

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Flávio de Melo Pereira

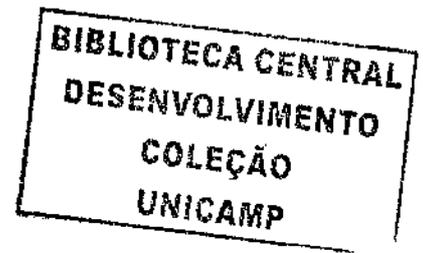
MODELAGEM, POLICIAMENTO E ESCALONAMENTO EM REDES ETHERNET PON

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Telecomunicações e Telemática.

Orientador: Dalton Soares Arantes

Co-orientador: Nelson Luis Saldanha da Fonseca



Campinas

2006

Este exemplar corresponde à redação final da tese defendida por: Flávio de Melo Pereira e aprovada pela Comissão Julgada em 25/07/06 Sauir Ar
[Signature] Orientador

Modelagem, policiamento e escalonamento em redes Ethernet PON.
© 2006 por Flávio de Melo Pereira. Todos os direitos reservados.

Para a elaboração de parte deste trabalho, o autor contou com o suporte financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo entre março de 2002 e abril de 2003 (Processo nº 14379/4).

UNIDADE ABC
Nº CHAMADA T/UNICAMP
P414m
V _____ EX _____
TOMBO BC/ 70710
PROC. 16-P 00133.06
C _____ \$ _____
PREÇO 11.00
DATA 22/11/06
BIB-ID 392101

Ficha catalográfica elaborada pela Biblioteca da Área de Engenharia e Arquitetura – BAE – Unicamp

| | |
|-------|---|
| P414m | Pereira, Flávio de Melo Modelagem, policiamento e escalonamento de tráfego em redes Ethernet PON / Flávio de Melo Pereira. – Campinas, SP: [s.n.], 2006. Orientadores: Dalton Soares Arantes, Nelson Luis Saldanha da Fonseca. Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação. 1. Redes de computação. 2. Telecomunicações – Tráfego. 3. Comunicações óticas. 4. Sistemas de comunicação em banda larga. I. Arantes, Dalton Soares. II. Fonseca, Nelson Luis Saldanha da. III. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título. |
|-------|---|

Título em Inglês: Traffic modelling, policing and scheduling in Ethernet passive optical networks
Palavras-chave em Inglês: Computer networks, Optical networks, Broadband communication systems, Access networks, Telecommunication systems – Traffic
Área de concentração: Telecomunicações e Telemática
Titulação: Doutor em Engenharia Elétrica
Banca Examinadora: Célio Vinícius Neves de Albuquerque, Hélio Waldman, Maurício Ferreira Magalhães, Otto Carlos Muniz Bandeira Duarte, Paulo Cardieri e Shusaburo Motoyama
Data da defesa: 25/07/2006

ERRATA

- pág. 8, parágrafo 4, linha 2. Onde se lê "...cuja transmissão é ilustrada na Figura 1.3", leia-se "...cuja transmissão é ilustrada na Figura 1.3 [7]".
- pág. 8, parágrafo 5, linha 2. Onde se lê "...mostra a Figura 1.4", leia-se "...mostra a Figura 1.4 [7]".
- pág. 10, parágrafo 3, linha 3. Onde se lê "...por meio do protocolo MPCP", leia-se "...por meio do protocolo MPCP [20]".
- pág. 10, parágrafo 3, linha 4. Onde se lê "...na Figura 1.5", leia-se "...na Figura 1.5 [20]".
- pág. 10, parágrafo 4, linha 6. Onde se lê "...na Figura 1.6", leia-se "...na Figura 1.6 [20]".
- pág. 13, parágrafo 1, linha 6. Onde se lê "...na Figura 1.8", leia-se "...na Figura 1.6 [18,19]".
- pág. 77, parágrafo 1, linha 3. Onde se lê "...encontradas em [20,84,85].", leia-se "...encontradas em [84,85,91].", sendo a referência 91 a seguinte:
[91] M.P. McGarry, M. Maier e M. Reisslein, "Ethernet PONs : a survey of dynamic bandwidth allocation DBA algorithms", *IEEE Commun. Mag.*, vol. 42, n° 8, págs. s8-s15, agosto de 2004.
- pág. 79, parágrafo 4, linha 2. Onde se lê "...serviço ou fluxos individuais.", leia-se "...serviços ou fluxos individuais [92 apud 91].", sendo a referência 92 a seguinte:
[92] L. Zhang, E. An, C. Youn, H. Yeo e S. Yang, "Dual DEB-GPS Scheduler for Delay Constraint Applications in Ethernet Passive Optical Networks", *IEICE Trans. Commun.*, vol. E86-B, n° 5, págs. 1575-1584, maio de 2003.
- pág. 79, parágrafo 4, linha 6. Onde se lê "...efetivas determinísticas", leia-se "...efetivas determinísticas [91].".
- pág. 79, parágrafo 6, linha 6. Onde se lê "...de curta duração", leia-se "...de curta duração [91].".

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Flávio de Melo Pereira

MODELAGEM, POLICIAMENTO E ESCALONAMENTO EM REDES ETHERNET PON

Tese de Doutorado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Telecomunicações e Telemática.
Aprovada em 25/07/2006

Banca Examinadora:

Prof. Dr. Célio Vinícius Neves de Albuquerque - UFF
Prof. Dr. Dalton Soares Arantes - UNICAMP
Prof. Dr. Hélio Waldman - UNICAMP
Prof. Dr. Mauricio Ferreira Magalhães - UNICAMP
Prof. Dr. Otto Carlos Muniz Bandeira Duarte - UFRJ
Prof. Dr. Paulo Cardieri - UNICAMP
Prof. Dr. Shusaburo Motoyama - UNICAMP

Campinas

2006

RESUMO

Neste trabalho, apresentamos uma nova proposta de modelagem, policiamento e escalonamento de tráfego em redes Ethernet com elementos ópticos passivos. Esta proposta visa ao suporte a garantias de desempenho individuais para os fluxos. Propomos um processo envelope denominado Processo de Chegada com Limitante Fracionário para a representação de tráfego multimídia com características fractais. Com base neste modelo, propomos o uso do algoritmo do balde furado para o policiamento do tráfego. Para o escalonamento de tráfego, propomos uma nova política em que a rede reserva uma parcela mínima de banda para cada fluxo, sendo que os fluxos com demanda reprimida recebem uma parcela justa dos recursos ociosos. Para esta política, denominada Compartilhamento Proporcional com Reserva de Carga, estabelecemos as condições de estabilidade da rede e apresentamos expressões limitantes para a justiça e para o atraso dos fluxos.

Palavras-chave: redes de computação, telecomunicações – tráfego, comunicações ópticas, sistemas de computação em banda larga, redes de acesso.

ABSTRACT

We propose new framework for traffic modelling, policing and scheduling in Ethernet Passive Optical Networks. This framework aims at providing performance guarantees to individual flows in access networks. We propose a new envelope process called Fractional Bounded Arrival Process for self-similar multimedia traffic representation. The policing of such a kind of traffic by using the leaky bucket algorithm is also proposed. We present a novel discipline for scheduling the upstream flows in EPON networks. In such a discipline, which is called Proportional Sharing with Load Reservation, each flow is guaranteed a minimum share of bandwidth and the unused bandwidth is fairly distributed among them in proportion to their own demands and priorities. We analyze the performance of the proposed discipline and provide fairness and delay bounds on a per-flow basis.

Keywords: computer networks, optical networks, broadband communication systems, access networks, telecommunication systems – traffic.

AGRADECIMENTOS

Ao concluir o presente trabalho, completo doze anos contínuos na Faculdade de Engenharia Elétrica e de Computação da Unicamp. A oportunidade de fazer parte da comunidade da FEEC tem sido para mim uma grande honra, e é com muita felicidade que tenho contado com a amizade e com o apoio de muitos dos seus professores e funcionários ao longo de todos esses anos. Por isso, gostaria de registrar a minha gratidão a esta instituição, que sempre me acolheu com muita dedicação e respeito. Espero que minhas modestas contribuições tenham sido, de alguma maneira, úteis ao seu desenvolvimento.

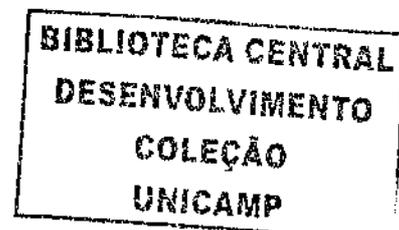
Registro também a minha gratidão aos meus orientadores, Prof. Dr. Dalton Soares Arantes e Prof. Dr. Nelson Luis Saldanha da Fonseca, não só pela orientação, mas também pelo respeito e pela confiança que ambos demonstraram durante todo o período em que trabalhamos juntos. Ainda me surpreendo com o quanto eles acreditaram na minha capacidade de concluir esta tese apesar das muitas dificuldades enfrentadas ao longo do curso, sobretudo após o meu ingresso no Banco Central do Brasil. Espero haver empenhado-me o suficiente para que a tese chegasse a bom termo e o crédito fosse merecido.

Gostaria ainda de agradecer a todos aqueles que, de alguma maneira, colaboraram para a execução deste trabalho. Em especial, gostaria de agradecer aos membros da banca examinadora pela revisão do texto e pelas importantes sugestões e correções; ao Dr. Glen Kramer da Universidade da Califórnia em Davis pela ajuda com o gerador de tráfego utilizado nas simulações e por ceder o código C++ do escalonamento IPACT; a Fundação de Amparo à Pesquisa do Estado de São Paulo, pelo suporte financeiro¹; e aos colegas do fretado Paulínia-Unicamp-São Paulo que muitas vezes me cederam as fileiras vazias do ônibus para que eu pudesse redigir ao longo do trajeto entre a minha casa e o trabalho. Deixo sinceros agradecimentos pelo apoio e incentivo também aos colegas da pós-graduação da FEEC, do Centro de Ensino de Línguas da Unicamp, do Banco Central do Brasil e da minha ex-república em Campinas.

Finalmente, agradeço à minha família e aos meus amigos pelo apoio e, principalmente, por terem sido sempre compreensivos diante da falta de tempo para os encontros e os passeios. Em especial, deixo carinhosos agradecimentos à minha esposa Manuella que, sempre zelosa e incentivadora, acompanhou-me nas qualificações, na redação e na defesa desta tese; meus sinceros agradecimentos aos irmãos Alexandre e Rafael, que sempre incentivaram e torceram pelo meu sucesso; e sobretudo os mais que devidos agradecimentos aos meus pais Jorge e Marluce, principais responsáveis pela minha educação e sem os quais nenhuma das minhas realizações teria sido possível. A vocês, dedico esta tese.

– F.M.P. , setembro de 2006.

¹cf. Processo nº01/14379-4.



*À minha esposa, Manuella,
aos meus pais, Jorge e Marluce, e
aos meus irmãos, Alexandre e Rafael.*

SUMÁRIO

| | |
|---|-------------|
| Lista de Figuras | xiii |
| Lista de Tabelas | xv |
| Lista de Acrônimos | xvi |
| 1 Introdução | 1 |
| 1.1 Tecnologias de acesso fixo a pequenos usuários | 1 |
| 1.2 Redes de acesso ópticas | 3 |
| 1.2.1 Arquitetura FTTH | 4 |
| 1.2.2 Redes ópticas com elementos passivos | 5 |
| 1.3 Redes Ethernet PON | 8 |
| 1.3.1 Protocolo MPCP | 8 |
| 1.3.2 Acesso por polling | 10 |
| 1.3.3 Sinalização interposta | 12 |
| 1.3.4 Qualidade de serviço em redes Ethernet | 13 |
| 1.3.5 Penalização de baixa carga | 14 |
| 1.3.6 Escalonamento em redes EPON | 15 |
| 1.4 Proposta deste trabalho | 15 |
| 2 Modelagem de tráfego em redes | 19 |
| 2.1 Processos <i>on/off</i> de cauda pesada | 20 |
| 2.2 Processo $M/G/\infty$ de cauda pesada | 21 |
| 2.3 Movimento browniano fracionário | 22 |
| 2.4 Processos α -estáveis e auto-similares | 24 |
| 2.5 Estimação do parâmetro H | 26 |
| 2.5.1 Estimador de Whittle | 26 |
| 2.5.2 Gráfico variância-tempo | 27 |
| 2.5.3 Gráfico da estatística R/S | 27 |
| 2.5.4 Estimador baseado em wavelets | 27 |
| 2.6 Processo de chegadas com limitante fracionário | 28 |
| 2.6.1 Definição | 29 |
| 2.6.2 Estimação dos parâmetros FBAP para tráfego real | 30 |
| 2.6.3 Armazenamento do tráfego FBAP | 37 |
| 2.6.4 Processo de chegadas com limitante multifractal | 38 |
| 3 Policiamento de tráfego fractal | 45 |
| 3.1 Algoritmo do balde furado | 46 |
| 3.2 Algoritmo do balde furado fractal | 46 |
| 3.3 Policiamento de tráfego fractal com o algoritmo do balde furado | 48 |
| 3.4 Comparação entre o balde furado e o balde furado fractal | 52 |

| | | |
|----------|---|------------|
| 4 | Escalonamento PSLR para redes EPON | 57 |
| 4.1 | Notação e modelo do sistema | 57 |
| 4.2 | Critérios de justiça | 58 |
| 4.3 | Métricas de justiça | 59 |
| 4.4 | Compartilhamento proporcional com reserva de carga | 60 |
| 4.5 | Algoritmo computacional | 62 |
| 4.6 | Exemplo numérico | 66 |
| 4.7 | Análise da política PSLR | 68 |
| 4.7.1 | Convergência e complexidade computacional | 68 |
| 4.7.2 | Estabilidade | 69 |
| 4.7.3 | Justiça | 70 |
| 4.7.4 | Latência e atraso | 72 |
| 5 | Coletânea de algoritmos de escalonamento EPON | 77 |
| 5.1 | Alocação dinâmica de banda para multimídia | 77 |
| 5.2 | Polling intercalado com tempo de ciclo adaptativo | 78 |
| 5.3 | GPS com banda efetiva determinística | 79 |
| 5.4 | Polling com garantia de banda | 79 |
| 5.5 | Compartilhamento justo baseado em envelope de serviço | 81 |
| 5.6 | Comparação com a política PSLR | 83 |
| 6 | Conclusão | 85 |
| | Referências bibliográficas | 92 |
| A | Provas de Teoremas | 93 |
| A.1 | Prova do Teorema 4.4 | 93 |
| A.2 | Prova do Teorema 4.6 | 94 |
| A.3 | Prova do Teorema 4.7 | 95 |
| A.4 | Prova do Teorema 4.8 | 96 |
| A.5 | Prova do Teorema 4.10 | 97 |
| B | Aprimoramento do algoritmo PSLR | 101 |
| B.1 | Algoritmo computacional | 101 |
| B.2 | Exemplo numérico | 105 |
| B.3 | Análise da política PSLR | 106 |
| B.3.1 | Convergência e complexidade computacional | 106 |
| B.3.2 | Estabilidade | 108 |
| B.3.3 | Justiça | 111 |
| B.3.4 | Latência e atraso | 113 |

LISTA DE FIGURAS

| | | |
|------|---|-----|
| 1.1 | Largura de banda requerida por aplicações multimídia. | 2 |
| 1.2 | Exemplo de rede de acesso PON. | 6 |
| 1.3 | Transmissão de fluxos de descida em uma rede EPON. | 9 |
| 1.4 | Transmissão de fluxos de subida em uma rede EPON. | 10 |
| 1.5 | Estratégia <i>poll-and-stop</i> | 11 |
| 1.6 | Estratégia de <i>polling</i> intercalado. | 11 |
| 1.7 | Estratégia de <i>polling</i> intercalado com parada. | 12 |
| 1.8 | Funcionamento da sinalização interposta. | 13 |
| | | |
| 2.1 | Validação do modelo FBAP para um ruído gaussiano fracionário. | 31 |
| 2.2 | Representação da série BOND utilizando o envelope FBAP. | 34 |
| 2.3 | Representação da série FUSS1 utilizando o envelope FBAP. | 35 |
| 2.4 | Representação da série DECPKT4 utilizando o envelope FBAP. | 36 |
| 2.5 | Representação do traço ATP utilizando os envelopes FBAP e MFBAP. | 41 |
| 2.6 | Representação do traço TALK1 utilizando os envelopes FBAP e MFBAP. | 42 |
| 2.7 | Representação do traço LBLTCP3 utilizando os envelopes FBAP e MFBAP. | 43 |
| | | |
| 3.1 | Restrição imposta ao tráfego pelo algoritmo do Balde Furado. | 47 |
| 3.2 | Condição em que o balde furado fractal permite oferecer limitantes. | 48 |
| 3.3 | Condição em que o balde furado fractal não permite oferecer limitantes. | 49 |
| 3.4 | Tamanho de balde como função da taxa de drenagem para $\mu_s = 0,800$, $k_s\gamma_s = 0,595$ e $H_s = 0,800$ | 50 |
| 3.5 | Balde furado com tráfego auto-similar. Violação do tráfego médio. | 52 |
| 3.6 | Balde furado com tráfego auto-similar. Violação do fator de escala. | 53 |
| 3.7 | Balde furado com tráfego auto-similar. Violação do fator H | 54 |
| 3.8 | Policimento de tráfego FBAP real. Envelope FBAP para a série DECPKT2-TCP. | 54 |
| 3.9 | Policimento de tráfego FBAP real. Limitante de latência. | 55 |
| 3.10 | Policimento de tráfego FBAP real. Taxa de descarte. | 55 |
| | | |
| 4.1 | Operação da política PSLR. Vazão da rede de acesso. | 67 |
| 4.2 | Operação da política PSLR. Duração do ciclo de <i>polling</i> | 68 |
| 4.3 | Operação da política PSLR. Tempo de passagem total da rede. | 69 |
| 4.4 | Operação da política PSLR. Resultados para os fluxos da ONU 1. | 71 |
| 4.5 | Operação da política PSLR. Resultados para os fluxos da ONU 1. | 72 |
| 4.6 | Operação da política PSLR. Resultados para os fluxos da ONU 9. | 73 |
| 4.7 | Operação da política PSLR. Resultados para os fluxos da ONU 9. | 74 |
| | | |
| A.1 | Ciclos de serviço considerados na prova do Teorema 4.10. | 97 |
| | | |
| B.1 | Operação da política PSLR. Vazão da rede de acesso. | 105 |
| B.2 | Operação da política PSLR. Duração do ciclo de <i>polling</i> | 106 |
| B.3 | Operação da política PSLR. Tempo de passagem total da rede. | 107 |
| B.4 | Operação da política PSLR. Resultados para os fluxos da ONU 1. | 108 |

| | | |
|-----|--|-----|
| B.5 | Operação da política PSLR. Resultados para os fluxos da ONU 1. | 109 |
| B.6 | Operação da política PSLR. Resultados para os fluxos da ONU 9. | 110 |
| B.7 | Operação da política PSLR. Resultados para os fluxos da ONU 9. | 111 |
| B.8 | Ciclos de serviço considerados na prova do Teorema B.7. | 114 |

LISTA DE TABELAS

| | | |
|-----|---|----|
| 1.1 | Capacidade das principais variantes da tecnologia DSL. | 3 |
| 1.2 | Principais características dos padrões PON. | 7 |
| 2.1 | Parâmetros do envelope FBAP correspondente às séries analisadas. | 33 |
| 2.2 | Traços utilizados na validação do modelo MFBAP. | 40 |
| 4.1 | Exemplo de operação da política PSLR. Parâmetros da EPON simulada. | 66 |
| 4.2 | Exemplo de operação da política PSLR. Parâmetros dos fluxos. | 66 |
| 4.3 | Exemplo de operação da política PSLR. Banda teórica a ser destinada a cada fluxo. | 70 |
| 4.4 | Comparação entre o RFB da política PSLR e outras políticas de escalonamento. | 75 |
| 4.5 | Comparação entre a latência da política PSLR e de outras políticas da literatura. | 75 |

LISTA DE ACRÔNIMOS

| | |
|---------|--|
| AFB | Limitante de justiça absoluto (<i>Absolute Fairness Bound</i>) |
| RFB | Limitante de justiça relativo (<i>Relative Fairness Bound</i>) |
| APON | Rede óptica de elementos passivos com ATM (<i>ATM Passive Optical Network</i>) |
| ATM | Modo de transferência assíncrono (<i>Asynchronous Transfer Mode</i>) |
| FLB | Balde furado fractal (<i>Fractal Leaky Bucket</i>) |
| BGP | Polling com garantia de banda (<i>Bandwidth Guaranteed Polling</i>) |
| BPON | Rede óptica de elementos passivos de banda larga (<i>Broadband Passive Optical Network</i>) |
| BRI | Interface de taxa básica (<i>Basic Rate Interface</i>) |
| CCITT | Comitê Consultivo Internacional para Telefonia e Telegrafia (<i>Comité Consultatif International Téléphonique et Télégraphique</i>) |
| CSMA/CD | Protocolo de acesso múltiplo com medição de portadora e detecção de colisão (<i>Carrier Sense Multiple Access/Collision Detection</i>) |
| DEB-GPS | GPS com banda efetiva determinística (<i>Deterministic Effective Bandwidth</i>) |
| DRR | <i>Deficit Round Robin</i> |
| DWDM | Multiplexagem densa por divisão em comprimento de onda (<i>Dense WDM</i>) |
| EDA | Algoritmo de distribuição uniforme (<i>Even Distribution Algorithm</i>) |
| EPON | Rede Ethernet óptica de elementos passivos (<i>Ethernet Passive Optical Network</i>) |
| FBAP | Processo de chegadas com limitante fracionário (<i>Fractional Bounded Arrival Process</i>) |
| FBM | Movimento browniano fracionário (<i>Fractional Brownian Motion</i>) |
| FQSE | Compartilhamento justo baseado em envelope de serviço (<i>Fair Queuing with Service Envelopes</i>) |
| FSAN | <i>Full Service Access Network committee</i> |
| FTTB | Fibra até o negócio (<i>Fiber to the Business</i>) |
| FTTC | Fibra até o meio fio (<i>Fiber to the Curb</i>) |
| FTTCAB | Fibra até o gabinete (<i>Fiber to the Cabinet</i>) |
| FTTH | Fibra até a residência (<i>Fiber to the Home</i>) |

| | |
|---------|--|
| FTTP | Fibra até o local (<i>Fiber to the Premise</i>) |
| GPON | Rede gigabit óptica de elementos passivos (<i>Gigabit Passive Optical Network</i>) |
| GPS | Compartilhamento generalizado de processador (<i>Generalized Processor Sharing</i>) |
| HDTV | TV de alta definição (<i>High Definition TV</i>) |
| IEEE | <i>The Institute of Electrical and Electronic Engineers</i> |
| IP | Protocolo Internet (<i>Internet Protocol</i>) |
| IPACT | <i>Polling</i> intercalado com tempo de ciclo adaptativo (<i>Interleaved Polling with Adaptive Cycle Time</i>) |
| ITU-T | União Internacional de Telecomunicações - Setor de Padronização de Telecomunicações (<i>International Telecommunications Union - Telecom Standardization Sector</i>) |
| LBAP | Processo de chegadas com limitante linear (<i>Linear Bounded Arrival Process</i>) |
| LLID | Identidade lógica de enlace (<i>Logical Link ID</i>) |
| LRD | Dependência de longa duração (<i>Long Range Dependency</i>) |
| MAC | Controle de acesso ao meio (<i>Medium Access Control</i>) |
| MFBAP | Processo de chegada com limitante multifractal (<i>Multifractal Bounded Arrival Process</i>) |
| MPCP | Protocolo de controle multiponto (<i>Multipoint Control Protocol</i>) |
| MPEG | <i>Motion Picture Expert Group</i> |
| OLT | Terminal de linha óptica (<i>Optical Line Terminal</i>) |
| ONU | Unidade de rede óptica (<i>Optical Network Unit</i>) |
| PHB | Comportamento do tráfego por nó (<i>Per-Hop Behavior</i>) |
| PON | Rede óptica com elementos passivos (<i>Passive Optical Networks</i>) |
| PRI | Interface de taxa primária (<i>Primary Rate Interface</i>) |
| PSLR | Compartilhamento proporcional com reserva de carga (<i>Proportional Sharing with Load Reservation</i>) |
| RDSI | Rede digital de serviços integrados |
| RDSI-FE | Rede digital de serviços integrados de faixa estreita |
| RTT | Tempo de ida e volta (<i>Round Trip Time</i>) |
| SCFQ | <i>Self-Clocked Fair Queueing</i> |
| SDH | Hierarquia digital síncrona (<i>Synchronous Digital Hierarchy</i>) |
| SONET | Rede óptica síncrona (<i>Synchronous Optical Network</i>) |
| SRD | Dependência de curta duração (<i>Short Range Dependency</i>) |

| | |
|---------|--|
| TCP | Protocolo de controle Telnet (<i>Telnet Control Protocol</i>) |
| TDM | Multiplexagem por divisão no tempo (<i>Time Division Multiplex</i>) |
| TDMA | Acesso múltiplo por divisão por tempo (<i>Time Division Multiple Access</i>) |
| VOIP | Voz sobre IP (<i>Voice-over-IP</i>) |
| WDM | Multiplexação por divisão em comprimento de onda (<i>Wavelength Division Multiplexing</i>) |
| WDM-PON | Rede óptica com elementos passivos baseada em WDM (<i>WDM Passive Optical Network</i>) |
| WFQ | <i>Weighted Fair Queueing</i> |
| WRR | <i>Weighted Round Robin</i> |
| WWW | <i>World-Wide Web</i> |

CAPÍTULO 1

INTRODUÇÃO

Hoje, as operadoras de serviços de telecomunicações consideram essencial a oferta de serviços integrados de voz, dados e vídeo como meio de aumentar a lucratividade dos seus negócios. Corrobora para esta percepção o interesse cada vez maior dos usuários por aplicações tais como telefonia sobre IP, transmissão de imagens por *webcam* e jogos interativos pela rede, além da demanda por novos serviços como, por exemplo, a difusão de TV em alta definição (HDTV) e a transmissão de vídeo sob demanda. Neste cenário, espera-se que o advento das redes de serviços multimídia integrados permita que as operadoras componham pacotes de serviços com os seguintes componentes essenciais:

- *Telefonia*. Diversas empresas já utilizam a tecnologia de voz sobre IP (VOIP) para reduzir os custos da sua rede interna de telefonia. Algumas operadoras utilizam esta tecnologia para transmitir voz entre centrais telefônicas tradicionais ou para oferecer telefonia convencional a usuários que dispõem de acesso em banda larga;
- *Internet de alta velocidade*. A maior largura de banda disponível torna possível aumentar a interatividade dos sítios Internet, aprimorando serviços como a educação à distância, o comércio eletrônico e os jogos *on-line*;
- *Vídeo*. As redes devem transmitir programação de TV em sinal digital em baixa e alta definição, além de suportar serviços como TV interativa, vídeo-sob-demanda, videoconferência e telemedicina.

Para estas aplicações, os valores típicos de largura de banda em função da aplicação e da qualidade do serviço são apresentados na Figura 1.1 [1]. Note que, dependendo dos serviços e desempenho contratados, a demanda de cada usuário pode facilmente chegar a dezenas de Mb/s.

Atualmente, apenas grandes empresas podem dispor de tal volume de banda, por meio de tecnologias como o gigabit Ethernet, o WDM¹ e os protocolos síncronos de transmissão (SDH² e SONET³). Por outro lado, pequenas empresas e usuários domésticos não suportam os elevados custos destas soluções e contam apenas com tecnologias de acesso com capacidade de centenas de kb/s a poucos Mb/s. Oferecer banda larga e serviços multimídia a estes clientes, que representam uma parcela importante do mercado de telecomunicações, é hoje uma das principais metas da indústria e objeto de intensa pesquisa acadêmica.

1.1 TECNOLOGIAS DE ACESSO FIXO A PEQUENOS USUÁRIOS

Para os usuários residenciais e para pequenas empresas, a forma mais básica de acesso ainda é a conexão discada. O acesso discado é uma conexão ponto a ponto realizada por meio de canais telefônicos analógicos em que se atinge uma capacidade máxima de 56kb/s.

¹Multiplexação por Divisão em Comprimento de Onda (*Wavelength Division Multiplexing*).

²Protocolo de Hierarquia Digital Síncrona (*Synchronous Digital Hierarchy*).

³Protocolo de Rede Óptica Síncrona (*Synchronous Optical Network*).

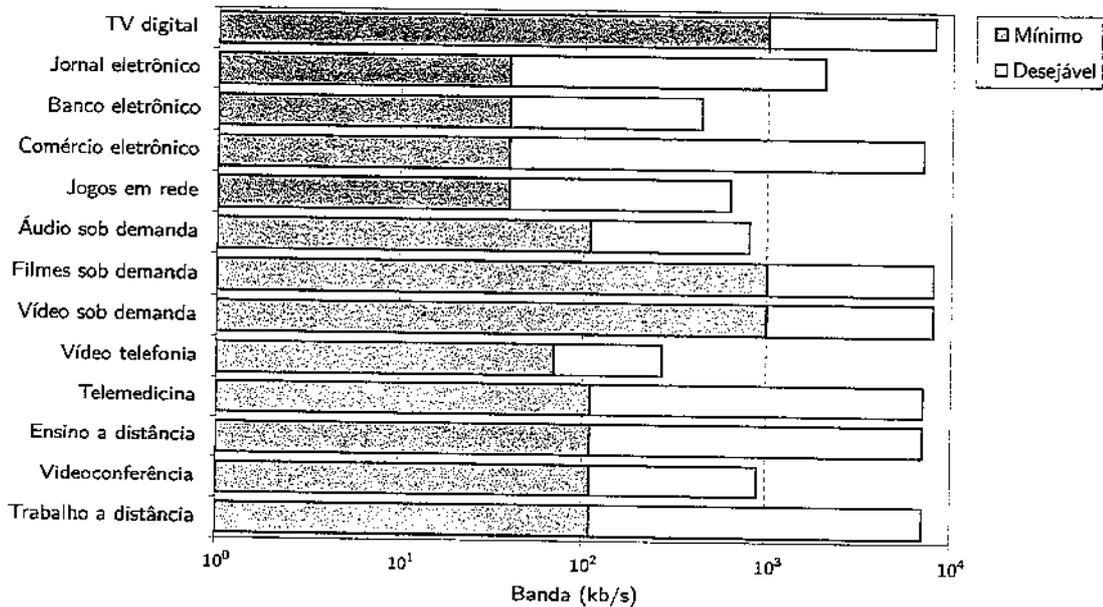


Figura 1.1: Largura de banda requerida por aplicações multimídia.

No Brasil, assim como em outros países emergentes, o acesso discado ainda atende uma parcela relevante dos usuários. Evidentemente, sua capacidade não é suficiente para prover serviços multimídia com mínimo desempenho. A primeira alternativa apresentada pela indústria foi a Rede Digital de Serviços Integrados (RDSI). A RDSI é uma rede ponto-a-ponto que utiliza dois canais telefônicos digitais de 64kb/s (canais B) para oferecer aos usuários um acesso a 128kb/s e um canal de 16kb/s (canal D) para sinalização e controle [2]. Este acesso é denominado Interface de Taxa Básica (BRI⁴). Extensões utilizando canais D de 64kb/s e um maior número de canais B, conhecidas por Interface de Taxa Primária (PRI⁵), alcançam taxas de até 1,544Mb/s (em canais telefônicos T1) e 2,048Mb/s (em canais telefônicos E1). Cabe observar que muitos autores se referem às RDSI como RDSI de faixa estreita (RDSI-FE) para não confundí-las com as redes de serviços multimídia de banda larga, que são conhecidas pela mesma sigla.

Mesmo em países desenvolvidos, a tecnologia RDSI-FE teve utilização restrita dado o seu custo elevado e pouca popularidade, na época, dos serviços proporcionados. Vale lembrar que as primeiras especificações da RDSI-FE feitas pelo ITU-T (então CCITT) remetem ao início dos anos 80, quando a popularização dos computadores apenas se iniciava, a Internet era restrita e a *World-Wide Web* sequer existia. No Brasil, onde o uso de tecnologia digital para telefonia só se difundiu a partir da segunda metade da década de 90, a tecnologia RDSI-FE praticamente não foi utilizada.

Em meados da década de 90, a popularização dos computadores e da Internet fizeram a demanda por acesso em banda larga crescer espantosamente. Isto fez a indústria atuar em duas frentes. Na primeira, buscavam-se alternativas de custo inferior ao das RDSI-FE e que fossem capazes prover banda suficiente para as aplicações multimídia ora disponíveis. Isto atenderia os usuários imediatamente, despertando-lhes o interesse em adquirir novas funcionalidades. Na segunda, desenvolviam-se redes capazes de suportar os serviços multimídia mais sofisticados.

Tecnologias como o acesso assimétrico via satélite e telefone e os enlaces dedicados via rádio foram duas das primeiras propostas apresentadas pela indústria para atender os usuários domésticos. Entretanto, estas eram alternativas de alto custo cuja viabilidade dependia muito do número de

⁴ Basic Rate Interface.

⁵ Primary Rate Interface.

| | Capacidade [kb/s] | | Distância [m] |
|-------------|-------------------|--------|---------------|
| | Descida | Subida | |
| HDSL | 768 | 768 | 4.000 |
| ADSL G.Lite | 1.544 | 512 | 1.500 |
| ADSL | 8.000 | 1.024 | 1.500 |
| ADSL2 | 12.000 | 1.024 | 2.500 |
| ADSL2+ | 24.000 | 1.024 | 1.500 |
| VDSL | 52.000 | 12.000 | 300 |

Tabela 1.1: Capacidade das principais variantes da tecnologia DSL.

usuários atendidos pelo mesmo enlace.

A solução mais eficaz foi proporcionada, de fato, por tecnologias que reutilizavam a rede física de telefonia ou de TV. A primeira delas, denominada acesso em linha de assinante digital (DSL⁶), utiliza o mesmo par metálico dos sinais telefônicos comuns para transmissão de dados. Em um acesso DSL, o usuário utiliza um modem para transmitir dados a um concentrador denominado Multiplexador de Acesso para Linha de Assinante Digital (DSLAM⁷) localizado na empresa de telefonia. A segunda tecnologia é baseada no uso da rede de distribuição de TV a cabo para transportar também sinais de dados, que também utiliza transmissão analógica. Esta tecnologia é conhecida pelo nome genérico de *cable modem*.

A principal vantagem da tecnologia DSL é o seu baixo custo de implantação, dado que os usuários normalmente já dispõem do par telefônico em suas residências ou escritórios. Contudo, esta tecnologia só é capaz de oferecer banda larga a pequenas distâncias. Os limites de banda e de distância para as principais variantes do DSL são apresentadas na Tabela 1.1 [2]. Para distâncias superiores àquelas indicadas, a banda disponível cai rapidamente. Por isso, as variantes de maior capacidade exigem que o DSLAM seja posicionado perto dos usuários, caso em que se empregam fibras ópticas ou cabos dedicados para o transporte dos sinais até a central telefônica. Vale observar que esta exigência tende a elevar os custos de implantação e a reduzir a escalabilidade da rede, dado que os DSLAM passam a ser compartilhados por um número menor de usuários.

A tecnologia de *cable modem* suplanta as restrições de distância impostas pelo DSL por meio de uma arquitetura híbrida de sinais elétricos e ópticos. Nesta arquitetura, denominada Híbrido de Cabo Coaxial e Fibra Óptica (HFC⁸), um grande número de residências compartilha o mesmo cabo coaxial de TV para transmitir dados. Este cabo funciona como um *hub* e é acoplado a uma fibra óptica que leva os sinais até o centro de distribuição da operadora. Tipicamente, a parte metálica do HFC suporta uma taxa de até 36Mb/s que é compartilhada por até dois mil usuários [3]. Desta maneira, a largura de banda disponível a cada usuário depende significativamente da multiplexação estatística dos fluxos na rede. O limite de banda imposto pela parte metálica do HFC representa um obstáculo significativo e que se torna ainda mais relevante quando se considera o custo elevado de implantação de um novo cabo HFC para ampliar a base de clientes.

1.2 REDES DE ACESSO ÓPTICAS

A partir das experiências com a RDSI-FE e das tecnologias subseqüentes, tanto a indústria como a academia concluíram que o fornecimento de serviços integrados a residências não poderia ser feito por meio de redes baseadas em condutores metálicos. Os principais motivos são:

- *custo elevado*. Além do elevado custo dos condutores, as redes metálicas requerem maior número de amplificadores de regeneração e, dependendo da arquitetura, multiplexadores ativos ao longo do percurso;

⁶ *Digital Subscriber Line*.

⁷ *Digital Subscriber Line Access Multiplexer*.

⁸ *Hybrid Fiber Coaxial*.

- *pouca largura de banda.* A baixa relação sinal/ruído e as perdas ao longo dos condutores limitam o produto entre a banda disponível e a distância entre as terminações da rede;
- *baixa escalabilidade.* A ampliação da capacidade de uma rede metálica geralmente requer a troca de equipamentos terminais e a instalação de novos condutores, o que pode ser inviável do ponto de vista econômico.

Diante destes fatores, tanto a indústria como a academia voltaram-se ao desenvolvimento de tecnologias capazes de substituir o uso condutores metálicos nas redes de acesso. Nesta linha, as redes de acesso ópticas surgiram como principal alternativa de acesso fixo. Dentre suas vantagens significativas, podemos destacar:

- *grande largura de banda.* A largura de banda disponível em redes ópticas é significativamente maior do que a das redes metálicas. No caso específico de redes de acesso, a capacidade típica é de 1Gb/s, que pode ser compartilhada ou não entre os usuários, dependendo da arquitetura escolhida. Redes de acesso com capacidade de 10Gb/s já se encontram em fase final de desenvolvimento;
- *baixo custo.* Os custos de implantação e de manutenção de redes ópticas são, no pior caso, comparáveis aos de uma rede metálica comum. No entanto, o retorno proporcionado pelo acesso óptico é, ao longo do tempo, superior ao das tecnologias DSL e HFC [4]. Além disso, a maior capacidade da rede torna o custo por Mb/s substancialmente menor. Vale observar que, dependendo da distância, as redes ópticas podem empregar apenas elementos passivos ao longo do percurso do sinal, reduzindo ainda mais os custos de implantação e de manutenção;
- *alta escalabilidade.* A tecnologia WDM permite utilizar a mesma fibra para transportar diversos fluxos simultaneamente, associando a cada um deles um comprimento de onda específico. Desta maneira, novos comprimentos de onda podem ser utilizados para aumentar a capacidade da rede praticamente sem qualquer modificação física;
- *domínio tecnológico.* Tecnologias ópticas como as redes ATM ópticas, o SDH e o SONET já são há muito utilizadas em redes metropolitanas e de grande abrangência, onde se requer elevada capacidade de transmissão. Novas tecnologias como o WDM, o DWDM⁹ e as redes gigabit Ethernet ópticas têm se difundido rapidamente, elevando ainda mais a capacidade dos enlaces.

1.2.1 ARQUITETURA FTTH

Hoje, as tecnologias DSL e HFC já podem empregar fibras ópticas para conectar os pontos de distribuição aos equipamentos de comutação das operadoras. Arquiteturas assim constituídas são conhecidas por fibra até o meio fio (FTTC¹⁰) ou fibra até o gabinete (FTTCAB¹¹) pois a fibra não chega até o usuário final.

Neste caso, não há pleno aproveitamento das vantagens da arquitetura óptica, dado que os problemas de escalabilidade e de custo apresentados pelo DSL e pelo HFC se mantêm. Para obter todas as vantagens da arquitetura óptica, é necessário levar a fibra até o usuário final. Esta arquitetura é conhecida como fibra até a residência (FTTH¹²). Termos como fibra até o negócio (FTTB¹³) e fibra até o local (FTTP¹⁴) são igualmente comuns na literatura. O acesso FTTH pode ser implementado de uma das seguintes maneiras:

⁹ Multiplexagem Densa por Divisão em Comprimento de Onda (*Dense WDM*).

¹⁰ *Fiber to the Curb.*

¹¹ *Fiber to the Cabinet.*

¹² *Fiber to the Home.*

¹³ *Fiber to the Business.*

¹⁴ *Fiber to the Premise.*

- *Ponto-a-multiponto com concentrador.* Nesta arquitetura, os equipamentos da operadora são conectados por fibra óptica a um concentrador remoto que, por sua vez, é conectado a cada usuário por uma fibra óptica independente. A principal vantagem é a economia de fibra, já que apenas um enlace é utilizado na maior parte do percurso do sinal. Além disso, é possível compartilhar este enlace entre acessos FTTH e por rede metálica. Contudo, o custo de implementação, operação e manutenção podem ser elevados, dada a necessidade do concentrador e de $2N$ transceptores para atender apenas N usuários.
- *Ponto-a-ponto.* Em uma rede óptica ponto-a-ponto, os equipamentos da operadora são conectados diretamente a cada usuário por meio de uma fibra óptica independente. A principal vantagem é a largura de banda oferecida a cada usuário, que pode ser aproveitada ao máximo caso o conteúdo multimídia seja fornecido pela mesma operadora que controla a rede. Para atender N usuários, esta arquitetura requer $2N$ transceptores e N enlaces ao longo de todo o percurso da rede, o que leva a maiores custos frente às demais arquiteturas.
- *Ponto-a-multiponto com elementos passivos.* Nesta arquitetura, também conhecida como PON¹⁵, os equipamentos da operadora são conectados à rede de acesso por uma fibra principal. Os usuários se conectam por meio de fibras dedicadas, acopladas por divisores e combinadores passivos. Assim, esta arquitetura requer apenas $N + 1$ transceptores para atender N usuários e utiliza uma única fibra ao longo de praticamente todo o percurso, que pode chegar a até 20km.

A principal vantagem das redes ponto-a-multiponto com concentrador e das redes ponto-a-ponto é o seu maior alcance, sobretudo quando se utilizam elementos ativos ao longo do percurso. Contudo, estudos indicam que as PON são capazes de cobrir distâncias suficientes para atingir quase a totalidade dos consumidores [5]. Diante de suas vantagens econômicas, a arquitetura PON é hoje considerada como sendo a alternativa ideal para implementação da maior parte das redes de acesso fixo de banda larga.

1.2.2 REDES ÓPTICAS COM ELEMENTOS PASSIVOS

As redes de acesso PON são redes ponto-multiponto que têm, em seu interior, apenas elementos ópticos unidirecionais e passivos tais como combinadores, divisores e acopladores ópticos. As transmissões dos fluxos de subida e de descida se fazem por meio de comprimentos de onda independentes, transportadas por fibras monomodo [3].

Um exemplo deste tipo de rede é apresentado na Figura 1.2. Toda transmissão ocorre entre um terminal de linha óptica (OLT¹⁶) e uma ou mais unidades de rede óptica (ONU¹⁷). A OLT é disposta junto ao provedor de serviços de rede e atua como ponto de acesso da rede óptica às redes de grande abrangência, tais como as metropolitanas ou a Internet. Por sua vez, as ONU são dispostas junto aos usuários finais e fornece-lhes acesso aos serviços da rede.

Na transmissão dos fluxos de descida, os pacotes enviados pela OLT passam por um divisor óptico passivo e são encaminhados a todas as ONU, que recebem apenas os pacotes que lhe são destinados. No fluxo de subida, as ONU competem pelo uso do meio compartilhado para transmitir os seus pacotes à OLT, sendo, portanto, necessário o uso de mecanismos de controle de acesso. Em geral, utiliza-se o acesso múltiplo por divisão no tempo (TDMA¹⁸), que hoje apresenta a relação entre complexidade, custo e eficiência mais favorável [6].

As PON apresentam vantagens significativas frente às tecnologias concorrentes [6, 7]. Podemos citar que:

¹⁵ *Passive Optical Network.*

¹⁶ *Optical Line Terminal.*

¹⁷ *Optical Network Unit.*

¹⁸ *Time Division Multiple Access.*

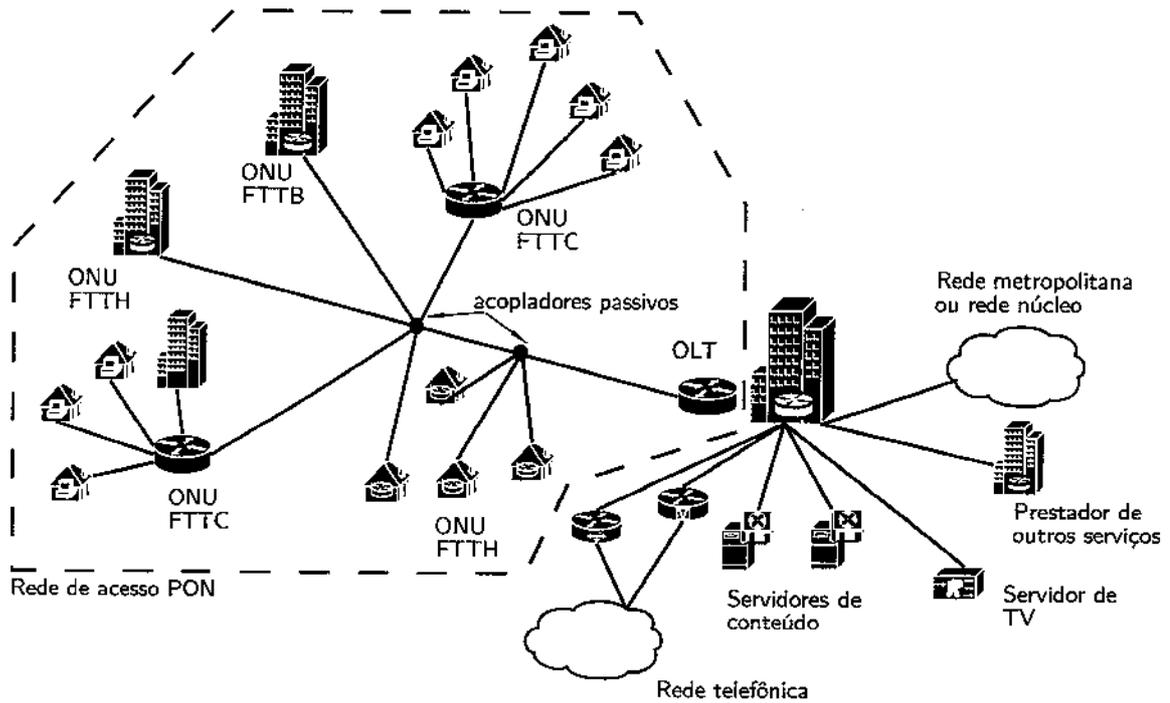


Figura 1.2: Exemplo de rede de acesso PON.

- podem prover maior banda aos usuários finais, pois utilizam, ao longo de todo o percurso, sinais ópticos que estão sujeitos a menor perda e dispersão do que os sinais elétricos utilizados no DSL e no HFC;
- têm alcance de até 20km, o que é suficiente para atingir 98% dos consumidores [5]. A tecnologia DSL, por outro lado, só é capaz de oferecer banda larga a pequenas distâncias que, no caso do VDSL, não passam de algumas centenas de metros;
- minimizam o uso de transceptores e de fibras entre os elementos terminais, o que reduz o seu custo frente às redes ópticas ponto-a-ponto;
- dispensam o uso de multiplexadores, demultiplexadores e transdutores nos pontos de acesso, o que reduz os custos de instalação, operação e manutenção frente às demais tecnologias;
- são escaláveis, sendo que se pode aumentar a capacidade do acesso e o número de usuários atendidos por meio da técnica WDM.

Vale observar que as PON podem ser ligadas em anel, barramento ou árvore, topologias estas que são as mais comuns em redes de acesso. As ONU podem ser instaladas no interior das residências e escritórios (arquitetura FTTH) ou em armários ópticos externos a que os usuários se conectam por meio de fibras independentes ou enlaces DSL (arquitetura FTTC).

As PON suportam ainda diferentes protocolos de camada de enlace. Atualmente, três propostas prevalecem na literatura:

- ATM PON ou *Broadband PON* (APON/BPON). Esta proposta foi definida pelo FSAN¹⁹, um comitê constituído por operadoras de telefonia européias, americanas e asiáticas. Em 1998, foi reconhecida pelo ITU-T por meio das recomendações ITU G.983. Transporta os dados em

¹⁹ Full Service Access Network committee.

| | BPON | GPON | EPON |
|------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Padronização | ITU-T G.983 | ITU-T G.984 | IEEE 802.3 |
| Enlace | ATM | SDH/SONET | Ethernet |
| Descida | 1,25Gb/s (máx.) | 2,5Gb/s (máx.) | 1,25Gb/s (máx.) |
| Subida | 622Mb/s (máx.) | 2,5Gb/s (máx.) | 1,25Gb/s (máx.) |
| Número de ONUs | 32 (típ.) 64 (máx.) | 64 (típ.) 128 (máx.) | 16 (típ.) 32 (máx.) |
| Número de fluxos | 256 | 4096 | variável |
| Acesso | TDMA | TDMA | TDMA |
| Eficiência | ≈80% | ≈90% | ≈50% |
| Laser | 1490nm (descida) 1310nm (subida) | 1490nm (descida) 1310nm (subida) | 1550nm (descida) 1310nm (subida) |
| Interface | ATM | ATM, Ethernet e TDM | Ethernet |
| TDM | TDM sobre ATM | Nativo | TDM sobre Ethernet |
| Alcance | 20km | 20km entre ONUs 60km (máx.) | 20km |
| Criptografia | Sim | Sim | Não |
| Custo | Baixo | Médio | Muito baixo |

Tabela 1.2: Principais características dos padrões PON.

células ATM e tem capacidade total de até 1,25Gb/s no fluxo de descida e de 622Mb/s no fluxo de subida. Prevê um comprimento de onda exclusivo para RF e difusão de vídeo (TV e HDTV). Define ainda um algoritmo de criptografia para garantir a segurança dos fluxos de descida.

- *Gigabit PON (GPON)*. Esta proposta, cujo objetivo era ampliar a capacidade das redes BPON, foi definida pelo FSAN entre 2001 e 2003. Foi reconhecida pelo ITU-T por meio das recomendações ITU G.984. Utiliza SONET/SDH na camada de enlace e o mecanismo de encapsulamento para transportar pacotes de outros protocolos, incluindo Ethernet. Atinge uma capacidade de até 2.5Gb/s tanto no fluxo de descida como no de subida. Também aplica criptografia nos fluxos de descida e prevê um comprimento de onda exclusivo para RF e difusão de vídeo (TV e HDTV).
- *Ethernet PON (EPON)*. Esta proposta foi definida pela força tarefa IEEE 802.3ah “Ethernet na primeira milha” em 2001. Foi incorporada às normas IEEE 802.3. Utiliza o protocolo gigabit Ethernet na camada de enlace, sendo as transmissões feitas em full-duplex. Atinge uma capacidade de até 1,25Gb/s tanto no fluxo de descida como no de subida. Implementações de 10Gb/s estão em estudo [8]. Não prevê criptografia para os fluxos, nem enlaces exclusivos para RF e difusão de vídeo.

As principais características de cada tecnologia estão relacionadas na Tabela 1.2 [9–11]. Outra proposta relevante que já encontra lugar na literatura é a PON baseada em WDM (WDM-PON) [9]. Nesta rede, cada ONU utiliza um comprimento de onda independente para transmissão e, desta maneira, pode utilizar toda a capacidade do canal. Aplicações experimentais indicam que uma WDM-PON pode chegar a 10Gb/s para cada ONU [9, 12]. Apesar das altas taxas obtidas, o uso de WDM em PON requer transceptores mais sofisticados, o que hoje a torna inviável. Abordagens híbridas que empregam simultaneamente WDM e TDMA têm sido consideradas como forma de reduzir os custos [9, 12].

Implementações comerciais de redes de acesso PON já estão em operação sobretudo na América do Norte e Ásia [3, 9, 10, 13]. A maior parte delas é baseada no padrão APON/BPON, o que se deve, em grande medida, ao fato de as operadoras terem se envolvido diretamente no seu desenvolvimento. Hoje, contudo, o interesse por este padrão tem caído devido ao seu elevado custo frente às demais

alternativas. Ademais, o transporte sobre ATM de pacotes IP, protocolo ubíquo no tráfego, requer camadas de adaptação que reduzem a eficiência da rede.

