

# Roteamento multicaminho para provisão eficiente de recursos interdomínio com qualidade de serviço em redes ópticas WDM

André C. Drummond<sup>1</sup>, Nelson L. S. da Fonseca<sup>1</sup>, Xiaomin Chen, Admela Jukan<sup>2</sup>

<sup>1</sup>Instituto de Computação – UNICAMP, Campinas, SP

<sup>2</sup>Technische Universität Carolo-Wilhelmina zu Braunschweig

{andred, nfonseca}@ic.unicamp.br, {chen, jukan}@ida.ing.tu-bs.de

**Abstract.** *The present paper studies the dynamic traffic grooming problem in a scenario of multi-domain optical networks with high speed traffic demands and strict QoS requirements required by real-time and bulk-data transfer applications. Tailored solutions are proposed to each type of application which are capable to work efficiently in both advanced static topology reservation and in shared topology scenario. The proposed algorithms delivered a better overall traffic balance and also mitigated the impact of the intra-domain traffic upon the inter-domain one along the crossed domains.*

**Resumo.** *Este artigo estuda o problema de agregação dinâmica de tráfego em um cenário de redes ópticas multidomínio com demandas de tráfego de alta capacidade e requisitos de QoS. Aplicações de tempo-real e de transferência de dados podem demandar diferentes requisitos de QoS. Nesse estudo, são propostas soluções específicas para cada cenário de tráfego capazes de atuar tanto em um ambiente multidomínio de alocação antecipada de recursos, quanto em um cenário no qual o tráfego interdomínio é forçado a disputar os recursos de cada domínio com o seu tráfego local. Os algoritmos propostos foram capazes de prover um melhor balanceamento do tráfego e também a mitigação do impacto do tráfego interdomínio sobre o tráfego local em cada domínio atravessado.*

## 1. Introdução

Recentemente, a crescente demanda das aplicações científicas e comerciais de banda larga pelo estabelecimento de caminhos multidomínio com garantias de Qualidade de Serviço (QoS) tem capitalizado um grande esforço em pesquisa e padronização para o desenvolvimento de arcabouços de roteamento interdomínio capazes de prover grande quantidade de banda passante para atender aplicações sensíveis ao atraso. Apesar de o roteamento interdomínio na Internet ser um tema bastante investigado, o serviço do tipo melhor esforço torna as redes IP inadequadas para as aplicações que exigem garantias de QoS fim-a-fim. Isso leva a novos paradigmas de roteamento interdomínio, especialmente em redes ópticas WDM, por ser a tecnologia de escolha que proverá a quantidade de banda passante necessária para as demandas futuras na Internet.

Em redes ópticas WDM o transporte dos dados é realizado através de circuitos ou caminhos ópticos que constituem a topologia virtual de um domínio. Topologias virtuais agregadas tem sido utilizada nas redes com comutação por circuito para facilitar o roteamento interdomínio e tem sido, também, propostas para simplificar o estabelecimento (*peering*) e aprovisionamento de conexões interdomínio [Chamania and Jukan 2009]. De

forma a eliminar a influência do tráfego intradomínio nas conexões interdomínio, foi introduzido o mecanismo de reserva antecipada, que aloca os túneis entre os nós de borda do domínio em questão [Chamania et al. 2009]. Esse túneis pré-reservados conectando os nós de borda do domínio são representados como uma topologia virtual e são anunciados para a computação de caminhos interdomínio. Banda passante é reservada nos enlaces exclusivamente para o tráfego interdomínio.

A topologia virtual com reserva antecipada traz estabilidade para o provisionamento de conexões interdomínio, pois elimina as atualizações constantes impostas pelo tráfego intradomínio [Chamania and Jukan 2009], todavia, esta técnica pode ser ineficiente na prática. A super-alocação de banda passante pode levar a uma baixa utilização dos túneis e ao aumento do bloqueio do tráfego intradomínio, enquanto que a sub-alocação pode resultar em altos valores de bloqueio para o tráfego interdomínio. A topologia virtual pode ser calculada alternativamente em tempo-real mediante a chegada de uma chamada. O conhecido mecanismo de computação inversa recursiva de rota (BRPC) [Vasseur et al. 2009] para o estabelecimento de rotas interdomínio pode iniciar a computação de uma árvore virtual a partir do nó destino para uma chamada. A árvore virtual é enviada para o domínio anterior que, por sua vez, calcula os caminhos entre seus nós de borda de forma a estender a árvore virtual. A construção da topologia virtual termina quando ela alcança o nó de origem da chamada. Esse tipo de arquitetura é capaz de utilizar melhor os recursos da rede, todavia o método de construção da topologia virtual por requisição introduz complexidade para o arcabouço de provisionamento de conexões interdomínio.

Dada a sua simplicidade, o roteamento de caminho único é comumente utilizado nos mecanismos de roteamento interdomínio, o qual encontra o melhor caminho seja na topologia virtual com reserva antecipada de túneis ou na topologia virtual construída dinamicamente. Por outro lado, o roteamento de caminho único pode causar congestionamento nos enlaces gargalo da rede, além de enfrentar os novos desafios impostos pela aplicações emergentes com grande demanda por banda passante [Deelman et al. 2005]. Exemplos típicos são as aplicações de e-Ciência, que podem requerer o transporte de grandes quantidades de dados entre hosts geograficamente distantes ou solicitar uma conexão com garantias de QoS entre dois hospedeiros para aplicações em tempo-real. Pode-se citar o caso da aplicação Montage [Deelman et al. 2005], que é uma aplicação de processamento de imagens que compõe mosaicos de imagens do céu para pesquisa astronômica. O tamanho do seu fluxo de trabalho (*workflow*) depende da área do céu a ser processada para gerar o mosaico. Por exemplo, para uma área de 1 grau quadrado do céu, um fluxo de trabalho com 232 tarefas é executado. Para uma imagem de 10 graus quadrados, um fluxo de trabalho com 20.652 tarefas é executado, tratando com uma massa de dados próxima a 100 Gigabytes.

O roteamento de caminho único tradicional é incapaz de servir esse tipo de aplicação, mesmo em redes ópticas WDM que possuem uma enorme capacidade de banda passante agregada provida pelos múltiplos comprimento de onda, como já estudado no trabalho [Chen et al. 2009] dos mesmos autores deste artigo. Aplicações triviais também se beneficiam do mecanismo de roteamento multicaminho pela sua capacidade de distribuição do tráfego, provendo um melhor balanceamento na utilização de recursos da rede.

Dado que o processo de atualização das capacidade da rede em múltiplos domínios para lidar com demandas de alta capacidade é um procedimento extremamente custoso, o projeto de novos mecanismos de roteamento interdomínio é necessário. Os resultados apresentados em [Chen et al. 2009] motivam a utilização do roteamento multicaminho em redes ópticas WDM.

