

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial

DISSERTAÇÃO
apresentada à UTFPR
para obtenção do título de

MESTRE EM CIÊNCIAS

por

EDMUNDO ZUCHOWSKI FILHO

**UMA ABORDAGEM PARA ANÁLISE DE DESEMPENHO DE
FLUXOS VoIP EM REDES DE SERVIÇOS DIFERENCIADOS**

Banca Examinadora

Presidente e Orientadora

PROF^a DR^a ENG^a KEIKO VERÔNICA ONO FONSECA

UTFPR

Examinadores

PROF. DR. ENG. PAULO ROBERTO GUARDIEIRO

UFU

PROF. DR. ENG. EMÍLIO CARLOS GOMES WILLE

UTFPR

Curitiba, Outubro de 2005

EDMUNDO ZUCHOWSKI FILHO

**UMA ABORDAGEM PARA ANÁLISE DE DESEMPENHO DE FLUXOS VoIP EM
REDES DE SERVIÇOS DIFERENCIADOS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do título de “Mestre em Ciências” – Área de Concentração: Telemática.

Orientadora:

Profª Drª Engª Keiko Verônica Ono Fonseca

Curitiba

2005

Ficha catalográfica elaborada pela Biblioteca da UTFPR — Campus Curitiba

Z94a Zuchowski Filho, Edmundo

Uma abordagem para análise de desempenho de fluxos VoIP em redes de serviços diferenciados / Edmundo Zuchowski Filho. Curitiba. UTFPR, 2005.

xxi, 125 f. : il.; 30 cm.

Orientador: Prof^a. Dr^a. Keiko Verônica Ono Fonseca

Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. Curitiba, 2005.

Bibliografia: f. 121-125.

1. Telecomunicações. 2. Redes de computadores. 3. Codificadores de voz.

I. Fonseca, Keiko Verônica Ono, orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial.

CDD: 621.382

AGRADECIMENTOS

Inicialmente, agradeço à Pontifícia Universidade Católica do Paraná - PUC-PR por permitir a utilização do convênio com a Universidade de Paris VI através do qual foi possível empregar a ferramenta de simulação OPNET Modeler 10.

Agradeço a minha esposa e companheira, Zaze, por compreender minha ausência em muitas oportunidades de lazer e pelo incentivo durante todo o tempo. Devo agradecer a todos os colegas do mestrado pelo espírito de luta e companheirismo e a todos os professores das disciplinas do mestrado pelos novos conhecimentos e pela amizade. Devo agradecer também a colaboração especial de amigos e colegas com críticas proveitosas e ajuda nas simulações. Obrigado, Carlos Pedroso. Obrigado, Thiago Pereira. Finalmente, quero expressar meus sinceros agradecimentos à Professora Keiko, pelas orientações, pelas discussões, pela paciência, pela dedicação e pela amizade. Sua participação foi fundamental. Obrigado, Keiko.

SUMÁRIO

LISTA DE FIGURAS.....	vii
LISTA DE TABELAS.....	xiii
LISTA DE ABREVIATURAS E SIGLAS.....	xv
RESUMO.....	xix
ABSTRACT.....	xxi
CAPÍTULO 1 INTRODUÇÃO.....	1
CAPÍTULO 2 FUNDAMENTOS TEÓRICOS.....	5
2.1 ARQUITETURA DE SERVIÇOS DIFERENCIADOS (DiffServ).....	5
2.2 QUALIDADE DE SERVIÇO PARA REDES IP.....	7
2.3 ACORDO DE NÍVEL DE SERVIÇO (SLA).....	10
2.4 PARÂMETROS DE QoS E OBJETIVOS DE DESEMPENHO DE REDE IP.....	13
2.5 CLASSES DE SERVIÇO IP DENTRO DA ARQUITETURA DiffServ.....	17
2.6 MEDIÇÃO DE FLUXOS SINTÉTICOS EM UM AMBIENTE DiffServ.....	18
CAPÍTULO 3 VOZ SOBRE IP.....	21
3.1 CODIFICADORES DE VOZ.....	21
3.2 MODELAGEM DO TRÁFEGO VoIP.....	27
3.3 TRABALHOS CORRELACIONADOS.....	36
CAPÍTULO 4 DESENVOLVIMENTO DO TRABALHO.....	41
4.1 MOTIVAÇÃO DO TRABALHO.....	42
4.2 MODELO SISTÊMICO PROPOSTO.....	42
4.3 PROPOSTA DE TRABALHO.....	43
4.4 PREMISSAS.....	44
4.5 AMBIENTE DE SIMULAÇÃO.....	45
4.6 DESCRIÇÃO DO PROCESSO DE SIMULAÇÃO.....	45
4.7 RESULTADOS DAS SIMULAÇÕES.....	47
4.7.1 Resultados do cenário de tráfego homogêneo.....	49
4.7.2 Resultados do cenário de tráfego heterogêneo.....	50
CAPÍTULO 5 CONCLUSÕES.....	57
ANEXO 1 – DISTRIBUIÇÕES ESTATÍSTICAS MAIS EMPREGADAS NA MODELAGEM DE TRÁFEGO VoIP.....	63

ANEXO 2 – COEFICIENTE DE CORRELAÇÃO.....	69
ANEXO 3 – RESULTADOS TRÁFEGO HOMOGÊNEO G.711.....	73
ANEXO 4 - RESULTADOS TRÁFEGO HOMOGÊNEO G.723.1.....	79
ANEXO 5 - RESULTADOS TRÁFEGO HOMOGÊNEO G.729B.....	85
ANEXO 6 – RESULTADOS TRÁFEGO HETEROGÊNEO G.711 E G.723.1.....	91
ANEXO 7 – RESULTADOS TRÁFEGO HETEROGÊNEO G.711 E G.729B.....	101
ANEXO 8 – RESULTADOS TRÁFEGO HETEROGÊNEO G.729B E G.723.1.....	111
REFERÊNCIAS BIBLIOGRÁFICAS.....	121

LISTA DE FIGURAS

Figura 2.1	Classificação e condicionamento de tráfego em DiffServ	7
Figura 2.2	O modelo geral de QoS com as abordagens da ETSI/ITU e da IETF.....	8
Figura 2.3	Relações entre SLA, SLS, TCA e TCS.....	12
Figura 2.4	Requisitos de QoS centrados no usuário.....	14
Figura 2.5	Curvas de tolerância para o eco do locutor.....	15
Figura 2.6	Satisfação dos usuários quanto ao limite de retardo absoluto.....	16
Figura 3.1	Comparação entre traços de tráfego Ethernet LAN e de rede convencional...	30
Figura 4.1	Diagrama sistêmico da rede sob simulação.....	42
Figura A3.1	Retardo mínimo no fluxo de controle G.711 em função do carregamento....	73
Figura A3.2	Retardo médio no fluxo de controle G.711 em função do carregamento.....	73
Figura A3.3	Retardo máximo no fluxo de controle G.711 em função do carregamento....	74
Figura A3.4	<i>Jitter</i> mínimo no fluxo de controle G.711 em função do carregamento.....	74
Figura A3.5	<i>Jitter</i> médio no fluxo de controle G.711 em função do carregamento.....	75
Figura A3.6	<i>Jitter</i> máximo no fluxo de controle G.711 em função do carregamento.....	75
Figura A3.7	Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.711.....	76
Figura A3.8	Variação do retardo máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.711.....	76
Figura A3.9	Variação do <i>jitter</i> médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.711.....	77
Figura A3.10	Variação do <i>jitter</i> máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.711.....	77
Figura A4.1	Retardo mínimo no fluxo de controle G.723.1 em função do carregamento..	79
Figura A4.2	Retardo médio no fluxo de controle G.723.1 em função do carregamento....	79
Figura A4.3	Retardo máximo no fluxo de controle G.723.1 em função do carregamento.	80
Figura A4.4	<i>Jitter</i> mínimo no fluxo de controle G.723.1 em função do carregamento.....	80
Figura A4.5	<i>Jitter</i> médio no fluxo de controle G.723.1 em função do carregamento.....	81
Figura A4.6	<i>Jitter</i> máximo no fluxo de controle G.723.1 em função do carregamento.....	81
Figura A4.7	Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.723.1.....	82

Figura A4.8	Variação do retardo máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.723.1.....	82
Figura A4.9	Variação do jitter médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.723.1.....	83
Figura A4.10	Variação do jitter máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.723.1.....	83
Figura A5.1	Retardo mínimo no fluxo de controle G.729B em função do carregamento..	85
Figura A5.2	Retardo médio no fluxo de controle G.729B em função do carregamento.....	85
Figura A5.3	Retardo máximo no fluxo de controle G.729B em função do carregamento..	86
Figura A5.4	<i>Jitter</i> mínimo no fluxo de controle G.729B em função do carregamento.....	86
Figura A5.5	<i>Jitter</i> médio no fluxo de controle G.729B em função do carregamento.....	87
Figura A5.6	<i>Jitter</i> máximo no fluxo de controle G.729B em função do carregamento.....	87
Figura A5.7	Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.729B.....	88
Figura A5.8	Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.729B.....	88
Figura A5.9	Variação do jitter médio normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.729B.....	89
Figura A5.10	Variação do jitter máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando codec G.729B.....	89
Figura A6.1	Retardo mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	91
Figura A6.2	Retardo médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	92
Figura A6.3	Retardo máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	92
Figura A6.4	Jitter mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	93
Figura A6.5	Jitter médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	93
Figura A6.6	Retardo mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	93
Figura A6.7	Variação do retardo médio normalizado em relação ao fluxo de controle	

	G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	94
Figura A6.8	Variação do retardo máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	94
Figura A6.9	Variação do jitter médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	95
Figura A6.10	Variação do jitter máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	95
Figura A6.11	Retardo mínimo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	96
Figura A6.12	Retardo médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	96
Figura A6.13	Retardo máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	96
Figura A6.14	Jitter médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	97
Figura A6.15	Jitter máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	97
Figura A6.16	Variação do retardo médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	98
Figura A6.17	Variação do retardo máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	98
Figura A6.18	Variação do jitter médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	99
Figura A6.19	Variação do jitter máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.723.1.....	99

Figura A7.1	Retardo mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	101
Figura A7.2	Retardo médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	102
Figura A7.3	Retardo máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	102
Figura A7.4	Jitter mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	103
Figura A7.5	Jitter médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	103
Figura A7.6	Jitter máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	103
Figura A7.7	Variação do retardo médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	104
Figura A7.8	Variação do retardo máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	104
Figura A7.9	Variação do jitter médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	105
Figura A7.10	Variação do jitter máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	105
Figura A7.11	Retardo mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	106
Figura A7.12	Retardo mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	106
Figura A7.13	Retardo máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	106
Figura A7.14	Jitter médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	107
Figura A7.15	Jitter máximo no fluxo de controle G.729B em função do carregamento	

	com fluxos VoIP empregando codecs G.711 e G.729B.....	107
Figura A7.16	Variação do retardo médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	108
Figura A7.17	Variação do retardo máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	108
Figura A7.18	Variação do jitter médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	109
Figura A7.19	Variação do jitter máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.711 e G.729B.....	109
Figura A8.1	Retardo mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	111
Figura A8.2	Retardo médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	112
Figura A8.3	Retardo máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	112
Figura A8.4	Jitter mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	113
Figura A8.5	Jitter médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	113
Figura A8.6	Jitter máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	113
Figura A8.7	Variação do retardo médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	114
Figura A8.8	Variação do retardo máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	114
Figura A8.9	Variação do jitter médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs	

	G.723.1 e G.729B.....	115
Figura A8.10	Variação do jitter máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	115
Figura A8.11	Retardo mínimo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	116
Figura A8.12	Retardo médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	116
Figura A8.13	Retardo máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	117
Figura A8.14	Jitter médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	117
Figura A8.15	Jitter máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	118
Figura A8.16	Variação do retardo médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	118
Figura A8.17	Variação do retardo máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	119
Figura A8.18	Variação do jitter médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	119
Figura A8.19	Variação do jitter máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando codecs G.723.1 e G.729B.....	120

LISTA DE TABELAS

Tabela 3.1	Dados de três <i>codecs</i> de voz padronizados pela ITU-T.....	26
Tabela 4.1	Parâmetros dos <i>codecs</i> para VoIP utilizados nas simulações.....	45
Tabela 4.2	Resumo dos valores das métricas – tráfego heterogêneo G.711/G.723.1.....	51
Tabela 4.3	Resumo dos valores das métricas – tráfego heterogêneo G.711/G.729B.....	52
Tabela 4.4	Resumo dos valores das métricas – tráfego heterogêneo G.729B/G.723.1....	54
Tabela 5.1	Coeficientes de correlação obtidos sobre os valores médios das métricas retardo e jitter do fluxo de controle em relação aos demais fluxos do agregado.....	57
Tabela A1.1	Distribuição Exponencial.....	63
Tabela A1.2	Distribuição Gamma.....	64
Tabela A1.3	Distribuição Lognormal.....	65
Tabela A1.4	Distribuição de Pareto.....	66
Tabela A1.5	Distribuição de Weibull.....	67

LISTA DE ABREVIATURAS E SIGLAS

ACELP	- <i>Algebraic Code-Excited Linear-Prediction</i>
ADPCM	- <i>Adaptive Differential PCM</i>
AF PHB	- <i>Assured Forwarding PHB</i>
ATC	- <i>Adaptive Transform Coding</i>
BE	- <i>best effort</i>
Bm	- <i>Brownian motion</i>
CBQ	- <i>Class Based Queuing</i>
CBR	- <i>Constant Bit Rate</i>
CDF	- <i>Cumulative Distribution Function</i>
CELP	- <i>Code Excited Linear Prediction</i>
CL	- <i>Controlled Load Service</i>
CNG	- <i>Comfort Noise Generator</i>
codec	- <i>codifier and decodifier</i>
CoS	- <i>Class of Service</i>
cRTP	- <i>compressed RTP</i>
CS-ACELP	- <i>Conjugate Structure ACELP</i>
DiffServ	- <i>Differentiated Services</i>
DSCP	- <i>Differentiated Services Code Point</i>
DSI	- <i>Digital Speech Interpolation</i>
DSP MIPS	- <i>Digital Signal Processor – Millions of Instructions Per Second</i>
DTX	- <i>Discontinuous Transmission</i>
EF PHB	- <i>Expedited Forwarding PHB</i>
ETSI	- <i>European Telecommunications Standards Institute</i>
F-ARIMA	- <i>Fractional Autoregressive Integrated Moving Average</i>
fBm	- <i>fractional Brownian motion</i>
FCFS	- <i>First Come First Served</i>
FTP	- <i>File Transfer Protocol</i>
GPRS	- <i>General Packet Radio Service</i>
GPS	- <i>Global Positioning System</i>
GS	- <i>Guaranteed Service</i>

GT	- <i>Gateway</i>
HTTP	- <i>Hyper Text Transport Protocol</i>
IETF	- <i>Internet Engineering Task Force</i>
IntServ	- <i>Integrated Services</i>
IP	- <i>Internet Protocol</i>
IPPM	- <i>IP Performance Metrics</i>
ISDN	- <i>Integrated Service Digital Network</i>
ISP	- <i>Internet Service Provider</i>
ITU-T	- <i>International Telecommunication Union – Telecom. Standardization Sector</i>
LAN	- <i>Local Area Network</i>
LD-ACELP	- <i>Low Delay ACELP</i>
LPC	- <i>Linear Predictive Coding</i>
LRD	- <i>long-range dependence</i>
MNB	- <i>Measuring Normalizing Blocks</i>
MOS	- <i>Mean Opinion Score</i>
MP-MLQ	- <i>MultiPulse, MultiLevel Quantization</i>
NP	- <i>network performance</i>
NS-2	- <i>Network Simulator, version 2</i>
NTP	- <i>Network Time Protocol</i>
PAMS	- <i>Perceptual Analysis Measurements Systems</i>
PCM	- <i>Pulse Code Modulation</i>
PDF	- <i>Probability Distribution Function</i>
pdf	- <i>probability density function</i>
PHB	- <i>per-hop behavior</i>
pmf	- <i>probability mass function</i>
PPP	- <i>Point-to-Point Protocol</i>
PQ	- <i>Priority Queuing</i>
PESQ	- <i>Perceptual Evaluation of Speech Quality</i>
PSQM	- <i>Perceptual Speech Quality Measure</i>
PSTN	- <i>Public Switched Telephone Network</i>
QoS	- <i>Quality of Service</i>
RAM	- <i>Randon Access Memory</i>
RFC	- <i>Request for Comment</i>

RMD	- <i>Random Midpoint Displacement</i>
RTP	- <i>Real-time Transport Protocol</i>
SBC	- <i>Sub-band Coding</i>
SID	- <i>Silence Insertion Descriptor</i>
SLA	- <i>Service Level Agreement</i>
SLS	- <i>Service Level Specification</i>
SP	- <i>Strict Priority</i>
SRD	- <i>short-range dependence</i>
TASI	- <i>Time Assigned Speech Interpolation</i>
TCP	- <i>Transmission Control Protocol</i>
TELR	- <i>Talker Echo Loudness Rating</i>
TIPHON	- <i>Telecommunication an IP Harmonization Over Networks</i>
TCA	- <i>Traffic Conditioning Agreement</i>
TCS	- <i>Traffic Conditioning Specification</i>
UDP	- <i>User Datagram Protocol</i>
VAD	- <i>Voice Activity Detector</i>
VBR	- <i>Variable Bit Rate</i>
Vocoder	- <i>VOice CODER</i>
VoIP	- <i>Voice over IP</i>
VPN	- <i>Virtual Private Network</i>
WAN	- <i>Wide Área Network</i>
WFQ	- <i>Weighted Fair Queuing</i>
WRR	- <i>Weighted Round Robin scheduling</i>
www	- <i>World Wide Web</i>
%GoB	- <i>% Good or Better</i>
%PoW	- <i>% Poor or Worse</i>

RESUMO

O presente trabalho apresenta uma análise de viabilidade do emprego de um fluxo de controle sintético VoIP para inferir sobre a performance de fluxos individuais de um fluxo agregado pertencente a um EF PHB em uma rede de serviços diferenciados. A abordagem proposta visa estabelecer através de simples verificação de performance quanto ao atendimento do SLA relacionado a alguns requisitos do fluxo VoIP. Os resultados poderão ser utilizados para alimentar especificações e requisitos para o projeto de ferramentas, por exemplo, para capacitar atividades de planejamento e ações de gerência de rede. O tráfego VoIP foi classificado como homogêneo (todos os pacotes do fluxo são criados pelo mesmo tipo de codec) e como heterogêneo (pacotes originados por mais de um tipo de *codec*) durante a realização dos experimentos. Os experimentos verificaram a hipótese de que a performance do fluxo de controle possa ser relacionada de alguma forma com a performance dos fluxos individuais de um fluxo agregado sob as suposições e métricas definidas. As métricas retardo, *jitter* e perda de pacotes foram estimadas por simulação tanto para o tráfego homogêneo quanto para o tráfego heterogêneo, em diversas condições de carga controlada. Os resultados permitem concluir quanto a viabilidade da abordagem para estimar o retardo e com limitações de confiança, quanto ao *jitter*, dependendo do tipo de tráfego (heterogêneo) e tipo de *codec*.

ABSTRACT

This work presents a viability analysis of the use of a synthetic VoIP control flow to infer about the performance of individual flows of a flow aggregate belonging to an EF PHB in a DiffServ network. The proposed approach aims to establish a simple performance verification of SLA accomplishment related to the some of the VoIP flow requirements. The results should be used to feed requirements specifications for tool design, for example, to capacity planning activities and management actions. We classify the VoIP traffic as homogeneous (all flow packets created by a same codec type) and heterogeneous (packets originated from more than one codec type) to carry out the experiments. The experiments checked the hypothesis that the control flow performance can be somehow related to the performance of individual flows of a flow aggregate under the agreed assumptions and metrics. The metrics one-way delay, jitter and packet loss were evaluated by simulation for both homogeneous and heterogeneous traffic at several network-controlled load. The results let us conclude about the viability of the approach to evaluate one-way delay and with confidence limitations, also the jitter, depending on the traffic type (heterogeneous) and codec type.

CAPÍTULO 1

INTRODUÇÃO

Atualmente é aceito que as redes comutadas a circuito e as redes comutadas a pacote estão gradualmente convergindo para uma infra-estrutura baseada no Protocolo Internet (*Internet Protocol* - IP). Essa infra-estrutura transportará tanto o tráfego da rede telefônica pública comutada (*Public Switched Telephone Network* – PSTN) quanto o tráfego das aplicações da Internet. Esse cenário tem grande apelo na redução de custos, através da consolidação da tecnologia e do crescimento da indústria, com a criação de novos serviços. Contudo, essa convergência tem sido lenta na sua materialização. Do ponto de vista técnico, a maior dificuldade tem sido a qualidade de serviço (*Quality of Service* - QoS) [SN2003], aqui entendida como *a capacidade de fornecer garantia de recursos e a diferenciação de serviços em uma rede* [WZ2001].

Nas tradicionais redes IP o paradigma de qualidade é o de melhor esforço (*best effort* - BE). Esse paradigma proporciona aos seus usuários o livre compartilhamento dos recursos disponíveis na rede, mas não assegura que qualquer nível de desempenho particular seja atingido. O paradigma BE tem sido bem sucedido quando trata com aplicações de dados em tempo não real, tais como *e-mail* e transferência de arquivos, e em aplicações multimídia próximas ao tempo real, tais como fluxos de áudio e vídeo e *Web browsing* [SN2003]. No entanto, para realizar os benefícios totais da convergência, as futuras redes baseadas em IP necessitam implementar paradigmas de compartilhamento de recursos capazes de fornecer, confiavelmente, QoS diferenciada para um grande e diverso conjunto de aplicações de usuário final, incluindo voz sobre IP (*Voice over IP* - VoIP). Esse compartilhamento de recursos deverá ser coordenado através de jurisdições independentes, denominadas multi-domínios, e em um ambiente multi-fornecedor. Portanto, o fornecimento de QoS fim-a-fim através de uma infra-estrutura heterogênea [ET2003] - especificamente a Internet, uma enorme coleção de redes operadas por diferentes organizações - é um desafio novo às operadoras tradicionais de sistemas de telefonia fixa cuja infra-estrutura física, baseada em comutação de circuitos, é bastante homogênea.

As arquiteturas de serviços integrados (*Integrated Services* - IntServ) [BR1994] e de serviços diferenciados (*Differentiated Services* – DiffServ) [BS1998] [GD2002] foram propostas pela IETF (*Internet Engineering Task Force*) para tratar de QoS. A IntServ opera

em uma base por fluxo provendo um modelo de serviço que capacita fortes garantias de QoS por fluxo mas sofre de falta de escalabilidade. Por outro lado, a DiffServ mantém a informação por fluxo somente no nó de ingresso (ou de borda ou de fronteira) de um domínio e a partir dali cria fluxos agregados em um conjunto limitado de classes de tráfego dentro da rede, resolvendo o problema de escalabilidade a custo de perda de garantia de QoS [ME2003].

Nas redes DiffServ, os roteadores processam o tráfego agregado que pertence a várias classes de serviço (*Class of Service – CoS*) conforme as políticas pré-definidas de QoS. A rede oferece vários tipos de serviço tais como: serviços em tempo real, serviços para Redes Virtuais Permanentes (*Virtual Private Network – VPN*), serviços BE, etc., onde o tráfego de cada classe de serviço apresenta requisitos específicos de QoS bem como características particulares de geração de carga para o sistema [AA2003]. Conseqüentemente, em termos da QoS desejada, o serviço oferecido deve ser quantificado e especificado previamente através de um ou mais parâmetros de tráfego e parâmetros de desempenho.

Para regular o relacionamento entre o cliente e o provedor de infra-estrutura ou o provedor de serviços IP ou entre os provedores é empregada a figura do acordo de nível de serviço (*Service Level Agreement – SLA*). Esse contrato deve conter uma clara especificação do nível de serviço (*Service Level Specification - SLS*) e a forma de verificação do seu atendimento. O termo SLA é bastante abrangente, tendo definições diferentes em função do ponto de vista, envolvendo características técnicas e parâmetros do serviço, bem como aspectos legais e financeiros.

A verificação pelo cliente do atendimento da especificação de QoS estabelecida no SLA tem motivado o desenvolvimento de ferramentas para este fim. Da mesma forma, também são necessárias ferramentas para que o provedor de serviço se assegure que pode cumprir e que esteja cumprindo os contratos assumidos em relação a QoS acordada sob pena de ter sua credibilidade e seu negócio comprometidos [AA2003].

Atualmente no mundo IP, as ferramentas de monitoração nas atuais redes de serviços BE têm principalmente um papel de diagnóstico [MF2005], não sendo projetadas para verificação de atendimento de SLAs. Elas avaliam o estado da rede, analisam o comportamento da rede durante certo período de tempo e fornecem relatórios para um sistema de gerência. A informação da medição é normalmente coletada tanto na base por fluxo de tráfego para contabilidade comercial quanto na base por enlace para diagnóstico técnico da rede.

Dada a variada gama de serviços com diferentes requisitos de QoS, a informação da medição necessita ser coletada em uma granularidade mais fina do que a do par fonte-destino

quando existir tráfego de diferentes classes de serviço, como em DiffServ, isto é, a classe de serviço necessita ser levada em conta. Conseqüentemente, uma grande quantidade de informação é necessária para monitoração do nível de serviço e para engenharia de tráfego de grandes redes IP. Também é igualmente importante haver adaptação do processo de monitoração às freqüentes mudanças no estado da rede e na configuração da rede.

Assim, uma ferramenta simples que possa dar indicação em tempo real do desempenho de uma classe DiffServ é um desafio de engenharia. Especificamente, a avaliação de desempenho refere-se aqui à verificação e comparação das principais métricas de SLA, tais como: perda de pacotes, retardo na entrega de pacotes e variação do retardo na entrega dos pacotes, em relação à especificação contratada.

Nesse contexto, o presente trabalho aborda a questão de avaliação de desempenho em tempo real de fluxos pertencentes a classe de encaminhamento expresso do comportamento por lance (*Expedited Forwarding per-hop behavior* – EF PHB) de serviços diferenciados que transporta pacotes do serviço VoIP que são multiplexados dentro de um fluxo agregado. Esta avaliação deve levar em conta a utilização dos vários tipos de codificadores de voz utilizados em VoIP [TA2000] [DT2003].

O objetivo principal deste trabalho é investigar o emprego de um fluxo colocado dentro do agregado para inferir sobre o desempenho dos demais fluxos VoIP desse agregado da classe EF, visando verificar o atendimento do SLA. Para consecução deste objetivo, propõe-se um método baseado na inserção de um fluxo de controle no agregado de fluxos VoIP. Esse fluxo, inserido no roteador de ingresso da rede DiffServ, é então multiplexado juntamente com os demais fluxos VoIP da mesma classe. No roteador de egresso da rede, o desempenho do fluxo de controle é comparado com o desempenho dos demais fluxos do agregado daquela classe. Investiga-se a partir daí como os resultados das medições das métricas realizadas neste fluxo de controle permitiriam ou não inferir sobre o desempenho dos demais fluxos do agregado. Em resumo, o trabalho busca contribuir com o estabelecimento de dois métodos:

- um método baseado na inserção de um fluxo de controle no agregado de uma classe e que servirá de referência quanto ao desempenho dos demais fluxos do agregado.
- um método de análise de parâmetros de desempenho do fluxo de controle que permita orientar a avaliação de atendimento de SLA para tráfego VoIP.

Considerando que, na prática, para obtenção de VoIP, é possível o emprego de vários tipos de codificadores de voz que funcionam de forma diferente, gerando taxas de transmissão diferentes e tamanhos de pacotes diferentes, é necessário entender o funcionamento e as características técnicas desses dispositivos. O tráfego VoIP é aqui caracterizado em dois tipos: homogêneo e não homogêneo. Na condição de tráfego homogêneo, todos os fluxos VoIP têm a mesma característica, ou seja, empregam o mesmo tipo de codificador. No tráfego não homogêneo, ocorre a combinação de fluxos VoIP gerados por codificadores diferentes. Portanto, o fluxo de controle deve operar nessas duas condições de tráfego e deve ser capaz de fornecer informações sobre o desempenho de todos os fluxos VoIP do agregado.

As inferências sobre o desempenho dos demais fluxos do agregado a partir do desempenho do fluxo de controle, foram constatadas através da análise dos resultados obtidos em simulações [OPNET]. Os ambientes estabelecidos para os experimentos procuraram refletir os diferentes cenários envolvendo redes DiffServ, tráfego VoIP homogêneo e não homogêneo.