Para novas implementações, a indústria está dividida entre o padrão GPON e o EPON. Ambos são capazes de transportar pacotes IP. O padrão GPON tem hoje a vantagem de prover maior banda, de transportar múltiplos protocolos e de oferecer melhor suporte a garantias de desempenho. O padrão EPON, por outro lado, conta com a simplicidade, a maturidade e a ampla aceitação do protocolo Ethernet, que hoje predomina nas redes locais. As garantias de desempenho, antiga deficiência no protocolo Ethernet, começam a ser contempladas pelos padrões IEEE [14, 15]. Estudos voltados à implementação de EPON de 10Gb/s tendem aumentar ainda mais a sua competitividade frente à tecnologia GPON [8].

Dado que a maior parte dos estudos acadêmicos em redes PON se voltam ao padrão EPON, e o apontam como sendo bastante promissor, concentraremos o presente trabalho nesta tecnologia.

1.3 REDES ETHERNET PON

Como em toda PON, as transmissões em uma EPON se realizam exclusivamente entre a OLT e as ONU. Os dados são transmitidos em quadros Ethernet compatíveis com a norma IEEE 802.3, sendo que os 8 bytes de preâmbulo são modificados para incluir um campo de 2 bytes (6° e 7° bytes) destinado a identificar a ONU envolvida na transmissão de cada quadro e a sinalizar transmissões em difusão (*broadcasting* ou *multicasting*). Este campo é denominado Identidade Lógica de Enlace (LLID²⁰) e é definido pela OLT durante a inicialização ou quando uma nova ONU é adicionada à rede [16]. Dois outros bytes do preâmbulo são modificados: um para marcar o início do campo LLID (SLD, 3° byte do preâmbulo, de valor constante igual a 0xd5) e outro de verificação de integridade deste mesmo campo (CRC8, 8° byte do preâmbulo).

As transmissões em uma EPON se fazem, então, da seguinte maneira. Para os fluxos de descida, cuja transmissão é ilustrada na Figura 1.3, os quadros enviados pela OLT são encaminhados a todas as ONU, que selecionam os seus com base no campo LLID. Caso o bit mais significativo deste campo (chamado *modo de recepção*) seja igual a 0, os demais bits do campo identificam uma única ONU destino. Neste caso, apenas a ONU com o mesmo LLID deve aceitar o quadro. Caso o modo de recepção seja igual a 1, todas as ONU cujo valor de LLID seja diferente daquele indicado no quadro devem aceitá-lo. O valor de LLID 0x7fff com modo de recepção igual a 1 é reservado para sinalizar transmissões em *broadcasting* [16].

Para os fluxos de subida, o acesso das ONU ao meio compartilhado é arbitrado por TDMA, como mostra a Figura 1.4. Neste caso, cada ONU pode transmitir apenas durante janelas de tempo exclusivas, que são alocadas pela OLT por meio de protocolo específico. A OLT também aloca janelas de tempo específicas para a detecção e registro de novas ONU na rede. Vale observar que a OLT também utiliza o campo LLID na recepção, recebendo apenas os quadros de ONU consideradas como em operação normal ou em fase de registro [16].

1.3.1 PROTOCOLO MPCP

Em redes Ethernet tradicionais, o controle de acesso é baseado no protocolo CSMA/CD. Este protocolo é de difícil implementação em redes EPON, dado que o uso de elementos ópticos unidirecionais e o elevado produto entre a largura de banda e o tempo de propagação dos sinais entre a OLT e as ONU tornam a detecção de colisões pouco eficiente.

A alternativa adotada pela norma IEEE 802.3ah para controlar o acesso das ONU ao canal de subida é o TDMA [6, 7, 17]. Em redes EPON, esta técnica de acesso é implementada por meio de um protocolo denominado Protocolo de Controle Multiponto (MPCP²¹). Neste protocolo, a OLT utiliza uma mensagem de permissão para liberar o acesso de uma ONU ao meio compartilhado, e as ONU utilizam mensagens de requisição para enviar a sua demanda ao algoritmo de alocação de banda

²⁰ Logical Link ID.

²¹ Multipoint Control Protocol.

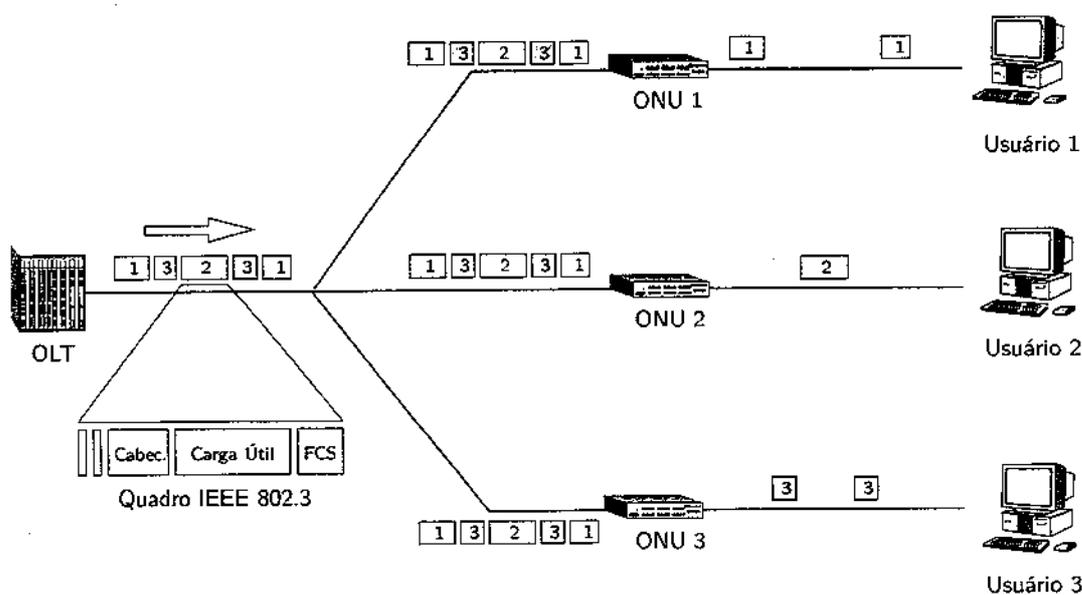


Figura 1.3: Transmissão de fluxos de descida em uma rede EPON.

da OLT. Estas mensagens são denominadas *gate* e *report*. O protocolo MPCP ainda define outras três mensagens para o processo de busca e registro de uma nova ONU na rede (*register_request*, *onu_register* e *register_ack*). Todas estas mensagens são encapsuladas em quadros Ethernet pela camada de enlace e transmitidas no mesmo canal em que fluem os pacotes de dados [17].

Por meio da mensagem *gate*, a OLT informa a uma dada ONU o instante em que se inicia a sua próxima janela de transmissão e a sua duração. Esta informação é denominada permissão (*grant*). Cada mensagem *gate* pode levar até seis permissões à mesma ONU, cada qual composta por quatro octetos que indicam o instante em que se inicia uma janela de transmissão (*grant start time*) e dois que indicam a sua duração (*grant length*). Há ainda um octeto destinado à seleção de filas dentro da ONU (*grant level*), que pode ser utilizado pela OLT para controlar também o escalonamento dos fluxos dentro da ONU.

A ONU também utiliza a mensagem *gate* para sincronizar o seu relógio interno com o da OLT, o que é feito com base no seu tempo de envio. A partir do início da sua janela de transmissão, a ONU passa a enviar os quadros Ethernet com os pacotes dos usuários. As transmissões são feitas de modo a garantir que não haja fragmentação de quadros, postergando o envio daqueles que não caibam na janela de transmissão corrente.

A mensagem *report* é utilizada pela ONU para informar a sua demanda à OLT, especificando o tamanho desejado para a próxima janela de transmissão. A demanda deve levar em conta, além dos dados, o preâmbulo de 64bits, o intervalo entre quadros de 96bits e qualquer outra informação necessária à transmissão. Cada mensagem *report* pode levar consigo a demanda de até oito filas distintas, que podem corresponder a diferentes usuários, sessões ou classes de serviço. Em geral, a mensagem *report* é enviada automaticamente junto com os quadros Ethernet correspondentes aos dados.

A informação de demanda recebida pela OLT é utilizada para determinar a divisão da capacidade da rede entre as ONU. A mensagem *report* ainda é utilizada para calcular o tempo de ida e volta (RTT²²) da ONU, que é utilizado no processo de sincronismo de relógio como na detecção de falhas. Variações significativas no RTT de uma mesma ONU sugerem problemas na transmissão ou de sincronismo, o que requer ações corretivas para prevenir a degradação do serviço oferecido aos usuários das outras ONU.

²² *round-trip time*.

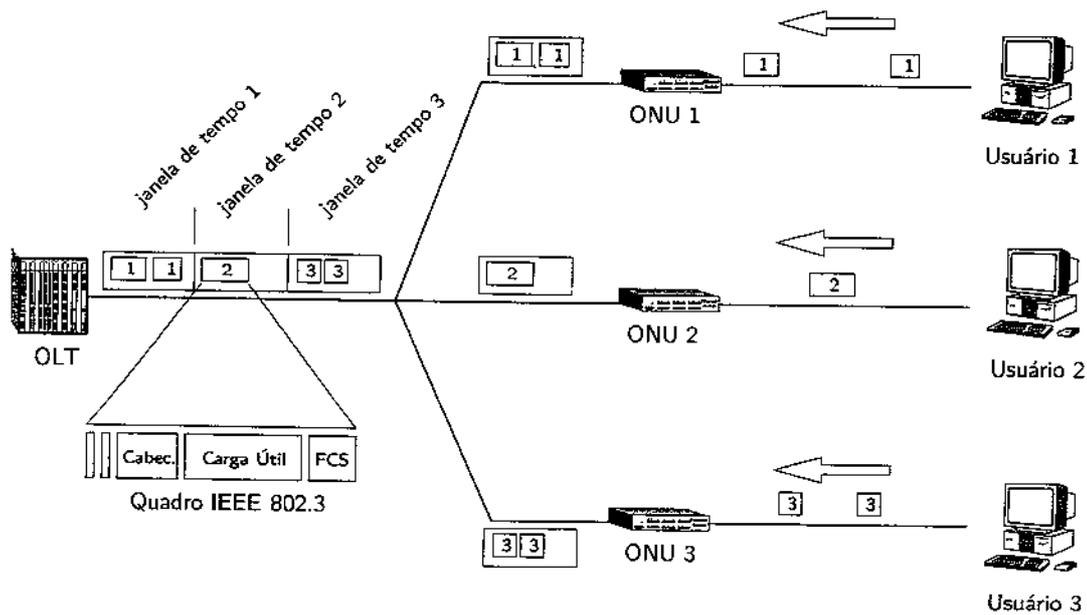


Figura 1.4: Transmissão de fluxos de subida em uma rede EPON.

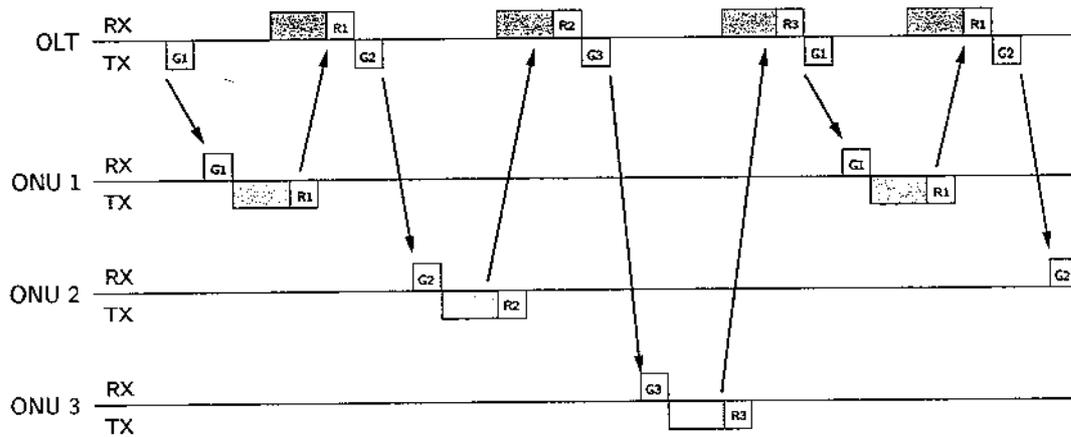
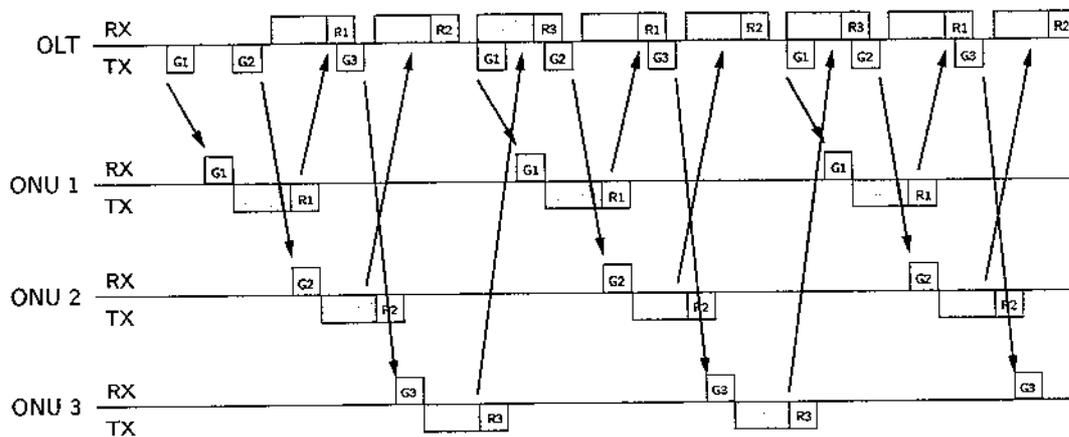
1.3.2 ACESSO POR POLLING

A definição do protocolo MPCP sugere que as EPON não utilizam uma política de acesso TDMA tradicional. Embora mecanismos de sincronização sejam utilizados para manter um relógio comum na rede, o acesso das ONU ao meio compartilhado é, de fato, assíncrono e coordenado pela OLT. O uso do relógio comum se restringe quase que somente a permitir que a OLT possa antecipar a alocação das janelas de transmissão de cada ONU.

Na literatura, a maioria dos estudos não utiliza o mecanismo de relógio comum, dado o custo computacional de se assegurar o perfeito sincronismo entre a OLT e as ONU. Considera-se, como alternativa, que a janela de transmissão de uma dada ONU se inicia logo após o recebimento da mensagem gate correspondente, o que caracteriza um acesso por *polling*.

A estratégia de *polling* utilizada pelo controle de acesso tem influência direta no aproveitamento da capacidade da rede. A literatura apresenta basicamente três estratégias que podem ser implementadas por meio do protocolo MPCP. A mais simples delas é denominada *poll-and-stop* e é ilustrada na Figura 1.5. Nesta estratégia, a OLT inicia a transmissão enviando uma mensagem gate a uma das ONU. A ONU acionada, ao receber esta mensagem, envia os pacotes correspondentes à permissão recebida. Ao final da transmissão, a ONU envia uma mensagem report com a informação de demanda que a OLT utilizará para definir o tamanho da próxima janela de transmissão. Somente após o recebimento desta mensagem a OLT envia uma nova mensagem gate à próxima ONU. Ao final da transmissão da última ONU, a OLT redimensiona as janelas de transmissão com base nas informações de demanda recebidas, e o processo se reinicia.

Note que a alocação de janelas na estratégia *poll-and-stop* é feita com base na informação de demanda de todas as ONU, o que torna possível distribuir a banda de maneira mais eficiente e justa. Apesar disso, essa estratégia não propicia o aproveitamento eficiente da capacidade do canal de subida, que se mantém ocioso em grande parte do tempo. Isto reduz a banda disponível para as transmissões e aumenta o atraso dos pacotes. Esta desvantagem pode ser contornada por meio da estratégia de *polling* intercalado [18, 19], que é ilustrada na Figura 1.6. Nesta estratégia, a OLT pode enviar uma mensagem gate a uma ONU antes que se recebam os dados e as mensagens report das ONU anteriores. Note que esta estratégia é factível pois os canais de subida e de descida são separados, de modo que a mensagem gate pode ser transmitida pela OLT ao mesmo tempo em que

Figura 1.5: Estratégia *poll-and-stop*.Figura 1.6: Estratégia de *polling* intercalado.

dados são recebidos das ONU. Ademais, o *polling* intercalado pressupõe que a OLT dispõha do RTT e da demanda de cada ONU, para que seja possível determinar o instante em que as mensagens gate devem ser transmitidas para que os dados de uma ONU cheguem à OLT logo após o fim da transmissão da ONU anterior, sem que haja colisão. Neste caso, maximiza-se a utilização do canal, que fica ocioso apenas por pequenos intervalos de tempo entre as transmissões de cada ONU. Estes intervalos, denominados intervalos de guarda, são necessários para evitar colisões devido a variações no tempo de processamento e a pequenas flutuações no RTT das ONU, e para que a OLT possa reajustar a sensibilidade dos seus receptores [19].

Em [18, 19], demonstra-se que a estratégia de *polling* intercalado efetivamente aumenta o desempenho da rede em termos de utilização do canal e de atraso médio dos pacotes. Apesar disso, esta estratégia tem uma desvantagem importante. Ao alocar a janela a uma dada ONU, a OLT não dispõe da demanda de todas as ONU, mas apenas daquelas que já puderam transmitir os seus dados. Com isso, torna-se mais difícil distribuir a banda de maneira eficiente e justa, sobretudo se não for possível prever a demanda das ONU com precisão.

Uma maneira de contornar a desvantagem do *polling* intercalado é utilizar uma variação denominada *polling* intercalado com parada [20]. Nesta estratégia, que é ilustrada na Figura 1.7, a OLT não reinicia o ciclo de *polling* antes que todas as ONU tenham enviado seus dados [20]. Desta maneira, a OLT dispõe da informação de todas as ONU para realizar a alocação das janelas do ciclo seguinte.

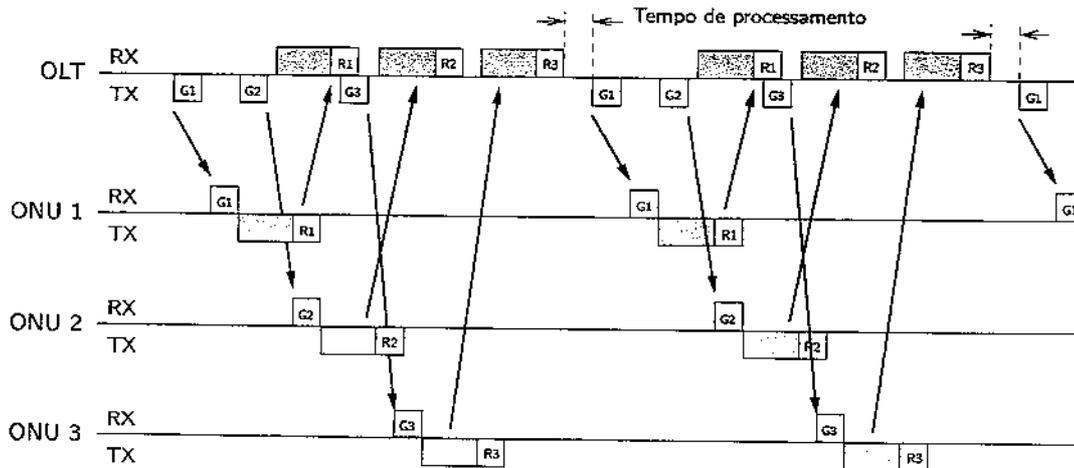


Figura 1.7: Estratégia de *polling* intercalado com parada.

Na Figura 1.7, pode-se observar que o *polling* intercalado é mais eficiente do que o *polling* intercalado com parada, em que não se pode antecipar o envio da primeira mensagem gate do ciclo. Com isso, o canal se mantém ocioso entre os ciclos, por um tempo igual ao tempo de processamento do algoritmo de alocação somado ao RTT da primeira ONU. A menor utilização do canal constitui, então, a contrapartida à disponibilidade da demanda de todas as ONU no momento da alocação de banda para o ciclo subsequente.

É possível reduzir o tempo ocioso entre ciclos, assim como os demais tempos de passagem, ordenando-se as ONU de maneira apropriada. A abordagem mais simples é atendê-las em ordem crescente de RTT, o que permite antecipar a mensagem gate das ONU mais distantes. É fácil ver, contudo, que a ordem ótima depende não só do RTT das ONU, mas também do tamanho da janela a ser alocada a cada uma delas. Assim, a ordem de atendimento ótima pode variar ao longo dos ciclos, e obtê-la pode ser computacionalmente custoso.

1.3.3 SINALIZAÇÃO INTERPOSTA

Em [18, 19], Kramer et al. argumentam que as mensagens gate levam mais informação do que a necessária para o controle do acesso ao canal de subida. Supondo que a OLT não realiza diretamente o escalonamento dos fluxos internos de cada ONU e que as janelas de transmissão se iniciam imediatamente após a recepção das mensagens gate, esta mensagem precisa levar apenas a duração da janela que será alocada. O mesmo vale para as mensagens report, que precisam levar apenas a demanda global da ONU, e não de cada fluxo ou grupo de fluxos. Em ambos os casos, a carga útil das mensagens de controle é inferior aos 48 bytes mínimos transportados por um quadro Ethernet, o que evidencia desperdício de banda. Este desperdício é ainda mais relevante quando uma ou mais ONU têm pouco tráfego a transmitir, condição em que o fluxo de mensagens de controle se torna mais intenso.

Note ainda que as mensagens de controle não precisam levar o endereço MAC da OLT, já que este elemento sempre está envolvido nas transmissões. Ademais, o endereço MAC da ONU é substituído por uma identidade lógica, representada por um único byte. Isto torna ainda maior o desperdício de banda decorrente do encapsulamento daquelas mensagens.

O encapsulamento também é desvantajoso quando o tráfego de descida é muito intenso. Dado que essas mensagens concorrem com o tráfego no canal, uma mensagem gate pode ser bloqueada pela transmissão de um pacote muito longo. Isto retarda a transmissão da ONU correspondente, reduzindo a utilização do canal. Esta dificuldade poderia ser contornada se as transmissões de controle fossem feitas em um canal exclusivo, solução que elevaria a complexidade e o custo da rede.

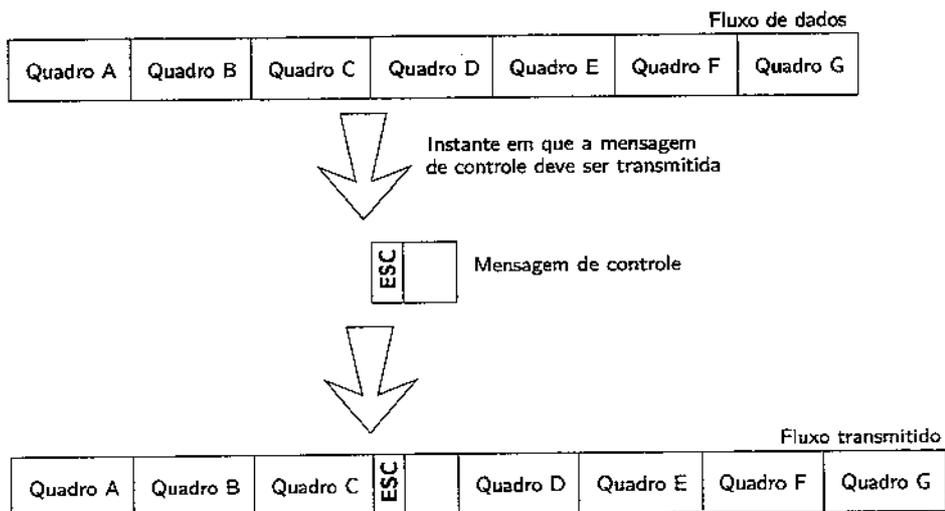


Figura 1.8: Funcionamento da sinalização interposta.

A solução apontada por Kramer et al. é permitir que as mensagens de controle sejam transportadas sem o encapsulamento Ethernet. Isto é feito interpondo-se as mensagens nos fluxos por meio de seqüências de escape [18, 19]. As redes EPON utilizam uma codificação 8B/10B para detecção de erros. Códigos de 10 bits sem correspondência poderiam, então, ser utilizados como seqüência de escape para marcar o início das mensagens de controle dentro do fluxo. Esta abordagem, denominada sinalização interposta, é ilustrada na Figura 1.8. Neste caso, a camada de enlace da OLT ou da ONU extrai a informação de controle ao receber um destes códigos, sem afetar os demais bytes do fluxo. As mensagens podem, portanto, ser interpostas entre quadros Ethernet diferentes ou mesmo entre os bytes de um mesmo quadro, o que praticamente elimina os problemas apontados.

Embora a sinalização interposta tenha sido definida considerando que a OLT não toma parte do escalonamento intra-ONU, é possível utilizá-la também no caso oposto. Para que isso seja possível, a OLT deve ser capaz de enviar às ONU informações que lhes permitam determinar, indiretamente, a demanda global da rede e o tamanho da janela a ser alocado a cada um dos seus fluxos. A OLT, por sua vez, deve ser capaz de antever a janela que será utilizada pela ONU sem dispor da demanda individual de cada fluxo. Um protocolo deste tipo foi proposto por Kramer et al. em [21], sendo que a abordagem apresentada requer mensagens de controle tão longas quanto aquelas especificadas pelo protocolo MPCP. No Capítulo 4 do presente trabalho, apresentamos um protocolo mais eficiente, que requer a transmissão de apenas um número inteiro pela mensagem gate e de dois pela mensagem report, podendo então ser implementado com maior eficiência por meio de sinalização interposta.

1.3.4 QUALIDADE DE SERVIÇO EM REDES ETHERNET

As normas da família IEEE 802 definem diretrizes para o suporte a garantias de qualidade de serviço em redes Ethernet. Neste caso, adota-se a abordagem baseada em classes de serviço em que fluxos com requisitos de banda e de desempenho semelhantes são agrupados em um número limitado de classes, que recebem tratamento diferenciado ao longo da rede. Este tratamento é conhecido como comportamento do tráfego por nó (PHB²³).

Segundo a norma IEEE 802.1d, as pontes (no caso as ONU) devem segregar o fluxo em classes de serviço, atendendo-as segundo uma política de prioridade estrita [14]. Neste caso, destina-se a cada classe uma fila exclusiva, que recebe os quadros dos fluxos correspondentes em ordem de chegada. Esses mesmos quadros são selecionados para transmissão em ordem tal que o tráfego de

²³ *per-hop behavior*.

uma dada classe só é atendido quando não houver tráfego em espera nas filas correspondentes às classes de maior prioridade.

Para suportar a política baseada em prioridades, a norma IEEE 802.1q introduz uma extensão que permite ao quadro Ethernet levar a informação sobre a prioridade dos pacotes no seu cabeçalho [15, 22]. A norma IEEE 802.1d prevê as seguintes classes de serviço, em ordem de prioridade:

- *Controle de rede*. Tráfego de máxima prioridade, destinado a operação e manutenção da rede;
- *Voz*. Requer um atraso inferior a 10ms, sendo que a rede deve suportar jitter significativo;
- *Video*. Requer um atraso inferior a 100ms;
- *Carga controlada*. Aplicações que requerem reserva de banda, estando sujeitas a alguma forma de controle na admissão;
- *Excelente esforço*. Tráfego de melhor esforço, de maior prioridade;
- *Melhor esforço*. Tráfego de LAN tradicional;
- *Segundo plano*. Tráfego sem requisitos de desempenho, que pode utilizar recursos da rede desde que não afete o desempenho percebido pelas demais aplicações.

A norma IEEE 802.1d faculta o agrupamento das classes de serviço de acordo com o número de filas independentes disponíveis. Este agrupamento é apresentado em [22, Tabela 1]. Finalmente, cabe destacar que a norma IEEE 802.1d permite que se utilizem políticas de prioridade não-estrita e algoritmos adicionais para o escalonamento, desde que não se inverta a transmissão de quadros de igual origem, destino e prioridade.

1.3.5 PENALIZAÇÃO DE BAIXA CARGA

Além da estagnação dos fluxos de menor prioridade, uma outra desvantagem que decorre do uso de políticas de prioridade estrita em redes EPON é conhecida por penalização de baixa carga [22]. Sob baixa carga, o volume de tráfego que chega a ONU entre o envio de uma mensagem report e a recepção da mensagem gate seguinte torna-se significativo frente à demanda informada. Quando este tráfego tem alta prioridade, ele se apropria da banda que seria alocada aos fluxos de menor prioridade, modificando a alocação prevista. Com isso, o tráfego de menor prioridade tende a sofrer maior atraso, que pode chegar a vários ciclos. Ao mesmo tempo, apropriação de banda pelo tráfego de maior prioridade ainda tende a causar um resíduo não utilizado no final das janelas de transmissão.

Em [22], Kramer et al. propõem o uso de um sistema de enfileiramento em dois estágios para eliminar a penalização de baixa carga. Neste sistema, a ONU classifica os quadros recebidos dos usuários segundo a prioridade e os armazena em filas individuais, que compõem o primeiro estágio. Quando a ONU recebe da OLT a mensagem de *polling*, os quadros armazenados nessas filas são encaminhados a uma fila única em ordem estrita de prioridade. A ONU então transmite os quadros desta fila, mantendo armazenados aqueles que não cabem na janela de transmissão corrente. Estes quadros serão os primeiros a serem transmitidos no ciclo de *polling* seguinte, o que garante a oportunidade de serviço ao tráfego de baixa prioridade mesmo quando o tráfego de alta prioridade é intenso. Com isso, elimina-se, além da penalização de baixa carga, o problema de estagnação de fluxos inerente a abordagem de prioridade estrita. Em contrapartida, aumenta-se a latência e o atraso dos fluxos de maior prioridade.

Outra alternativa apontada pelos autores é utilizar um sistema de créditos, em que se tenta prever o tráfego de alta prioridade adicional e alocar-lhe os recursos necessários. Embora esta abordagem seja capaz de mitigar a penalização, ela requer o conhecimento prévio das características do tráfego, o que nem sempre é possível.

1.3.6 ESCALONAMENTO EM REDES EPON

Em uma rede de dados, cada servidor deve estabelecer a ordem em que os pacotes de cada fluxo são transmitidos, definindo, assim, a maneira pela qual se compartilham os recursos disponíveis. Esta ordem é estabelecida por um algoritmo ou um conjunto de regras que é denominado política de escalonamento.

No caso das redes EPON, o escalonamento envolve apenas ações sobre os fluxos de subida dado que os fluxos de descida são transmitidos por difusão. A hierarquia que existe entre a OLT e as ONU divide o escalonamento em duas etapas. Na primeira, denominada escalonamento inter-ONU, a OLT aloca janelas de transmissão a cada ONU com base na demanda agregada de seus fluxos. Obviamente, o tamanho das janelas é, frente ao tamanho do ciclo de *polling*, proporcional à parcela de banda destinada a cada ONU. A segunda etapa do escalonamento corresponde à divisão da banda recebida pela ONU entre os seus fluxos individuais. Esta etapa é denominada escalonamento intra-ONU.

A norma IEEE 802.3ah, que define como a OLT deve controlar o acesso das ONU ao canal de subida, não estabelece que políticas a rede deve adotar para realizar o escalonamento inter-ONU e o escalonamento intra-ONU. Na literatura, há autores que consideram que a rede de acesso se comporta como uma ponte entre a rede local do usuário e as redes de grande abrangência e que, por isso, deve respeitar a abordagem por classes de serviço e prioridade estrita que é proposta pela norma IEEE 802.1d [22, 23]. Esta abordagem foi apresentada na Seção 1.3.4.

A literatura apresenta diversas propostas de políticas que são baseadas na abordagem por classes de serviço de prioridade estrita. Apesar de propiciar maior escalabilidade, este tipo de política pode levar à estagnação dos fluxos de baixa prioridade e à penalização de baixa carga. Além disso, a abordagem por classes de serviço não permite oferecer garantias de desempenho individuais aos fluxos, que podem ser prejudicados pelo tráfego de outros fluxos da mesma classe. Atualmente, a única política de escalonamento voltada ao tratamento individualizado dos fluxos foi proposta por Kramer et al. [21] e é discutida, de maneira resumida, no Capítulo 5 do presente trabalho.

1.4 PROPOSTA DESTE TRABALHO

Diante do que foi exposto neste capítulo, podemos resumir o atual cenário das redes de acesso em banda larga nos seguintes pontos:

- é crescente a demanda por serviços integrados de voz, dados e vídeo;
- as operadoras consideram que a oferta destes serviços é essencial para aumentar a lucratividade do negócio;
- as redes de acesso PON são consideradas como a tecnologia mais apropriada de acesso fixo, dada a largura de banda requerida pelos serviços multimídia;
- as tecnologias GPON e EPON disputam pelas novas implementações, sendo crescente o interesse pela tecnologia EPON;
- com o desenvolvimento de redes EPON de 10Gb/s, a principal deficiência desta tecnologia frente a GPON passa a ser o limitado suporte à garantias de QoS, especialmente no se que relaciona ao escalonamento dos fluxos de subida.

Como em toda rede de dados, a escolha de uma política apropriada para o escalonamento dos fluxos em uma rede EPON é essencial para que se possa garantir a QoS dos usuários [24]. Apesar disso, nenhuma política foi estabelecida pela norma IEEE 802.3ah. Isto mantém o escalonamento em EPON como um tema de intensa pesquisa tanto na indústria como no meio acadêmico e constitui a principal motivação deste trabalho.

O nosso principal objetivo é estabelecer uma política que seja capaz de garantir banda mínima para os fluxos e de distribuir a banda ociosa entre si segundo a sua prioridade. Esta política, que

denominamos Compartilhamento Proporcional com Reserva de Carga (PSLR²⁴), tem como fundamentos a:

- *orientação fluxo-a-fluxo*, para que os requisitos de desempenho sejam atingidos individualmente e não de forma agregada em classes;
- *eficiência*, para utilizar ao máximo o uso dos recursos da rede;
- *proteção*, para assegurar o desempenho contratado pelos usuários independente de flutuações de tráfego na rede, dos excessos de tráfego gerado por outros usuários e do volume de tráfego de melhor esforço;
- *justiça*, para assegurar que os usuários são atendidos segundo os requisitos contratados e que a relação de prioridade é respeitada independente da sua localização geográfica ou da ONU a que pertençam;
- *flexibilidade*, para suportar diferentes requisitos de desempenho e tratar diferentes perfis de tráfego;
- *simplicidade*, para que tenha um baixo custo computacional;
- *escalabilidade*, para que seja possível ampliar o número de usuários com o mínimo custo; e
- *analiticidade*, para que seja possível relacionar os requisitos de desempenho aos recursos da rede que sejam necessários para satisfazê-los.

No caso de redes com garantias de QoS, essas expressões constituem um instrumento fundamental para determinar se um novo fluxo pode ou não ser aceito pela rede e para destinar-lhe recursos suficientes (banda e prioridade) para que seus requisitos de desempenho sejam atendidos. Em geral, estes requisitos são dados em termos de valores máximos de atraso, de variação de atraso e de taxa de perda de pacotes.

Neste trabalho nos concentramos no requisito de atraso introduzido nos fluxos. Na literatura, diversos estudos são voltados a relacionar este atraso aos parâmetros de tráfego dos fluxos e aos recursos de rede a eles destinados. Consideramos aqui os estudos desenvolvidos por Stiliadis e Varma, que apresentam limitantes supondo que o tráfego é policiado na entrada da rede pelo algoritmo do balde furado e que os fluxos são atendidos por uma política pertencente a uma classe denominada classe dos servidores latência-taxa [25]. Demonstramos que a política PSLR pertence à classe dos servidores latência-taxa e obtemos uma expressão limitante para o atraso com base nos referidos estudos.

Vale dizer que o policiamento do tráfego em redes de serviços integrados não é trivial. Isto se deve à natureza fractal do tráfego multimídia, comprovada por diversos estudos [26–36]. Para contemplar esta característica no limitante de atraso, propomos, primeiro, um modelo de tráfego capaz de representar o tráfego fractal com acurácia. A partir dele, estudamos o problema do policiamento de tráfego fractal por meio do algoritmo do balde furado. Estabelecemos então a relação entre os parâmetros do modelo e os do policiamento, o que permite obter um limitante específico para o atraso de fluxos com tráfego auto-similar sob a política PSLR.

Em resumo, consideramos que as contribuições deste trabalho são:

- Apresentar uma política de escalonamento para os fluxos de subida de uma rede EPON capaz de garantir banda mínima e de atender os fluxos individuais de maneira justa;
- Definir uma implementação computacional desta política;
- Demonstrar que a política pertence à classe dos servidores latência-taxa e obter uma expressão limitante para o atraso introduzido pela rede, supondo que o tráfego é policiado pelo algoritmo do balde furado;

²⁴do inglês *Proportional Sharing with Load Reservation*.

- Estudar a natureza fractal do tráfego multimídia e estabelecer um modelo capaz de representá-lo com acurácia e a partir de poucos parâmetros;
- Estudar o policiamento de tráfego fractal, estabelecendo as relações entre os parâmetros do modelo de tráfego proposto e os do algoritmo do balde furado;
- Obter um limitante específico para o atraso de fluxos com tráfego auto-similar sob a política proposta.

O restante deste trabalho está organizado da seguinte maneira. No Capítulo 2, discutimos a natureza fractal do tráfego em redes e apresentamos alguns dos principais modelos disponíveis na literatura. Propomos ainda um novo processo envelope denominado Processo de Chegada com Limitante Fracionário (FBAP²⁵) para representar o tráfego auto-similar. Desenvolvemos a análise de sistemas de filas para o tráfego FBAP obtemos limitantes para a latência e para o atraso introduzido nos fluxos quando o serviço tem taxa constante. Apresentamos ainda a banda efetiva do tráfego FBAP. Finalmente, os resultados são estendidos para o caso do tráfego multifractal.

No Capítulo 3, tratamos do problema do policiamento de tráfego fractal. Analisamos o policiamento deste tráfego por meio do algoritmo do balde furado fractal proposto por [37] e do balde furado tradicional. Demonstramos que o balde furado fractal não limita a explosividade do tráfego e, por isso, pode não permitir oferecer garantias de desempenho aos fluxos. Propomos então o uso do balde furado para o policiamento do tráfego FBAP, o que é ilustrado por meio de exemplos baseados em simulação.

No Capítulo 4, definimos a política de escalonamento PSLR. Propomos a sua implementação por meio de um algoritmo de baixa complexidade, cuja convergência é demonstrada analiticamente. Demonstramos também que o algoritmo é capaz de promover o compartilhamento justo dos recursos entre fluxos individuais, independente da ONU a que pertençam.

Ainda no Capítulo 4, analisamos o desempenho da política PSLR. Obtemos as condições de estabilidade da rede e dos fluxos individuais. Apresentamos ainda expressões limitantes para a latência e o atraso dos fluxos, pressupondo que o tráfego é policiado pelo algoritmo do balde furado.

No Capítulo 5, apresentamos uma revisão das principais políticas de escalonamento EPON disponíveis na literatura. Dentre elas, apenas a política FQSE proposta por Kramer et al. em [21] tem os mesmos objetivos da PSLR. Demonstramos, nesse capítulo, que menor complexidade computacional, a maior justiça e a existência de limitantes de desempenho são as principais vantagens do PSLR frente ao FQSE.

Finalmente, no Capítulo 6 são apresentadas as conclusões deste trabalho.

²⁵do inglês *Fractional Bounded Arrival Process*.

CAPÍTULO 2

MODELAGEM DE TRÁFEGO EM REDES

Estudos recentes indicam que o tráfego em redes apresenta propriedades fractais [26–36, 38, 39]. Em outras palavras, as propriedades estatísticas do tráfego se preservam sob agregação temporal e não é possível apontar uma escala de tempo dominante para a sua caracterização e controle [26, 29, 39].

O primeiro estudo a apresentar evidências da natureza fractal do tráfego em redes foi realizado por Leland et al., que analisaram as estatísticas do tráfego da rede local e dos enlaces de saída do *Bellcore Morristown Research and Engineering Center* [39]. Esses autores observaram que o tráfego era caracterizado por uma alternância de períodos de surtos e de suavidade, o que não dependia da resolução temporal em que as medidas eram tomadas. Posteriormente, outros estudos encontraram evidências da natureza fractal do tráfego em transmissões de vídeo a taxa variável, em aplicações cliente-servidor, em transferência e compartilhamento de arquivos e em aplicações *World-Wide Web* (www) [26–36].

Em geral, o tráfego fractal apresenta alta variabilidade e dependência de longa duração, propriedades estas que têm efeitos nocivos no desempenho da rede. O principal deles é o decaimento lento da distribuição probabilística da ocupação das filas dos comutadores, o que aumenta a perda de pacotes e o atraso na transmissão [40–47]. Matematicamente, a alta variabilidade decorre do fato de os incrementos de tráfego seguirem uma distribuição marginal com decaimento algébrico, i.e.,

$$\mathbb{P}\{\mathbf{X}_\tau[n] > x\} \sim x^{-\alpha}, \quad 1 < \alpha < 2, \quad (2.1)$$

em que $\mathbf{X}_\tau[n]$ é a variável aleatória que representa o incremento de tráfego no intervalo $[n\tau; (n+1)\tau)$. Neste caso, a probabilidade de surgirem surtos de tráfego de grande magnitude é relevante em qualquer instante. A dependência de longa duração, por sua vez, decorre do fato de a distribuição marginal de $\mathbf{X}_\tau[n]$ depender significativamente dos valores passados em um horizonte muito distante. Matematicamente, esta dependência é caracterizada por uma estrutura de autocorrelação com decaimento não somável, i.e.,

$$\sum_{\lambda} |\mathcal{R}_X(\lambda)| \rightarrow \infty. \quad (2.2)$$

Assim, quando há dependência de longa duração, o comportamento de $\mathbf{X}_\tau[n]$ se preserva ao longo de grandes intervalos de tempo. O tráfego fractal é então marcado pela alternância de longos períodos de silêncio e de longas rajadas de tráfego com surtos de grande magnitude que não pode ser representada por meio dos modelos de tráfego tradicionais. Por isso, diversos estudos têm sido direcionados ao desenvolvimento de modelos fractais e à sua aplicação na análise de desempenho e no controle de tráfego em redes.

Muitos autores argumentam que a natureza fractal do tráfego é consistente com o conceito de auto-similaridade estatística [26, 27, 29, 39]. Um processo $\mathbf{Z}(t)$ com incrementos estacionários é auto-similar se a relação

$$\mathbf{Z}(t) \stackrel{d}{=} m^{-H} \mathbf{Z}(mt), \quad m, t \in \mathbb{R}_+, \quad (2.3)$$

é válida para algum $0 < H < 1$, parâmetro que é denominado índice de escala ou fator de Hurst. O símbolo $\stackrel{d}{=}$ indica igualdade em distribuição, e o termo m^{-H} é denominado lei de escala. Note que os processos estocásticos auto-similares são particularmente interessantes para o fim de modelagem de tráfego, pois permitem resumir toda a natureza fractal em um único parâmetro.

Muitos estudos demonstram que os processos auto-similares são capazes de representar a dependência de longa duração e a alta variabilidade do tráfego [26–29, 32, 33, 39, 48]. Entretanto, há evidências de que o tráfego pode apresentar propriedades fractais mais complexas, que não podem ser representadas com acurácia por meio da auto-similaridade [29, 34, 35, 47, 49]. Em alguns casos, há necessidade de modelos mais sofisticados, como os que são baseados em processos multifractais.

Na literatura, a definição do que é um processo multifractal não é única [29, 35, 47, 50]. Em [50], um processo $\mathbf{Z}(t)$ com incrementos estacionários é dito ser multifractal se

$$\mathbf{Z}(mt) \stackrel{d}{=} C(m)\mathbf{Z}(t), \quad m, t \in \mathbb{R}_+, \quad (2.4)$$

em que $\mathbf{Z}(t)$ e $C(m)$ são funções aleatórias independentes. Podemos definir um índice de escala generalizado para o processo $\mathbf{Z}(t)$, dado por

$$\mathbf{H}(m) = \log_m C(m). \quad (2.5)$$

De (2.4),

$$\mathbf{Z}(mt) \stackrel{d}{=} m^{\mathbf{H}(m)}\mathbf{Z}(t), \quad m, t \in \mathbb{R}_+. \quad (2.6)$$

Note que a auto-similaridade representa um caso particular de multifractalidade em que $\mathbf{H}(m)$ é determinístico e constante para todas as escalas de tempo. O uso de processos multifractais para modelagem de tráfego ainda é um tema recente. Um tutorial sobre modelagem de tráfego por meio de processos multifractais é apresentado em [51]. Em [34], Riedi e Véhel analisaram diversos traços de tráfego TCP, e demonstraram que eles satisfazem relações de multifractalidade. Em [35], Feldmann et al. propuseram o uso de processos multifractais baseados em cascatas multiplicativas para a modelagem do tráfego em redes. Em [47], Molnár et al. analisaram o comportamento assintótico de um sistema de filas alimentado por tráfego multifractal.

Embora os processos multifractais possam, em muitos casos, representar o tráfego com acurácia maior do que os auto-similares, seu tratamento analítico é bem mais difícil. Os modelos multifractais tendem a ser não-paramétricos, o que elimina a sua parcimônia e torna a identificação um problema complexo. Mesmo nos casos em que se impõe uma estrutura paramétrica ao modelo, a estimação dos parâmetros não costuma ser tarefa simples. Por isso, seu uso em estudos de desempenho de redes ainda é bastante limitado.

Os modelos auto-similares, por outro lado, são extremamente parcimoniosos, de fácil identificação e de tratamento analítico mais simples. Além disso, a literatura demonstra que eles são capazes de prover, na maioria dos casos, uma representação bastante satisfatória do tráfego fractal [29]. Por estes motivos, o presente capítulo se volta primordialmente à representação de tráfego fractal por meio de modelos auto-similares. Nas próximas seções, são apresentados os principais modelos de tráfego auto-similar disponíveis na literatura e, no fim deste capítulo, apresentamos um resumo das principais técnicas para a estimação do fator de auto-similaridade H .

2.1 PROCESSOS *on/off* DE CAUDA PESADA

O processo *on/off* é bastante utilizado para representar os incrementos de tráfego gerados por uma fonte isolada. Ele consiste em uma cadeia que representa a alternância de estados de atividade e de silêncio. No estado de atividade, o processo assume um valor constante, dado por r . No estado de silêncio, seu valor é nulo. A duração destes estados é representada pelas variáveis aleatórias independentes τ_{on} e τ_{off} .

Um processo *on/off* $\mathbf{X}(t)$ tem distribuição marginal binomial, tal que

$$\mathbb{E}\mathbf{X}(t) = rp \quad (2.7)$$

$$\text{var } \mathbf{X}(t) = r^2 p(1-p), \quad (2.8)$$

em que

$$p = \frac{\mathbb{E}\tau_{\text{on}}}{\mathbb{E}\tau_{\text{on}} + \mathbb{E}\tau_{\text{off}}}. \quad (2.9)$$

No modelo de cauda pesada, τ_{on} ou τ_{off} tem distribuição com decaimento algébrico, i.e.,

$$\mathbb{P}\{\tau_{\text{on}} > \tau\} = \tau^{-\alpha_{\text{on}}} L_{\text{on}}(\tau), \text{ ou} \quad (2.10)$$

$$\mathbb{P}\{\tau_{\text{off}} > \tau\} = \tau^{-\alpha_{\text{off}}} L_{\text{off}}(\tau), \quad (2.11)$$

em que $\alpha_{\text{on}} > \alpha_{\text{off}}$, e $L_{\text{on}}(\tau)$ e $L_{\text{off}}(\tau)$ representam funções de decaimento lento. Neste caso, é possível demonstrar que a função de autocorrelação de $\mathbf{X}(t)$ é tal que [52]

$$\lim_{\lambda \rightarrow \infty} \mathcal{R}_{\mathbf{X}}(\lambda) \sim \frac{(\mathbb{E}\tau_{\text{off}})^2}{(\alpha_{\text{on}} - 1)(\mathbb{E}\tau_{\text{on}} + \mathbb{E}\tau_{\text{off}})^3} \eta^{(\alpha_{\text{on}} - 1)} L_{\text{on}}(\eta), \quad (2.12)$$

ou seja, o processo $\mathbf{X}(t)$ tem dependência de longa duração. Resultado similar pode ser obtido para o caso de $\alpha_{\text{on}} \leq \alpha_{\text{off}}$. Em [32, 52, 53], demonstra-se que a acumulação de N processos *on/off* independentes, identicamente distribuídos e somados pode convergir, para $N \rightarrow \infty$, a um movimento browniano fracionário (FBM¹) com

$$H = \frac{3 - \min(\alpha_{\text{on}}; \alpha_{\text{off}})}{2}, \quad (2.13)$$

ou a um movimento α -estável de Lévy, com

$$\alpha = \min(\alpha_{\text{on}}; \alpha_{\text{off}}), \quad (2.14)$$

$$H = \frac{1}{\alpha}. \quad (2.15)$$

A convergência a um ou a outro processo depende do quanto a soma dos processos é capaz de compensar a alta variabilidade e a dependência de longa duração de cada um deles [52]. Em ambos os casos, este resultado pode ser visto como consequência do Teorema do Limite Central Generalizado, que demonstra que a soma de um número muito grande de variáveis aleatórias independentes converge assintoticamente a uma distribuição α -estável [54].

2.2 PROCESSO M/G/∞ DE CAUDA PESADA

Os incrementos de tráfego fractal também podem ser representados por meio de um processo M/G/∞ de cauda pesada. Neste processo, subjaz um sistema de filas em que as chegadas têm distribuição de Poisson de taxa η e a duração do serviço segue uma distribuição de decaimento algébrico. O serviço e as chegadas são estatisticamente independentes entre si.

Em um dado instante de tempo, o valor de um processo M/G/∞ de cauda pesada $\mathbf{X}(t)$ é dado pela latência do sistema de filas subjacente. Seja \mathbf{T} o tempo de serviço deste sistema, cuja distribuição é tal que

$$\mathbb{P}\{\mathbf{T} > \tau\} = \tau^{-\alpha} L(\tau), \quad (2.16)$$

¹ *fractional brownian motion.*

em que $L(\tau)$ representa uma função de decaimento lento. Normalmente, considera-se que o tempo de serviço segue a distribuição de Pareto, dada por

$$\mathbb{P}\{\mathbf{T} > \tau\} = \begin{cases} \left(\frac{\tau}{\delta}\right)^{-\alpha}, & \tau \geq \delta \\ 1, & \text{c.c.} \end{cases} \quad (2.17)$$

Neste caso, o valor médio do processo $\mathbf{X}(t)$ é dado por [55]

$$\mathbb{E}\mathbf{X}(t) = \frac{\eta\delta\alpha}{\alpha-1}, \quad 1 < \alpha < 2. \quad (2.18)$$

Em [52], demonstra-se que a função de autocorrelação de $\mathbf{X}(t)$ satisfaz o limite

$$\lim_{\eta \rightarrow \infty} \mathcal{R}_{\mathbf{X}}(\lambda) \sim \lambda^{(\alpha-1)} L(\lambda), \quad (2.19)$$

de modo que processo M/G/ ∞ apresenta longa duração. A variância do processo $\mathbf{X}(t)$ é dada por [55]

$$\text{var } \mathbf{X}(t) = \begin{cases} 2\eta t^2 \left[\frac{\delta}{2} \left(1 - \frac{1}{1-\alpha}\right) - \frac{t}{6} \right], & 0 \leq t \leq \delta \\ 2\eta f_1(\alpha, \delta) + f_2(\alpha, \delta)(t - \delta) + f_3(\alpha, \delta) \cdot t^{3-\alpha}, & t > \delta, \end{cases} \quad (2.20)$$

em que

$$\begin{aligned} f_1(\alpha, \delta) &= \delta^3 \left[\frac{1}{3} - \frac{1}{2-2\alpha} + \frac{1}{(1-\alpha)(2-\alpha)(3-\alpha)} \right] \\ f_2(\alpha, \delta) &= \delta^2 \left[\frac{1}{2} - \frac{1}{1-\alpha} + \frac{1}{(1-\alpha)(2-\alpha)} \right] \\ f_3(\alpha, \delta) &= \frac{\delta^\alpha}{(\alpha-1)(2-\alpha)(3-\alpha)}. \end{aligned}$$

Note que, para valores suficientemente grandes de t , a variância é dominada pelo termo $f_3(\alpha, \delta) \cdot t^{3-\alpha}$. Desta maneira, o processo $\mathbf{X}(t)$ é assintoticamente auto-similar, com

$$H = \frac{3-\alpha}{2}. \quad (2.21)$$

2.3 MOVIMENTO BROWNIANO FRACIONÁRIO

Em [56], Norros propôs um modelo auto-similar denominado tráfego browniano fracionário, que é dado por

$$\mathbf{A}(t) = \mu t + \gamma \mathbf{Z}(t), \quad (2.22)$$

em que $\mathbf{A}(t)$ representa o tráfego acumulado até o instante t , μ é a sua taxa média, γ é um parâmetro associado ao seu desvio-padrão e $\mathbf{Z}(t)$ é um movimento browniano fracionário, com média zero, fator de auto-similaridade $1/2 \leq H < 1$ e variância $\text{var } \mathbf{Z}(t) = |t|^{2H}$. O movimento browniano fracionário é dado por

$$\mathbf{Z}(t) = \frac{1}{\Gamma(H + \frac{1}{2})} \int_{\mathbb{R}} \left[(t-u)_+^{H-1/2} - (-u)_+^{H-1/2} \right] d\mathbf{B}(u), \quad \frac{1}{2} \leq H < 1 \quad (2.23)$$

em que $\mathbf{B}(t)$ é um movimento browniano puro e $(x)_+ = x$, se $x > 0$, ou zero se $x \leq 0$. Note que o movimento browniano fracionário tem média nula, é não-estacionário e se degenera no próprio movimento browniano se $H = 1/2$.

