferiores a abordagem HTA.

Este trabalho é organizado conforme segue. Na Seção 2, apresenta-se o trabalho que introduz o conceito de tempo de duração conhecido para a agregação dinâmica de tráfego [Tornatore et al. 2008]. Na Seção 3, introduz-se o algoritmo proposto (HTBalancing). Na Seção 4, compara-se o desempenho do HTBalancing com o HTA. Finalmente,

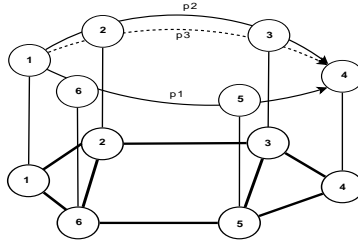


Figura 1. Exemplo: cálculo da custo de utilização de caminhos ópticos (modificado de [Tornatore et al. 2008])

na Seção 5 são delineadas as conclusões.

2. Agregação dinâmica com consciência da duração das demandas

Com as especificações de SLAs, a informação sobre a duração da chamadas é conhecida no momento do estabelecimento da conexão, consequentemente, é possível determinar o momento da finalização da conexão.

Um caminho óptico é explicitamente estabelecido no ato de aceitação da solicitação e é mantido durante toda existência da conexão. A agregação de tráfego faz a multiplexação de várias conexões no mesmo caminho óptico, ou seja, o tempo de vida deste caminho é determinado pela conexão que tem o maior tempo de término agregada a este caminho óptico. Assim, com o conhecimento do tempo de término das conexões, é possível determinar o tempo de vida do caminho óptico usado por um conjunto de conexões, o que possibilita que se tente aumentar a utilização do caminho óptico.

A agregação com conhecimento do tempo de duração das conexões consiste em agregar requisições explorando a informação sobre o tempo de duração das conexões e o tempo de vida dos caminhos ópticos, maximizando a utilização dos caminhos ópticos e minimizando a taxa de rejeição de solicitações futuras [Tornatore et al. 2007b, Tornatore et al. 2007a, Tornatore et al. 2008].

Uma demanda será aceita se houver um caminho óptico entre o par de nós origem-destino com recurso suficiente para fornecer a banda passante solicitada pela demanda. Podem existir vários possíveis caminhos entre um par de nós origem-destino com banda disponível para atender a demanda. Para determinar o caminho a ser escolhido, o custo de utilização dos caminhos é calculado com base no tempo de duração da demanda e no tempo de vida dos caminhos ópticos da seguinte forma:

$$ht(p_i) = \begin{cases} p_i \times \epsilon & \text{se } H_i \geq h \\ p_i \times \epsilon + p_i \times \Delta_t & \text{se } H_i < h \\ p_i \times h & \text{se } p_i \text{ é um novo caminho} \end{cases} \quad (1)$$

onde: $ht(p_i)$ é o custo do i -ésimo caminho; h é o tempo de duração da demanda; H_i é o tempo de vida do i -ésimo caminho; $\Delta_t = h - H_i$; $\epsilon = 10^{-5}$ é uma constante definida em [Tornatore et al. 2008]; p_i é o número de saltos ao longo do i -ésimo caminho.

Caso não haja nenhum caminho estabelecido, deve-se criar um para satisfazer a solicitação de estabelecimento da conexão. Caminhos ópticos já estabelecidos, ou seja, em utilização por outras demandas, recebem custo menor do que os novos caminhos. A

ideia principal da abordagem em [Tornatore et al. 2008] engloba dois princípios: utilizar o menor número possível de novos recursos para as novas solicitações e minimizar o tempo de vida adicional de um caminho já estabelecido para atender novas solicitações.

Para exemplificar o funcionamento desta abordagem, considere a rede apresentada na Figura 1, a qual recebe uma nova demanda $d_1(1,4,2,30)$ requisitando 2 Mbps de banda passante durante 30 segundos (*holding-time*) entre o par origem-destino (1 – 4). Os caminhos p_1 , p_2 e p_3 possuem, respectivamente, 4Mbps, 6Mbps e 1Gbps e são candidatos à atender d_1 . Enquanto p_1 e p_2 são caminhos ópticos já estabelecidos, com tempo de vida de 10s e 20s, p_3 é um novo caminho. A demanda pode ser atendida pelo caminho óptico p_1 estendendo seu tempo de vida por 20s, pois durante os 10s iniciais o caminho óptico agrega demandas que chegaram anteriormente e estão sendo atendidas, e d_1 pode ser agregada à este caminho com um custo adicional mínimo. Assim, o custo de utilização de p_1 é determinado pelo custo de utilização dos três comprimentos de onda em p_1 ($1 \rightarrow 6$, $6 \rightarrow 5$, $5 \rightarrow 4$) multiplicado pelos 20s finais de d_1 ($3 \times 20 = 60$) (linha 2 da equação). De maneira semelhante, se p_2 é escolhido, o seu tempo de vida precisa ser estendido pelos 10s finais de d_1 e o custo de utilização é dado pelos comprimentos de onda em p_2 ($1 \rightarrow 2$, $2 \rightarrow 3$, $3 \rightarrow 4$) multiplicado pelo tempo adicional necessário para transmitir pelo caminho óptico ($3 \times 10 = 30$). Por outro lado, p_3 é um novo caminho, ou seja, ele não está sendo utilizado por outras demandas e seu custo de utilização é o custo de estabelecê-lo através dos comprimentos de onda $1 \rightarrow 2$, $2 \rightarrow 3$ e $3 \rightarrow 4$, durante os 30s de duração de d_1 ($3 \times 30 = 90$) (linha 3 na equação). Portanto, levando-se em consideração o tempo de duração da demanda e o tempo de vida dos caminhos ópticos candidatos é possível verificar que o caminho óptico p_2 apresenta o menor custo e deve atender a demanda d_1 .