O desenvolvimento do trabalho requereu um estudo sobre a modelagem das fontes de voz levando em consideração os codificadores e a superposição probabilística dessas fontes através do emprego de funções de distribuição representativas do comportamento do tráfego de pacotes VoIP [AA1997] [JW2000] [CC2000] [BB2002] [DT2003]. Os resultados deste estudo possibilitam configurar e analisar o impacto dos diferentes parâmetros das fontes de tráfego VoIP sobre os parâmetros de desempenho. O trabalho de experimentação com os modelos estudados e os cenários escolhidos conduz a conclusões interessantes sobre aspectos de modelagem de tráfego IP e que representam novas contribuições sobre este tema.

A estrutura da dissertação é a seguinte: no Capítulo 2 são apresentados os conceitos básicos necessários para o desenvolvimento e compreensão do trabalho tais como: arquitetura de serviços diferenciados, qualidade de serviço, acordo de nível de serviço, parâmetros de qualidade de serviço e objetivos de desempenho, classe de serviço, e medições ativas; no Capítulo 3 são apresentados os conceitos de VoIP, são descritas as características dos codificadores e decodificadores de voz, são apresentados os conceitos envolvidos na modelagem de tráfego VoIP e no final são apresentados os trabalhos relacionados; no Capítulo 4 é apresentada a contribuição deste trabalho, um modelo para monitoração do desempenho de fluxos VoIP através de um fluxo de controle injetado quando da formação do fluxo agregado e também são apresentados os cenários e os resultados obtidos quando da simulação do problema proposto e no Capítulo 5 são apresentadas as conclusões obtidas e é comentada a validação do modelo proposto.

CAPÍTULO 2

FUNDAMENTOS TEÓRICOS

A terminologia associada com QoS em redes IP pode ser enquadrada em três grupos relacionados com as arquiteturas para QoS, com a definição de QoS e de CoS; e com a especificação de um contrato entre o usuário e o provedor de serviço, denominado SLA.

Este capítulo está organizado da seguinte maneira: na seção 2.1, são apresentadas as informações básicas sobre a arquitetura de Serviços Diferenciados; na seção 2.2, é descrita a conceituação de QoS para redes IP conforme os pontos de vista dos órgãos de padronização internacionais; na seção 2.3 são descritas as partes constituintes do SLA; na seção 2.4 são apresentados os parâmetros de QoS e os objetivos de desempenho de rede IP; na seção 2.5 são tratadas as classes de serviço IP na arquitetura DiffServ; e na seção 2.6, são descritos os cuidados na medição de fluxos IP sintéticos em um ambiente DiffServ.

2.1 ARQUITETURA DE SERVIÇOS DIFERENCIADOS (DiffServ)

Um passo em direção a QoS nas redes IP é a capacitação dessas redes para diferenciar serviços. O paradigma DiffServ estabelece uma estrutura para obtenção dessa diferenciação através do manuseio do tráfego dentro de uma plataforma IP. Essa diferenciação se dá com o tratamento de pacotes baseado em classes de serviço conforme codificado no cabeçalho do pacote IP. O provedor de serviço estabelece com cada usuário um SLA, o qual, entre outras coisas, especifica quanto tráfego o usuário pode enviar dentro de qualquer classe de serviço dada. O tráfego é então policiado nas fronteiras (bordas) da rede do provedor de serviço. Uma vez que o tráfego tenha entrado na rede, os roteadores fornecem para esse tráfego um tratamento diferenciado por classe de serviço. Os blocos relevantes para a abordagem DiffServ incluem: gerenciamento de *buffer*, marcação de pacotes, medição e registro de tráfego, policiamento e formatação de tráfego e ordenação de pacotes [LH2003]. A garantia de desempenho nas redes DiffServ é fornecida por uma combinação de provisionamento de recursos, priorização de tráfego e controle de admissão.

A diretriz básica da arquitetura DiffServ é alcançar escalabilidade no núcleo (*core*) da rede. Um número limitado de serviços fornecidos pela rede e uma arquitetura simplificada de

nós no núcleo são os fatores chave da escalabilidade da rede DiffServ. Nenhuma informação de estado sobre os fluxos é mantida no núcleo da rede.

Na arquitetura DiffServ, o tráfego entrante na rede é classificado e condicionado somente nas fronteiras da rede e designado a diferentes agregados de comportamento [BS1998] [GD2002]. Um agregado de comportamento DiffServ é uma coleção de pacotes com o mesmo DSCP (*Differentiated Services Code Point*), atravessando um enlace em uma direção particular. Por consequência, a garantia de QoS é aplicada somente em uma direção da transmissão de dados.

Os fluxos em uma rede DiffServ são encaminhados para a classe de serviço pré-definida, sendo servidos da mesma forma como os demais fluxos que tem o mesmo DSCP [GJ2003]. Os fluxos (ou pacotes) servidos pela mesma classe de serviço são agregados e experimentam o nível de QoS dado pelo agregado. O processamento dos pacotes de agregado por um nó da rede é denominado PHB (*per-hop behavior*), o qual é definido como *comportamento de encaminhamento externamente observável, aplicado em um nó aderente a arquitetura de serviços diferenciados para um agregado DiffServ* [BS1998] [GD2002]. Um PHB pode ser especificado em termos dos recursos de rede (isto é, *buffer*, banda), prioritariamente em relação a outros PHBs, ou às suas características de tráfego observáveis (isto é, retardo, perda).

Foram padronizados dois tipos de PHB, denominados: encaminhamento garantido (*Assured Forwarding PHB – AF PHB*) e encaminhamento expresso (*Expedited Forwarding PHB – EF PHB*). O AF PHB, especificado na RFC 2597 [HJ1999] [GD2002], tem quatro classes de encaminhamento, cada uma com sua própria alocação de banda. Dentro de cada classe de encaminhamento AF, existem três níveis de prioridade de descarte. O EF PHB, especificado na RFC 2598 [JV1999] [AG2002], provê um tratamento de encaminhamento de retardo mínimo e perda mínima. Ao ser utilizado o encaminhamento EF a taxa de saída dos pacotes em qualquer nó deve ser limitada, menor do que a taxa de saída do roteador.

Conforme [GJ2003], os PHBs na rede DiffServ não são estritamente definidos e podem ser diferentes de domínio para domínio DiffServ. Um domínio DiffServ é um conjunto contíguo de nós DiffServ que tem implementado o mesmo mecanismo PHB e que opera com um conjunto de políticas de provisionamento de serviço comuns. Todos os nós em um domínio DiffServ servem os fluxos de dados da mesma forma, dependendo da CoS que o agregado atende. Um domínio DiffServ tem uma fronteira bem definida consistindo de nós de borda DiffServ que classificam e possibilitam a condição do tráfego de ingresso. Uma região DiffServ é um conjunto de um ou mais domínios DiffServ contíguos em cooperação.

As funções que os nós de borda realizam são geralmente referidas como classificação de tráfego e condicionamento de tráfego, conforme visto na Figura 2.1 [WZ2001].

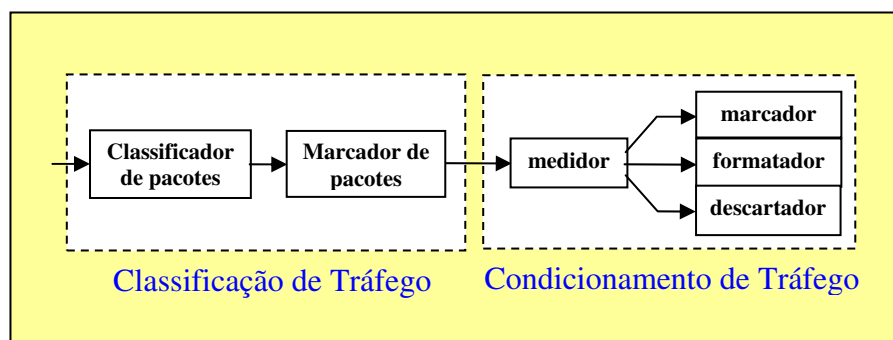


Figura 2.1: Classificação e condicionamento de tráfego em DiffServ [WZ2001]

Há um bom número de mecanismos de ordenação de fila que podem ser utilizados nos roteadores de borda, tais como: fila prioritária – (*Priority Queuing* – PQ), com a prioridade mais alta e um policiador de taxa; ordenador com tabela de pesos (*Weighted Round Robin scheduling* – WRR) e fila baseada em classe (*Class Based Queuing* – CBQ) [WZ2001]. Não é abordado neste trabalho o impacto do uso desses mecanismos no desempenho fim-a-fim do fluxo agregado.

A arquitetura DiffServ focaliza o gerenciamento de recursos em um domínio simples. A questão do gerenciamento de recursos fim-a-fim não está resolvida – os usuários são capazes de perceber uma melhora nos serviços somente quando existe suporte de QoS em todos os caminhos entre emissor e receptor. Atualmente não existe consenso em como obter gerenciamento de recursos fim-a-fim [WZ2001].

2.2 QUALIDADE DE SERVIÇO PARA REDES IP

Conforme [GJ2003], um serviço em ambiente IP, ou serviço baseado em IP, é definido pela ITU-T (*International Telecommunication Union – Telecommunications Standardization Sector*) como *um serviço fornecido pelo plano de serviço para um usuário final (isto é, um host – sistema terminal – ou um elemento de rede) e que utiliza as capacitações de transferência IP e funções de controle e de gerenciamento associadas, para entregar a informação do usuário com a qualidade especificada pelo acordo de nível de serviço - SLA* [Rec. Y.1241].

Na visão da ETSI (*European Telecommunications Standards Institute*), ITU e IETF, existem três noções definidas de QoS: a intrínseca, a percebida e a avaliada, que constituem o modelo geral mostrado na Figura 2.2 [GJ2003].

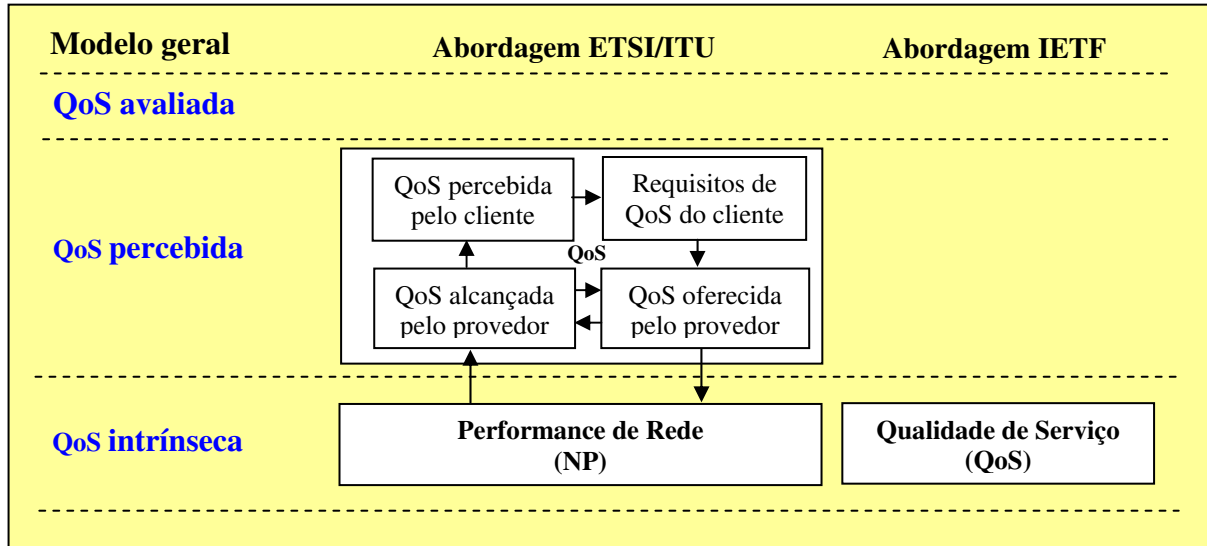


Figura 2.2: O modelo geral de QoS com as abordagens da ETSI/ITU e da IETF [GJ2003].

As características do serviço descritas sob a forma de aspectos técnicos pertencem à noção de QoS intrínseca, que é determinada pelo projeto da rede de transporte e pelo provisionamento da rede de acesso, terminações e conexões [HW2001]. A QoS intrínseca é obtida pela comparação das características de desempenho medida e esperada. A percepção do usuário do serviço não influencia no *rating* da QoS intrínseca. Entre outras coisas, a QoS requerida é alcançada pela seleção apropriada dos protocolos de transporte, dos mecanismos de garantia de QoS e dos valores relativos dos parâmetros.

A QoS percebida reflete a experiência do usuário empregando um serviço particular. Esta QoS percebida é influenciada pelas expectativas do usuário quando ele compara o desempenho observado do serviço. Por sua vez, as expectativas pessoais são usualmente afetadas pela experiência do usuário com um serviço similar prestado por outro provedor e pela opinião de outros usuários. Assim, a QoS com as mesmas características intrínsecas pode ser percebida diferentemente por vários usuários. Disso segue que, mesmo assegurando parâmetros de serviço (rede) particulares pode não ser suficiente para satisfazer usuários que não sabem como o serviço é fornecido.

A QoS oferecida por um provedor deve refletir a QoS intrínseca mesmo com alguns parâmetros não técnicos que são ótimos para o usuário e relevantes para um conjunto particular de expectativas.

A QoS avaliada começa a ser notada quando o usuário decide se continua utilizando o serviço ou não. Essa decisão depende da qualidade percebida, preço do serviço e respostas do provedor a queixas e problemas submetidos. Disso segue que, mesmo uma atitude no serviço de atendimento a cliente pode ser um importante fator na obtenção de pontos no *rating* da QoS avaliada [GJ2003].

A declaração de um nível satisfatório para QoS intrínseca, para QoS percebida e para QoS avaliada pode ser considerada separadamente. A primeira declaração é de responsabilidade do provedor de rede e depende da arquitetura da rede, do planejamento e do gerenciamento. É principalmente um problema técnico tratado por engenheiros, projetistas e operadores.

O uso apropriado das capacitações da QoS intrínseca ajustadas para um serviço particular oferecido, juntamente com análises de mercado, são necessárias para assegurar um alto nível na QoS percebida. Esta obrigação é do provedor de serviço e esforços de propaganda e de marketing têm impacto positivo na QoS percebida [GJ2003].

A ETSI e a ITU-T [Rec. E.800] [Rec. G.1000] [Rec. I.350] [Rec. Y.1541] abordam a terminologia relacionada com QoS de forma quase igual. A definição de QoS dada na Recomendação G.800 é '*o efeito coletivo do desempenho do serviço que determina o grau de satisfação de um usuário do serviço*'. Como a definição acima sugere, a abordagem da ETSI/ITU é mais aderente ao conceito da QoS percebida do que ao conceito da QoS intrínseca. Diante disso, a ITU-T introduziu a noção de desempenho de rede (*network performance* - NP) para cobrir os aspectos técnicos. Então, é feita uma clara distinção entre QoS, entendida como algo focado nos efeitos percebíveis pelo usuário, e NP, englobando todas as funções de rede essenciais para prover o serviço. Conforme a Recomendação I.350 da ITU-T, os parâmetros de QoS são orientados ao usuário e não são diretamente transladados em parâmetros de rede. Por outro lado, os parâmetros de NP determinam a QoS observada pelos usuários mas que não são necessariamente significativos para eles. Deve existir um mapeamento consistente entre QoS e os parâmetros de NP. As relações entre QoS e NP são também mostradas no modelo geral dado na Figura 2.2 [GJ2003].

A NP, como acima mencionado, corresponde a QoS intrínseca. Esse aspecto está definido na Recomendação E.800 como *a habilidade de uma rede ou porção de rede em prover as funções relacionadas às comunicações entre usuários*. A NP é definida e tratada em

termos de parâmetros de elementos de rede particulares envolvidos no fornecimento de um serviço. Esses parâmetros são a chave para a eficiência e a eficácia da rede no fornecimento de um serviço [GJ2003].

Para cobrir os vários pontos de vista de QoS por parte do usuário e por parte do provedor, a ITU-T e a ETSI distinguem quatro definições particulares, que são [GJ2003]: requisitos de QoS do usuário; QoS oferecida pelo provedor; QoS alcançada pelo provedor e QoS percebida pelo usuário.

Por outro lado, a IETF focaliza a QoS intrínseca e não trabalha com a QoS percebida. O conceito de QoS é entendido pela IETF como sendo *um conjunto de requisitos de serviço a serem satisfeitos pela rede enquanto transportando um fluxo* [CE1998]. Esta definição é quase equivalente à noção de NP definida pela ITU-T/ETSI, sendo definida em termos de parâmetros (métricas).

Adicionalmente, a QoS intrínseca pode ter os seguintes atributos dependendo da arquitetura de rede tanto quanto das demandas da aplicação:

- I. Fim-a-fim (no modelo IntServ) ou limitado a um domínio particular ou limitado a vários domínios (no modelo DiffServ);
- II. Aplicado em todo o tráfego ou somente em uma sessão particular ou sessões;
- III. Unidirecional ou bidirecional;
- IV. Garantido ou estatístico.

Diante do exposto, o presente trabalho enfoca o conceito de QoS conforme a visão da IETF, ou seja, uma vez acordado com o usuário (ou cliente) um conjunto de requisitos de serviços, a rede (provedor) deve atender tais requisitos enquanto estiver transportando os fluxos objeto desse acordo, limitado ao domínio DiffServ. Portanto, agora é oportuno introduzir o conceito de SLA.

2.3 ACORDO DE NÍVEL DE SERVIÇO (SLA)

Um conjunto separado de conceitos é utilizado para descrever o contrato entre um usuário e um provedor de serviço ou operador de rede. Conforme [GJ2003], o termo mais geral utilizado na denominação desse contrato é SLA. As partes de um SLA são: a SLS, o acordo de condicionamento de tráfego (*Traffic Conditioning Agreement – TCA*) e a especificação do condicionamento de tráfego (*Traffic Conditioning Specification – TCS*).

A noção de SLS, foi introduzida para separar a parte técnica do corpo do SLA. A SLS é definida como *um conjunto de parâmetros e seus valores que juntos definem o serviço oferecido para um dado tráfego* [GD2002]. Em outras palavras, a SLS especifica um conjunto de valores dos parâmetros de rede relacionados a um serviço em particular. Os serviços de transporte IP são tecnicamente descritos pela SLS. Estão em progresso trabalhos direcionados para especificar o conjunto de parâmetros básicos que comporá o conteúdo da SLS [GJ2003].

Conforme a definição da ITU-T na Recomendação Y.1241, o SLA, é *um acordo negociado entre um cliente e o provedor de serviço sobre os níveis das características do serviço e o conjunto associado de métricas. O conteúdo do SLA varia dependendo do serviço oferecido e inclui os atributos requeridos para o acordo negociado.* O SLA pode ser colocado na forma de um documento contendo os nomes das partes que assinam o contrato, devendo ser composto pelos objetivos de nível de serviço, pelos componentes de monitoração do serviço e pelos componentes de compensação financeira. Os objetivos do nível de serviço englobam os parâmetros de QoS da CoS provida, a disponibilidade do serviço, os objetivos de autenticação, a data de término do SLA, e assim por diante. A monitoração do serviço especifica a forma de medição da QoS e as métricas utilizadas para apontar se o serviço está em conformidade com o SLA. Esse documento pode incluir também um acordo quanto à forma e frequência de entrega de relatórios referentes ao uso e desempenho do serviço. A componente financeira pode incluir as opções de cobrança, penalidades por quebra do contrato, e assim por diante [Rec. Y.1241].

A IETF define o SLA de forma similar, como sendo *um contrato de serviço entre um cliente e um provedor de serviço que especifica o serviço que um cliente deverá receber* [BS1998] [GD2002]. Um SLA deve ser expresso de forma inteligível para o cliente. Ele engloba as características básicas do serviço e critérios não ambíguos e bem definidos de verificação se o serviço entregue é consistente com o contrato. Por outro lado, os limites impostos ao cliente devem ser claros. Um SLA deve consistir de regras de responsabilidade para quebra do contrato pelo provedor de serviço tanto quanto pelo cliente. Com relação à definição da IETF, o SLA também pode incluir regras de condicionamento de tráfego que, em parte, constituem o TCA.

Os SLAs podem ser estáticos ou dinâmicos. A maioria dos SLAs é negociada entre usuários e provedores de serviço antes dos serviços serem iniciados, com mudanças ou revisões ocasionais de tempos em tempos. Os SLAs podem também ser dinamicamente alterados e negociados. Por exemplo, um usuário pode querer pedir banda adicional para eventos especiais ou para fluxos de tráfego especiais. Tal SLA dinâmico requer sistemas de

gerenciamento que possam automatizar essas mudanças e incluir negociações baseadas em políticas acordadas e em regras estabelecidas entre as partes [WZ2001].

Um TCA especifica as regras de classificação de pacotes e de perfis de tráfego como uma descrição das propriedades temporais do fluxo de tráfego, tais como a taxa de rajada e o tamanho da rajada. O cliente é obrigado a ajustar o fluxo do tráfego gerado ao perfil contratado. Para forçar que o tráfego do cliente seja aderente a um perfil particular são definidas regras para medição, para marcação, para descarte e para formatação desse tráfego. O tratamento dos pacotes fora do perfil é também especificado no TCA. Adicionalmente, de acordo com a definição da IETF, *o TCA engloba todas as regras de condicionamento de tráfego explicitamente especificadas dentro do SLA e todas as regras implícitas dos requisitos relevantes do serviço e/ou da política de provisionamento de serviço no domínio DiffServ* [BS1998] [GD2002].

A TCS é um conjunto de parâmetros com valores designados que, de forma não ambígua, especifica um conjunto de regras de classificação e um perfil de tráfego. A TCS é a parte técnica do TCA. A TCS é também um elemento integrante do SLS [GD2002].

As relações entre SLA, SLS, TCA e TCS são mostradas de forma simplificada na Figura 2.3 [GJ2003].

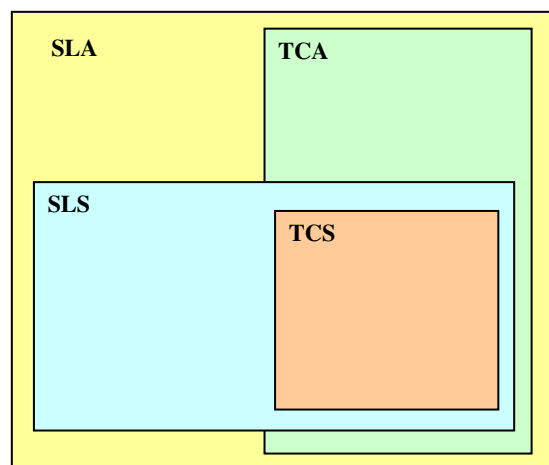


Figura 2.3: Relações entre SLA, SLS, TCA e TCS [GJ2003].

Assim, após ter sido descrita a parte técnica constituinte do SLA, formada pela TCS (que especifica as regras de classificação de tráfego, o perfil de tráfego do cliente e as regras de descarte de pacotes), e pela SLS (que especifica o conjunto de valores dos parâmetros de rede relacionados a um serviço em particular), é oportuno descrever os parâmetros de QoS e os objetivos de desempenho de rede IP que normalmente fazem parte da SLS.

2.4 PARÂMETROS DE QoS E OBJETIVOS DE DESEMPENHO DE REDE IP

Conforme [GJ2003], empregando os conceitos da IETF, a QoS intrínseca em redes de pacotes é expressa ao menos pelo seguinte conjunto de parâmetros que são significativos para a maioria dos serviços baseados em IP:

- i. Taxa de bit de transferência de dados (*throughput*) disponível para o usuário no serviço em questão ou objetivo de vazão de dados que pode ser alcançado;
- ii. Retardo unidirecional (*one-way delay*) experimentado pelos pacotes enquanto estão passando através da rede. Pode ser considerado em uma relação fim-a-fim ou com relação a um elemento de rede particular.
- iii. Variações do retardo (*jitter*) de transferência de pacotes IP. Pode ser considerado em uma relação fim-a-fim ou com relação a um elemento de rede particular.
- iv. Taxa de perda de pacotes (*packet loss*), usualmente definida como a relação entre o número de pacotes não entregues e o total de pacotes transmitidos.

A QoS é usualmente uma característica fim-a-fim da comunicação entre *hosts*.

A QoS pode ser assegurada ao longo de todo o trajeto entre parceiros (*peers*), podendo o trajeto passar por vários sistemas autônomos pertencentes a vários provedores de rede. Então o desempenho de todos os sistemas autônomos contribui para a qualidade do serviço final. No presente trabalho é considerado um único domínio DiffServ, tendo em vista o emprego de um fluxo de controle sintético inserido no nó de ingresso da rede e retirado do fluxo agregado no nó de egresso do domínio para comparação com o comportamento dos demais fluxos constituintes do agregado, de forma que seja possível obter informações gerenciais, em tempo real, para atuação imediata na rede, caso necessário.

Na Recomendação G.1010 da ITU-T são apresentados os fatores chave (retardo, variação do retardo e perda de pacotes) que influenciam a QoS, vistos pela perspectiva do usuário final, para diversas aplicações. A Figura 2.4, apresenta oito agrupamentos de aplicações conforme descrito na recomendação supracitada. Dentro desses agrupamentos existe uma segregação primária no eixo vertical entre aplicações que podem tolerar alguma perda de informação e aplicações que não toleram qualquer perda de informação, e uma outra segregação realizada no eixo horizontal onde são mostradas quatro faixas de tolerância de retardo, desde milissegundos até segundos. Deve ser observado que estes resultados independem da tecnologia de rede empregada [Rec. G.1010].

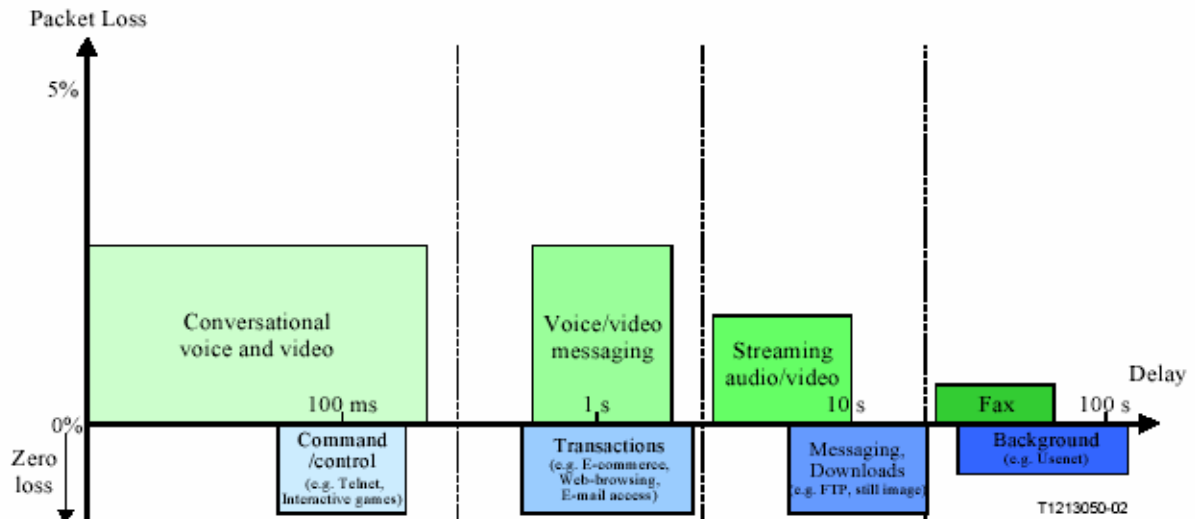


Figura 2.4: Requisitos de QoS centrados no usuário [Rec. G.1010]

Uma conexão VoIP pode ser realizada em três cenários básicos onde, em um lado está o conjunto de fones de ouvido e microfone adaptado à cabeça do usuário (*headset*), instalado em um computador pessoal: (a) sem a presença de terminal PSTN ou terminal da rede digital de serviços integrados (*Integrated Services Digital Network - ISDN*), (b) com um terminal ISDN empregando um GT (*Gateway*) e (c) com um terminal analógico ligado na PSTN.

Os fatores mais significativos da degradação da qualidade da voz sobre IP são: perda de pacotes, atraso e *jitter* produzidos na rede IP: o atraso fim-a-fim e a perda de quadros produzidos nos *buffers* de *jitter*; o eco acústico produzido nas interfaces acústicas (microfone e alto-falante); o eco elétrico produzido nos pontos de transformação de 2 para 4 fios dos circuitos telefônicos [HO2002].