Neste artigo, propõe-se algoritmos para o estabelecimento (*peering*) e aprovisionamento de conexões interdomínio em redes ópticas WDM com topologias virtuais estáticas. A solução proposta engloba tanto o cenário de topologias virtuais com reserva antecipada para conexões interdomínio, quanto o cenário sem reserva que considera a interação entre o tráfego inter e intradomínio. Para amenizar a complexidade de troca de mensagens imposta pelo cenário sem reserva antecipada, utiliza-se um mecanismo que troca apenas informações de conectividade entre os domínios. As informações referentes a outras métricas de QoS como a quantidade de banda passante disponível são obtidas e atualizadas mediante a chegada de uma requisição. O mecanismo proposto distingue-se dos mecanismos de roteamento interdomínio existentes ao permitir o roteamento multicaminho no aprovisionamento de serviços interdomínio com o intuito de superar a ineficiência e falta de escalabilidade da topologia virtual estática. O tráfego pode ser separado em múltiplos caminhos para evitar a superalocação realizada para se lidar com cargas de pico ou com o congestionamento. O mecanismo contribui também para manter a escalabilidade da arquitetura de roteamento interdomínio dado que ele pode eliminar as atualizações frequentes causadas pela insuficiência de recursos nas topologias virtuais pré-calculadas.

## 2. Trabalhos relacionados

A arquitetura de elementos de computação de rota (PCE) foi proposta para as redes orientadas à conexões, e.g. GMPLS e *Carrier Ethernet*, para computar caminhos multidomínio e multi-camada. Os caminhos que atravessam múltiplos domínios são definidos através da cooperação dos PCEs de cada domínio. O algoritmo BRPC é um mecanismo bastante utilizado para a computação de caminhos multidomínio que inicia o estabelecimento da rota a partir do nó de destino e calcula caminhos ótimos entre os nós de egresso e ingresso do domínio. Uma topologia virtual composta pelos caminhos calculados no domínio é enviada para o PCE do domínio anterior que, por sua vez, repete o mesmo procedimento até que o domínio destino seja atingido e o nó de origem alcançado. Esse, por sua vez, decide pelo melhor caminho a ser tomado baseado na topologia virtual agregada fornecida [Vasseur et al. 2009]. A topologia virtual pré-definida com reserva de banda passante foi introduzida com o intuito de reduzir a sinalização interdomínio na Internet em [Pan et al. 2000, Sofia et al. 2003], no qual a banda passante é reservada entre os nós de borda dos domínios para o tráfego interdomínio.

A utilização do roteamento multicaminho para conexões multidomínio tem sido estudada para a provisão de caminhos diversos no nível de sistemas autônomos (AS) [Xu and Rexford 2006], enquanto que a solução para a camada 2 são em maior parte limitadas a um único domínio. Nesse cenário, os benefícios do roteamento multicaminho, tais como balanceamento de carga e aumento da vazão da rede, já foram demonstrados na literatura, em especial, foram evidenciadas as vantagens do roteamento multicaminho sobre o roteamento de caminho único, considerando-se o tempo de estabelecimento das conexões [I.Cidon et al. 1999]. Em [Gao et al. 2002], é proposto um mecanismo de ba-

lançamento baseado no retardo dos caminho em redes MPLS, utilizando o roteamento multicaminho.

O artigo [Huang et al. 2008] é um dos poucos artigos que abordam a agregação de tráfego interdomínio em redes ópticas WDM. Nele, foi proposto um algoritmo que atua em conjunto com os algoritmos RWA dos domínios. São propostas duas políticas de gerenciamento intradomínio responsáveis pelo roteamento e alocação dos recursos. Elas se diferenciam na forma da escolha dos comprimentos de onda a serem utilizados na transmissão. Esta abordagem é interessante para o roteamento dinâmico das chamadas, que tipicamente trabalha com pouco conhecimento sobre a topologia da rede. No entanto, ela não considera o princípio fundamental da independência administrativa dos domínios.

Este artigo diferencia-se da literatura por apresentar o roteamento multicaminho como um mecanismo capaz de viabilizar o uso de topologias virtuais agregadas estáticas no provimento de recursos multidomínio para aplicações com demandas de alta capacidade e requisitos de QoS .

### 3. Algoritmos para o roteamento multicaminho entre domínios

Nesta seção, são apresentados 2 algoritmos para o roteamento multicaminho interdomínio, um projetado para aplicações em tempo-real e o outro para aplicações de transferência de dados. A notação utilizada na definição dos algoritmos é apresentada a seguir:  $R(s, d)$  é uma requisição de conexão do nó  $s$  para o nó  $d$ ;  $B_R$  é a demanda de banda passante da requisição  $R$ ;  $D_l$  é o prazo para processamento dos dados da requisição  $R$ , para aplicação em tempo-real;  $D_R$  é a restrição de retardo para a requisição  $R$ , sendo o prazo para transferência dos dados para aplicações de transferência de dados ou o retardo fim-a-fim máximo para aplicações em tempo-real;  $C$  é a quantidade de dados a ser transmitida da requisição  $R$ , para aplicações de transferência de dados;  $p_i$  é o  $i$ -ésimo caminho  $\mathcal{P}$ ;  $f_{p_i}$  é a banda passante disponível no caminho  $p_i$ ;  $b_{p_i}$  é a capacidade de banda passante do caminho  $p_i$ ;  $t_{p_i}$  é a banda passante reservada no caminho  $p_i$ ;  $PD_{p_i}$  é o retardo fim-a-fim do caminho  $p_i$ , no qual  $PD = \text{TempoTotalPropagação} + (n - 1) \times \text{TempoConversãoOEO}$  e  $n =$  número de saltos;  $M_d$  é a capacidade da memória no nó de destino  $d$ ;  $\beta$  é a máxima utilização da banda passante disponível nos caminhos ópticos.

#### 3.1. Retardo diferencial

O retardo diferencial é causado pela utilização de múltiplos caminhos que possuam diferentes retardos de propagação, o que pode acarretar na desordenação dos pacotes no nó de destino, o que requer memória para o re-sequenciamento dos pacotes. O retardo diferencial entre dois caminhos  $p_i$  e  $p_j$  pode ser definido como  $dd(p_i, p_j) = |d_{p_i} - d_{p_j}|$  [Ahuja et al. 2004].