O movimento browniano fracionário é o único processo não-estacionário que tem realizações contínuas, distribuição marginal gaussiana e incrementos estacionários. Teoricamente, é possível

obtê-lo a partir da agregação de um grande número de fontes *on/off* identicamente distribuídas se, para cada uma delas, os períodos de atividade e de silêncio são independentes e ao menos um deles segue distribuição de cauda pesada [53]. Aplicações multimídia e de transferências de arquivos têm estas características, o que até hoje justifica o uso do processo (2.23) como modelo de tráfego em redes.

Podemos constatar a auto-similaridade do processo $\mathbf{Z}(t)$ a partir de sua variância, que satisfaz a relação

$$\begin{aligned}\text{var } \mathbf{Z}(t) &= \mathbb{E} \mathbf{Z}^2(t) \\ &= t^{2H} \sigma_{\mathbf{Z}}^2,\end{aligned}\quad (2.24)$$

em que $\sigma_{\mathbf{Z}}^2$ é a variância de $\mathbf{Z}(1)$. A não-estacionariedade do processo $\mathbf{Z}(t)$ é evidenciada pela sua função de autocovariância, dada por

$$C_{\mathbf{Z}}(s, t) = \frac{\sigma_{\mathbf{Z}}^2}{2} \left(|s|^{2H} + |t|^{2H} - |s - t|^{2H} \right). \quad (2.25)$$

O processo de incrementos discretos $\mathbf{X}_{\tau}[n]$ associado ao processo $\mathbf{Z}(t)$, denominado ruído Gaussiano fracionário, é dado por

$$\begin{aligned}\mathbf{X}_{\tau}[n] &= \mathbf{Z}(t + \tau) - \mathbf{Z}(t) \\ &= \frac{1}{\Gamma(H + \frac{1}{2})} \int_{\mathbb{R}} \left\{ |(n+1)\tau - u|_+^{H-1/2} - [n\tau - u]_+^{H-1/2} \right\} d\mathbf{B}(u), \quad \frac{1}{2} \leq H < 1.\end{aligned}\quad (2.26)$$

Este processo é estacionário, tem média nula e função de autocovariância dada por

$$C_{\mathbf{X}_{\tau}}(n, n + \lambda) = \frac{\sigma_{\mathbf{Z}}^2 \tau^{2H}}{2} \left(|\lambda + 1|^{2H} - 2|\lambda|^{2H} + |\lambda - 1|^{2H} \right), \quad (2.27)$$

que, para $H = 1/2$, é somável. Contudo, para $H > 1/2$,

$$\lim_{\eta \rightarrow \infty} C_{\mathbf{X}_{\tau}}(\lambda) \sim H(2H - 1)\lambda^{2H-2}, \quad (2.28)$$

que não é somável. Assim, os processos $\mathbf{X}_{\tau}[n]$ e $\mathbf{Z}(t)$ apresentam dependência de longa duração se $H > 1/2$ [36].

O tráfego browniano fracionário pode ser considerado um modelo parcimonioso, pois requer apenas três parâmetros para caracterizar o tráfego. Em [40, 56], demonstrou-se que o tráfego browniano fracionário é capaz de representar o tráfego em redes locais. Neste caso, a acurácia do modelo se deve às características gaussianas que advêm da agregação do tráfego de um grande número de fontes independentes. Em contrapartida, o mesmo estudo demonstra que o tráfego browniano fracionário pode não ser adequado à representação do tráfego em enlaces de saída pois, neste caso, as estatísticas podem ser fortemente não-gaussianas. Isto evidencia a principal desvantagem desse processo, que é o fato de não representar adequadamente o tráfego quando há alta variabilidade.

Em [40, 56], Norros analisou o problema do armazenamento do tráfego browniano fracionário em um sistema de filas simples. Seja \mathbf{Q}^* a variável aleatória que representa o máximo comprimento da fila em qualquer instante. A distribuição de \mathbf{Q}^* é, então, dada por

$$\mathbb{P} \{ \mathbf{Q}^* > x \} = \mathbb{P} \left\{ \max_{t \geq 0} [\mathbf{A}(t) - gt] > x \right\}, \quad (2.29)$$

em que g é a capacidade do servidor. Supondo que $\{ \mathbf{Q}^* > x \}$ é um evento raro, a Teoria dos Grandes Desvios leva a [41]

$$\mathbb{P} \{ \mathbf{Q}^* > x \} \simeq \max_{t \geq 0} \mathbb{P} \{ \mathbf{A}(t) - gt > x \}. \quad (2.30)$$

Esta relação pode ser utilizada para caracterizar a distribuição de \mathbf{Q}^* para valores grandes de x . Neste caso,

$$\begin{aligned}
\mathbb{P}\{\mathbf{Q}^* > x\} &\simeq \max_{t \geq 0} \mathbb{P}\{\mu t + \gamma \mathbf{Z}(t) - gt > x\} = \\
&= \max_{t \geq 0} \mathbb{P}\{\mu t + \gamma t^H \mathbf{Z}(1) - gt > x\} = \\
&= \max_{t \geq 0} \mathbb{P}\left\{\mathbf{Z}(1) > \frac{(g - \mu)t + x}{\gamma t^H}\right\} = \\
&= \max_{t \geq 0} \bar{G}\left(\frac{(g - \mu)t + x}{\gamma t^H}\right)
\end{aligned} \tag{2.31}$$

lembrando que $\mathbf{Z}(t) = t^H \mathbf{Z}(1)$. A função $\bar{G}(y)$ representa a distribuição residual gaussiana. Dado que

$$\lim_{y \rightarrow \infty} \bar{G}(y) \sim \exp\left(-\frac{y^2}{2}\right), \tag{2.32}$$

a relação (2.31) pode ser aproximada por

$$\mathbb{P}\{\mathbf{Q}^* > x\} \simeq \max_{t \geq 0} \exp\left\{-\frac{1}{2} \left[\frac{(g - \mu)t + x}{\gamma t^H}\right]^2\right\}. \tag{2.33}$$

A solução de (2.33) é obtida para

$$t^* = \frac{xH}{(g - \mu)(1 - H)}. \tag{2.34}$$

Note que os transbordos são mais prováveis nesta escala de tempo, que é denominada escala de tempo crítica [43]. É importante observar que esta escala de tempo está associada somente com a ocorrência de transbordo da fila no instante em que ele é mais provável e não permite, portanto, caracterizar a probabilidade de transbordo em outros instantes de tempo [45]. Substituindo (2.34) em (2.33), obtemos [40, 56]

$$\mathbb{P}\{\mathbf{Q}^* > x\} \simeq \exp\left\{-\frac{(g - \mu)^{2H}}{2\gamma^2 H^{2H} (1 - H)^{2-2H}} x^{2-2H}\right\}. \tag{2.35}$$

Assim, a variável \mathbf{Q}^* segue uma distribuição Weibull quando $x \rightarrow \infty$. Isto significa que valores elevados de latência são obtidos com probabilidade não-negligenciável no caso do tráfego auto-similar, sendo que a distribuição residual de \mathbf{Q}^* decai mais lentamente conforme aumenta o valor de H .

A relação (2.35) pode também ser utilizada para dimensionar a taxa do servidor em função da probabilidade de transbordo da fila, quando esta tem uma capacidade finita [40, 56]. Sendo ε a probabilidade de transbordo, a taxa, que representa a banda efetiva do tráfego fBm, é dada por

$$g^* = \mu + (-2\gamma^2 \ln \varepsilon)^{\frac{1}{2H}} H(1 - H)^{\frac{1-H}{H}} x^{\frac{H-1}{H}}. \tag{2.36}$$

2.4 PROCESSOS α -ESTÁVEIS E AUTO-SIMILARES

Estudos recentes indicam que o modelo (2.22) também pode ser utilizado para representar tráfego com distribuição marginal α -estável [48, 57]. Neste caso, é possível representar o tráfego fractal mesmo quando há alta variabilidade, tal que a distribuição marginal tenha fortes características não-gaussianas.

As distribuições α -estáveis surgiram dos estudos de Paul Lévy sobre a soma de variáveis aleatórias independentes e identicamente distribuídas. Uma variável aleatória V tem distribuição α -estável se a sua função característica é dada por²

$$\Phi(\omega) = \mathbb{E} \exp(j\omega V) = \exp \left\{ j\mu\omega - \gamma^\alpha |\omega|^\alpha \left[1 - j\beta \frac{\omega}{|\omega|} \theta(\omega, \alpha) \right] \right\}, \quad (2.37)$$

em que

$$\theta(\omega, \alpha) = \begin{cases} \tan \frac{\alpha\pi}{2}, & \text{se } \alpha \neq 1, \\ -\frac{2}{\pi} \ln |\omega| & \text{se } \alpha = 1. \end{cases} \quad (2.38)$$

O parâmetro $\alpha \in (0; 2]$ é denominado expoente característico e está relacionado com o decaimento da distribuição. O parâmetro $\beta \in [-1; 1]$ é denominado parâmetro de simetria. Para $\beta = 0$, a distribuição é simétrica. Para valores negativos de β , a distribuição torna-se assimétrica à esquerda e, para valores positivos, assimétrica à direita. Os parâmetros μ e γ são denominados parâmetros de localização e de escala.

As distribuições de Lévy, de Cauchy e gaussiana são casos particulares da distribuição α -estável que se obtêm para $\alpha = 1/2$, 1 e 2, respectivamente. Para $0 < \alpha < 2$, a distribuição α -estável tem decaimento algébrico, dado por [54]

$$\lim_{v \rightarrow \infty} \mathbb{P} \{V > v\} \sim v^{-\alpha}. \quad (2.39)$$

As distribuições α -estáveis generalizam a distribuição gaussiana no que tange o fato de a forma da distribuição se preservar sob a soma. Sejam $V_1 \sim S(\alpha, \beta_1, \gamma_1, \delta_1)$ e $V_2 \sim S(\alpha, \beta_2, \gamma_2, \delta_2)$ duas variáveis aleatórias α -estáveis independentes. A soma destas variáveis é dada por $V = V_1 + V_2 \sim S(\alpha, \beta, \gamma, \delta)$, em que

$$\beta = \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \quad \gamma^\alpha = \gamma_1^\alpha + \gamma_2^\alpha, \quad \delta = \delta_1 + \delta_2 \quad (2.40)$$

No contexto de modelagem de tráfego fractal, consideram-se processos estocásticos α -estáveis tais que [48, 57, 58]

$$Z(t) = \frac{1}{\Gamma(H + \frac{1}{\alpha})} \int_{\mathbb{R}} \left[(t-u)_+^{H-1/\alpha} - (-u)_+^{H-1/\alpha} \right] M_S(du), \quad \frac{1}{\alpha} \leq H < 1, \quad 1 < \alpha \leq 2, \quad (2.41)$$

sendo o processo $Z(t)$ denominado movimento linear estável fracionário. Para $\alpha = 2$, $Z(t)$ equivale ao movimento browniano fracionário. Para $0 < \alpha < 2$, a distribuição marginal de $Z(t)$ é uma variável aleatória α -estável, cujo decaimento é algébrico e dado por (2.39).

A auto-similaridade do processo $Z(t)$ pode ser constatada por meio do seu parâmetro de escala, que satisfaz a igualdade

$$\gamma_{Z(t)} = t^H \gamma_{Z(1)}. \quad (2.42)$$

O processo de incrementos discretos $X_\tau[n]$ associado ao processo $Z(t)$ é denominado ruído α -estável fracionário e dado por

$$X_\tau[n] = \frac{1}{\Gamma(H + \frac{1}{\alpha})} \int_{\mathbb{R}} \left[((n+1)\tau - u)_+^{H-1/\alpha} - (n\tau - u)_+^{H-1/\alpha} \right] M_S(du), \quad \frac{1}{\alpha} \leq H < 1. \quad (2.43)$$

²A parametrização (2.37) e (2.38) não é a única possível para uma distribuição α -estável, embora seja a mais comum.

Para $0 < \alpha < 2$, as distribuições α -estáveis não têm momentos de segunda ordem, de modo que a autocovariância dos processos $\mathbf{Z}(t)$ e $\mathbf{X}_\tau[n]$ não pode ser definida. As dependências temporais destes processos podem, contudo, serem avaliadas por meio da função de codiferença. A codiferença entre duas variáveis aleatórias \mathbf{V}_1 e \mathbf{V}_2 é definida como [59]

$$C_{\mathbf{V}_1, \mathbf{V}_2} = \{\ln \mathbb{E} \exp(j\omega \mathbf{V}_1) + \ln \mathbb{E} \exp(-j\omega \mathbf{V}_2) - \ln \mathbb{E} \exp[j\omega (\mathbf{V}_1 - \mathbf{V}_2)]\}_{\omega=1}. \quad (2.44)$$

Quando \mathbf{V}_1 e \mathbf{V}_2 têm distribuição gaussiana, a codiferença equivale à covariância. Para um movimento linear estável fracionário $\mathbf{Z}(t)$ em que $M_S(du)$ é simétrica, a função de autocodiferença dos incrementos é dada por

$$C_{\mathbf{X}_\tau}(n, n + \lambda) = \gamma_{\mathbf{Z}(1)}^\alpha \tau^{\alpha H} \left(|\lambda + 1|^{\alpha H} - 2|\lambda|^{\alpha H} + |\lambda - 1|^{\alpha H} \right). \quad (2.45)$$

Para $H > 1/\alpha$,

$$\lim_{\eta \rightarrow \infty} C_{\mathbf{X}}(\eta) \sim |\eta|^{\alpha(H-1)}, \quad (2.46)$$

de modo que os processos $\mathbf{X}_\tau[n]$ e $\mathbf{Z}(t)$ têm dependência de longa duração [60].

Em [48, 57], analisa-se o armazenamento do tráfego auto-similar α -estável. Ao contrário do que ocorre no caso do tráfego browniano fracionário, não é possível caracterizar o comportamento assintótico da distribuição probabilística da latência das filas, dado que o decaimento algébrico da distribuição marginal de $\mathbf{Z}(t)$ torna o transbordo um evento não-raro, tal que a Teoria dos Grandes Desvios não se aplica. Contudo, é possível demonstrar que a probabilidade de a máxima latência em qualquer instante tem distribuição residual tal que [48, 57]

$$\mathbb{P}\{\mathbf{Q}^* > x\} \geq K(\alpha, \gamma, \mu, r) x^{-\alpha(1-H)},$$

de modo que, sob tráfego auto-similar α -estável, o decaimento da ocupação das filas nos comutadores é, na melhor das hipóteses, algébrico.

2.5 ESTIMAÇÃO DO PARÂMETRO H

Nesta seção, apresentamos um resumo das principais técnicas para estimar o parâmetro de auto-similaridade H de uma série temporal. Outras técnicas de interesse são apresentadas em [61].

2.5.1 ESTIMADOR DE WHITTLE

Este método supõe que tanto a estrutura como os parâmetros do processo de incrementos $\mathbf{X}[n]$ são conhecidos, à exceção da variância e do parâmetro H . A partir de amostras da função de autocorrelação deste processo, a sua densidade espectral de potência pode ser definida como

$$S_{\mathbf{X}}(\omega, H) = \sum_{\eta} \mathcal{R}_{\mathbf{X}}[\eta] \exp(-i\omega\eta). \quad (2.47)$$

Seja $S_{\mathbf{X}}^*(\omega, H)$ uma forma normalizada de $S_{\mathbf{X}}(\omega, H)$ em que a variância de $\mathbf{X}[n]$ é unitária. Seja ainda $I_N(\omega)$ o periodograma obtido a partir de N amostras do processo $\mathbf{X}[n]$, sendo este dado por

$$I_N(\omega) = \frac{1}{2\pi N} \left| \sum_{k=1}^N \mathbf{X}[k] \exp(i\omega k) \right|^2. \quad (2.48)$$

O parâmetro de auto-similaridade pode, então, ser obtido resolvendo o seguinte problema de otimização:

$$H = \arg \min_H \int_{-\pi}^{\pi} \frac{I_N(\omega)}{S_{\mathbf{X}}^*(\omega, H)} d\omega. \quad (2.49)$$

Se o processo $\mathbf{X}[n]$ é reduzido a um conjunto de N amostras, a integral em (2.49) pode ser entendida como uma soma discreta para $\omega = \frac{2\pi}{N}, \frac{4\pi}{N}, \dots, \frac{2(n-1)\pi}{N}$. As definições utilizadas pelo estimador de Whittle permitem ainda obter diretamente a variância da estimativa, que é dada por

$$\sigma_H^2 = 4\pi \left[\int_{-\pi}^{\pi} \left(\frac{\partial}{\partial H} \log S_{\mathbf{X}}^*(\omega, H) \right)^2 d\omega \right]^{-1} \quad (2.50)$$

Embora o estimador de Whittle permita uma análise bastante refinada dos dados, ele tem a desvantagem de ser um método paramétrico, podendo levar a resultados incorretos caso os dados não se ajustem ao espectro $S_{\mathbf{X}}^*(\omega, H)$.

2.5.2 GRÁFICO VARIÂNCIA-TEMPO

Este método recorre à relação entre as variâncias do processo original e do processo obtido por agregação, que é dada por

$$\log \left(\text{var } \mathbf{X}^{(m)}[n] \right) = \log \left(\text{var } \mathbf{X}[n] \right) + (2H - 2) \log m. \quad (2.51)$$

Portanto, para um conjunto de amostras de uma série temporal, a curva que representa a relação entre $\log \left(\text{var } \mathbf{X}^{(m)}[n] \right)$ e $\log m$ aproxima-se de uma reta de inclinação $2H - 2$. Esta reta pode ser obtida facilmente a partir dos processos agregados para diversos valores de m . Note que o método não pode, teoricamente, ser aplicado quando a distribuição marginal de $\mathbf{X}[n]$ tem cauda pesada pois, neste caso, a variância é infinita.

2.5.3 GRÁFICO DA ESTATÍSTICA R/S

Este método é baseado na definição da estatística R/S para processos estocásticos discretos, e sua principal vantagem é a independência com relação à distribuição marginal do processo em análise. A estatística R/S representa a razão entre a medida da faixa do processo $\mathbf{X}(t)$ e o desvio padrão amostrado, sendo definida como

$$R/S(N) = \frac{\max_{1 \leq j \leq N} \left\{ \sum_{n=1}^j [\mathbf{X}[n] - M_N] \right\} - \min_{1 \leq j \leq N} \left\{ \sum_{n=1}^j [\mathbf{X}[n] - M_N] \right\}}{\sqrt{\frac{1}{N} \sum_{n=1}^N [\mathbf{X}[n] - M_N]^2}}, \quad (2.52)$$

em que $M_N = \frac{1}{N} \sum_{l=1}^N \mathbf{X}[l]$. No caso de processos auto-similares, esta razão apresenta a seguinte característica [38, 39]:

$$R/S(N) \sim N^H. \quad (2.53)$$

Desta maneira, a relação entre $\log(R/S)$ e $\log N$ aproxima de uma reta de inclinação H .

2.5.4 ESTIMADOR BASEADO EM WAVELETS

O parâmetro de auto-similaridade de um processo pode ser estimado utilizando técnicas baseadas na transformada de *wavelets*. O desenvolvimento em *wavelets* de um processo auto-similar $\mathbf{Z}(t)$ até a escala J é dado por

$$\mathbf{Z}(t) = 2^{-J/2} \sum_{n=-\infty}^{\infty} \mathbf{a}_J[n] \phi(2^{-J}t - n) + \sum_{j=1}^J 2^{-j/2} \sum_{n=-\infty}^{\infty} \mathbf{d}_j[n] \psi(2^{-j}t - n), \quad (2.54)$$

em que $\psi(t)$ a função *wavelet* utilizada no desenvolvimento, e $\phi(t)$ a correspondente função de escala. Os processos $\mathbf{a}_j[n]$ e $\mathbf{d}_j[n]$ são denominados coeficiente de aproximação e coeficiente de detalhamento da transformada. Para o caso de o processo $\mathbf{Z}(t)$ ser um movimento α -estável linear e fracionário, é possível demonstrar que o processo $\mathbf{d}_j[n]$ é um processo estacionário, cuja codiferença satisfaz a relação [60]

$$|\mathcal{C}_{\mathbf{d}_j}(1, n, n')| \leq C |2^j n - 2^j n'|^{-\frac{\alpha}{2}(N-H)}, \quad (2.55)$$

em que N é o número de momentos nulos da *wavelet* $\psi(\cdot)$. Desta maneira, podem ser obtidas seqüências $\mathbf{d}_j[n]$ sem dependência utilizando uma *wavelet* com N suficientemente grande para que seja rápido o decaimento de (2.55). Neste caso, as seqüências $\mathbf{d}_j[n]$ satisfazem a relação [60]

$$\mathbb{E} \log_2 |\mathbf{d}_j[n]| = \left(H + \frac{1}{2}\right) j + \mathbb{E} \log_2 |\mathbf{d}_0[n]|. \quad (2.56)$$

Portanto a relação entre $\mathbb{E} \log_2 |\mathbf{d}_j[n]|$ e a escala j é representada por uma reta de inclinação $H + 1/2$. Quando $\mathbf{Z}(t)$ é gaussiano, a variância de $\mathbf{d}_j[n]$ satisfaz a relação

$$\text{var } \mathbf{d}_j[n] = (2H + 1) j + \text{var } \mathbf{d}_0[n]. \quad (2.57)$$

2.6 PROCESSO DE CHEGADAS COM LIMITANTE FRACIONÁRIO

Em geral, os modelos de tráfego baseados em processos estocásticos auto-similares têm tratamento matemático difícil. Uma alternativa é utilizar funções determinísticas, denominadas processos envelope, que limitam o volume de tráfego em um dado intervalo de tempo. Um processo envelope é determinístico se a probabilidade de violação pelo tráfego é nula em qualquer tempo. Caso contrário, a função determinística e a correspondente probabilidade de violação constituem, juntas, um envelope estatístico. Um processo envelope pode ainda ser acumulativo, se ele limita o volume de tráfego até um dado instante de tempo, ou incremental, se ele limita o volume em um intervalo qualquer.

Em [45], propõe-se um processo envelope para a modelagem de tráfego auto-similar. Este processo, denominado processo envelope do movimento browniano fracionário, é dado por

$$\widehat{A}(t) = \mu t + k \gamma t^H, \quad (2.58)$$

sendo μ , γ e H obtidos dos parâmetros do tráfego browniano fracionário correspondente. O parâmetro k determina a probabilidade de que o processo $\mathbf{A}(t)$ exceda o envelope $\widehat{A}(t)$ no instante t , que é dada por

$$\mathbb{P} \left\{ \mathbf{A}(t) > \widehat{A}(t) \right\} = \mathbb{P} \left\{ \mathbf{Z}(1) > k \right\} = \overline{G}(k), \quad (2.59)$$

em que $\overline{G}(\cdot)$ representa a função distribuição gaussiana residual.

Note que, por definição, o processo envelope do movimento browniano fracionário pressupõe que o tráfego tem distribuição marginal gaussiana, o que nem sempre é verdadeiro [48, 52, 57]. Além disso, este processo não representa adequadamente a estacionariedade dos incrementos de tráfego pois, neste caso,

$$\begin{aligned} \mathbb{P} \left\{ \mathbf{A}(t + \tau) - \mathbf{A}(t) > \widehat{\mathbf{A}}(t + \tau) - \widehat{\mathbf{A}}(t) \right\} &= \mathbb{P} \left\{ \mu\tau + \gamma\mathbf{Z}(\tau) > \widehat{\mathbf{A}}(t + \tau) - \widehat{\mathbf{A}}(t) \right\} \\ &= \mathbb{P} \left\{ \mathbf{Z}(1) > k \frac{(t + \tau)^H - t^H}{\tau^H} \right\}, \end{aligned} \quad (2.60)$$

ou seja, a probabilidade de violação não é constante no tempo.

Nesta seção, propomos um envelope incremental denominado Processo de Chegada com Limitante Fracionário (FBAP³) para a representação do tráfego auto-similar. O envelope FBAP é capaz de representar tanto o tráfego acumulado como os seus incrementos, e generaliza o processo envelope do movimento browniano fracionário por não pressupor qualquer distribuição marginal específica. Analisamos o armazenamento de tráfego FBAP e obtemos limitantes para a latência e o atraso em um sistema de filas com taxa de serviço constante. Finalmente, propomos uma extensão do envelope FBAP para o caso multifractal, o que permite representar acuradamente o tráfego de natureza fractal mais complexa.

2.6.1 DEFINIÇÃO

Seja o tráfego acumulado até o instante t dado por (2.22), em que $\mathbf{Z}(t)$ é um processo auto-similar arbitrário, não necessariamente gaussiano. Supondo ainda que os incrementos de $\mathbf{Z}(t)$ são estacionários, o incremento do processo $\mathbf{A}(t)$ no intervalo $[t; t + \tau]$ é dado por

$$\begin{aligned} \mathbf{A}(t; t + \tau) &= \mathbf{A}(0; t + \tau) - \mathbf{A}(0; t) \\ &= \mu\tau + \gamma\mathbf{Z}(\tau). \end{aligned} \quad (2.61)$$

Note que o processo $\mathbf{A}(t; t + \tau)$ também é estacionário. Por simplicidade, ele será representado por $\mathbf{A}(\tau)$. Um processo envelope para $\mathbf{A}(\tau)$ pode ser definido como [62–65]

$$\widehat{\mathbf{A}}(\tau) = \mu\tau + k\gamma\tau^H, \quad \tau \in \mathbb{R}_+. \quad (2.62)$$

A probabilidade de violação do envelope $\widehat{\mathbf{A}}(\tau)$ é dada por

$$\begin{aligned} \mathbb{P} \left\{ \mathbf{A}(\tau) > \widehat{\mathbf{A}}(\tau) \right\} &= \mathbb{P} \left\{ \mathbf{Z}(1) > k \right\} \\ &= \overline{F}_{\mathbf{Z}}(k). \end{aligned} \quad (2.63)$$

A probabilidade de violação é, então, constante no tempo. Este resultado difere daquele que é obtido no caso do processo envelope do movimento browniano fracionário, dado por (2.60). Dada a suposição de que os incrementos de tráfego são estacionários, podemos concluir que o envelope FBAP provê melhor representação. Note ainda que o envelope FBAP é capaz de representar tanto o tráfego acumulado como os seus incrementos, e generaliza o processo envelope do movimento browniano fracionário por não presumir qualquer distribuição marginal para o tráfego.

Podemos obter os parâmetros do envelope FBAP estatístico da seguinte maneira. Seja $\mathbf{A}[n]$ o processo discreto para o qual desejamos obter o envelope, i.e.,

$$\mathbf{A}[n] = \mathbf{A}(nT), \quad n \in \mathbb{N},$$

sendo T uma constante real e positiva que representa a menor escala de tempo disponível para análise. Os incrementos discretos de $\mathbf{A}[n]$ são então dados por

³ *Fractional Bounded Arrival Process.*

$$\begin{aligned} \mathbf{A}_{(i)}[n] &= \mathbf{A}[(n+1)T] - \mathbf{A}[nT] \\ &= \mu T i + \gamma \mathbf{Z}[i], \quad n, i \in \mathbb{N} \end{aligned} \quad (2.64)$$

em que $\mathbf{Z}[i]$ é o processo discreto que representa a amostragem de $\mathbf{Z}(t)$ nos instantes iT , $i \in \mathbb{N}$. Por simplicidade, o índice (i) será omitido sempre que for igual a 1. A realização de $\mathbf{A}_{(i)}[n]$ é dada por $\mathbf{A}_{(i)}[n]$. Supondo que a distribuição marginal de $\mathbf{A}_{(i)}[n]$ é conhecida, podemos obter os parâmetros μ e γ facilmente. Por exemplo, se a distribuição marginal é gaussiana, estes parâmetros são dados por

$$\mu = \frac{1}{T} \mathbb{E} \mathbf{A}[n] \quad (2.65)$$

$$\gamma = \frac{1}{T^H} \sqrt{\text{var } \mathbf{A}[n]}. \quad (2.66)$$

No caso de uma distribuição marginal α -estável, o parâmetro γ é dado pelo parâmetro de escala, normalizado por $1/T^H$. Para estimar o parâmetro H , podemos recorrer aos métodos apresentados na Seção 2.5. Finalmente, devemos escolher o parâmetro k de tal modo que obtenhamos a probabilidade de violação desejada, i.e.,

$$\varepsilon = \bar{F}_{\mathbf{Z}}(k),$$

em que $\bar{F}_{\mathbf{Z}}$ representa a distribuição residual de $\mathbf{Z}[1]$, dada por $\mathbf{A}[n] - \mu T$.

Ilustramos o uso do modelo FBAP na Figura 2.1. A série temporal representada corresponde a 10^6 amostras de um ruído gaussiano fracionário que foi gerado por meio do método proposto em [66]. Por simplicidade, supomos que $T = 1$. A média e o desvio padrão da série temporal são 0,800 e 0,160, respectivamente. O parâmetro de auto-similaridade H foi estimado por meio do método de *wavelets* proposto em [67], sendo obtido o valor de 0,794. Para o envelope FBAP, os parâmetros μ e γ são dados pela média e pelo desvio padrão da série, respectivamente. O parâmetro k é igual a 3,719, do que resulta uma probabilidade de violação de 10^{-4} .

Na Figura 2.1, podemos ver que o modelo FBAP representa a série acuradamente, mesmo para escalas de tempo grandes. A diferença entre o envelope FBAP e o comportamento do tráfego ao longo das escalas de tempo é devida a imprecisões na geração da série e na estimação dos seus parâmetros.

2.6.2 ESTIMAÇÃO DOS PARÂMETROS FBAP PARA TRÁFEGO REAL

Para tráfego real, podemos obter o envelope FBAP por meio do envelope empírico. Esta abordagem é especialmente interessante quando há evidências de que o tráfego pode se desviar da estacionariedade ou da auto-similaridade em determinadas escalas de tempo. Seja ε a probabilidade de violação desejada para o envelope FBAP. Seja ainda o envelope empírico para a série $\mathbf{A}_{(i)}[n]$, dado por

$$\widehat{\mathbf{A}}_{(i)}^{(e)}[n] = \mu i T + Z^{(e)}[i], \quad (2.67)$$

em que o parâmetro μ é dado por (2.65), e $Z^{(e)}[i]$ é dado por

$$Z^{(e)}[i] = \{x \in \mathbb{R} : \mathbb{P} \{ \mathbf{A}_{(i)}[n] - \mu i T > x \} = \varepsilon \}, \quad (2.68)$$

sendo ε a probabilidade de violação desejada. Para cada escala de tempo, estimamos a probabilidade $\mathbb{P} \{ \mathbf{A}_{(i)}[n] - \mu i T > x \}$ a partir da série temporal. Supondo que a série é suficientemente longa, esta probabilidade pode ser aproximada pela frequência relativa do evento sempre que a distribuição marginal do tráfego não for conhecida. Opcionalmente, $Z^{(e)}[i]$ pode ser dada por $\max_n \mathbf{A}_{(i)}[n] - \mu i T$, caso em que o envelope empírico se torna determinístico e igual ao descrito em [68].

Podemos obter o produto $k\gamma$ e o parâmetro H ajustando o termo $k\gamma(iT)^H$ de modo a limitar $Z^{(e)}[i]$ por cima em um intervalo suficientemente grande de escalas de tempo. Contudo, o envelope

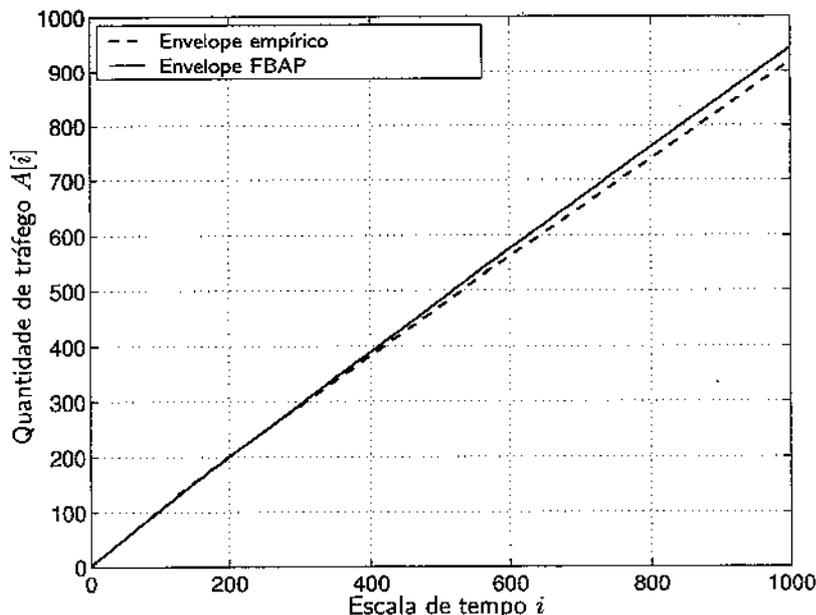


Figura 2.1: Validação do modelo FBAP para um ruído gaussiano fracionário.

FBAP pode se tornar demasiado conservador. Uma abordagem mais realista é tolerar uma pequena violação do envelope (2.67). Isto pode ser obtido fazendo

$$\{k\gamma, H\} = \arg \min_{k\gamma, H} \left\{ \frac{1}{2} \sum_{i=1}^N \left[f \left(\log Z^{(e)}[i] - \log k\gamma - H \log i \right) \right]^2 \right\}, \quad (2.69)$$

em que

$$f(x) = \begin{cases} x, & x < 0 \\ wx, & x \geq 0, w > 1. \end{cases}$$

O valor de w deve ser tal que leve a uma solução de compromisso entre a violação expressa por (2.67) e um menor conservadorismo no envelope. Em alguns casos, podemos obter melhores resultados associando (2.69) com outro procedimento de estimação mais adequado a uma dada série temporal. Por exemplo, o produto $k\gamma$ pode ser dado por

$$k\gamma = \max_n A[n], \quad (2.70)$$

e o parâmetro H pela solução de

$$H = \arg \min_H \left\{ \frac{1}{2} \sum_{i=1}^N \left[f \left(\log Z^{(e)}[i] - \log k\gamma - H \log i \right) \right]^2 \right\}. \quad (2.71)$$

Em [64, 65], utilizamos o método do envelope empírico para validar o modelo FBAP para tráfego real. Os resultados são resumidos na Tabela 2.1. As séries BCPAUG89, BCPOCT89, BCPOCT89EXT, BCPOCT89EXT4, LBLTCP3, LBLPKT4-5 e DECPKT2-4 correspondem a amostras de tráfego de redes reais, que apresentam características fractais e que são normalmente utilizadas no estudo de modelagem de tráfego em redes [27, 38, 39, 69, 70]. As séries originais foram pré-processadas para criar seqüências com até 1×10^6 amostras, que representam a quantidade de tráfego (em bytes) que

fluiu através da rede durante um dado intervalo de tempo. Para as séries BCPAUG89, BCPOCT89, BCOCT89EXT e BCOCT89EXT4, o intervalo foi de 10ms. Para as demais, o intervalo foi de 100ms.

As demais séries indicadas na Tabela 2.1 correspondem a seqüências de vídeo MPEG, que também apresentam comportamento fractal [28, 31]. Cada série consiste em uma seqüência discreta de 4×10^4 amostras, sendo que cada amostra corresponde ao tamanho (em bytes) de um quadro de vídeo MPEG completo. Dado que os quadros são gerados a uma taxa de 20 quadros por segundo, cada amostra corresponde à quantidade de tráfego (em bytes) que fluiu através da rede durante um intervalo de 50ms.

Escolhemos os parâmetros do modelo FBAP de modo a obter uma probabilidade de violação de 10^{-4} . Obtivemos o parâmetro μ utilizando (2.65). Para o produto $k\gamma$ e o parâmetro H , utilizamos (2.69) ou (2.70) e (2.71). Nos resultados, referimo-nos a estas abordagens por Método 1 e Método 2, respectivamente. Em ambos os casos, consideramos que $w = 10$.

Na Tabela 2.1, são indicados os parâmetros FBAP obtidos para cada série, bem como o método utilizado para obter $k\gamma$ e H . As Figuras 2.2 a 2.4 apresentam a relação $\log Z^{(e)}[i] \times \log i$ correspondente às séries BOND, FUSS1 e DECPKT4, e os envelopes empíricos e FBAP correspondentes. Os gráficos obtidos para as demais séries podem ser encontrados na referência [64]. Observe que o modelo FBAP representa bem o comportamento fractal do tráfego, sobretudo quando ele é consistente com o conceito de auto-similaridade estatística.

| Série | Método | μ | $k\gamma$ | H | Série | Método | μ | $k\gamma$ | H |
|------------|--------|--------------------|--------------------|-------|------------|--------|--------------------|--------------------|-------|
| bcoc89ext | 2 | 7.49×10^2 | 5.12×10^4 | 0.576 | lblpkt5 | 2 | 2.61×10^4 | 2.30×10^5 | 0.718 |
| bcoc89ext4 | 2 | 6.22×10^3 | 1.90×10^5 | 0.949 | lbltcp3 | 2 | 3.39×10^4 | 3.76×10^5 | 0.814 |
| bcpaug89 | 1 | 1.38×10^5 | 4.54×10^5 | 0.759 | movie | 2 | 3.57×10^5 | 1.04×10^6 | 0.641 |
| bcpoc89 | 2 | 3.63×10^5 | 6.24×10^5 | 0.911 | mrbean | 2 | 4.41×10^5 | 1.59×10^6 | 0.777 |
| asterix | 1 | 5.59×10^5 | 1.56×10^6 | 0.740 | mtv1 | 2 | 6.15×10^5 | 2.20×10^6 | 0.797 |
| atp | 2 | 5.47×10^5 | 1.54×10^6 | 0.699 | mtv2 | 2 | 4.95×10^5 | 2.61×10^6 | 0.769 |
| bond | 2 | 6.08×10^5 | 1.41×10^6 | 0.659 | news1 | 2 | 5.17×10^5 | 1.47×10^6 | 0.694 |
| decpkt2 | 2 | 2.38×10^5 | 3.79×10^5 | 0.793 | news2 | 1 | 3.84×10^5 | 2.00×10^6 | 0.733 |
| decpkt3 | 2 | 1.81×10^5 | 3.04×10^5 | 0.766 | race | 1 | 7.69×10^5 | 2.03×10^6 | 0.732 |
| decpkt4 | 2 | 2.63×10^5 | 3.09×10^5 | 0.789 | sbowl | 2 | 5.88×10^5 | 1.20×10^6 | 0.727 |
| dino | 2 | 3.27×10^5 | 9.66×10^5 | 0.723 | simpsons | 2 | 4.64×10^5 | 1.19×10^6 | 0.702 |
| fuss1 | 1 | 6.78×10^5 | 1.55×10^6 | 0.639 | star2 | 2 | 2.33×10^5 | 1.16×10^6 | 0.729 |
| fuss2 | 2 | 6.28×10^5 | 1.92×10^6 | 0.774 | talk1 | 2 | 3.63×10^5 | 8.96×10^5 | 0.717 |
| lams | 2 | 1.83×10^5 | 1.02×10^6 | 0.710 | talk2 | 2 | 4.48×10^5 | 1.22×10^6 | 0.766 |
| lblpkt4 | 1 | 3.64×10^4 | 2.57×10^5 | 0.736 | terminator | 1 | 2.73×10^5 | 5.11×10^5 | 0.626 |

Tabela 2.1: Parâmetros do envelope FBAP correspondente às séries analisadas.

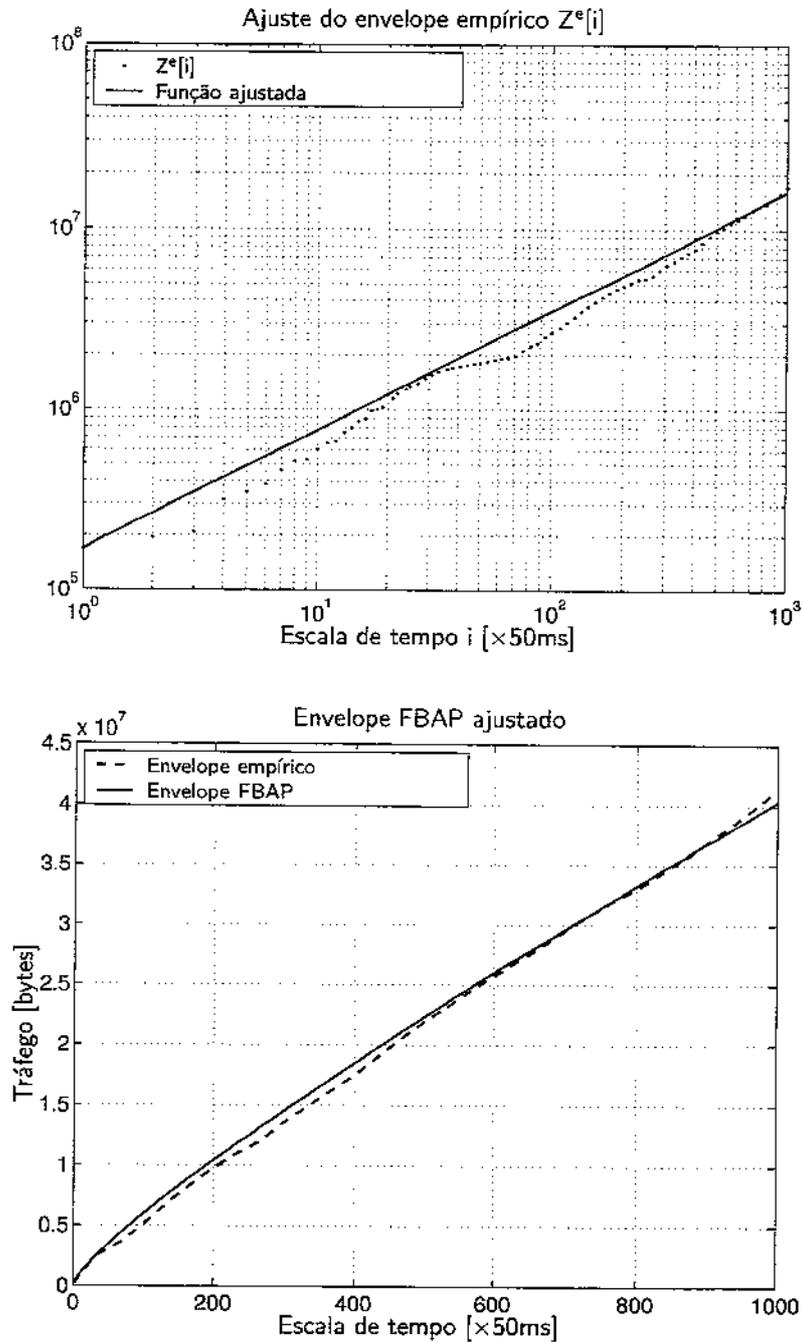


Figura 2.2: Representação da série BOND utilizando o envelope FBAP.

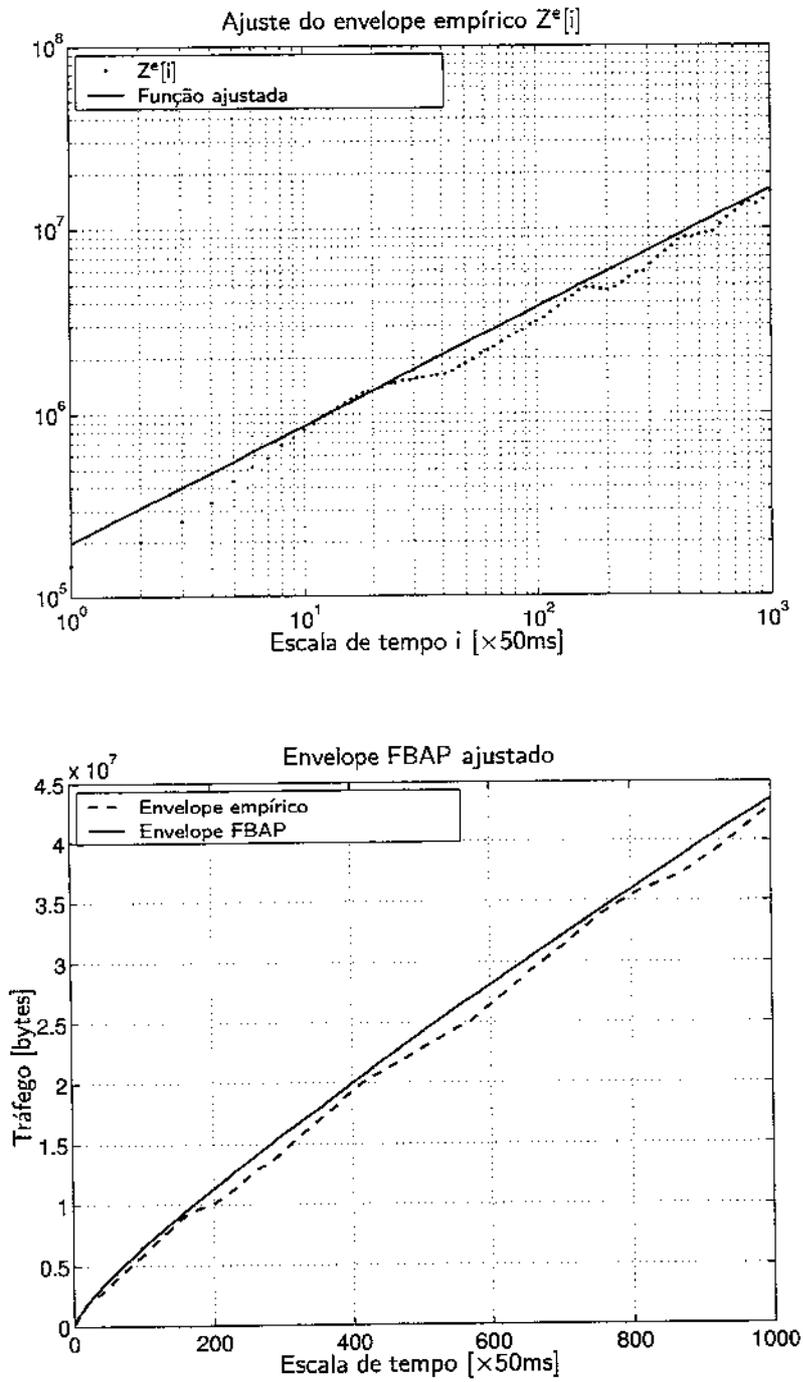


Figura 2.3: Representação da série FUSS1 utilizando o envelope FBAP.

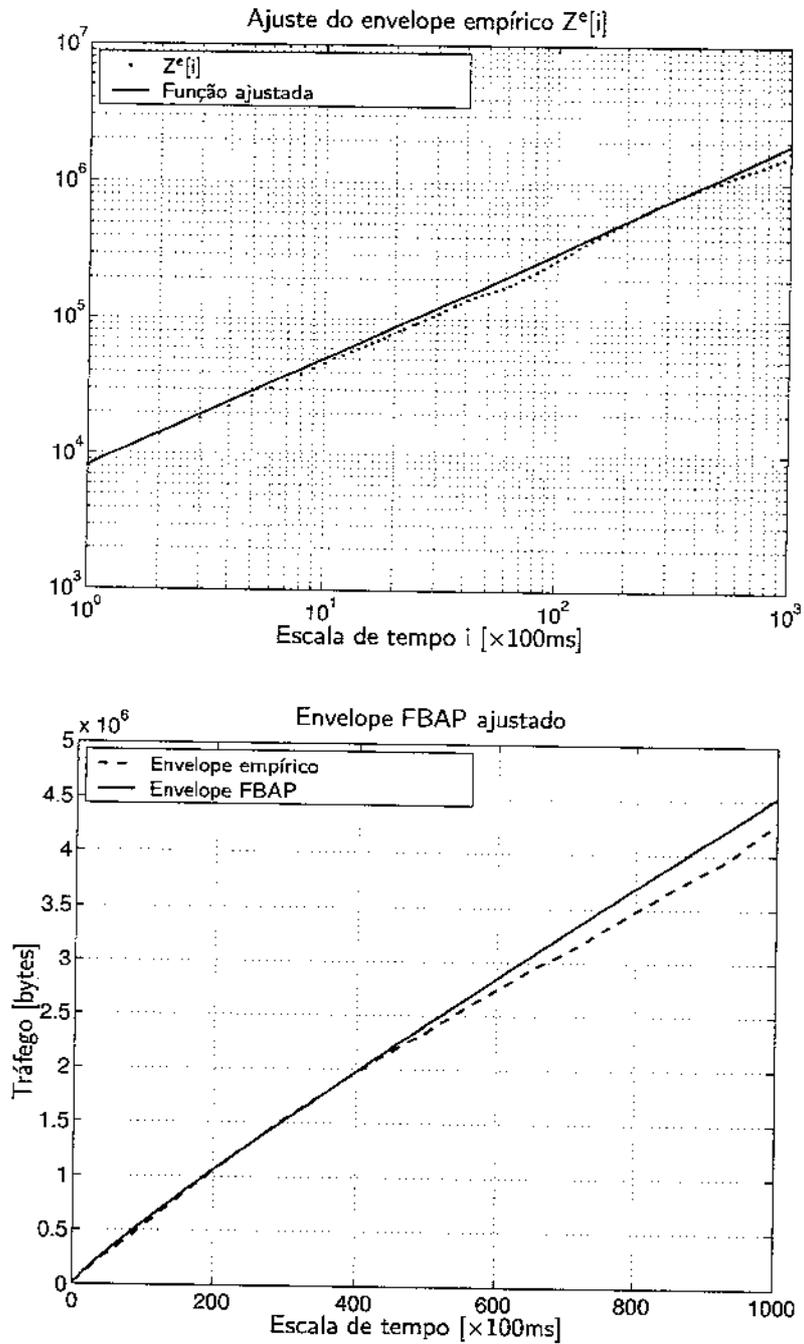


Figura 2.4: Representação da série DECPKT4 utilizando o envelope FBAP.

2.6.3 ARMAZENAMENTO DO TRÁFEGO FBAP

Seja um sistema de filas com servidor com taxa constante g e latência dada por $\mathbf{Q}(t)$. O tráfego de entrada é representado pelo modelo FBAP dado por (2.62) e (2.63). De modo a garantir a estabilidade do sistema, vamos supor que a taxa do servidor é maior do que a taxa média do tráfego de entrada, dada por μ .