O grau de insatisfação devido ao eco de locutor em uma conversação depende tanto da quantidade de retardo (atraso) quanto da diferença de nível entre os sinais de voz e de eco sentidos pelo usuário na recepção. Essa diferença de nível entre os sinais é caracterizada pela medida de volume de eco de locutor (*Talker Echo Loudness Rating – TELR*) como descrito na Recomendação G.122 e no anexo A da Recomendação G.111 da ITU-T.

A Recomendação G.131 da ITU-T indica os requisitos mínimos de TELR como função do tempo médio de transmissão unidirecional, T , conforme visto na Figura 2.5.

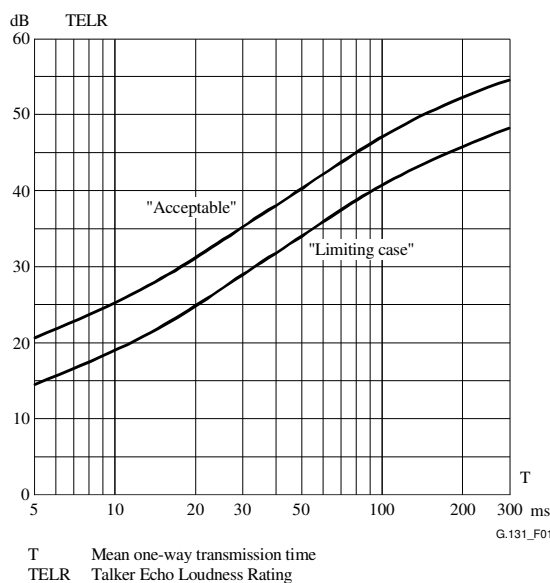


Figura 2.5: Curvas de tolerância para o eco do locutor [Rec. G.131 ITU-T]

Na prática, o valor aceitável (*acceptable*) é de 1%, que significa que, em média, 1% dos usuários irão se queixar do eco percebido na ligação. A curva para 10% é um limite extremo (*limiting case*), que deve ser permitido apenas em casos excepcionais. Analisando a Figura 2.5 é constatado que o eco torna-se mais audível à medida que o atraso aumenta.

Quando há um retardo elevado na conversação, o locutor tende a pensar que o ouvinte não escutou ou não prestou atenção no que foi dito e conseqüentemente tende a repetir o que falou, podendo ser interrompido pela resposta atrasada, proveniente do lado de destino da ligação. Ambos podem parar de falar – e recomeçarem simultaneamente. Esse procedimento prejudica a interatividade da comunicação.

Geralmente um retardo unidirecional abaixo de 100 ms fornece uma boa interatividade e um retardo entre 100 e 250 ms fornece uma interatividade aceitável (ligação via satélite).

A ETSI, através do TIPHON (*Telecommunication and IP Harmonization Over Networks*), desenvolveu o modelo-E para atender às necessidades de planejamento de redes. Esse modelo permite que a qualidade subjetiva de uma conversação seja avaliada como ela seria percebida por um usuário. O modelo-E associa cada fator de degradação na qualidade da voz percebida a um valor chamado de fator de prejuízo. Os fatores de prejuízo são então processados pelo modelo-E, que fornece como saída uma avaliação R entre 0 e 100. O valor R pode ser mapeado em um valor denominado índice de opinião médio (*Mean Opinion Score – MOS*) ou em valores percentuais bom ou melhor (*% Good or Better – %GoB*) ou ruim ou pior (*% Poor or Worse – %PoW*) utilizando tabelas. Um valor R = 50% é muito ruim, enquanto um valor R = 90% é muito bom.

O modelo-E leva em conta parâmetros que não são considerados na Recomendação G.131 da ITU-T, tais como a qualidade do codificador de voz. A maioria dos codificadores de voz foi avaliada com relação ao fator de prejuízo sem perda de quadros e, conseqüentemente, o modelo-E pode ser utilizado para avaliar a qualidade da voz percebida através de uma rede de telefonia IP sem perda de pacotes e *jitter* baixo.

A conexão de telefonia IP está sujeita à mesma relação eco/retardo como qualquer outra conexão telefônica. Mesmo com um codificador de qualidade telefônica (*toll quality*), os tipos de retardos unidirecionais encontrados na telefonia IP exigem TLR ≥ 55 dB. Para conversação com fala simultânea o cancelamento de eco pode ser reduzido a cerca de 30 dB.

Na Recomendação G.114 da ITU-T é analisado o efeito do retardo unidirecional fim-a-fim em aplicações de voz empregando o modelo-E e expressando em termos do parâmetro R. Sob o ponto de vista de planejamento de rede, é recomendado que o retardo seja mantido o menor possível e um limite aceitável seria de 150 ms quando então a maioria das aplicações não seria afetada significativamente. Um limite superior de 400 ms já causaria algum grau de insatisfação, como visto na Figura 2.6 [Rec. G.114].

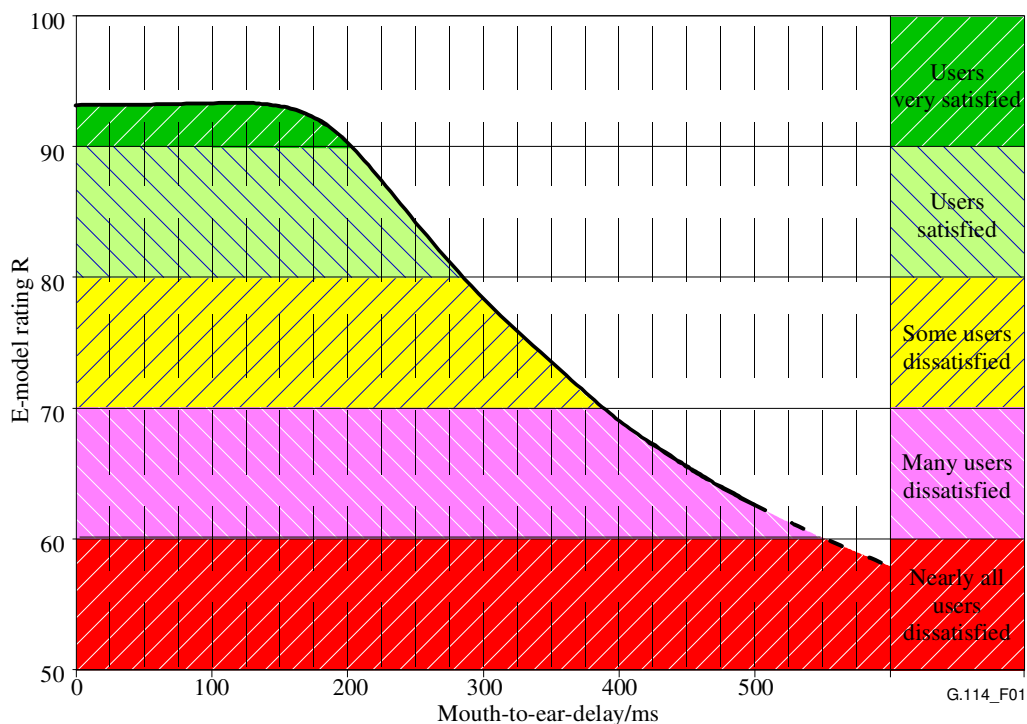


Figura 2.6: Satisfação dos usuários quanto ao limite de retardo absoluto [Rec. G.114]

Os processos de medição da qualidade da voz mais comumente empregados são: MOS, PSQM (*Perceptual Speech Quality Measure*), PESQ (*Perceptual Evaluation of Speech*

Quality), MNB (*Measuring Normalizing Blocks*), PAMS (*Perceptual Analysis Measurements Systems*) e E-model.

2.5 CLASSES DE SERVIÇO IP DENTRO DA ARQUITETURA DiffServ

A CoS é um termo abrangente que descreve um conjunto de características disponíveis em um serviço específico. Para a IETF, a CoS é definida como *as definições da semântica e dos parâmetros de um tipo específico de QoS* [CE1998]. Os serviços pertencentes à mesma classe são descritos pelo mesmo conjunto de parâmetros, que podem ser valores quantitativos ou qualitativos.

Atualmente, classes de serviço concretas são definidas dentro das arquiteturas de QoS IP propostas pela IETF, tal como IntServ e DiffServ. As três classes definidas dentro da arquitetura IntServ são: *Guaranteed Service* (GS), *Controlled Load Service* (CL) e *Best Effort* (BE). Para DiffServ foram também definidas três classes: *Premium*, *Olympic* e *Best Effort*, mas esta classificação tem apenas importância histórica [GJ2003].

Na DiffServ, cada PHB tem suas próprias características e entre os PHB padronizados estão o AF, EF e BE [WZ2001]. A abordagem DiffServ utiliza o conceito de fluxos agregados, que são encaminhados utilizando o mesmo PHB. Por exemplo, os pacotes VoIP de um cliente são agregados com os pacotes VoIP de outros clientes enquanto servidos pelo mesmo EF PHB. Dado que o tráfego EF PHB é prioritário, é esperado que as filas destinadas a esta classe estejam quase sempre vazias em todas as circunstâncias.

O encaminhamento expresso, EF PHB, pode ser utilizado para criar serviços de baixa perda, baixo retardo e banda garantida. O EF PHB foi originalmente proposto para caracterizar um tratamento de encaminhamento similar àquele de uma fila com prioridade simples. Segundo [WZ2001], o EF PHB é definido como um tratamento de encaminhamento para o tráfego agregado onde a taxa de saída dos pacotes de agregado no roteador de qualquer nó DiffServ deve ser igual ou exceder a taxa configurável. O tráfego EF pode receber esta taxa independente da intensidade de qualquer outro tráfego tentando transitar pelo nó. Pode ser argüido que tal tratamento de encaminhamento também recai na definição de uma classe AF. Contudo, é implicitamente assumido que o tráfego EF tem prioridade em relação aos demais tráfegos AF.

Como o tráfego agregado EF recebe uma taxa configurável independente de qualquer outro tráfego tentando transitar pelo nó, as características de encaminhamento são fornecidas

ao agregado, não aos fluxos individuais. Assim, não existe limite superior para o retardo e para a variação de retardo (*jitter*) para um fluxo tributário individual (micro fluxo).

Portanto, o tráfego que entra na rede DiffServ pelo nó de ingresso é servido por pelo menos duas filas, uma para o EF PHB e a outra para, por exemplo, o BE PHB. Quando surge necessidade de individualizar mais comportamentos com diferentes QoS, mais classes podem ser criadas para atender os requisitos de QoS do cliente, empregando o grupo de AF PHB que tem quatro classes independentes com três níveis de prioridade de descarte em cada uma.

2.6 MEDIÇÃO DE FLUXOS SINTÉTICOS EM UM AMBIENTE DiffServ

Considerando que a rede deve estar apta a oferecer vários tipos de serviço tais como: serviços em tempo real, VPN, serviços BE, etc., a monitoração do serviço torna-se muito importante para o fornecimento de QoS fim-a-fim e para a garantia do serviço. Então, nesse contexto, a monitoração não tem somente o papel de diagnóstico mas é uma ferramenta de assessoramento da operação de rede fornecendo um serviço de auditoria tanto para os serviços tradicionais quanto para os serviços de valor adicionado. Além disso, o tráfego de cada tipo de serviço ou classe de serviço tem certos requisitos e exibe certo comportamento como, por exemplo, o serviço VoIP, que requer baixo retardo e baixa perda de pacotes, enquanto que os serviços de dados podem tolerar um retardo médio [AA2003].

Para não prejudicar a operação dinâmica da rede, a arquitetura de medição deve ser capaz de capturar o estado operacional da rede sem degradar a desempenho da rede e sem gerar uma grande quantidade de dados, causando um *overhead* desnecessário.

A monitoração pode ocorrer em diferentes níveis: nível de aplicação, nível de usuário, nível de agregado de tráfego, nível de nó, etc. A quantidade de dados das medidas aumenta com o número de classes, isto é, de acordo com o número de diferentes filas, conforme a quantidade de interfaces e conforme o número de roteadores de borda por classe que deve ser monitorado.

A quantidade de medições pode ser relacionada com a quantidade de itens especificados na SLS da classe ou PHB, ou seja, injetando um tráfego de teste no nó de ingresso e medindo seu desempenho no nó de egresso, podemos inferir quanto ao atendimento das várias SLSs da classe.

Duas abordagens podem ser consideradas nas medições de desempenho. As medições podem ser realizadas entre dois nós de fronteira do domínio DiffServ, de ingresso e de

egresso, que é a medição borda-a-borda ou entre dois nós adjacentes dentro do domínio DiffServ, que é a medição lance-a-lance. Na segunda abordagem, o estado de cada enlace entre roteadores adjacentes poderá ser conhecido, mas ocorrerá imprecisão quando da concatenação das medidas lance-a-lance para estimar os valores de desempenho entre os nós de ingresso e de egresso.

Existem dois métodos para realização de medições em uma rede: medições ativas e medições passivas. Na medição dita ativa é injetado um tráfego sintético de teste na rede com a finalidade de captar o comportamento da rede. Normalmente, o procedimento nas medições ativas exige a cooperação entre os dois nós envolvidos para sincronização e também a existência de uma sessão contínua entre esses dois nós durante todo o procedimento da medição. Nos casos práticos pode ser necessário implementar a sincronização empregando o sistema de posicionamento global (*Global Positioning System* – GPS) ou então utilizar o protocolo de temporização de rede (*Network Time Protocol* – NTP). Por outro lado, na medição dita passiva, é observado o tráfego real, sem injeção de tráfego extra na rede. Enquanto a medição passiva não exige a cooperação entre os pontos de medição, ela requer uma coleta contínua de dados e deve monitorar a ocorrência da carga plena no enlace, o que pode ser problemático em enlaces de alta velocidade [AA2003].

A escolha do método de amostragem da medição e dos intervalos de amostragem determina o nível de precisão da medição (dentro de um dado intervalo de confiança) e a confiabilidade da medição. Um intervalo de amostragem é definido como sendo o período de leitura para medição passiva ou como sendo o intervalo de tempo médio sobre o qual dois pacotes de teste estão separados. O método de amostragem periódico é normalmente utilizado na medição passiva. Uma vez que os dados na medição passiva estão disponíveis em qualquer tempo, o intervalo de amostragem pode ser especificado apropriadamente.

Conforme [AA2003], nas medições ativas, os requisitos básicos para o tráfego de teste são:

- 1) O tráfego de teste deve ser pequeno quando comparado com a carga da conexão sob teste. Caso contrário, o tráfego de teste afetará o desempenho e a medição não refletirá a condição real da rede [MF2005].
- 2) Os intervalos de amostragem devem ser pequenos o suficiente para captarem as flutuações no desempenho da rede.
- 3) Como a carga na rede varia com o tempo, a quantidade e o tipo do tráfego de teste devem ser configuráveis.

- 4) Como descrito na recomendação de métricas de desempenho IP (*IP Performance Metric – IPPM*) [PV1998], deverá ser empregada taxa de amostragem pseudo-aleatória de Poisson de forma que as medições sejam aleatoriamente distribuídas prevenindo a sincronização de eventos.

A quantidade de tráfego de teste (carga de tráfego e taxa de pacotes) injetada na rede para medição do retardo e perda de pacotes, durante um intervalo de tempo especificado, dependerá do número de fluxos, do número de PHB, do número de diferentes tamanhos de pacotes de teste utilizados e dos intervalos de rajada.

No presente trabalho a rede DiffServ é formada por dois roteadores, um de ingresso e outro de egresso, sendo utilizada a classe EF PHB para agregar o tráfego VoIP que ocupará no máximo 30% da capacidade do enlace de saída do roteador. Com isso, não haverá possibilidade do fluxo de controle alterar a condição de congestionamento da rede. Quando da aplicação de tráfego homogêneo, haverá somente um fluxo de controle. Para o caso de tráfego heterogêneo, serão usados dois fluxos de controle, um para cada tipo de codificador VoIP empregado na simulação. Esse procedimento é necessário para verificação da influência do tamanho dos pacotes do fluxo de controle para obtenção de valores de retardo e de jitter de transferência de pacotes mais aderentes aos respectivos valores dos demais fluxos do agregado.

Dado que a topologia da rede consiste de somente dois roteadores, que estão muito próximos, não é esperada a ocorrência de perda de pacotes, sendo desnecessário analisar esta métrica.

Assim, durante as simulações, serão gerados fluxos VoIP onde o fluxo de controle terá função semelhante àquela desempenhada pelo fluxo sintético usado no método de medição ativa, também denominado método intrusivo.

No próximo Capítulo serão apresentadas as características dos codificadores para VoIP e os conceitos para modelagem do tráfego VoIP.

CAPÍTULO 3

VOZ SOBRE IP

Pela sua atual importância como aplicação de convergência de redes, escolheu-se a aplicação de VoIP para o tema de estudo. Neste capítulo é realizada uma breve introdução do que é a aplicação VoIP, são descritas as características dos principais codificadores de voz empregados em VoIP e são apresentados os modelos de tráfego para VoIP. Finalizando o Capítulo, são apresentados e resumidos os trabalhos relacionados que orientam a linha da dissertação.

3.1 CODIFICADORES DE VOZ

O objetivo de todos os sistemas de codificação de voz é transmitir a voz com a maior qualidade possível empregando a mínima capacidade de canal. Estes objetivos devem ser alcançados na concepção do *codec* mantendo aceitáveis os níveis de complexidade de implementação do dispositivo e de retardo na comunicação de voz. [RT1996].

Muitos algoritmos têm sido padronizados para codificação dos sinais de voz com taxas de bit variando desde 5,3 Kb/s até 64 Kb/s, sem considerar o processamento dos cabeçalhos dos protocolos de transporte [TA2000] [CR1997].

Quanto à forma de tratamento do sinal de voz, existem dois tipos básicos de codificadores: os codificadores de forma de onda (*waveform coders*) e os codificadores de fonte (*source coders*). Os codificadores de forma de onda codificam a amplitude da forma de onda no domínio do tempo ou no domínio da frequência.

Os codificadores no domínio do tempo se dividem em dois tipos: os não diferenciais como o PCM (*Pulse Code Modulation*) e os diferenciais como o Delta PCM e o ADPCM (*Adaptive Differential PCM*).

Os codificadores no domínio da frequência se dividem em dois tipos: os codificadores SBC (*Sub-band Coding*) e os codificadores ATC (*Adaptive Transform Coding*).

Os codificadores de forma de onda se dividem em *Vocoders (VOICE CODERS – Codificadores de Voz)* e os codificadores LPC (*Linear Predictive Coders*). Dentro da

categoria *Vocoders* temos: *Channel Vocoders*, *Formant Vocoders*, *Cepstrum Vocoders* e *Voice-Excited Vocoders*.

Genericamente, os *Vocoders* analisam o sinal de voz no lado de transmissão, transmitem parâmetros derivados dessa análise e então sintetizam a voz na recepção empregando tais parâmetros. Esses *Vocoders* trabalham no domínio da frequência determinando o envelope do sinal de voz para um número de bandas de frequência e então realizam a amostragem e a codificação, multiplexando essas informações nas saídas codificadas de filtros. A amostragem é realizada a cada 10 ms (até 30 ms). Juntamente com a informação da energia de cada banda de frequência são transmitidos também o limiar de decisão voz/não voz e o *pitch* de frequência para sintetização da voz.

Os codificadores LPC utilizam informações paramétricas referentes ao sinal de voz original. Nessa categoria se incluem: o LPC (*Linear Predictive Coding*), o ACELP (*Algebraic Code-Excited Linear-Prediction*) e o MP-MLQ (*MultiPulse, MultiLevel Quantization*) [RT1996].

Os algoritmos de codificação padronizados pela ITU-T são regidos pelas seguintes recomendações [CR1997] [KT1998]: G.711 em 64 Kb/s empregando PCM; G.726 utilizando ADPCM em 40, 32, 24 e 16 Kb/s; G.728 empregando LD-ACELP (*Low Delay ACELP*) em 16 Kb/s; G.729A e G.729B empregando CS-ACELP (*Conjugate Structure ACELP*) em 8 Kb/s; e G.723.1 que é baseado na tecnologia MP-MLQ com duas taxas de transmissão, 5,3 e 6,3 Kb/s.

Conforme [HO2002], durante uma conversação, é usado em média apenas 35% do tempo. Assim a compressão ou supressão de silêncio é uma característica importante pois pode acarretar até 50% de redução de banda em ligações ponto a ponto e valores bem maiores em conferências *multicast*.

A compressão de silêncio inclui três blocos funcionais [CR1997]:

- a) Detector de Presença de Voz (*Voice Activity Detector – VAD*), que é responsável por determinar quando o usuário está conversando e quando ele está em silêncio. O detector deve ser bastante sensível de forma a não perder a primeira palavra ou inserir um silêncio inútil no final das sentenças, mas não pode ser disparado por ruído de fundo. A detecção de presença de voz avalia a energia das amostras que chegam e ativa o canal de mídia se essa energia estiver acima de um certo valor mínimo. De forma similar, quando a energia cai abaixo desse limiar durante algum tempo, o canal de mídia é silenciado. Se o módulo VAD descartasse todas as amostras até que a energia média das amostras que chegam atingisse o limiar, o

início do período de fala ativa seria cortado. Portanto, as implementações VAD exigem um certo esquema de previsão, isto é, elas retêm na memória cerca de 150 ms de amostras para iniciar a ativação do canal de mídia antes do período de fala ativa. Isso geralmente acrescenta um retardo adicional no processo de codificação. O VAD também é útil para economizar energia, no caso de dispositivos que funcionam a bateria, suspendendo a transmissão durante os períodos em que não há voz presente.

- b) Transmissão Descontínua (*Discontinuous Transmission – DTX*), é a capacidade do codificador suspender a transmissão de quadros quando o VAD tiver detectado um período de silêncio. Se a transmissão for interrompida completamente, deve ser acionado o bit marcador M do primeiro pacote do RTP (*Real-time Transport Protocol*) após o período de silêncio. Alguns codificadores avançados não interrompem a transmissão completamente mas, em vez disso, vão para um modo de silêncio no qual usam muito menor largura de banda e enviam apenas os parâmetros mínimos (intensidade, etc.) para que o receptor possa reconstruir o ruído de fundo.
- c) Gerador de Ruído de Conforto (*Comfort Noise Generator – CNG*), é usado para recriar algum tipo de ruído de fundo. Nos codificadores mais antigos, quando a parte chamadora não estivesse falando e houvesse silêncio na linha, sendo esse período de silêncio suficientemente grande para ser detectado pelo VAD, a transmissão era desligada completamente, ficando na recepção a sensação de que a ligação estava interrompida. Para evitar isso, o CNG insere algum ruído aleatório de fundo com um nível deduzido a partir dos níveis mínimos obtidos durante períodos de fala ativa. Os codificadores mais avançados, como o G.723.1 (Anexo A) ou o G.729 (Anexo B), possuem opções para enviar informações que permitem ao decodificador remoto reconstruir um ruído de ambiente próximo do ruído de fundo original.

Os codificadores de áudio atuais podem melhorar a qualidade da voz reproduzindo o ruído de fundo do usuário que está falando. Essa característica é suportada por um quadro especial denominado SID (*Silence Insertion Descriptor*), gerado no lado do usuário que está falando e que descreve as características principais do ruído de fundo. Os quadros SID são gerados durante os períodos de inatividade da voz e seu esquema de codificação difere daquele dos quadros de voz (quadros ACT). Os quadros SID são gerados pelos algoritmos

DTX dos codificadores conforme ocorram mudanças na energia do ruído de fundo ou também podem ser gerados por alguma regra específica dependendo da implementação do codificador [EA2003] [EA2004].

A voz humana consiste de surtos de fala e intervalos de silêncio, também conhecidos como padrões ON/OFF. A existência de surtos e intervalos de silêncio permite a supressão do silêncio e os principais benefícios dessa supressão são [JW2000]:

- permitir maior utilização da banda através da multiplexação;
- permitir ajuste do retardo *playout* por surto de voz;
- capacitar a supressão de eco baseada na saída do detector de silêncio.

Os algoritmos de detecção de atividade de voz em *Vocoders* usam uma largura de banda significativamente menor em períodos de silêncio do que em períodos de voz ativa. Assim, para a maioria dos codificadores, a taxa de transmissão de bits em atividade é constante e em períodos de silêncio, os codificadores mais simples possuem taxa de transmissão de bits nula, mas os codificadores G.723.1 e G.729, enviam informação nos períodos de silêncio para descrever o nível de ruído de fundo e dados de outros parâmetros. Por exemplo, o codificador G.723.1 (5,3 Kb/s) trabalha com 4 quadros por pacote e sua taxa de transmissão nos períodos de atividade é de 8 Kb/s e nos períodos de silêncio é de 3,73 Kb/s. Deve ser observado que a taxa de transmissão de 5,3 Kb/s desse codificador corresponde à taxa de transmissão na saída do codificador e a taxa de 8 Kb/s, acima citada, inclui o *overhead* dos cabeçalhos dos protocolos de transporte.

A maioria dos codificadores de voz se baseia em blocos denominados quadros. Isso significa que eles comprimem blocos contendo um número fixo de amostras quantizadas linearmente, ao invés de comprimirem amostra por amostra. Assim sendo, o fluxo de dados de áudio precisa ser acumulado até que atinja o tamanho do bloco antes de ser processado pelo codificador. Esse processo adiciona retardo na comunicação. Outros codificadores precisam saber um pouco a respeito das amostras que vêm imediatamente após as amostras que ele está codificando naquele quadro. Esse esquema é denominado previsão (*lookahead*). Nesse caso, o atraso mínimo introduzido por uma seqüência de codificação e de decodificação, denominado atraso de algoritmo, é igual ao comprimento do quadro acrescido do tempo extra devido ao esquema de previsão [HO2002].

Aparentemente, para reduzir retardos em uma rede ideal, o *codec* escolhido deve ter um comprimento de quadro pequeno. No entanto, os codificadores com tamanhos de quadro

maiores tendem a ser mais eficiente e a ter melhores taxas de compressão. Ocorre que o quadro é transmitido através da rede com uma quantidade considerável de *overhead*, acrescentado pelos protocolos de transporte em cada pacote transmitido. Se cada quadro de voz comprimida for transmitido em um pacote único, esse *overhead* será acrescido no pacote e em alguns casos o *overhead* terá tamanho comparável, senão maior, que os dados úteis. Assim, para diminuir o *overhead* a um nível aceitável, a maioria das implementações transmite vários quadros em cada pacote. Deve ser observado que se todos os quadros acumulados no pacote pertencerem ao mesmo fluxo de áudio, haverá um aumento no retardo de acumulação. Portanto, em termos de *overhead* e de retardo de acumulação, usar um codificador com um tamanho de quadro igual a ζ e três quadros por pacote é equivalente a usar um codificador com um tamanho de quadro igual a 3ζ e um quadro por pacote [HO2002].

Os fluxos de pacotes em conexões VoIP são codificados quase que em taxa de bit constante (*Constant Bit Rate* – CBR). Logo, em um *codec*, o intervalo dos pacotes e o tamanho dos pacotes são constantes. O que deve ser considerado é que, diferentes codificadores produzem diferentes intervalos e diferentes tamanhos de pacotes, mesmo se a banda efetiva for a mesma. Por exemplo, um codificador G.711 produz um fluxo de 64 Kb/s. Cada pacote deve conter pelo menos 8 bits, que é a quantidade de bits para uma amostra. A amostra deve ser enviada a cada $125\mu\text{s}$ para produzir uma taxa de 64 Kb/s, como na ISDN.

O *overhead* total da rede de transporte formado pelos protocolos IP, UDP (*User Datagram Protocol*) e RTP para um pacote com *payload* de 8 bits (1 Byte) é cerca de 40 Bytes. Além disso, nesse cenário 8.000 pacotes devem ser transmitidos dentro de 1 segundo, consequentemente, mais do que uma amostra é transferida em um único pacote, normalmente de 30 a $(20\text{ms}/125\mu\text{s}) = 160$. A quantidade exata depende do codificador empregado. Isso significa que cada fluxo enviará de 22 a 100 p/s (pacotes por segundo). Para resolver o problema do enorme cabeçalho (*header*) comparado com o *payload*, é realizada uma compressão (cRTP) nos cabeçalhos (IP + UDP + RTP). Essa compressão reduz o cabeçalho de 40 Bytes para 2 ou 4 Bytes. No entanto, essa compressão não é largamente empregada devido a incompatibilidade de *hardware* [HO2003].