Assuma que  $k$  caminhos são utilizados por uma conexão qualquer e que o caminho com o maior retardo dentre todos os caminhos é  $\tilde{p}$ . Dado que o tráfego no caminho  $\mathcal{P}$  é denotado por  $t_p$ , e que o retardo do  $i$ -ésimo caminho  $p_i$  é  $d_{p_i}$ , então, a quantidade de memória necessária no nó de destino  $d$  é dada por:

$$M = \sum_{p_i, i=1,2,\dots,k} t_{p_i} \cdot (d_{\tilde{p}} - d_{p_i}) \quad (1)$$

O retardo diferencial afeta principalmente aplicações em tempo-real que exigem que os seus pacotes sejam entregues em ordem no destino [Mao et al. 2006], dado sua

sensibilidade em relação ao atraso, o que leva a elaboração de algoritmos especiais para essa classe de aplicação.

### 3.2. Heurísticas multicaminho

Nesta seção, são apresentados os algoritmos propostos para o roteamento multicaminho para aplicações em tempo-real e de transferência de dados, respectivamente. Em ambos os algoritmos, é definido um parâmetro  $\beta$  que expressa a quantidade de banda passante que poderá ser utilizada em cada caminho óptico considerado, o que visa alcançar dois objetivos: melhorar o balanceamento de carga na rede; e promover o espaço residual nos caminhos ópticos utilizados de forma a acomodar dados de outros fluxos.

#### Roteamento multicaminho para aplicações multimídia de tempo-real

O primeiro algoritmo proposto é denominado MRRTM e é dedicado para aplicações em tempo-real que demandam memória de re-sequenciamento no nó receptor para efetuar a reordenação dos pacotes e absorver o *jitter* gerado. À cada pacote está associado um prazo de processamento  $D_l$ , no qual o pacote deve ser extraído da memória para ser encaminhado [Mao et al. 2006]. Para tais aplicações, um pacote ficará na memória de re-sequenciamento por no máximo  $D_l$  unidades de tempo. A perda de pacotes é causada pela falta de memória disponível no nó receptor. O roteamento multicaminho é capaz de atender a necessidade de banda passante das aplicações de alto desempenho, porém o retardo diferencial torna o problema muito mais desafiador. O algoritmo proposto tenta servir a conexão utilizando apenas a quantidade de memória existente no receptor.

A heurística capitaliza na suposição sobre o valor de retardo do caminho  $\tilde{p}$  (maior retardo), e então determina o conjunto solução correspondente. Na heurística proposta, não há uma busca exaustiva das possíveis soluções, pelo contrário, o algoritmo para ao encontrar a primeira solução que possa satisfazer a requisição. O algoritmo MRRTM é mostrado na Figura 1 e seu funcionamento é composto de dois passos. Dada uma requisição de conexão interdomínio  $R(s, d, B_r, D_l, D_R)$ , define-se, no passo 1, o conjunto de caminhos candidatos, conforme descrito a seguir. Calcula-se os  $K$  menores caminhos disjuntos de  $s$  para  $d$  (linha 2), e seleciona-se os  $M$  caminhos que possuam retardo fim-a-fim menores do que  $D_R$ , formando o conjunto  $\mathcal{P} = \{p_1, p_2, \dots, p_M\}$  (linha 3). Ordena-se os elementos de  $\mathcal{P}$  em ordem decrescente da quantidade de banda passante disponível (linha 4). Avalia-se a possibilidade de uma solução de caminho único considerando apenas o primeiro caminho do conjunto  $\mathcal{P}$  (linha 5); o parâmetro  $\beta$  é utilizado como limitador da banda passante disponível. Caso seja possível a alocação, então, a chamada é aceita e o algoritmo termina, caso contrário executa-se o passo 2 do algoritmo.

No passo 2, o algoritmo faz a seleção dos caminhos de interesse e reserva banda passante. O laço externo do algoritmo (linha 9) inicia com o caminho de maior capacidade disponível do conjunto  $\mathcal{P}$  e assume que ele é, também, o caminho de maior retardo fim-a-fim  $\tilde{p}$ . O primeiro laço interno (linha 13) seleciona todos os caminhos de  $\mathcal{P}$  com retardo inferior a  $\tilde{p}$  e constrói o conjunto  $\mathcal{P}'$ . No segundo laço interno (linha 18), calcula-se a soma das capacidades disponíveis dos caminhos selecionados em  $\mathcal{P}'$  ponderada pelo parâmetro  $\beta$  sendo que, antes de se considerar um caminho  $\mathcal{P}$  para a soma, o retardo diferencial gerado entre ele e o caminho  $\tilde{p}$  é comparado ao retardo máximo de tempo de



## Algoritmo MRRTM

---

**Entrada:** Topologia virtual  $G(V, E)$  e  $\beta$ .  
**Entrada:** Requisição  $R(s, d, B_R, D_l, D_R)$ .  
**Saída:** Solução multicaminho para  $R$ .

- 1: **Passo 1:** *Preprocessamento*
- 2: Calcula  $K$  caminhos através do algoritmo de  $K$  menores caminhos.
- 3: Seleciona os caminhos com retardo menor do que  $D_R$  e insere no conjunto  $\mathcal{P}$ .
- 4: Ordena  $\mathcal{P}$  em ordem decrescente da quantidade de banda passante disponível  $\mathcal{P} = p_1, p_2 \dots p_M, M \leq K$ .
- 5: **Se**  $f_{p_1} \cdot \beta \geq B_R$  **então**
- 6:     Gera a solução  $p_1$  e termina.
- 7: **Fim Se**
- 8: **Passo 2:** *Seleção de caminhos e reserva de banda passante*
- 9: **Para**  $i = 1, \dots, M, p_i \in \mathcal{P}$  **faça**
- 10:      $Resv = 0$
- 11:      $\mathcal{P}' = \emptyset$
- 12:      $\tilde{p} = p_i$
- 13:     **Para**  $j = 1, \dots, M, p_j \in \mathcal{P}$  **faça**
- 14:         **Se**  $PD_{p_j} \leq PD_{\tilde{p}}$  **então**
- 15:              $\mathcal{P}' \leftarrow p_j$
- 16:         **Fim Se**
- 17:     **Fim Para**
- 18:     **Para**  $p_k \in \mathcal{P}'$  **faça**
- 19:         **Se**  $PD_{\tilde{p}} - PD_{p_k} \leq D_l$  **então**
- 20:              $Resv = Resv + f_{p_k} \cdot \beta$
- 21:             **Se**  $Resv \geq B_R$  **então**
- 22:                 interrompe laço;
- 23:             **Fim Se**
- 24:         **Fim Para**
- 25:     **Fim Para**
- 26:     **Se**  $Resv \geq B_R$  **então**
- 27:         **Para** cada caminho  $p_k$  dentre os caminhos selecionados **faça**
- 28:              $t_{p_k} = B_R \cdot \frac{f_{p_k}}{Resv}$
- 29:             **Fim Para**
- 30:             **Se**  $\sum_{p_k} t_{p_k} \cdot (PD_{\tilde{p}} - PD_{p_k}) \leq M_d$  **então**
- 31:                 Gera a solução e termina..
- 32:             **Fim Se**
- 33:         **Fim Para**
- 34:     **Fim Para**

Figura 1. Algoritmo MRRTM

re-senquenciamento  $D_l$ , e apenas os caminhos que passarem nessa restrição serão somados. Se a soma atingir (ou ultrapassar) a quantidade de banda passante requerida  $B_r$ , o laço é interrompido, caso contrário, o laço externo será executado novamente e um novo caminho será escolhido como  $\tilde{p}$ . O algoritmo executa até que um grupo de caminhos seja encontrado com capacidade de banda passante agregada maior ou igual a  $B_r$ , caso contrário o algoritmo termina.