Suponha que o i -ésimo ciclo ativo do servidor inicie no instante t_i . Entendemos por ciclo ativo o período em que o tráfego de entrada é continuamente acumulado na fila. Assim,

$$\mathbf{Q}(t_i + \tau) = \mathbf{A}(t_i; t_i + \tau) - g\tau, \quad 0 \leq \tau < \tau_i^{busy},$$

em que τ_i^{busy} é a duração do i -ésimo ciclo ativo do servidor.

Sendo o tráfego de entrada representado pelo modelo FBAP, podemos definir um limitante superior para $\mathbf{Q}(t_i + \tau)$ como

$$\begin{aligned} \widehat{Q}(\tau) &= \widehat{A}(\tau) - g\tau \\ &= (\mu - g)\tau + k\gamma\tau^H, \quad 0 \leq \tau < \tau^{busy}. \end{aligned} \quad (2.72)$$

em que τ^{busy} é dado por $\arg_{\tau>0} \{\widehat{Q}(\tau) = 0\}$. A probabilidade de que $\mathbf{Q}(t_i + \tau)$ exceda $\widehat{Q}(\tau)$ é dada por

$$\begin{aligned} \mathbb{P}\{\mathbf{Q}(t_i + \tau) > \widehat{Q}(\tau)\} &= \mathbb{P}\{\mathbf{A}(\tau) > \widehat{A}(\tau)\} \\ &= \overline{F}_{\mathbf{Z}}(k). \end{aligned} \quad (2.73)$$

Para $0 < H < 1$, o envelope $\widehat{Q}(\tau)$ é \cap -convexo em $\tau \geq 0$. Podemos então definir um limitante superior para a máxima latência do sistema de filas, que é dado por

$$\begin{aligned} Q^* &= \max_{0 \leq \tau \leq \tau^{busy}} \widehat{Q}(\tau) \\ &= (g - \mu)^{\frac{H}{H-1}} (k\gamma)^{\frac{1}{1-H}} H^{\frac{H}{1-H}} (1 - H). \end{aligned} \quad (2.74)$$

A escala de tempo para a qual obtemos Q^* é denominada escala de tempo máxima [45]. A probabilidade de violação do limitante Q^* no instante $t_i + \tau$ é dada por

$$\begin{aligned} \mathbb{P}\{\mathbf{Q}(t_i + \tau) > Q^*\} &\leq \mathbb{P}\{\mathbf{Q}(t_i + \tau) > \widehat{Q}(\tau)\} \\ &= \overline{F}_{\mathbf{Z}}(k). \end{aligned} \quad (2.75)$$

Por outro lado, a probabilidade de violação do limitante Q^* em qualquer instante de tempo durante o ciclo ativo é dada por [62]

$$\begin{aligned} \mathbb{P}\bigcup_{\tau \geq 0} \{\mathbf{Q}(t_i + \tau) > Q^*\} &= \mathbb{P}\left\{\sup_{\tau \geq 0} \mathbf{Q}(t_i + \tau) > Q^*\right\} \\ &\geq \sup_{\tau \geq 0} \mathbb{P}\{\mathbf{Q}(t_i + \tau) > Q^*\} \\ &= \overline{F}_{\mathbf{Z}}(k). \end{aligned} \quad (2.76)$$

No caso do tráfego browniano fracionário, a relação acima converge para uma igualdade se Q^* é suficientemente grande, tal que a violação possa ser considerada um evento raro [40, 41, 56]. No caso

do tráfego α -estável com $\alpha < 2$, entretanto, o decaimento da ocupação é, na melhor das hipóteses, algébrico. Assim, a aproximação (2.76) não é válida e k deve ser escolhido tal que o envelope FBAP seja determinístico, para que o limitante Q^* não seja violado.

Dado o limitante (2.74), o máximo atraso introduzido no tráfego pelo sistema de filas é limitado a $D^* = Q^*/g$. Além disso, podemos utilizar a relação (2.74) para calcular a taxa de servidor que limita a latência a Q^* com probabilidade de violação dada por (2.76). Esta taxa é dada por

$$g^* = \mu + (k\gamma)^{\frac{1}{H}} H(1-H)^{\frac{1-H}{H}} (Q^*)^{\frac{H-1}{H}}. \quad (2.77)$$

Nesta relação, o valor de g^* pode ser considerado como uma medida de banda efetiva do tráfego FBAP. Cabe observar que a mesma relação também representa a banda efetiva do processo envelope do FBM [68].

2.6.4 PROCESSO DE CHEGADAS COM LIMITANTE MULTIFRACTAL

Como foi dito no capítulo anterior, há na literatura evidências de que o tráfego em redes pode apresentar propriedades fractais mais complexas, que não podem ser representadas com acurácia por meio da auto-similaridade [29, 34, 35, 47, 49]. Neste caso, há necessidade de modelos mais sofisticados, como os que são baseados em processos multifractais.

Nesta seção, apresentamos uma extensão do modelo FBAP que permite representar o tráfego com características multifractais. Seja o tráfego acumulado até o instante t dado por

$$\mathbf{A}(t) = \mu t + \gamma \mathbf{Z}(t), \quad \mathbf{Z}(0) = 0, \quad (2.78)$$

em que $\mathbf{Z}(t)$ é um processo estocástico multifractal com incrementos estacionários. Em particular, consideramos que um processo $\mathbf{Z}(t)$ com incrementos estacionários é um processo multifractal se [50]

$$\mathbf{Z}(mt) \stackrel{d}{=} C(m)\mathbf{Z}(t), \quad m, t \in \mathbb{R}_+, \quad (2.79)$$

em que $C(m)$ é um processo estocástico independente de $\mathbf{Z}(t)$. A partir de (2.79), podemos concluir que

$$\mathbf{Z}(t) \stackrel{d}{=} C(t)\mathbf{Z}(1), \quad t \in \mathbb{R}_+. \quad (2.80)$$

Note que, sendo $\mathbf{Z}(0) = 0$, temos que $C(0) = 0$. Além disso, dado que os incrementos de $\mathbf{Z}(t)$ são estacionários,

$$\begin{aligned} \mathbf{Z}(t+\tau) - \mathbf{Z}(t) &\stackrel{d}{=} \mathbf{Z}(\tau) - \mathbf{Z}(0) \\ &\stackrel{d}{=} \mathbf{Z}(1)[C(\tau) - C(0)] \\ &\stackrel{d}{=} \mathbf{Z}(1)C(\tau), \quad \tau, t \geq 0. \end{aligned}$$

Os incrementos do processo $\mathbf{A}(t)$ no intervalo $[t, t+\tau]$ são, portanto, dados por

$$\begin{aligned} \mathbf{A}(t; t+\tau) &= \mathbf{A}(t+\tau) - \mathbf{A}(t) \\ &= \mu\tau + \gamma\mathbf{Z}(1)C(\tau). \end{aligned} \quad (2.81)$$

Note que os incrementos do processo $\mathbf{A}(t)$ também são estacionários. Podemos definir um envelope para este processo como [63–65]

$$\widehat{\mathbf{A}}(\tau) = \mu\tau + k\gamma\widehat{C}(\tau), \quad \forall \tau \geq 0, \quad (2.82)$$

em que $\widehat{C}(\tau)$ é um envelope para o processo $C(\tau)$. O processo envelope $\widehat{A}(\tau)$ é denominado Processo de Chegada com Limitante Multifractal (MFBAP⁴). Note que podemos considerar este processo como sendo uma generalização do envelope FBAP, que é obtido quando $\widehat{C}(\tau) = \tau^H$, $0 < H < 1$.

Para o cálculo da probabilidade de violação do envelope $\widehat{A}(\tau)$, vamos supor que o envelope $\widehat{C}(\tau)$ é determinístico. Embora seja aparentemente restritiva do ponto de vista matemático, tal suposição é razoável pois, na prática, os incrementos de tráfego são limitados pela capacidade dos enlaces. Isto significa que a distribuição marginal de $C(\tau)$ tem domínio finito para τ finito e pode ser limitada pelo envelope considerado. A probabilidade de que os incrementos de tráfego excedam o envelope $\widehat{A}(\tau)$ é então dada por

$$\begin{aligned} \mathbb{P} \{ \mathbf{A}(t; t + \tau) > \widehat{A}(\tau) \} &= \\ &= \mathbb{P} \{ \mu\tau + \gamma \mathbf{Z}(1)C(\tau) > \mu\tau + k\gamma\widehat{C}(\tau) \} \\ &\leq \mathbb{P} \{ \mathbf{Z}(1) > k \} = \\ &= \overline{F}_{\mathbf{Z}}(k), \end{aligned} \tag{2.83}$$

em que $\overline{F}_{\mathbf{Z}}(\cdot)$ representa a função distribuição residual da variável aleatória $\mathbf{Z}(1)$.

Podemos obter o envelope MFBAP para uma dada série de tráfego a partir da definição de envelope empírico apresentada na Seção 2.6.2. Neste caso, o parâmetro μ é dado por (2.65). Para o envelope $\widehat{C}(\tau)$, escolhemos uma função determinística que reproduza a relação entre $\log Z^{(e)}[i]$ e $\log i$ observada empiricamente.

Para a validação do modelo MFBAP, foram utilizadas as mesmas séries de tráfego real da Seção 2.6.2. Para os envelopes, escolhemos uma probabilidade de violação ε igual a 10^{-4} . As Figuras 2.5 a 2.7 mostram a aproximação obtida no caso do modelo FBAP. É possível concluir que, para garantir a probabilidade de violação desejada, o modelo FBAP tende a ser conservador. Por outro lado, a relação $\log Z^{(e)}[i] \times \log i$ de cada uma das séries pode ser representada utilizando um modelo MFBAP, em que

$$\widehat{C}(\tau) = \tau^{H(\tau)}, \tag{2.84}$$

sendo

$$H(\tau) = H_0 + \gamma_H \exp \left\{ - \frac{[\ln(\tau) - \mu_H]^2}{2\sigma_H^2} \right\}. \tag{2.85}$$

As Figuras 2.5 a 2.7 mostram, para cada uma das séries citadas, a aproximação obtida no caso do modelo MFBAP proposto. Os parâmetros do modelo para estes e para os demais traços analisados se encontram na Tabela 2.2. Os gráficos obtidos para os demais traços podem ser encontrados na referência [64].

Note que o envelope MFBAP é capaz de representar o tráfego de maneira bem mais acurada do que o envelope FBAP. Não obstante, vale lembrar que o seu tratamento matemático não é trivial e a sua complexidade depende muito da função $\widehat{C}(\tau)$ escolhida. Além disso, a estimação dos parâmetros por meio do método do envelope empírico baseado em (2.69) ou em (2.70) também tende a ser computacionalmente muito custosa e sujeita a mínimos locais. Ademais, não há como garantir a unicidade dos parâmetros no caso geral.

Em relação ao armazenamento do tráfego MFBAP, vale observar que o problema de otimização

$$Q^* = \arg \max_{\tau} \widehat{A}(\tau) - g\tau \tag{2.86}$$

geralmente não tem solução analítica, sendo que a convergência e a unicidade da solução obtida por meio de métodos numéricos geralmente não pode ser assegurada. Atualmente, a falta de resultados

⁴ *Multifractal Bounded Arrival Process.*

| Série | μ | $k\gamma$ | H_0 | γ_H | σ_H | μ_H |
|-------------|--------------------|--------------------|-------|------------|--------------------|---------|
| bcoct89ext | 7.49×10^2 | 5.30×10^4 | 0.378 | 0.406 | 3.514 | -1.569 |
| bcoct89ext4 | 6.22×10^3 | 1.86×10^5 | 0.489 | 0.486 | 14.681 | 0.623 |
| bcpaug89 | 1.38×10^5 | 4.52×10^5 | 0.393 | 0.374 | 11.625 | -0.428 |
| bcpact89 | 3.63×10^5 | 6.27×10^5 | 0.485 | 0.483 | 5.256 | -1.381 |
| asterix | 5.59×10^5 | 1.58×10^6 | 0.480 | 0.347 | 5.185 | -1.555 |
| atp | 5.47×10^5 | 1.14×10^6 | 0.468 | 0.557 | 2.130 | 0.317 |
| bond | 6.08×10^5 | 1.39×10^6 | 0.391 | 0.293 | 4.63×10^6 | 0.000 |
| decpkt2 | 2.38×10^5 | 3.85×10^5 | 0.419 | 0.403 | 5.955 | -1.965 |
| decpkt3 | 1.81×10^5 | 3.10×10^5 | 0.401 | 0.403 | 6.202 | -1.630 |
| decpkt4 | 2.63×10^5 | 3.12×10^5 | 0.400 | 0.434 | 5.330 | -1.335 |
| dino | 3.27×10^5 | 8.48×10^5 | 0.491 | 0.363 | 3.204 | 0.324 |
| fuss1 | 6.78×10^5 | 1.61×10^6 | 0.405 | 0.304 | 5.505 | -0.745 |
| fuss2 | 6.28×10^5 | 1.81×10^6 | 0.491 | 0.455 | 2.500 | -0.605 |
| lambs | 1.83×10^5 | 7.15×10^5 | 0.385 | 0.497 | 3.934 | 1.759 |
| lblpkt4 | 3.64×10^4 | 2.58×10^5 | 0.409 | 0.345 | 5.472 | -2.381 |
| lblpkt5 | 2.61×10^4 | 2.33×10^5 | 0.393 | 0.394 | 5.407 | -1.259 |
| ibltcp3 | 3.39×10^4 | 3.65×10^5 | 0.464 | 0.437 | 4.845 | -1.232 |
| movie | 3.57×10^5 | 9.93×10^5 | 0.392 | 0.328 | 4.778 | 0.473 |
| mrbean | 4.41×10^5 | 1.39×10^6 | 0.543 | 0.368 | 3.214 | 0.313 |
| mtv1 | 6.15×10^5 | 2.12×10^6 | 0.571 | 0.413 | 2.454 | -0.427 |
| mtv2 | 4.95×10^5 | 2.51×10^6 | 0.460 | 0.372 | 5.625 | 0.476 |
| news1 | 5.17×10^5 | 1.43×10^6 | 0.564 | 0.292 | 2.214 | -0.373 |
| news2 | 3.84×10^5 | 2.11×10^6 | 0.493 | 0.486 | 1.932 | -1.278 |
| race | 7.69×10^5 | 2.00×10^6 | 0.411 | 0.496 | 3.255 | -1.051 |
| sbowl | 5.88×10^5 | 1.17×10^6 | 0.453 | 0.429 | 3.147 | -0.159 |
| simpsons | 4.64×10^5 | 1.17×10^6 | 0.443 | 0.305 | 5.103 | -0.341 |
| star2 | 2.33×10^5 | 7.51×10^5 | 0.462 | 0.465 | 3.134 | 1.546 |
| talk1 | 3.63×10^5 | 6.50×10^5 | 0.369 | 0.608 | 3.077 | 0.870 |
| talk2 | 4.48×10^5 | 1.03×10^6 | 0.470 | 0.571 | 2.413 | 0.157 |
| terminator | 2.73×10^5 | 5.12×10^5 | 0.420 | 0.281 | 4.174 | -1.038 |

Tabela 2.2: Traços utilizados na validação do modelo MFBAP.

analíticos e de condições de unicidade para os parâmetros do modelo e para os limitantes de latência e de atraso são as principais restrições ao uso do modelo MFBAP na análise e no controle de tráfego em redes.

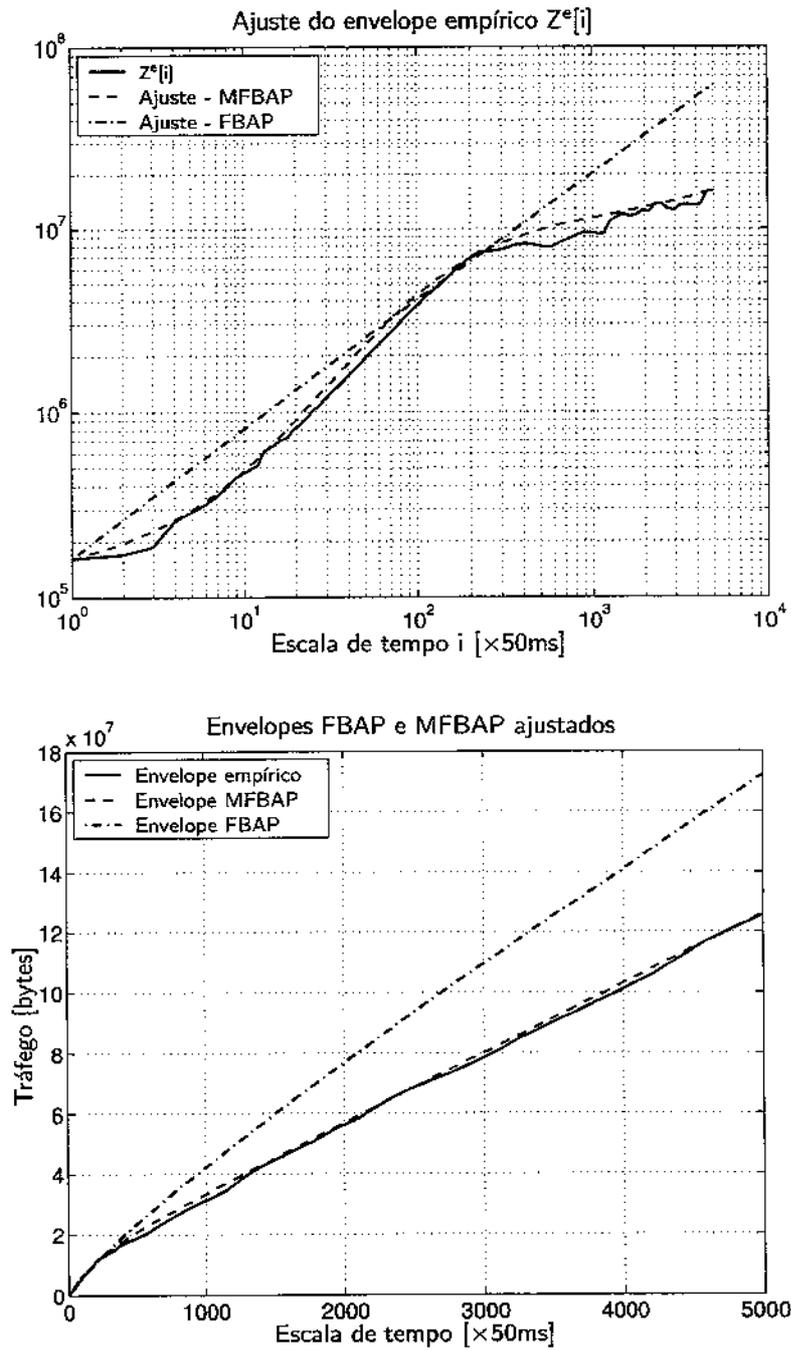


Figura 2.5: Representação do traço ATP utilizando os envelopes FBAP e MFBAP.

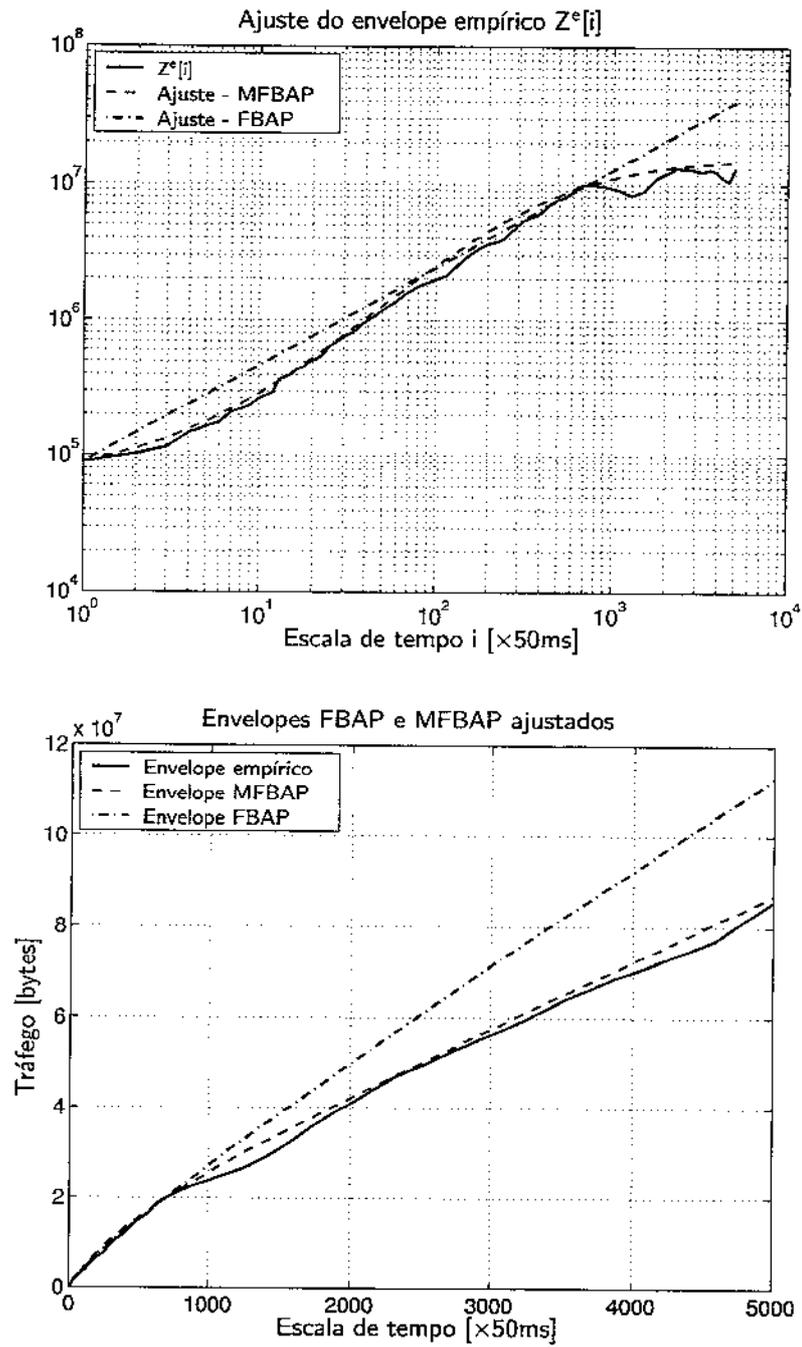


Figura 2.6: Representação do traço TALK1 utilizando os envelopes FBAP e MFBAP.

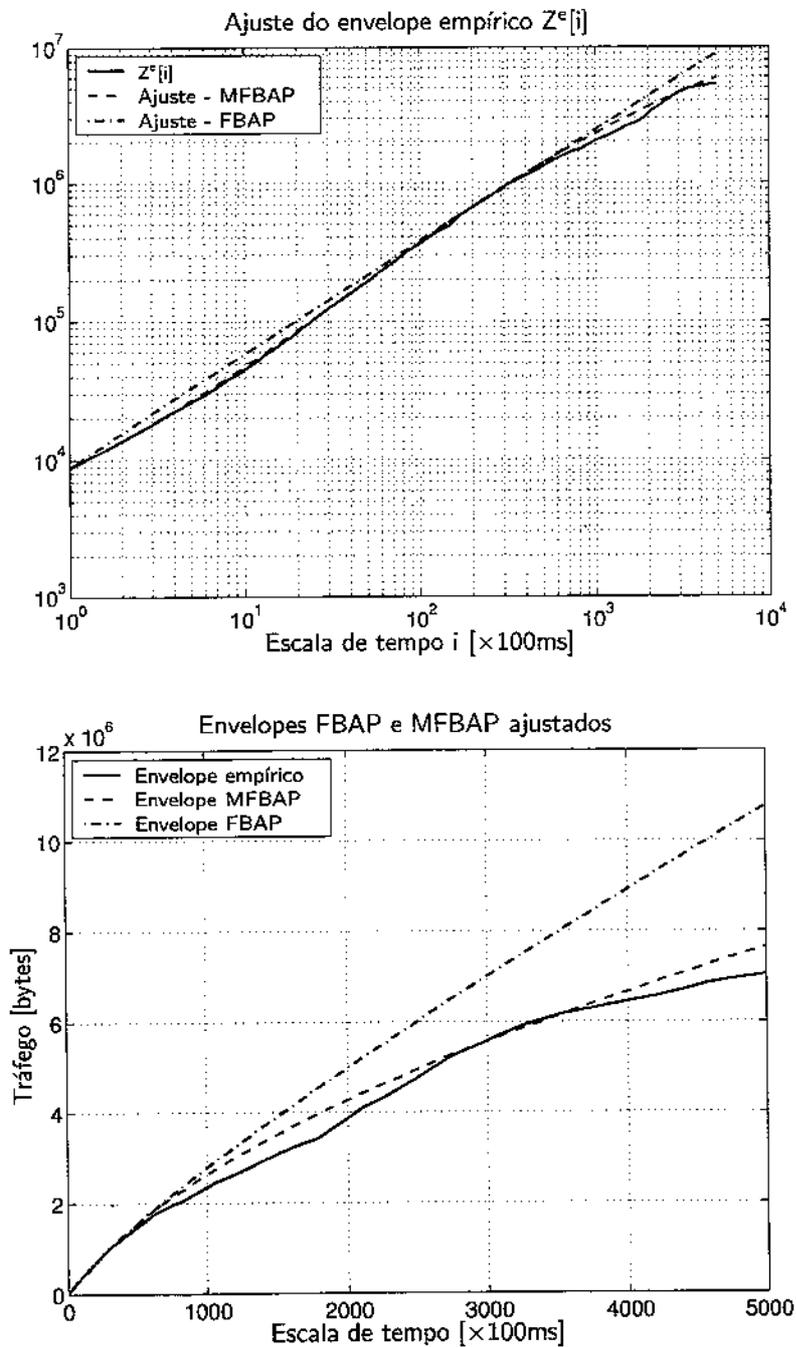


Figura 2.7: Representação do traço LBLTCP3 utilizando os envelopes FBAP e MFBAP.

CAPÍTULO 3

POLICIAMENTO DE TRÁFEGO FRACTAL

Em redes de serviço garantido, as métricas de desempenho contratadas pelos usuários somente podem ser atingidas se os parâmetros do tráfego estiverem em conformidade com os valores negociados na fase de admissão. Para assegurar-se desta conformidade, e para garantir que o tráfego de usuários mal-comportados não prejudique o serviço contratado pelos demais, a rede emprega mecanismos de policiamento que são capazes de detectar e de tratar as violações de maneira adequada.

O policiamento de tráfego tem papel primordial na capacidade da rede de oferecer garantias de desempenho aos usuários. Na literatura, há estudos que se propõem a obter expressões limitantes para métricas de desempenho sob diferentes políticas de escalonamento [25, 71–73]. Tais expressões quase sempre pressupõem que o tráfego é policiado na entrada da rede.

No caso do tráfego fractal, o policiamento geralmente não é trivial. Suponhamos que o tráfego seja consistente com a auto-similaridade estatística, que é o caso mais simples de fractalidade. Neste caso, vale a relação

$$\Delta\widehat{A}(m\tau) \leq m\Delta\widehat{A}(\tau), \quad (3.1)$$

sendo que a igualdade é atingida somente para $m = 1$ ou para $H = 0$. Tal relação indica que não é suficiente policiar m intervalos consecutivos de tamanho τ para garantir que o tráfego em todo intervalo de tamanho $m\tau$ está em conformidade. Assim, podemos concluir que um mecanismo de policiamento baseado em janelas de tamanho fixo não pode ser utilizado para policiar o tráfego auto-similar. Em [68], foi proposto um mecanismo de policiamento de tráfego auto-similar denominado algoritmo do balde furado fractal (FLB¹). Este algoritmo é um mecanismo de policiamento baseado em janelas que restringe o tráfego a um envelope determinístico cumulativo dado por

$$\widehat{A}(t) = \mu t + \psi t^H, \quad t \geq 0, \quad (3.2)$$

em que os parâmetros do algoritmo são dados pelos parâmetros do processo envelope do FBM, sendo ψ dado pelo produto $k\gamma$.

Uma alternativa ao balde furado fractal é usar o algoritmo do balde furado tradicional para policiar tráfego fractal. Em [37], Viana Neto estudou o uso deste algoritmo para o policiamento de fluxos modelados por meio do processo envelope do FBM. Em [74], Procissi et al. aplicaram o balde furado para policiar o tráfego browniano fracionário.

Esta abordagem tem como principal vantagem a de permitir o uso de vários resultados de análise disponíveis na literatura, que estabelecem limitantes para as métricas de desempenho dos fluxos

¹ *Fractal Leaky Bucket.*

pressupondo que eles são policiados pelo algoritmo do balde furado. Dentre estes resultados, incluem-se os limitantes para a latência e para o atraso que foram obtidos por Stiliadis e Varma em [25]. No presente trabalho, utilizamos estes limitantes para caracterizar o desempenho dos fluxos em uma rede EPON que opera sob a política de escalonamento PSLR, que definiremos no Capítulo 4.

Neste capítulo, estudamos o policiamento de tráfego supondo que o tráfego fractal é representado por meio do modelo FBAP. Apresentamos o algoritmo do balde furado e o algoritmo do balde furado fractal e demonstramos, por meio de um exemplo simples, que o algoritmo do balde furado fractal pode não ser sempre capaz de limitar a explosividade do tráfego e que, por isso, não permite oferecer garantias de desempenho aos fluxos. Analisamos então o policiamento de tráfego FBAP por meio do algoritmo do balde furado. Os parâmetros do algoritmo e os do modelo FBAP são relacionados matematicamente, e o funcionamento do algoritmo é ilustrado por meio de simulação.

3.1 ALGORITMO DO BALDE FURADO

O algoritmo do balde furado é o mecanismo de policiamento mais conhecido e estudado na literatura. Ele é baseado no conceito de permissões, que são geradas a uma taxa constante denominada taxa de drenagem. As permissões são armazenadas dentro de um reservatório de tamanho fixo e um pacote só é aceito na rede se ele puder obter um número suficiente de permissões do mecanismo de policiamento. Do contrário, ele é descartado ou marcado como prioritário para descarte.

Sejam a taxa de drenagem e o tamanho do reservatório representados por ρ e σ , respectivamente. Do ponto de vista analítico, o algoritmo do balde furado restringe o tráfego de um fluxo a um envelope denominado processo de chegadas com limitante linear (LBAP²), dado por

$$\widehat{A}(t; t + \tau) = \rho\tau + \sigma. \quad (3.3)$$

A restrição que o algoritmo do Balde Furado impõe ao tráfego é ilustrada na Figura 3.1, em que $s(t)$ representa o volume de permissões disponíveis no reservatório no instante t , e $l(t)$ o volume de permissões adicionado ao reservatório no intervalo $(0; t]$ [72]. Note que o fato de reservatório de permissões ter capacidade fixa confere ao algoritmo do balde furado uma importante propriedade. Suponha que, no instante inicial t , o reservatório de permissões esteja completamente cheio. Suponha ainda que o volume de tráfego que chega ao policiamento no intervalo $[t; t + \tau - \varepsilon]$, $\varepsilon \rightarrow 0^+$, seja desprezível. Neste caso, o máximo volume de tráfego que pode ser aceito pelo algoritmo do balde furado durante o intervalo $(t + \tau - \varepsilon; t + \tau]$ nas condições estabelecidas é limitado por σ , e não por $\rho\tau + \sigma$ como poderíamos supor.

Podemos estender este raciocínio para o caso do reservatório não estar completamente cheio no instante t' , mas contiver um volume $s(t')$ de permissões. Neste instante, o maior surto que pode ser aceito é dado justamente por $s(t')$ que, por sua vez, é menor que σ . Isto é ilustrado na Figura 3.1.

A partir desta análise, podemos concluir que o parâmetro σ representa um limitante para o maior surto que o policiamento pode admitir em qualquer instante de tempo. Dizemos então que o algoritmo do balde furado limita a explosividade do tráfego. Esta característica do balde furado é extremamente importante, dado que ela impõe, em última análise, um limite ao crescimento da ocupação das filas nos comutadores. Sob certas condições, é possível demonstrar esse limite está relacionado à máxima latência e ao atraso do tráfego no sistema [25, 71, 72].

3.2 ALGORITMO DO BALDE FURADO FRACTAL

O algoritmo do balde furado fractal foi proposto em [45, 68] para policiar tráfego auto-similar. Este algoritmo restringe o tráfego ao processo envelope do FBM, dado por

$$\widehat{A}(t) = \mu t + \psi t^H \quad (3.4)$$

²Linear Bounded Arrival Process.

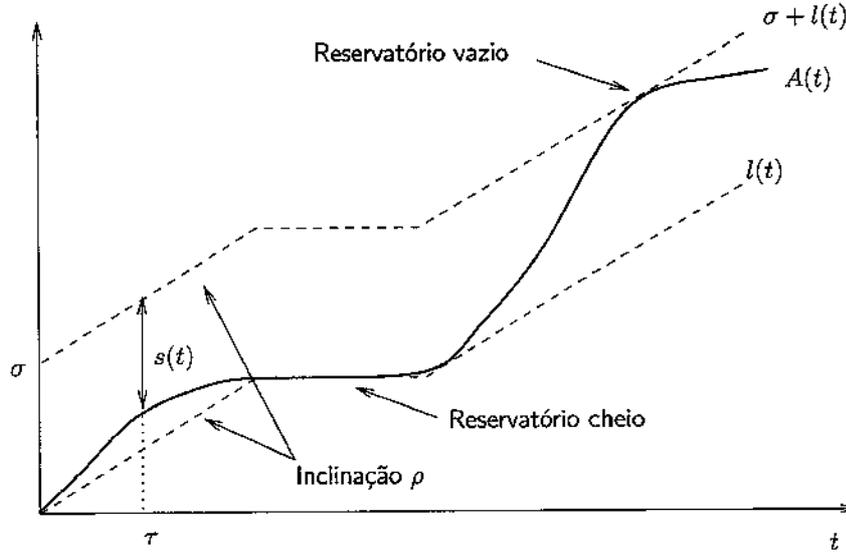


Figura 3.1: Restrição imposta ao tráfego pelo algoritmo do Balde Furado.

em que μ é a taxa média do tráfego, H é o parâmetro de auto-similaridade e ψ é equivalente ao produto $k_i \sigma_i$ no modelo (2.58).

O algoritmo do balde furado fractal é um mecanismo de policiamento por janelas que funciona da seguinte maneira. Seja uma janela de tempo inicial de tamanho Δ . Se, durante esta janela, o tráfego excede o valor médio contratado (dado por $\mu\Delta$), então todos os pacotes que violam o processo envelope $\hat{A}(t)$ neste período são descartados, e a janela é ampliada em Δ unidades de tempo. Esta nova janela se inicia no instante em que houve a violação. O processo é, então, repetido enquanto o tráfego viola a taxa média declarada dentro da janela. No entanto, como alguns pacotes já foram descartados na janela anterior, o número de pacotes a ser descartado é, agora, igual ao número de pacotes que violaram o envelope na janela atual, menos o número de pacotes já descartados nas janelas anteriores. Quando o número médio de chegadas cai abaixo do valor declarado, o tamanho da janela é novamente reduzido a Δ unidades de tempo, e o algoritmo se reinicia.

Resultados de simulação apresentados em [37] demonstram que o algoritmo do balde furado fractal pode policiar o tráfego se houver consistência com o conceito de auto-similaridade estatística.

Quando o tráfego não é exatamente auto-similar, o algoritmo do balde furado fractal pode não ser capaz de realizar o policiamento de maneira eficaz [62, 64]. Isto ocorre porque o algoritmo não é capaz de limitar a explosividade do tráfego e, dependendo de como este se distribui no tempo, ele pode admitir surtos maiores do que os previstos a partir do envelope (3.4).

Podemos ilustrar este problema por meio de um exemplo. Considere um sistema de filas em que o servidor tem taxa constante g e o tráfego de entrada e a latência são dados por $A(t)$ e $Q(t)$. De modo a garantir a estabilidade, supomos que g é maior do que a taxa média do tráfego de entrada. Suponha que o tráfego de entrada seja policiado pelo algoritmo do balde furado fractal, sendo os parâmetros μ , ψ e H especificados pelo usuário. Como foi dito no Capítulo 2, a máxima latência deste sistema é limitada, de maneira determinística, por

$$Q^* = \hat{A}(t^*) - gt^*, \quad (3.5)$$

em que t^* é dado por

$$t^* = \arg \max_{t \geq 0} \hat{A}(t) - gt. \quad (3.6)$$

Suponha agora que o tráfego que chega ao policiamento seja dado por

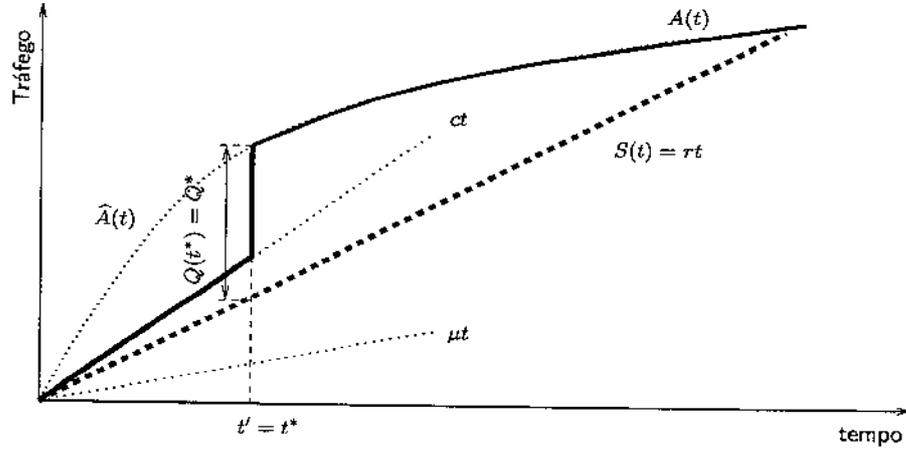


Figura 3.2: Condição em que o balde furado fractal permite oferecer limitantes.

$$A(t) = \begin{cases} ct, & 0 \leq t < t' \\ \mu t + \psi t^H, & t \geq t'. \end{cases} \quad (3.7)$$

A descontinuidade de (3.7) pressupõe que os enlaces têm capacidade infinita. Tal suposição apenas simplifica a análise, que é válida desde que a capacidade dos enlaces seja bem maior do que g , o que geralmente é verdadeiro.

Como é ilustrado na Figura 3.2, a latência máxima do sistema é inferior a (3.5) quando $c > g$. Contudo, suponha que $\mu < c < g$ e que $t' = t^*$. Esta condição é mostrada na Figura 3.3. O volume total de tráfego que chega ao sistema até o instante t^* é dado por ct^{max} . Dado que $c < g$, a fila fica vazia até o instante imediatamente anterior a t^* . Neste instante, a janela do algoritmo do balde furado fractal ocupa todo o intervalo $[0; t^*)$, pois o tráfego chega com taxa superior a μ e inferior ao envelope (3.5) desde $t = 0$. O volume de tráfego que o algoritmo aceita em $t = t^*$ é, então, dado por $\hat{A}(t^*) - ct^*$. Note que este valor é superior a Q^* , de modo que o algoritmo não é capaz de prevenir a violação do limitante calculado a partir do seu próprio envelope.

A partir deste exemplo simples, podemos concluir que o uso do algoritmo do balde furado fractal não permite obter limitantes robustos para métricas de desempenho. Isto restringe a sua aplicação em redes com serviço garantido, para as quais o algoritmo do balde furado ainda é a alternativa mais apropriada.

3.3 POLICIAMENTO DE TRÁFEGO FRACTAL COM O ALGORITMO DO BALDE FURADO

Para estudarmos o policiamento de tráfego fractal por meio do algoritmo do balde furado, consideramos o fato de que este algoritmo pode ser representado por um sistema de filas em que o reservatório é representado por uma fila, de comprimento máximo σ , que é continuamente atendida por um servidor de taxa igual a ρ . Assim, quando uma quantidade qualquer de tráfego chega ao mecanismo de policiamento, a mesma quantidade de permissões é inserida na fila. Se a fila não transborda, o tráfego está em conformidade. Caso contrário, a quantidade de tráfego correspondente ao transbordo é descartada.

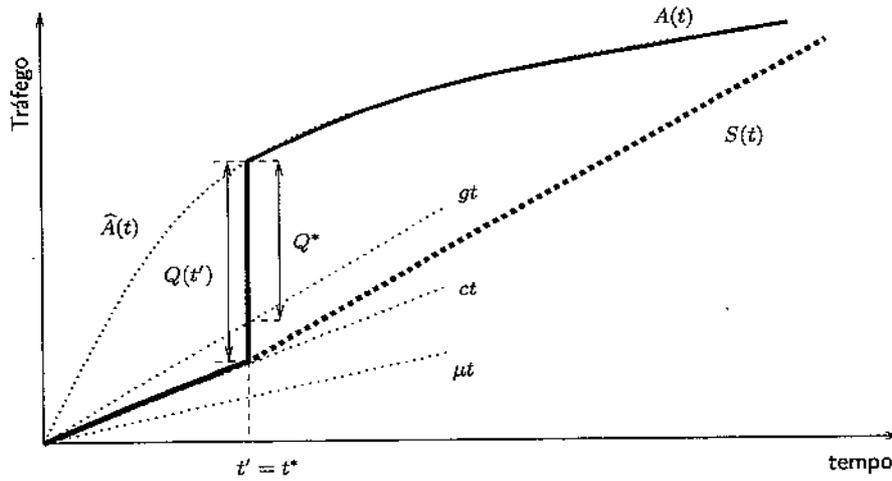


Figura 3.3: Condição em que o balde furado fractal não permite oferecer limitantes.

Suponha que o tráfego que chega ao policiamento seja especificado por um envelope FBAP dado por $\hat{A}_s(\tau) = \mu_s \tau + k_s \gamma_s \tau^{H_s}$. Da análise de filas desenvolvida no Capítulo 2, concluímos que uma dada probabilidade de descarte de tráfego pode ser obtida fazendo a taxa de drenagem igual à banda efetiva do tráfego, e o tamanho do reservatório igual ao correspondente limitante de latência. De (2.74),

$$\sigma = (\rho - \mu_s)^{\frac{H_s}{H_s-1}} (k_s \gamma_s)^{\frac{1}{1-H_s}} H_s^{\frac{H_s}{1-H_s}} (1 - H_s). \quad (3.8)$$

O tráfego que é aceito pelo algoritmo do balde furado é, então, dado por um envelope LBAP de taxa média ρ e de explosividade σ . Em [68], esta relação foi obtida para o processo envelope do movimento browniano fracionário por meio de outra abordagem. Para policiar o tráfego, os parâmetros do algoritmo do balde furado devem ser escolhidos de modo a satisfazer (3.8). Sendo o tráfego representado por um modelo FBAP, devemos então ajustar o algoritmo do balde furado de modo que a taxa de drenagem ρ seja igual à banda efetiva supondo um limitante de latência igual ao tamanho do balde, dado por σ .

Na Figura 3.4 mostramos um exemplo da relação entre a taxa de drenagem e o tamanho do balde. Os parâmetros nominais do tráfego, que tem distribuição marginal gaussiana, são $\mu_s = 0,800$, $k_s \gamma_s = 0,595$ e $H_s = 0,800$. Observe que o tamanho do balde cai rapidamente à medida que a taxa de drenagem aumenta. Podemos observar que valores relativamente baixos de tamanho de balde são obtidos mesmo para taxas de drenagem pouco maiores do que a taxa média do tráfego.

Segundo a relação (3.8), os parâmetros ρ e σ mantêm uma relação inversa entre si quando o tráfego de entrada é FBAP. Note que aquela relação não define completamente o par $(\rho; \sigma)$ que devemos utilizar no algoritmo. Para obtê-lo, podemos representar a escolha dos parâmetros como um problema de otimização, i.e.,

$$\begin{aligned} (\rho^*; \sigma^*) &= \arg \min f(\rho; \sigma) \\ &\text{sujeito a} \\ &\sigma = (\rho - \mu_s)^{\frac{H_s}{H_s-1}} (k_s \gamma_s)^{\frac{1}{1-H_s}} H_s^{\frac{H_s}{1-H_s}} (1 - H_s), \end{aligned} \quad (3.9)$$

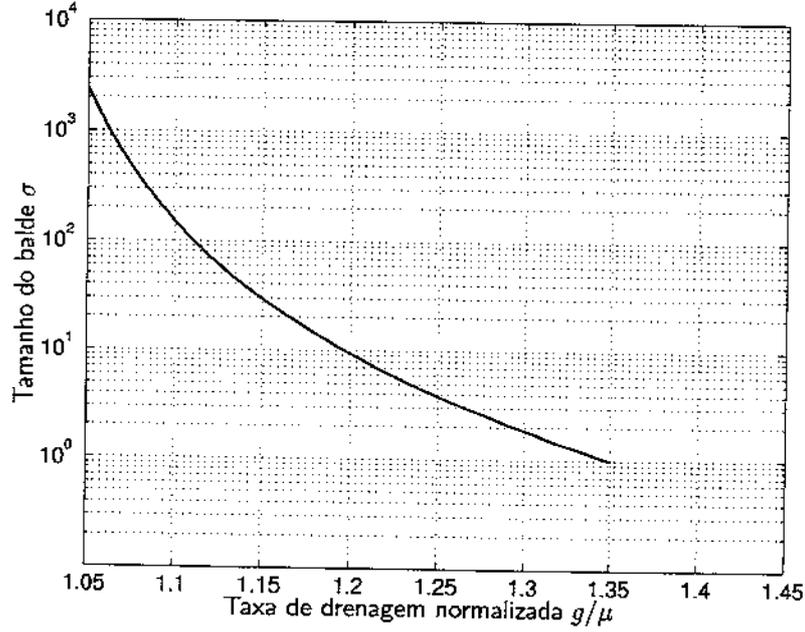


Figura 3.4: Tamanho de balde como função da taxa de drenagem para $\mu_s = 0,800$, $k_s\gamma_s = 0,595$ e $H_s = 0,800$.

em que a função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ representa o compromisso entre a banda requerida pelo tráfego e a sua explosividade.

Esta abordagem foi proposta por Procissi et al. para ajustar os parâmetros do algoritmo do balde furado ao tráfego browniano fracionário [74]. Uma função simples e que permite obter expressões analíticas para ρ^* e σ^* é dada por [74]

$$f(\rho; \sigma) = \sigma + a \cdot \rho, \quad (3.10)$$

em que $a > 0$ é uma medida da preferência por reduzir a banda requerida pelo tráfego frente ao aumento da explosividade. Para essa função, a solução de (3.9) é dada por

$$\rho^* = \mu_s + k_s\gamma_s H_s a^{H_s-1} \quad (3.11)$$

$$\sigma^* = k_s\gamma_s a^{H_s}(1 - H_s). \quad (3.12)$$

Note que a esta solução satisfaz a relação

$$\sigma^* = \frac{a - aH_s}{H_s} (\rho^* - \mu_s). \quad (3.13)$$

Deste modo, o tamanho do reservatório σ_s^* cai com o fator auto-similaridade H_s . Este resultado indica que a latência cai mais rapidamente com o aumento de banda para tráfego com maior auto-similaridade, confirmando assim o que geralmente se verifica na prática.

Para avaliar a eficácia do algoritmo do balde furado, seja o tráfego que chega ao mecanismo de policiamento dado por $\mathbf{A}_i(\tau) = \mu_i\tau + \gamma_i\mathbf{Z}_i(\tau)$ e \mathcal{P}_i a probabilidade de descarte de pacotes em qualquer instante. É fácil concluir que $\mathcal{P}_i = 1$ se $\mu_i \geq \mu_s$. Quando $\mu_i < \mu_s$, os resultados apresentados no Capítulo 2 levam a

$$\begin{aligned}
\mathcal{P}_i &= \mathbb{P} \left\{ \sup_{\tau \geq 0} \mathbf{Q}_i(t + \tau) > \sigma \right\} \\
&\geq \sup_{\tau \geq 0} \mathbb{P} \left\{ \mathbf{Z}_i(1) > \frac{\sigma + (\rho - \mu_i)\tau}{\gamma_i \tau^{H_i}} \right\} = \\
&= \bar{F}_{\mathbf{Z}_i} \left[\frac{(\rho - \mu_i)^{H_i}}{\gamma_i^{H_i} (1 - H_i)^{1-H_i}} \sigma^{1-H_i} \right]. \tag{3.14}
\end{aligned}$$

Assim, quando os parâmetros μ_i , γ_i e H_i assumem seus valores nominais, a probabilidade de descarte é dada por

$$\mathcal{P}_i \geq \bar{F}_{\mathbf{Z}_i}(k_s),$$

em que $\bar{F}_{\mathbf{Z}_i}$ é a distribuição residual do tráfego de entrada. Quando a probabilidade de violação é suficientemente pequena, podemos supor que

$$\mathcal{P}_i \simeq \bar{F}_{\mathbf{Z}_i}(k_s). \tag{3.15}$$

Escolhemos então o parâmetro k_s com base nesta relação, supondo ser conhecida a distribuição $\bar{F}_{\mathbf{Z}_i}$. Note que a probabilidade de descarte cresce à medida em que crescem os parâmetros μ_i , γ_i e H_i . Note ainda que o algoritmo do balde furado também policia (ainda que de maneira implícita) a distribuição $\bar{F}_{\mathbf{Z}_i}$ pois, se k_s é especificado com base em uma distribuição de cauda mais leve do que a real, a probabilidade de descarte de fato será maior do que aquela dada por (3.15).

Cabe observar que a sensibilidade do algoritmo a variações nos parâmetros especificados depende dos parâmetros ρ e σ escolhidos, já que o tráfego que é aceito pelo policiamento não tem envelope FBAP, mas sim um envelope linear dado por (3.3). Em [75], demonstramos que este envelope é maior do que o envelope FBAP para todo $\tau > 0$, exceto por um ponto de tangência. Por isso, o algoritmo tende a aceitar mais tráfego do que o inicialmente especificado. Este é o principal argumento contrário ao uso do balde furado no policiamento de tráfego fractal [37, 75].

Contudo, vale observar que o tráfego aceito pelo balde furado tem taxa média igual à banda efetiva do envelope FBAP e também a mesma banda efetiva que seria obtida caso o policiamento restringisse o tráfego a este envelope. Como o volume de tráfego aceito pelo balde furado é maior do que o aceito por um policiamento FBAP ideal, podemos concluir que esse algoritmo não só policia o tráfego, mas também possibilita o uso mais eficiente dos recursos da rede.

Para ilustrar o funcionamento do algoritmo do balde furado com tráfego auto-similar, utilizamos uma série de ruído gaussiano fracionário com 10^6 amostras, gerado pelo método proposto em [66]. Cada amostra corresponde à quantidade de tráfego que chega ao mecanismo de policiamento em um intervalo de tempo fixo. Os parâmetros do tráfego são $\mu_i = 1,470$, $\gamma_i = 0,399$ e $H_i = 0,800$. Supomos que o tráfego seja distribuído uniformemente dentro de cada intervalo, o que o torna mais suave e dificulta a detecção de violações nos parâmetros.

Para o envelope FBAP nominal, escolhemos um parâmetro k_s igual a 3,719, o que leva a uma probabilidade de violação nominal de 10^{-4} . Dado que os parâmetros μ_i , γ_i e H_i são fixos, as violações de taxa média, parâmetro de escala e parâmetro de auto-similaridade são obtidas variando os parâmetros nominais μ_s , γ_s e H_s . Escolhemos, para o algoritmo do balde furado, uma taxa de drenagem igual a $1,2\mu_s$, sendo o tamanho do reservatório dado por (3.8).

As Figuras 3.5–3.7 mostram a probabilidade de descarte correspondente a violações de taxa média, de parâmetro de escala e de parâmetro de auto-similaridade. Aproximamos esta probabilidade pela frequência relativa dos ciclos ativos em que há descarte de tráfego. Note que o valor aumenta à medida que os parâmetros reais do tráfego crescem frente aos nominais, o que confirma a eficácia do policiamento.