Portanto, na maioria das implementações são enviados múltiplos quadros por pacote e o comprimento de quadro real a ser levado em conta é, a soma de todos os quadros agrupados em um único pacote IP. Esse artifício é limitado, por questões de eco e de interatividade, a 120 ms de voz codificada em cada pacote IP. Assim, para a maioria das implementações,

quanto menor o tamanho do quadro, maior será o número de quadros em um pacote IP e não haverá influência sobre o atraso.

Também é considerado o fato de que as amostras devem ser atrasadas (retardadas) no tempo quando elas forem obtidas durante a espera para transmissão. Esse período é denominado retardo endógeno (*endogenous delay*) que é constante para cada fluxo VoIP e poderá unicamente ser encurtado a custa de intervalos de pacotes menores e portando causando um *overhead* de pacote maior.

A Tabela 3.1 apresenta as características dos *codecs* padronizados pela ITU-T e que, são objeto do presente trabalho por serem os mais utilizados no mercado.

Tabela 3.1: Dados de três *codecs* de voz padronizados pela ITU-T, empregados nas simulações [TA2000]

CODEC	G.711	G.723.1	G.729B
Velocidade de codificação [Kb/s]	64	5,3	8
Tamanho do quadro [ms]	20	30	10
Quadro na saída do codec (<i>payload</i>) [Bytes]	160	20	20
Quadro com PPP+IP+UDP+RTP [Bytes]	206	66	66
Quadro SID [Bytes]	--	4	--
Quadro de Ruído de Conforto [Bytes]	--	--	2
Pacote sem compressão [Bytes]	120	70	60
Pacote com compressão [Bytes]	82	32	22
Retardo de processamento [ms]	20	30	10
Retardo lookahead [ms]	0	7,5	5
DSP MI PS	0,34	16	20

No item 4 do Capítulo 2 foram apresentados os parâmetros de QoS e os objetivos de desempenho da rede IP. Após a descrição acima, referente ao funcionamento do *codec* de voz, aqueles conceitos podem ser expandidos para aplicações VoIP conforme [TA2000]:

- 1) Retardo na entrega de pacotes. Dá origem a dois problemas, o eco e a superposição dos sinais de voz dos interlocutores (*talker overlap*). Portanto, desde que o retardo seja grande é necessário o cancelamento de eco em um sistema VoIP. O problema *talker overlap* torna-se significativo para retardo unidirecional maior do que 400 ms. É importante manter o retardo fim-a-fim em um mínimo. Para aplicações VoIP, os requisitos de retardo variam com os usuários e com as aplicações. Um retardo dentro da faixa de 150-200 ms é aceitável para boa qualidade de voz [KT1998]. Um retardo de 300 ms pode ser tolerado em uma conversação aceitável. Conforme a Tabela 3.1, os retardos devido a codificação e ao *framing* variam entre 40-80 ms para os

vários tipos de *codec*. O restante do retardo é devido ao retardo da rede e ao retardo devido à compensação de *jitter*. Para uma conversação de voz ter boa qualidade, o retardo da rede, incluindo a compensação de *jitter*, não deve exceder a faixa de 70-160 ms. Com o retardo de rede entre 220-260 ms a qualidade da conversação de voz é apenas aceitável.

- 2) Variação do retardo: desde que os pacotes experimentam variação dos retardos em uma rede, o tempo interpacotes no lado do receptor não é constante, mesmo que esse tempo seja constante no lado do emissor. Para retirar o *jitter* é empregada a técnica de *play-out buffer*, que causa um retardo adicional nos pacotes.
- 3) Perda de pacotes. Os pacotes de voz são sensíveis ao tempo. Logo, um atraso ou perda de pacote irá afetar a qualidade da voz recebida. Para recuperar da perda de pacotes, a voz pode ser interpolada na recepção a partir do último pacote recebido de forma a ser preenchido o lugar de um pacote perdido. Quando é utilizado o EF PHB para VoIP, pacotes podem ser perdidos devido a grandes retardos. Deve ser observado que no presente trabalho a perda de pacotes esperada é praticamente nula devido às condições de simulação como tamanho do buffer, ocupação da banda disponível e comprimento do enlace entre os roteadores.

3.2 MODELAGEM DO TRÁFEGO VoIP

Um bom conhecimento dos modelos de tráfego é importante para realizar simulações de qualquer tipo de rede [FV1994]. O conhecimento das características do tráfego para um conjunto de aplicações permite estabelecer modelos de tráfego e, através deles, estudar o comportamento das redes. Assim, torna-se possível, através de simulações baseadas em tráfego sintético, inferir sobre requisitos que possam ser necessários para o projeto da rede ou para o gerenciamento de tráfego de forma que a rede suporte adequadamente a carga de tráfego e entregue a QoS adequada ao usuário final [HO2003].

Nas redes baseadas em pacotes, a modelagem do tráfego emprega funções distribuição de probabilidade (*Probability Distribution Function* - PDF) que permitem estabelecer previsões sobre eventos discretos tais como: chegada (criação) de pacotes (de tamanho variável ou não) que possam ser produzidos por uma combinação de protocolos de aplicação e

de transporte para uma fonte única; bem como o fluxo de pacotes resultante. Matematicamente, as chegadas de pacotes podem ser descritas como um processo temporal de eventos discretos. As chegadas das rajadas de pacotes podem também ser consideradas como um processo discreto que relaciona estatísticas de eventos específicos de protocolos de alto nível tais como sessões ou conexões. Nesse caso, a natureza destes processos para a modelagem do tráfego depende de características particulares dos protocolos de transporte e de aplicação.

Modelos de tráfego agregado tentam prever as chegadas e tamanhos de pacotes pela combinação de diferentes fluxos no enlace. A modelagem do tráfego agregado é importante como uma simplificação dos efeitos do vasto número de fluxos de tráfego em um enlace. O tráfego agregado tem sido estudado rigorosamente em cenários de rede específicos [HO2003].

A maioria dos modelos matemáticos de tráfego pode ser dividida em duas categorias: processos probabilísticos sem memória, tal como o emprego de uma função densidade de probabilidade (*probability density function* - pdf) de decaimento exponencial em um modelo de Poisson para descrever as chegadas de pacotes, e processos estocásticos com memória. Os processos sem memória podem ser aplicados como uma aproximação aos tempos de interchegadas de pacotes e aos tamanhos dos pacotes, mas podem também ser aplicados em períodos ON/OFF dos ciclos de voz.

A maioria das análises de desempenho de sistemas de voz por pacotes está baseada no modelo ON/OFF clássico desenvolvido por Brady em 1965 [BP1965] [BP1968] [BP1969]. Esse modelo foi criado para o estudo de detectores de voz empregando a técnica TASI (*Time Assigned Speech Interpolation*) também conhecida como DSI (*Digital Speech Interpolation*), onde a atividade do sinal de áudio era modelada alternando entre dois estados: surto de voz (*talkspurt*) ou ON e período de silêncio (*silence gap*) ou OFF. Conforme [BP1965], um *talkspurt* é o período de tempo entendido pelo ouvinte como tendo uma seqüência de sons de voz não interrompida por uma pausa e uma pausa é um período de tempo entendido pelo ouvinte como sendo um período sem fala, que não seja provocado por uma parada consonantal ou por uma leve hesitação ou por um leve respiro. Por outro lado, *spurts* e *gaps* provêm do detector de voz, sendo por definição eventos mensuráveis fisicamente, enquanto *talkspurts* e *pauses* são determinados subjetivamente.

Conforme [PK2000], recentemente foi possível coletar uma grande quantidade de medições de tráfego em redes de comunicações. Muitos desses conjuntos de dados foram utilizados para validar as tradicionais suposições estatísticas feitas quando da análise de tais redes. Essas suposições partiram da premissa de que o tráfego na rede poderia ser descrito por

modelos markovianos. Isso implicava que as autocorrelações no tráfego da rede tinham decaimento exponencial e de forma rápida. Assim, essa espécie de tráfego deveria se comportar de forma plana (intensidade aproximadamente constante) sobre grandes escalas de tempo. No entanto, foi constatado que essas tradicionais suposições markovianas não eram sempre satisfeitas. Uma cuidadosa análise estatística realizada por Leland *et al* em [LW1994] mostrou que o tráfego nas redes *Ethernet* de área local no *Bellcore* exibiam propriedades como auto similaridade (*self-similarity*) e LRD (*long-range dependence*). Em particular, esse tráfego comportava-se de forma extremamente variável (*bursty*) em uma grande faixa de escalas de tempo, conforme visto na Figura 3.1.

O lado esquerdo da figura mostra traços reais do tráfego Ethernet LAN (*Local Area Network*). Começando com uma unidade de tempo de 100 segundos, cada gráfico subsequente no sentido vertical foi obtido do gráfico acima pela redução do intervalo de tempo e amplitude por um fator de 10 e pela ampliação do intervalo escolhido aleatoriamente, mostrado pela área escura na figura [LW1994].

A Figura 3.1 mostra também que o traço do tráfego medido tem grande variação em todas as escalas de tempo, diferentemente do que ocorre com o traço obtido por simulação a partir do modelo convencional de tráfego, baseado na distribuição exponencial, tendo mesma intensidade de chegada e mesmo tamanho de pacote médio. No lado direito da figura pode ser visto que o tráfego do modelo convencional é praticamente plano nas mesmas escalas de tempo.

Análises estatísticas da correlação do tráfego medido na *Ethernet LAN* mostraram que sua função de autocorrelação decai muito lentamente. Esta propriedade está proximamente relacionada com a noção de LRD.

Para definir LRD, é empregado um processo estocástico estritamente estacionário dado por $\chi = \{X(t), t \geq 0\}$ onde $X(t)$ pode ser pensado como a taxa de tráfego da rede gerada no tempo t . A quantidade cumulativa de tráfego no tempo t é dada por $T(t) = \int_0^t X(u) \cdot du$. e a autocorrelação é definida pela função $c(t) = Cov\{X(s), X(s+t)\} / Var\{X(s)\}$.

Assim, χ é dita ser SRD (*short-range dependent*) se $\int_0^\infty |c(t)| \cdot dt < \infty$ e, caso o resultado dessa integral seja ∞ , χ é dita ser LRD. Existem outras definições para LRD, como por exemplo, χ é LRD se a função de autocorrelação $c(\cdot)$ mostrar um tipo particular de comportamento conforme a lei de potência: $c(t) \approx c_0 \cdot t^{-\alpha}$, $0 < \alpha < 1$.

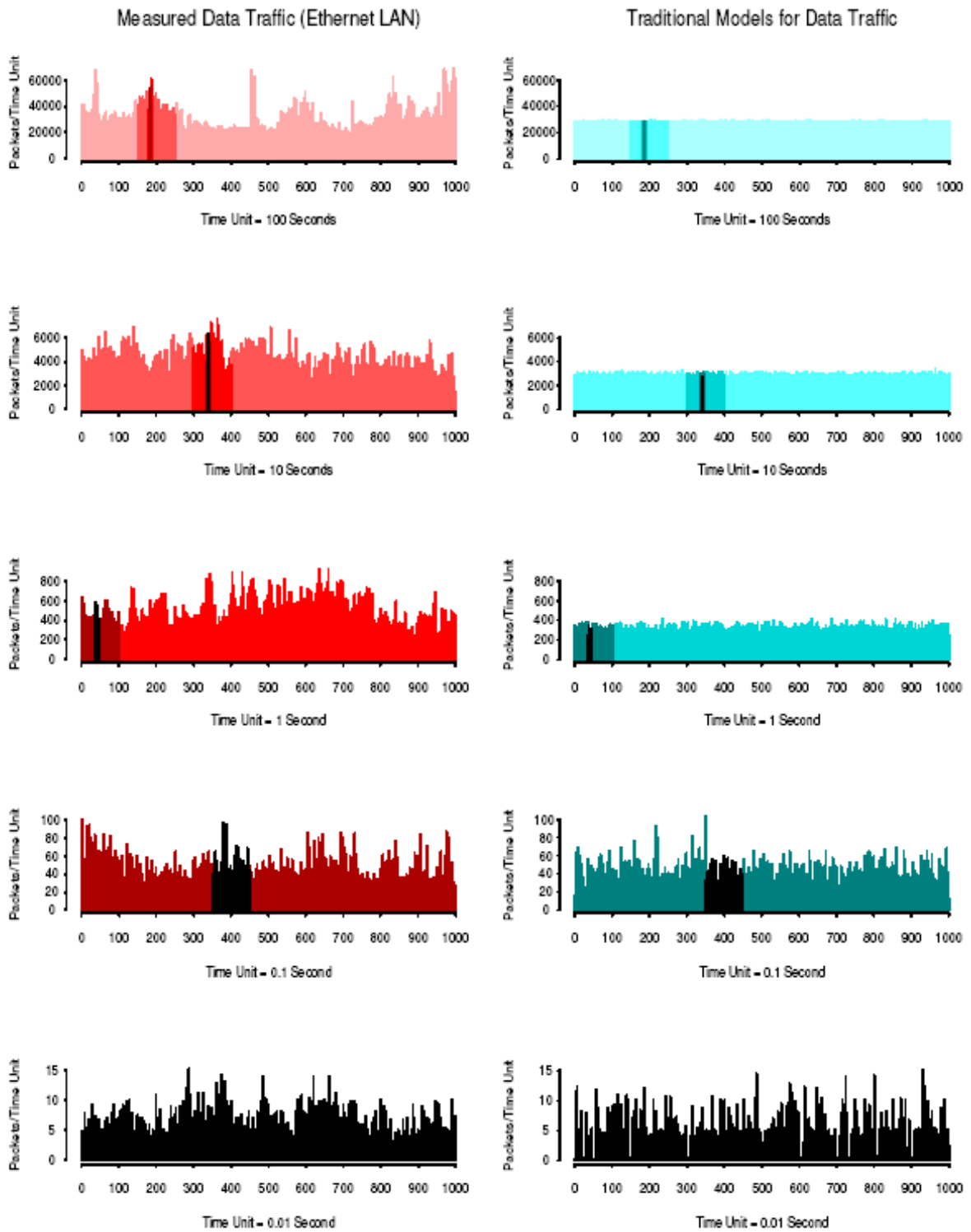


Figura 3.1: Comparação entre traços de tráfego de *Ethernet LAN* e de redes convencionais.

As suposições tradicionais tipicamente admitem que $c(t)$ decresce conforme uma exponencial negativa em t dada por: $c(t) \approx c_0 \cdot e^{-\lambda t}$. A definição de LRD dada acima está intimamente relacionada ao comportamento da variância do processo de tráfego cumulativo

$T(t)$. A identidade $Var\{T(t)\} = 2 \int_0^t \int_0^u c(v) \cdot dv \cdot du$, mostra que a variância de $T(t)$ tem comportamento linear em t se χ for SRD, e tem comportamento super linear em t se χ for LRD. Em particular, se $c(t) \approx c_0 \cdot t^{-\alpha}$, $0 < \alpha < 1$ for aplicável, então $Var\{T(t)\} \approx c_1 \cdot t^{2-\alpha}$, onde c_1 pode ser expresso em termos de α e c_0 .

Uma segunda propriedade importante é a auto similaridade que é definida como segue. Um processo estocástico $\chi = \{X(t), t \geq 0\}$ é (estritamente) auto similar com parâmetro H se $\chi = \{X(t), t \geq 0\}$ e $\{\gamma^{-H} \cdot X(\gamma t), t \geq 0\}$ tiverem as mesmas distribuições de dimensão finita para qualquer $\gamma > 0$. Deve ser observado que um processo auto similar é não estacionário.

Se χ for auto similar com parâmetro H , então $Var\{X(t)\} = t^{2H} \cdot Var\{X(1)\}$. Qualquer processo χ satisfazendo essa propriedade é chamado de auto similar de segunda ordem. Uma forma mais fraca de auto similaridade é a auto similaridade de segunda ordem assintótica, que é também definida para um processo de tempo discreto. É dito que χ satisfaz esta propriedade se uma apropriada versão normalizada e centrada de $\{\chi(\gamma t), t \geq 0\}$ convergir para um processo auto similar quando $\gamma \rightarrow \infty$.

As noções de auto similaridade e LRD são relacionadas em alguns exemplos, mas não são equivalentes. Por exemplo, o Bm (*Brownian motion*) é auto similar com $H = 1/2$ mas não é LRD. De modo inverso, existem também processos LRD que não são auto similares. Se $H > 1/2$, então as definições de auto similaridade assintótica e de LRD são equivalentes [PK2000]. Uma explanação intuitiva é que se $T(\cdot)$ é auto similar com parâmetro de Hurst, H , então vale dizer que $Var\{T(t)\} \approx c_2 \cdot t^{2H}$. Nesse caso, a variância de $T(t)$ tem comportamento super linear em t se $H > 1/2$ e as constantes H e α são relacionadas pela identidade $H = (2 - \alpha)/2$. Em vista dessa equivalência geralmente não é usado explicitamente o termo auto similaridade mas é usado LRD.

Como dito anteriormente, existem agora muitas evidências que o tráfego da rede é auto similar e LRD. Além do trabalho de Leland *et al* [LW1994], outros trabalhos mostraram essas propriedades tais, como: Willinger *et al* [WW1995] para o tráfego em LAN, Paxson e Floyd [PV1995] para o tráfego em WAN (*Wide Area Network*) e Beran *et al* [BJ1995] para o tráfego VBR (*Variable Bit Rate*) de vídeo [PK2000].

Todas essas propriedades são examinadas no nível de pacotes. Um número de estudos explanou esses resultados examinando quantidades relacionadas ao tráfego da rede em níveis mais altos de agregação, particularmente no nível de aplicação onde as entidades são o

tamanho das filas, o tempo de conexão, o tempo de transmissão, etc. Vários estudos nesse nível indicaram que LRD podia ser causada pela distribuição com cauda pesada de certas características do tráfego. Assim, Crovella e Bestavros [CM1997] mostraram que os tamanhos das filas e os tempos de transmissão das filas na Internet têm distribuições cujas caudas obedecem a lei de potência com variância infinita, ou seja, se Y for um tamanho de fila genérico ou o tempo de transmissão, então $P\{Y > t\} \approx c_3 \cdot t^{-\alpha}$, $0 < \alpha < 2$.

Distribuições de cauda pesada e LRD estão intimamente relacionadas. O processo canônico ON/OFF, por exemplo, é LRD se e somente se os tempos de ON e de OFF tiverem variância infinita. Uma importante observação é que a LRD pode ser originada pela forma de cauda pesada de muitas distribuições de características relacionadas ao usuário tal como o comprimento das filas, que é uma consequência da demanda do usuário, ou dos mecanismos de controle de tráfego, como o TCP (*Transmission Control Protocol*) utilizado na Internet..

Na realidade, o interesse não está apenas nas distribuições de cauda pesada, mas em todas as distribuições nas quais a cauda decresce mais lentamente do que exponencialmente, formando várias classes de distribuições de cauda pesada.

Sejam $X, X_i, i \geq 1$, variáveis aleatórias não negativas e independentes com função distribuição de probabilidade $F(x) = P\{X \leq x\}$ e seja $\bar{F}(x) = 1 - F(x)$, então:

a) F é uma distribuição de cauda pesada se, para todo $\varepsilon > 0$, $E\{e^{\varepsilon X}\} = \infty$, ou equivalentemente, se para todo $\varepsilon > 0$, $\frac{P\{X > x\}}{e^{-\varepsilon x}} \rightarrow \infty$.

b) F é uma de distribuições de cauda longa, se para qualquer valor fixo $y > 0$ e $x \rightarrow \infty$, $P\{X > x + y | X > x\} = \frac{\bar{F}(x + y)}{\bar{F}(x)} \rightarrow 1$. A classe de distribuições de cauda longa é representada por \mathcal{L} .

c) F é subexponencial se $\frac{\bar{F}^{2*}(x)}{\bar{F}(x)} = \frac{P\{X_1 + X_2 > x\}}{P\{X > x\}} \rightarrow 2$, $x \rightarrow \infty$, onde F^{n*} é a

convolução de dobra n de F , isto é $F^{n*}(x) = \int_{u=0}^x F^{(n-1)*}(x-u) \cdot dF(u)$. A classe de funções subexponenciais é representada por \mathcal{S} . Deve ser observado que se $F \in \mathcal{S}$ então $P\{X_1 + \dots + X_n > x\} \approx n \cdot P\{X_1 > x\}$ para algum $n \geq 2$. Da mesma forma F é subexponencial se $P\{X_1 + \dots + X_n > x\} \approx P\{\max\{X_1, \dots, X_n\} > x\}$ para algum $n \geq 2$.

Algumas bem conhecidas distribuições de probabilidade subexponenciais são: Pareto, Lognormal e Weibull, que estão descritas no Anexo 1.

As funções distribuição de probabilidade fornecem a estatística de primeira ordem, ou seja, o número relativo de contagens simples de um evento (considerado assintoticamente). Exemplos da estatística de primeira ordem são os tempos de interchegadas determinados pela distribuição exponencial em um processo de Poisson [HO2003].

De especial interesse para este trabalho são as distribuições de cauda pesada e variância infinita. A indicação de que uma distribuição é de cauda pesada é obtida através da forma assintótica da distribuição seguindo uma lei de potência, conforme

$$P\{X > x\} \approx x^{-a} \text{ com } a \rightarrow \infty.$$

Se $a \leq 2$ a distribuição tem variância ∞ e se $a \leq 1$ então a distribuição também tem média ∞ . Assim, enquanto a decresce, uma grande porção da massa (valores discretos) de probabilidades está presente na cauda da distribuição.

Entre as características de tráfego da rede que devem ser representadas pelo modelo de tráfego destaca-se, pelo seu impacto no desempenho das aplicações, a variação das rajadas de tráfego (*traffic burstiness*).

Os fluxos de pacotes na rede de computadores ou inter-redes geralmente ocorrem em rajadas. O conhecimento da frequência de ocorrências e do nível de variação das rajadas em fonte única e no tráfego de agregado é de vital importância para o projetista da rede uma vez que estas variações podem causar sobrecarga na memória de armazenamento (*buffer overflow*).

Existem vários indicadores de *burstiness*, dados em relação ao tempo. Uma medida comum da variação das rajadas de tráfego é o parâmetro de Hurst, o qual pode também ser considerado como uma medida para auto-similaridade. Como visto, a condição de tráfego auto-similar indica a existência de tráfego com comportamento similar observado em escalas de tempo diferentes – essa é uma característica de fenômeno fractal. A LRD, tradicionalmente definida como *o descaimento lento das correlações em uma série temporal e referida como um processo de memória longa*, conduz à condição de auto-similaridade. Tem sido mostrado que a auto-similaridade afeta fortemente o desempenho dos protocolos de rede [CM1997] e que ela pode ter origem em fontes de tráfego isoladas ou na agregação dessas fontes. Considera-se auto-similar o tráfego cujo valor do parâmetro de Hurst é maior que 0,5 (não auto-similar no todo) até 1 (auto-similar) [BB2002].

O emprego de tráfego gerado computacionalmente (tráfego sintético) com variação de rajadas e auto-similaridade é geralmente uma boa aproximação da realidade em simulações.

Nesse caso, é necessária uma forma eficiente para geração das variações nas rajadas de tráfego simulado ou na geração de tráfego auto-similar.

Um método para obtenção de variação de rajadas e auto-similaridade é a superposição de fontes ON/OFF de tráfego [TM1997]. No modelo ON/OFF, os fluxos de pacotes transmitidos são comutados para o estado ON ou para o estado OFF pelas fontes enquanto elas geram conjuntos de pacotes para filas separadas ou enquanto usuários esperam entre pedidos de várias filas ou mídias. O modelo ON/OFF para VoIP considera distribuições de surtos de voz e de intervalos de silêncio produzidas pelos codificadores. Foi mostrado em [TM1997] que a superposição de um grande número de fontes ON/OFF, onde as fontes alternam períodos ON e OFF que tem alta variabilidade, conduz para a formação assintótica de tráfego auto-similar. Nesse caso, se os fluxos constituintes tiverem variância infinita o tráfego poderia ser auto-similar, mas se esses fluxos fossem simplesmente de cauda pesada (*heavy tail*) sem a variância infinita o tráfego criado poderia simplesmente ter rajadas; isso corresponde aos parâmetros de Hurst $H \leq 0,5$ (tráfego em rajadas), e se o parâmetro de Hurst $H > 0,5$ o tráfego é auto-similar [PK1996].

Uma distribuição efetiva para usar nesse método é a distribuição de Pareto [CM1997]. Como demonstrado por Bodamer et al., o parâmetro de Hurst do modelo m-Pareto resultante vale $H = \frac{3-a}{2}$, onde a é o fator de forma da distribuição de Pareto.

Deve ser observado que a variação de rajadas tal que $1/2 \leq H < 1$ sugere tráfego auto-similar, onde o grau de auto-similaridade aumenta quando $H \rightarrow 1$, que implica em $1 < a \leq 2$ com a auto-similaridade aumentando quando $a \rightarrow 1$. Mas, $1 < a \leq 2$ é também o critério para variância infinita da distribuição de Pareto, portanto para obtenção de fontes ON/OFF de Pareto independentes e identicamente distribuídas (i.i.d.) devemos ter variância infinita [HO2003].

Em síntese, uma fonte de tráfego LRD tem autocorrelação que decai em uma taxa menor do que a da distribuição exponencial, implicando na existência de forte conexão entre o comportamento da fonte em um certo tempo t_0 e seu comportamento em $t_{0+\tau}$, onde τ pode ser grande. Uma fonte auto-similar mantém a mesma estrutura de autocorrelação através de diferentes escalas de tempo. Essas duas características estatísticas têm influência sobre a variação das rajadas do tráfego multiplexado, desde que a duração dos períodos de rajada possa ser muito grande com uma probabilidade não desprezível, que caracteriza as distribuições de cauda pesada.

Em [PV1995] foram investigados seis métodos computacionais para a criação de tráfego auto-similar, a saber: superposição de fontes ON/OFF de cauda pesada; emprego de fila M/G/ ∞ com tempos de serviço distribuídos com cauda pesada; RMD (*Random Midpoint Displacement*); computação de coeficientes *wavelet* de transformadas *wavelet* de fBm (*fractional Brownian motion*); processo F-ARIMA (*Fractional Autoregressive Integrated Moving Average*); e fBm (*fractional Brownian motion*) [HO2003].

Os modelos de tráfego podem ser estacionários e não estacionários. Os modelos estacionários podem ser classificados em geral em duas classes: SRD (*short-range dependence*) e LRD. Os modelos SRD incluem os processos de Markov e modelos de Regressão. Esses modelos de tráfego têm uma estrutura de correlação que é significativa para τ relativamente pequenos. Os modelos LRD tais como F-ARIMA e fBm têm significantes correlações mesmo para grandes valores de τ .

Sriram e Whitt [SK1986] produziram um modelo ON/OFF de simulação para um processo de chegada de pacotes de voz. Um enlace na saída do roteador transportava somente tráfego de voz e utilizava VoIP com N chamadas simultâneas em progresso. O tamanho dos pacotes era fixo e esses tamanhos dependiam do *codec* utilizado. A conversação de voz era formada por períodos de silêncio e de surtos de voz, exponencialmente distribuídos, com média de 650 ms e 352 ms, respectivamente. Se T em [ms] é o período de transformação da voz em pacotes, então o número esperado de pacotes, $E\{p\}$, em um *talk spurt* é $E\{p\} = 352/T$. Para uma conversação, o tempo de interchegadas de pacotes é T com probabilidade $p = (E\{p\} - 1)/E\{p\}$ e é $(650 + T)$ com probabilidade $(1-p)$. Assim, o tempo de interchegadas de pacotes para uma conversação é $\lambda^{-1} = T \cdot p + (650 + T) \cdot (1 - p)$, sendo T determinado pelo tipo de codec usado. Os autores mostraram que usando $T = 16$ ms, um enlace com taxa $T1$ (1,54 Mb/s) e tamanho de pacote fixo de 512 bits, o modelo de fila M/D/1 forneceu bom resultado até $N = 100$ conversações simultâneas de voz, com fator de utilização 0,73.

Sabendo que λ é a taxa de chegada de pacotes para uma conversação de voz,, então para N conversações simultâneas, a taxa poderia ser $N\lambda$, e assim o processo de superposição poderia aumentar e não se tornar esparso. O trabalho mostrou que o modelo de Poisson para o processo de chegadas de pacotes era bom para utilização entre 60 e 70%. Para altas utilizações, o modelo subestimou o desempenho pois os retardos reais são maiores do que aqueles obtidos no modelo M/D/1, na realidade os fluxos não são esparsos e as covariâncias

de longa duração entre os tempos de interchegadas dos pacotes tornam-se mais significativas. Assim, foi abandonado esse modelo que é uma das opções apresentadas pelo OPNET.