Associando custo elevado para os novos caminhos, os caminhos ópticos já estabelecidos ganham prioridade para atender as demandas solicitadas e deste modo um grande número de conexões são agregadas aos caminhos já estabelecidos, minimizado a subutilização de recursos ópticos.

Manter elevado o custo para caminhos ópticos com tempo de vida curto é também benéfico para melhorar a utilização dos recursos e minimizar a probabilidade de bloqueio, pois tão logo finalize o tempo de vida do caminho óptico, ele é liberado para atender demandas futuras com caminho fim a fim diferente do estabelecido. Por outro lado, um caminho óptico com valor elevado de tempo de vida será mantido por longo período de tempo até que todas as conexões a ele agregadas terminem de transmitir, logo é oportuno usar o recurso disponível neste caminho para atender as novas solicitações.

3. O Algoritmo Holding Time and Traffic Balancing

Apresentado na Seção 2, a abordagem de agregação dinâmica consciente do tempo de duração das chamadas tem como principal objetivo maximizar a agregação de demandas em caminhos ópticos já estabelecidos. Ao tentar minimizar a utilização de novos recursos para as novas demandas, o algoritmo HTA dificulta a distribuição balanceada do tráfego na rede, pois caminhos ópticos que possuem tempo de vida longo recebem custos menores e, conseqüentemente, o tráfego é concentrado nestes caminhos ópticos, os quais podem ter sua capacidade exaurida e passam a rejeitar solicitações futuras.

Para ilustrar as consequências decorrentes da negligência na distribuição balanceada de tráfego considere o exemplo hipotético da Figura 2. (S_1, D_1) , (S_2, D_2) e (S_3, D_3)

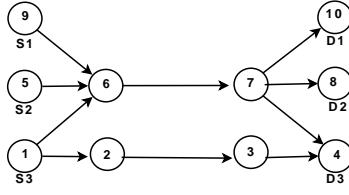


Figura 2. Exemplo de desbalanceamento de carga

são os possíveis pares origem-destino; o enlace $6 \rightarrow 7$ tem apenas uma unidade de banda passante disponível, enquanto os demais enlaces possuem disponibilidade de cinco unidades de banda passante. No instante atual t_0 , uma nova demanda $d_1(S_3, D_3, 1, 25)$, com duração de 25 segundos e requisito de 1 unidade de banda, chega na rede para ser provisionada. Entre o par S_3-D_3 há os caminhos p_1 (rota 1, 6, 7, 4) e p_2 (rota 1, 2, 3, 4) já estabelecidos, com tempo de vida, respectivamente, de 20s e 5s. Levando-se em consideração apenas o critério de tempo de vida para o cálculo do custo de utilização dos caminhos [Tornatore et al. 2008], o caminho p_1 seria escolhido para atender a demanda d_1 já que possui maior tempo de vida. Ao atender a demanda pelo caminho p_1 , o enlace entre os nós 6 e 7 tem toda sua banda consumida e, conseqüentemente, todas as demandas entre os pares (S_1, D_1) e (S_2, D_2) serão bloqueadas até que demandas que ocupam caminhos que passam pelo enlace entre os nós 6 e 7 terminem sua transmissão e o enlace volte a ter disponibilidade de banda. Logo, o enlace entre os nós 6 e 7 faz parte do caminho crítico (gargalo) dos pares origem-destino apresentados e sua utilização deve ser feita de maneira cautelosa.

Fica evidente, portanto, que além de minimizar o custo ao atender uma demanda através da agregação, é de fundamental importância, também, minimizar o impacto que este provisionamento causará sobre o estado da rede para maximizar a probabilidade de aceitação de demandas futuras.

Para superar a deficiência do HTA, propõe-se o algoritmo Holding-Time and Traffic Balancing (HTBalancing), que introduz o conceito de balanceamento de tráfego à abordagem de agregação dinâmica com *holding-time-aware*. Desta forma, a motivação principal do HTBalancing é agregar as demandas de forma a equalizar a distribuição de carga na rede com o intuito de minimizar a probabilidade de rejeição das demandas futuras.

A carga em uma rede é considerada balanceada se o tráfego é distribuído de forma proporcional por todos os enlaces desta rede, o que é equivalente a dizer que a banda passante disponível nos enlaces é utilizada de forma equitativa. Para agregar conexões de forma a balancear carga na rede, considera-se uma função de custo de utilização dos caminhos ópticos que leva em consideração o tempo de duração das demandas e a disponibilidade de banda nestes caminhos ópticos. A função de custo de utilização dos caminhos ópticos adotada é dada pela Equação 2:

$$C(p_i) = (ht(p_i) \times \beta) + \left(\frac{1}{bw(p_i)} \times (1 - \beta) \right) \quad (2)$$

onde: $C(p_i)$ é o custo de utilização do caminho óptico p_i ; $ht(p_i)$ representa o custo de

p_i calculado pela função de *Holding-Time-Aware* (Equação 1); β é um valor no intervalo $[0, 1]$ e determina o peso do *Holding-Time-Aware* para o custo $C(p_i)$; bw representa a banda disponível no caminho p_i ; $(1 - \beta)$ determina o peso da disponibilidade de banda no custo de $C(p_i)$.