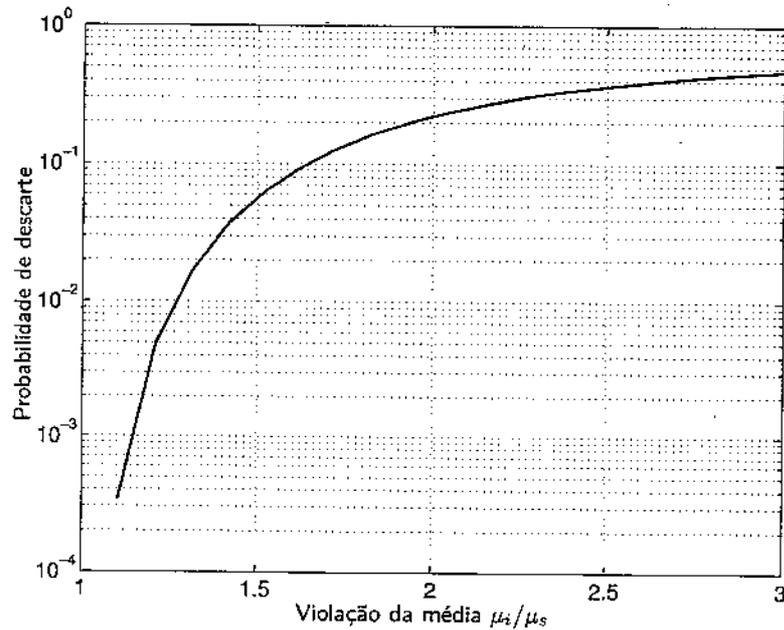


Figura 3.5: Balde furado com tráfego auto-similar. Violação do tráfego médio.

3.4 COMPARAÇÃO ENTRE O BALDE FURADO E O BALDE FURADO FRACTAL

Nesta seção, comparamos o uso do algoritmo do balde furado ao do balde furado fractal para realizar policiamento de tráfego real com características fractais. Por meio de simulações, demonstramos que é possível garantir que o limitante de latência obtido na Seção 2.6.3 é válido quando o tráfego é policiado pelo algoritmo do balde furado, e que isso não é possível quando utilizamos o algoritmo do balde furado fractal.

Para realizar as simulações, utilizamos as mesmas séries de tráfego real que foram apresentadas na Seção 2.6.2. As amostras da série representam o volume de tráfego em intervalos fixos de tempo, conforme foi descrito naquela seção. Por simplicidade, consideramos que o mecanismo de policiamento trata o tráfego como se este fosse infinitamente divisível e não discretizado em pacotes. Consideramos ainda que o tráfego é uniformemente distribuído ao longo do intervalo correspondente a cada amostra, o que o torna mais suave e dificulta a detecção de violações.

Para cada série, simulamos o comportamento da fila de um sistema com um servidor de taxa constante g e que é alimentado pelo tráfego policiado. O objetivo é determinar como a máxima latência da fila varia com g quando o tráfego é policiado pelo balde furado e pelo balde furado fractal, e comparar os resultados obtidos com a relação teórica (2.74) que obtivemos na Seção 2.6.3. No algoritmo do balde furado, ajustamos a taxa de drenagem ρ para que seja igual a g . O tamanho do balde σ é, então, dado por (3.8). No caso do algoritmo do balde furado fractal, os parâmetros são dados pelos parâmetros nominais do envelope FBAP de cada série, que são apresentados na Tabela 2.1.

Nas Figuras 3.8–3.10 são apresentados os resultados obtidos para a série DECPKT2-TCP. O envelope FBAP correspondente é mostrado na Figura 3.8. Os parâmetros do modelo são $\mu_s = 2,38 \times 10^5$, $k_s \gamma_s = 3,85 \times 10^5$ e $H_s = 0,796$, o que resulta em uma probabilidade de violação de 10^{-4} .

A relação entre o limitante teórico para o máxima latência e a taxa de serviço é mostrada na Figura 3.9, que também apresenta os valores de máxima latência obtidos quando o tráfego é policiado

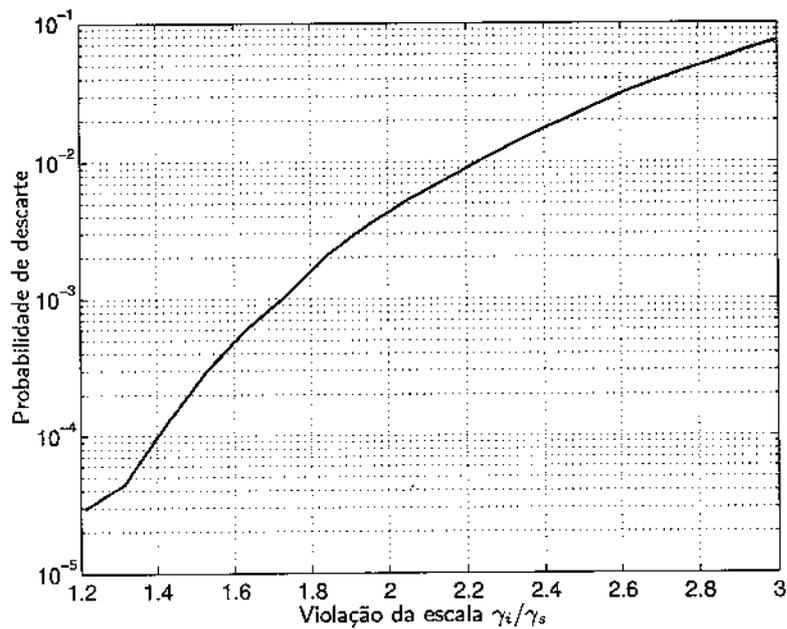


Figura 3.6: Balde furado com tráfego auto-similar. Violação do fator de escala.

por meio do balde furado e do balde furado fractal. As taxas de descarte correspondentes aos dois algoritmos são mostradas na Figura 3.10.

A partir da Figura 3.9, podemos observar que o limitante de latência nunca é violado quando o tráfego é policiado pelo algoritmo do balde furado. Por outro lado, pode haver violação quando utilizamos o algoritmo do balde furado fractal. Ademais, a Figura 3.10 indica que este algoritmo leva a taxas de descarte sempre muito maiores do que o algoritmo do balde furado, mesmo não garantindo aquele limitante. Este resultado confirma a análise apresentada na Seção 3.2. Para as demais séries, os resultados obtidos estão disponíveis em [64]. Em todos os casos, confirmamos a eficácia do uso do algoritmo do balde furado no policiamento de tráfego fractal.

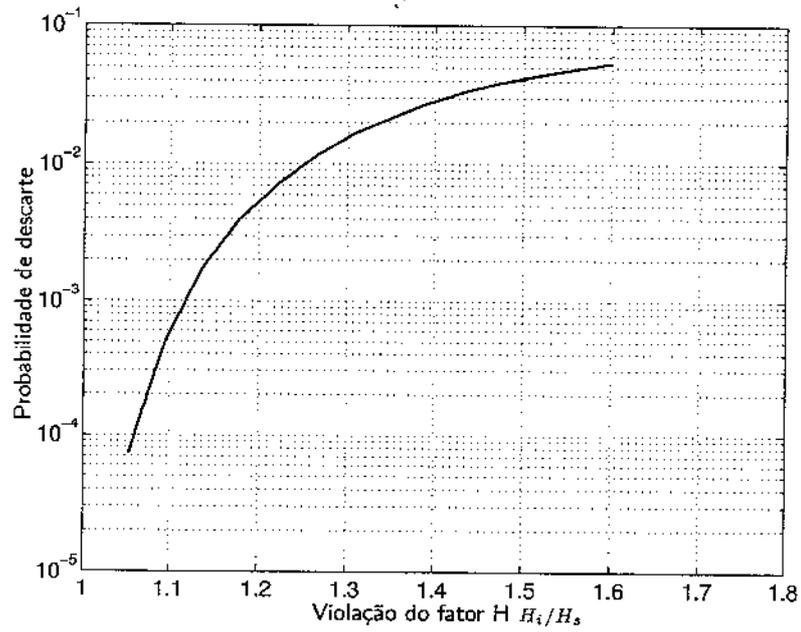


Figura 3.7: Balde furado com tráfego auto-similar. Violação do fator H .

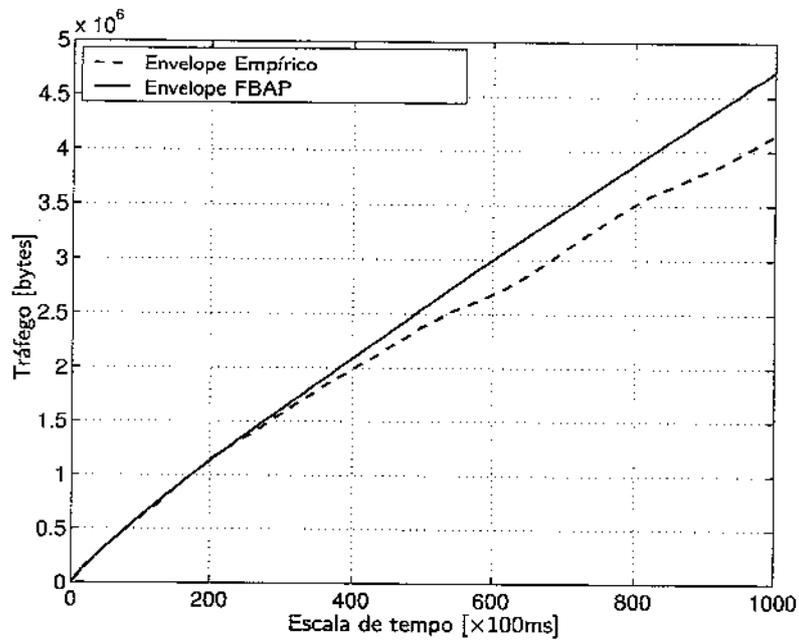
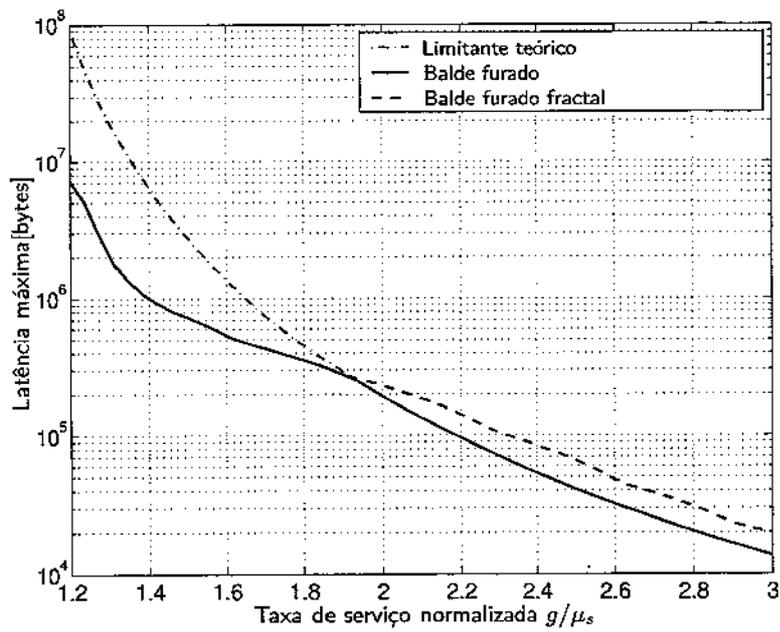


Figura 3.8: Policiamento de tráfego FBAP real. Envelope FBAP para a série DECPKT2-TCP.



(a) Latência máxima

Figura 3.9: Policiamento de tráfego FBAP real. Limitante de latência.

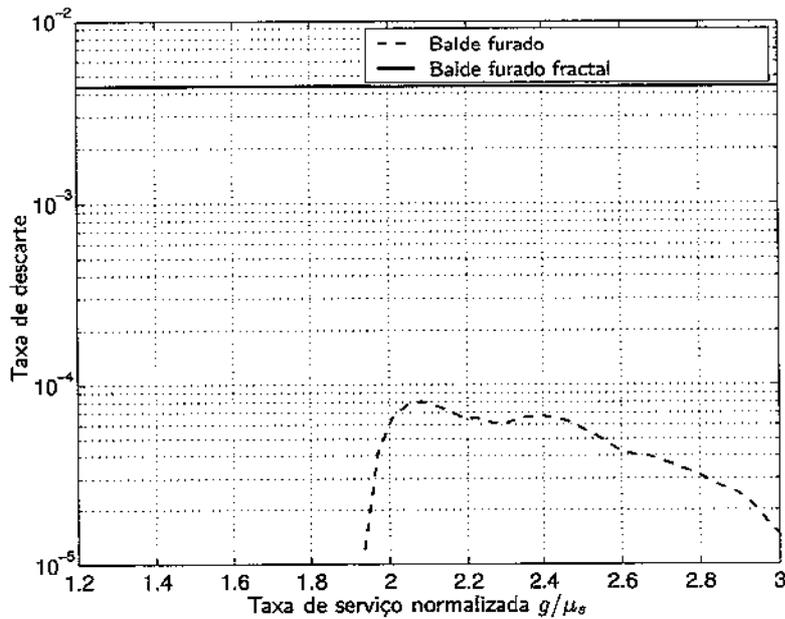


Figura 3.10: Policiamento de tráfego FBAP real. Taxa de descarte.

CAPÍTULO 4

ESCALONAMENTO PSLR PARA REDES EPON

Neste capítulo, definimos uma nova política de escalonamento para redes EPON. Esta política, denominada Compartilhamento Proporcional com Reserva de Carga (PSLR¹) atende os fluxos internos de cada ONU individualmente, reservando-lhes uma parcela mínima de banda e distribuindo entre si os recursos ociosos da rede. Desta maneira, busca-se compartilhar a capacidade da rede entre os fluxos de maneira justa, tal que cada um deles seja atendido segundo a sua demanda e prioridade, esta refletida em um fator de ponderação numérico.

Apresentamos ainda um algoritmo computacional para a política PSLR. Propomos um protocolo de *polling* baseado em uma mensagem de permissão e outra de requisição. Este protocolo pode ser implementado por sinalização interposta ou por meio de encapsulamento em mensagens do protocolo MPCP padrão.

Finalmente, analisamos o desempenho desta política. Apontamos as condições de estabilidade dos fluxos em relação à banda mínima reservada e ao fator de ponderação. Obtemos limitantes de justiça e de atraso para os fluxos individuais, sendo que os últimos pressupõem o policiamento pelo algoritmo do balde furado. Finalmente, analisamos a complexidade computacional do algoritmo proposto para implementar a política PSLR demonstrando, assim, o seu baixo custo.

4.1 NOTAÇÃO E MODELO DO SISTEMA

Ao longo deste capítulo, consideramos que o canal de subida de uma rede de acesso EPON pode ser representado por um sistema de filas com N fluxos e com um servidor de capacidade τ . A rede contém J ONU, que mantêm filas individuais para cada um dos seus N_j fluxos. Para simplificar a notação, o i -ésimo fluxo da ONU j é dito ser o ij -ésimo fluxo.

A cada fluxo ij , atribuímos uma banda mínima de ρ_{ij} e um fator de ponderação de $w_{ij} \in \mathbb{R}_+$, que representa a sua prioridade. Seja μ_{ij} a taxa média do tráfego gerado pelo fluxo ij e $Q_{ij}(t)$ o volume de tráfego armazenado na ONU no instante t . Por simplicidade, consideramos que $Q_{ij}[n] = Q_{ij}\left(t_{ij}^{(n)}\right)$, em que $t_{ij}^{(n)}$ é o instante de tempo em que se inicia o atendimento do fluxo ij durante o n -ésimo ciclo de *polling*.

Durante a operação da rede, a OLT aloca janelas de transmissão às ONU por meio da estratégia de *polling* intercalado com parada. Cada ONU, por sua vez, divide a sua janela entre os seus fluxos de maneira não-intercalada. No n -ésimo ciclo de *polling*, a OLT destina uma janela de tempo equivalente a $C_j[n]$ bits à j -ésima ONU. Dentro desta janela, a onu transmite $c_{ij}[n]$ bits do fluxo ij . Note que esta quantidade inclui os bits de preâmbulo e de espaçamento entre quadros. Por ora, supomos que a rede é capaz de fracionar os pacotes, evitando levar a sua transmissão ao próximo ciclo.

Supomos ainda que o tempo de passagem entre fluxos de uma mesma ONU é desprezível. Seja h_j o tempo de passagem entre a ONU j e a próxima ONU na ordem de atendimento. Em bits, a duração

¹do inglês *Proportional Sharing with Load Reservation*.

do n -ésimo ciclo é então dada por

$$b[n] = h + \sum_j c_j[n]. \quad (4.1)$$

em que h é o tempo de passagem total da rede, que supomos ser constante e dado por $h = \sum_j h_j$. O instante em que se inicia o n -ésimo ciclo de *polling* é então dado por

$$t^{(n)} = \frac{1}{r} \sum_{k=0}^{n-1} b[k]. \quad (4.2)$$

4.2 CRITÉRIOS DE JUSTIÇA

A função primordial do escalonamento é dividir a capacidade da rede entre os fluxos de modo que cada um deles receba uma parcela justa segundo a sua demanda instantânea, seus requisitos mínimos de desempenho e a sua prioridade. No n -ésimo ciclo de *polling* de uma rede EPON, esta divisão é feita alocando a cada fluxo ij uma janela de transmissão de tamanho $c_{ij}[n]$, de tal modo que nenhum fluxo receba mais serviço do que a sua demanda e que a capacidade da rede não seja excedida. Matematicamente,

$$c_{ij}[n] \leq Q_i[n] \quad (4.3)$$

$$\sum_{ij} c_{ij}[n] \leq B - h. \quad (4.4)$$

Seja $\mathbf{c}[n] = \langle c_{11}[n], \dots, c_{NJ}[n] \rangle$ o vetor que representa o tamanho da janela alocada a cada um dos fluxos da rede. As restrições (4.3) e (4.4) definem o conjunto de alocações factíveis, dado por

$$\mathcal{C}[n] = \{ \mathbf{c}[n] \in \mathbb{R}_+^{NJ} \mid (c_{ij}[n] < Q_{ij}[n]) \wedge (\mathbf{c}[n] \cdot \mathbf{1}^T \leq B - h) \}. \quad (4.5)$$

Note que o conjunto $\mathcal{C}[n]$ é fechado para todo n . Para determinar o vetor $\mathbf{c}^*[n] \in \mathcal{C}[n]$ que representa a divisão justa dos recursos entre os fluxos, precisamos definir um critério de justiça que reflita as relações entre a prioridade e a demanda de cada um dos fluxos. Na literatura, os principais critérios de justiça são o max-min, o proporcional e o de utilidade [76]. Matematicamente, o critério max-min é dado pela seguinte definição:

Definição 4.1 (Critério max-min). Um vetor $\mathbf{c}[n] \in \mathcal{C}[n]$ representa uma alocação max-min se, para dois fluxos ij e lm que têm demanda reprimida,

$$\frac{c_{lm}[n]}{w_{lm}} \leq \frac{c_{ij}[n]}{w_{ij}} \wedge \frac{c'_{lm}[n]}{w_{lm}} < \frac{c_{lm}[n]}{w_{lm}}, \quad (4.6)$$

para qualquer outro vetor $\mathbf{c}'[n] \in \mathcal{C}[n]$ em que $c'_{ij}[n] > c_{ij}[n]$.

Em [77], demonstra-se que a alocação max-min é única. O critério max-min leva então a uma alocação em que não se pode melhorar o atendimento a um fluxo sem prejudicar outro mesma ou inferior prioridade. Isto leva a rede a satisfazer os fluxos em ordem tal que os fluxos de menor demanda sejam plenamente atendidos e os fluxos de maior demanda sejam reprimidos. Por isso, diz-se que o critério max-min atribui maior prioridade aos fluxos que têm menor demanda. Esta prioridade é eliminada pelo critério proporcional, que é matematicamente definido como:

Definição 4.2 (Critério proporcional). Um vetor $\mathbf{c}[n] \in \mathcal{C}[n]$ representa a alocação justa pelo critério proporcional se e somente se, para qualquer outro vetor $\mathbf{c}'[n] \in \mathcal{C}[n]$,

$$\sum_{ij} w_{ij} \frac{c'_{ij}[n] - c_{ij}[n]}{c_{ij}[n]} < 0. \quad (4.7)$$

No critério de utilidade, associa-se a cada fluxo ij uma função utilidade $U_{ij}(c_{ij}[n])$ que representa o nível de desempenho ou de satisfação atingido quando o fluxo recebe uma janela de transmissão de tamanho $c_{ij}[n]$. Matematicamente, o critério é definido como:

Definição 4.3 (Critério de utilidade). Um vetor $c[n] \in \mathcal{C}[n]$ representa a alocação justa pelo critério de utilidade se ela resulta da solução do problema de otimização

$$c^*[n] = \arg \max_{c[n] \in \mathcal{C}[n]} \sum_{ij} w_{ij} U_{ij}(c_{ij}[n]), \quad (4.8)$$

em que $U_{ij}(c_{ij}[n])$ é uma função crescente que representa o benefício alcançado pelo fluxo ij quando ele recebe um serviço de $c_{ij}[n]$ no ciclo n .

Dado que o conjunto $\mathcal{C}[n]$ é fechado, pode-se garantir a existência e a unicidade de uma solução para (4.8) desde que $U_{ij}(c_{ij}[n])$ seja contínua e \cap -convexa para todo fluxo ij . Note que o critério de utilidade representa uma generalização dos critérios max-min e proporcional, que são obtidos, respectivamente, para

$$U_{ij}(c_{ij}[n]) = \begin{cases} c_{ij}[n], & ij = \arg \min_{ij} c_{ij}[n] \\ 0, & \text{c.c.}, e \end{cases} \quad (4.9)$$

$$U_{ij}(c_{ij}[n]) = \log c_{ij}[n]. \quad (4.10)$$

Na prática, a função utilidade varia de acordo com o tipo de tráfego. Cabe ressaltar, contudo, que a suposição de \cap -convexidade é bastante razoável, dado que o benefício marginal tende a ser menor à medida em que a banda alocada cresce. Isto é particularmente verdadeiro para aplicações com demanda elástica, tais como *e-mail* e transferência de arquivos. No caso de aplicações multimídia, a função utilidade tende a apresentar uma forma sigmoïdal, que evidencia o baixo desempenho quando se alocam recursos insuficientes para assegurar o desempenho mínimo exigido [76]. Neste caso, contudo, podemos restringir o conjunto $\mathcal{C}[n]$ e recorrer a uma função limitante para $U_{ij}(c_{ij}[n])$ que seja \cap -convexa, de modo a assegurar a unicidade da solução de (4.8) ainda que ela seja, de fato, sub-ótima.

4.3 MÉTRICAS DE JUSTIÇA

Para determinar o quão justa é a divisão de recursos promovida por uma política de escalonamento real, precisamos definir alguma medida que reflita a discrepância entre o serviço que ela oferece aos fluxos e o que deveria ser oferecido no caso ideal. Na literatura, uma das métricas mais utilizadas é o limitante de justiça absoluto (AFB²), que é definido como [76, 78]

$$\text{AFB} = \max_{ij \in \mathcal{B}(\tau; t)} \max_{\tau; t} \left| \frac{S_{ij}^{(I)}(\tau; t)}{w_{ij}} - \frac{S_{ij}(\tau; t)}{w_{ij}} \right|, \quad (4.11)$$

em que $\mathcal{B}(\tau; t)$ é o conjunto de fluxos latentes no intervalo $[\tau; t]$, $S_{ij}^{(I)}(\tau; t)$ é o serviço oferecido ao fluxo $ij \in \mathcal{B}(\tau; t)$ pela política ideal e $S_{ij}(\tau; t)$ o serviço oferecido pela política real. Em geral, o AFB é calculado com a pressuposição do critério max-min, sendo então a política ideal dada pelo Compartilhamento Generalizado de Processador (GPS³) [76, 78].

Para algumas políticas reais pode ser difícil obter o valor do AFB correspondente. Nestes casos, é usual caracterizar a justiça por meio da comparação entre os serviços oferecidos a dois fluxos ij e $lm \in \mathcal{B}(\tau; t)$ pela mesma política. Define-se, assim, o conceito de limitante de justiça relativa (RFB⁴) que é dado por [76, 78]

² Absolute Fairness Bound.

³ do inglês Generalized Processor Sharing.

⁴ Absolute Fairness Bound.

$$\text{RFB} = \max_{ij, lm \in \mathcal{B}(\tau; t)} \max_{\tau; t} \left| \frac{S_{ij}(\tau; t)}{w_{ij}} - \frac{S_{lm}(\tau; t)}{w_{lm}} \right|, \quad (4.12)$$

Em [78], demonstra-se que, para uma política com conservação de trabalho que pressuponha o critério max-min, o RFB e o AFB satisfazem a relação

$$\frac{1}{2} \text{RFB} \leq \text{AFB} \leq \left(1 - \frac{\min_{ij} w_{ij}}{\sum_{ij} w_{ij}} \right) \text{RFB}. \quad (4.13)$$

No caso de uma política que reserve uma banda mínima para os fluxos, o RFB deve ser calculado de modo a desconsiderar o serviço correspondente à banda reservada aos fluxos ij e lm , já que esta parcela não advém da divisão dos recursos ociosos. Assim, o presente trabalho redefine o RFB como

$$\text{RFB} = \max_{ij, lm \in \mathcal{B}(\tau; t)} \max_{\tau; t} \left| \frac{S_{ij}(\tau; t) - \phi \rho_{ij}}{w_{ij}} - \frac{S_{lm}(\tau; t) - \phi \rho_{lm}}{w_{lm}} \right|, \quad (4.14)$$

em que ρ_{ij} e ρ_{lm} representam a banda reservada pelos fluxos ij e lm , respectivamente, e $\phi \geq 0$ é o maior valor para o qual as frações são ambas positivas. Note que (4.14) é igual a (4.12) para $\rho_{ij} = \rho_{lm} = 0$, de modo que é fácil comparar o RFB de uma política com reserva de banda ao de uma política sem reserva.

4.4 COMPARTILHAMENTO PROPORCIONAL COM RESERVA DE CARGA

Suponha que um contrato de serviço é estabelecido entre a rede e cada um dos fluxos no instante da sua admissão. Neste contrato, a rede se compromete a garantir uma banda mínima de ρ_{ij} a cada fluxo ij , e a dar-lhe prioridade de acordo com um fator de ponderação w_{ij} . Seja B a duração máxima (em bits) desejada para ciclos de *polling*. De modo a evitar que os usuários reservem mais banda do que a rede dispõe, vamos supor que

$$\sum_{ij} \rho_{ij} < r \left(1 - \frac{h}{B} \right). \quad (4.15)$$

Como veremos mais adiante, o valor de B está relacionado ao máximo atraso percebido pelos fluxos durante as transmissões. Neste cenário, vamos considerar que a rede escala os fluxos segundo o critério de utilidade, com $U_{ij}(c_{ij}[n])$ dada por

$$U_{ij}(c_{ij}[n]) = \begin{cases} \log(c_{ij}[n] - P_{ij}), & Q_{ij}[n] \geq P_{ij} \\ 0, & \text{c.c.} \end{cases} \quad (4.16)$$

em que $P_{ij} = \rho_{ij} B / r$. A alocação justa $\mathbf{c}^*[n]$ pode ser obtida, então, por meio do seguinte problema de otimização:

$$\mathbf{c}^*[n] = \arg \max_{\mathbf{c}[n] \in \mathcal{C}[n]} \sum_{ij \in \mathcal{B}[n]} \log(c_{ij}[n] - P_{ij}) \quad (4.17)$$

sujeito a

$$\sum_{ij} c_{ij}[n] \leq B - h \quad (4.18)$$

$$c_{ij}[n] \leq \tilde{Q}_{ij}[n] \quad (4.19)$$

$$c_{ij}[n] \geq P_{ij}, \text{ se } ij \in \mathcal{B}[n] \quad (4.20)$$

$$c_{ij}[n] = \tilde{Q}_{ij}[n], \text{ se } ij \notin \mathcal{B}[n], \quad (4.21)$$

em que $\tilde{Q}_{ij}[n] = Q_{ij}[n-1] - c_{ij}[n-1]$ é a latência do fluxo ij no instante em que a ONU j envia a sua requisição para o ciclo $n+1$. O conjunto $\mathcal{B}[n]$ é o conjunto de fluxos ij para os quais $P_{ij} < \tilde{Q}_{ij}[n]$.

A restrição (4.18) limita o tamanho dos ciclos de *polling* a B . Note que a condição de igualdade é atingida sempre que houver demanda reprimida, i.e., houver algum fluxo i para o qual $c_i[n] < Q_i[n]$. A restrição (4.19), que decorre da estratégia de *polling*, estabelece que o tamanho da janela alocada ao fluxo ij não pode exceder o volume de tráfego que ele tem latente no interior da ONU j quando esta envia a mensagem de requisição. Por fim, as restrições (4.20) e (4.21) asseguram que a rede garante ao fluxo ij uma banda mínima de ρ_{ij} sempre que a sua demanda exceder este valor.

Para resolver o problema de otimização (4.17)–(4.21), seja

$$\mathcal{L}[n] = \sum_{ij \in \mathcal{B}[n]} w_{ij} \log(c_{ij}[n] - P_{ij}) - \lambda[n] \left(\sum_{ij} c_{ij}[n] - B + h \right), \quad (4.22)$$

em que $\lambda[n] \geq 0$ é o multiplicador de Lagrange associado com a restrição (4.18). Seja $\mathcal{A}[n]$ o conjunto de fluxos para os quais a restrição (4.19) está ativa, i.e., o conjunto de fluxos têm a sua demanda completamente atendida pela rede. Dado que essa restrição está ativa sempre que (4.21) também está, é fácil concluir que $\mathcal{A}[n] \supset \mathcal{B}[n]$. Assim, para todo fluxo $ij \in \mathcal{A}[n]$ temos que $c_{ij}[n] = \tilde{Q}_{ij}[n]$. Para os demais fluxos, as condições de otimalidade estabelecem que

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda} &= \sum_{ij} c_{ij}[n] - B + h \\ &= \sum_{ij \in \mathcal{A}[n]} \tilde{Q}_{ij}[n] + \sum_{ij \notin \mathcal{A}[n]} c_{ij}[n] - B + h = 0, \end{aligned} \quad (4.23)$$

$$\frac{\partial \mathcal{L}}{\partial c_{ij}} = \frac{w_{ij}}{c_{ij}[n] - P_{ij}} + \lambda[n] = 0, \quad ij \in \mathcal{A}[n]. \quad (4.24)$$

Note que $\lambda[n] = 0$ se $\bar{\mathcal{A}}[n] = \emptyset$. Resolvendo o sistema (4.23)–(4.24), obtemos

$$c_{ij}[n] = \begin{cases} \tilde{Q}_{ij}[n], & ij \in \mathcal{A}[n] \\ P_{ij} + w_{ij}\xi[n], & ij \notin \mathcal{A}[n] \end{cases} \quad (4.25)$$

em que

$$\xi[n] = \frac{1}{\sum_{lm \notin \mathcal{A}[n]} w_{lm}} \left\{ B - h - \sum_{ij \notin \mathcal{A}[n]} P_{ij} - \sum_{ij \in \mathcal{A}[n]} \tilde{Q}_{ij}[n] \right\} \quad (4.26)$$

representa o estado da rede no ciclo n .

Note que $b[n] = B$ sempre que $\bar{\mathcal{A}}[n] \neq \emptyset$. Assim, um fluxo $ij \notin \mathcal{A}[n]$ é atendido com uma taxa média dada por

$$\begin{aligned} s_{ij}[n] &= \frac{c_{ij}[n]}{B} \\ &= \rho_{ij} + w_{ij} \frac{\xi[n]}{B}. \end{aligned} \quad (4.27)$$

Assim, a rede destina a um fluxo com demanda reprimida uma banda mínima de ρ_{ij} e uma parcela adicional que é proporcional ao seu fator de ponderação. Por isso, a política dada por (4.25) é chamada de Compartilhamento Proporcional com Reserva de Carga (PSLR⁵).

⁵do inglês *Proportional Sharing with Load Reservation*.

4.5 ALGORITMO COMPUTACIONAL

Apesar da simplicidade de (4.25), a implementação computacional da política PSLR não é trivial. Isto porque não há como determinarmos a priori quais são os fluxos que compõem o conjunto $\mathcal{A}[n]$. Isto impede que utilizemos essa equação para determinar a parcela justa de banda de cada fluxo.

No caso de políticas pacote-a-pacote, é usual encontrarmos relações de alocação semelhantes a (4.25). Contudo, a falta de uma expressão analítica que permita obter $\mathcal{A}[n]$ pode ser contornada por meio da ordenação apropriada dos pacotes a serem atendidos. Intercalando os pacotes de diferentes fluxos, estas políticas atingem implicitamente a regra de alocação.

Em redes EPON, entretanto, o elevado produto entre a largura de banda e o tempo de propagação torna inviável intercalar pacotes de diferentes ONU. A alternativa é, então, utilizar um algoritmo adaptativo que convirja assintoticamente para a solução dada por (4.25). Em [79], Lapsley e Low estudam um problema de controle de fluxo no tempo contínuo que leva a condições de otimalidade semelhantes a (4.23) e (4.24). Esses autores propõem a solução do problema por meio do algoritmo do gradiente projetado [80] em que $\lambda(t)$ é continuamente ajustado na direção oposta a $\partial\mathcal{L}/\partial\lambda$.

Lapsley and Low demonstram que o algoritmo converge para alocação ótima sempre que o conjunto de alocações factíveis $\mathcal{C}[n]$ for fechado e, nele, a função utilidade for contínua, \cap -convexa e monotonicamente crescente [79]. Visto que (4.16) satisfaz estas propriedades, o algoritmo do gradiente projetado pode ser utilizado também para implementar a política PSLR.

O pseudo-código do escalonamento PSLR é apresentado nos Algoritmos 4.1 e 4.2. O algoritmo pode ser descrito da seguinte maneira. Suponha que a OLT destine janelas de transmissão às ONU por meio de um mecanismo de *polling* intercalado com parada, conforme este foi apresentado na Seção 1.3.2. No início do n -ésimo ciclo, a OLT inicia as transmissões enviando uma mensagem de permissão a cada ONU j . Esta mensagem é composta por dois campos:

- a identidade da ONU de destino, dada por $j \in \mathbb{Z}_+$;
- o estado da rede, $\xi[n] \in \mathbb{Z}_+$.

Para inicializar o algoritmo, supomos que $\xi[0] = 0$. As mensagens de permissão podem ser enviadas por sinalização interposta ou por meio de encapsulamento em uma mensagem gate do protocolo MPCP. Quando a ONU j recebe uma mensagem de permissão, ela calcula a quantidade de tráfego que cada fluxo ij deve enviar no $(n+1)$ -ésimo ciclo. Esse valor é dado por

$$c_{ij}[n+1] = \max \left\{ \tilde{Q}_{ij}[n] - c_{ij}[n], [c_{ij}^e[n] + d_{ij}[n]]^+ \right\}, \quad (4.28)$$

em que $[z]^+ = \max(0; z)$ e

$$c_{ij}^e[n] = \left[\frac{\rho_{ij}B}{r} + w_{ij}\xi[n] \right]^+, \quad (4.29)$$

sendo que a operação $[z]^+$ arredonda z para cima, de tal modo que não haja fragmentação de pacotes caso apenas z bits fossem transmitidos. Neste caso, o estado $d_{ij}[n]$ representa a quantidade (positiva ou negativa) de serviço que deve ser somada no ciclo $n+1$ para compensar a divisão injusta de recursos causada por esse arredondamento.

Após calcular o valor de $c_{ij}[n+1]$, a ONU atualiza o valor de $d_{ij}[n]$ fazendo

$$d_{ij}[n+1] = P_{ij} + w_{ij}\xi[n] - c_{ij}[n+1] + d_{ij}[n]. \quad (4.30)$$

Para evitar o acúmulo de resíduos, fazemos $d_{ij}[n+1] = 0$ sempre que $d_{ij}[n+1] > 0$ e a janela destinada ao fluxo ij no ciclo $n+1$ for suficiente para atender toda a sua demanda, i.e.,

$$c_{ij}[n+1] = \tilde{Q}_{ij}[n-1] - c_{ij}[n]. \quad (4.31)$$

A ONU então marca os quadros a serem transmitidos durante o ciclo $n+1$ e, então, transmite os quadros marcados na iteração anterior, que totalizam $c_{ij}[n]$ bits. Ao recebê-los, a OLT os envia diretamente ao enlace de saída da rede. Finalmente, a ONU encerra a transmissão do ciclo corrente enviando à OLT uma mensagem de requisição composta dos seguintes campos:

- a identidade lógica da ONU, dada por $j \in \mathbb{Z}_+$;
- a demanda total para o ciclo $n+1$, dada por

$$C_j[n+1] = \sum_i c_{ij}[n+1], \quad (4.32)$$

- a soma dos fatores de ponderação dos fluxos ativos após o serviço do ciclo $n+1$, i.e.,

$$S_j[n+1] = \sum_i w_{ij} \mathbb{I} \{ \tilde{Q}_{ij}[n+1] \neq 0 \}, \quad (4.33)$$

em que a função $\mathbb{I}(\cdot)$ vale 1 quando o seu argumento é verdadeiro ou zero, caso contrário.

A mensagem de requisição pode ser enviada por sinalização interposta ou por encapsulamento em uma mensagem report do protocolo MPCP. Note que a ONU não envia à OLT qualquer informação de fluxos individuais. Este tipo de informação é mantida somente pelas ONU. Assim, a OLT mantém informações de estado apenas para cada ONU. Isto permite tratar os fluxos individualmente e obter, ao mesmo tempo, uma baixa complexidade computacional.

Após receber a requisição da última ONU do ciclo, a OLT calcula o valor de $b[n]$ e, caso haja ONU com fluxos ativos (i.e., $\sum_j S_j[n] > 0$), atualiza o valor de $\xi[n]$ por meio de

$$\tilde{\xi}[n+1] = \left[\tilde{\xi}[n] + \eta(B - b[n]) \right]^+ \quad (4.34)$$

$$\xi[n] = \frac{1}{\sum_j S_j} \tilde{\xi}[n] \quad (4.35)$$

em que $0 < \eta < 1$ é o passo do algoritmo. Note que $\tilde{\xi}[0] = 0$ dado que supusemos $\xi[0] = 0$. Além disso a OLT determina o instante, durante o ciclo $n+1$, em que se deve enviar a mensagem de permissão da ONU j . Este instante é dado por

$$t_j[n+1] = \max \{ t_r[n]; t'_j[n+1] \}, \quad (4.36)$$

em que $t'_j[n+1] = 0$ para $j = 0$ e

$$t'_j[n+1] = t_{j-1}[n+1] + \text{RTT}_{j-1} + \frac{1}{r} C_{j-1}[n+1] + G - \text{RTT}_j, \quad j > 0, \quad (4.37)$$

em que $t_r[n]$ é o instante de tempo em que a OLT recebe a requisição da última ONU na ordem de serviço, RTT_j é o tempo de ida e volta da ONU j e G é o intervalo de guarda necessário para a preparar a camada física para a transmissão e para compensar flutuações no tempo de ida e volta das ONU [6, 17, 19]. Dispondo dos valores de $t_j[n+1]$, a OLT programa o envio das permissões de cada ONU, e um novo ciclo de *polling* se inicia.

Algoritmo 4.1: Algoritmo para implementar a política PSLR — Parte da OLT.

Dados: Passo do algoritmo η e conjunto de ONU \mathcal{O} .

$\tilde{\xi} \leftarrow 0;$
 $\xi \leftarrow 0;$

/ Obtém o RTT e a demanda inicial de cada ONU */*

para cada ONU j faça

- $t \leftarrow \text{relogio} ();$
- enviaPermissão (j, ξ);
- espera por requisição da ONU j ;
- $(C_j; S_j) \leftarrow \text{recebeRequisição} ();$
- $\text{RTT}_j \leftarrow \text{relogio} () - t;$

$t_r \leftarrow t_1 \leftarrow \text{relogio} ();$

$\mathcal{O} \leftarrow \text{ordenaPorRTT}(\mathcal{O});$

/ Laço principal */*

repita

- /* Envia a cada ONU uma mensagem de permissão */*
- para cada ONU $j \in \mathcal{O}$ faça

 - agendaPara ($t_j, \text{enviaPermissão} (j, \xi)$);
 - se $j < N_j$ então

 - $t_{j+1} \leftarrow t_j + \text{RTT}_j + \frac{1}{r} C_j + G - \text{RTT}_{j+1};$
 - se $t_{j+1} < t_r$ então $t_{j+1} \leftarrow t_r;$

/ Recebe os quadros de cada ONU */*

para cada ONU $j \in \mathcal{O}$ faça

- repita

 - | Recebe quadro da ONU j ;
 - até que sejam recebidos C_j bits de dados ;
 - $(C_j; S_j) \leftarrow \text{recebeRequisição} ();$
 - $\text{RTT}_j \leftarrow \text{relogio} () - t_j;$

/ Atualiza o estado da rede ξ */*

$b \leftarrow \text{relogio} () - t_r;$

se $\sum S_j > 0$ então

- $\tilde{\xi} \leftarrow \tilde{\xi} + \eta (B - b);$
- se $\tilde{\xi} < 0$ então $\tilde{\xi} \leftarrow 0;$
- $\xi \leftarrow \tilde{\xi} / \sum_j S_j;$

/ Atualiza a referência de tempo */*

$t_r \leftarrow t_1 \leftarrow \text{relogio} ();$

para sempre

Algoritmo 4.2: Algoritmo para implementar a política PSLR — Parte da ONU.

Dados: Para cada fluxo i da ONU j , os parâmetros de SLA w_{ij} e ρ_{ij} .

$d_{ij} \leftarrow 0 \forall i$;

/* Laço principal */

repita

 espera por permissão da OLT;

$\xi \leftarrow \text{recebePermissão}()$;

 para cada fluxo i faça

$\tilde{Q}_{ij} \leftarrow$ demanda informada no ciclo anterior;

$c_{ij} \leftarrow$ pacotes marcados no ciclo anterior;

 envia os pacotes do fluxo i marcados no ciclo anterior;

$c_{ij}^E \leftarrow P_{ij} + w_{ij}\xi$;

 se $c_{ij}^E - d_{ij} < 0$ então $c_{ij}^E \leftarrow 0$;

 senão

$c_{ij}^E \leftarrow c_{ij}^E + d_{ij}$;

 arredonda c_{ij}^E para não truncar pacotes e não exceder $\tilde{Q}_{ij} - c_{ij}$;

 marca os pacotes correspondentes a c_{ij}^E bits;

 para cada fluxo i faça

$d_{ij} \leftarrow d_{ij} + P_{ij} + w_{ij}\xi - c_{ij}^E$;

 if $d_{ij} > 0 \wedge \tilde{Q}_{ij} - c_{ij} - c_{ij}^E = 0$ then $d_{ij} \leftarrow 0$;

/* Envia requisição */

$C_j \leftarrow \sum_i c_{ij}^E$;

$S_j \leftarrow \sum_i w_{ij} \mathbb{I} \{ \tilde{Q}_{ij} - c_{ij} - c_{ij}^E \neq 0 \}$;

 enviaRequisição ($C_j; S_j$);

para sempre

| Descrição | Parâmetro | Valor |
|-----------------------------|-----------|-----------|
| Capacidade da EPON | r | 1Gb/s |
| Passo do algoritmo PSLR | η | 0,1 |
| Número de ONU | J | 16 |
| Número de filas por ONU | N_j | 48 |
| Tamanho máximo de cada fila | Q_{max} | 500kb |
| Intervalo de guarda | G | 1 μ s |
| Máximo comprimento de ciclo | B | 1Mb |

Tabela 4.1: Exemplo de operação da política PSLR. Parâmetros da EPON simulada.

| Fluxo | g_{ij} [Mb/s] | Tipo | H_{ij} | μ_{ij} [Mb/s] | ONU 1...8 t_{on} [s] | t_{off} [s] | ONU 9...16 t_{on} [s] | t_{off} [s] | Classe | ρ_i [Mb/s] | w_i |
|---------|--------------------|------|----------|----------------------|---------------------------|---------------|----------------------------|---------------|--------|--------------------|-------|
| 1...6 | 4 | LRD | 0,8 | 2 | 0 | 20 | 0 | 30 | 1 | 1 | 0 |
| 7...12 | 6 | LRD | 0,8 | 3 | 0 | 20 | 0 | 30 | 2 | 1 | 1 |
| 13...18 | 6 | SRD | - | 3 | 10 | 20 | 10 | 30 | 3 | 1 | 2 |
| 19...24 | 6 | LRD | 0,6 | 3 | 0 | 20 | 0 | 30 | 4 | 2 | 0 |
| 25...30 | 10 | LRD | 0,8 | 5 | 10 | 30 | 10 | 30 | 5 | 2 | 1 |
| 31...36 | 10 | SRD | - | 5 | 0 | 30 | 0 | 30 | 6 | 2 | 2 |
| 37...42 | 4 | LRD | 0,8 | 2 | 0 | 30 | 0 | 30 | 7 | 0 | 1 |
| 43...48 | 8 | SRD | - | 4 | 0 | 30 | 0 | 30 | 8 | 0 | 2 |

Tabela 4.2: Exemplo de operação da política PSLR. Parâmetros dos fluxos.

4.6 EXEMPLO NUMÉRICO

Nesta seção, ilustramos a operação da política PSLR por meio de um exemplo. Consideramos uma rede de acesso EPON com 16 ONU, cada uma com 48 filas independentes. A distância entre a OLT e as ONU é aleatória e segue uma distribuição uniforme entre 1km e 20km, tal que o tempo de propagação dos pacotes varia entre 5 μ s e 100 μ s. Os demais parâmetros da EPON são apresentados na Tabela 4.1.

Em cada ONU, as filas recebem fluxos de taxa variável independentes. Os parâmetros de tráfego de cada fluxo são apresentados na Tabela 4.2. A capacidade dos enlaces de entrada é representada por g_{ij} . Para cada fluxo, o tráfego corresponde à agregação de 256 fontes *on/off* independentes, cada qual com um tempo médio em atividade de 4kB. As fontes permanecem ativas apenas durante o intervalo [t_{on} ; t_{off}]. Para fluxos com dependência de longa duração (LRD ⁶), os tempos de atividade e silêncio das fontes seguem uma distribuição de Pareto com índice $3 - 2H_{ij}$, em que H_{ij} é o parâmetro de auto-similaridade do tráfego. Os valores médios destes tempos são tais que o fluxo agregado tem taxa média de μ_{ij} . Para os demais fluxos, a distribuição é exponencial e há dependência de curta duração (SRD ⁷). O tamanho dos quadros gerados por cada fluxo segue uma distribuição empírica obtida por Sala e Gummalla no enlace de subida de uma rede de acesso típica [81].

A rede EPON foi simulada por meio do software Omnet++, em que implementamos classes C++ para as ONU, a OLT e o enlace óptico [82]. Para o gerador de tráfego, implementamos uma classe de interfaceamento entre o Omnet++ e o gerador de tráfego implementado por Kramer [83]. A vazão da rede ao longo do intervalo de zero a 30s é mostrada na Figura 4.1. Podemos observar que a vazão média se mantém constante em cerca de 975Mb/s, o que resulta em uma eficiência de 97,5% no aproveitamento da capacidade total da rede. A duração dos ciclos de *polling* é mostrada na Figura 4.2 e o tempo de passagem total da rede é apresentado na Figura 4.3.

A partir da vazão total, podemos calcular a parcela justa que deve ser destinada pela rede a cada um dos fluxos. Sendo r' o valor da vazão, cada fluxo deve receber uma largura de banda dada por

⁶ Long Range Dependency.

⁷ Short Range Dependency.

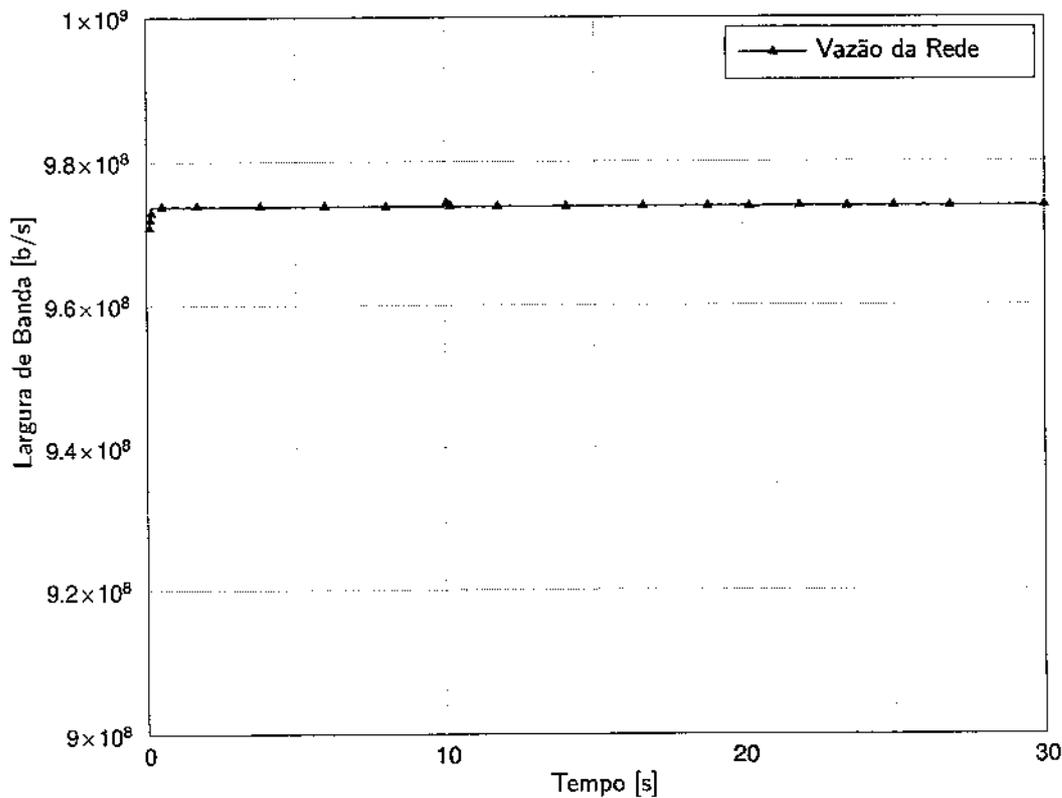


Figura 4.1: Operação da política PSLR. Vazão da rede de acesso.

$$s_{ij} = \rho_{ij} + \frac{w_{ij}}{\sum_{lm \in \mathcal{B}(t)} w_{lm}} r^t, \quad (4.38)$$

em que $\mathcal{B}(t)$ é o conjunto de fluxos que têm tráfego em espera no instante t . Vale observar que os valores médios dos fluxos são tais que todos apresentam demanda reprimida enquanto estão gerando tráfego. Caso contrário, a relação (4.38) não se aplicaria e deveríamos recorrer a (4.25) para realizar o cálculo.

Os valores teóricos de banda de cada fluxo são indicados na Tabela 4.3. Os resultados experimentais obtidos para os fluxos da ONU 1 e 9 são apresentados nas Figuras 4.4–4.7. Nestas figuras, os valores de banda indicados foram obtidos a partir do volume de tráfego de cada fluxo que é atendido em intervalos não-sobrepostos de 0,1s. Podemos observar que os resultados experimentais são bastante próximos daqueles obtidos por meio da relação (4.38). As oscilações observadas nos resultados empíricos se devem à ação do mecanismo de arredondamento que previne o fracionamento de pacotes.

A comparação entre os resultados teóricos e empíricos nos permite concluir que a política PSLR é efetivamente capaz de garantir banda mínima aos fluxos e de proporcionar o compartilhamento justo dos recursos ociosos da rede.