Através das referências bibliográficas citadas no próximo item, foi constatado que a obtenção de dados para a construção de distribuições estatísticas que representem os períodos ON e OFF poderiam ser conseguidas a partir de trilhas obtidas por gravação de diversos tipos de conversações empregando os diversos tipos de *codecs* VoIP em: aplicações de telefonia IP, leituras *multicast* e sessões de conferência multimídia que envolvam múltiplos usuários falando em diferentes dinâmicas de interação além de conversações bidirecionais. Assim, após o tratamento das chamadas, elas devem ser parametrizadas com o intuito de obter a média e a variância da distribuição. A seguir, é realizado o teste de aderência de Kolmogorov-Smirnov (*goodness-of-fit test*), ou teste KS, com o qual é procurada a distribuição teórica que melhor representa a distribuição real, sendo então determinados os parâmetros da distribuição teórica.

Na seqüência, são apresentados os trabalhos relacionados que forneceram os dados de tráfego VoIP e as respectivas distribuições utilizadas no presente trabalho.

3.3 TRABALHOS RELACIONADOS

Dentro do extenso conjunto de referências bibliográficas foram selecionados e resumidos os artigos que tem relação direta com o presente trabalho, a saber:

- [TA2000] realiza uma simulação em um cenário formado por dois roteadores de fronteira de um domínio DiffServ interligados por um enlace T1 (1,54 Mb/s). A disciplina de fila usada é CBQ. Define que 30% da capacidade (450 Kb/s) desse enlace é reservada para o fluxo agregado EF PHB formado por fluxos VoIP. Simula as condições de tráfego homogêneo obtido com cada tipo de codificador (G.711, G.723.1 e G.729B) e posteriormente simula as condições de tráfego heterogêneo obtido pela combinação de dois codificadores, a carga do enlace é variada entre 15% e 85%. No caso de tráfego heterogêneo, a quantidade de fluxos G.711 é variada entre 15% e 85% de 450 Kb/s e a carga restante é preenchida com fluxos de um outro codificador. Verifica a influência do tamanho dos pacotes, do número de fluxos e das taxas de transmissão dos codificadores nas características do agregado. No nó de egresso verifica os efeitos no retardo e no *jitter* em cada fluxo dos codificadores. Os resultados obtidos nesse experimento foram os seguintes:

- 1) O *jitter* e o retardo devido a fila foram muito pequenos em termos percentuais da taxa de pacote. O *jitter* foi cerca de 1% do tempo da taxa de pacote para o 50º percentil e foi cerca de 5% para o 90º percentil, considerando a rede com pouca carga. Para alto carregamento da rede esses percentis subiram para 5% e 20%, respectivamente. Deve ser observado que essas pequenas variações de retardo (*jitter*) podem ser retiradas pelo *play-out buffer* no terminal do receptor.
 - 2) O número de fluxos em um agregado afeta o retardo e o *jitter* dos fluxos individuais dentro do agregado. Enquanto o número de fluxos dentro do agregado aumenta, o *jitter* do fluxo individual também aumenta. Em um agregado heterogêneo, o retardo aumenta com o aumento do número de fluxos.
 - 3) A variação do tamanho dos pacotes de um fluxo afeta o *jitter* e o retardo experimentado por esse fluxo. No agregado de tráfego homogêneo, um fluxo com tamanho de pacote grande experimenta retardos (e *jitter*) mais longos em comparação com o retardo (*jitter*) experimentado por um fluxo de tamanho de pacote menor. No agregado de tráfego heterogêneo, os fluxos de tamanho de pacotes pequeno experimentam *jitter* mais longos em comparação ao *jitter* experimentado por um fluxo de tamanho de pacote maior. Mas o fluxo com tamanho de pacote grande experimenta *jitter* mais longo em termos de percentagem do tempo da taxa de pacotes subscrita quando comparado ao fluxo de tamanho de pacote menor.
 - 4) Um agregado de tráfego homogêneo ou de tráfego heterogêneo, quando transportado dentro de uma classe EF PHB, experimenta retardo e *jitter* que estão dentro dos limites para boa qualidade de conversação.
- [CC2000] faz análise de 70 trilhas empregando codificador G.711 de VoIP. Consta que as distribuições obtidas divergem do modelo exponencial tradicional, sendo observadas distribuições de cauda pesada para os períodos ON/OFF. Verifica que a distribuição teórica que melhor representa a distribuição real nos períodos de ON e de OFF é a distribuição de Weibull.
 - [JW2000] investiga as distribuições dos períodos ON/OFF de VoIP produzidas pelos codificadores G.729 com detector de atividade de voz. Os resultados indicam que as

distribuições ON/OFF são bastante sensíveis tanto ao volume do som quanto ao tipo de detector de silêncio e que essas distribuições são mais cauda pesada do que o modelo tradicional da distribuição exponencial.

- [NE2001] investiga uma arquitetura de acesso IP que tem disciplinas de fila FCFS (*First Come First Served*), WFQ (*Weighted Fair Queuing*) e SP (*Strict Priority*). Emprega uma distribuição de Pareto para modelar o tráfego IP (Web, http, FTP, VoIP).
- [BB2002] investiga o emprego de VoIP em redes móveis GPRS (*General Packet Radio Service*). A partir de 25 trilhas de voz codificadas com G.729 constrói um modelo ON/OFF onde os períodos ON obedecem a distribuição Gamma e os períodos OFF obedecem a distribuição de Weibull.
- [CE2002], a partir de trilhas de VoIP empregando vários codificadores investiga o comportamento do modelo ON/OFF. O resultado indica que a distribuição Lognormal produz melhor aderência tanto para os períodos ON quanto para os períodos OFF para todos os codificadores.
- [LX2003] faz um estudo sobre a agregação de fluxos ON/OFF de cauda pesada e mostra que esses fluxos podem exibir multifractalidade em adição ao comportamento de LRD. Assume que os períodos ON e OFF obedecem a distribuição de Pareto.
- [SJ2003] estuda a modelagem ON/OFF para agregação de tráfego VoIP a partir de parâmetros de teletráfego. Examina os codificadores de voz G.711, G.723.1 e G.729.
- [DT2003] realizada uma análise fractal do tráfego VoIP tanto ao nível de chamada quanto ao nível de pacotes a partir de trilhas de voz empregando codificadores G.723.1 e G.729B. Constata que as distribuições reais se aproximam da distribuição de Pareto nos períodos de ON e de OFF, tanto para o tempo de retenção das chamadas quanto para os tempos interpacotes. Estuda o comportamento fractal dos fluxos agregados. Discute a propriedade auto-similar no tráfego VoIP.

- [EA2004] propõe para o serviço VoIP um novo modelo de geração de quadro para os codificadores G.729B e G.723.1 que empregam os quadros SID.
- [AA2003] realiza uma análise sobre procedimentos de medição ativa e passiva e relata cuidados que devem ser tomados quando da realização dessas medições.
- [MF2005] faz um tutorial sobre a medição da desempenho de redes. Explana os princípios das medições ativas.

A concepção da idéia do presente trabalho é uma junção do conceito de medição ativa em tempo real empregando um fluxo de controle, com a possibilidade de obter desempenho semelhante entre um fluxo de controle e os demais fluxos VoIP que estão multiplexados dentro do mesmo agregado EF PHB.

Não foi encontrado trabalho com finalidade semelhante ao agora proposto, mas foram encontrados alguns trabalhos que contribuíram com fragmentos para a concepção final. Assim, em [TA2000] através do emprego de tráfego homogêneo e tráfego heterogêneo, foi verificada a influência do tamanho dos pacotes, do número de fluxos e das taxas de transmissão dos codificadores nas características do agregado, observadas no nó de egresso, pela avaliação das métricas retardo e *jitter*. Porém, o trabalho só apresentou comparações qualitativas que não forneceram subsídios para inferir sobre o comportamento entre os fluxos. Assim, desse trabalho foi aproveitada a concepção de tráfego, a forma de carregamento, o uso de EF PHB com até 30% de carga no enlace de saída do roteador de ingresso.

Para a determinação do modelo de tráfego ON/OFF, para estudo do funcionamento dos codecs e para a escolha das distribuições foram analisados os trabalhos [JW2000], [NE2001], [BB2002], [CE2002], [LX2003], [SJ2003] e [EA2004]. Vários desses artigos citam que a distribuição Exponencial, não consegue refletir os efeitos de LRD para VoIP. Aqueles trabalhos que relataram simulações não forneceram dados para serem empregados no presente trabalho. Assim, foram adotados os dados obtidos em [DT2003] para a distribuição de Pareto empregando os codecs G.723.1 e G.729B e os dados obtidos em [CC2000] para a distribuição de Weibull empregando o codec G.711.

O emprego das distribuições de Pareto e de Weibull para modelamento do tráfego VoIP exigiu estudo específico dos parâmetros envolvidos, conforme descrito no Anexo 1.

Em [AA2003] e [MF2005] foram descritos os conceitos de medição ativa através da inserção de fluxo sintético, porém não ficou claro se a posição do fluxo de controle na fila tem

importância ou não sobre os resultados. Com a limitação em 30% da carga do enlace de saída para o agregado EF PHB, ficou garantida a não sobrecarga da rede pelo fluxo de controle.

Então, a questão a ser respondida é se existe aderência das métricas (retardo e *jitter*) obtidas no fluxo de controle em relação aos demais fluxos do agregado, tendo em vista que a perda de pacotes esperada é nula na topologia proposta para a simulação.

Em princípio, para a condição de tráfego homogêneo é esperado um comportamento do fluxo de controle, traduzido pelas métricas retardo e *jitter*, mais aderente ao comportamento dos demais fluxos do que na condição de tráfego heterogêneo. Além disso, na condição de tráfego heterogêneo, conforme os resultados de [TA2000], a relação entre os tamanhos dos pacotes do fluxo de controle e dos demais fluxos VoIP, bem como a quantidade relativa entre os tipos de fluxos, deve influenciar os resultados.

A disciplina de fila adotada, FCFS, pode influenciar os resultados na condição de tráfego heterogêneo, porém não será avaliado esse item tendo em vista que seria necessário um conjunto de simulações específico envolvendo outras disciplinas.

CAPÍTULO 4

DESENVOLVIMENTO DO TRABALHO

Neste capítulo é descrita e apresentada uma proposta de solução para o problema da avaliação do desempenho em tempo real de fluxos pertencentes à mesma classe EF PHB de serviços diferenciados, transportando pacotes do serviço VoIP que estão multiplexados dentro de um fluxo agregado, através de medições em um fluxo de controle inserido no agregado para tal finalidade. A avaliação de desempenho consiste na verificação e comparação das métricas de maior importância do SLA, tais como: retardo na entrega de pacotes e variação do retardo na entrega dos pacotes (*jitter*). Assim sendo, o fluxo de controle inserido no roteador de ingresso da rede DiffServ é multiplexado juntamente com os demais fluxos VoIP formando o fluxo agregado EF PHB e no roteador de egresso da rede, esse fluxo de controle é verificado, servindo seu desempenho como uma indicação quanto ao desempenho dos demais fluxos VoIP, que estão no agregado em questão.

Considerando que, na prática, para obtenção de VoIP, é possível o emprego de vários tipos de codificadores de voz que funcionam de forma diferente, gerando taxas de transmissão diferentes e tamanhos de pacotes diferentes, é necessário simular o sistema em duas condições: com tráfego homogêneo e com tráfego não homogêneo. Na condição de tráfego homogêneo, todos os fluxos VoIP têm a mesma característica, ou seja, empregam o mesmo tipo de codificador e na condição de tráfego heterogêneo, ocorre a combinação de diversos tipos de codificadores, que no presente caso, serão apenas dois tipos de *codecs* VoIP por etapa de simulação. Portanto, o fluxo de controle deve operar nessas duas condições de tráfego e ser capaz de fornecer informações, com precisão adequada, sobre o desempenho dos demais fluxos e conseqüentemente, se o SLA está sendo atendido.

O presente Capítulo está organizado da seguinte forma: na seção 4.1 é apresentada a motivação para o presente trabalho, na seção 4.2 é apresentado o modelo sistêmico da rede DiffServ, na seção 4.3 é detalhada a proposta de trabalho, na seção 4.4 é elencado o conjunto de premissas do trabalho, na seção 4.5 é descrito o ambiente de simulação, na seção 4.6 são indicados os detalhes da implementação do simulador OPNET Modeller 10.0 e na seção 4.7 são apresentados os resultados obtidos nas simulações.

4.1 MOTIVAÇÃO DO TRABALHO

Quando da formação dos fluxos agregados em uma classe EF PHB pode ocorrer a degradação dos requisitos de QoS à medida que forem sendo aglutinados mais fluxos ou a ocorrência de eventos de congestionamento no interior da rede. A motivação deste trabalho é examinar a possibilidade do emprego de um fluxo sintético de controle, injetado no nó de ingresso da rede juntamente com os demais fluxos VoIP, tendo a função de informar em tempo real, no nó egresso do domínio, a aderência ou não aos valores de QoS definidos no SLS e no TCS dos demais fluxos constituintes do agregado. As informações obtidas, se confirmada a validade da correspondência do resultado de um fluxo de controle para os demais fluxos do agregado, auxiliariam a tomada de decisão para agregar ou não mais fluxos no enlace de saída do roteador do nó de ingresso ou prover mais banda nos enlaces do núcleo da rede.

A criação de um fluxo de controle se trata de um procedimento de simples implementação, podendo servir aos pequenos provedores de serviço IP (*Internet Service Provider – ISP*) como uma ferramenta para auxiliar decisões sobre a carga e desempenho da rede.

4.2 MODELO SISTÊMICO PROPOSTO

A Figura 4.1 mostra a configuração sistêmica da rede a ser simulada. A descrição é apresentada no decorrer dos próximos itens.

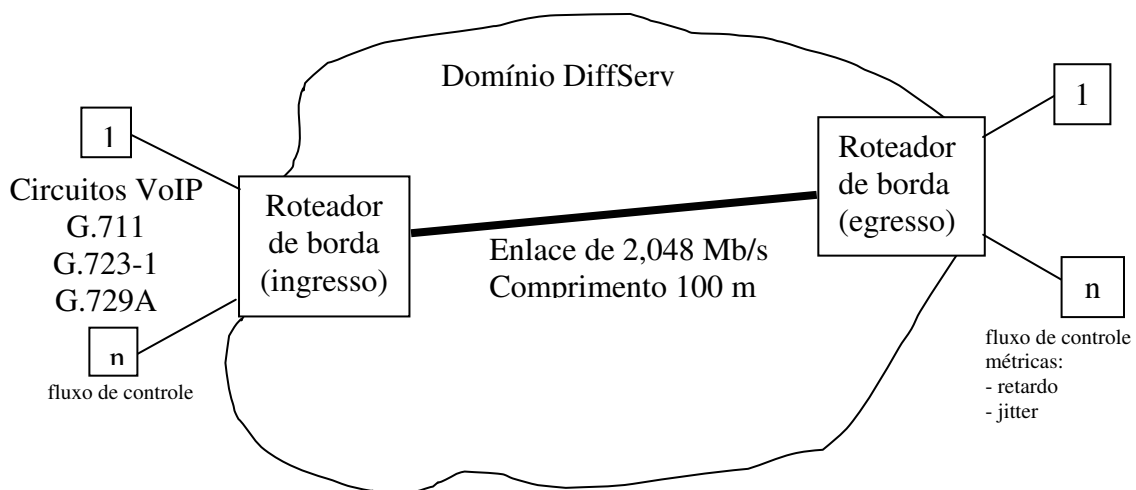


Figura 4.1: Diagrama sistêmico da rede sob simulação

4.3 PROPOSTA DE TRABALHO

Simular um ambiente DiffServ empregando EF PHB exclusivo para serviço VoIP tendo um enlace E1 (2,048 Mb/s) entre os roteadores de ingresso e de egresso e modelar um tráfego com cauda pesada, originado por fontes ON/OFF, tendo codificadores de voz do tipo G.711, G.723.1 e G.729B, considerando tráfego homogêneo e tráfego heterogêneo. Obter os dados das métricas de QoS como retardo e de variação do retardo (*jitter*) no nó de egresso de todos os fluxos, comparando com os respectivos valores das métricas do fluxo de controle.

A proposta acima tem como aplicação a supervisão em tempo real do desempenho de fluxos VoIP de forma a verificar a aderência ao SLA dos valores obtidos nas métricas, visando auxiliar no processo de decisão quanto à aceitação ou não de mais fluxos VoIP no fluxo agregado ou a ampliação da capacidade do *core*.

O modelo de tráfego ON/OFF adotado segue resultados de trabalhos realizados onde, a partir de trilhas de conversações gravadas, empregando os codificadores tipo G.711, G.723.1 e G.729B, foram obtidas as respectivas distribuições. Assim, em [CC2000], a partir de 70 trilhas de diversos serviços de voz empregando o *codec* G.711, sendo 24 trilhas de conversações, foi obtida a distribuição de Weibull; em [DT2003], a partir de um grande número de ligações empregando *codec* G.723.1 ou G.729B foram obtidas duas distribuições de Pareto diferentes para cada codificador, uma para os períodos de surtos de voz e outra para os períodos de silêncio.

A questão chave é se o fluxo de controle fornece informações como retardo e variação do retardo que poderão ser estendidas aos demais fluxos do agregado. Espera-se que a condição mais difícil de atender seja aquela com tráfego heterogêneo. Além disso, quando o número de fluxos cresce pode haver degradação de retardo ou da variação do retardo não uniforme em função da disciplina de fila. Da mesma forma, a relação entre os tamanhos dos pacotes e a proporção de fluxos dentro do agregado pode causar degradação de retardo ou da variação do retardo não uniforme.

A maior dificuldade prevista para conclusão do presente trabalho é a forma de avaliação para caracterizar se há aderência entre os valores das métricas obtidas no fluxo de controle com os valores obtidos nos demais fluxos. Ou seja, até quanto um valor pode se afastar do outro de forma que ainda haja consenso de que são valores semelhantes e que com isso, a leitura do valor indicado pela métrica do fluxo de controle espelha o comportamento dos demais fluxos do agregado.

4.4 PREMISSAS

As premissas assumidas no desenvolvimento do presente trabalho são:

- a. O EF PHB é destinado somente para tráfego VoIP [TA2000].
- b. Taxa de transmissão de 2,048 Mb/s (E1) no enlace de ligação entre os roteadores de ingresso e de egresso do domínio DiffServ com distância de até 100 m. Com essa distância é eliminado do elenco de variáveis o retardo de propagação. Os pequenos ISPs operam com taxas *E1* ou com *n E1*.
- c. O limite de 30% da capacidade do enlace para EF PHB é um dado prático, também utilizado em [TA2000]. Com essa limitação de carga são evitados os congestionamentos no enlace.
- d. Considerando que no máximo 30% da capacidade do enlace entre roteadores está sendo ocupada com tráfego EF poderia ser pensado no completamento da carga até 100% com tráfego BE. Tendo em vista que o tráfego EF tem prioridade, caso ele necessite de mais banda numa condição de congestionamento da rede, esta banda será automaticamente fornecida pelo descarte de pacotes do tráfego BE. Como o objetivo do trabalho é verificar o comportamento do agregado do tráfego EF, não será utilizada na simulação uma carga complementar com tráfego BE.
- e. A disciplina de fila adotada é a FCFS. Esse é o pior caso para a simulação tendo em vista a condição de tráfego heterogêneo. Alguns autores utilizaram disciplinas PQ, CBQ ou WFQ. A influência da disciplina de fila poderá ser verificada em trabalhos futuros.
- f. O comprimento de fila do *buffer* de 500 pacotes é o valor *default* do simulador [OPNET].
- g. Os *codecs* de voz escolhidos, G.711, G.723.1 e G.729B, são muito utilizados comercialmente e também são aqueles que produzem maior diferença de tamanho de pacotes [TA2000] [DT2003], possibilitando o pior caso para a simulação de tráfego heterogêneo.
- h. O tráfego de pacotes assumido para as fontes ON/OFF tem comportamento LRD, tendo parâmetro de Hurst entre $(0,5 < H < 1)$ [TM1997] [DT2003].
- i. As distribuições de Weibull [CC2000] e de Pareto [DT2003] para os períodos ON/OFF são distribuições sub-exponenciais apresentando comportamento de cauda pesada.

4.5 AMBIENTE DE SIMULAÇÃO

Durante a simulação foi empregado um *laptop* com as seguintes características:

- processador Pentium 4[®]; relógio 2,4 GHz;
- RAM (*Random Access Memory*) de 512 MB;
- disco rígido 60 MB;
- Sistema Operacional Windows XP[®].

4.6 DESCRIÇÃO DO PROCESSO DE SIMULAÇÃO

Com o auxílio do simulador OPNET Modeller 10.0, inicialmente é criado o ambiente de um domínio DiffServ tendo um nó de ingresso e outro de egresso, ligados por um enlace de 2,048 Mb/s, considerando um enlace de escritório com distância máxima de 100 m.

A fila no roteador de ingresso obedece a disciplina FCFS com tamanho de fila de 500 pacotes.

Em conformidade com [TA2000] a taxa de bits do enlace de saída do roteador de ingresso será limitada em 30% (614,4 Kb/s) e disponibilizada para VoIP usando EF PHB.

Os codificadores e decodificadores de voz têm as características conforme mostrado na Tabela 4.1.

Tabela 4.1: Parâmetros dos *codecs* para VoIP utilizados nas simulações

parâmetro	G.711	G.723.1	G.729B
Frame size [ms]	20	30	10
Codec rate [Kb/s]	64	5,3	8
VAD ativa (supressão de silêncio)	ok	ok	ok

Os parâmetros das distribuições de tráfego VoIP empregadas nas simulações são:

- codec G.711
Distribuição de Weibull conforme [CC2000] tendo para os períodos ON e de OFF: *scale parameter* $a = 1,05 s$ e *shape parameter* $b = 0,335$.
- codec G.723.1
Distribuição de Pareto conforme [DT2003] tendo período ON com média $2,2822 s$, variância $12,77 s^2$, *location parameter* $c = 1,238448$ e *shape*

parameter a = 2,1865 e período OFF com média 1,4849 s, variância 5,9242 s²; location parameter c = 0,801057 e shape parameter a = 2,171405.

- codec G.729B

Distribuição de Pareto conforme [DT2003] tendo período ON com média 2,3651 s, variância 18,3193, *location parameter c = 1,261211 e shape parameter a = 2,142517* e períodos de OFF com média 1,5621 s, variância 5,192, *location parameter c = 0,856043 e shape parameter a = 2,212429.*

A simulação foi dividida em dois cenários: o primeiro usando tráfego homogêneo e o segundo usando tráfego heterogêneo. Cada cenário é formado por três fases de simulação. No cenário de tráfego homogêneo, a primeira fase só emprega voz codificada com G.711, a segunda fase só emprega voz codificada com G.723.1 e a terceira fase só emprega voz codificada com G.729B. No cenário de tráfego heterogêneo, a primeira fase emprega uma mistura de voz codificada com G.711 e com G.723.1, a segunda fase emprega uma mistura de voz codificada com G.711 e com G.729B e a terceira fase emprega uma mistura de voz codificada com G.729B e com G.723.1.

Em cada etapa do cenário de tráfego homogêneo, os fluxos são inseridos um a um até o completamento de 30% da carga do enlace E1 destinada ao EF PHB.

O primeiro fluxo inserido é o fluxo de controle. Na prática, por questões operacionais esse seria o procedimento para evitar novos trabalhos e alteração nas tabelas de conexão. Eventualmente, poderia se questionar a existência da posição do fluxo de controle ótimo, aquele que desse o melhor resultado de aderência aos demais fluxos. Essa questão está fora do escopo do presente trabalho. Intuitivamente pode-se aceitar a idéia porém, na prática, essa procura causaria transtornos e gasto de tempo que não justificariam a procura tendo em vista a aleatoriedade temporal do processo e a quantidade de variáveis envolvidas.

Em cada passo de incremento do número de fluxos são obtidas as métricas: retardo *jitter* e utilização do enlace.

No cenário de tráfego heterogêneo, é considerado o número de fluxos VoIP com codificadores G.711 para a condição de 30% de carga do enlace sendo então retirado um fluxo G.711 e colocado o número necessário de fluxos com codificador G.723.1. A seguir o processo é repetido retirando de um em um fluxo G.711 até restar apenas um e os demais fluxos sendo G.723.1. Na segunda fase, o processo repete a primeira fase utilizando o

codificador G.729B no lugar do codificador G.723.1. Na terceira fase é considerado o número de fluxos VoIP com codificadores G.729B para a condição de 30% de carga do enlace sendo então retirado um fluxo G.729B e colocado o número necessário de fluxos com codificador G.723.1. A seguir o processo é repetido retirando de um em um fluxo G.729B até restar apenas um e os demais fluxos sendo G.723.1. Em cada passo da simulação são obtidas as métricas: retardo e *jitter*.

Para obtenção do nível de confiança mínimo de 95%, cada rodada de simulação foi replicada 10 vezes, com sementes variando de 128 até 704 em passos de 64. O período de cada simulação é de 600 segundos, sendo definidos os seguintes parâmetros no simulador: *start time* de 20 segundos e *start time offset* entre 5 e 30 segundos. Ou seja, no período entre 25 e 50 segundos todas as conexões VoIP são realizadas. O período de obtenção das métricas ocorre a partir de 50 segundos até 600 segundos.

4.7 RESULTADOS DAS SIMULAÇÕES

Tendo em vista que em todas as fases dos cenários da simulação, a perda de pacotes é praticamente nula, são apresentados somente os resultados das métricas retardo e variação do retardo (*jitter*).

Cada processo de simulação tem duração de 600 segundos, sendo repetido 10 vezes. Como resultado, cada fluxo do agregado, apresenta variações temporais nos valores da métrica analisada (retardo e *jitter*) que são representados por uma média, um máximo e um mínimo em cada condição de carga do agregado. Assim, para uma carga de i fluxos VoIP, os valores de retardo e de *jitter* estão dispersos no tempo ao redor dos correspondentes valores do fluxo de controle. Então, para capturar estas relações é feita uma normalização em relação ao correspondente valor da métrica do fluxo de controle em cada condição de carga i , sendo obtidos três conjuntos de dados:

- variação dos valores mínimos normalizados para retardo e variação do retardo;
- variação dos valores médios normalizados para retardo e variação do retardo;
- variação dos valores máximos normalizados para retardo e variação do retardo.

A variação dos valores mínimos foi desprezada tendo em vista não ter significado prático, além de ocasionar, para o caso da métrica *jitter*, muitos erros de divisão por zero quando do processo de normalização.

Exemplificando para a condição de tráfego homogêneo empregando codificador C. A média da métrica (retardo ou *jitter*) tem um valor em cada um dos i fluxos VoIP do agregado, que resume os 10 ciclos de simulação de 600 segundos. Dentro desse conjunto de i médias temos o valor mínimo da média, o valor médio das médias e o valor máximo da média. Por outro lado, a correspondente métrica no fluxo de controle apresenta um valor mínimo, um valor médio e um valor máximo em cada condição de carga (i fluxos no agregado). Então, para capturar essas relações é feita a normalização da métrica em relação ao correspondente valor do fluxo de controle, sendo obtidos:

- 1) o valor mínimo da média da métrica para os i fluxos, normalizado em relação ao valor mínimo da média da mesma métrica no fluxo de controle na condição de i fluxos, representado nas figuras dos anexos por *valor mín do retardo (jitter) médio*;
- 2) o valor máximo da média da métrica para os i fluxos, normalizado em relação ao valor médio da mesma métrica no fluxo de controle na condição de i fluxos, representado nas figuras dos anexos por *valor máx do retardo (jitter) médio*.

A indicação *mínimo* e *máximo* serve para informar que, no caso, as médias dos i fluxos estão dispersas entre esses valores em cada condição de carga.

Para obtenção do valor absoluto de cada métrica na condição de i fluxos, basta multiplicar os valores normalizados pelo valor absoluto da correspondente métrica do fluxo de controle para a condição de i fluxos.

Os resultados das simulações referentes ao cenário de tráfego homogêneo estão descritos no Anexo 3 (codec G.711), no Anexo 4 (codec G.723.1) e Anexo 5 (codec G.729B) e os resultados para o tráfego heterogêneo estão no Anexo 6 (codec G.711 e codec G.723.1), no Anexo 4 (codec G.723.1) e Anexo 5 (codec G.729B).