Se a quantidade de banda passante tiver sido atingida (linha 26), a banda passante requerida  $B_r$  é dividida, proporcionalmente, entre os caminhos selecionados (linha 28). Finalmente, a quantidade de memória necessária pelo conjunto de caminho selecionados é comparada com a quantidade de memória disponível no de destino  $M_d$  (linha 30). Caso a memória requerida seja menor ou igual a quantidade disponível, o algoritmo termina pois uma solução foi encontrada. Caso contrário, o laço externo é executado novamente e um novo caminho será escolhido como  $\tilde{p}$ . O algoritmo heurístico pára apenas quando uma solução é encontrada ou quando todos os caminhos existentes em  $\mathcal{P}$  forem considerados. No primeiro caso, a requisição  $R(s, d, B_r, D_l, D_R)$  é estabelecida utilizando apenas os caminhos selecionados de  $\mathcal{P}'$ , no segundo a requisição é bloqueada.

## Algoritmo MRDT

---

**Entrada:** Topologia virtual  $G(V, E)$  e  $\beta$ .  
**Entrada:** Requisição  $R(s, d, C, D_R)$ .  
**Saída:** Solução multicaminho para  $R$ .

- 1: **Passo 1:** *Preprocessamento*
- 2: Calcula  $K$  caminhos através do algoritmo de  $K$  menores caminhos.
- 3: Seleciona os caminhos com retardo menor do que  $D_R$  e insere no conjunto  $\mathcal{P}$ .
- 4: Ordena  $\mathcal{P}$  em ordem decrescente da quantidade de banda passante disponível  $\mathcal{P} = p_1, p_2 \dots p_M, M \leq K$ .
- 5: **Passo 2:** *Seleção de caminhos e reserva de banda passante*
- 6: Encontra o caminho com o maior retardo  $\bar{p}$  em  $\mathcal{P}$ .
- 7:  $Resv = 0$
- 8: **Enquanto**  $\mathcal{P} \neq \emptyset$  **faça**
- 9:      $Resv = Resv + f_{p_1} \cdot \beta$
- 10:    **Se**  $(Resv \geq C / (D_R - PD_{\bar{p}}))$  **então**
- 11:     interrompe laço
- 12:    **Fim Se**
- 13:    Atualiza a banda passante disponível dos caminhos em  $\mathcal{P}$ .
- 14:    Remove os caminhos sem capacidade disponível de  $\mathcal{P}$ .
- 15:    Ordena  $\mathcal{P}$  em ordem decrescente da quantidade de banda passante disponível.
- 16: **Fim Enquanto**
- 17: **Se**  $(Resv \geq C / (D_R - PD_{\bar{p}}))$  **então**
- 18:    **Para** cada caminho  $p_k$  dentre os caminhos selecionados **faça**
- 19:      $t_{p_k} = f_{p_k} \cdot \beta$
- 20:    **Fim Para**
- 21: **Fim Se**

Figura 2. Algoritmo MRDT

## Roteamento multicaminho para aplicações de transferência de dados

O segundo algoritmo proposto é denominado MRDT e é voltado para aplicações de transferência de dados. Aplicações de transferência de dados são aplicações que possuem restrição de tempo para a transferência dos dados entre um par de nós na rede. Um exemplo clássico é a transferência de grandes massas de dados gerada pela transmissão de dados entre tarefas sequenciais, ou pela migração de tarefas de uma aplicação de e-Ciência na qual terabytes de dados são demandados por cada execução da aplicação. Assumindo que o intervalo entre duas tarefas de uma aplicação é  $D_R$ , o principal objetivo de um algoritmo de roteamento que pretenda suportar tais aplicações é o de prover uma conexão com retardo fim-a-fim menor ou igual a  $D_R$ , e ao mesmo tempo seja capaz de evitar o congestionamento na rede.

Para evitar os congestionamentos na rede, o parâmetro  $\beta$  pode ser utilizado para forçar a distribuição do tráfego por vários caminhos. A Figura 2 apresenta o algoritmo proposto (MRDT). Este, também, é composto de 2 passo, sendo o primeiro equivalente ao do algoritmo MRRTM que define um conjunto  $\mathcal{P}$  de caminhos válidos e ordenados em ordem decrescente em relação a quantidade de banda passante disponível. No passo 2, executa-se um laço (linha 8) cujas iterações percorrem os caminhos de  $\mathcal{P}$ , acumulando o valor da banda passante disponível ponderada pelo valor do parâmetro  $\beta$ . As iterações terminam quando for atingida uma quantidade de banda passante suficiente para que a requisição possa ser transmitida em um tempo inferior a  $D_R$ . Note que para garantir a restrição de tempo, o valor de retardo do caminho  $\bar{p}$  é considerado (linha 10). A cada iteração, a quantidade de banda passante disponível dos caminhos de  $\mathcal{P}$  é atualizada (linha 13), e os caminhos sem recursos são removidos do conjunto (linha 14). O conjunto é então reordenado, segundo as capacidades disponíveis atualizadas (linha 15). O algoritmo termina quando todos os caminhos são considerados, ou quando a quantidade de banda passante mínima requerida para garantir  $D_R$  é atingida. No primeira caso, a chamada é bloqueada por falta de recursos na rede, no segundo caso, a chamada é aceita e é reservada a quantidade de banda passante necessária para o seu estabelecimento (linha 19).