Ao considerar o inverso da banda passante no segundo termo da Equação 2, os caminhos ópticos com menor disponibilidade de banda recebem custos mais elevados, ou seja, a utilização destes caminhos é evitada, de forma a não exaurir a banda disponível. Este cuidado na escolha do caminho é importante pois a transmissão de dados entre certos pares origem-destino deve ser essencialmente feita por caminhos específicos, logo estes caminhos são gargalos e se toda sua banda for consumida as demandas solicitadas serão necessariamente rejeitadas. Por outro lado, um caminho óptico com grande disponibilidade de recurso recebe um custo menor, ou seja, uma vez que há grande quantidade de banda passante este caminho recebe alta prioridade para atender a demanda, deixando os caminhos com menor quantidade de banda livres para atender demandas que os tem como um caminho gargalo.

O emprego da disponibilidade de banda passante na função de custo é motivado pelo fato de que esta medida fornece uma visão do estado da rede e é, portanto, de fundamental importância considerá-la nas estratégias de tomada de decisões de agregação de tráfego.

O funcionamento do algoritmo HTBalancing é apresentado no Pseudocódigo 1. Toda vez que uma nova solicitação chega a rede, é calculado o custo de utilização de todos os caminhos ópticos que possuem banda passante suficiente (caminhos candidatos) para atender a solicitação da demanda. O cálculo do custo de utilização dos caminhos candidatos é realizado de duas maneiras distintas que levam em consideração o fato de um caminho estar ou não em utilização. Um caminho já estabelecido com demandas a ele agregadas tem seu custo determinado conforme a linha 3 do Algoritmo 1. Nos caso em que o caminho candidato é um novo caminho, o seu custo é definido conforme a linha 6 do HTBalancing, na qual o termo $\frac{1}{bw(p_i)}$ da Equação 2 é substituído por 1. Esta ação é necessária porque o objetivo é priorizar a agregação, logo os novos caminhos precisam receber os custos mais elevados.

Para cada caminho candidato o custo é calculado e as arestas do grafo G correspondentes ao caminho em questão são atualizadas com peso igual a $C(p_i)$. Finalmente, para determinar o caminho óptico de menor custo, o tradicional algoritmo *Shortest-path* é empregado.

4. Resultados Numéricos

Para avaliar o desempenho do algoritmo HTBalancing foram realizadas simulações e comparações com o algoritmo HTA [Tornatore et al. 2008]. Para este fim, utilizou-se um simulador de eventos discretos desenvolvido em Java. O algoritmo de RWA utilizado pelos algoritmos HTBalancing e HTA é o de roteamento fixo-alternado com 5 rotas alternativas, e a política de alocação de comprimento de onda utilizada é a *First-Fit*. Na Equação 2, utilizou-se $\beta = 0.5$.

As topologias consideradas nas simulações foram as das redes NSF com 16 nós e 25 enlaces bidirecionais (Fig. 3(a)), Pan-European com 28 nós e 82 enlaces bidirecionais

Algoritmo 1 HTBalancing: Holding-Time and Traffic-Balancing

Entrada: Grafo $G = (V, E)$ representando a rede; demanda $d_j(s, d, b, ht)$ solicitando b unidades de banda passante por ht unidades de tempo entre os nós origem-destino (s, d) .

Saída: Um caminho óptico entre s e d ao qual d_j é agregada

- 1: **for** (todos caminhos ópticos candidatos p_i) **do**
 - 2: **if** p_i é um caminho já estabelecido **then**
 - 3: $C(p_i) \leftarrow (ht(p_i) \times \beta) + \left(\frac{1}{bw(p_i)} \times (1 - \beta) \right)$ (Equação 2)
 - 4: **else**
 - 5: **if** p_i é um novo caminho **then**
 - 6: $C(p_i) \leftarrow (ht(p_i) \times \beta) + (1 \times (1 - \beta))$
 - 7: **end if**
 - 8: **end if**
 - 9: **end for**
 - 10: Aplica o algoritmo *Shortest-path* para determinar o caminho de menor custo
 { **onde:**
 $C(p_i)$ é o custo de utilização do caminho óptico p_i ; $ht(p_i)$ representa o custo de p_i calculado pela função de *Holding-Time-Aware* (Equação 1); β determina o peso do *Holding-Time-Aware* para o custo $C(p_i)$; bw representa a banda disponível no caminho p_i ; $(1 - \beta)$ determina o peso da disponibilidade de banda no custo de $C(p_i)$ }.
-

(Fig. 3(b)), USA com 24 nós e 43 enlaces bidirecionais (Fig. 3(d)) e Manhattan Street (grafo grade 5X5) com 25 nós e 40 enlaces bidirecionais.