Com auxílio do coeficiente de correlação, explicado no Anexo 2, são cotejados os resultados referentes aos valores médios da métrica retardo (*jitter*) indicados pelo fluxo de controle e os respectivos valores médios obtidos para as mesmas métricas em cada fluxo do agregado, sendo então obtido um valor entre 0 e 1 que indica uma menor ou uma maior aderência, respectivamente, entre as distribuições de valores. Ou seja, um coeficiente de correlação próximo de 1 significa que os valores da métrica considerada estão muito bem retratados pelas indicações do fluxo de controle.

4.7.1 Resultados do cenário de tráfego homogêneo

Na Fase 1, empregando codificadores G.711, a carga de 30% de 2,048 Mb/s foi atingida com 16 fluxos VoIP mais o fluxo de controle. No Anexo 3 constam os resultados das distribuições do retardo (e do *jitter*) no fluxo de controle em função do número de fluxos VoIP no agregado e as variações do retardo (e do *jitter*) desde 1 até 16 fluxos no agregado, normalizadas em relação ao valor do retardo (*jitter*) do fluxo de controle.

Os resultados da variação da métrica retardo normalizado às diversas condições de carga, considerando tráfego homogêneo G.711, dados nas Figuras A3.7 e A3.8 do Anexo 3, indicam que os valores médios estão no intervalo $[0,80; 1,23]$ e os valores máximos estão no intervalo $[0,54; 1,54]$. O coeficiente de correlação calculado entre as médias da métrica retardo no fluxo de controle G.711 e nos demais fluxos de tráfego homogêneo indicou 0,3786.

Por outro lado, os resultados da variação do *jitter* normalizado, mostrados nas Figuras A3.9 até A3.10 do Anexo 3, indicam que os valores médios estão no intervalo $[0,02; 10,2]$ e os valores máximos estão no intervalo $[0,05; 10,4]$. O coeficiente de correlação calculado sobre os valores médios da métrica *jitter* no fluxo de controle G.711 de tráfego homogêneo indicou 0,4282.

Conforme visto nas Figuras A3.1 até A3.3 do Anexo 3, o retardo indicado no fluxo de controle tem ordem de grandeza de $10^{-3}s$ enquanto que os valores absolutos de *jitter*, mostrados nas Figuras A3.4 até A3.6 do Anexo 3, tem picos com ordem de grandeza de $10^{-7}s$.

Na Fase 2, empregando codificadores G.723.1, a carga de 30% de 2,048 Mb/s foi atingida com 80 fluxos VoIP mais o fluxo de controle. No Anexo 4 constam os resultados das distribuições da métrica retardo (e da métrica *jitter*) no fluxo de controle em função do número de fluxos VoIP no agregado e as variações da métrica retardo (e da métrica *jitter*) desde 1 até 80 fluxos no agregado, normalizadas em relação ao valor do retardo (*jitter*) do fluxo de controle.

Os resultados da variação da métrica retardo normalizado às diversas condições de carga, considerando tráfego homogêneo G.723.1, dados nas Figuras A4.7 e A4.8 do Anexo 4, indicam que os valores médios estão no intervalo $[0,91; 1,06]$ e os valores máximos estão no intervalo $[0,71; 2,00]$. O coeficiente de correlação calculado entre as médias da métrica retardo no fluxo de controle G.723.1 e nos demais fluxos de tráfego homogêneo indicou 0,8718.

Por outro lado, os resultados da variação do *jitter* normalizado, mostrados nas Figuras A4.9 e A4.10 do Anexo 4, indicam que os valores médios estão no intervalo $[0,31; 5,30]$ e os valores máximos estão no intervalo $[0,11; 14,75]$. O coeficiente de correlação calculado sobre os valores médios da métrica *jitter* no fluxo de controle G.711 de tráfego homogêneo indicou 0,6780.

Conforme visto nas Figuras A4.1 até A4.3 do Anexo 4, o retardo indicado no fluxo de controle tem ordem de grandeza de $10^{-4}s$ enquanto que os valores absolutos de *jitter*, mostrados nas Figuras A4.4 até A4.6 do Anexo 4, tem picos com ordem de grandeza de $10^{-8}s$.

Na Fase 3, empregando codificadores G.729B, a carga de 30% de 2,048 Mb/s foi atingida com 35 fluxos VoIP mais o fluxo de controle. No Anexo 5 constam os resultados das distribuições do retardo (e do *jitter*) no fluxo de controle em função do número de fluxos VoIP no agregado e as variações da métrica retardo (e da métrica *jitter*) desde 1 até 35 fluxos no agregado, normalizadas em relação ao valor do retardo (*jitter*) do fluxo de controle.

Os resultados da variação da métrica retardo normalizado às diversas condições de carga, considerando tráfego homogêneo G.729B, dados nas Figuras A5.7 e A5.8 do Anexo 5, indicam que os valores médios estão no intervalo $[0,94; 1,06]$ e os valores máximos estão no intervalo $[0,75; 1,84]$. O coeficiente de correlação calculado entre as médias da métrica retardo no fluxo de controle G.729B e nos demais fluxos de tráfego homogêneo indicou 0,8341.

Por outro lado, os resultados da variação do *jitter* normalizado, mostrados nas Figuras A5.9 e A5.10 do Anexo 5, indicam que os valores médios estão no intervalo $[0; 3,07]$ e os valores máximos estão no intervalo $[0; 5,18]$. O coeficiente de correlação calculado sobre os valores médios da métrica *jitter* no fluxo de controle G.729B de tráfego homogêneo indicou 0,7748.

Conforme visto nas Figuras A5.1 até A5.3 do Anexo 4, o retardo indicado no fluxo de controle tem ordem de grandeza de $10^{-4}s$ enquanto que os valores absolutos de *jitter*, mostrados nas Figuras A5.4 até A5.6 do Anexo 5, tem picos com ordem de grandeza de $10^{-8}s$.

4.7.2 Resultados do cenário de tráfego heterogêneo

Na Fase 1, empregando codificadores G.711 e G.723.1, foram simuladas as configurações desde 15 fluxos G.711, com 1 fluxo G.723.1, sendo então retirado um fluxo

G.711 e acrescidos de um até cinco fluxos G.723.1 até alcançar 1 fluxo G.711 com 1 até 70 fluxos G.723.1. Para verificar a melhor condição de fluxo de controle, foram utilizados um fluxo de controle G.711 e outro G.723.1 em todos os casos.

No Anexo 6, são mostrados os resultados das simulações para cada fluxo de controle, G.711 e G.723.1. Os resultados obtidos estão resumidos na Tabela 4.2.

Tabela 4.2: Resumo dos valores das métricas – tráfego heterogêneo G.711/G.723.1

Fluxo de controle	Extremos da variação	métrica			
		Retardo normalizado		Jitter normalizado	
		média	máximo	média	máximo
G.711	Valor mínimo	0,27	0,22	0	0
	Valor máximo	1,7	2,66	359,5	317,7
G.723.1	Valor mínimo	0,83	0,42	0	0
	Valor máximo	4,53	4,15	131	37,3

O coeficiente de correlação calculado sobre os valores médios da métrica retardo no fluxo de controle G.711 indicou 0,2429 e o coeficiente de correlação calculado sobre os valores médios da métrica *jitter* indicou 0,9193.

O coeficiente de correlação calculado sobre os valores médios da métrica retardo no fluxo de controle G.723.1 indicou 0,3232 e o coeficiente de correlação calculado sobre os valores médios da métrica *jitter* indicou 0,9193.

Conforme visto nas Figuras A6.3 e A6.13 do Anexo 6, o retardo máximo medido nos fluxos de controles G.711 e G.723.1, respectivamente, tem ordem de grandeza $10^{-3}s$ enquanto que os valores absolutos de *jitter*, mostrados nas Figuras A6.6 e A6.15 do Anexo 6, respectivamente, são muito pequenos porém apresentam um pico com ordem de grandeza de $10^{-3}s$ e $10^{-4}s$. Dada a quantidade de medidas realizadas e o número de ocorrências de tais valores, esses picos podem ser ignorados.

Comparando as Figuras A6.7 e A6.8 com as Figuras A6.16 e A6.17, percebe-se que há muito pouca variação no retardo normalizado em relação ao fluxo de controle G.711 e ao fluxo de controle G.723.1, respectivamente, porém, à medida que é reduzida a quantidade de fluxos G.711 e aumentada a quantidade de fluxos G.723.1 no agregado, em ambos os casos, o valor do retardo normalizado aumenta lentamente. A diferença é que os valores normalizados para o fluxo de controle G.723.1 são pouco maiores do que para o fluxo de controle G.711. Isso mostra que a quantidade de fluxos G.723.1 está influenciando nos valores da métrica retardo.

Comparando as Figuras A6.9 e A6.10 com as Figuras A6.18 e A6.19, percebe-se que à medida que é reduzida a quantidade de fluxos G.711 e aumentada a quantidade de fluxos G.723.1 ocorre um aumento no valor do *jitter* normalizado em relação ao valor do *jitter* do respectivo fluxo de controle. A diferença é que os valores para o fluxo de controle G.711 são muito maiores do que para o fluxo de controle G.723.1. Isso significa que, sob o ponto vista do *jitter*, o fluxo de controle G.711 não consegue capturar o comportamento do *jitter* dos demais fluxos G.723.1 e conseqüentemente, ele deve ser preterido em relação ao fluxo de controle G.723.1.

Na Fase 2, empregando codificadores G.711 e G.729B, foram simuladas as configurações desde 15 fluxos G.711 com 1 fluxo G.729B, 14 fluxos G.711 com 1 até 4 fluxos G.729B, sendo então retirado um fluxo G.711 e acrescidos dois fluxos G.729B até alcançar 1 fluxo G.711 com 1 até 30 fluxos G.729B. Foram utilizados dois fluxos de controle, um do tipo G.711 e outro do tipo G.729B.

No Anexo 7, são mostrados os resultados das simulações para cada fluxo de controle, G.711 e G.729B. Os resultados obtidos estão resumidos na Tabela 4.3.

Tabela 4.3: Resumo dos valores das métricas – tráfego heterogêneo G.711/G.729B

Fluxo de controle	Extremos da variação	métrica			
		Retardo normalizado		Jitter normalizado	
		média	máximo	média	máximo
G.711	Valor mínimo	0,27	0,14	0	0,02
	Valor máximo	3,97	7,97	77,6	353
G.729B	Valor mínimo	0,81	0,01	0	0,01
	Valor máximo	3,85	4,70	7,44	16,4

O coeficiente de correlação calculado sobre os valores médios da métrica retardo no fluxo de controle G.711 indicou 0,2878 e o coeficiente de correlação calculado sobre os valores médios da métrica *jitter* indicou 0,8655.

O coeficiente de correlação calculado sobre os valores médios da métrica retardo no fluxo de controle G.729B indicou 0,4371 e o coeficiente de correlação calculado sobre os valores médios da métrica *jitter* indicou 0,8655.

Conforme visto nas Figuras A7.3 e A7.13 do Anexo 7, o retardo máximo medido nos fluxos de controles G.711 e G.729B, respectivamente, tem ordem de grandeza 10^{-3} s enquanto que os valores absolutos de *jitter*, mostrados nas Figuras A7.6 e A7.15 do Anexo 7,

respectivamente, são muito pequenos porém apresentam um pico com ordem de grandeza de $10^{-3}s$ e $10^{-4}s$. Dada a quantidade de medidas realizadas e o número de ocorrências de tais valores, esses picos podem ser ignorados.

Comparando as Figuras A7.7 e A7.8 com as Figuras A7.16 e A7.17, percebe-se que há muito pouca variação no retardo normalizado em relação ao fluxo de controle G.711 e ao fluxo de controle G.729B, respectivamente, porém, à medida que é reduzida a quantidade de fluxos G.711 e aumentada a quantidade de fluxos G.729B no agregado, em ambos os casos, o valor do retardo normalizado aumenta muito lentamente. A diferença é que os valores normalizados para o fluxo de controle G.729B são pouco maiores do que para o fluxo de controle G.711. Isso mostra que a quantidade de fluxos G.729B está influenciando pouco nos valores da métrica retardo.

Comparando as Figuras A7.9 e A7.10 com as Figuras A7.18 e A7.19, percebe-se que à medida que é reduzida a quantidade de fluxos G.711 e aumentada a quantidade de fluxos G.729B ocorre um aumento no valor do *jitter* normalizado em relação ao valor do *jitter* do respectivo fluxo de controle. A diferença é que os valores para o fluxo de controle G.711 são muito maiores (pelo menos uma ordem de grandeza) do que para o fluxo de controle G.729B. Isso significa que, sob o ponto vista do *jitter*, o fluxo de controle G.711 não consegue capturar o comportamento do *jitter* dos demais fluxos G.729B e conseqüentemente, ele deve ser preterido em relação ao fluxo de controle G.729B.

Além disso, comparando os resultados obtidos nas Fases 1 e 2, o fluxo de controle G.729B forneceu melhores indicações do comportamento de pico para o *jitter* do que o fluxo de controle G.723.1, em condições de tráfego heterogêneo com codificadores G.711.

De qualquer forma, ainda não está evidente se a variação de *jitter* é função somente da quantidade de fluxos e se a relação entre os tamanhos de pacotes tem influência no processo. Nesse momento poderia ser verificado se a mudança da disciplina de fila traria melhora no desempenho do fluxo de controle em relação a indicação dos valores dos demais fluxos do agregado.

Na Fase 3, empregando codificadores G.729B e G.723.1, foram simuladas as configurações desde 34 fluxos G.729B com 1 fluxo G.723.1, sendo então retirado um fluxo G.729B e acrescentados 2 ou 1 fluxo G.723.1 até alcançar a condição de 1 fluxo G.729B com 1 até 48 fluxos G.723.1. Foram utilizados dois fluxos de controle, um do tipo G.729B e outro do tipo G.723.1.

No Anexo 8, são mostrados os resultados das simulações para cada fluxo de controle, G.729B e G.723.1. Os resultados obtidos estão resumidos na Tabela 4.4.

Tabela 4.4: Resumo dos valores das métricas – tráfego heterogêneo G.729B/G.723.1

Fluxo de controle	Extremos da variação	métrica			
		Retardo normalizado		Jitter normalizado	
		média	máximo	média	máximo
G.729B	Valor mínimo	0,92	0,58	0	0
	Valor máximo	2,28	2,14	8,1	12,8
G.723.1	Valor mínimo	0,43	0,43	0	0
	Valor máximo	2,08	2,25	59	84

O coeficiente de correlação calculado sobre os valores médios da métrica retardo no fluxo de controle G.729B indicou $0,19$ e o coeficiente de correlação calculado sobre os valores médios da métrica *jitter* indicou $0,9783$.

O coeficiente de correlação calculado sobre os valores médios da métrica retardo no fluxo de controle G.723.1 indicou $0,2030$ e o coeficiente de correlação calculado sobre os valores médios da métrica *jitter* indicou $0,9878$.

Conforme visto nas Figuras A8.3 e A8.13 do Anexo 8, o retardo máximo medido nos fluxos de controles G.729B e G.723.1, respectivamente, tem ordem de grandeza $10^{-4}s$ enquanto que os valores absolutos de *jitter*, mostrados nas Figuras A8.6 e A8.15 do Anexo 8, respectivamente, são muito pequenos porém apresentam um pico com ordem de grandeza de $10^{-4}s$. Dada a quantidade de medidas realizadas e o número de ocorrências de tais valores, esses picos podem ser ignorados.

Comparando as Figuras A8.7 e A8.8 com as Figuras A8.16 e A8.17, percebe-se que há variação no retardo normalizado em relação ao fluxo de controle G.729B e ao fluxo de controle G.723.1, respectivamente, porém, à medida que é reduzida a quantidade de fluxos G.729B e aumentada a quantidade de fluxos G.723.1 no agregado, em ambos os casos, o valor do retardo normalizado aumenta lentamente. No entanto, não há diferença nos comportamentos. Isso evidencia novamente que a quantidade de fluxos G.723.1 está influenciando nos valores da métrica retardo.

Comparando as Figuras A8.9 e A8.10 com as Figuras A8.18 e A8.19, percebe-se que à medida que é reduzida a quantidade de fluxos G.729B e aumentada a quantidade de fluxos G.723.1 ocorre um aumento no valor do *jitter* normalizado em relação ao valor do *jitter* do respectivo fluxo de controle. A diferença é que os valores para o fluxo de controle G.723.1

são muito maiores do que para o fluxo de controle G.729B. Isso significa que, sob o ponto de vista do *jitter*, o fluxo de controle G.729B tem melhor desempenho do que o fluxo de controle G.723.1.

CAPÍTULO 5

CONCLUSÕES

Neste capítulo são apresentadas as análises, as avaliações e as conclusões sobre os resultados, são explicitadas as contribuições e são sugeridas as possibilidades para desenvolvimento de trabalhos futuros através de idéias geradas durante o presente trabalho.

Os resultados obtidos conduzem às seguintes conclusões:

- i) Considerando os coeficientes de correlação calculados para os valores médios das métricas retardo e *jitter* dados na Tabela 5.1, vemos que o fluxo de controle G.711 apresenta o pior desempenho para a métrica retardo, indicando a influência do tamanho dos pacotes no retardo e também indicando a fraca aderência da métrica retardo do fluxo de controle G.711 em relação aos valores dos demais fluxos do agregado.

Tabela 5.1: Coeficientes de correlação obtidos sobre os valores médios das métricas retardo e *jitter* do fluxo de controle em relação aos demais fluxos do agregado

tipo de tráfego	tipo de codec	coeficiente de correlação do fluxo de controle*					
		métrica retardo			métrica <i>jitter</i>		
		G.711	G.723.1	G.729B	G.711	G.723.1	G.729B
homogêneo	G.711	0,3786	-	-	0,4282	-	-
	G.723.1	-	0,8718	-	-	0,6780	-
	G.729B	-	-	0,8341	-	-	0,7748
heterogêneo	G.711	0,2429	-	-	0,9193	-	-
	G.723.1	-	0,3232	-	-	0,9193	-
	G.711	0,2878	-	-	0,8655	-	-
	G.729B	-	-	0,4371	-	-	0,8655
	G.723.1	-	0,2030	-	-	0,9878	-
	G.729B	-	-	0,19	-	-	0,9783

* método de cálculo apresentado no Anexo 2

- ii) Na condição de tráfego homogêneo, os fluxos de controle G.711, G.723.1 e G.729B representam razoavelmente o desempenho do retardo dos demais fluxos do agregado, pois:

- (1) No agregado VoIP formado por até 16 fluxos G.711 a variação da indicação da métrica retardo normalizado no fluxo de controle se situa no intervalo $[0,80; 1,23]$ para os valores médios e no intervalo $[0,54; 1,54]$ para os valores máximos. Coeficiente de correlação 0,3786.
 - (2) No agregado VoIP formado por até 80 fluxos G.723.1 a variação da indicação da métrica retardo normalizado no fluxo de controle se situa no intervalo $[0,91; 1,06]$ para os valores médios e no intervalo $[0,71; 2,0]$ para os valores máximos. Coeficiente de correlação 0,8718.
 - (3) No agregado VoIP formado por até 35 fluxos G.729B a variação da indicação da métrica retardo normalizado no fluxo de controle se situa no intervalo $[0,95; 1,06]$ para os valores médios e no intervalo $[0,75; 1,84]$ para os valores máximos. Coeficiente de correlação 0,8341.
- iii) Na condição de tráfego homogêneo empregando codificadores G.711, G.723.1 e G.729B a separação entre o valor indicado pelo respectivo fluxo de controle e os valores dos demais fluxos, para a métrica retardo, aumenta com o aumento do número de fluxos. Estes resultados estão em conformidade com os resultados de Tyagi *et al.* publicados em [TA2000].
- iv) Na condição de tráfego heterogêneo, onde o agregado é formado por dois conjuntos de fluxos, tendo cada conjunto um tipo de codificador VoIP, sob o ponto de vista da métrica retardo, cada fluxo de controle apresentou comportamento diferenciado.
- (1) No par G.711/G.723.1, conforme resumo dado na Tabela 4.2, o fluxo de controle empregando o codec G.711 apresentou menores variações na métrica retardo normalizado, mas o coeficiente de correlação obtido foi 0,2429 enquanto que o fluxo de controle com codec G.723.1 embora apresentasse maior variação na métrica retardo normalizado, obteve um coeficiente de correlação melhor, 0,3232.
 - (2) No par G.711/G.729B, conforme resumo dado na Tabela 4.3, o fluxo de controle empregando o *codec* G.711 apresentou maiores variações na métrica retardo normalizado e o coeficiente de correlação obtido foi 0,2878 enquanto que o fluxo de controle com *codec* G.729B apresentou

menor variação na métrica retardo normalizado, obtendo um coeficiente de correlação melhor, $0,4371$.

- (3) No par G.729B/G.723.1, conforme resumo dado na Tabela 4.4, o fluxo de controle empregando o *codec* G.729B apresentou menos variações que os casos anteriores na métrica retardo normalizado, tendo coeficiente de correlação de $0,19$ enquanto que o fluxo de controle com *codec* G.723.1 embora apresentasse maior variação levemente maior do que o fluxo de controle com *codec* G.729B na métrica retardo normalizado, obteve um coeficiente de correlação melhor, $0,3232$.
 - (4) Analisando a aderência da métrica retardo sob o ponto de vista do coeficiente de correlação, o fluxo de controle empregando o *codec* G.723.1 apresenta o melhor desempenho nos diversos cenários.
- v) Quanto a métrica *jitter*, as conclusões referentes a aderência da informação obtida no fluxo de controle em relação aos demais fluxos do agregado são as seguintes:
- (1) Para tráfego homogêneo, a indicação da métrica *jitter* normalizado, dada pelo fluxo de controle, no pior caso, apresenta diferenças de até uma década e meia (*codec* G.723.1). Sob o ponto de vista do coeficiente de correlação, o melhor desempenho foi obtido pelo fluxo de controle com *codec* G.729B com $0,7748$, seguido pelo *codec* G.723.1 com $0,6780$.
 - (2) Para tráfego heterogêneo, considerando todos os cenários, sob o ponto de vista do coeficiente de correlação, a indicação mais aderente é aquela dada pelo fluxo de controle empregando codificador de menor taxa de transmissão, na ordem G.723.1 ($0,9878$ e $0,9193$), G.729B ($0,9783$ e $0,8655$) e G.711 ($0,9193$ e $0,8655$). Sob o ponto de vista da variação do *jitter* normalizado, as indicações dadas pelo fluxo de controle podem alcançar grandes diferenças como registrado nas Tabelas 4.2, 4.3 e 4.4.
 - (3) Conforme visto nas Figura A6.9, A6.10, A7.9 e A7.10, na condição de tráfego heterogêneo, o fluxo de controle empregando codificador G.711 apresentou forte dependência do aumento do número de fluxos do outro tipo de codificador da composição, tornando impraticável o emprego do fluxo de controle com codificador G.711.

- vi) Sob o ponto de vista prático, obedecida a conclusão anterior, a informação dada pelo fluxo de controle quanto aos valores de pico, a métrica *jitter* pode ser encarada simplesmente como uma indicação de comportamento.
- (1) Em todos os casos analisados, os valores da métrica *jitter* são muito pequenos em relação aos valores de retardo, conseqüentemente não causam perda de pacotes ou mesmo distorção do sinal de voz decodificado.
 - (2) Muitas aplicações VoIP têm dispositivos para eliminação de pequenos retardos, ou seja, se os valores da métrica *jitter* forem pequenos, uma indicação com duas décadas de erro não causará problema.
- vii) Considerando que o emprego de fluxos de controle visa o ambiente de pequenos provedores de serviço VoIP, os resultados obtidos no experimento indicam ser viável o emprego desse dispositivo, obedecidas as premissas iniciais dadas no item 4 do Capítulo 4.
- (1) Não haveria risco de perda substancial de banda com o emprego do fluxo de controle e os benefícios da informação em tempo real do comportamento dos fluxos de agregado tornam o gerenciamento da rede mais fácil.
 - (2) Deve ser observado que na prática não são empregados intervalos de tempo de 1 segundo como realizado na simulação, mas intervalos da ordem de 15 minutos, 1 hora, etc., dependendo da estratégia operacional e do custo para tratamento das informações. Assim sendo, avaliar um único fluxo pode ser significativamente menos oneroso para o ISP considerando a contrapartida de perda de precisão no processo. A resposta está na relação custo-benefício de cada empresa.
 - (3) O objetivo do emprego do fluxo de controle é servir como ferramenta para prevenir a falta de banda no *core* da rede através da degradação das métricas entre o nó de ingresso e o nó de egresso do domínio DiffServ, eventualmente servir de referência para a admissão de mais fluxos no agregado e indicar a degradação das métricas que possam afetar a QoS e conseqüentemente os SLAs acordados, empregando limiares de degradação. Não se procura um indicador exato do que está acontecendo em cada fluxo que compõem o agregado, mesmo porque cada fluxo tem um desempenho específico. O que se deseja é uma ferramenta que possa indicar o comportamento médio dos fluxos

do agregado ou o efeito causado pela multiplexação desses fluxos no agregado. Assim, com as simulações realizadas, pode ser visto o efeito do carregamento do agregado ser capturado pelo fluxo de controle nos diversos cenários.

Os resultados obtidos foram analisados com o objetivo de estabelecer uma prova de conceito sobre o emprego da abordagem proposta, isto é, de que é viável o uso de dados de um fluxo de controle para averiguação de desempenho de fluxos individuais de um agregado de fluxos VoIP. Para tanto foram consideradas as hipóteses de tráfego agregado homogêneo e de tráfego agregado heterogêneo. Embora não conclusivos para grandes agregados de tráfego, os resultados obtidos permitem avaliar qualitativamente o retardo e o *jitter*. Acredita-se que sua aplicação em redes de topologia e de recursos controlados, como redes corporativas e pequenos provedores de infra-estrutura, podem ser beneficiados com o seu emprego, uma vez que não requer grandes investimentos de software, podendo ser avaliados através do emprego de *sniffers* para obtenção dos dados para análise e de simuladores para planejamento de capacidade.

O trabalho contribuiu com uma abordagem de uso de fluxo de controle como uma ferramenta para avaliação de desempenho de agregado de fluxos VoIP. O contexto do seu uso foi imaginado para pequenos provedores de serviço e gerência de redes corporativas, já que o uso de DiffServ ainda não é padrão na Internet.

O modelo de simulação desenvolvido contribui com um detalhado estudo de modelos de tráfego para VoIP e as implicações do uso de diferentes *codecs* em agregados desse tipo de tráfego. Os resultados do estudo de modelos de tráfego podem ser diretamente aplicados para geração de tráfego sintético VoIP a ser utilizado em novas simulações, por exemplo, com o emprego do simulador NS-2 (*Network Simulator, version 2*).

Embora ainda de forma inicial, os resultados obtidos permitiriam inferir de forma qualitativa sobre o agregado de tráfego VoIP com relação às métricas estudadas e diferentes *codecs*. O tratamento estatístico aplicado aos dados revelou padrões de aderência aceitáveis para a base de dados simulada.

No decorrer do trabalho foi discutida a possibilidade de emprego de outras técnicas de disciplina de fila para verificar se é possível melhorar a aderência das métricas indicadas pelo

fluxo de controle. Esse tema não foi investigado e não se conhece o impacto da aplicação de diferentes disciplinas de fila sobre o desempenho dos fluxos VoIP.

Também no decorrer do trabalho foi comentada a possibilidade de verificar a existência de uma posição dentro da fila para obtenção do fluxo de controle ótimo, de forma a melhorar a aderência das métricas indicadas por esse fluxo. Esse tema não foi investigado e não se sabe o impacto dessa aplicação.

Outra possibilidade de desenvolvimento da idéia para otimizar o fluxo de controle é utilizar um fluxo sintético diferente daquele produzido pelos *codecs* de voz padronizados, com tamanho de pacotes variável em função da formação do agregado e analisar conjuntamente com as demais técnicas de fila o efeito sobre diferentes ambientes VoIP, considerando agregados com misturas de mais de dois *codecs*.

Sugere-se também a análise de tráfegos de voz e de vídeo-telefonia empregando a mesma classe DiffServ conforme o trabalho [AR2004]. Essa é mais uma aplicação que envolve a modelagem de tráfego e as diversas disciplinas de fila. Várias referências bibliográficas tratam, também, da modelagem dos diversos tipos de tráfego da Internet, aplicado ao segmento de sistemas móveis celulares.

Outra possibilidade é a realização de estudos de modelos de tráfego para obtenção de distribuições estatísticas e a provável constatação de que existem vários serviços de voz para os quais poderia haver modelagens de tráfego diferentes.