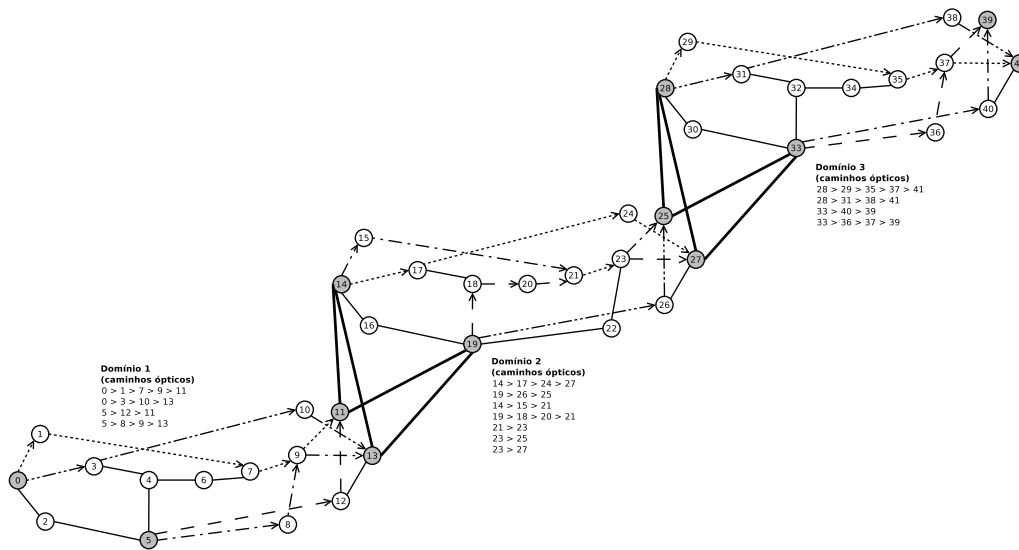


Figura 3. Rede multidomínio linear composta por 3 topologias NSF interligadas.

#### 4. Exemplos numéricos

A rede óptica WDM multidomínio utilizada nas simulações é composta de três domínios NSF, como pode ser visto na Figura 3. A capacidade dos comprimentos de onda nos enlaces interdomínio é de 40Gbps, as capacidades dos comprimentos de onda nos enlaces intradomínio para o primeiro e terceiro domínios é 10Gbps, a capacidade do segundo domínio é de 40Gbps. Desta forma, o segundo domínio opera como um domínio de trânsito e, portanto, suporta uma maior quantidade de tráfego. Cada nó da rede é um OXC com capacidade total de agregação sem conversores de comprimentos de onda. A topologia virtual agregada estática considerada é composta pela união dos caminhos ópticos de cada domínio. Foi criado, também, um caminho óptico para cada enlace interdomínio. Os nós de borda de cada domínio estão preenchidos na cor cinza.

A alocação dos recursos na topologia virtual é avaliada em dois cenários, o cenário de reserva adiantada, no qual a capacidade de banda passante dos caminhos ópticos é reservada exclusivamente para o tráfego interdomínio, e o cenário sem reserva no qual os recursos dos caminhos ópticos podem ser utilizados para o estabelecimento de conexões intradomínio, ou seja, conexões locais de cada domínio. No caso do cenário com reserva, apenas as conexões interdomínio competem pelos recursos dos caminhos ópticos, já no caso do cenário sem reserva, o tráfego interdomínio deve concorrer com o tráfego intradomínio de cada um dos domínios atravessados.

Para qualquer tipo de tráfego, o processo de chegada das requisições de conexão segue a distribuição de *Poisson*. Para o tráfego intradomínio o tempo médio de duração de uma chamada foi de uma unidade de tempo e as granularidades para requisição de banda passante foram escolhidas no conjunto [100, 200, 400, 800 e 1600] Mbps. As chegadas de requisições de conexão estão distribuídas de forma inversamente proporcional a requisição de banda, i.e. a probabilidade de chegada de uma chamada que requisita 200Mbps é 8 vezes maior do que a de uma chamada de 1600Mbps. Para o tráfego interdomínio das aplicações multimídia de tempo real as chamadas possuem um tempo médio de duração de uma unidade de tempo e a demanda de banda passante é escolhidas no conjunto [250, 500, 1000, 2000, 4000, e 8000] Mbps. As chegadas são distribuídas de



forma inversamente proporcional. Por último, para o tráfego interdomínio formado pelas aplicações de transferência de dados, todas chamadas possuem a requisição de transferência de uma massa de dados de 10Gb, variando apenas a requisição de retardo máximo no conjunto [600, 900, 1200, 1500 e 1800] milisegundos.

Os algoritmos de agregação de tráfego multicaminho utilizados nas simulações foram o algoritmo MRRTM (Fig. 1) para o tráfego interdomínio de aplicações multimídia de tempo-real e o algoritmo MRDT (Fig. 2) para o tráfego interdomínio de aplicações de transferência de dados. Para avaliar o mecanismo de roteamento de caminho único, os mesmos algoritmos foram utilizados, porém com a escolha do caminho sendo o melhor caminho encontrado no primeiro passo de cada algoritmo. O algoritmo de roteamento fixo de menor caminho foi utilizado para o estabelecimento do tráfego intradomínio nos cenários de alocação dinâmica de recursos.

As simulações foram realizadas utilizando-se o simulador WDMSim [Drummond 2010]. As métricas utilizadas na avaliação são a probabilidade de bloqueio (BP), a taxa de bloqueio de banda (BBR), e a utilização média dos caminhos ópticos. Os intervalos de confiança com 95% de confiabilidade calculados foram omitidos nos gráficos pois não ultrapassaram 5% dos valores médios calculados.

### Aplicações em tempo-real

Primeiramente, avalia-se o desempenho do uso do roteamento multicaminho em comparação com o algoritmo de roteamento de caminho único, utilizando as métricas BBR e utilização média dos caminhos ópticos, para os cenários com e sem reserva, como mostrado nas Figuras 4 e 5. O valor do parâmetro  $\beta = 1$  foi utilizado para permitir a comparação. Pode-se ver que o uso do roteamento multicaminho é capaz de diminuir o bloqueio e, simultaneamente, aumentar a utilização da rede em ambos os cenários, o que implica no fato de que a utilização do roteamento multicaminho para conexões interdomínio pode aumentar a vazão da rede. Pode-se observar também que a BBR das conexões interdomínio no cenário em que competem com o tráfego intradomínio é muito maior do que a BBR para o caso da topologia virtual com reserva antecipada, o que evidencia o compartilhamento dos recursos entre o tráfego inter e intradomínio. O aumento na utilização dos caminhos ópticos sugeridos pelo algoritmo multicaminho é mitigado no cenário sem reserva pela mesma razão. No entanto, a vantagem de se utilizar o multicaminho no roteamento interdomínio pode ser facilmente constatada, especialmente para valores menores de carga, nos quais, para o cenário com reserva, a BBR chega a ser até uma ordem de grandeza menor do que os valores apresentados pelo algoritmo de caminho único, uma diferença muito significativa. Com o aumento da carga da rede, torna-se mais difícil encontrar caminhos alternativos para rotear o tráfego, portanto o desempenho de ambos os algoritmos tende a se aproximar.