Os nós utilizados são OXCs com capacidade restrita de agregação e sem capacidade de conversão, possuindo 32 pares (input, output) de portas de agregação. Cada fibra óptica comporta 16 comprimentos de onda e cada comprimento de onda tem uma capacidade de transmissão OC-192 [Tornatore et al. 2008].

A taxa de chegada de chamadas e o tempo de duração das mesmas seguem, respectivamente, a distribuição de Poisson e a distribuição exponencial negativa. A média para o tempo de duração da conexão é de uma unidade de tempo. A carga em Erlangs é definida como a taxa média de chegada \times a duração da chamada \times a banda passante da chamada normalizada pela taxa OC-192. As requisições de conexões são uniformemente distribuídas entre todos os pares de nós com probabilidade 6/19 para OC-3(165Mbps), OC-12 e OC-48; e 1/19 para OC-192(10Gbps) [Zhu et al. 2002].

A métrica considerada nessa avaliação é a taxa de bloqueio de banda (BBR), ou seja, a porcentagem da quantidade de banda bloqueada sobre o total de banda requerida de todas as conexões, durante todo o período de simulação. Todos os gráficos apresentados nesta seção, exceto os de balanceamento da taxa de bloqueio, mostram valores médios obtidos a partir da execução de 10 simulações, utiliza-se intervalo de confiança com nível de confiança 95%, derivados através do método da replicação independente. A reta horizontal presente nos gráficos de balanceamento de bloqueio indica a BBR média apresentada pelo algoritmo HTA para o cenário indicado.

Os resultados obtidos para a topologia NSF são apresentados na Figura 4. Os va-

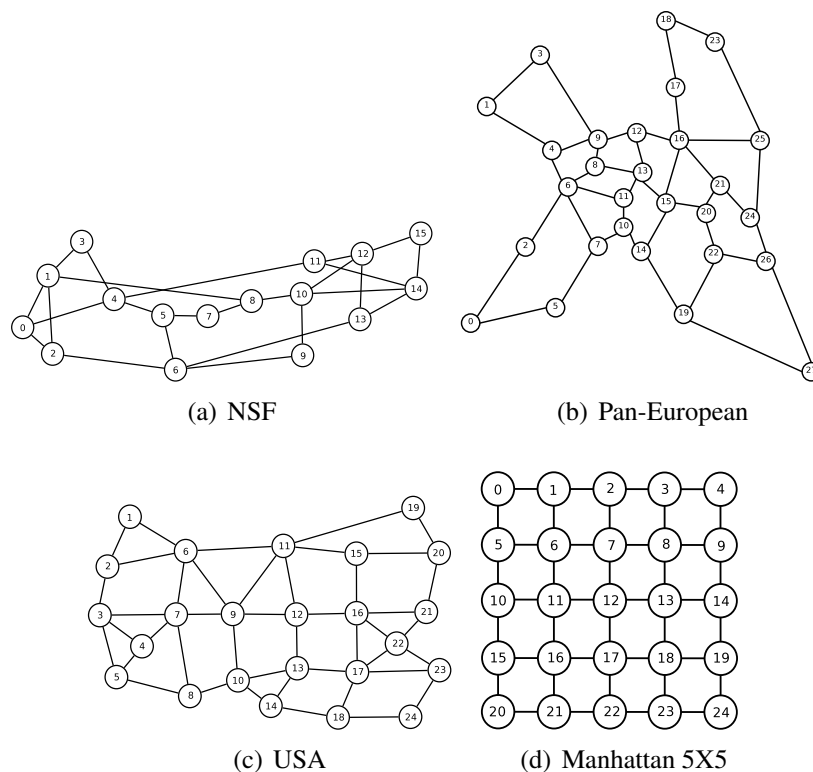


Figura 3. Topologias

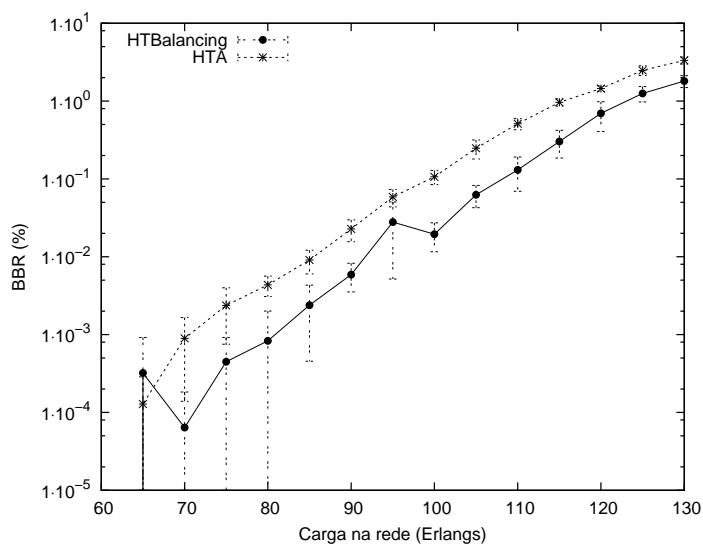


Figura 4. BBR do HTBalancing e HTA *vs* a carga (NSF)

lores de BBR produzidos pelo algoritmo HTBalancing são consideravelmente menores do que os valores obtidos pela utilização do algoritmo HTA. Esta diferença é aproximadamente uma ordem de grandeza para a maioria das cargas. O BBR produzido pelo algoritmo HTBalancing para carga de 70 Erlangs é 93% menor do que o valor gerado pela abordagem HTA. Para cargas elevadas, a diferença entre os valores de BBR diminui,