Considerando as dificuldades enfrentadas no decorrer do presente trabalho para modelagem de tráfego VoIP, é sugerida a abertura dessa linha de investigação, especialmente considerando a tendência do mercado atual para este tipo de aplicação.

ANEXO 1

DISTRIBUIÇÕES ESTATÍSTICAS MAIS USADAS NA MODELAGEM DE TRÁFEGO VoIP

Conforme Jain [JR1991], as distribuições estatísticas citadas neste trabalho são definidas e especificadas como:

A1.1 DISTRIBUIÇÃO EXPONENCIAL

A distribuição exponencial é extensivamente utilizada em modelos de filas. As principais características da distribuição estão na Tabela A1.1:

Tabela A1.1 – Distribuição exponencial - e^a

1. Parâmetros: a = parâmetro de escala = média, $a > 0$.
2. Faixa: $0 \leq x \leq \infty$.
3. função densidade de probabilidade (pdf): $f(x) = \frac{1}{a} e^{-\frac{x}{a}}$.
4. função distribuição cumulativa (CDF): $F(x) = 1 - e^{-\frac{x}{a}}$.
5. inversa da CDF: $-a \cdot \ln(u)$.
6. média: a .
7. variância: a^2 .

A distribuição exponencial é a única distribuição contínua com a propriedade de não ter memória de forma a lembrar o tempo desde o último evento, não ajudando na predição do tempo do próximo evento. Se os tempos de interchegadas das tarefas forem exponencialmente distribuídos, com média $1/\lambda$, a distribuição de probabilidade é

$$F(\tau) = P(\tau < t) = 1 - e^{-\lambda t}, t \geq 0.$$

Se for observada uma chegada e nesse instante o relógio for acionado em $t = 0$, o tempo médio para a próxima chegada é também $1/\lambda$. Supondo que aquela chegada não tenha sido observada até $t = \tau$, a distribuição do tempo remanescente até a próxima chegada será $P(\tau - x < t | \tau > x) = 1 - e^{-\lambda x}$, que é idêntica a situação em $t = 0$. Em particular, o tempo médio da próxima chegada é ainda $1/\lambda$. Assim, o tempo esperado para a próxima chegada é sempre $1/\lambda$. Um argumento similar se aplica a qualquer outra variável exponencial. Essa

propriedade sem memória, que simplifica consideravelmente a análise, é a razão para a popularidade da distribuição exponencial.

A distribuição exponencial é empregada para modelar o tempo entre eventos sucessivos, particularmente se os eventos forem causados por um grande número de fatores independentes, como por exemplo,

- o tempo entre chegadas de pedidos sucessivos para um dispositivo ou
- o tempo entre falhas de um equipamento.

Os tempos de serviço em um dispositivo são também modelados como exponencialmente distribuídos.

As variáveis exponenciais podem ser geradas utilizando a transformação inversa, sendo gerado um número aleatório u em $U(0,1)$ e retornando $-a \ln(u)$ como e^u , esse procedimento é baseado no método da transformação inversa.

A1.2 DISTRIBUIÇÃO GAMMA

A distribuição Gamma é uma generalização da distribuição de Erlang, permitindo parâmetro de forma não inteiro. Como a distribuição Exponencial e de Erlang, a distribuição Gamma é usada na modelagem de filas, sendo é empregada para modelar tempos de serviço em fila de dispositivos em modelos de rede e para calcular tempos de reparo de forma similar ao emprego da distribuição de Erlang. As características principais da distribuição Gamma estão relacionadas na Tabela A1.2:

Tabela A1.2 – Distribuição Gamma - $\gamma(a,b)$

1. Parâmetros: a = parâmetro de escala, $a > 0$;
 b = parâmetro de forma, $b > 0$.
2. Faixa: $0 \leq x \leq \infty$.

3. função densidade de probabilidade (pdf): $f(x) = \frac{\left(\frac{x}{a}\right)^{b-1} \cdot e^{-\frac{x}{a}}}{a \cdot \Gamma(b)}$.

$\Gamma(*)$ é a função Gamma

4. função distribuição cumulativa (CDF): não tem.
5. média: $a \cdot b$
6. variância: $a^2 \cdot b$

A variável Gamma pode ser gerada como segue:

- a) se b for um número inteiro, a soma da variável exponencial b tem uma distribuição Gamma. Esse procedimento produz a seguinte expressão para o gerador da variável Gamma a partir da variável $U(0,1)$:

$$\gamma(a,b) \approx -a \cdot \ln \left[\prod_{i=1}^b u_i \right]$$

- b) para $b < 1$, gerar uma variável β tal que $x \approx \beta(b, 1-b)$ e uma variável exponencial $y \approx e^1$. O produto $a \cdot x \cdot y$ tem uma distribuição $\gamma(a,b)$.
- c) para outros valores de b , criar e somar duas variáveis Gamma correspondendo a parte inteira e a parte fracionária de b para produzir a variável Gamma desejada:

$$\gamma(a,b) \approx \gamma(a, \lfloor b \rfloor) + \gamma(a, b - \lfloor b \rfloor)$$

A1.3 DISTRIBUIÇÃO LOGNORMAL

O logaritmo de uma variável normal tem uma distribuição lognormal. As principais características da distribuição Lognormal estão na Tabela A1.3:

Tabela A1.3 – Distribuição Lognormal - $LN(\mu, \sigma)$

1. Parâmetros: $\mu =$ média de $\ln(x)$, $\mu > 0$;
 $\sigma =$ desvio padrão de $\ln(x)$, $\sigma > 0$.
2. Faixa: $0 \leq x \leq \infty$.
3. função densidade de probabilidade (pdf): $f(x) = \frac{1}{\sigma \cdot x \cdot \sqrt{2\pi}} \cdot e^{-\frac{(\ln x - \mu)^2}{2 \cdot \sigma^2}}$.
4. função distribuição cumulativa (CDF): não tem
5. média: $e^{\mu + \frac{\sigma^2}{2}}$
6. variância: $e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$

O produto de um grande número de variáveis aleatórias positivas tende a ter uma distribuição lognormal aproximada.

Essa função é utilizada para modelar erros que são um produto de efeitos de um grande número de fatores.

A variável lognormal pode ser gerada utilizando um \log de uma variável normal, gerando $x \approx N(0,1)$ e retornando $e^{\mu + \sigma x}$.

Outra forma de representar a função distribuição de probabilidade (PDF) lognormal é: $P\{X > x\} = P\{e^{\mu + \sigma U} > x\}$, $\mu \in \mathfrak{R}$, $\sigma > 0$, com U sendo uma variável aleatória normal padronizada.

A1.4 DISTRIBUIÇÃO DE PARETO

A função distribuição cumulativa (CDF) de Pareto é uma curva de potência que pode ser facilmente traçada para dados observados. As principais características da distribuição estão na Tabela A1.4:

Tabela A1.4 – Distribuição de Pareto - $Pareto(a)$

1. Parâmetros: a = parâmetro de forma, $a > 0$.
2. Faixa: $1 \leq x \leq \infty$.
3. função densidade de probabilidade (pdf): $f(x) = a \cdot x^{-(a+1)}$.
4. função distribuição cumulativa (CDF): $F(x) = 1 - x^{-a}$.
5. inversa da CDF: $\frac{1}{u^{1/a}}$
6. média: $\frac{a}{a-1}$, fornecendo $a > 1$
7. variância: $\frac{a}{(a-1)^2 \cdot (a-2)}$, fornecendo $a > 2$.

A distribuição de Pareto é útil para traçar uma distribuição de dados observados. Dada uma amostragem de n observações $\{x_1, x_2, \dots, x_n\}$, a máxima probabilidade estimada do parâmetro a é

$$a = \frac{1}{\frac{1}{n} \cdot \sum_{i=1}^n \ln x_i}.$$

O método de transformação inverso é a forma mais fácil de gerar a variável de Pareto. Gerar $u \approx U(0,1)$ e retornar $1/u^{1/a}$.

A função distribuição de probabilidade de Pareto pode ser escrita como:

$$P\{X > x\} = \left(\frac{a}{a+x} \right)^v, \quad a, v > 0.$$

Observação: na biblioteca do simulador OPNET a pdf de Pareto tem a forma

$$f(x) = \frac{a \cdot c^a}{x^{a+1}}, \quad c \leq x \leq \infty$$

$$\text{média} \quad E(x) = \frac{a \cdot c}{a-1}, \quad a > 1$$

$$\text{variância} \quad \sigma_x^2 = \left[\frac{a \cdot c^2}{(a-1)^2 \cdot (a-2)} \right], \quad a > 2$$

$$a > 0 \quad \text{shape parameter}$$

$$c > 0 \quad \text{location parameter}$$

Fazendo $c = 1$ nas fórmulas acima são obtidas as expressões dadas em [JR1991].

Nos anexos B e C de [PV1995] estão descritas a distribuição de Pareto e a forma de obtenção dos tempos inter pacotes da distribuição de Pareto.

A1.5 DISTRIBUIÇÃO DE WEIBULL

A distribuição de Weibull é comumente usada em análise de confiabilidade. As principais características dessa distribuição estão na Tabela A1.5:

Tabela A1.5 – Distribuição de Weibull - $Weibull(a, b)$

1. Parâmetros: a = parâmetro de escala, $a > 0$;
 b = parâmetro de forma, $b > 0$.
2. Faixa: $0 \leq x \leq \infty$.
3. função densidade de probabilidade (pdf): $f(x) = \frac{b \cdot x^{b-1}}{a^b} \cdot e^{-\left(\frac{x}{a}\right)^b}$.
4. função distribuição cumulativa (CDF): $F(x) = 1 - e^{-\left(\frac{x}{a}\right)^b}$.
5. inversa da CDF: $a \cdot (\ln u)^{1/b}$
6. média: $\frac{a}{b} \cdot \Gamma(1/b)$
7. variância: $\frac{a^2}{b^2} \{2 \cdot b \cdot \Gamma(2/b) - [\Gamma(1/b)]^2\}$

Se $b = 3,602$, a distribuição de Weibull é aproximadamente uma distribuição normal. Para $b > 3,602$, a distribuição de Weibull tem uma cauda esquerda longa e no caso de $b < 3,602$, ela tem uma cauda direita longa. Para $b \leq 1$, a função densidade de probabilidade de Weibull tem um pico agudo na moda.

A distribuição de Weibull é utilizada para modelar tempo de vida de componentes. Para $b < 1$, a distribuição de Weibull fornece a taxa de falhas aumentando com o tempo. Para $b > 1$, a taxa de falhas decresce com o tempo. Em $b = 1$, a taxa de falhas é constante e os tempos de vida são exponencialmente distribuídos.

A técnica de transformação inversa pode ser utilizada para gerar a variável Weibull. Gerar $u \approx U(0,1)$ e retornar $a.(ln u)^{1/b}$ como *Weibull*(a,b).

A função distribuição de probabilidade de Weibull pode ser escrita como:

$$P\{X > x\} = e^{-ax^b}, \quad a > 0, \quad 0 < b < 1.$$

ANEXO 2

COEFICIENTE DE CORRELAÇÃO

Conforme apresentado em [SM1977], são descritos a seguir os conceitos necessários para a compreensão do coeficiente de correlação que é empregado nesse trabalho para informar o grau de aderência dos valores médios das métricas retardo e *jitter* indicados no fluxo de controle aos valores médios medidos nos demais fluxos do agregado EF PHB para a aplicação VoIP.

A2.1 VARIÂNCIA E DESVIO PADRÃO

A esperança de uma variável aleatória X é freqüentemente designada por *média* e representada por μ .

A *variância* é um número não negativo, definida como sendo $Var(X) = E[(X - \mu)^2]$.

O *desvio padrão* é dado por $\sigma_x = \sqrt{Var(X)} = \sqrt{E[(X - \mu)^2]}$, ou seja, $Var(X) = \sigma_x^2$.

Se X for uma variável aleatória discreta com função de probabilidade $f(x)$, então a variância de X é dada por $\sigma_x^2 = E[(X - \mu)^2] = \sum_{j=1}^n (x_j - \mu)^2 \cdot f(x_j) = \sum (x - \mu)^2 \cdot f(x)$.

Se X for uma variável aleatória contínua com função de densidade $f(x)$, então a variância de X é dada por $\sigma_x^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) \cdot dx$.

A variância (ou desvio padrão) é uma medida de dispersão dos valores da variável aleatória ao redor da média μ . Se os valores tendem a concentrarem-se próximos da média, a variância é pequena; mas se os valores tendem a afastar-se, a variância é grande.

A2.2 VARIÂNCIA DE DISTRIBUIÇÕES CONJUNTAS

No caso de duas variáveis aleatórias contínuas, X e Y , com função densidade conjunta $f(x,y)$, as médias ou esperanças de X e de Y são:

$$\mu_X = E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f(x, y) \cdot dx \cdot dy \quad e \quad \mu_Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \cdot f(x, y) \cdot dx \cdot dy$$

e as variâncias são:

$$\sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f(x, y) \cdot dx \cdot dy$$

$$\sigma_Y^2 = E[(Y - \mu_Y)^2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_Y)^2 \cdot f(x, y) \cdot dx \cdot dy$$

A covariância das variáveis X e Y é $\sigma_{XY} = Cov(X, Y) = E[(X - \mu_X) \cdot (Y - \mu_Y)]$, ou seja, em termos da função densidade conjunta $f(x, y)$, a covariância é

$$\sigma_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X) \cdot (y - \mu_Y) \cdot f(x, y) \cdot dx \cdot dy$$

No caso de duas variáveis aleatórias discretas, temos:

$$\mu_X = \sum_x \sum_y x \cdot f(x, y) \quad e \quad \mu_Y = \sum_x \sum_y y \cdot f(x, y)$$

$$\sigma_{XY} = \sum_x \sum_y (x - \mu_X) \cdot (y - \mu_Y) \cdot f(x, y)$$

onde os somatórios se estendem sobre todos os valores discretos de X e de Y .

A2.3 COEFICIENTE DE CORRELAÇÃO

Se X e Y forem independentes então $Cov(X, Y) = \sigma_{XY} = 0$. Por outro lado, se X e Y forem completamente dependentes, $X = Y$, então $Cov(X, Y) = \sigma_{XY} = \sigma_X \sigma_Y$.

Isso conduz a uma medida da dependência das variáveis X e Y dada por $\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$, que é uma quantidade adimensional, denominada *coeficiente de correlação*, que está contida no intervalo $-1 \leq \rho \leq 1$.

Se $\rho = 0$, a covariância é zero, e as variáveis X e Y são não correlacionadas e em tal caso, as variáveis podem ou não ser independentes.

A2.4 A MÉDIA AMOSTRAL E A DISTRIBUIÇÃO AMOSTRAL

Sejam X_1, X_2, \dots, X_n variáveis aleatórias para uma amostra de tamanho n . Então, a média da amostra, ou *média amostral*, é uma variável aleatória definida por

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Se $f(x)$ for a função distribuição de probabilidade de certa população da qual extraímos uma amostra de tamanho n , então a distribuição de probabilidade da estatística amostral \bar{X} , é chamada de distribuição amostral da média da amostra ou *distribuição amostral das médias*. Aqui valem as seguintes relações:

$$E(\bar{X}) = \mu_{\bar{X}} = \mu, \quad \text{onde } \mu \text{ é a média da população}$$

$$E[(\bar{X} - \mu)^2] = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \quad \text{onde } \sigma^2 \text{ é a variância da população}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \quad \text{onde } N \text{ é o tamanho da população, amostragem sem reposição e } n \leq N, \text{ é o tamanho da amostra}$$

Se a população da qual se extraem amostras tem uma distribuição de probabilidade, não necessariamente normal, com média μ e variância σ^2 , então a variável padronizada associada a \bar{X} , dada por

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

é assintoticamente normal, ou seja

$$\lim_{n \rightarrow \infty} P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} \cdot du$$

A variação da amostra ou *variância amostral* é dada por

$$S^2 = \frac{(X_1 - \bar{X}) + (X_2 - \bar{X}) + \dots + (X_n - \bar{X})}{n}$$

Sempre que o valor esperado de uma estatística for igual ao correspondente parâmetro da população é dito que a estatística é um estimador não tendencioso.

Aqui, $E(S^2) = \mu_{S^2} = \frac{n-1}{n} \cdot \sigma^2$ é muito próximo de σ^2 somente para grandes valores de n , tais como $n \geq 30$, de forma que o estimador não tendencioso desejado é definido por

$$\hat{S}^2 = \frac{n-1}{n} S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1}$$

e assim, $E(\hat{S}^2) = \sigma^2$.

Conhecida a variância σ^2 da população, a variável padronizada Z é normalmente distribuída se a população da qual se extraem as amostras tiver distribuição normal, enquanto que é assintoticamente normal se a população não for normal, desde que $n \geq 30$.

Se não for conhecida a variância da população, um procedimento é substituir σ pela variável aleatória \hat{S} que dá o desvio padrão amostral e determinar a distribuição da estatística correspondente, designada por

$$T = \frac{\bar{X} - \mu}{\frac{\hat{S}}{\sqrt{n}}} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n-1}}}$$

onde T tem distribuição de Student-t com $n-1$ graus de liberdade.

A2.5 O COEFICIENTE DE CORRELAÇÃO LINEAR

A partir da reta dos mínimos quadrados pode-se definir o coeficiente de correlação amostral como sendo $r = S_{xy}/S_x S_y$ que desenvolvendo produz

$$r = \frac{n \cdot \sum x \cdot y - (\sum x)(\sum y)}{\sqrt{[n \cdot \sum x^2 - (\sum x)^2][n \cdot \sum y^2 - (\sum y)^2]}} = \frac{\text{variação explicada}}{\text{variação total}}$$

Onde, para o presente trabalho, x representa os n valores da métrica obtidos no fluxo de controle e y representa os respectivos valores obtidos nos demais n fluxos do agregado.

Assim, r^2 pode ser interpretado como a fração da variação total que é explicada pela reta de regressão de mínimos quadrados, ou seja, r mede quão bem a reta de regressão de mínimos quadrados se ajusta aos dados amostrais. Se a variação total é toda explicada pela reta de regressão, $r^2 = 1$ ou $r = \pm 1$, dizemos que há correlação linear perfeita. Por outro lado, se a variação explicada é zero, ou seja, a variação total é toda ela não-explicada, $r = 0$.

O coeficiente de correlação linear pode ser positivo ou negativo. Se r for positivo, y tende a aumentar com x enquanto que se r for negativo, y tende a decrescer enquanto x cresce (o coeficiente angular da reta de mínimos quadrados é negativo).

ANEXO 3

RESULTADOS TRÁFEGO HOMOGÊNEO G.711

A3.1 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE

Nas Figuras A3.1 até A3.3 são apresentadas as distribuições do retardo no fluxo de controle G.711 em função do carregamento do agregado EF PHB (1 a 16 fluxos VoIP).

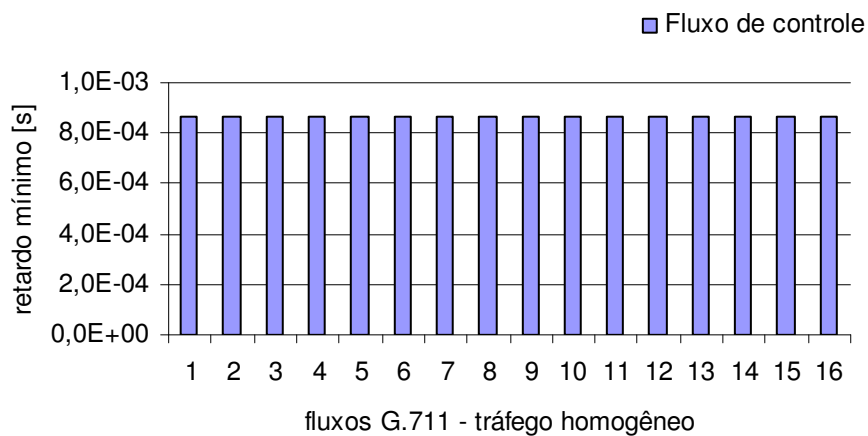


Figura A3.1: Retardo mínimo no fluxo de controle G.711 em função do carregamento

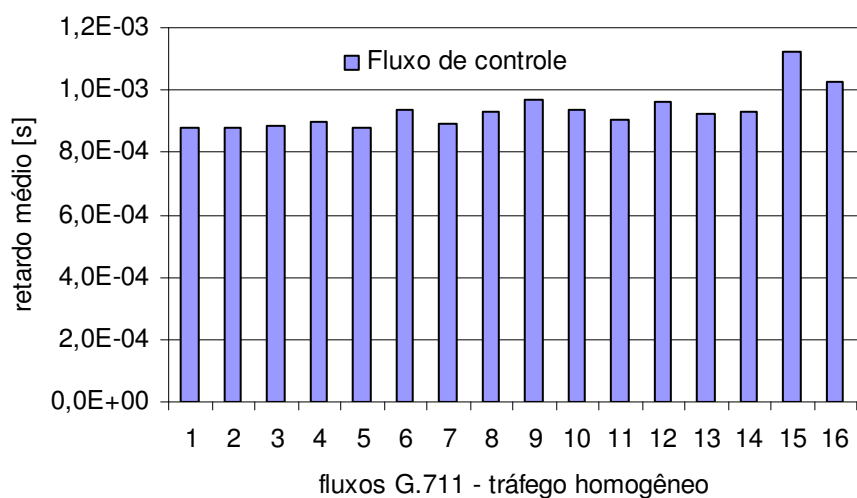


Figura A3.2: Retardo médio no fluxo de controle G.711 em função do carregamento

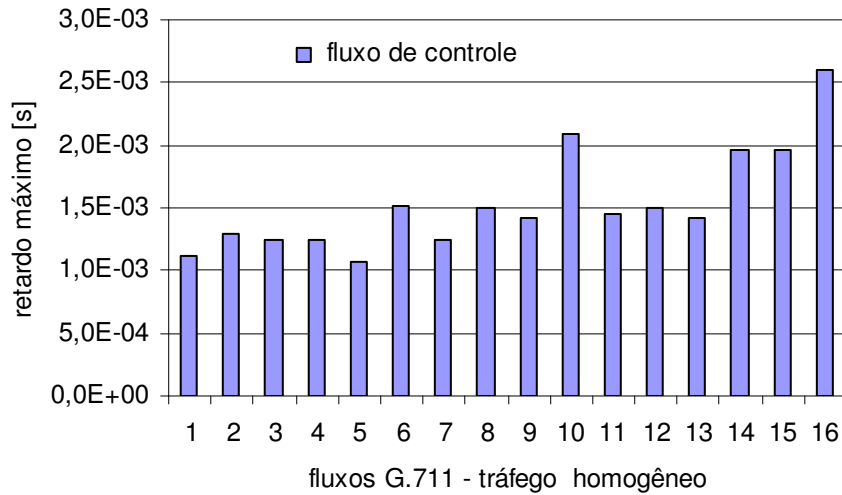


Figura A3.3: Retardo máximo no fluxo de controle G.711 em função do carregamento

A3.2 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE

Nas Figuras A3.4 até A3.6 são apresentadas as distribuições do *jitter* no fluxo de controle G.711 em função do carregamento do agregado EF PHB (1 a 16 fluxos VoIP com *codec* G.711).

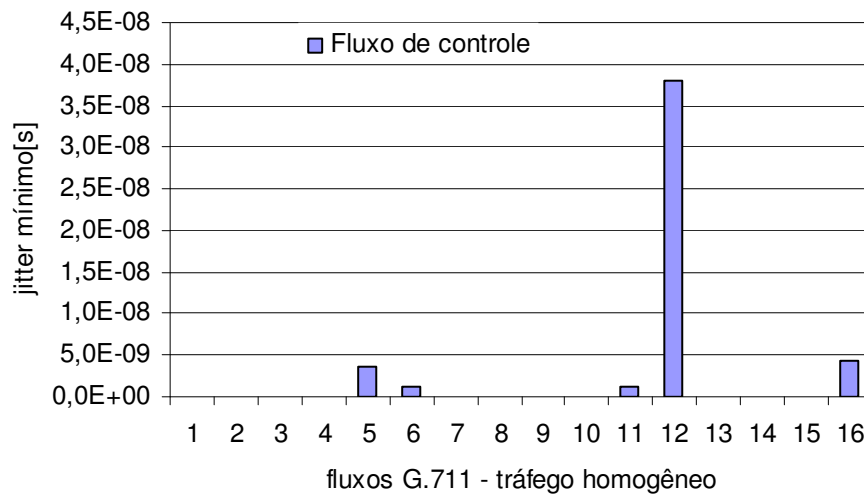


Figura A3.4: *Jitter* mínimo no fluxo de controle G.711 em função do carregamento

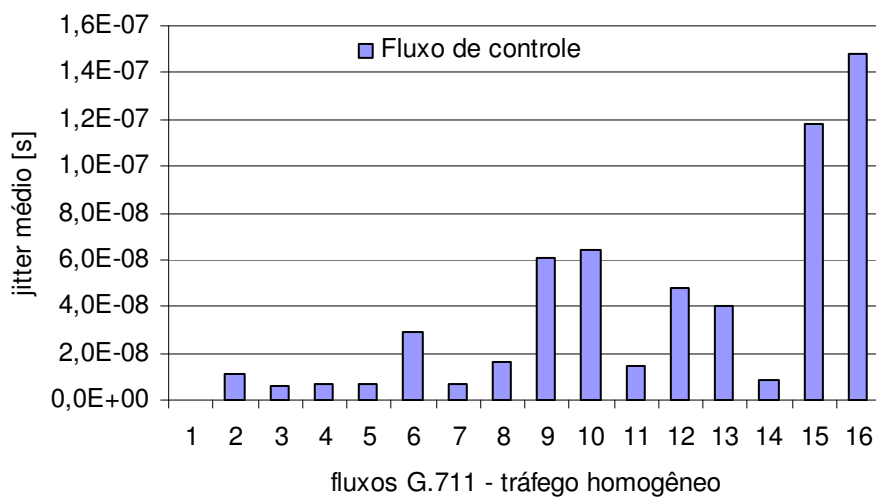


Figura A3.5: *Jitter* médio no fluxo de controle G.711 em função do carregamento

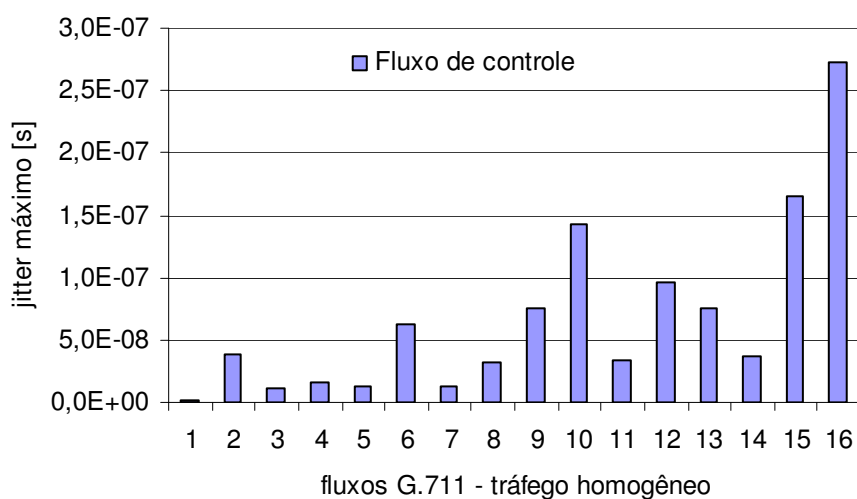


Figura A3.6: *Jitter* máximo no fluxo de controle G.711 em função do carregamento

A3.3 VARIAÇÃO DO RETARDO NORMALIZADO

Nas Figuras A3.7 e A3.8 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.711 em função do carregamento do agregado EF PHB (1 a 16 fluxos VoIP).