O desempenho do algoritmo de roteamento multicaminho em relação ao balanceamento do tráfego é evidenciado na Figura 6, na qual são mostrados os resultados para o algoritmo considerando diferentes valores para o parâmetro  $\beta$ . Pode-se ver que, no cenário com reserva, o desempenho do algoritmo multicaminho é inferior ao do algoritmo de caminho único (Fig. 5) apenas para  $\beta = 0,7$ , o que significa que mesmo limitando a banda passante alocável pelas requisições interdomínio em mais de 20%, o algoritmo multicaminho é capaz de manter uma vazão maior do que o algoritmo de caminho único. Para o cenário sem reserva, o valor de  $\beta$  precisa ser menor do que 0,9 para que o desempenho

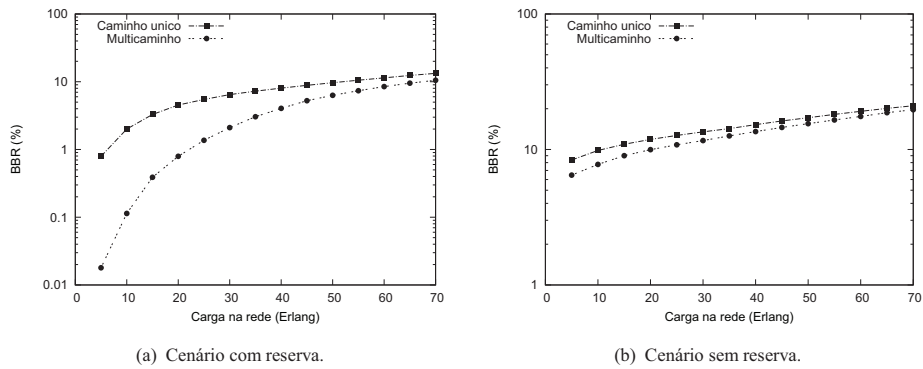


Figura 4. Taxa de bloqueio de banda (BBR).

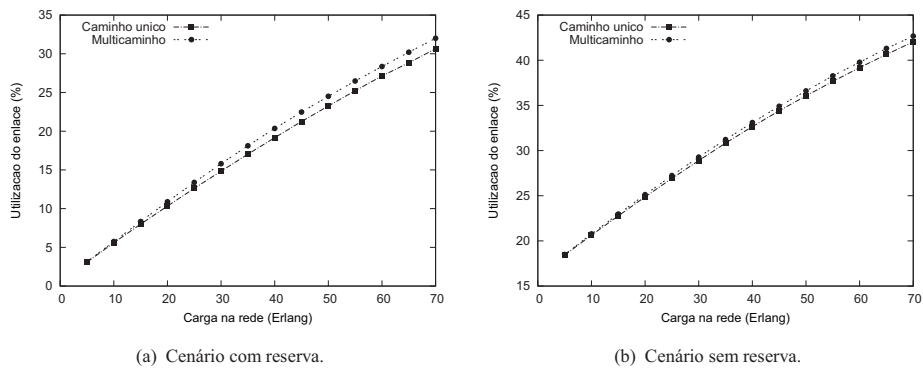


Figura 5. Utilização média dos caminhos ópticos.

do algoritmo multicaminho se iguale ao do algoritmo de caminho único, ou seja, mesmo competindo pelos recursos dos caminhos ópticos, e com alocação parcial da banda passante disponível, ainda assim, o algoritmo multicaminho é capaz de garantir uma maior vazão para o tráfego interdomínio. Na prática, isso implica em um menor bloqueio e uma maior vazão para o tráfego interdomínio, bem como a minimização do impacto sobre o tráfego local de cada domínio.

O custo do algoritmo multicaminho é medido através da quantidade de memória requerida no destino para ordenação dos pacotes, o que é causada pelo retardo diferencial entre os caminhos. Nesse estudo, todas as requisições interdomínio servidas pelo algoritmo multicaminho utilizaram entre 2 e 3 caminhos. A Figura 7 mostra o retardo diferencial médio para o algoritmo multicaminho com diferentes valores para o parâmetro  $\beta$ . Pode-se ver que valores menores de  $\beta$  levam a mais divisões do tráfego com maior frequência para o cenário com reserva, o que resulta em valores maiores do retardo diferencial. Com o aumento da carga da rede a situação se inverte devido a dificuldade em se encontrar caminhos alternativos, o que aumenta o bloqueio para os algoritmos com valores de  $\beta$  menores. Os valores de retardo diferencial para o cenário com reserva, em comparação com o cenário sem reserva, foram superiores pois no cenário sem reserva a quantidade de recursos disponível é mais escassa, o que impediu o algoritmo de utilizar vários caminhos alternativos, portanto, o roteamento multicaminho opera com mais eficiência no cenário com reserva. O que é importante notar nesse resultados é que o custo da utilização do algoritmo, representado pela quantidade de memória requerida, para o al-

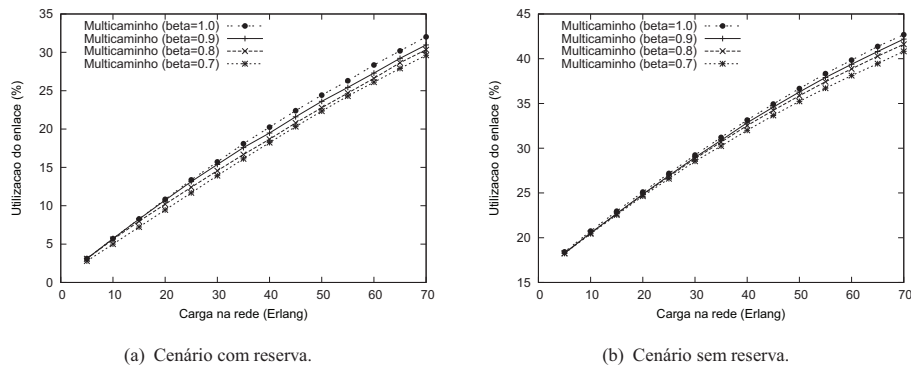


Figura 6. Utilização média dos caminhos ópticos considerando vários valores de  $\beta$ .

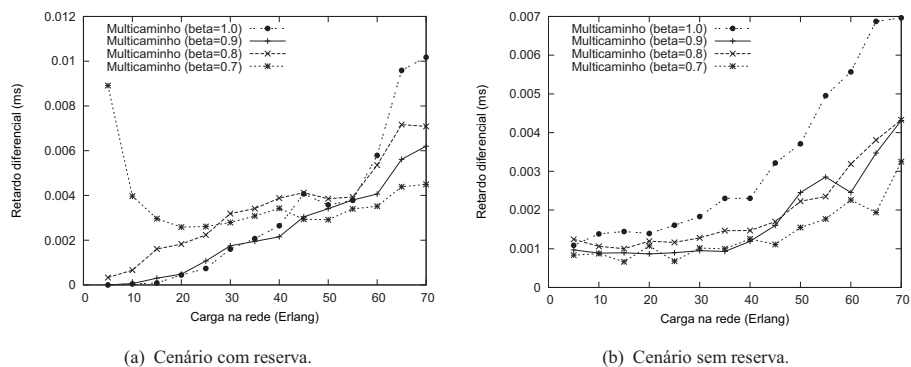


Figura 7. Retardo diferencial considerando vários valores de  $\beta$ .

goritmo com o parâmetro  $\beta = 1$  permaneceu estável para toda a faixa de carga no cenário com reserva, o que facilita a sua adoção em uma rede real onde, de fato, existem restrições de recursos. Os resultados para o cenário sem reserva mostram uma variação total e valores absolutos ainda menores, o que é consequência da limitação do uso de múltiplos caminhos devido a escassez de banda passante disponível.