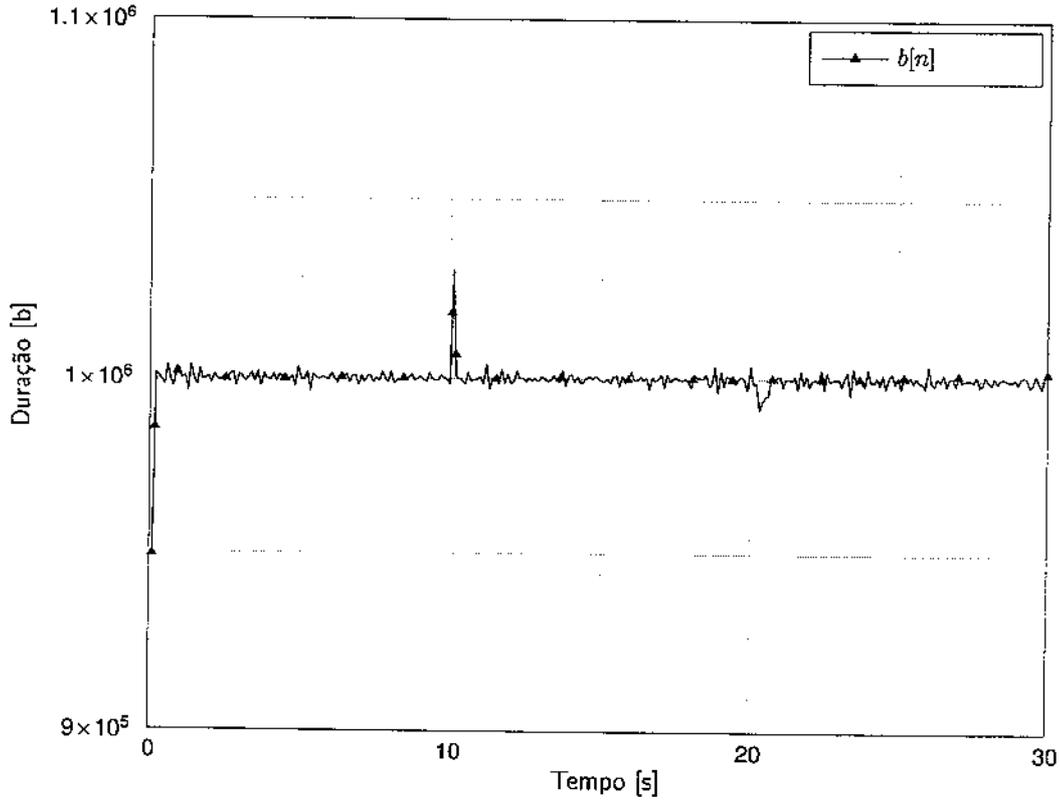


Figura 4.2: Operação da política PSLR. Duração do ciclo de *polling*.

4.7 ANÁLISE DA POLÍTICA PSLR

4.7.1 CONVERGÊNCIA E COMPLEXIDADE COMPUTACIONAL

Na seção anterior, demonstramos empiricamente que a política PSLR é capaz de atender os fluxos segundo o critério de justiça dado por (4.17)–(4.21). Não obstante, é importante demonstrarmos tal capacidade também por meios analíticos, de modo a garantir a validade do resultado para qualquer cenário. Seja então o seguinte teorema:

Teorema 4.4. *A solução de (4.17)–(4.21) pode ser obtida, de forma adaptativa, por meio dos Algoritmos 4.1 e 4.2 com $0 < \eta < 1$. Ademais, a convergência para a solução é linear e de taxa $1 - \eta$.*

Demonstração. Veja Apêndice A.1. □

Para que seja possível implementar a política PSLR em redes reais, precisamos garantir ainda que o algoritmo apresente baixo custo computacional. Do contrário, o tempo necessário para que a OLT e a ONU determinem quais pacotes devem ser transmitidos pode se tornar tão grande diante do aumento do número de elementos (fluxos ou ONU) que a operação da rede se tornaria inviável. Assim, seja o seguinte teorema:

Teorema 4.5. *A complexidade computacional da política PSLR implementada por meio dos Algoritmos 4.1 e 4.2 é linear com o número de fluxos e de ONU, i.e., $\mathcal{O}[(N_{max} + 1)J]$, em que N_{max} é o número máximo de fluxos e J é o número de ONU na rede.*

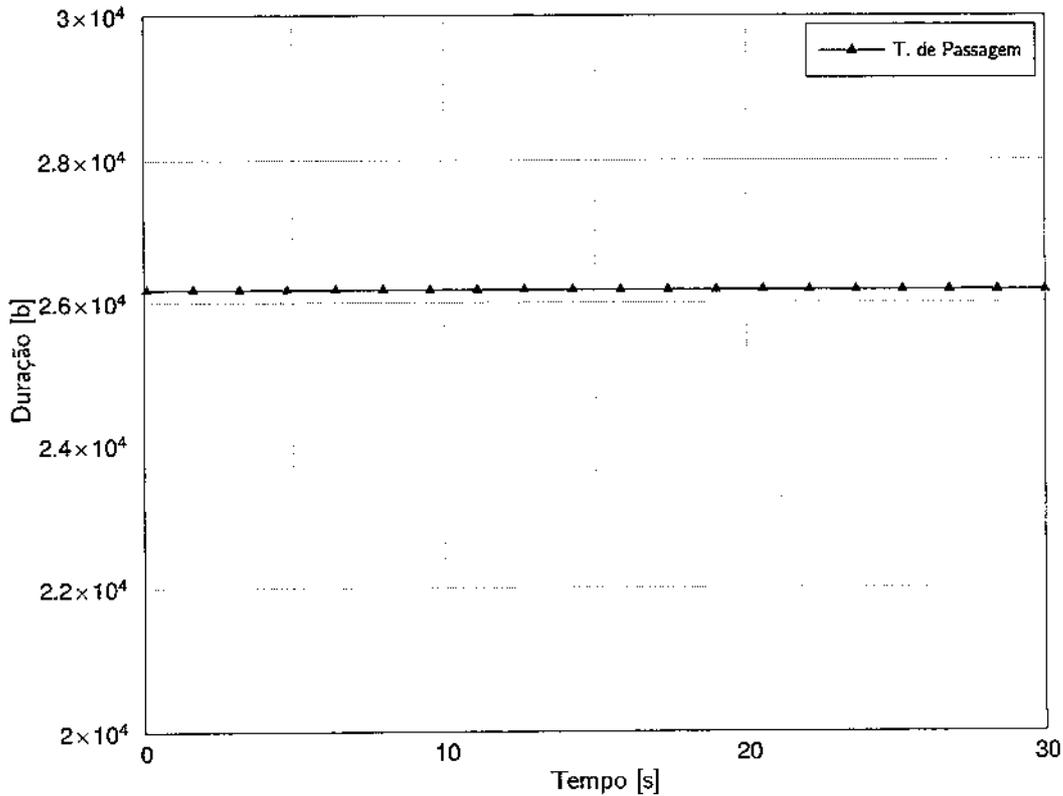


Figura 4.3: Operação da política PSLR. Tempo de passagem total da rede.

Demonstração. A cada ciclo de *polling*, a ONU precisa executar os cálculos (4.28)–(4.30), transmitir os pacotes marcados no ciclo anterior e marcar os pacotes a serem transmitidos no próximo ciclo. Estas operações são feitas uma vez para cada fluxo. Dado que estas relações envolvem apenas escalares, sua complexidade é linear com N_{max} . A ONU precisa ainda realizar as operações (4.29)–(4.33) uma única vez por ciclo, sendo que estas operações envolvem o somatório de até N_{max} escalares. Neste caso, a complexidade das operações também é linear. Concluimos, portanto, que a complexidade computacional do algoritmo da ONU é $\mathcal{O}(N_{max})$.

No caso da OLT, é preciso executar uma única vez a iteração (4.36), o que envolve o somatório de até J escalares (os valores de S_j). Ademais, é necessário determinar, por meio de (4.36) e de (4.37), o instante de tempo em que devem ser enviadas as permissões do ciclo subsequente. Estas operações, juntamente com o próprio envio das permissões, têm complexidade $\mathcal{O}(J)$.

Assim, dado que o algoritmo da ONU é executado J vezes por ciclo e o da OLT apenas uma vez, temos que a complexidade computacional do escalonamento PSLR é a indicada no teorema. \square

4.7.2 ESTABILIDADE

Para analisar a estabilidade da política PSLR, consideramos um cenário de pior caso em que todos os fluxos têm demanda reprimida, i.e., nenhum fluxo pertence a $\mathcal{A}[n]$. Assim, no ciclo n , a rede garante a cada fluxo ij um serviço mínimo dado por

$$c_{ij}^{(min)}[n] = P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left\{ B - h - \sum_{lm} P_{lm} \right\}. \quad (4.39)$$

Como há demanda reprimida, o tamanho dos ciclos converge para B quando $n \rightarrow \infty$. Neste caso,

| Fluxo | Classe | Taxa justa s_{ij} [Mb/s] | | | | | |
|---------|--------|----------------------------|----------|----------|------------|----------|----------|
| | | ONU 1...8 | | | ONU 9...16 | | |
| | | 0...10s | 10...20s | 20...30s | 0...10s | 10...20s | 20...30s |
| 1...6 | 1 | 1,0 | 1,0 | 0 | 1,0 | 1,0 | 1,0 |
| 7...12 | 2 | 1,7 | 1,1 | 0 | 1,7 | 1,1 | 1,5 |
| 13...18 | 3 | 0 | 1,3 | 0 | 0 | 1,3 | 2,0 |
| 19...24 | 4 | 2,0 | 2,0 | 0 | 2,0 | 2,0 | 2,0 |
| 25...30 | 5 | 0 | 2,1 | 2,5 | 0 | 2,1 | 2,5 |
| 31...36 | 6 | 3,4 | 2,3 | 3,0 | 3,4 | 2,3 | 3,0 |
| 37...42 | 7 | 0,7 | 0,1 | 0,5 | 0,7 | 0,1 | 0,5 |
| 43...48 | 8 | 1,4 | 0,3 | 1,0 | 1,4 | 0,3 | 1,0 |

Tabela 4.3: Exemplo de operação da política PSLR. Banda teórica a ser destinada a cada fluxo.

a rede garante ao fluxo ij uma taxa média de serviço s_{ij} , dada por

$$\begin{aligned}
 s_{ij}^{(min)} &= \frac{c_{ij}^{(min)} r}{B} \\
 &= \rho_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left\{ r - \frac{h}{B} - \sum_{lm} \rho_{lm} \right\}. \quad (4.40)
 \end{aligned}$$

Se $s_{ij}^{(min)} > \mu_{ij}$, o fluxo ij é considerado isoladamente estável pois a sua estabilidade não depende da estabilidade dos demais fluxos. Note, contudo, que este critério de estabilidade não é muito forte, dado que o fluxo ainda poderia se tornar instável caso a reserva de banda total da rede varie. Isto pode acontecer, por exemplo, se um novo fluxo é admitido.

Cabe observar que um fluxo pode ser estável ainda que não o seja isoladamente. Isto porque ele pode utilizar uma parcela da banda ociosa da rede para satisfazer a sua demanda. Assim, é suficiente garantir que a rede seja estável para que todos os fluxos apresentem a mesma propriedade. Seja então o seguinte teorema:

Teorema 4.6. *Sob a política PSLR, um dado fluxo i será localmente estável se*

$$\rho_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(r - \frac{r \cdot h}{B} - \sum_{lm} \rho_{lm} \right) \geq \mu_{ij}. \quad (4.41)$$

em que μ_{ij} é a taxa média do fluxo ij . Todos os fluxos serão estáveis se

$$\sum_{ij} \mu_{ij} < r \left(1 - \frac{h}{B} \right). \quad (4.42)$$

Demonstração. Veja Apêndice A.2. □

4.7.3 JUSTIÇA

Para medir a justiça da política PSLR, utilizamos a métrica de RFB apresentada na Seção 4.3. Seja o seguinte teorema:

Teorema 4.7. *Para a política PSLR, o limitante de justiça relativo satisfaz a desigualdade*

$$RFB \leq \frac{1}{\min_{ij} w_{ij}} \max_{ij} [P_{ij} + B - h + 2(L - 1)]. \quad (4.43)$$

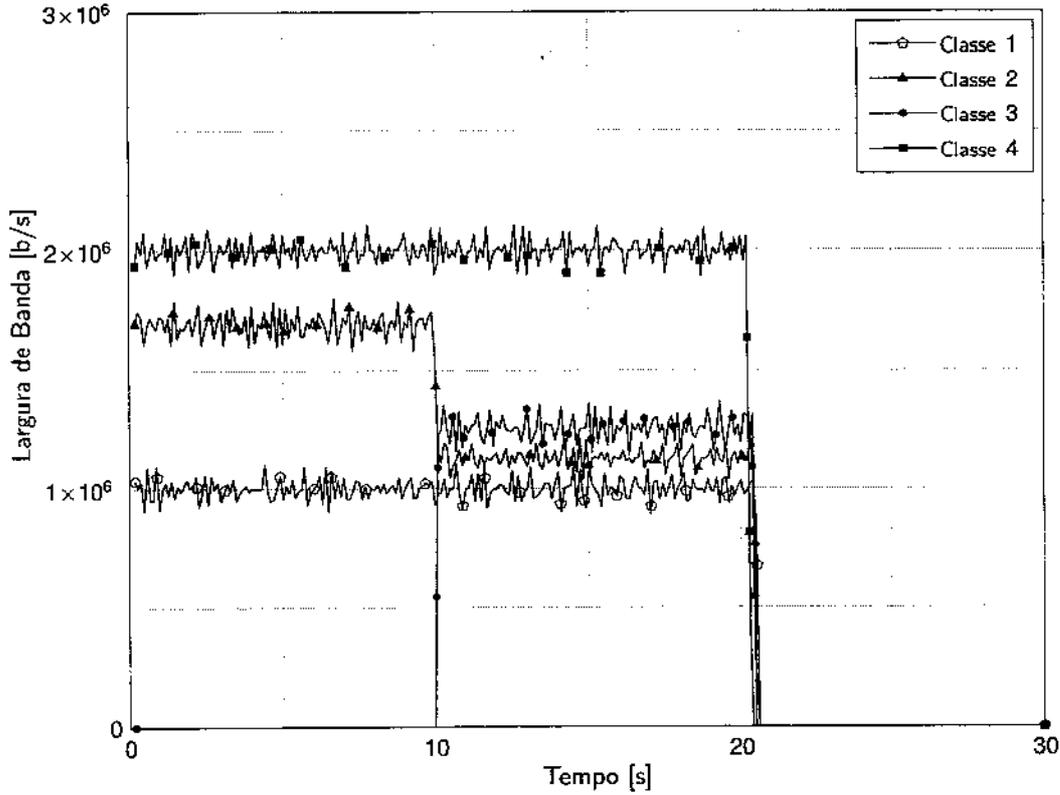


Figura 4.4: Operação da política PSLR. Resultados para os fluxos da ONU 1.

Demonstração. Veja Apêndice A.3. □

Note que o RFB da política PSLR não depende do intervalo de tempo considerado no cálculo. A Tabela 4.4 apresenta o RFB desta e de outras políticas de escalonamento disponíveis na literatura [76]. Para a disciplina PSLR, a tabela apresenta o RFB calculado para reserva de banda e tempos de passagem nulos pois as expressões correspondentes às demais políticas foram assim calculados. Podemos concluir que, se $w_{ij} > 1$ para todo fluxo ij , a política PSLR apresenta um RFB comparável ao das demais políticas apresentadas. Note que, dentre as políticas apresentadas na Tabela 4.4, a política PSLR é a única que pode ser empregada diretamente em redes EPON. Justificamos a comparação com estas disciplinas com o fato de não havermos encontrado, na literatura, outra política de escalonamento EPON para a qual o RFB tenha sido calculado.

A diferença entre os serviços oferecidos a dois fluxos ij e lm quaisquer em um intervalo de n ciclos de *polling* é dado pelo seguinte teorema:

Teorema 4.8. *Seja $B[n_1; n_2]$ o conjunto de fluxos latentes durante todos os ciclos $n_1 \leq n \leq n_2$. Seja ainda o limitante de justiça relativo para ciclos completos definido como*

$$\text{RFB}_{\text{CC}} = \max_{ij, lm \in B[n_1; n_2]} \max_{n_1; n_2} \left| \frac{S_{ij}[n_1; n_2] - \phi \rho_{ij}}{w_{ij}} - \frac{S_{lm}[n_1; n_2] - \phi \rho_{lm}}{w_{lm}} \right|. \quad (4.44)$$

em que $\phi = (n_2 - n_1 + 1)B/\tau$. Para a política PSLR, o RFB_{CC} satisfaz a desigualdade

$$\text{RFB}_{\text{CC}} \leq \frac{2}{\min_{ij} w_{ij}} (L - 1). \quad (4.45)$$

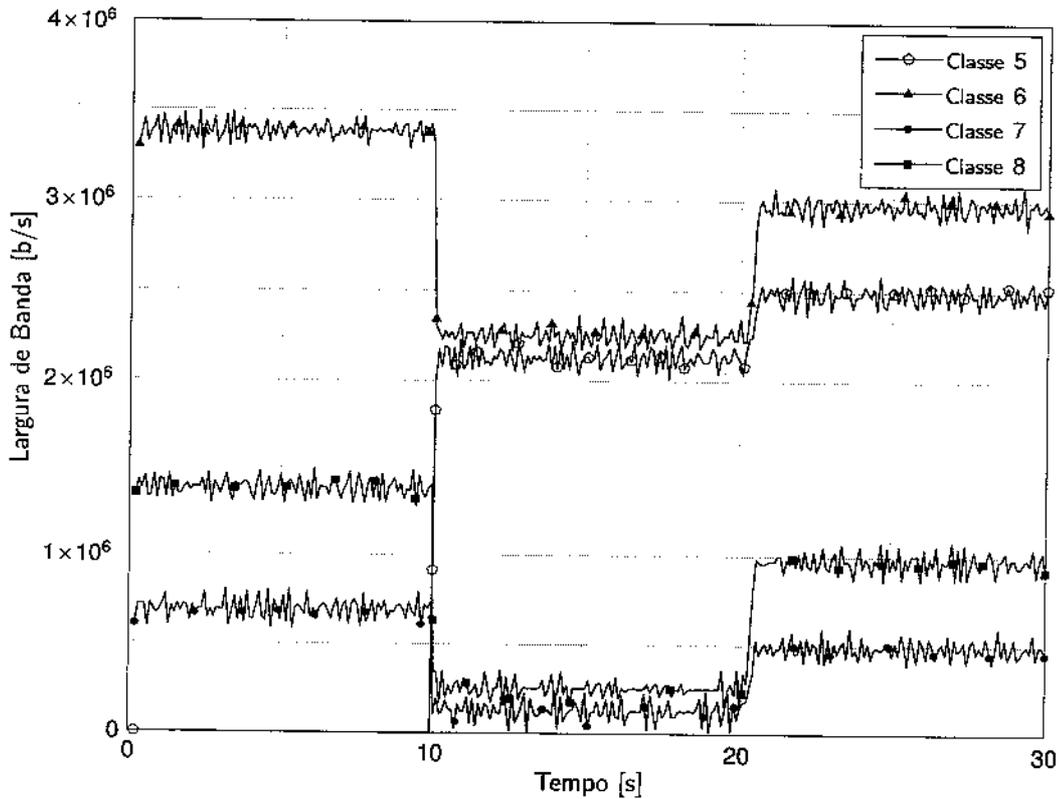


Figura 4.5: Operação da política PSLR. Resultados para os fluxos da ONU 1.

Demonstração. Veja Apêndice A.4. □

Assim, para intervalos de tempo que correspondem a ciclos completos, a política PSLR assegura que a maior quantidade de serviço que um fluxo pode receber além do que corresponde à sua banda reservada não difere da recebida por um outro fluxo por mais do que o termo à direita de (4.45). Note que esta diferença não depende do número de ciclos, o que indica que o escalonamento não pretere ostensivamente um fluxo em favor de outro.

4.7.4 LATÊNCIA E ATRASO

Em redes com serviço garantido, é necessário relacionar os parâmetros de tráfego e as métricas de desempenho dos fluxos. É com esta relação que a rede pode determinar se um novo fluxo pode ser atendido sem degradar o serviço oferecido aos demais, bem como ajustar os parâmetros da política de escalonamento da maneira apropriada.

As principais métricas utilizadas para caracterizar o desempenho percebido pelos usuários são o atraso, a variação do atraso e a taxa de perda de pacotes [24]. Obter expressões que as relacionem aos parâmetros de tráfego e do escalonamento nem sempre é trivial. Em [25], Stiliadis e Varma definem uma classe de políticas de escalonamento denominada servidores latência-taxa para as quais podem ser obtidos limitantes determinísticos para o atraso e para a latência de fluxos individuais. Uma política pertence à classe dos servidores latência-taxa com parâmetros $(\rho_i; \theta_i)$ se o serviço oferecido a qualquer fluxo i que esteja latente em todo o intervalo $[\tau; t]$ satisfaz a relação

$$S_i(\tau; t) \geq [\rho_i(t - \tau - \theta_i)]^+, \quad (4.46)$$

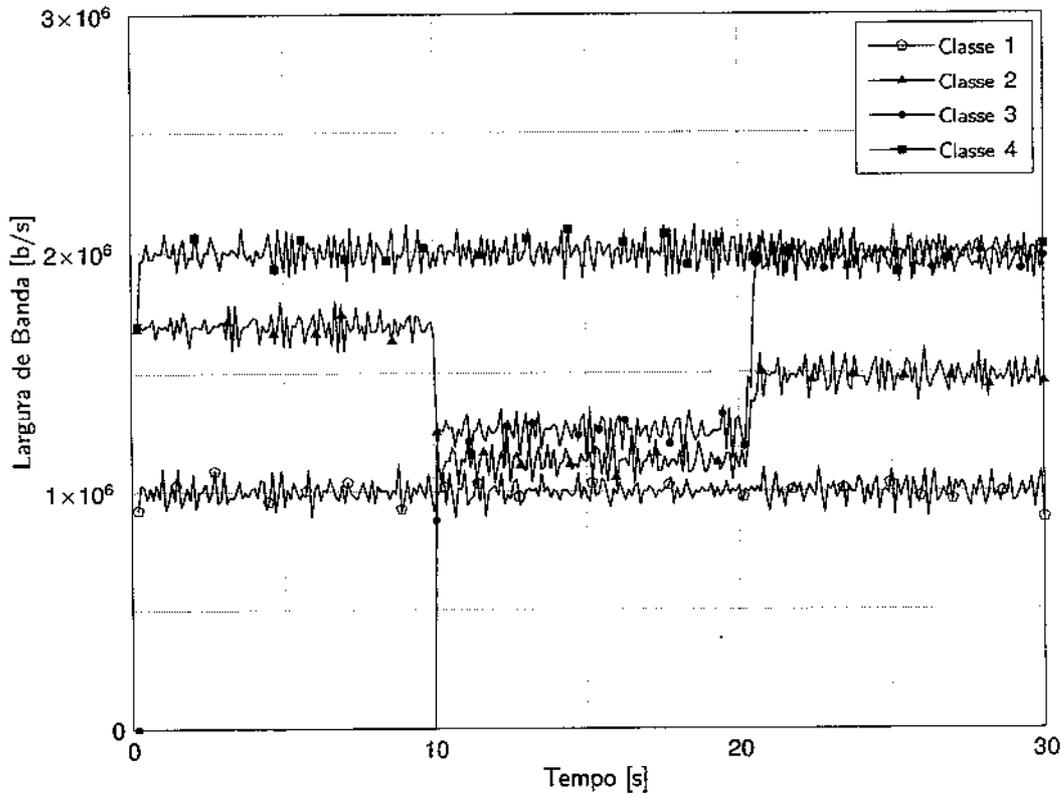


Figura 4.6: Operação da política PSLR. Resultados para os fluxos da ONU 9.

em que ρ_i é a taxa de serviço de longo prazo proporcionada pela política ao fluxo i e θ_i é o seu parâmetro de latência temporal. Este parâmetro representa o maior tempo necessário para que o escalonamento passe a atender o fluxo i continuamente à taxa ρ_i .

Para uma política que pertença à classe dos servidores latência-taxa, é possível obter limitantes para a latência e para o atraso dos fluxos se o tráfego for policiado pelo algoritmo do balde furado. Seja o seguinte teorema [25]:

Teorema 4.9. *Seja um fluxo i cujo tráfego é policiado pelo algoritmo do balde furado com parâmetros ρ_i e σ_i . Se, para este fluxo, o servidor atua como um servidor latência-taxa com parâmetros $(\rho_i; \theta_i)$, a latência e o atraso satisfazem*

$$Q_i(t) \leq \sigma_i + \rho_i \theta_i, \quad (4.47)$$

$$D_i(t) \leq \frac{\sigma_i}{\rho_i} + \theta_i. \quad (4.48)$$

Podemos demonstrar que a política PSLR pertence à classe dos servidores latência-taxa. Seja o seguinte teorema:

Teorema 4.10. *Seja ρ_{ij} a banda reservada a um fluxo ij pela política PSLR. Neste caso, o escalonamento pode ser representado por um servidor latência-taxa $(\rho_{ij}; \theta_{ij})$ em que*

$$\theta_{ij} \leq \frac{(4 + \eta^{-1})B - 2(\phi_{ij} + P_{ij}) + NL}{r}, \quad (4.49)$$

sendo $\phi_{ij} = w_{ij}B / \sum_{lm} w_{lm}$.

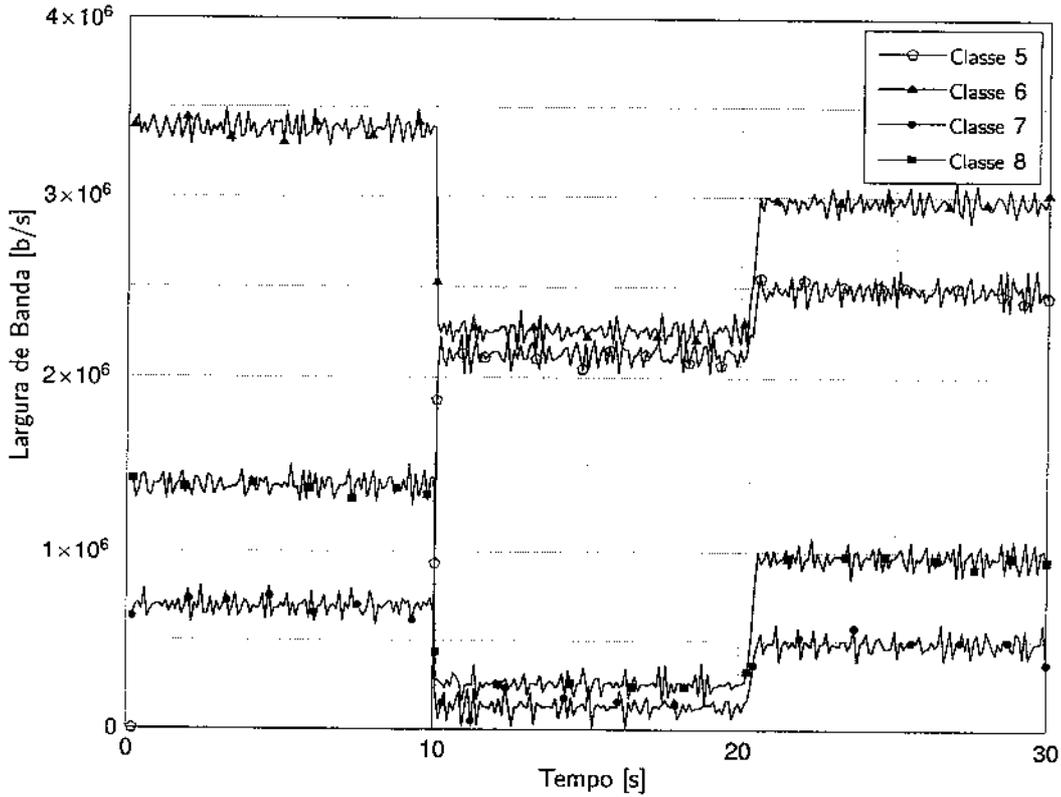


Figura 4.7: Operação da política PSLR. Resultados para os fluxos da ONU 9.

Demonstração. Veja Apêndice A.5. □

Note que o valor de θ_{ij} cresce à medida em que diminui o parâmetro η do algoritmo que implementa a política PSLR. Isto era esperado, pois o algoritmo converge mais rapidamente quanto maior for o valor desse parâmetro. Contudo, vale observar que a escolha do valor de η envolve um compromisso entre a latência temporal (ou a velocidade de convergência) e a oscilação que pode ocorrer quando fazemos η muito próximo da unidade.

A Tabela 4.5 apresenta limitantes para o valor de θ_i para algumas das principais políticas de escalonamento disponíveis na literatura [25]. Note que, dentre as políticas apresentadas na Tabela 4.5, a política PSLR é a única que pode ser empregada diretamente em redes EPON. Não foi possível encontrar, na literatura, uma outra política de escalonamento em EPON que tenha sido caracterizada como latência-taxa e para a qual um limitante para o valor de θ_i tenha sido obtido.

Dado o Teorema 4.9, a latência e o atraso de um fluxo ij que seja policiado pelo algoritmo do balde furado com parâmetros ρ_{ij} e σ_{ij} são limitados por

$$Q_{ij}(t) \leq \sigma_{ij} + \frac{\rho_{ij}}{r} \left[(4 + \eta^{-1}) B - 2 \left(\phi_{ij} + \frac{\rho_{ij} B}{r} \right) + NL \right], \quad (4.50)$$

$$D_{ij}(t) \leq \frac{\sigma_{ij}}{\rho_{ij}} + \frac{1}{r} \left[(4 + \eta^{-1}) B - 2 \left(\phi_{ij} + \frac{\rho_{ij} B}{r} \right) + NL \right]. \quad (4.51)$$

Assim, a política PSLR garante valores máximos para a latência e o atraso dos fluxos individuais. Cabe observar, contudo, que os limitantes (4.50) e (4.51) são bastante conservadores, dado que são de

| Disciplina | RFB |
|------------------------------|---|
| WFQ | $\frac{NL}{\min_{ij} w_{ij}}$ |
| SCFQ | $\frac{2NL}{\min_{ij} w_{ij}}$ |
| Virtual Clock | ∞ |
| WRR | B |
| DRR | $3B$ |
| PSLR ($\rho_{ij} = h = 0$) | $\frac{1}{\min_{ij} w_{ij}} [B + 2(L - 1)]$ |

Tabela 4.4: Comparação entre o RFB da política PSLR e outras políticas de escalonamento.

| Disciplina | Latência θ_i |
|---------------|--|
| WFQ | $\frac{L}{\rho_i} + \frac{L}{r}$ |
| SCFQ | $\frac{L}{\rho_i} + \frac{L}{r}(N - 1)$ |
| Virtual Clock | $\frac{L}{\rho_i} + \frac{L}{r}$ |
| WRR | $\frac{B - \phi_i - L}{r}$, sendo $\phi_i = \frac{w_i}{\sum_j w_j} B$ |
| DRR | $\frac{3B - 2\phi_i}{r}$ |
| PSLR | $\frac{(4 + \eta^{-1})B - 2(\phi_i + P_i) + NL}{r}$ |

Tabela 4.5: Comparação entre a latência da política PSLR e de outras políticas da literatura.

natureza determinística e que pressupõem que os fluxos são atendidos somente pela banda reservada ρ_{ij} , que é inferior mesmo à taxa mínima garantida pela política PSLR, dada por (4.40).

Embora sejam conservadores, os limitantes (4.50) e (4.51) constituem uma primeira aproximação para o desempenho percebido pelos fluxos sob a política PSLR. Quando o tráfego é FBAP, estes limitantes podem ser ainda utilizados junto com os resultados do Capítulo 3 para estimar o valor mínimo de banda reservada ρ_{ij} que a rede deve destinar a um fluxo de prioridade ϕ_{ij} para que não se violem determinados valores máximos de latência ou de atraso. Ademais, obtemos também o valor do parâmetro σ_{ij} do algoritmo do balde furado, lembrando que a taxa de drenagem do algoritmo é, para os limitantes (4.50) e (4.51), coincidente com a banda reservada.

Seja então o tráfego de um fluxo ij qualquer dado pelo processo FBAP $\tilde{A}_{ij}(\tau) = \mu_{ij}\tau + k_{ij}\gamma_{ij}\tau^{H_{ij}}$. No Capítulo 3, demonstramos que os parâmetros ρ_{ij} e σ_{ij} do algoritmo do balde furado ajustado a este tráfego podem ser dados por

$$\rho_{ij} = \mu_{ij} + k_{ij}\gamma_{ij}H_{ij}a^{H_{ij}-1} \quad (4.52)$$

$$\sigma_{ij} = k_{ij}\gamma_{ij}a^{H_{ij}}(1 - H_{ij}), \quad (4.53)$$

em que a é uma constante que representa o compromisso entre a redução da banda e da explosividade do tráfego policiado. Note que

$$\sigma_{ij} = \frac{a - aH_{ij}}{H_{ij}} (\rho_{ij} - \mu_{ij}). \quad (4.54)$$

Substituindo esta relação em (4.50) e (4.51), temos que

$$Q_{ij}^* = \frac{a - aH_{ij}}{H_{ij}} (\rho_{ij} - \mu_{ij}) + \frac{\rho_{ij}}{r} \left[(4 + \eta^{-1}) B - 2 \left(\phi_{ij} + \frac{\rho_{ij}B}{r} \right) + NL \right], \quad (4.55)$$

$$D_{ij}^* = \frac{a - aH_{ij}}{H_{ij}\rho_{ij}} (\rho_{ij} - \mu_{ij}) + \frac{1}{r} \left[(4 + \eta^{-1}) B - 2 \left(\phi_{ij} + \frac{\rho_{ij}B}{r} \right) + NL \right], \quad (4.56)$$

em que Q_{ij}^* e D_{ij}^* são os limitantes de latência e de atraso para o fluxo i . Note que as expressões (4.55) e (4.56) relacionam estas métricas de desempenho à reserva de banda ρ_{ij} , à medida de prioridade relativa $\phi_{ij} = w_{ij}B / \sum_{lm} w_{lm}$ e ao parâmetro a . Por conseguinte, elas podem ser utilizadas para estabelecer garantias de desempenho absolutas aos fluxos.

Vale lembrar, mais uma vez, que estes resultados são conservadores. Expressões mais acuradas devem ser obtidas por meio de análises baseadas em curvas de serviço, tais como as que foram apresentadas para a política GPS em [72, 73].

CAPÍTULO 5

COLETÂNEA DE ALGORITMOS DE ESCALONAMENTO EPON

O escalonamento dos fluxos de subida em uma rede EPON ainda é tema de intensa pesquisa. Neste capítulo, apresentamos as principais propostas disponíveis na literatura e buscamos destacar as suas vantagens e desvantagens. Outras propostas de interesse podem ser encontradas em [20, 84, 85].

5.1 ALOCAÇÃO DINÂMICA DE BANDA PARA MULTIMÍDIA

A alocação dinâmica de banda para aplicações multimídia foi proposta por Choi e Huh [23] considerando a recomendação de prioridade estrita da norma IEEE 802.1d. Nessa política, cada ONU classifica internamente o seu tráfego em três classes de prioridade (alta, média e baixa) e mantém uma fila independente para cada uma delas. A cada ciclo de *polling*, que é regulado pelo protocolo MPCP, a ONU envia uma mensagem report contendo a latência de cada uma dessas filas. A latência, neste caso, representa a demanda da classe de serviço. Com base nesta informação, a OLT determina o número de permissões que devem ser enviadas a cada ONU no próximo ciclo. As permissões de cada classe são enviadas separadamente, mas em uma mesma mensagem gate, e são calculadas considerando que a OLT reserva uma parcela fixa da capacidade útil da rede para o tráfego de alta prioridade de cada ONU. No ciclo n , a OLT destina ao tráfego de alta prioridade de uma dada ONU i um volume de permissões dado por

$$G_i^{(A)}[n] = B_i^{(A)}, \quad (5.1)$$

em que $B_i^{(A)}$ é o tamanho da janela a que corresponde a parcela de banda reservada ao tráfego de alta prioridade daquela ONU. Para o tráfego de média prioridade, a OLT aloca um volume de permissões dado por

$$G_i^{(M)}[n] = \min \left[R_i^{(M)}[n-1]; \frac{R_i^{(M)}[n-1]}{\sum_j R_j^{(M)}[n-1]} \left(B - \sum_i G_i^{(A)}[n] \right) \right], \quad (5.2)$$

em que $R_i^{(M)}[n-1]$ é a demanda dos fluxos de média prioridade que a ONU i informou no final da transmissão do ciclo $n-1$, e B é o tamanho da janela (em bytes) a que corresponde a parcela de banda reservada para o atendimento do tráfego dos usuários. Esta parcela é dada pela capacidade total da rede, deduzida da banda utilizada pelas mensagens do protocolo MPCP e de gerência de rede. Finalmente, o volume de permissões que a OLT destina ao tráfego de baixa prioridade é dado por

$$G_i^{(B)}[n] = \frac{R_i^{(B)}[n-1]}{\sum_j R_j^{(B)}[n-1]} \left(B - \sum_i G_i^{(A)}[n] + G_i^{(M)}[n] \right). \quad (5.3)$$

Assim, a política de alocação dinâmica de banda para multimídia distribui a capacidade da rede segundo uma política de prioridade estrita baseada na demanda de cada classe. O tráfego de alta prioridade é atendido de modo a satisfazer a sua demanda ao máximo, e as duas outras classes são atendidas com a banda remanescente, segundo sua prioridade. Para estas classes, a OLT destina a cada ONU uma parcela de banda proporcional à demanda. Como a OLT determina esta parcela com base na demanda de todas as ONU, podemos dizer que esta política promove o compartilhamento justo dos recursos entre as classes. Em contrapartida, esta política tem as mesmas desvantagens que são inerentes à priorização estrita e que foram citados no início desta seção. Além disso, é difícil estender a política para tratar um maior número de classes ou mesmo fluxos individuais, dado que a OLT e as ONU precisam trocar informações sobre cada classe ou fluxo separadamente, o que resultaria em uma sobrecarga da rede com mensagens de controle.

5.2 POLLING INTERCALADO COM TEMPO DE CICLO ADAPTATIVO

A política de *polling* intercalado com tempo de ciclo adaptativo (IPACT¹) considera que os escalonamentos intra-ONU e inter-ONU são feitos separadamente [18, 19]. Nessa política, a OLT aloca seqüencialmente janelas de transmissão para as ONU com base na sua demanda, dada pelo volume total de tráfego em espera que cada ONU informa ao final de cada janela de transmissão. Se uma ONU não tem tráfego a transmitir em um dado ciclo, a OLT envia-lhe uma permissão com tamanho nulo, de modo a garantir que ela possa informar a sua demanda nos ciclos subseqüentes.

Dado que o escalonamento inter-ONU é feito com base na demanda total de cada ONU, a quantidade de informação que as mensagens de controle precisam transportar é pequena. Por isso, a política IPACT adota a sinalização interposta no lugar do protocolo MPCP como forma de reduzir o *overhead* do controle de acesso. Ademais, ela adota a estratégia de *polling* intercalado ilustrada na Figura 1.6, o que aumenta significativamente a utilização do canal [18, 19]. Nesta estratégia, a OLT mantém uma tabela com o tempo de ida e volta e com a demanda de cada ONU, informações que são utilizadas para determinar o instante em que a OLT deve enviar a permissão da ONU i no próximo ciclo. Para o ciclo n , este instante é dado por

$$t_n^{(i)} = \max \left\{ t_{n-1}^{(i)} + \text{RTT}^{(i)}; t_n^{(i-1)} + \text{RTT}^{(i-1)} - \text{RTT}^{(i)} + W_n^{(i-1)} + G \right\}, \quad (5.4)$$

em que $(i-1)$ representa a ONU que antecede a ONU i na ordem de serviço, $\text{RTT}^{(i)}$ é o tempo de ida e volta da ONU i , $W_n^{(i-1)}$ é a janela (em unidades de tempo) alocada a ONU $(i-1)$ e G é um intervalo de guarda.

Na política IPACT, o tamanho do ciclo de *polling* não é fixo, mas varia com a demanda das ONU. Para evitar que uma ONU monopolize o canal de subida, a OLT limita o tamanho da janela a um valor máximo dado por W_{MAX} . Ademais, as janelas podem ser alocadas de acordo com uma das seguintes estratégias:

- *Serviço limitado.* A OLT aloca uma janela de tamanho igual à demanda da ONU, mas limitada ao valor máximo W_{MAX} ;
- *Serviço fixo.* A OLT ignora a informação de demanda e aloca sempre uma janela de tamanho W_{MAX} ;
- *Serviço com crédito.* A OLT aloca uma janela de tamanho igual à demanda, somada de um crédito fixo ou proporcional que visa a atender o tráfego que chega entre o instante em que a ONU informa a sua demanda e o início da janela;

¹ *interleaved polling with adaptive cycle time.*

- *Serviço elástico.* A OLT aloca as janelas de modo que as N últimas tenham, somadas, tamanho inferior a $N \cdot W_{MAX}$. Desta maneira, uma dada ONU pode receber a parcela de banda não utilizada pelas anteriores, o que pode levar a uma janela de tamanho até $N \cdot W_{MAX}$.

Em [19], são apresentados resultados de simulação que indicam que o serviço fixo leva a valores de latência e de atraso de pacotes bem maiores do que as demais estratégias que, geralmente, apresentam resultados semelhantes.

A política IPACT tem como principal vantagem o maior aproveitamento do canal, que se obtém pelo uso da estratégia de *polling* intercalado e da sinalização interposta. Em contrapartida, essa política não assegura o compartilhamento justo dos recursos da rede, pois o *polling* intercalado exige que as janelas de transmissão sejam alocadas antes que a OLT disponha da informação de demanda de todas as ONU. Finalmente, cabe destacar que a política não prevê o suporte ao tratamento diferenciado dos fluxos, pois não estabelece como o escalonamento intra-ONU deve ser realizado.

5.3 GPS COM BANDA EFETIVA DETERMINÍSTICA

Na política de GPS com banda efetiva determinística (DEB-GPS²), cada ONU mantém filas individuais para classes de serviço ou fluxos individuais. As filas são classificadas em filas de serviço garantido ou filas de melhor esforço, de acordo com os requisitos de desempenho de cada fluxo. No caso das filas de serviço garantido, supõe-se que os fluxos são policiados por meio do algoritmo do balde furado. Neste caso, especifica-se ainda o atraso máximo tolerado, que é utilizado para determinar a banda mínima a cada fluxo. Para isso, utilizam-se resultados da teoria de bandas efetivas determinísticas.

Durante um ciclo de *polling*, a OLT aloca janelas de transmissão para cada ONU com base na banda efetiva total das filas de serviço garantido. Para assegurar que a capacidade total da rede não é excedida, emprega-se um mecanismo de controle de admissão. Em cada ONU, as filas são atendidas segundo uma aproximação da política GPS com ponderação [71, 72], sendo que o peso de cada fila é calculado de modo a garantir que os fluxos com serviço garantido recebam a banda mínima necessária para satisfazer o seu requisito de desempenho. Este cálculo é atualizado sempre que um novo fluxo é admitido na rede. Se, em um dado ciclo, um desses fluxos não utiliza toda a banda mínima, a parcela remanescente é redistribuída pelo GPS entre as demais filas de serviço garantido e as de melhor esforço.

As principais vantagens da política DEB-GPS são o tratamento individualizado dos fluxos e o suporte a garantias determinísticas de desempenho. A sua principal desvantagem é a divisão injusta dos recursos ociosos entre fluxos de ONU diferentes. Isto porque, se um fluxo não utiliza toda a banda que lhe é destinada, o excedente é redistribuído apenas entre os outros fluxos da mesma ONU. Outra desvantagem desta política é a necessidade do controle de admissão para novos fluxos, o que pode resultar em um *overhead* significativo no caso de fluxos de curta duração.

5.4 POLLING COM GARANTIA DE BANDA

Na política de *polling* com garantia de banda (BGP³), as ONU são divididas entre ONU com serviço de melhor esforço e ONU com serviço garantido. As ONU com serviço garantido especificam uma banda mínima por meio de um contrato de prestação de serviços [86]. Assim como a política IPACT, o BGP considera que os escalonamentos inter-ONU e intra-ONU são realizados separadamente, sendo que o último não é especificado na proposta original.

Para o escalonamento inter-ONU, a OLT considera que a capacidade do canal é dividida em “unidades de banda”, de número superior ao número de ONU na rede. A OLT mantém duas tabelas de *polling*, uma para as ONU com serviço garantido e outra para as ONU com serviço de melhor esforço. Em ambas as tabelas, cada registro contém um campo de identificação da ONU e o tempo

² *Deterministic Effective Bandwidth GPS.*

³ *bandwidth guaranteed polling.*

de propagação do sinal entre ela e a OLT. A tabela de ONU com serviço de melhor esforço, por sua vez, não é limitada em relação ao número de registros. Por outro lado, a tabela de ONU com serviço garantido tem número de registros igual ao número de unidades de banda. Cada ONU ocupa, nesta tabela, tantos registros quanto forem necessários para satisfazer a banda mínima requerida. No caso de uma ONU utilizar mais de um registro, cada um deles é disposto de modo que todos sejam distribuídos uniformemente ao longo da tabela.

Durante o ciclo de *polling*, a OLT aloca as janelas primeiro para as ONU com serviço garantido, na ordem em que aparecem na tabela correspondente. Para isso a OLT envia uma mensagem gate contendo permissões em número equivalente a uma unidade de banda. Este número é representado por G . Se um dado registro da tabela está vago, a OLT destina a janela às ONU de melhor esforço, na ordem em que aparecem na tabela correspondente. Ao receber a mensagem gate, a ONU envia imediatamente uma mensagem report com o número de permissões efetivamente utilizado, representado por B , seguida dos dados. Se B é inferior a um limiar pré-estabelecido, a OLT aloca uma nova janela de tamanho $G - B$ à próxima ONU de melhor esforço na ordem da tabela, enviando-lhe uma mensagem gate no instante apropriado. Caso contrário, a OLT espera o fim da transmissão da ONU e, então, continua a alocar as janelas para as ONU de serviço garantido.

A principal vantagem da política BGP é garantir que as ONU recebam a largura de banda mínima contratada e que a capacidade remanescente da rede seja distribuída uniformemente entre as demais. Além disso, garante-se que o intervalo entre duas oportunidades de serviço é limitado para as ONU com serviço garantido. Isto permite obter limitantes para o atraso introduzido no seu tráfego agregado. Em relação a esta métrica, resultados de simulação obtidos em [86] indicam que, para ONU com serviço garantido que ocupem quatro ou mais registros da tabela, o uso da política BGP leva a atrasos de uma ordem de magnitude inferiores ao obtido sob a política IPACT. Para os demais casos, o atraso pode ser uma ordem de magnitude maior no caso do BGP quando a carga da rede é alta, e muitas ordens maior quando a carga é baixa. Resultados similares foram obtidos para a latência nas ONU. A vazão da rede também tende a ser menor no caso da política BGP, sobretudo quando a carga é alta. Essas desvantagens estão relacionadas à discretização da banda em unidades fixas, o que leva à fragmentação das transmissões e, conseqüentemente, a um maior atraso e a um maior *overhead* por mensagens de controle.

Cabe ainda lembrar que a política BGP garante apenas a banda agregada contratada pela ONU, não determinando, contudo, como ela é distribuída entre os fluxos. Em [87], é proposta uma extensão da política BGP que busca incorporar o tratamento de fluxos individuais baseado em classes de serviço. Um algoritmo de controle de admissão é empregado para determinar se uma dada ONU terá serviço garantido ou de melhor esforço. Para isso, cada ONU especifica, com base nas exigências dos seus fluxos, o requisito mínimo de banda e o atraso máximo desejado. A OLT, por sua vez, avalia se este desempenho pode ser garantido, caso em que a ONU é classificada como de serviço garantido. Caso contrário, a OLT tenta negociar novos parâmetros de tráfego ou, se nenhuma solução é obtida, classifica a ONU como sendo de melhor esforço. Com base nesta classificação, a OLT utiliza um algoritmo denominado algoritmo de distribuição uniforme (EDA⁴ para preencher as tabelas de *polling*.

Para o escalonamento intra-ONU, a proposta apresentada em [87] considera que os fluxos são classificados em três classes de prioridade. Emprega-se um sistema de enfileiramento em dois estágios, semelhante ao proposto em [22], e que já foi citado no início desta seção quando abordamos as técnicas para mitigar a penalização de baixa carga. Com isso, garante-se serviço aos fluxos de menor prioridade de cada ONU, eliminando o problema da estagnação.

A política proposta em [87] tem como vantagens a garantia de banda mínima para cada ONU e a possibilidade de oferecer garantias de desempenho relativas aos fluxos. Apesar disso, ela não permite garantir que os recursos da rede são compartilhados de maneira justa entre os fluxos de ONU diferentes, já que o escalonamento inter-ONU é feito com base no tráfego agregado. Finalmente, a política mantém a principal desvantagem do BGP, que é a fragmentação das transmissões.

⁴ Even Distribution Algorithm.

5.5 COMPARTILHAMENTO JUSTO BASEADO EM ENVELOPE DE SERVIÇO

O compartilhamento justo baseado em envelope de serviço (FQSE⁵) proposto por Kramer et al. [21] é uma política de escalonamento hierárquico implementada por meio de um algoritmo cuja execução é distribuída entre os nós de comutação. Essa política, cuja aplicação não se restringe às redes EPON, busca oferecer suporte a garantias absolutas de desempenho a fluxos individuais ou a classes de serviço, e promover o compartilhamento justo dos recursos ociosos da rede.

Na política FQSE, a banda total da rede é distribuída entre os fluxos (ou classes) de modo a garantir-lhes uma banda mínima contratada. A banda ociosa é redistribuída, junto com a parcela da banda mínima que não é utilizada pelos fluxos, entre os fluxos com demanda reprimida. Assim, a cada ciclo n , o FQSE busca destinar ao fluxo i uma janela de transmissão de tamanho dado por

$$G_i[n] = \begin{cases} Q_i[n], & i \notin \mathcal{B}[n] \\ G_i^{(MIN)} + \frac{w_i}{\sum_{m \in \mathcal{B}[n]} w_m} \left(T - \sum_{j \in \mathcal{B}[n]} G_j^{(MIN)} - \sum_{l \notin \mathcal{B}[n]} Q_l[n] \right), & i \in \mathcal{B}[n], \end{cases} \quad (5.5)$$

em que $\mathcal{B}[n]$ é o conjunto de fluxos latentes durante o ciclo i , T é o tamanho do ciclo de *polling* (em bits), $G_i^{(MIN)}$ é a janela de transmissão a que corresponde a banda mínima requerida pelo fluxo i , w_i é o peso (ou prioridade) deste fluxo, e $Q_i[n]$ é a latência do fluxo no início do ciclo n .

Para atingir o compartilhamento dado por (5.5), a política FQSE utiliza um algoritmo baseado no conceito de envelope de serviço. Um envelope de serviço representa o tamanho da janela a ser alocada a um dado fluxo ou grupo de fluxos (e.g., fluxos de uma mesma ONU) como função de um indicador denominado parâmetro de satisfação. Este parâmetro mede o quanto a rede pode satisfazer a demanda daquele fluxo.

No caso de redes EPON, um envelope de serviço é associado a cada fluxo. Este envelope pode ser calculado diretamente pela ONU ou determinado pelo usuário, caso em que é informado a ONU por meio de uma estratégia de *polling* intercalado definida em [21]. Para um dado fluxo i , o envelope de serviço do ciclo n é dado por

$$E_i(n, s) = \begin{cases} w_i s + W_i^{(MIN)}, & s < \frac{1}{w_i} (Q_i[n] - W_i^{(MIN)}) \\ Q_i[n], & \text{c.c.}, \end{cases} \quad (5.6)$$

em que $s \geq 0$ é o parâmetro de satisfação.

Na política FQSE, alternam-se uma fase de requisição e uma de permissão. Para uma dada ONU, a fase de requisição se inicia no final da janela de transmissão de um dado ciclo de *polling*. Nesta fase, a ONU define o seu envelope de serviço, que é dado pela soma dos envelopes de cada um dos seus fluxos, e o envia à OLT por meio de uma mensagem de requisição.