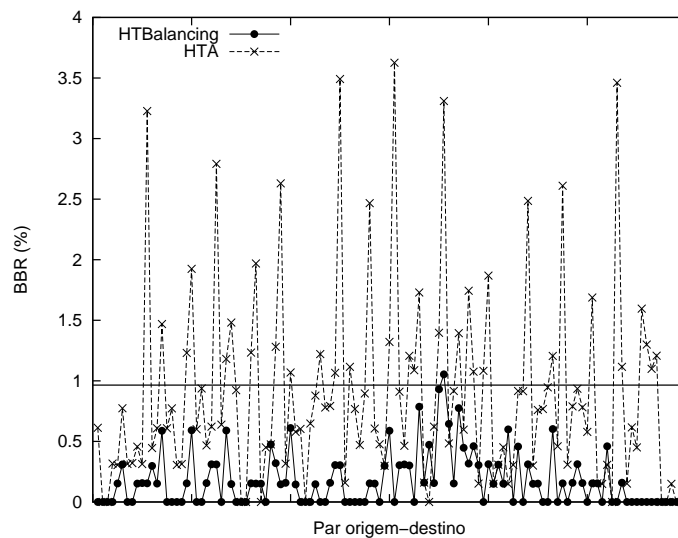


Figura 5. BBR média de cada par origem-destino para uma carga de 115 Erlangs (NSF)

entretando, os valores obtidos pelo algoritmo HTBalancing são ainda 46% menores para cargas de 130 Erlangs. O baixo grau de conectividade de alguns caminhos da topologia NSF, por exemplo, o caminho composto pelos OXCs 4, 5, 7, 8 e 10, levam a criação de gargalos mesmo para as cargas mais baixas. A ideia principal do algoritmo HTA é utilizar o menor número possível de recursos para as novas solicitações e minimizar a quantidade de tempo de vida adicional dos caminhos ópticos já estabelecidos em decorrência de novas solicitações, ou seja, esta abordagem tende a concentrar o tráfego em poucos caminhos ópticos. Portanto, utilizar o algoritmo HTA em topologias semelhantes a topologia NSF pode levar a uma degradação severa no desempenho da rede, o que fica evidenciado pelos valores de BBR produzidos pelo algoritmo HTA. Por outro lado, a abordagem HTBalancing tenta distribuir a carga para prevenir gargalos e, conseqüentemente, apresenta taxas de bloqueio menores que o algoritmo HTA.

A Figura 5 apresenta a distribuição da BBR de cada par origem-destino para a carga de 115 Erlangs para a topologia NSF. O algoritmo HTA propiciou a alguns dos pares taxas até 3.7 vezes maiores do que a sua BBR média de 0.96% e até 12 vezes maiores do que a BBR média de 0.30% do algoritmo HTFBalancing. Estes valores demonstram a incapacidade do algoritmo HTA em distribuir o bloqueio de chamadas de forma equitativa entre todos os pares origem-destino, o que é de grande importância para a eficiência da rede.

Na Figura 6, apresenta-se os resultados obtidos para a topologia Pan-European. Para a carga de 70 Erlangs, o BBR produzido pelo algoritmo HTBalancing é, aproximadamente, 2 ordens de grandeza menor do que o valor produzido pelo algoritmo HTA. À medida que a carga na rede aumenta, a diferença entre o algoritmo HTBalancing e algoritmo HTA diminui, mas mesmo assim, o algoritmo HTBalancing produz valores de BBR menores do que os valores produzidos pelo algoritmo HTA. Para carga de 130 Erlangs o valor de BBR obtido pelo HTBalancing é 29% menor do que o valor obtido pelo

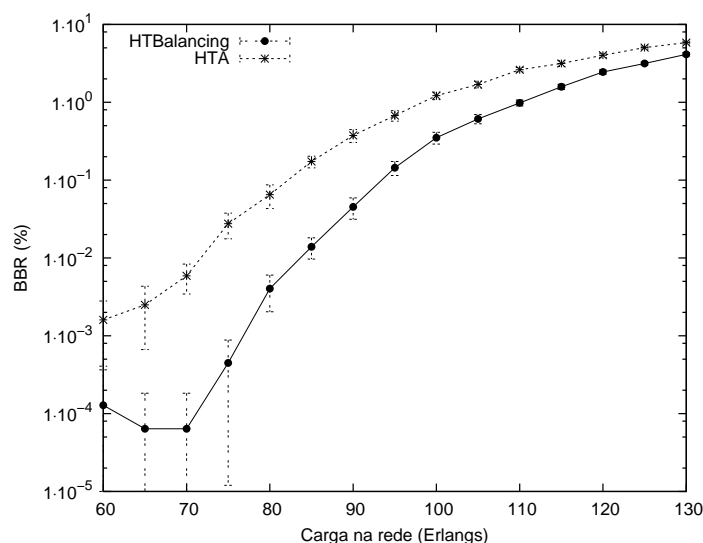


Figura 6. BBR do HTBalancing e HTA vs a carga (Pan-European)

algoritmo HTA. Os valores obtidos neste experimento são influenciados pela topologia Pan-European que apresenta vários ciclos, como por exemplo o ciclo 6, 2, 0, 5, 7, 6, os quais tendem a dificultar a obtenção de rotas alternativas pelo algoritmo HTBalancing.