### Aplicações de transferência de dados

A primeira avaliação diz respeito ao ganho da vazão da rede com a utilização do algoritmo de multicaminho em relação ao algoritmo de caminho único. Como pode ser visto na Figura 8, o algoritmo multicaminho é capaz de prover às requisições de conexão interdomínio a banda passante necessária para atender os requisitos de retardo; o mesmo não ocorre com o algoritmo de caminho único. O baixo bloqueio pode ser observado para ambos os cenários, no caso do algoritmo multicaminho.

Quando o tráfego intradomínio é considerado, o bloqueio aumenta para ambos os algoritmos devido a competição por recursos existente entre os tráfegos inter e intradomínio (Fig. 8(b)), mas mesmo assim os resultados do algoritmo multicaminho mantêm-se praticamente 50% menores do que os valores do bloqueio apresentados pelo algoritmo de caminho único. Em contraste com os resultados para as aplicações em tempo-real, que requerem conexões de duração relativamente longa, quando o roteamento multicaminho é utilizado, uma maior quantidade de banda passante agregada pode ser reservada, o que permite a conclusão da massa de dados em períodos mais curtos, liberando os recursos

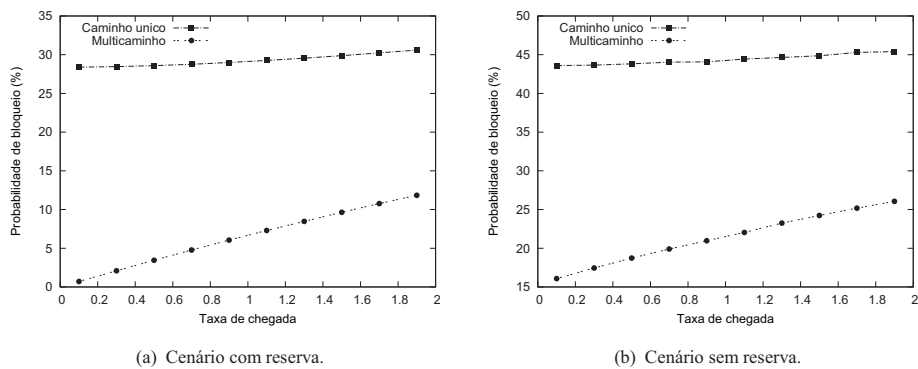


Figura 8. Probabilidade de bloqueio.

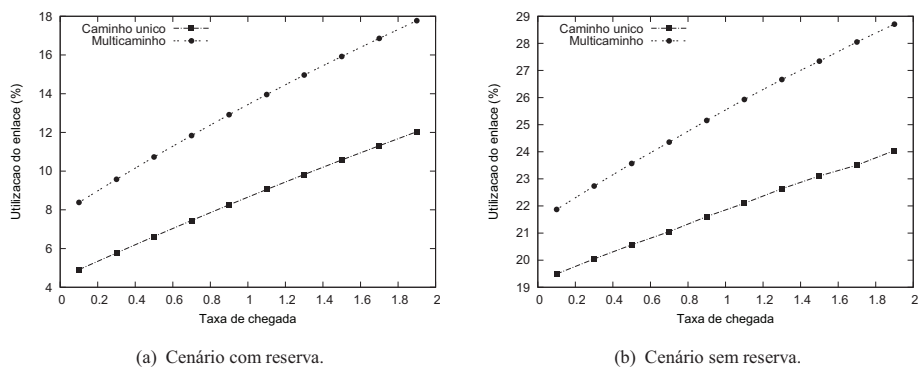


Figura 9. Utilização dos caminhos ópticos.

dos caminhos ópticos. Portanto, a utilização dos caminhos ópticos é reduzida significativamente quando o roteamento multicaminho é utilizado em ambos os cenários, como pode ser visto na Figura 9.

O algoritmo multicaminho proposto para as aplicações de transferência de dados é capaz de acomodar muito mais requisições de conexão do que o algoritmo de caminho único, e utiliza muito menos os recursos da rede, o que acaba por minimizar o seu impacto no tráfego local dos domínios atravessados pelas conexões interdomínio.

Ao avaliar as possibilidades resultadas do ajuste do valor do parâmetro  $\beta$  para o balanceamento da carga (Figura 10(a)), conclui-se que, para o cenário com topologia virtual estática, não houveram diferenças perceptíveis para os valores de  $\beta$  igual a 0,9 ou 1, apenas para  $\beta \leq 0,8$  ocorreram diferenças. O mesmo comportamento pode ser observado para os resultados de utilização média dos caminhos ópticos (Fig. 11(a)). Isso se deve ao fato de que no cenário com reserva não há competição com o tráfego local, existindo mais recursos disponíveis, e portanto a alocação de 90% da banda passante disponível ( $\beta = 0,9$ ) acaba sendo suficiente. Quando o valor de  $\beta$  é definido como 0,7 ou 0,8 os recursos são limitados, o que exige que o algoritmo encontre mais um caminho para atender a requisição de banda, e portanto ocorrem os saltos de bloqueio e de utilização dos caminhos ópticos observados nas Figuras 10(a) e 11(a). Todavia, mesmo para o pior dos casos do algoritmo multicaminho ( $\beta = 0,7$ ), a diferença em relação ao bloqueio obtido com o algoritmo de caminho único (Fig. 8(a)) é da ordem de 28%.

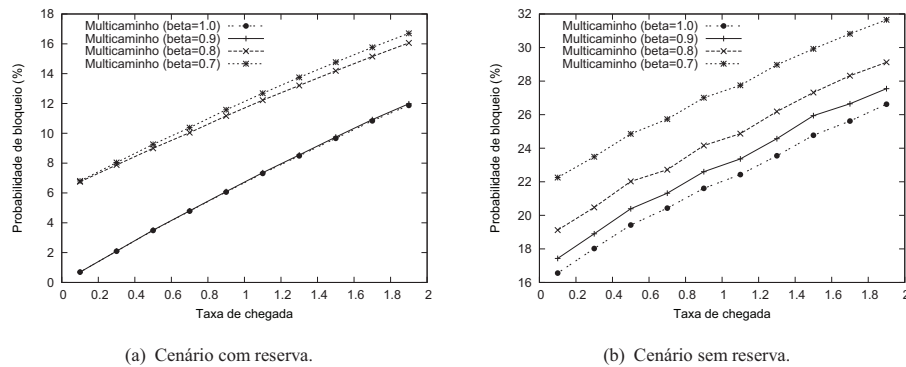


Figura 10. Probabilidade de bloqueio considerando valores de  $\beta$ .