A partir de (5.6), pode-se constatar que o envelope de serviço de uma dada ONU é uma função linear por partes, cujo número de pontos de inflexão pode chegar ao número de fluxos. Tipicamente, este número é grande e não pode ser transportado por uma mensagem de requisição de tamanho fixo. Em [21], é apresentado um algoritmo que é capaz de aproximar o envelope de serviço por um número pequeno de pontos, que podem ser transmitidos por meio de uma mensagem report do protocolo MPCP.

Ao final do ciclo de *polling*, a OLT calcula o envelope de serviço da rede com base nos envelopes enviados pelas ONU. Neste momento, inicia-se a fase de permissão. A partir do envelope de serviço da rede, dado por $E^*(n, s)$, a OLT calcula o parâmetro de satisfação por meio da relação

$$E^*(n, s) = T. \quad (5.7)$$

⁵ Fair Queuing with Service Envelopes.

A OLT inicia, então, um novo ciclo de *polling*, em que envia a cada ONU o parâmetro de satisfação da rede por meio de uma mensagem de permissão. Dado que a OLT conhece o envelope de serviço de cada ONU, ela pode calcular a duração da janela de transmissão que cada ONU ocupará, e assim enviar cada mensagem de permissão no instante apropriado, de modo a não haver colisão nas transmissões de diferentes ONU. Ao receber a mensagem de permissão, as ONU utilizam o parâmetro de satisfação para determinar, com base no envelope correspondente, a janela de transmissão de cada fluxo. Os quadros de cada fluxo são, então, transmitidos e o processo se reinicia.

Em [21], demonstra-se que a política FQSE garante banda mínima aos fluxos e assegura que cada um recebe uma parcela da banda ociosa que é proporcional ao peso (ou prioridade) atribuído a cada um deles. Esta proporção se mantém mesmo para fluxos de ONU diferentes. Além disso, o FQSE inclui ainda dois mecanismos baseados em contadores que visam a aumentar a utilização das janelas de transmissão e a mitigar o bloqueio que pode acontecer quando se aloca, a um fluxo qualquer, uma janela de tamanho inferior ao do primeiro pacote a ser transmitido (*head-of-line blocking*).

A principal vantagem da política FQSE é o fato de ela permitir o tratamento individualizado dos fluxos com razoável escalabilidade. Contudo, ela tem como desvantagens a falta de resultados analíticos para as métricas de desempenho individuais, o custo computacional de se aproximar o envelope de serviço em cada ONU, o que inclui uma complicada estratégia de *polling*, e o *overhead* introduzido pelo tamanho da mensagem de controle necessária para transportar o envelope de serviço de cada ONU até a OLT.

Em relação à complexidade computacional da política, Kramer et al. [21] demonstram que a fase de requisição tem complexidade $\mathcal{O}[K \cdot D(\log D + \log B)]$ para cada nó da rede (no caso de uma EPON, a OLT ou as ONU). Nesta expressão, B é a duração (em bits) do ciclo, D o número de nós descendentes e K o número de pontos que compõe a curva de serviço de cada um deles. Vale lembrar que os nós descendentes da OLT são as ONU que, por sua vez, têm como descendentes os próprios fluxos ou classes. Para a fase de permissão, Kramer et al. demonstram que a complexidade é $\mathcal{O}(D)$.

Com base nesse resultado, podemos calcular a complexidade da política FQSE para uma rede EPON com J ONU e N_{max} fluxos em cada. Como é explicado em [21], a curva de serviço dos fluxos (ou classes) individuais tem apenas dois pontos. Desta maneira, a complexidade da fase de requisição de uma ONU é $\mathcal{O}[2N_{max}(\log N_{max} + \log B)]$ e a da sua fase de permissão é $\mathcal{O}(N_{max})$. A curva de serviço de cada ONU têm $N_{max} + 1$ pontos. Para a OLT, então, as mesmas fases têm complexidade $\mathcal{O}[(N_{max} + 1)J(\log J + \log B)]$ e $\mathcal{O}(J)$, respectivamente.

A complexidade total do algoritmo FQSE é dada então por

$$\begin{aligned}
C_{fqse} &\equiv \mathcal{O}\{J \cdot (C_{req.onu} + C_{perm.onu}) + (C_{req.olt} + C_{perm.olt})\} \\
&\equiv \mathcal{O}\left\{2N_{max}J \left(\log N_{max} + \log B + \frac{1}{2}\right) + (N_{max} + 1)J \left[\log J + \log B + \frac{1}{N_{max} + 1}\right]\right\} \\
&\equiv \mathcal{O}\{2N_{max}J(\log N_{max} + \log B) + N_{max}J[\log J + \log B]\} \\
&\equiv \mathcal{O}[N_{max}J \log(B^3 N_{max}^2 J)].
\end{aligned} \tag{5.8}$$

Assim, a política tem complexidade do tipo $\beta \log \beta$ tanto em relação ao número de fluxos como ao de ONU. Vale lembrar que a complexidade da política PSLR que ora propomos é linear em ambos os parâmetros.

No que tange a justiça no compartilhamento da banda entre os fluxos, Kramer et al. demonstram que a diferença entre os serviços oferecidos a dois fluxos com igual reserva de banda e igual prioridade em um intervalo de n ciclos de *polling* satisfaz as relações [21]

$$|S_i[n] - S_j[n]| \leq \max\{|K_i^-(n) - K_j^+(n)|; |K_i^+(n) - K_j^-(n)|\}, \tag{5.9}$$

em que $S_i[n]$ é o serviço oferecido ao fluxo i até o final do ciclo n e $K_i^-(n)$ e $K_i^+(n)$ são dados por

$$\begin{aligned}
K_i^-(n) &= \max\{1 - L; -nP_i\} \\
K_i^+(n) &= \max\{L - 1; n(L - P_i)\},
\end{aligned}$$

em que L é o tamanho máximo dos pacotes e $P_i = \rho_i B/r$, sendo ρ_i a banda reservada pelo fluxo i e r a capacidade da rede. Em [21], Kramer et al. demonstram que a diferença $|S_i[n] - S_j[n]|/n$ é crescente em relação a n e que ela satura para $n \rightarrow \infty$. Portanto, a diferença entre as taxas de serviço oferecidas aos fluxos i e j cresce a curto prazo, mas é limitada quando se consideram escalas de tempo maiores. Note que a diferença $|S_i[n] - S_j[n]|$ deve ser sempre crescente em relação a n para que o resultado apontado por Kramer et al. seja válido. Isto significa que, ao longo dos ciclos, a política FQSE pretere certos fluxos em favor de outros, oferecendo menos serviço àqueles do que seria o justo. Sob este aspecto, e à luz da definição de justiça relativa que apresentamos no Capítulo 4, a política FQSE não pode ser considerada justa.

5.6 COMPARAÇÃO COM A POLÍTICA PSLR

Dentre as políticas apresentadas ao longo deste capítulo, apenas o DEB-GPS e o FQSE têm o mesmo objetivo da política PSLR, que é atender os fluxos individualmente. No entanto, o DEB-GPS não é capaz de garantir que a capacidade ociosa da rede seja dividida de maneira justa entre fluxos de ONU diferentes, dado que a banda que um fluxo deixa de usar é obrigatoriamente destinada a outro da mesma ONU.

A política FQSE, por outro lado, busca centralizar a distribuição da banda na OLT, que sintetiza as informações de demanda individuais em um único indicador. Este indicador, que representa a demanda global da rede, é então repassado às ONU para que as suas janelas de transmissão sejam divididas entre os fluxos de maneira justa. Este princípio é o mesmo da política PSLR. Contudo, vimos que a política FQSE não é capaz de atingir a justiça relativa definida no Capítulo 4, o que se deve principalmente ao mecanismo de prevenção de fracionamento de pacotes que ela adota. Apesar disso, a política FQSE pode ser considerada justa se observarmos apenas a diferença de taxas de serviço (e não do serviço em si) que, nesta política, é limitada.

Cabe ressaltar que a política FQSE tem complexidade computacional superior à da política PSLR, o que representa uma desvantagem sobretudo em termos de escalabilidade, já que o acréscimo de novos fluxos ou de novas ONU pode tornar a sua implementação proibitiva devido ao tempo computacional restrito de que o escalonamento dispõe. No nosso ponto de vista, a maior complexidade computacional e a falta de resultados de análise de desempenho são as principais desvantagens desta política frente ao PSLR.

CAPÍTULO 6

CONCLUSÃO

Neste trabalho, apresentamos uma proposta para a modelagem, policiamento e escalonamento de tráfego em redes EPON visando ao suporte a garantias de desempenho individuais para as fluxos. Em um contexto mais geral, abordamos a natureza fractal do tráfego em redes e a sua relevância no que tange o contexto de redes de serviço garantido. Propusemos o uso de um processo envelope denominado Processo de Chegada com Limitante Fracionário (FBAP) para a representação de tráfego auto-similar. Desenvolvemos a análise de sistemas de filas para o tráfego FBAP e obtivemos a banda efetiva e limitantes para a latência e para o atraso introduzido nos fluxos. Os resultados foram estendidos para o caso do tráfego multifractal, para o qual definimos um processo envelope denominado Processo de Chegada com Limitante Multifractal (MFBAP). Tanto o FBAP como o MFBAP foram validados utilizando amostras de tráfego real, e comprovamos a sua acurácia.

Em relação ao problema do policiamento de tráfego, estudamos o policiamento do tráfego fractal por meio do algoritmo do balde furado fractal proposto em [37] e do balde furado tradicional. Demonstramos que o balde furado fractal não é capaz de suportar garantias de desempenho aos fluxos por não limitar a explosividade do tráfego. Propusemos, então, o uso do balde furado para o policiamento do tráfego com base no modelo FBAP, o que ilustramos com exemplos baseados em simulação.

No caso específico de redes EPON, apresentamos uma revisão dos principais aspectos relacionados ao problema do escalonamento dos fluxos dos usuários. Propusemos uma política de escalonamento denominada Compartilhamento Proporcional com Reserva de Carga (PSLR) em que a rede reserva uma parcela mínima de banda para cada fluxo e a capacidade ociosa da rede é redistribuída entre os fluxos com demanda reprimida. Propusemos também um algoritmo distribuído e de baixa complexidade para a implementação prática do escalonamento PSLR, cuja operação foi ilustrada por meio de simulação.

Por meios analíticos, demonstramos que a política PSLR tem baixa complexidade computacional, garante banda mínima aos fluxos e assegura o compartilhamento justo dos recursos da rede entre eles. Obtivemos ainda as condições de estabilidade dos fluxos e expressões limitantes para a latência e para o atraso dos fluxos quando estes são policiados pelo algoritmo do balde furado.

Comparando com as principais políticas disponíveis na literatura, a política PSLR tem como principais vantagens a baixa complexidade computacional, o suporte a garantias de atraso e banda mínima aos fluxos e a capacidade de tratá-los individualmente e de maneira justa. Nenhuma daquelas políticas atinge simultaneamente todos estes objetivos.

Vale observar ainda que, embora a política pslr tenha sido definida no contexto de EPON, ela poderia ser utilizada, com algumas modificações, em outras redes em que o acesso seja controlado por *polling*. Consideramos os estudos envolvidos como uma possível extensão deste trabalho.

Como possíveis extensões deste trabalho, podemos apontar a análise de desempenho da política PSLR por meio de técnicas baseadas em curvas de serviço [71–73, 75, 88, 89] visando, com isso, obter limitantes de latência e de atraso mais acurados e a obtenção de limitantes para a latência e o atraso do tráfego MFBAP.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Industry Canada, “The new national dream: networking the nation for broadband access”, Report of the National Broadband Task Force, julho de 2003. [Internet]. Disponível em <http://broadband.gc.ca>
- [2] “Wikipedia, the free encyclopedia”, 2006, acessado em 10/09/2006. [Internet]. Disponível em <http://en.wikipedia.org/wiki/>
- [3] P. Takeuti, “Projeto e dimensionamento de redes ópticas passivas (PONs)”, Dissertação de mestrado, Universidade de São Paulo, 2005.
- [4] P. Garvey, “The economics of FTTH”, *Corning Guidelines Magazine*, vol. 8E, 2004. [Internet]. Disponível em http://www.corning.com/opticalfibre/guidelines_magazine/eguidelines/vol8/cover.aspx
- [5] International Engineering Consortium, “FTTH explained: delivering efficient customer bandwidth and enhanced services”, Tutorial eletrônico, 2005. [Internet]. Disponível em http://www.iec.org/online/tutorials/fiber_home/
- [6] G. Kramer, B. Mukherjee e G. Pesavento, “Ethernet PON (ePON): Design and analysis of an optical access network”, *Photonic Network Communications*, vol. 3, n°3, págs. 307–319, julho de 2001.
- [7] G. Kramer e G. Pesavento, “Ethernet Passive Optical Network (EPON): Building a next-generation optical access network”, *IEEE Commun. Mag.*, vol. 40, n°2, págs. 66–73, fevereiro de 2002.
- [8] IEEE Groups, “IEEE 802.310Gb/s PHY for EPON Study Group”, 2006. [Internet]. Disponível em http://grouper.ieee.org/groups/802/3/10GEPON_study/index.html
- [9] D. Gutierrez, K. S. Kim, S. Rotolo, F.-T. An e L. Kazovsky, “FTTH standards, deployments and research issues”, in *Proc. JCIS 2005*, Salt Lake City, USA, 2005. [Internet]. Disponível em <http://www.stanford.edu/degm/publications.html>
- [10] M. Howard, “PON & FTTx update”, Light Reading (<http://www.lightreading.com>), agosto de 2005. [Internet]. Disponível em http://www.lightreading.com/document.asp?doc_id=73507&print=true
- [11] Nortel Networks, “Ultra broadband access: unleashing the power of PON”, White paper, setembro de 2004. [Internet]. Disponível em <http://www.nortel.com/solutions/brdbndacss/collateral/nn109280-092304.pdf>
- [12] P. Chancelou, S. Gosselin, J. Palacios, V. Alvarez e E. Zouganeli, “Overview of the optical broadband access evolution: a joint article by operators in the IST network of excellence e-Photon/One”, *IEEE Commun. Mag.*, vol. 44, n°8, págs. 29–35, agosto de 2006.

- [13] International Engineering Consortium, “Asynchronous Transfer Mode (ATM) Passive Optical Networks PONs”, tutorial eletrônico, 2005. [Internet]. Disponível em http://www.iec.org/online/tutorials/atm_pon/
- [14] ANSI/IEEE, “IEEE 802.1d — IEEE standards for local and metropolitan area networks, part III: Media Access Control (MAC) bridges”, junho de 1998.
- [15] —, “IEEE 802.1q — IEEE standards for Local and Metropolitan Area Networks — Virtual Bridged Local Area Networks”, maio de 2003.
- [16] —, “IEEE 802.3ah — IEEE 802.3 standard amendment: Media access control parameters, physical layers and management parameters for subscriber access networks”, junho de 2004.
- [17] G. Kramer, *Ethernet Passive Optical Networks*. McGraw-Hill, 2005.
- [18] G. Kramer, B. Mukherjee e G. Pesavento, “IPACT: A dynamic protocol for an Ethernet PON (EPON)”, *IEEE Commun. Mag.*, vol. 40, n°2, págs. 74–80, fevereiro de 2002.
- [19] —, “Interleaved Polling with Adaptive Cycle Time (IPACT): A dynamic bandwidth distribution scheme in an optical access network”, *Photonic Network Communications*, vol. 4, n°1, págs. 89–107, janeiro de 2002.
- [20] J. Zheng e H. T. Mouftah, “Media Access Control for Ethernet Passive Optical Networks: an overview”, *IEEE Commun. Mag.*, vol. 43, n°2, págs. 145–150, fevereiro de 2005.
- [21] G. Kramer, A. Banerjee, N. Singhal, B. Mukherjee, S. Dixit e Y. Ye, “Fair Queuing with Service Envelopes (FQSE): A cousin-fair hierarchical scheduler for subscriber access networks”, *IEEE J. Select. Areas Commun.*, vol. 22, n°8, págs. 1497–1513, outubro de 2004.
- [22] G. Kramer, B. Mukherjee, S. Dixit, Y. Ye e R. Hirth, “Supporting differentiated classes of service in Ethernet Passive Optical Networks”, *OSA Journal of Optical Networking*, vol. 1, n°8/9, págs. 280–298, agosto de 2002.
- [23] S.-I. Choi e J.-D. Huh, “Dynamic bandwidth allocation algorithm for multimedia services over Ethernet PONs”, *ETRI Journal*, vol. 24, n°6, págs. 465–468, dezembro de 2002.
- [24] H. Zhang, “Service disciplines for guaranteed performance service in packet switching networks”, *Proc. IEEE*, vol. 83, n°10, págs. 1374–1396, 1995.
- [25] D. Stiliadis e A. Varma, “Latency-Rate servers: a general model for analysis of traffic scheduling algorithms”, *IEEE/ACM Transactions on Networking*, vol. 6, n°5, págs. 611–624, 1998.
- [26] M. W. Garret e W. Willinger, “Analysis, modeling and generation of self-similar VBR video traffic”, in *Proc. ACM/SIGCOMM*, agosto de 1994, págs. 269–280.
- [27] V. Paxson e S. Floyd, “Wide Area Traffic: the failure of Poisson modeling”, *IEEE Trans. Networking*, vol. 3, n°3, págs. 226–244, junho de 1995.
- [28] O. Rose, “Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems”, Tech. Report 101, Institute of Computer Sciences, University of Würzburg, fevereiro de 1995.
- [29] M. S. Taqqu, V. Teverovsky e W. Willinger, “Is network traffic self-similar or multifractal?” *Fractals*, vol. 5, n°1, págs. 63–73, março de 1997.
- [30] K. Park, G. Kim e M. Crovella, “On the relation between file sizes, transport protocols and self-similar network traffic”, in *Proc. IEEE Int’l. Conf. Network Protocols*, outubro de 1996, págs. 171–180.

- [31] O. Rose, "Simple and efficient models for variable bit rate MPEG video traffic", *Perf. Eval.*, vol. 30, págs. 69–85, julho de 1997.
- [32] W. Willinger, M. S. Taqqu, R. Sherman e D. V. Wilson, "Self-similarity through high-variability: statistical analysis of ethernet LAN traffic at the source level", *IEEE/ACM Trans. Networking*, vol. 5, n°1, págs. 71–86, fevereiro de 1997.
- [33] M. E. Crovella, "Self-similarity in WWW traffic: evidence and possible causes", *IEEE/ACM Trans. Networking*, vol. 5, n°6, págs. 835–845, fevereiro de 1997.
- [34] R. H. Riedi e J. L. Véhel, "TCP traffic is multifractal: a numerical study", Research Report 3129, INRIA, março de 1997.
- [35] A. Feldmann, A. C. Gilbert e W. Willinger, "Data networks as cascades: investigating the multifractal nature of internet WAN traffic", in *Proc. SIGCOMM*, setembro de 1998, págs. 42–55.
- [36] K. Park e W. Willinger, *Self-similar network traffic and performance evaluation*. Wiley, 2000.
- [37] C. A. Viana Neto, "Multiplexação e policiamento de tráfego auto-semelhante", Dissertação de Mestrado, Universidade Estadual de Campinas, novembro de 1999.
- [38] W. E. Leland e D. V. Wilson, "High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection", in *Proc. IEEE/INFOCOM*, abril de 1991, págs. 1360–1366.
- [39] W. E. Leland, W. Willinger, M. S. Taqqu e D. Wilson, "On the self-similar nature of ethernet traffic (extended version)", *IEEE/ACM Trans. Networking*, vol. 2, n°1, págs. 1–15, fevereiro de 1994.
- [40] I. Norros, "A storage model with self-similar input", *Queueing Systems*, vol. 16, págs. 387–396, setembro de 1994.
- [41] N. G. Duffield e N. O'Connell, "Large deviations and overflow probabilities for the general single-server queue, with applications", *Math. Proc. Cambridge Phil. Society*, vol. 118, n°2, págs. 363–374, 1995.
- [42] P. Morin, "The impact of self-similarity on network performance analysis", Tech. Report Computer Science 95.495, Carleton University, dezembro de 1995.
- [43] M. Montgomery e G. de Veciana, "On the relevance of time scales in performance oriented traffic characterizations", in *Proc. IEEE/INFOCOM*, março de 1996, págs. 513–520.
- [44] D. Heymand e T. Lakshman, "What are the implications of long range dependence for VBR video traffic engineering?" *IEEE/ACM Trans. Networking*, vol. 4, n°3, págs. 301–317, junho de 1996.
- [45] G. S. Mayor e J. A. Silvester, "Time scale analysis of an ATM queueing system with long-range dependent traffic", in *Proc. IEEE/INFOCOM*, abril de 1997, págs. 205–212.
- [46] K. Park, G. Kim e M. Crovella, "On the effect of traffic self-similarity on network performance", in *Proc. SPIE Int'l. Conf. Perf. and Control of Network Systems*, novembro de 1997, págs. 296–310.
- [47] S. Molnár, T. D. Dang e I. Maricza, "On the queue tail asymptotics for general multifractal traffic", in *Proc. 2nd Intl. IFIP-TC6 Networking Conference*, maio de 2002, págs. 105–116.
- [48] A. Karasaridis e D. Hatzinakos, "A non-gaussian self-similar process for broadband heavy-traffic modeling", in *Proc. IEEE/GLOBECOM*, novembro de 1998, págs. 2995–3000.

- [49] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi e D. Veitch, “Multiscale nature of network traffic”, *IEEE Signal Proc. Mag.*, vol. 19, págs. 28–46, maio de 2002.
- [50] B. Mandelbrot, A. Fisher e L. Calvet, “A multifractal model of asset returns”, Cowles Foundation Discussion Paper 1164, Yale University, setembro de 1997.
- [51] P. Chainais, P. Abry e D. Veitch, “Multifractal analysis and α -stable processes: a methodological contribution”, in *Proc. ICASSP*, junho de 2000.
- [52] T. Mikosch, S. Resnick, H. Rootzén e A. Stegeman, “Is network traffic approximated by stable Lévy motion or fractional brownian motion?” Tech. Report 1247, Cornell University, 1999.
- [53] M. S. Taqqu, R. Sherman e W. Willinger, “Proof of a fundamental result in self-similar traffic modeling”, *ACM/SIGCOMM Computer Communication Review*, vol. 27, n^o2, págs. 5–23, abril de 1997.
- [54] J. P. Nolan, “Stable distributions : models for heavy-tailed data”, 2000, versão preliminar do Capítulo 1. [Internet]. Disponível em <http://academic2.american.edu/~jpnolan/stable/chap1.pdf>
- [55] T. Neame, M. Zukerman e R. G. Addie, “A practical approach for multimedia traffic modeling”, in *Proc. Broadband Commun.*, novembro de 1999, págs. 73–82.
- [56] I. Norros, “On the use of fractional brownian motion in the theory of connectionless networks”, *IEEE J. Select. Areas Commun.*, vol. 13, n^o6, págs. 953–962, agosto de 1995.
- [57] A. Karasaridis e D. Hatzinakos, “Broadband heavy-traffic modeling using stable self-similar processes”, in *2nd Canadian Conference on Broadband Research*, junho de 1998, págs. 157–168.
- [58] F. Harmantzis, D. Hatzinakos e I. Katzela, “Tail probabilities for the multiplexing of fractional α -stable broadband traffic”, in *Proc. IEEE/ICC*, junho de 2001, págs. 2665–2669.
- [59] J. B. Levy e M. S. Taqqu, “The asymptotic codifference and covariation of log-fractional stable noise”, in *Bundesbank 2005 Conference*, novembro de 2005. [Internet]. Disponível em http://www.bundesbank.de/download/vfz/konferenzen/20051110_12_eltville/paper_taqqu.pdf
- [60] P. Abry, M. S. Taqqu e B. Pesquet-Popescu, “Wavelet based estimators for self-similar α -stable processes”, in *Proc. Int. Conf. On Signal Proc., 16th World Computer Congress*, agosto de 2000.
- [61] M. S. Taqqu, V. Teverovsky e W. Willinger, “Estimators for long-range dependence: an empirical study”, *Fractals*, vol. 3, n^o4, págs. 785–798, 1995.
- [62] F. M. Pereira, N. L. S. Fonseca e D. S. Arantes, “Fractional Bounded Arrival Process for self-similar traffic policing”, in *Proc. IEEE/ICC*, Anchorage, USA, 2003, págs. 1537–1542.
- [63] —, “Modelagem de tráfego utilizando o processo de chegadas com limitante multifractal”, in *Anais do XX Simpósio Brasileiro de Telecomunicações*, SBrT, Ed., Rio de Janeiro, Brasil, 2003.
- [64] —, “Fractal traffic modeling and policing”, Relatório Técnico 03-06, Instituto de Computação, Universidade Estadual de Campinas, 2003.
- [65] —, “Modelagem de tráfego fractal utilizando processos envelope”, *Revista da Sociedade Brasileira de Telecomunicações*, vol. 19, n^o2, págs. 94–106, agosto de 2004.
- [66] M. Chi, E. Neal e G. Young, “Practical applications of fractional brownian motion and noise to synthetic hydrology”, *Water Resources Research*, vol. 9, págs. 1523–1533, 1973.
- [67] P. Abry e D. Veitch, “Wavelet analysis of long-range dependent traffic”, *IEEE Trans. Inform. Theory*, vol. 44, n^o1, págs. 2–15, janeiro de 1998.

- [68] N. L. S. Fonseca, G. S. Mayor e C. A. V. Neto, “On the equivalent bandwidth of self-similar sources”, *ACM Trans. On Modeling and Computer Simulation*, vol. 10, n°2, págs. 104–124, 2000.
- [69] Lawrence Berkeley Laboratory, “The Internet Traffic Archive”. [Internet]. Disponível em <http://ita.ee.lbl.gov>
- [70] V. Paxson, “Fast approximation of self-similar network traffic”, Tech. Report LBL-36750, Lawrence Berkeley Laboratories, 1995.
- [71] A. Parekh e R. Gallager, “A Generalized Processor Sharing approach to flow control in integrated services networks: the multiple-node case”, *IEEE/ACM Trans. Networking*, vol. 2, n°2, págs. 137–150, abril de 1994.
- [72] —, “A Generalized Processor Sharing approach to flow control in integrated services networks: the single-node case”, *IEEE/ACM Trans. Networking*, vol. 1, n°3, págs. 344–357, junho de 1993.
- [73] F. M. Pereira, N. L. S. Fonseca e D. S. Arantes, “On the performance of Generalized Processor Sharing servers under long-range dependent traffic”, *Computer Networks*, vol. 40, n°3, págs. 413–431, 2002.
- [74] G. Procissi, A. Garg, M. Gerla e M. Y. Sanadidi, “Token bucket characterization of long-range dependent traffic”, *Computer Communications*, vol. 25, págs. 1009–1017, 2002.
- [75] F. M. Pereira, “Análise de desempenho da disciplina de serviço Generalized Processor Sharing sob tráfego auto-similar”, Dissertação de Mestrado, Universidade Estadual de Campinas, Campinas, Brasil, 2002.
- [76] Y. Zhou, “Resource allocation in computer networks: fundamental principles and practical strategies”, Tese de Doutorado, Drexel University, maio de 2003.
- [77] J. Le Boudec, “Rate adaptation, congestion control and fairness: A tutorial”, 2001. [Internet]. Disponível em <http://citeseer.ist.psu.edu/504527.html>
- [78] Y. Zhou e H. Sethu, “On the relationship between absolute and relative fairness bounds”, *IEEE Commun. Letters*, vol. 6, n°1, págs. 37–39, janeiro de 2002.
- [79] S. Low e D. Lapsley, “Optimization flow control — I: basic algorithm and convergence”, *IEEE/ACM Transactions on Networking*, vol. 7, n°6, págs. 861–874, 1999.
- [80] D. Bertsekas e J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, 1989.
- [81] D. Sala e A. Gummalla, “PON functional requirements: Services and performance”, julho de 2001. [Internet]. Disponível em http://grouper.ieee.org/groups/802/3/efm/public/jul01/presentations/sala_1_0701.pdf
- [82] A. Varga, “The Omnet++ simulator (version 3.0)”. [Internet]. Disponível em <http://www.omnetpp.org>
- [83] G. Kramer, “Generator of self-similar traffic (version 3)”. [Internet]. Disponível em http://wwwcsif.cs.ucdavis.edu/~kramer/code/trf_gen3.html
- [84] Y. Luo e N. Ansari, “Bandwidth allocation for multiservice access on EPONs”, *IEEE Commun. Mag.*, vol. 43, n°2, págs. S16–S21, fevereiro de 2005.
- [85] J. Xie, S. Jiang e Y. Jiang, “A dynamic bandwidth allocation scheme for differentiated services in EPONs”, *IEEE Commun. Mag.*, vol. 42, n°8, págs. 32–39, agosto de 2004.

- [86] M. Ma, Y. Zhu e T. H. Cheng, “A bandwidth guaranteed polling MAC protocol for Ethernet Passive Optical Networks”, in *Proc. IEEE/INFOCOM*, 2003.
- [87] Y. Zhu, M. Ma e T. H. Cheng, “Hierarchical scheduling to support differentiated services in Ethernet Passive Optical Networks”, *Computer Networks*, vol. 50, págs. 350–366, 2006.
- [88] R. Cruz, “A calculus for network delay. Part I: network elements in isolation”, *IEEE Trans. Inform. Theory*, vol. 37, n°1, págs. 114–131, 1991.
- [89] —, “A calculus for network delay. Part II: network analysis”, *IEEE Trans. Inform. Theory*, vol. 37, n°1, págs. 132–141, 1991.
- [90] S. Kanhere e H. Sethu, “On the latency bound of Deficit Round Robin”, in *Proc. Intl. Conf. on Computer Commun. and Networks*, Miami, FL, 2002.

APÊNDICE A

PROVAS DE TEOREMAS

A.1 PROVA DO TEOREMA 4.4

Teorema A.1. *A solução (4.17)–(4.21) pode ser obtida, de forma adaptativa, por meio dos Algoritmos 4.1 e 4.2. Ademais, a convergência para a solução é linear e de taxa $1 - \eta$.*

Demonstração. No algoritmo proposto, $-d_{ij}[n]$ é igual ao excesso de serviço recebido pelo fluxo ij para evitar diferimento de pacotes até o ciclo $n - 1$. Podemos então concluir que

$$-L < d_{ij}[n] \leq 0, \quad (\text{A.1})$$

em que L é o tamanho máximo dos pacotes. Dado que o algoritmo busca sempre descontar $-d_{ij}[n]$ do serviço a ser oferecido nos ciclos subsequentes, é fácil concluir que, para $n \rightarrow \infty$, o mecanismo que previne o fracionamento de pacotes tem efeito somente sobre a oscilação da janela alocada $c_{ij}[n]$ em torno do ponto de convergência, dado por $c_{ij}^*[n]$. Desta maneira, podemos analisar a convergência do algoritmo diretamente de (4.34).

Suponhamos que o conjunto $\mathcal{A}[n]$ varie lentamente em relação a n , ao menos bem mais lentamente do que a dinâmica da equação acima. Neste caso, podemos considerá-lo constante e dado por \mathcal{A}_∞ , de modo que

$$\begin{aligned} \tilde{\xi}[n+1] &= \tilde{\xi}[n] + \eta \left(B - h - \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n] - \sum_{ij \notin \mathcal{A}_\infty} P_{ij} - \tilde{\xi}[n] \right) \\ &= (1 - \eta) \tilde{\xi}[n] + \eta \left(B - h - \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n] - \sum_{ij \notin \mathcal{A}_\infty} P_{ij} \right) \end{aligned}$$

Para que os pólos desta equação sejam positivos e de módulo menor do que 1, temos que $0 < \eta < 1$. Assim, para $n \rightarrow \infty$, temos

$$\begin{aligned} \lim_{n \rightarrow \infty} \tilde{\xi}[n] &= B - h - \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n] - \sum_{ij \notin \mathcal{A}_\infty} P_{ij} \\ \Rightarrow \xi^*[n] &= \frac{1}{\sum_{ij \notin \mathcal{A}_\infty} w_{ij}} \left(B - h - \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n] - \sum_{ij \notin \mathcal{A}_\infty} P_{ij} \right), \end{aligned}$$

em que $\xi^*[n]$ é dado por (4.26). Assim, o algoritmo converge para o valor ótimo. Ademais, podemos ver que a taxa dada por

$$\frac{\tilde{\xi}[n+1] - \tilde{\xi}^*}{\tilde{\xi}[n] - \tilde{\xi}^*} = (1 - \eta),$$

ou seja, a convergência do algoritmo é linear e tem taxa $1 - \eta$. □

A.2 PROVA DO TEOREMA 4.6

Teorema. *O fluxo ij é localmente estável se*

$$\rho_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(r - \frac{r \cdot h}{B} - \sum_{lm} \rho_{lm} \right) \geq \mu_{ij}.$$

em que μ_{ij} é a taxa média do fluxo ij . A rede será estável se

$$\sum_{ij} \mu_{ij} < r - \frac{r \cdot h}{B}.$$

Demonstração. Suponhamos que $ij \notin \mathcal{A}[n]$ para todo fluxo ij . De (4.34),

$$\begin{aligned} \tilde{\xi}[n+1] &= \tilde{\xi}[n] + \eta(B - b[n]) \\ &= (1 - \eta)\tilde{\xi}[n] + \eta \left(B - h - \sum_{ij \notin \mathcal{A}[n]} P_{ij} - \sum_{ij \in \mathcal{A}[n]} \tilde{Q}_{ij} \right) \\ &\geq (1 - \eta)\tilde{\xi}[n] + \eta \left(B - h - \sum_{ij} P_{ij} \right), \end{aligned}$$

pois o serviço oferecido a um fluxo $ij \in \mathcal{A}[n]$ é inferior à parcela justa. Assim, para $n \rightarrow \infty$, temos que $\xi[n] \geq B - h - \sum_{ij} P_{ij}$.

$$c_{ij}[n] \geq P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(B - h - \sum_{lm} P_{lm} \right),$$

de modo que

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^k c_{ij}[n] \geq P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(B - h - \sum_{lm} P_{lm} \right).$$

Portanto, para que o fluxo ij seja localmente estável, é suficiente que

$$P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(B - h - \sum_{lm} P_{lm} \right) \geq \frac{\mu_{ij}}{r} B.$$

A partir da definição de $b[n]$, temos que

$$\lim_{k \rightarrow \infty} \sum_{n=0}^k b[n] = \lim_{k \rightarrow \infty} \sum_{n=0}^k \sum_{ij} c_{ij}[n] + \lim_{k \rightarrow \infty} h \cdot k.$$

Para que o fluxo ij seja estável, é suficiente que a rede seja estável. Ou seja,

$$\lim_{k \rightarrow \infty} \sum_{n=0}^k c_{ij}[n] > \frac{\mu_{ij}}{r} \lim_{k \rightarrow \infty} \sum_{n=0}^k b[n].$$

Reunindo as duas relações anteriores e dado que a convergência do algoritmo leva a $\lim_{n \rightarrow \infty} b[n] = B$, temos que

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^k b[n] &> h + \sum_{ij} \frac{\mu_{ij}}{r} \left(\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^k b[n] \right) \\ B &> h + \sum_{ij} \frac{\mu_{ij}}{r} B, \end{aligned}$$

o que conclui a prova. \square

A.3 PROVA DO TEOREMA 4.7

Teorema. Para a política PSLR, o limitante de justiça relativo satisfaz a desigualdade

$$RFB \leq \frac{1}{\min_{ij} w_{ij}} \max_{ij} [P_{ij} + B - h + 2(L-1)].$$

Demonstração. Sejam dois fluxos $ij, lm \in \mathcal{B}(\tau; t)$ quaisquer. Sejam ainda n_1 e n_2 o número de vezes em que o fluxo ij é atendido até os instantes τ e t , respectivamente. Supondo que o fluxo ij antecede o fluxo lm na ordem, o serviço oferecido a cada um no intervalo $(\tau; t)$ é dado por

$$\begin{aligned} S_{ij}(\tau; t) &= \sum_{k=n_1+1}^{n_2} c_{ij}[k] + d_{ij}[n_1] - d_{ij}[n_2], \\ S_{lm}(\tau; t) &\geq \sum_{k=n_1+1}^{n_2-1} c_{lm}[k] + d_{lm}[n_1] - d_{lm}[n_2]. \end{aligned}$$

em que $c_{ij}[k] = P_{ij} + w_{ij}\xi[k]$. Dado que $-L+1 \leq d_{ij}[n] \leq 0$,

$$\begin{aligned} S_{ij}(\tau; t) &\leq \sum_{k=n_1+1}^{n_2} c_i[k] + (L-1), \\ S_{lm}(\tau; t) &\geq \sum_{k=n_1+1}^{n_2-1} c_j[k] - (L-1). \end{aligned}$$

Seja $\phi = (n_2 - n_1)B/r$. Da definição de $\xi[n]$, temos que

$$\begin{aligned} S_{ij}(\tau; t) &\leq \rho_{ij}\phi + \sum_{k=n_1+1}^{n_2-1} w_{ij}\xi[k] + w_{ij}\xi[n_2] + (L-1), \\ S_{lm}(\tau; t) &\geq \rho_{lm}\phi - P_{lm} + \sum_{k=n_1+1}^{n_2-1} w_{lm}\xi[k] - (L-1). \end{aligned}$$

A partir de (4.34),

$$\begin{aligned} \tilde{\xi}[n+1] &= \tilde{\xi}[n] + \eta(B - b[n]) \\ &= (1-\eta)\tilde{\xi}[n] + \eta \left(B - h - \sum_{ij \notin \mathcal{A}[n]} P_{ij} - \sum_{j \in \mathcal{A}[n]} \tilde{Q}_j[n] \right) \\ &\leq (1-\eta)\tilde{\xi}[n] + \eta(B-h). \end{aligned}$$

Assim, temos que $\tilde{\xi}[n] \leq B - h$ e

$$\begin{aligned} \text{RFB} &\leq \max_{ij,lm} \left[\frac{P_{lm}}{w_{lm}} + \tilde{\xi}[n_2] + \frac{L-1}{w_{ij}} + \frac{L-1}{w_{lm}} \right] \\ &\leq \frac{1}{\min_{ij} w_{ij}} \max_{ij} [P_{ij} + B - h + 2(L-1)], \end{aligned}$$

o que conclui a prova. □

A.4 PROVA DO TEOREMA 4.8

Teorema. *Para a política PSLR, o limitante de justiça relativo para ciclos completos satisfaz a desigualdade*

$$\text{RFB}_{CC} \leq \frac{2}{\min_{ij} w_{ij}} (L-1).$$

Demonstração. Sejam dois fluxos $ij, lm \in \mathcal{B}[n_1; n_2]$ quaisquer. Supondo que o fluxo ij antecede o fluxo lm na ordem, o serviço oferecido a cada um destes fluxos no intervalo entre o início do ciclo n_1 e o final do ciclo n_2 é dado por

$$\begin{aligned} S_{ij}[n_1; n_2] &= \sum_{k=n_1}^{n_2} c_{ij}[k] + d_{ij}[n_1] - d_{ij}[n_2], \\ S_{lm}[n_1; n_2] &= \sum_{k=n_1}^{n_2} c_{lm}[k] + d_{lm}[n_1] - d_{lm}[n_2]. \end{aligned}$$

Dado que $-L + 1 \leq d_{ij}[n] \leq 0$,

$$\begin{aligned} S_{ij}[n_1; n_2] &\leq \sum_{k=n_1}^{n_2} c_{ij}[k] + (L-1), \\ S_{lm}[n_1; n_2] &\geq \sum_{k=n_1}^{n_2} c_{lm}[k] - (L-1). \end{aligned}$$

Seja $\phi = (n_2 - n_1 + 1)B/r$. Da definição de $\xi[n]$, temos que

$$\begin{aligned} S_{ij}[n_1; n_2] &\leq \rho_{ij}\phi + \sum_{k=n_1}^{n_2} w_{ij}\xi[k] + (L-1), \\ S_{lm}[n_1; n_2] &\geq \rho_{lm}\phi + \sum_{k=n_1}^{n_2} w_{lm}\xi[k] - (L-1). \end{aligned}$$

Assim,

$$\begin{aligned} \text{RFB}_{CC} &\leq \max_{ij,lm} \left[\frac{L-1}{w_{ij}} + \frac{L-1}{w_{lm}} \right] \\ &\leq \frac{2}{\min_{ij} w_{ij}} (L-1), \end{aligned}$$

o que conclui a prova. □

A.5 PROVA DO TEOREMA 4.10

Teorema. *Seja ρ_{ij} a banda reservada a um fluxo ij pela política PSLR. Neste caso, o escalonamento pode ser representado por um servidor latência-taxa $(\rho_{ij}; \theta_{ij})$ em que*

$$\theta_{ij} \leq \frac{(4 + \eta^{-1})B - 2(\phi_{ij} + P_{ij}) + NL}{r},$$

sendo $\phi_{ij} = w_{ij}B / \sum_{lm} w_{lm}$.

Demonstração. A prova que segue é semelhante àquela proposta em [90] para a política DRR. Suponha que o fluxo ij tenha se tornado ativo em um dado instante de tempo τ e que permaneça nesta condição até o instante de tempo t . Seja t_0 o início do primeiro ciclo de serviço após o instante de tempo τ e t_{k+1} o início do ciclo em que se encontra o instante t . Estes ciclos serão representados por n_0 e $n_0 + k + 1$.

O intervalo compreendido entre esses ciclos é ilustrado na Figura A.1. É fácil ver que a latência será máxima se τ for seguinte a uma oportunidade de serviço qualquer do fluxo ij , e t for imediatamente anterior à oportunidade de serviço para o ciclo $n_0 + k + 1$.

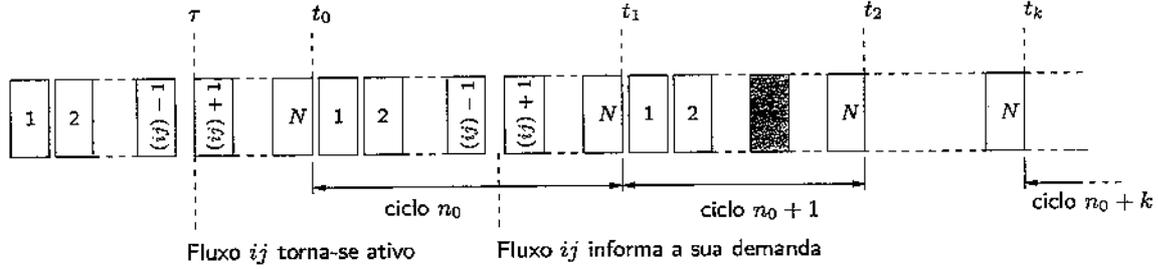


Figura A.1: Ciclos de serviço considerados na prova do Teorema 4.10.

Suponha que todos os fluxos $lm \neq ij$ estejam ativos para todo instante anterior a t . Sendo \mathcal{G} o conjunto de fluxos que sucedem o fluxo ij na ordem de serviço e dado que $\xi[n] \leq B$, temos que

$$\begin{aligned} r(t_0 - \tau) &= h_{ij} + \sum_{lm \in \mathcal{G}} h_{lm} + P_{lm} + \frac{w_{lm}}{\sum_{kp} w_{kp}} \xi[n_0] + d_{lm}[n_0 - 1] - d_{lm}[n_0] \\ &\leq h_{ij} + \sum_{lm \in \mathcal{G}} h_{lm} + P_{lm} + \frac{w_{lm}}{\sum_{kp} w_{kp}} B + d_{lm}[n_0 - 1] - d_{lm}[n_0], \end{aligned}$$

em que h_{ij} é o tempo de passagem (em bits) entre o fluxo ij e o próximo na ordem. Da mesma maneira,

$$r(t_1 - t_0) \leq h_{ij} + \sum_{lm \neq ij} h_{lm} + P_{lm} + \frac{w_{lm}}{\sum_{kp} w_{kp}} B + d_{lm}[n_0] - d_{lm}[n_0 + 1].$$

Dado que o fluxo ij transmite dados a partir de $n_0 + 1$, temos que

$$\begin{aligned} \xi[n + 1] &= \xi[n] + \eta(B - b[n]) \\ &= (1 - \eta)\xi[n] + \eta(B - P - h), \quad n \geq n_0 + 1, \end{aligned}$$

em que $P = \sum_{lm} P_{lm}$ e $h = \sum_{lm} h_{lm}$. A solução desta equação é dada por

$$\xi[n] = [\xi[n_0 + 1] - (B - P - h)](1 - \eta)^{n - n_0 - 1} + (B - P - h), \quad n \geq n_0 + 1.$$

Assim,

$$\begin{aligned} r(t_{k+1} - t_1) &= \sum_{n=n_0+1}^{n_0+k} P + h + \xi[n] + \sum_{n=n_0+1}^{n_0+k} \sum_{lm} d_{lm}[n] - d_{lm}[n+1] \\ &= \sum_{n=n_0+1}^{n_0+k} P + h + \xi[n] + \sum_{lm} d_{lm}[n_0 + 1] - d_{lm}[n_0 + k + 1] \\ &\leq k(P + h) + \sum_{n=n_0+1}^{n_0+k} \xi[n] + \sum_{lm} d_{lm}[n_0 + 1] - d_{lm}[n_0 + k + 1]. \end{aligned}$$

A partir da solução de $\xi[n]$, temos que

$$\begin{aligned} \sum_{n=n_0+1}^{n_0+k} \xi[n] &= \sum_{n=0}^{k-1} [\xi[n_0 + 1] - (B - P - h)](1 - \eta)^n + (B - P - h) \\ &\leq k(B - P - h) + (P + h) \sum_{n=0}^{k-1} (1 - \eta)^n \\ &= k(B - P - h) + \frac{P + h}{\eta} [1 - (1 - \eta)^k] \\ &\leq k(B - P - h) + \frac{P + h}{\eta} \end{aligned}$$

pois $\xi[n] \leq B$, $\forall n$. Desta maneira, sendo $B > P + h$,

$$\begin{aligned} r(t_{k+1} - t_1) &\leq kB + \frac{P + h}{\eta} + \sum_{lm} d_{lm}[n_0 + 1] - d_{lm}[n_0 + k + 1] \\ &= k(B - P - h) + \frac{P + h}{\eta} + \sum_{lm} d_{lm}[n_0 + 1] - d_{lm}[n_0 + k + 1] \\ &\leq kB + B\eta^{-1} + \sum_{lm} d_{lm}[n_0 + 1] - d_{lm}[n_0 + k + 1]. \end{aligned}$$

Finalmente, para o intervalo $[t_{k+1}; t]$ temos

$$\begin{aligned} r(t - t_{k+1}) &= \sum_{lm \notin \mathcal{G}} h_{lm} + P_{lm} + \frac{w_{lm}}{\sum_{kp} w_{kp}} \xi[n_0 + k + 1] + d_{lm}[n_0 + k + 1] - \\ &\quad d_{lm}[n_0 + k + 2] \\ &\leq \sum_{lm \notin \mathcal{G}} h_{lm} + P_{lm} + \frac{w_{lm}}{\sum_{kp} w_{kp}} B + d_{lm}[n_0 + k + 1] - \\ &\quad d_{lm}[n_0 + k + 2]. \end{aligned}$$

Assim, após alguma álgebra, temos que

$$\begin{aligned}
r(t-\tau) &\leq kB + 2h + \left[\eta^{-1} + 2 \left(1 - \frac{w_{ij}}{\sum_{lm} w_{lm}} \right) \right] B + 2 \sum_{lm \neq ij} P_{lm} + \\
&\quad \sum_{lm \in \mathcal{G}} d_{lm}[n_0 - 1] - d_{lm}[n_0 + k + 1] + d_{ij}[n_0 + 1] - d_{ij}[n_0 + k + 1] + \\
&\quad \sum_{lm \notin \mathcal{G}} d_{lm}[n_0] - d_{lm}[n_0 + k + 2] \\
&\leq kB + 2h + \left[\eta^{-1} + 2 \left(1 - \frac{w_{ij}}{\sum_{lm} w_{lm}} \right) \right] B + 2 \sum_{lm \neq ij} P_{lm} + NL.
\end{aligned}$$

Isolando o termo que contém o número de ciclos k , temos

$$kB \geq r(t-\tau) + 2h + 2 \sum_{lm \neq ij} P_{lm} + NL + \left[\eta^{-1} + 2 \left(1 - \frac{w_{ij}}{\sum_{lm} w_{lm}} \right) \right] B.$$

Dado que

$$\begin{aligned}
s_{ij}(\tau; t) &\geq kP_{ij} \\
&= k \frac{\rho_{ij} B}{r} \\
&\geq \frac{\rho_{ij}}{r} \left[r(t-\tau) + 2h + 2 \sum_{lm \neq ij} P_{lm} + NL + \left(\eta^{-1} + 2 - \frac{2w_{ij}}{\sum_{lm} w_{lm}} \right) B \right],
\end{aligned}$$

a latência θ_{ij} é dada por

$$\begin{aligned}
\theta_{ij} &= \frac{1}{r} \left[2h + 2 \sum_{lm \neq ij} P_{lm} + NL + \left(\eta^{-1} + 2 - \frac{2w_{ij}}{\sum_{lm} w_{lm}} \right) B \right] \\
&\leq \frac{1}{r} [(4 + \eta^{-1}) B - 2P_{ij} - 2\phi_{ij} + NL],
\end{aligned}$$

sendo $\phi_{ij} = w_{ij} B / \sum_{lm} w_{lm}$, o que conclui a prova. \square

APÊNDICE B

APRIMORAMENTO DO ALGORITMO PSLR

Neste apêndice, apresentamos uma versão aprimorada do algoritmo que implementa o escalonamento PSLR. Esta versão permite obter um limitante de latência θ_{ij} melhor do que (4.49), o que se traduz em limitantes mais precisos para o atraso introduzido nos fluxos.

B.1 ALGORITMO COMPUTACIONAL

O pseudo-código do escalonamento PSLR aprimorado é apresentado nos Algoritmos B.1 e B.2. O algoritmo pode ser descrito da seguinte maneira. Suponha que a OLT destine janelas de transmissão às ONU por meio de um mecanismo de *polling* intercalado com parada, conforme este foi apresentado na Seção 1.3.2. No início do n -ésimo ciclo, a OLT inicia as transmissões enviando uma mensagem de permissão a cada ONU j . Esta mensagem é composta por três campos:

- a identidade da ONU de destino, dada por $j \in \mathbb{Z}_+$;
- o estado da rede, $\xi[n] \in \mathbb{Z}_+$;
- a duração do ciclo anterior em bits, $b[n-1] \in \mathbb{Z}_+$.

Para inicializar o algoritmo, supomos que $\xi[n] = 0$ e que $b[n-1] = b[n] = B$ para $n = 0$. As mensagens de permissão podem ser enviadas por sinalização interposta ou por meio de encapsulamento em uma mensagem gate do protocolo MPCP.

Quando a ONU j recebe uma mensagem de permissão, ela calcula a quantidade de tráfego que cada fluxo ij deve enviar no $(n+1)$ -ésimo ciclo. Esse valor é dado por

$$c_{ij}[n+1] = \max \left\{ \hat{Q}_{ij}[n] - c_{ij}[n], [c_{ij}^e[n] + d_{ij}[n]]^+ \right\}, \quad (\text{B.1})$$

em que $[z]^+ = \max(0; z)$ e

$$c_{ij}^e[n] = \left[\frac{\rho_{ij}}{r} b[n-1] + w_{ij} \xi[n] \right]^+ \quad (\text{B.2})$$

e a operação $[z]^+$ arredonda z para cima, tal que não haja fragmentação de pacotes caso apenas z bits fossem transmitidos. Note que utilizamos $b[n-1]$ no lugar de B em (B.2), sendo esta a diferença em relação ao algoritmo PSLR apresentado no Capítulo 4.