A Figura 7 apresenta a distribuição da BBR de cada par origem-destino para a carga de 115 Erlangs para a rede Pan-European. Para alguns dos pares, o algoritmo HTA produziu taxas até 3.9 vezes maiores do que a sua BBR média de 3.15% e até 7.8 vezes maiores do que a BBR média de 1.58% do algoritmo HTFBalancing. Estes valores demonstram que, apesar das características da topologia, o algoritmo HTBalancing apresenta maior capacidade para balancear o tráfego na rede.

A Figura 8 descreve os resultados para a rede USA. Neste cenário, as diferenças entre os valores de BBR dos algoritmos HTBalancing e HTA são ainda mais acentuadas. Como pode ser observado no gráfico, para cargas menores que 100 Erlangs, a abordagem HTBalancing é livre de bloqueio de solicitações, enquanto que o algoritmo HTA chega a bloquear 756 das demandas solicitadas. Para carga de 100 Erlangs, o BBR produzido pelo algoritmo HTBalancing é 3 ordens de grandeza menor que os valores produzidos pelo algoritmo HTA. A menor diferença entre os dois algoritmos comparados é para a carga de 150 Erlangs, quando a abordagem proposta, neste trabalho, apresenta valores de BBR 94% menores que os valores obtidos pelo algoritmo HTA. Esta enorme diferença é fortemente influenciada pelas características da topologia USA, a qual apresenta um alto grau de conexidade entre os nós provendo, assim, vários caminhos alternativos na rede, o que evita a criação de gargalos. Além disso, o algoritmo HTBalancing concentra-se em utilizar os benefícios do conhecimento das informações da duração das demandas para realizar agregação ao mesmo tempo que distribui o tráfego para balancear a carga na rede, reduzindo as chances de criação de gargalos. Os excelentes resultados obtidos são, portanto, consequência direta da junção das características da topologia USA e das propriedades do algoritmo HTBalancing para prevenir gargalos, o que leva a redução da taxas de bloqueio, requisito este altamente almejado pelas aplicações. Este desempenho

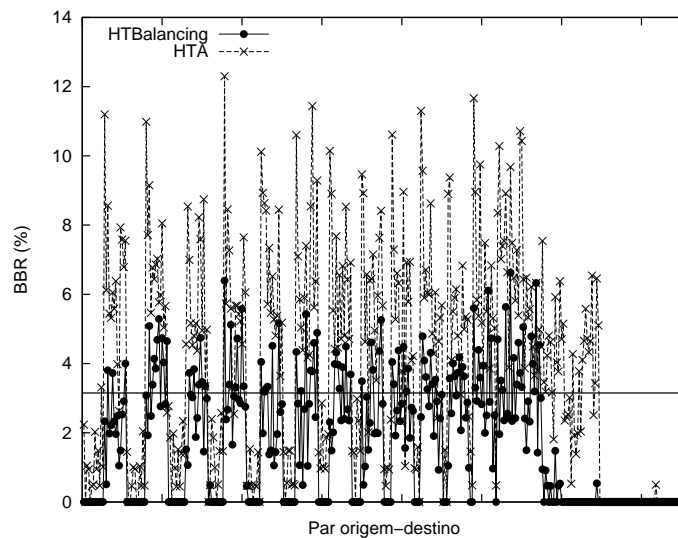


Figura 7. BBR média de cada par origem-destino para uma carga de 115 Erlangs (Pan-European)

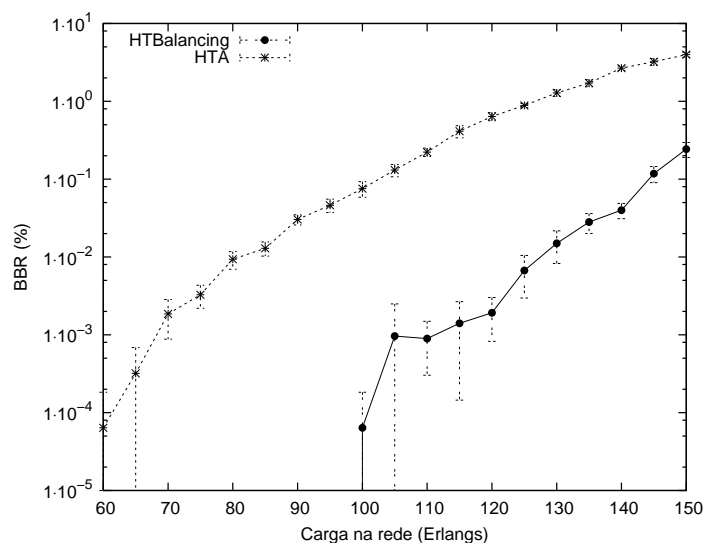


Figura 8. BBR do HTBalancing e HTA vs a carga (USA)

ressalta a importância do HTBalancing para a agregação dinâmica de tráfego.