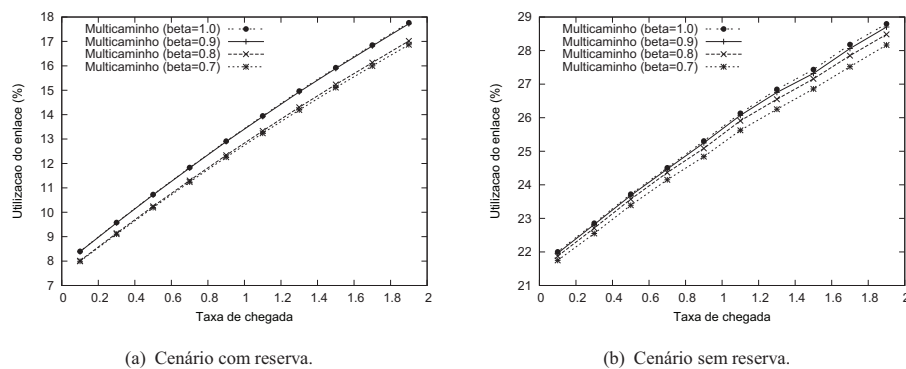


Figura 11. Utilização dos enlaces considerando vários valores de  $\beta$ .

Quando o tráfego intradomínio é considerado, a disputa pelos recursos aumenta e portanto o valor do parâmetro  $\beta$  passa a ter um impacto maior na alocação de caminhos realizada pelo algoritmo de multicaminhos. Quanto menor o valor de  $\beta$ , maior é o bloqueio experimentado na rede (Fig. 10(b)) e, conseqüentemente, menor é a utilização dos caminhos ópticos (Fig. 11(b)). Isso se deve ao fato de o algoritmo ter que buscar mais caminhos alternativos para poder servir a mesma chamada. Mesmo no pior caso, quando  $\beta = 0,7$ , ou seja, quando apenas 70% da banda passante disponível pode ser utilizada pelo algoritmo de multicaminhos, o bloqueio para a carga mais alta (2 Erlang) ficou em 32% que é menor que os 44% gerados pelo algoritmo de caminho único. Em resumo, dado que a diminuição do valor do parâmetro  $\beta$  acarreta no aumento do bloqueio, escolher valores menores para  $\beta$  nem sempre é a melhor opção. O valor de  $\beta$  deve ser escolhido buscando limiares como o observado nas Figuras 10(a) e 11(a).

## 5. Conclusões

Neste artigo, dois algoritmos para o provisionamento de conexões interdomínio em redes ópticas foram propostos. Os resultados apresentados mostram que as soluções multicaminho propostas superam a solução tradicional de caminho único em todos os cenários avaliados. O algoritmo especializado em aplicações multimídia foi capaz de reduzir o bloqueio em até uma ordem grandeza e aumentar a vazão da rede, mantendo o requisito de memória no destino estável, o que viabiliza a sua aplicação em redes reais. O algoritmo voltado para aplicações de transferência de dados reduzir o bloqueio em mais de 50% e manteve a utilização da rede baixa devido a sua capacidade de agregação de banda

passante. Ambos utilizam um parâmetro  $\beta$  que permite o ajuste do nível de banda passante alocável, mecanismo particularmente útil no cenário sem reserva, no qual busca-se minimizar o impacto da alocação de recursos no tráfego local dos domínios atravessados.

## Referências

- Ahuja, S., Korkmaz, T., and Krunz, M. (2004). Minimizing the differential delay for virtually concatenated ethernet over sonet systems. In *International Conference on Computer Communications and Networks*, pages 205–210.
- Chamania, M., Chen, X., and Jukan, A. (2009). An adaptive inter-domain pce framework to improve resource utilization and reduce inter-domain signaling. *Optical Switching and Networking*, 6(4):259–2677.
- Chamania, M. and Jukan, A. (2009). A survey of inter-domain peering and provisioning solutions for the next generation optical networks. *Communications Surveys & Tutorials, IEEE*, 11:33 – 51.
- Chen, X., Jukan, A., Drummond, A. C., and da Fonseca, N. L. S. (2009). A multipath routing mechanism in optical networks with extremely high bandwidth requests. In *IEEE Global Communications Conference*.
- Deelman, E., Singh, G., Su, M.-H., Blythe, J., Gil, Y., Kesselman, C., Mehta, G., Vahi, K., Berriman, G. B., Good, J., Laity, A., Jacob, J. C., and Katz, D. S. (2005). Pegasus: A framework for mapping complex scientific workflows onto distributed systems. *Scientific Programming Journal*, 13(3):219–237.
- Drummond, A. C. (2010). Wdmsim (<http://www.lrc.ic.unicamp.br/wdmsim>).
- Gao, D., Shu, Y., Liu, S., and Yang, O. (2002). Delay-based adaptive load balancing in mpls networks. In *IEEE International Conference on Communications*, volume 2, pages 1184–1188.
- Huang, S., Deng, Y., Luo, P., Li, B., and Gu, W. (2008). A new dynamic inter-domain routing and grooming algorithms under dedicated protection in intelligent optical networks. *Dynamics of Continuous, Discrete and Impulsive Systems*, 15(4b):589–604.
- I.Cidon, R.Rom, and Shavitt, Y. (1999). Analysis of multi-path routing. *IEEE/ACM TRANSACTIONS ON NETWORKING*, 7(6):885–896.
- Mao, S., Panwar, S., and Y.T., H. (2006). On minimizing end-to-end delay with optimal traffic partitioning. *IEEE Transactions on Vehicular Technology*, 55(2):681–690.
- Pan, P., Hahne, E., and Schulzrinne, H. (2000). The border gateway reservation protocol (bgrp) for tree-based aggregation of inter-domain reservations. *Journal of Communications and Networks*.
- Sofia, R., Guerin, R., and Veiga, P. (2003). A shared-segment inter-domain control aggregation protocol. In *High Performance Switching and Routing*.
- Vasseur, J.-P., Zhang, R., Bitar, N., and Le Roux, J. (2009). A backward-recursive pce-based computation (brpc) procedure to compute shortest constrained inter-domain traffic engineering label switched paths. <http://tools.ietf.org/html/rfc5441>.
- Xu, W. and Rexford, J. (2006). Miro: Multi-path interdomain routing. In *ACM SIGCOMM*, pages 171–182.