Após calcular o valor de $c_{ij}[n+1]$, a ONU atualiza o valor de $d_{ij}[n]$ fazendo

$$d_{ij}[n+1] = \frac{\rho_{ij}}{r} b[n-1] + w_{ij} \xi[n] - c_{ij}[n+1] + d_{ij}[n]. \quad (\text{B.3})$$

Para evitar o acúmulo de resíduos, fazemos $d_{ij}[n+1] = 0$ sempre que $d_{ij}[n+1] > 0$ e a janela destinada ao fluxo ij no ciclo $n+1$ for suficiente para atender toda a sua demanda, i.e.,

$$c_{ij}[n+1] = \tilde{Q}_{ij}[n-1] - c_{ij}[n]. \quad (\text{B.4})$$

A ONU então marca os quadros a serem transmitidos durante o ciclo $n+1$ e, então, transmite os quadros marcados na iteração anterior, que totalizam $c_{ij}[n]$ bits. Ao recebê-los, a OLT os envia diretamente ao enlace de saída da rede. Finalmente, a ONU encerra a transmissão do ciclo corrente enviando à OLT uma mensagem de requisição composta dos seguintes campos:

- a identidade lógica da ONU, dada por $j \in \mathbb{Z}_+$;
- a demanda total para o ciclo $n+1$, dada por

$$C_j[n+1] = \sum_i c_{ij}[n+1], \quad (\text{B.5})$$

- a soma dos fatores de ponderação dos fluxos ativos após o serviço do ciclo $n+1$, i.e.,

$$S_j[n+1] = \sum_i w_{ij} \mathbb{I} \{ \tilde{Q}_{ij}[n+1] \neq 0 \}, \quad (\text{B.6})$$

em que a função $\mathbb{I}(\cdot)$ vale 1 quando o seu argumento é verdadeiro ou zero, caso contrário.

A mensagem de requisição pode ser enviada por sinalização interposta ou por encapsulamento em uma mensagem report do protocolo MPCP. Após receber a requisição da última ONU do ciclo, a OLT calcula o valor de $b[n]$ e, caso haja ONU com fluxos ativos (i.e., $\sum_j S_j[n] > 0$), atualiza o valor de $\xi[n]$ por meio de

$$\tilde{\xi}[n+1] = \left[\tilde{\xi}[n] + \eta (B - b[n]) \right]^+ \quad (\text{B.7})$$

$$\xi[n] = \frac{1}{\sum_j S_j} \tilde{\xi}[n] \quad (\text{B.8})$$

em que η é o passo do algoritmo. Note que $\tilde{\xi}[0] = 0$ dado que supusemos $\xi[0] = 0$.

Após atualizar a informação de estado da rede, a OLT determina o instante em que se deve enviar a próxima mensagem de permissão de cada ONU j . Esta mensagem leva consigo os valores de $\xi[n]$ e de $b[n]$, sendo que enviamos $b[n] = B$ sempre que não houver fluxos ativos, de modo a reinicializar o algoritmo sempre que estivermos nesta situação. O instante em que a OLT envia a permissão de cada ONU j é dado por

$$t_j[n+1] = \max \{ t_r[n]; t'_j[n+1] \}, \quad (\text{B.9})$$

em que $t'_j[n+1] = 0$ para $j = 0$ e

$$t'_j[n+1] = t_{j-1}[n+1] + \text{RTT}_{j-1} + \frac{1}{\tau} C_{j-1}[n+1] + G - \text{RTT}_j, \quad j > 0, \quad (\text{B.10})$$

em que $t_r[n]$ é o instante de tempo em que a OLT recebe a requisição da última ONU na ordem de serviço, RTT_j é o tempo de ida e volta da ONU j e G é o intervalo de guarda necessário para a preparar a camada física para a transmissão e para compensar flutuações no tempo de ida e volta das ONU [6, 17, 19]. Dispondo dos valores de $t_j[n+1]$, a OLT programa o envio das permissões de cada ONU, e um novo ciclo de *polling* se inicia.

Algoritmo B.1: Algoritmo para implementar a política PSLR --- Parte da OLT.

Dados: Passo do algoritmo η e conjunto de ONU \mathcal{O} .

$\tilde{\xi} \leftarrow 0;$

$\xi \leftarrow 0;$

/* Obtém o RTT e a demanda inicial de cada ONU */

para cada ONU j faça

$t \leftarrow \text{relogio}();$
 enviaPermissão (j, ξ, B);
 espera por requisição da ONU j ;
 $(C_j; S_j) \leftarrow \text{recebeRequisição}();$
 $RTT_j \leftarrow \text{relogio}() - t;$

$t_r \leftarrow t_1 \leftarrow \text{relogio}();$

$\mathcal{O} \leftarrow \text{ordenaPorRTT}(\mathcal{O});$

$b \leftarrow B;$

/* Laço principal */

repita

/* Envia a cada ONU uma mensagem de permissão */

para cada ONU $j \in \mathcal{O}$ faça

agendaPara ($t_j, \text{enviaPermissão}(j, \xi, b)$);
 se $j < N_j$ então
 $t_{j+1} \leftarrow t_j + RTT_j + \frac{1}{r}C_j + G - RTT_{j+1};$
 se $t_{j+1} < t_r$ então $t_{j+1} \leftarrow t_r;$

/* Recebe os quadros de cada ONU */

para cada ONU $j \in \mathcal{O}$ faça

repita
 Recebe quadro da ONU j ;
 até que sejam recebidos C_j bits de dados ;
 $(C_j; S_j) \leftarrow \text{recebeRequisição}();$
 $RTT_j \leftarrow \text{relogio}() - t_j;$

/* Atualiza o estado da rede ξ */

$b \leftarrow \text{relogio}() - t_r;$

se $\sum S_j > 0$ então

$\tilde{\xi} \leftarrow \tilde{\xi} + \eta(B - b);$
 se $\tilde{\xi} < 0$ então $\tilde{\xi} \leftarrow 0;$
 $\xi \leftarrow \tilde{\xi} / \sum_j S_j;$

senão

$b \leftarrow B;$

/* Atualiza a referência de tempo */

$t_r \leftarrow t_1 \leftarrow \text{relogio}();$

para sempre

Algoritmo B.2: Algoritmo para implementar a política PSLR — Parte da ONU.

Dados: Para cada fluxo i da ONU j , os parâmetros de SLA w_{ij} e ρ_{ij} .

$d_{ij} \leftarrow 0 \forall i$;

/* Laço principal */

repita

 espera por permissão da OLT;

$(\xi; b) \leftarrow \text{recebePermissão}()$;

 para cada fluxo i faça

$\tilde{Q}_{ij} \leftarrow$ demanda informada no ciclo anterior;

$c_{ij} \leftarrow$ pacotes marcados no ciclo anterior;

 envia os pacotes do fluxo i marcados no ciclo anterior;

$c_{ij}^E \leftarrow \rho_{ij}b/r + w_{ij}\xi$;

 se $c_{ij}^E - d_{ij} < 0$ então $c_{ij}^E \leftarrow 0$;

 senão

$c_{ij}^E \leftarrow c_{ij}^E + d_{ij}$;

 arredonda c_{ij}^E para não truncar pacotes e não exceder $\tilde{Q}_{ij} - c_{ij}$;

 marca os pacotes correspondentes a c_{ij}^E bits;

 para cada fluxo i faça

$d_{ij} \leftarrow d_{ij} + \rho_{ij}b/r + w_{ij}\xi - c_{ij}^E$;

 if $d_{ij} < 0 \wedge \tilde{Q}_{ij} - c_{ij} - c_{ij}^E = 0$ then $d_{ij} \leftarrow 0$;

/* Envia requisição */

$C_j \leftarrow \sum_i c_{ij}^E$;

$S_j \leftarrow \sum_i w_{ij} \mathbb{I} \{ \tilde{Q}_{ij} - c_{ij} - c_{ij}^E \neq 0 \}$;

enviaRequisição ($C_j; S_j$);

para sempre

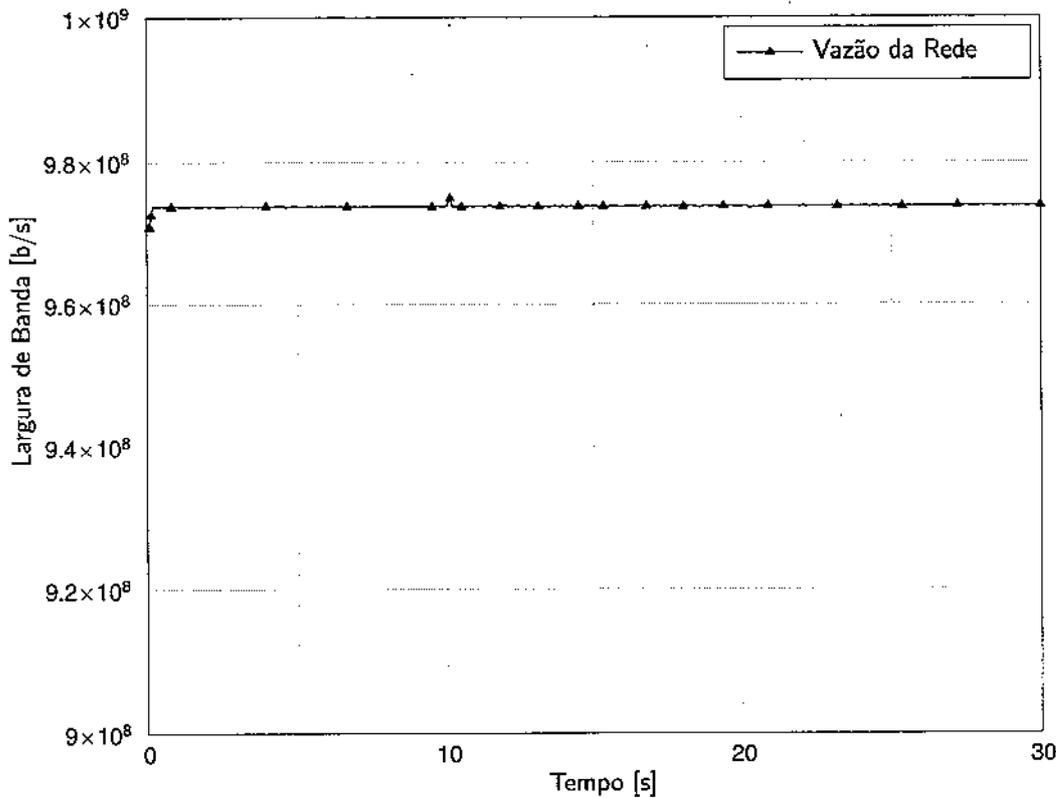


Figura B.1: Operação da política PSLR. Vazão da rede de acesso.

B.2 EXEMPLO NUMÉRICO

Nesta seção, ilustramos a operação da política PSLR implementada por meio dos Algoritmos B.1 e B.2. Para isso, consideramos a mesma rede de acesso EPON descrita na Seção 4.

A vazão da rede ao longo do intervalo de zero a 30s é mostrada na Figura B.1. Podemos observar que a vazão média se mantém constante em cerca de 975Mb/s, o que resulta em uma eficiência de 97,5% no aproveitamento da capacidade total da rede. A duração dos ciclos de *polling* é mostrada na Figura B.2 e o tempo de passagem total da rede é apresentado na Figura B.3.

A partir da vazão total, podemos calcular a parcela justa que deve ser destinada pela rede a cada um dos fluxos por meio de 4.38. Vale observar que os valores médios dos fluxos são tais que todos apresentam demanda reprimida enquanto estão gerando tráfego. Caso contrário, a relação (4.38) não se aplicaria e deveríamos recorrer a (4.25) para realizar o cálculo.

Os valores teóricos de banda de cada fluxo são indicados na Tabela 4.3. Os resultados experimentais obtidos para os fluxos da ONU 1 e 9 são apresentados nas Figuras B.4–B.7. Nestas figuras, os valores de banda indicados foram obtidos a partir do volume de tráfego de cada fluxo que é atendido em intervalos não-sobrepostos de 0,1s. Podemos observar que os resultados experimentais são bastante próximos daqueles obtidos por meio da relação (4.38). As oscilações observadas nos resultados empíricos se devem à ação do mecanismo de arredondamento que previne o fracionamento de pacotes.

A comparação entre os resultados teóricos e empíricos nos permite concluir que a política PSLR é efetivamente capaz de garantir banda mínima aos fluxos e de proporcionar o compartilhamento justo dos recursos ociosos da rede.

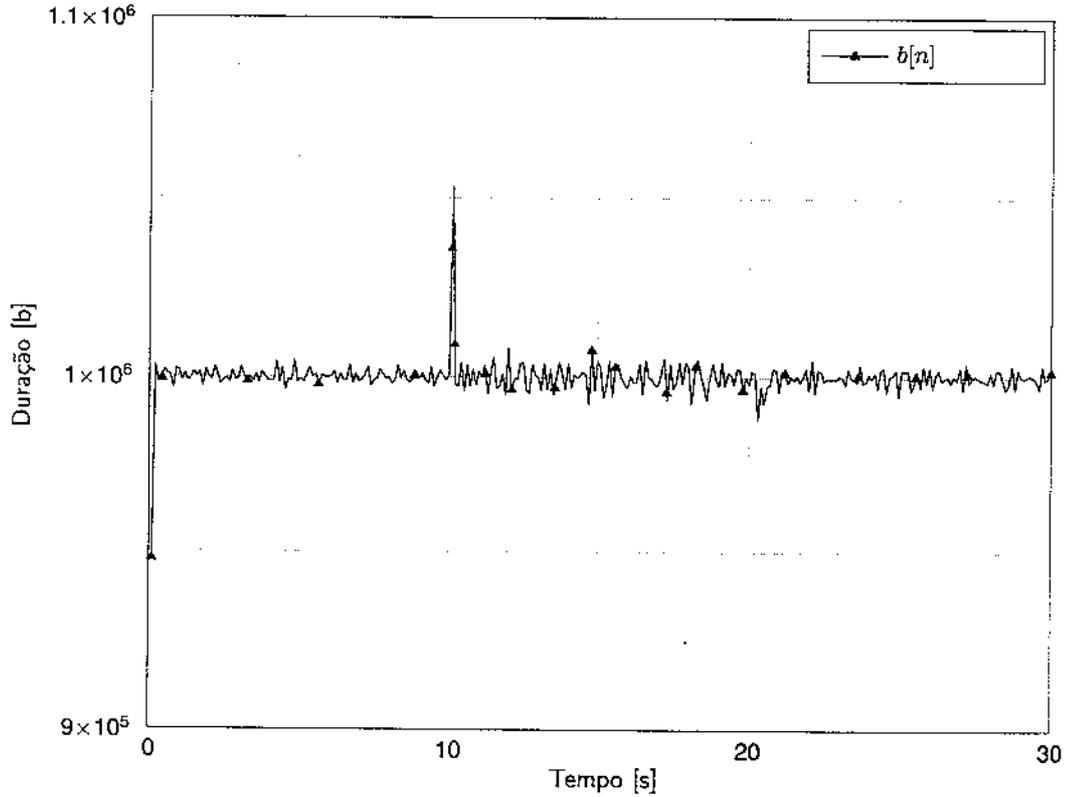


Figura B.2: Operação da política PSLR. Duração do ciclo de *polling*.

B.3 ANÁLISE DA POLÍTICA PSLR

B.3.1 CONVERGÊNCIA E COMPLEXIDADE COMPUTACIONAL

O seguinte teorema demonstra a convergência do algoritmo proposto na seção anterior:

Teorema B.1. *A solução de (4.17)–(4.21) pode ser obtida, de forma adaptativa, por meio dos Algoritmos B.1 e B.2.*

Demonstração. Segundo os mesmos argumentos que apresentamos na demonstração da convergência dos Algoritmos 4.1 e 4.2, podemos estudar a convergência dos Algoritmos B.1 e B.2 desconsiderando os efeitos de $d_{ij}[n]$.

Suponhamos que o conjunto $\mathcal{A}[n]$ varie lentamente em relação a n , ao menos bem mais lentamente do que a dinâmica da equação acima. Neste caso, podemos considerá-lo constante e dado por \mathcal{A}_∞ . Dado que

$$\begin{aligned}
 b[n+1] &= h + \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n+1] + \sum_{ij \notin \mathcal{A}_\infty} \frac{\rho_{ij}}{r} b[n] + \tilde{\xi}[n+1] \\
 &= h + \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n+1] + \sum_{ij \notin \mathcal{A}_\infty} \frac{\rho_{ij}}{r} b[n] + \tilde{\xi}[n] + \eta(B - b[n]) \\
 &= h + \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n+1] - \left(\eta - \sum_{ij \notin \mathcal{A}_\infty} \frac{\rho_{ij}}{r} \right) b[n] + \tilde{\xi}[n] + \eta B,
 \end{aligned}$$

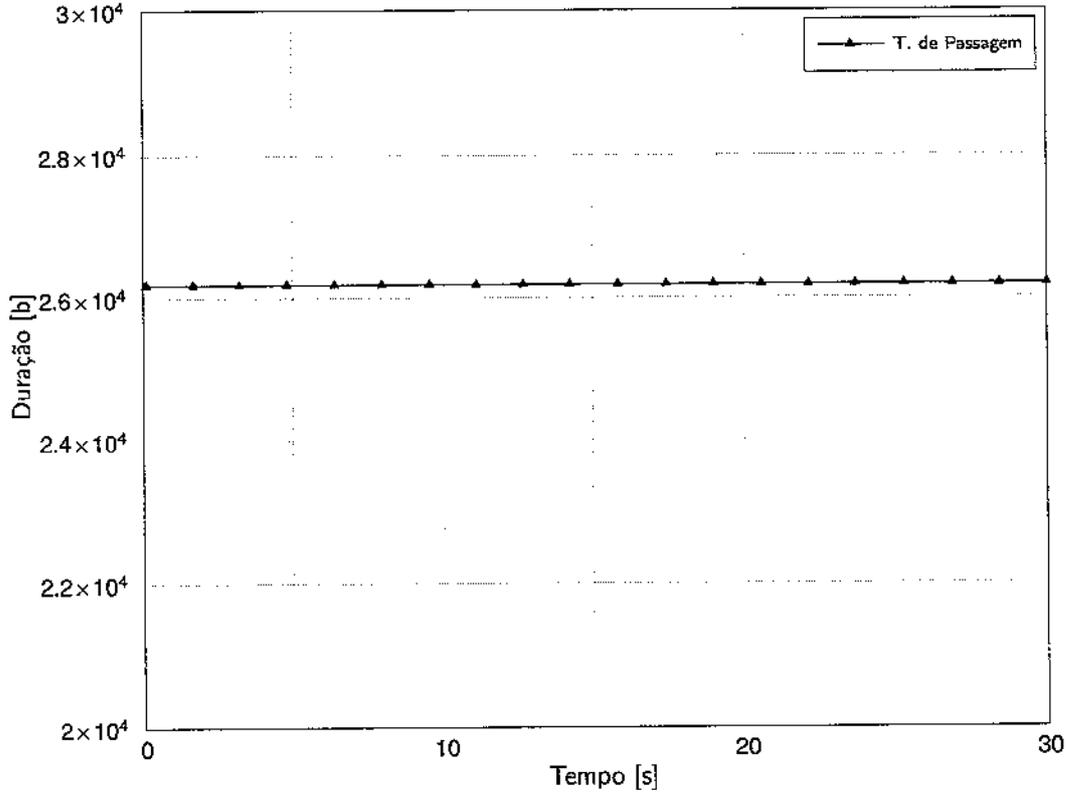


Figura B.3: Operação da política PSLR. Tempo de passagem total da rede.

temos então que

$$\begin{bmatrix} \tilde{\xi}[n+1] \\ \tilde{b}[n+1] \end{bmatrix} = \begin{bmatrix} 1 & -\eta \\ 1 & P-\eta \end{bmatrix} \begin{bmatrix} \tilde{\xi}[n] \\ \tilde{b}[n] \end{bmatrix} + \begin{bmatrix} \eta B \\ \eta B + h + \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij}[n+1] \end{bmatrix}, \quad (\text{B.11})$$

em que $P = \sum_{ij \in \mathcal{A}_\infty} \rho_{ij}/r$. Os pólos deste sistema são dados por

$$\begin{aligned} \begin{vmatrix} 1-\zeta & -\eta \\ 1 & P-\eta-\zeta \end{vmatrix} &= (1-\zeta)(P-\eta-\zeta) + \eta \\ &= \zeta^2 + (\eta-1-P)\zeta + P = 0. \end{aligned}$$

Para que os pólos desta equação sejam positivos e de módulo menor do que 1, tal que o sistema convirja, temos que $\eta < 1 + P$. Assim, para $n \rightarrow \infty$ e supondo que os fluxos $ij \in \mathcal{A}_\infty$ satisfazem $\lim_{n \rightarrow \infty} \tilde{Q}_{ij}[n] = \tilde{Q}_{ij}$, temos

$$\begin{bmatrix} \tilde{\xi}^* \\ \tilde{b}^* \end{bmatrix} = \begin{bmatrix} 1 & -\eta \\ 1 & P-\eta \end{bmatrix} \begin{bmatrix} \tilde{\xi}^* \\ \tilde{b}^* \end{bmatrix} + \begin{bmatrix} \eta B \\ \eta B + h + \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij} \end{bmatrix},$$

de modo que

$$\begin{bmatrix} \tilde{\xi}^* \\ \tilde{b}^* \end{bmatrix} = \begin{bmatrix} (1-P)B - h - \sum_{ij \in \mathcal{A}_\infty} \tilde{Q}_{ij} \\ B \end{bmatrix}.$$

Assim, demonstramos que o algoritmo converge para o valor ótimo. \square

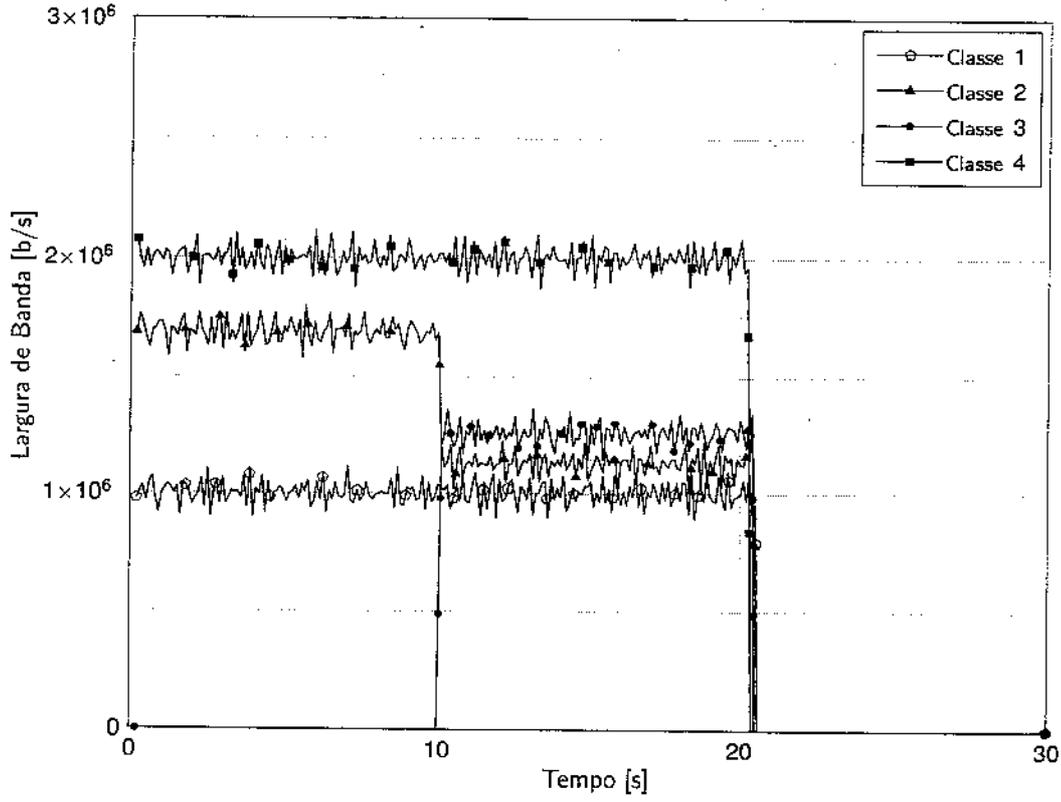


Figura B.4: Operação da política PSLR. Resultados para os fluxos da ONU 1.

A complexidade computacional do mesmo algoritmo é dada pelo seguinte teorema:

Teorema B.2. *A complexidade computacional da política PSLR implementada por meio dos Algoritmos B.1 e B.2 é linear com o número de fluxos e de ONU, i.e., $\mathcal{O}[(N_{max} + 1)J]$, em que N_{max} é o número máximo de fluxos e J é o número de ONU na rede.*

Demonstração. Veja a prova do Teorema 4.5. □

B.3.2 ESTABILIDADE

Para analisar a estabilidade da política PSLR, consideramos um cenário de pior caso em que todos os fluxos têm demanda reprimida, i.e., nenhum fluxo pertence a $\mathcal{A}[n]$. Assim, no ciclo $n \rightarrow \infty$, a rede garante a cada fluxo ij um serviço mínimo dado por

$$c_{ij}^{(min)} = P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left\{ B - h - \sum_{lm} P_{lm} \right\}. \quad (\text{B.12})$$

Como há demanda reprimida, o tamanho dos ciclos converge para B quando $n \rightarrow \infty$. Neste caso, a rede garante ao fluxo ij uma taxa média de serviço s_{ij} , dada por

$$\begin{aligned} s_{ij}^{(min)} &= \frac{c_{ij}^{(min)} \tau}{B} \\ &= \rho_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left\{ \tau - \frac{h}{B} - \sum_{lm} \rho_{lm} \right\}. \end{aligned} \quad (\text{B.13})$$

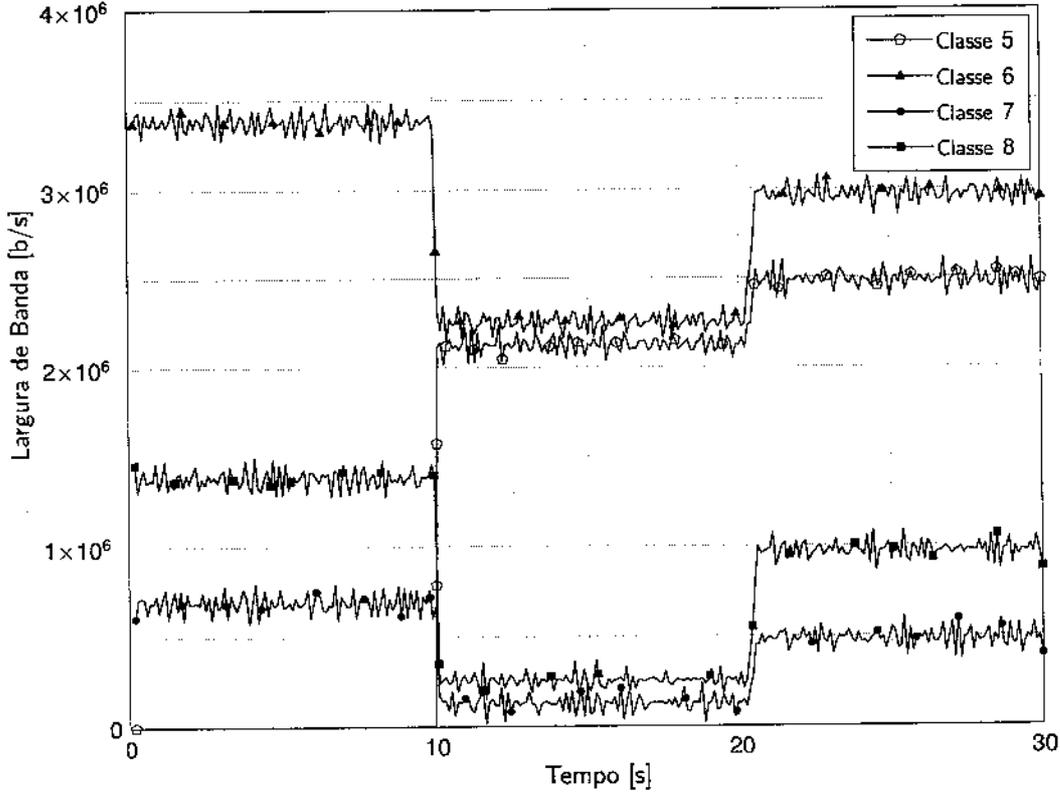


Figura B.5: Operação da política PSLR. Resultados para os fluxos da ONU 1.

Seja então o seguinte teorema:

Teorema B.3. *Sob a política PSLR, um dado fluxo i será localmente estável se*

$$\rho_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(\tau - \frac{r \cdot h}{B} - \sum_{lm} \rho_{lm} \right) \geq \mu_{ij}. \quad (\text{B.14})$$

em que μ_{ij} é a taxa média do fluxo ij . Todos os fluxos serão estáveis se

$$\sum_{ij} \mu_{ij} < \tau \left(1 - \frac{h}{B} \right). \quad (\text{B.15})$$

Demonstração. Suponhamos que $ij \notin \mathcal{A}[n]$ para todo fluxo ij . De (4.34),

$$\begin{aligned} \tilde{\xi}[n+1] &= \tilde{\xi}[n] + \eta(B - b[n]) \\ &= (1 - \eta)\tilde{\xi}[n] + \eta \left(B - h - \sum_{ij \notin \mathcal{A}[n]} \frac{\rho_{ij}}{r} b[n-1] - \sum_{ij \in \mathcal{A}[n]} \tilde{Q}_{ij} \right) \\ &\geq (1 - \eta)\tilde{\xi}[n] + \eta \left(B - h - \sum_{ij} \frac{\rho_{ij}}{r} b[n-1] \right), \end{aligned}$$

pois o serviço oferecido a um fluxo $ij \in \mathcal{A}[n]$ é inferior à parcela justa. Assim, para $n \rightarrow \infty$, temos que $\xi[n] \geq B - h - \sum_{ij} P_{ij}$, sendo $P_{ij} = \rho_{ij} B / r$, e

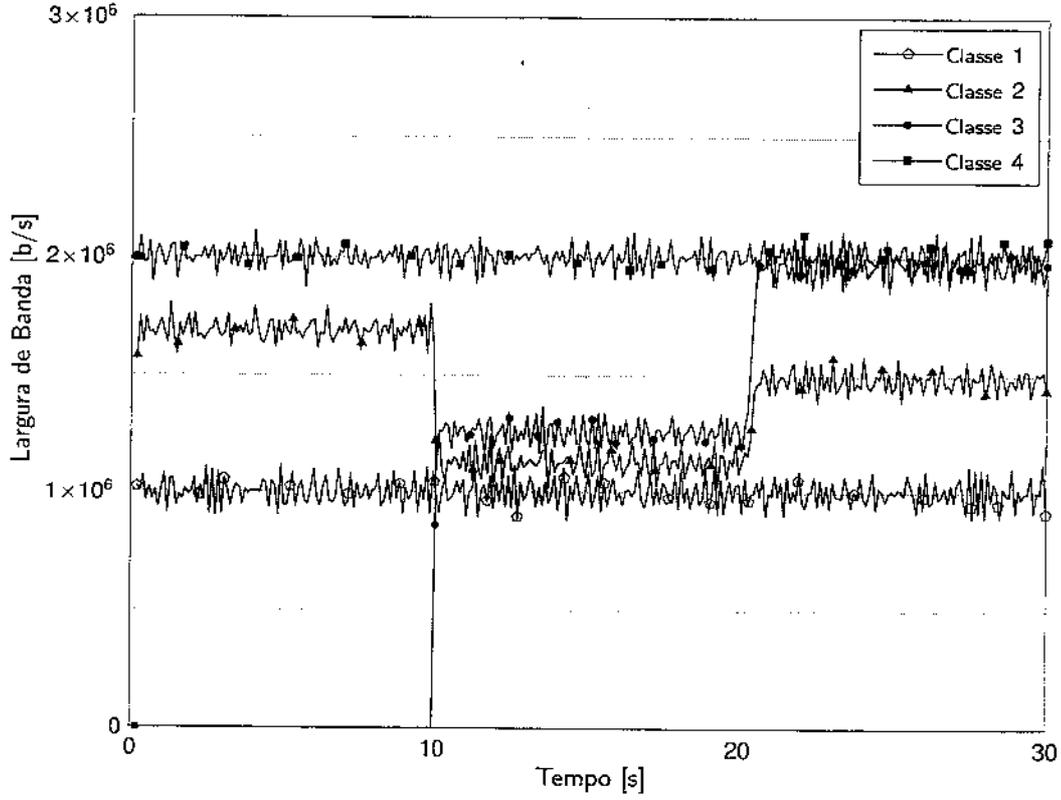


Figura B.6: Operação da política PSLR. Resultados para os fluxos da ONU 9.

$$c_{ij}[n] \geq P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(B - h - \sum_{lm} P_{lm} \right),$$

de modo que

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^k c_{ij}[n] \geq P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(B - h - \sum_{lm} P_{lm} \right).$$

Portanto, para que o fluxo ij seja localmente estável, é suficiente que

$$P_{ij} + \frac{w_{ij}}{\sum_{lm} w_{lm}} \left(B - h - \sum_{lm} P_{lm} \right) \geq \frac{\mu_{ij}}{r} B.$$

A partir da definição de $b[n]$, tem-se que

$$\lim_{k \rightarrow \infty} \sum_{n=0}^k b[n] = \lim_{k \rightarrow \infty} \sum_{n=0}^k \sum_{ij} c_{ij}[n] + \lim_{k \rightarrow \infty} h \cdot k.$$

Para que o fluxo ij seja estável, é suficiente que a rede seja estável. Ou seja,

$$\lim_{k \rightarrow \infty} \sum_{n=0}^k c_{ij}[n] > \frac{\mu_{ij}}{r} \lim_{k \rightarrow \infty} \sum_{n=0}^k b[n].$$

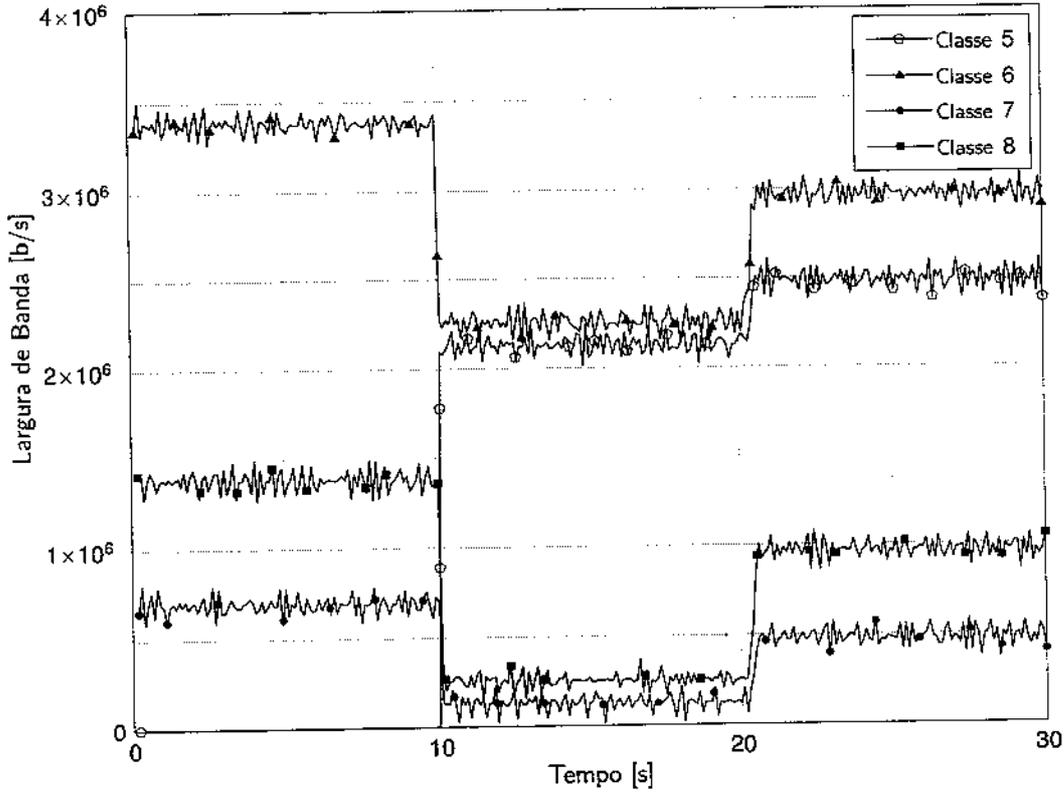


Figura B.7: Operação da política PSLR. Resultados para os fluxos da ONU 9.

Reunindo as duas relações anteriores e dado que a convergência do algoritmo leva a $\lim_{n \rightarrow \infty} b[n] = B$, tem-se que

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^k b[n] > h + \sum_{ij} \frac{\mu_{ij}}{r} \left(\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^k b[n] \right)$$

$$B > h + \sum_{ij} \frac{\mu_{ij}}{r} B,$$

o que conclui a prova. \square

B.3.3 JUSTIÇA

Para medir a justiça da política PSLR, utiliza-se a métrica de RFB apresentada na Seção 4.3. Seja o seguinte teorema:

Teorema B.4. Para a política PSLR, o limitante de justiça relativo satisfaz a desigualdade

$$RFB \leq \frac{1}{\min_{ij} w_{ij}} \max_{ij} \left[\rho_{ij} \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} + B - h + 2(L-1) \right]. \quad (B.16)$$

Demonstração. Sejam dois fluxos $ij, lm \in \mathcal{B}(\tau; t)$ quaisquer. Sejam ainda n_1 e n_2 o número de vezes em que o fluxo ij é atendido até os instantes τ e t , respectivamente. Supondo que o fluxo ij antecede o fluxo lm na ordem, o serviço oferecido a cada um no intervalo $(\tau; t)$ é dado por

$$S_{ij}(\tau; t) = \sum_{k=n_1+1}^{n_2} c_{ij}[k] + d_{ij}[n_1] - d_{ij}[n_2],$$

$$S_{lm}(\tau; t) \geq \sum_{k=n_1+1}^{n_2-1} c_{lm}[k] + d_{lm}[n_1] - d_{lm}[n_2].$$

em que $c_{ij}[k] = \rho_{ij}b[k-1]/r + w_{ij}\xi[k]$. Dado que $-L+1 \leq d_{ij}[n] \leq 0$,

$$S_{ij}(\tau; t) \leq \sum_{k=n_1+1}^{n_2} c_i[k] + (L-1),$$

$$S_{lm}(\tau; t) \geq \sum_{k=n_1+1}^{n_2-1} c_j[k] - (L-1).$$

No Teorema B.7 adiante, demonstraremos que

$$b[n] \leq \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} r, \forall n.$$

Assim, seja $\phi = \sum_{n=n_1}^{n_2} b[n]/r$. Da definição de $\xi[n]$, temos que

$$S_{ij}(\tau; t) \leq \rho_{ij}\phi + \sum_{k=n_1+1}^{n_2-1} w_{ij}\xi[k] + w_{ij}\xi[n_2] + (L-1),$$

$$S_{lm}(\tau; t) \geq \rho_{lm}\phi - \frac{\rho_{lm}}{r}b[n_2] + \sum_{k=n_1+1}^{n_2-1} w_{lm}\xi[k] - (L-1)$$

$$\geq \rho_{lm}\phi - \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} + \sum_{k=n_1+1}^{n_2-1} w_{lm}\xi[k] - (L-1)$$

A partir de (4.34),

$$\begin{aligned} \tilde{\xi}[n+1] &= \tilde{\xi}[n] + \eta(B - b[n]) \\ &= (1-\eta)\tilde{\xi}[n] + \eta \left(B - h - \sum_{ij \notin \mathcal{A}[n]} \frac{\rho_{ij}}{r} b[n] - \sum_{j \in \mathcal{A}[n]} \tilde{Q}_j[n] \right) \\ &\leq (1-\eta)\tilde{\xi}[n] + \eta(B - h). \end{aligned}$$

Assim, temos que $\tilde{\xi}[n] \leq B - h$ e

$$\begin{aligned} \text{RFB} &\leq \max_{ij, lm} \left[\frac{\rho_{lm}}{w_{lm}} \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} + \tilde{\xi}[n_2] + \frac{L-1}{w_{ij}} + \frac{L-1}{w_{lm}} \right] \\ &\leq \frac{1}{\min_{ij} w_{ij}} \max_{ij} \left[\rho_{ij} \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} + B - h + 2(L-1) \right], \end{aligned}$$

o que conclui a prova. \square

A diferença entre os serviços oferecidos a dois fluxos ij e lm quaisquer em um intervalo de n ciclos de *polling* é dado pelo seguinte teorema:

Teorema B.5. *Seja $\mathcal{B}[n_1; n_2]$ o conjunto de fluxos latentes durante todos os ciclos $n_1 \leq n \leq n_2$. Para a política PSLR, o RFB_{CC} satisfaz a desigualdade*

$$\text{RFB}_{\text{CC}} \leq \frac{2}{\min_{ij} w_{ij}} (L - 1). \quad (\text{B.17})$$

Demonstração. Similar à demonstração do Teorema 4.8. □

B.3.4 LATÊNCIA E ATRASO

Para uma política que pertença à classe dos servidores latência-taxa, é possível obter limitantes para a latência e para o atraso dos fluxos se o tráfego for policiado pelo algoritmo do balde furado. Seja o seguinte teorema [25]:

Teorema B.6. *Seja um fluxo i cujo tráfego é policiado pelo algoritmo do balde furado com parâmetros ρ_i e σ_i . Se, para este fluxo, o servidor atua como um servidor latência-taxa com parâmetros $(\rho_i; \theta_i)$, a latência e o atraso satisfazem*

$$Q_i(t) \leq \sigma_i + \rho_i \theta_i, \quad (\text{B.18})$$

$$D_i(t) \leq \frac{\sigma_i}{\rho_i} + \theta_i. \quad (\text{B.19})$$

Pode-se demonstrar que a política PSLR pertence à classe dos servidores latência-taxa. Seja o seguinte teorema:

Teorema B.7. *Seja ρ_{ij} a banda reservada a um fluxo ij pela política PSLR. Neste caso, o escalonamento pode ser representado por um servidor latência-taxa $(\rho_{ij}; \theta_{ij})$ em que*

$$\theta_{ij} \leq \frac{1}{r} \left[3B - h + 4N(L - 1) + \frac{B + 2N(L - 1)}{r - \sum_{kp} \rho_{kp}} \left(r + 2 \sum_{lm \neq ij} \rho_{lm} \right) \right], \quad (\text{B.20})$$

sendo $\phi_{ij} = w_{ij}B / \sum_{lm} w_{lm}$. Se os fluxos têm reserva de banda nula, a latência então satisfaz a relação

$$\theta_{ij} \leq \frac{1}{r} [4B - h + 6N(L - 1)]. \quad (\text{B.21})$$

Demonstração. Suponha que o fluxo ij tenha se tornado ativo em um dado instante de tempo τ e que permaneça nesta condição até o instante de tempo t . Seja t_0 o início do primeiro ciclo de serviço após o instante de tempo τ e t_{k+1} o início do ciclo em que se encontra o instante t . Estes ciclos serão representados por n_0 e $n_0 + k + 1$.

O intervalo compreendido entre esses ciclos é ilustrado na Figura B.8. É fácil ver que a latência será máxima se τ for seguinte a uma oportunidade de serviço qualquer do fluxo ij , e t for imediatamente anterior à oportunidade de serviço para o ciclo $n_0 + k + 1$.

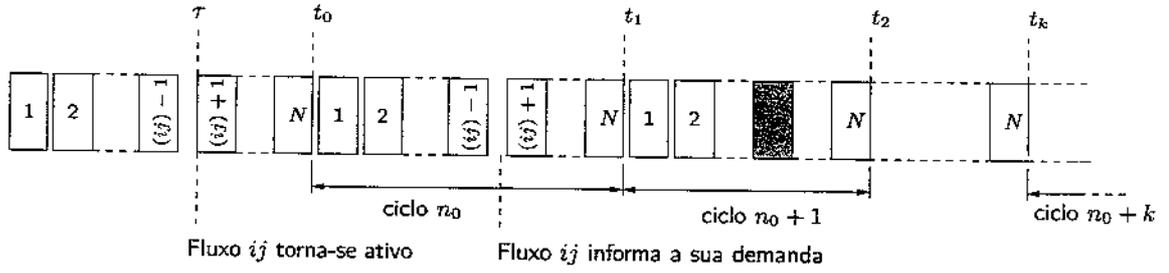


Figura B.8: Ciclos de serviço considerados na prova do Teorema B.7.

Para demonstrar o teorema, precisamos primeiro estabelecer um limitante para a duração do ciclo de *polling*. A partir de (B.7), temos que

$$\begin{aligned}\tilde{\xi}[n+1] &= \tilde{\xi}[n] + \eta(B - b[n]) \\ &= \tilde{\xi}[n] + \eta \left(B - h - \tilde{\xi}[n] - \sum_{lm \notin \mathcal{A}[n]} \frac{\rho_{lm}}{r} b[n-1] - \sum_{lm \in \mathcal{A}[n]} \tilde{Q}_{lm}[n] + \right. \\ &\quad \left. - \sum_{lm} d_{lm}[n] + \sum_{lm} d_{lm}[n+1] \right).\end{aligned}$$

Como $-(L-1) \leq d_{lm}[n] \leq 0$, temos

$$\tilde{\xi}[n+1] \leq (1-\eta)\tilde{\xi}[n] + \eta[B - h + N(L-1)],$$

e, sendo $\tilde{\xi}[0] < B - h + N(L-1)$, temos que $\tilde{\xi}[n] < B - h + N(L-1)$ para todo n . A partir da definição de $b[n]$, temos então que

$$\begin{aligned}b[n] &= h + \tilde{\xi}[n] + \sum_{lm \notin \mathcal{A}[n]} \frac{\rho_{lm}}{r} b[n-1] + \sum_{lm \in \mathcal{A}[n]} \tilde{Q}_{lm}[n] + \sum_{lm} d_{lm}[n] - \sum_{lm} d_{lm}[n+1] \\ &\leq h + \tilde{\xi}[n] + \sum_{lm} \frac{\rho_{lm}}{r} b[n-1] + N(L-1),\end{aligned}$$

pois os fluxos $lm \in \mathcal{A}[n]$ transmitem menos fluxos do que se estivessem na condição $lm \notin \mathcal{A}[n]$. Substituindo o limitante obtido para $\tilde{\xi}[n]$, temos então que

$$b[n] \leq \sum_{lm} \frac{\rho_{lm}}{r} b[n-1] + B + 2N(L-1).$$

Supondo que o sistema estivesse em equilíbrio em algum momento antes do intervalo $[\tau; t]$ considerado, tal que $b[n] = y$, teríamos

$$y \leq \sum_{lm} \frac{\rho_{lm}}{r} y + B + 2N(L-1),$$

de modo que

$$y \leq \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} r.$$

Como $b[0] = B < y$, temos que $b[n]$ não pode ultrapassar o valor de y calculado e este representa um limitante para o maior tamanho de ciclo possível. Assim, sejam \mathcal{G} e $\tilde{\mathcal{G}}$ os conjuntos de fluxos que sucedem e antecedem o fluxo ij na ordem de serviço, respectivamente. Temos então que

$$\begin{aligned}
r(t_0 - \tau) &= h_{ij} + \sum_{lm \in \mathcal{G}} h_{lm} + \frac{\rho_{lm}}{r} b[n_0 - 2] + \frac{w_{lm}}{\sum_{kp} w_{kp}} \tilde{\xi}[n_0 - 1] + d_{lm}[n_0 - 1] - d_{lm}[n_0] \\
&\leq h_{ij} + \sum_{lm \in \mathcal{G}} \left(h_{lm} + \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} + d_{lm}[n_0 - 1] - d_{lm}[n_0] \right) + [B - h + N(L-1)],
\end{aligned}$$

em que h_{ij} é o tempo de passagem (em bits) entre o fluxo ij e o próximo na ordem. Da mesma maneira,

$$\begin{aligned}
r(t_1 - t_0) &= h_{ij} + \sum_{lm \neq ij} h_{lm} + \frac{\rho_{lm}}{r} b[n_0 - 1] + \frac{w_{lm}}{\sum_{kp} w_{kp}} \tilde{\xi}[n_0] + d_{lm}[n_0] - d_{lm}[n_0 + 1] \\
&\leq h_{ij} + \sum_{lm \neq ij} \left(h_{lm} + \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} + d_{lm}[n_0] - d_{lm}[n_0 + 1] \right) + [B - h + N(L-1)].
\end{aligned}$$

A partir do ciclo $n_0 + 1$, o fluxo ij começa a transmitir. Temos então que

$$\begin{aligned}
r(t_k - t_1) &= \sum_{n=n_0+1}^{n_0+k-1} b[n] + \sum_{n=n_0+1}^{n_0+k-1} \sum_{lm} d_{lm}[n] - d_{lm}[n+1] \\
&= \sum_{n=n_0+1}^{n_0+k-1} b[n] + \sum_{lm} d_{lm}[n_0+1] - \sum_{lm} d_{lm}[n_0+k].
\end{aligned}$$

Finalmente, para o intervalo $[t_k; t]$ temos

$$\begin{aligned}
r(t - t_k) &= \sum_{lm \in \mathcal{G}} h_{lm} + \frac{\rho_{lm}}{r} b[n_0 + k] + \frac{w_{lm}}{\sum_{kp} w_{kp}} \tilde{\xi}[n] + d_{lm}[n_0 + k] - d_{lm}[n_0 + k + 1] \\
&\leq \sum_{lm \in \mathcal{G}} \left(h_{lm} + \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} + d_{lm}[n_0 + k] - d_{lm}[n_0 + k + 1] \right) + [B - h + N(L-1)].
\end{aligned}$$

Assim, temos que

$$\begin{aligned}
r(t - \tau) &\leq h_{ij} + \sum_{lm \in \mathcal{G}} \left(h_{lm} + \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} + d_{lm}[n_0 - 1] - d_{lm}[n_0] \right) + [B - h + N(L-1)] + \\
&h_{ij} + \sum_{lm \neq ij} \left(h_{lm} + \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} + d_{lm}[n_0] - d_{lm}[n_0 + 1] \right) + [B - h + N(L-1)] + \\
&\sum_{n=n_0+1}^{n_0+k-1} b[n] + \sum_{lm} d_{lm}[n_0+1] - \sum_{lm} d_{lm}[n_0+k+1] + \\
&\sum_{lm \in \mathcal{G}} \left(h_{lm} + \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{lm} \rho_{lm}} + d_{lm}[n_0+k] - d_{lm}[n_0+k+1] \right) + [B - h + N(L-1)].
\end{aligned}$$

Após alguma álgebra,

$$r(t - \tau) \leq \sum_{n=n_0+1}^{n_0+k-1} b[n] + h + 3B + 4N(L-1) + 2 \sum_{lm \neq ij} \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}}.$$

Isolando o termo com o somatório, temos

$$\sum_{n=n_0+1}^{n_0+k-1} b[n] \geq r(t-\tau) - 3B + h - 4N(L-1) - 2 \sum_{lm \neq ij} \rho_{lm} \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}}.$$

Dado que

$$\begin{aligned} S_{ij}(\tau; t) &\geq \frac{\rho_{ij}}{r} \sum_{n=n_0+1}^{n_0+k-1} b[n-1] \\ &= \frac{\rho_{ij}}{r} \sum_{n=n_0}^{n_0+k-2} b[n] \\ &= \frac{\rho_{ij}}{r} \left(b[n_0] - b[n_0+k-1] + \sum_{n=n_0+1}^{n_0+k-1} b[n] \right) \\ &\geq \frac{\rho_{ij}}{r} \left(-r \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} + \sum_{n=n_0+1}^{n_0+k-1} b[n] \right), \end{aligned}$$

temos

$$S_{ij}(\tau; t) \geq \rho_{ij} \left\{ t - \tau - \frac{1}{r} \left[3B - h + 4N(L-1) + \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} \left(r + 2 \sum_{lm \neq ij} \rho_{lm} \right) \right] \right\},$$

e a latência θ_{ij} é limitada então por

$$\theta_{ij} \leq \frac{1}{r} \left[3B - h + 4N(L-1) + \frac{B + 2N(L-1)}{r - \sum_{kp} \rho_{kp}} \left(r + 2 \sum_{lm \neq ij} \rho_{lm} \right) \right].$$

□