Na Figura 9, apresenta-se a distribuição da BBR de cada par origem-destino para carga de 150 Erlangs para a topologia USA. Para esta topologia, em alguns pares a taxa de bloqueio do algoritmo HTA é até 3.1 vezes maior do que a sua BBR média de 3.97% e até 51.29 vezes maiores do que a BBR média de 0.24% do algoritmo HTFBalancing. Na topologia USA, devido ao seu maior nível de conexidade, o algoritmo HTBalancing foi capaz de evitar a criação de gargalos na rede, o mesmo não ocorreu com o algoritmo HTA, que não foi capaz de prover recursos de forma justa.

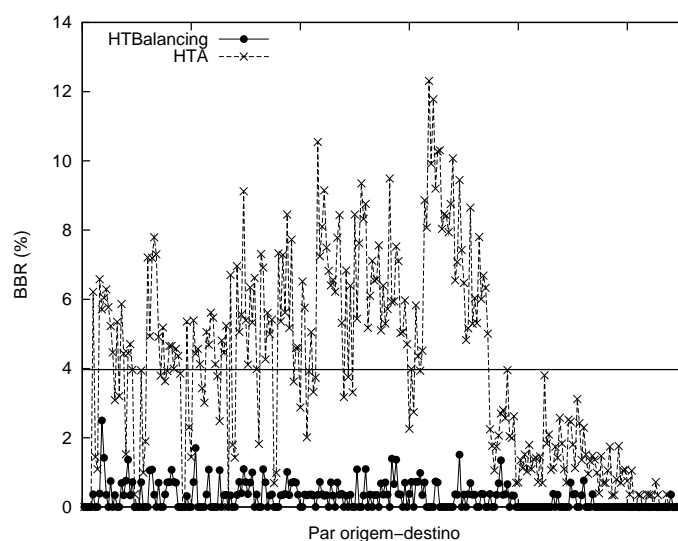


Figura 9. BBR média de cada par origem-destino para uma carga de 150 Erlangs (USA)

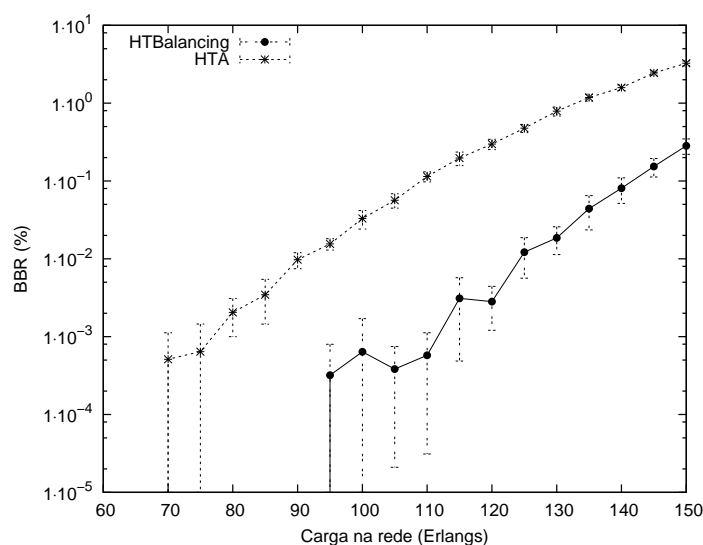


Figura 10. BBR do HTBalancing e HTA *vs* a carga (Manhattan)

A Figura 10 mostra os resultados para a rede Manhattan Street. Neste cenário, para cargas menores que 95 Erlangs o algoritmo HTBalancing é capaz de atender todas as demandas solicitadas, ao passo que o algoritmo HTA chega a rejeitar até 156 das demandas solicitadas. Para carga de 95 Erlangs, o algoritmo HTBalancing produz BBR de 1.5 ordens de grandeza menor que os valores produzidos pela abordagem HTA. Embora a diferença da BBR gerada pelos algoritmos comparados diminua com o aumento da carga na rede, o algoritmo HTA apresenta, ainda, uma BBR 91% maior do que o valor de BBR produzido pelo algoritmo HTBalancing para a carga de 150 Erlangs. Os resultados apresentados pelo algoritmo HTBalancing são impactados pela simetria da topologia

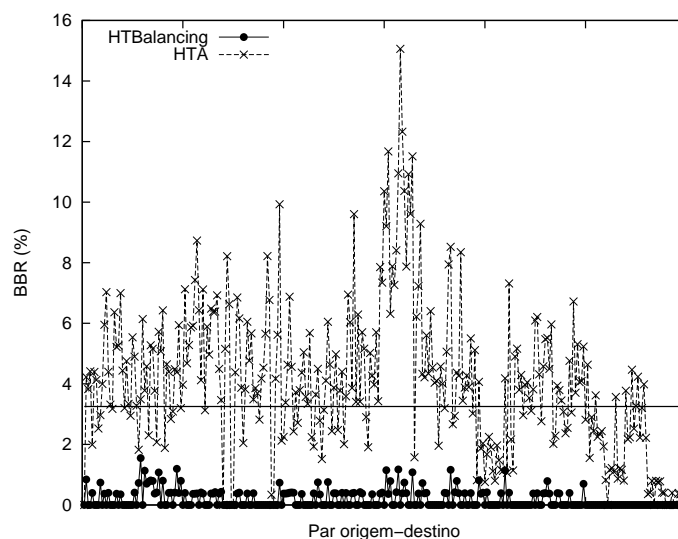


Figura 11. BBR média de cada par origem-destino para uma carga de 150 Erlangs (Manhattan)

Manhattan Street, uma vez que esta topologia possibilita a criação de vários caminhos concorrentes para os pares de comunicação da rede.

Na Figura 11, apresenta-se a distribuição da BBR de cada par origem-destino para carga de 150 Erlangs para a topologia Manhattan Street. Na topologia Manhattan, em alguns pares a taxa de bloqueio do algoritmo HTA é até 4.6 vezes maior do que a sua BBR média de 3.25% e até 53.7 vezes maiores do que a BBR média de 0.28% do algoritmo HTFBalancing. A topologia Manhattan apresenta, também, um alto nível de conectividade e, desta forma, o algoritmo HTBalancing apresenta maior eficiência do que o algoritmo HTA para evitar a criação de gargalos na rede. Observa-se, ainda, que a BBR apresentada pelo algoritmo HTBalancing é distribuída entre os pares de forma mais uniforme do que no algoritmo HTA, mostrando, portanto, que o algoritmo HTBalancing é um mecanismo mais justo.

5. Conclusão

A grande disponibilidade de banda proporcionada pela tecnologia WDM levou ao aumento das expectativas dos usuários com relação a qualidade dos serviços de rede fornecidos. Estas necessidades de QoS são explicitamente especificadas em contratos SLAs que têm surgido como uma nova tendência. Os SLAs possibilitam maior conhecimento de informações a respeito das solicitações das aplicações, tais como o tempo de duração (*holding-time*) das conexões, as quais podem ser utilizadas para gerenciar e controlar a utilização dos recursos ópticos.

Neste artigo, apresentou-se o algoritmo HTBalancing para agregação dinâmica de tráfego que utiliza eficientemente os recursos da rede. A principal contribuição deste trabalho está na utilização do conhecimento da duração das chamadas e da disponibilidade de recursos na rede, visando aproveitar ao máximo os recursos já estabelecidos, enquanto distribui o tráfego de forma balanceada evitando, assim, a criação de gargalos e o bloqueio

de solicitações.

Os resultados foram obtidos em cenários realistas a partir de simulações utilizando as topologias NSF, Pan-European, USA e Manhattan Street, clássicas em estudos de redes ópticas, compostas por OXCs com capacidade restrita de agregação e sem conversão de comprimento de onda. Nos três cenários, o algoritmo proposto apresentou resultados significativamente melhores. Ademais, os resultados obtidos para as topologia USA e Manhattan Street mostram que o algoritmo HTBalancing é um forte candidato para a solução do problema de agregação de tráfego com QoS nas futuras redes ópticas metropolitanas.

Pretende-se, como trabalho futuro, investigar como os parâmetros da heurística afetam o desempenho do sistema.

Agradecimentos

Este trabalho foi parcialmente financiado pelo CNPq sob o processo número 141246/2009-0 e pela FAPESP.

Referências

- Cavdar, C., Buzluca, F., Tornatore, M., and Mukherjee, B. (2009). Dynamic scheduling of survivable connections with delay tolerance in wdm networks. In *INFOCOM Workshops 2009, IEEE*, pages 1–6.
- Doumith, E. and Gagnaire, M. (2007). Impact of traffic predictability on wdm exc/oxc network performance. *Selected Areas in Communications, IEEE Journal on*, 25(5):895–904.
- Liu, S. S. and Chen, L. F. (2007). Deployment of carrier-grade bandwidth-on-demand services over optical transport networks: A verizon experience. In *National Fiber Optic Engineers Conference*, page NThC3. Optical Society of America.
- Munir, A., Tanwir, S., and Zaidi, S. (2009). Holding time aware dynamic bandwidth allocation algorithm for emerging bandwidth on demand multicast applications. pages 16–21.
- Systems, C. (2008). Approaching the zettabyte era. In *Cisco Withe Paper c11-481374*, volume c11-481374.
- Tornatore, M., Baruffaldi, A., Zhu, H., Mukherjee, B., and Pattavina, A. (2007a). Dynamic traffic grooming of subwavelength connections with known duration. In *Optical Fiber Communication and the National Fiber Optic Engineers Conference, 2007. OFC/NFOEC 2007. Conference on*, pages 1–3.
- Tornatore, M., Baruffaldi, A., Zhu, H., Mukherjee, B., and Pattavina, A. (2007b). Exploiting connection-holding time for an efficient dynamic traffic grooming. In *Design and Reliable Communication Networks, 2007. DRCN 2007. 6th International Workshop on*, pages 1–8.
- Tornatore, M., Baruffaldi, A., Zhu, H., Mukherjee, B., and Pattavina, A. (2008). Holding-time-aware dynamic traffic grooming. *Selected Areas in Communications, IEEE Journal on*, 26(3):28–35.
- Zhu, H., Zang, H., Zhu, K., and Mukherjee, B. (2002). Dynamic traffic grooming in wdm mesh networks using a novel graph model. volume 3, pages 2681 – 2685 vol.3.