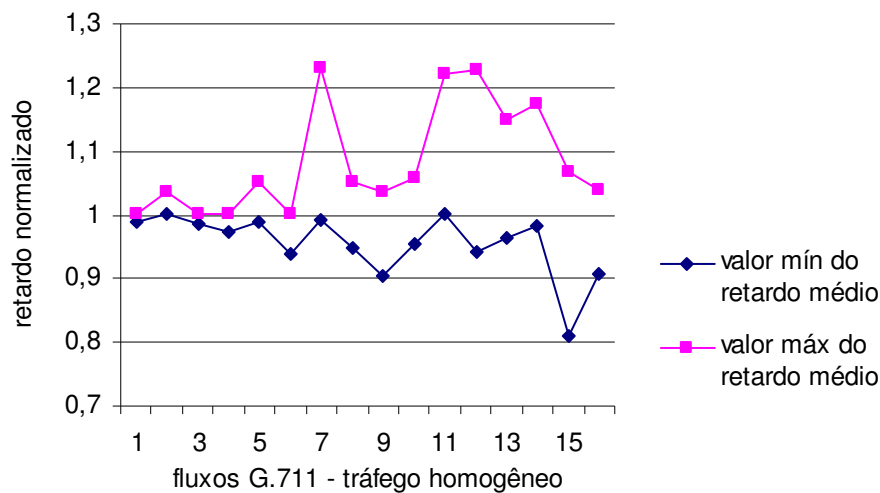


Figura A3.7: Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.711

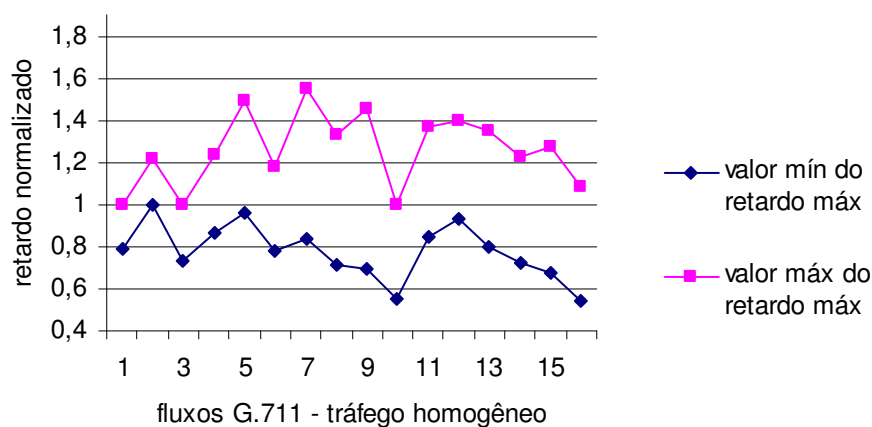


Figura A3.8: Variação do retardo máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.711

A3.4 VARIAÇÃO DO JITTER NORMALIZADO

Nas Figuras A3.9 e A3.10 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.711 em função do carregamento do agregado EF PHB (1 a 16 fluxos VoIP).

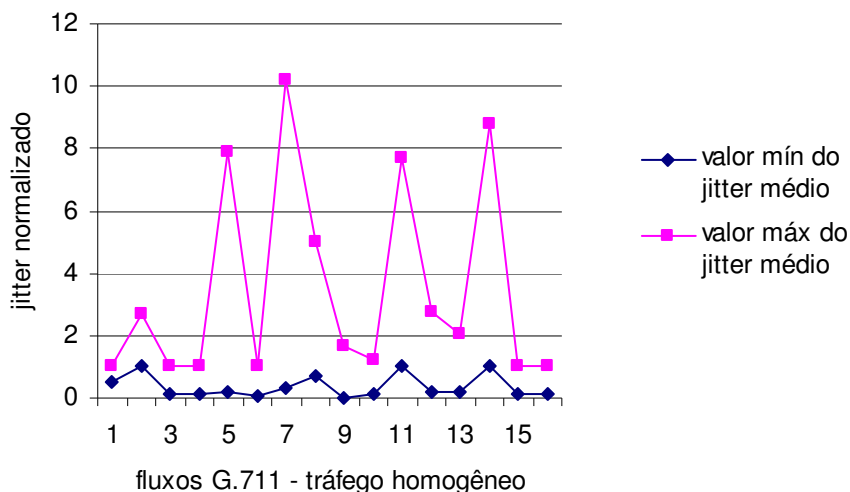


Figura A3.9: Variação do *jitter* médio normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.711

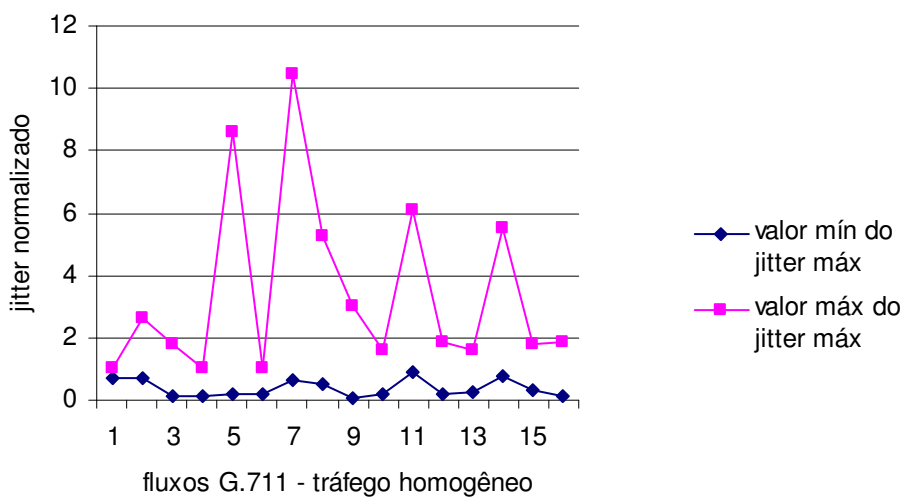


Figura A3.10: Variação do *jitter* máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.711

ANEXO 4

RESULTADOS TRÁFEGO HOMOGÊNEO G.723.1

A4.1 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE

Nas Figuras A4.1 até A4.3 são apresentadas as distribuições do retardo no fluxo de controle G.723.1 em função do carregamento do agregado EF PHB (1 a 80 fluxos VoIP).

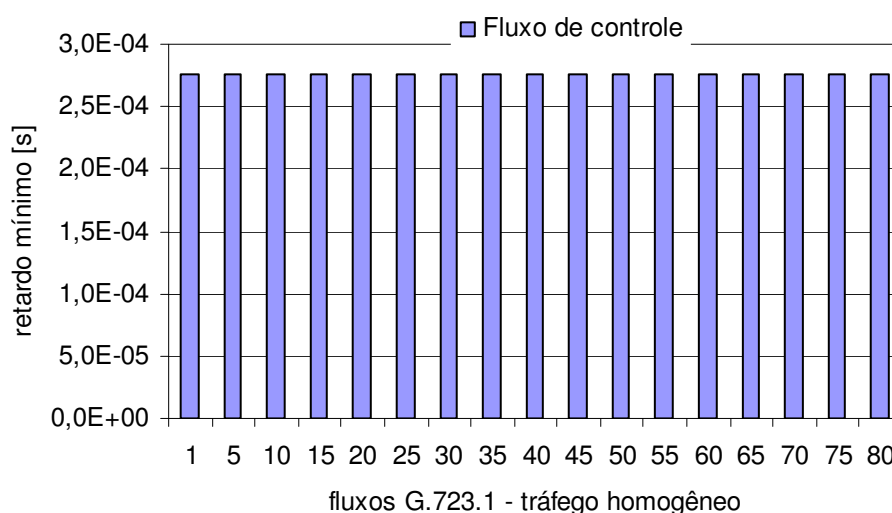


Figura A4.1: Retardo mínimo no fluxo de controle G.723.1 em função do carregamento

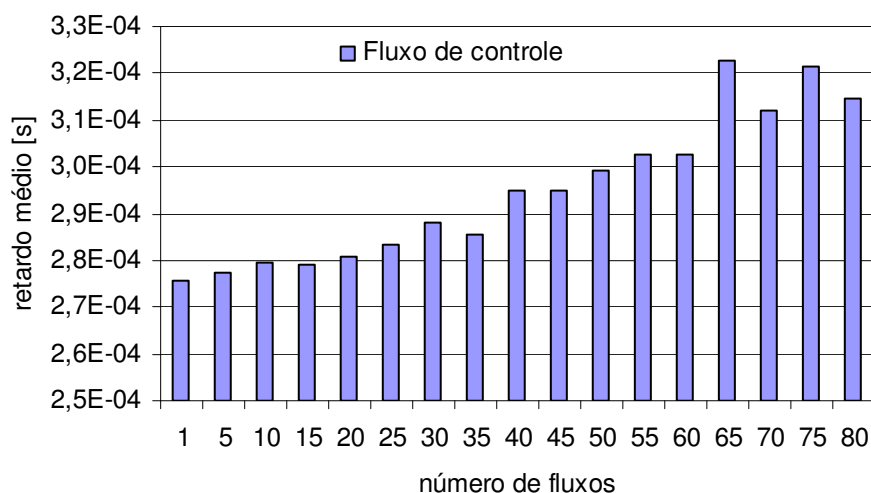


Figura A4.2: Retardo médio no fluxo de controle G.723.1 em função do carregamento

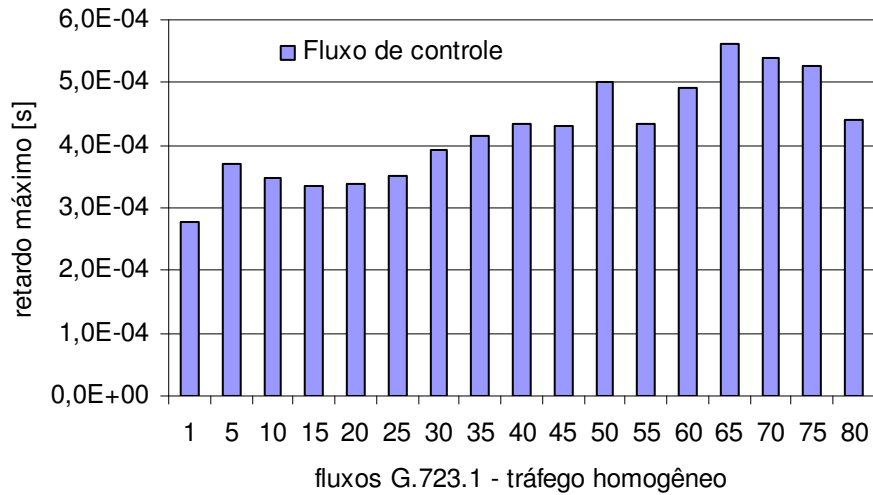


Figura A4.3: Retardo máximo no fluxo de controle G.723.1 em função do carregamento

A4.2 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE

Nas Figuras A4.4 até A4.6 são apresentadas as distribuições do *jitter* no fluxo de controle G.723.1 em função do carregamento do agregado EF PHB (1 a 80 fluxos VoIP).

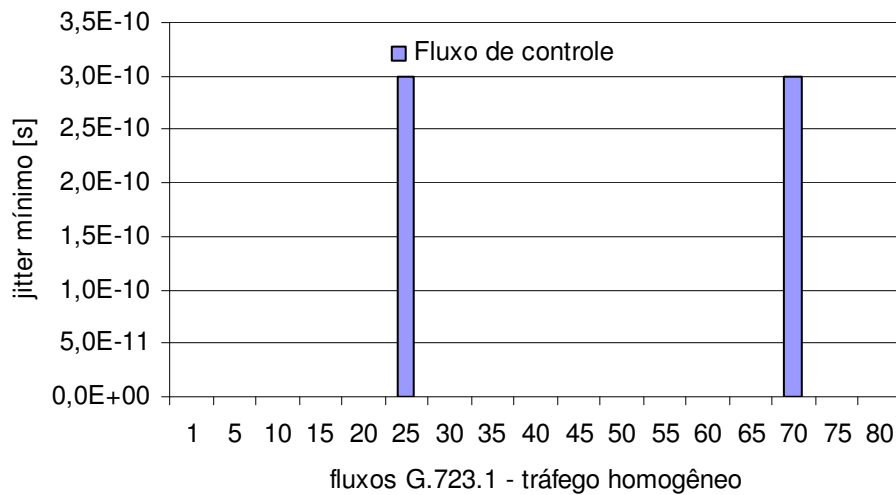


Figura A4.4: *Jitter* mínimo no fluxo de controle G.723.1 em função do carregamento

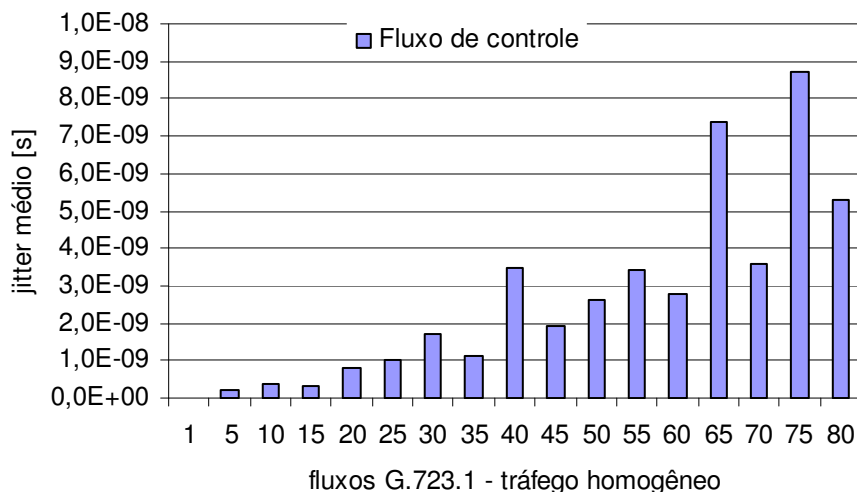


Figura A4.5: *Jitter* médio no fluxo de controle G.723.1 em função do carregamento

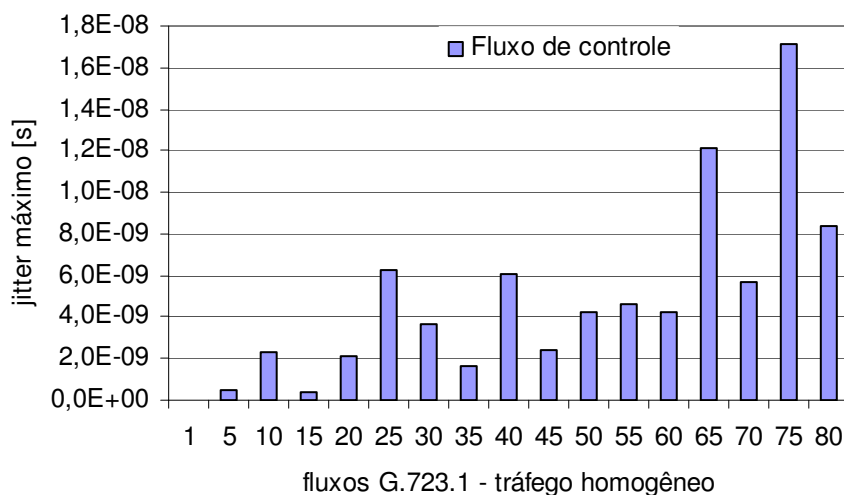


Figura A4.6: *Jitter* máximo no fluxo de controle G.723.1 em função do carregamento

A4.3 VARIAÇÃO DO RETARDO NORMALIZADO

Nas Figuras A4.7 e A4.8 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB (1 a 80 fluxos VoIP).

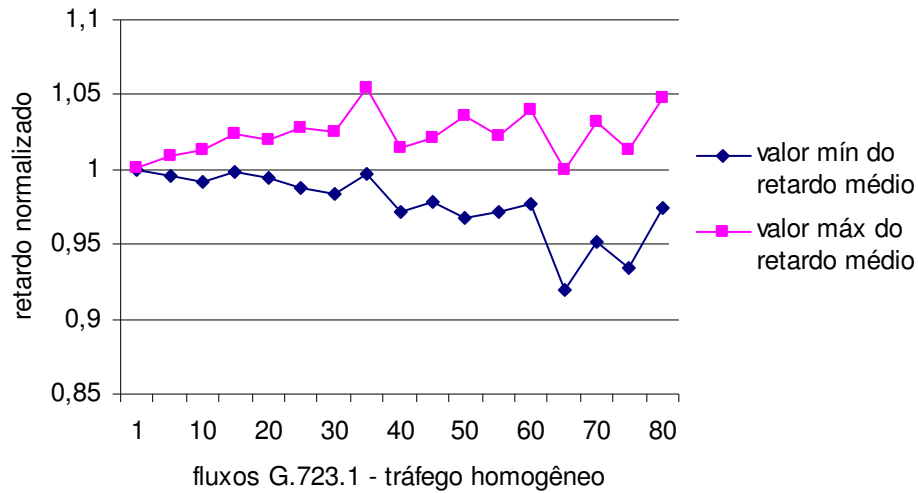


Figura A4.7: Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.723.1

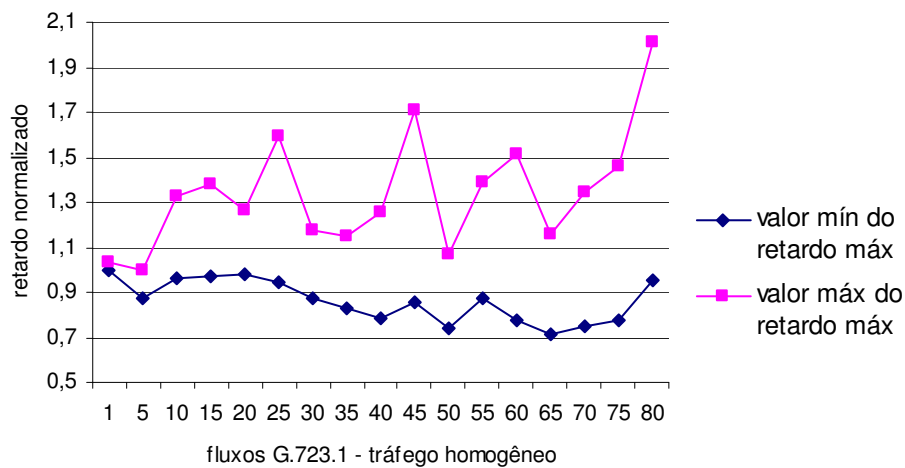


Figura A4.8: Variação do retardo máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.723.1

A4.4 VARIAÇÃO DO JITTER NORMALIZADO

Nas Figuras A4.9 e A4.10 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do fluxo agregado VoIP da classe EF PHB (1 a 80 fluxos VoIP).

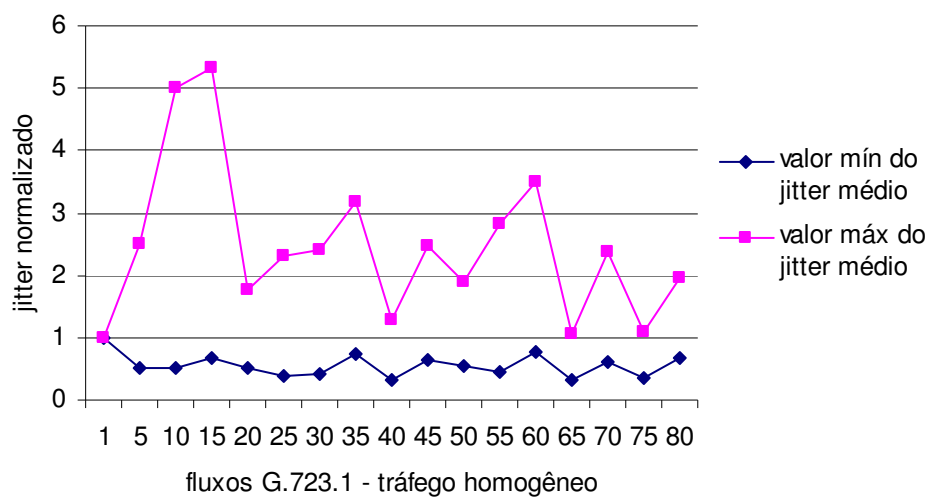


Figura A4.9: Variação do *jitter* médio normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.723.1

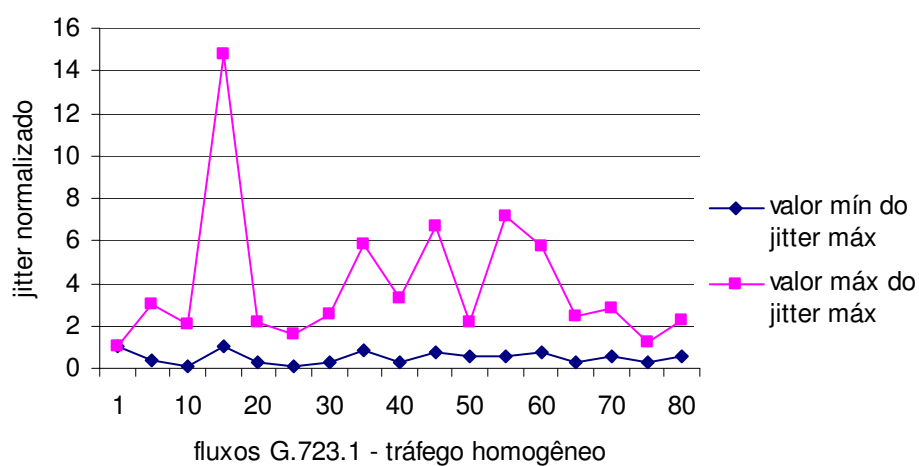


Figura A4.10: Variação do *jitter* máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.723.1

ANEXO 5

RESULTADOS TRÁFEGO HOMOGÊNIO G.729B

A5.1 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE

Nas Figuras A5.1 até A5.3 são apresentadas as distribuições do retardo no fluxo de controle G.729B em função do carregamento do agregado EF PHB (1 a 35 fluxos VoIP).

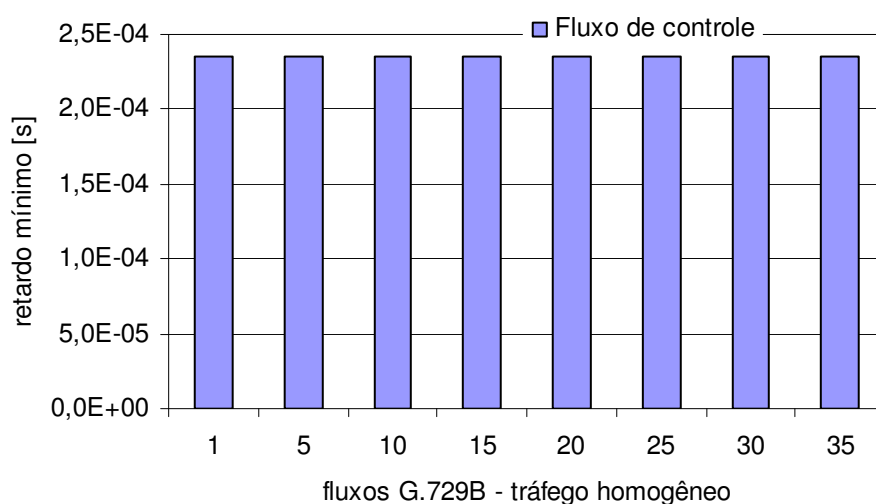


Figura A5.1: Retardo mínimo no fluxo de controle G.729B em função do carregamento

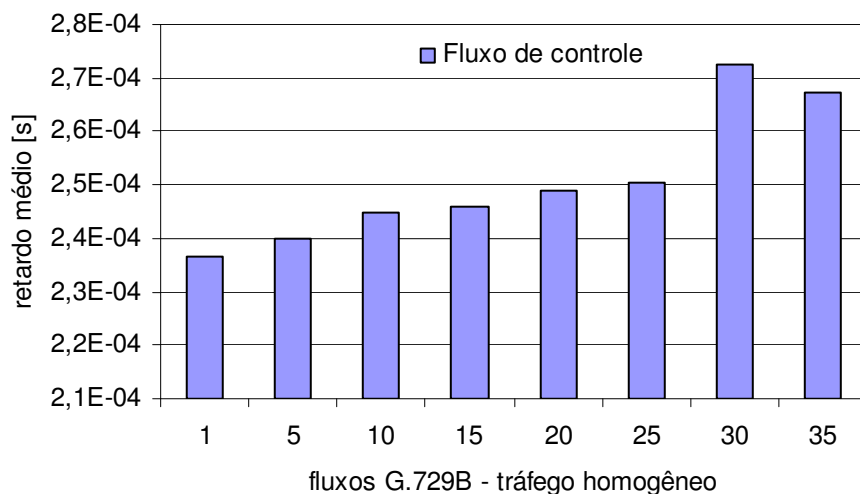


Figura A5.2: Retardo médio no fluxo de controle G.729B em função do carregamento

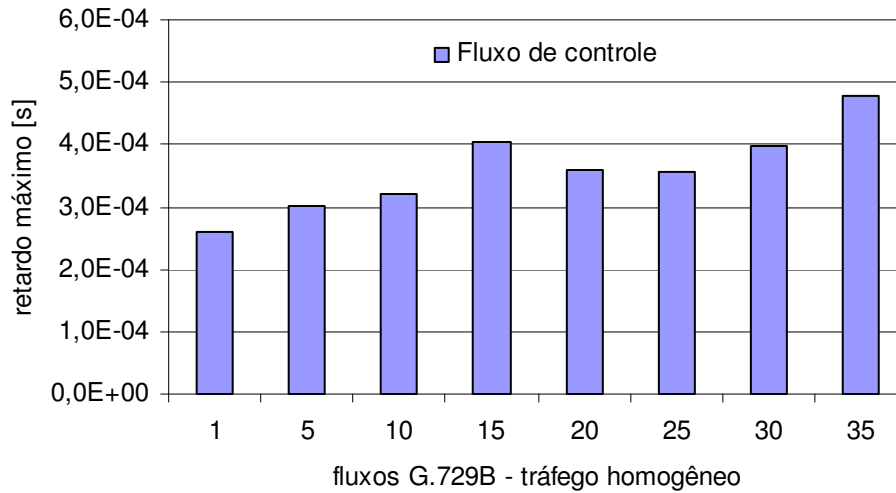


Figura A5.3: Retardo máximo no fluxo de controle G.729B em função do carregamento

A5.2 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE

Nas Figuras A5.4 até A5.6 são apresentadas as distribuições do *jitter* no fluxo de controle G.729.B em função do carregamento do agregado EF PHB (1 a 35 fluxos VoIP).

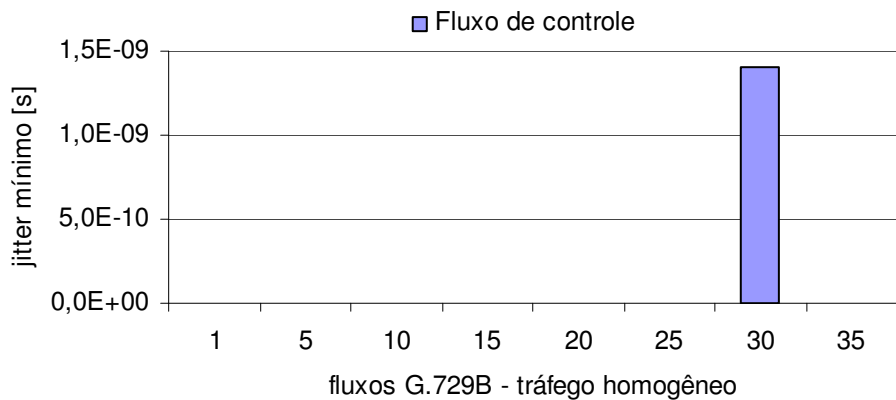


Figura A5.4: *Jitter* mínimo no fluxo de controle G.729B em função do carregamento

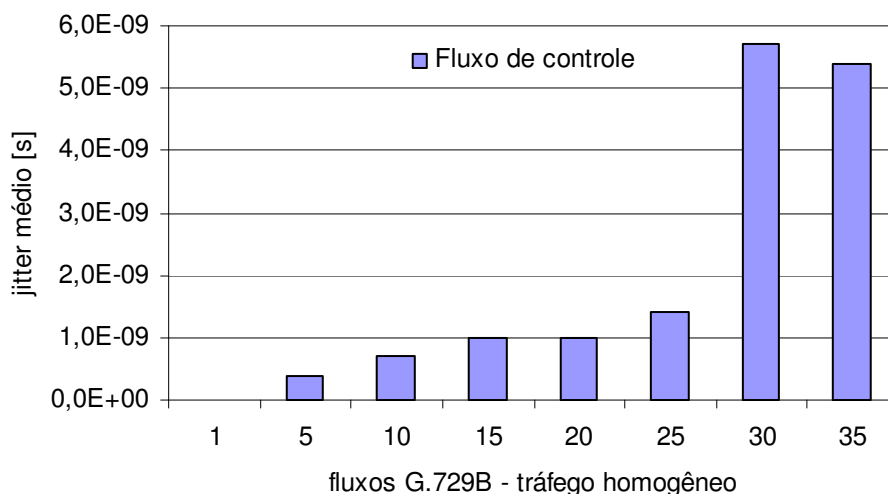


Figura A5.5: *Jitter* médio no fluxo de controle G.729B em função do carregamento

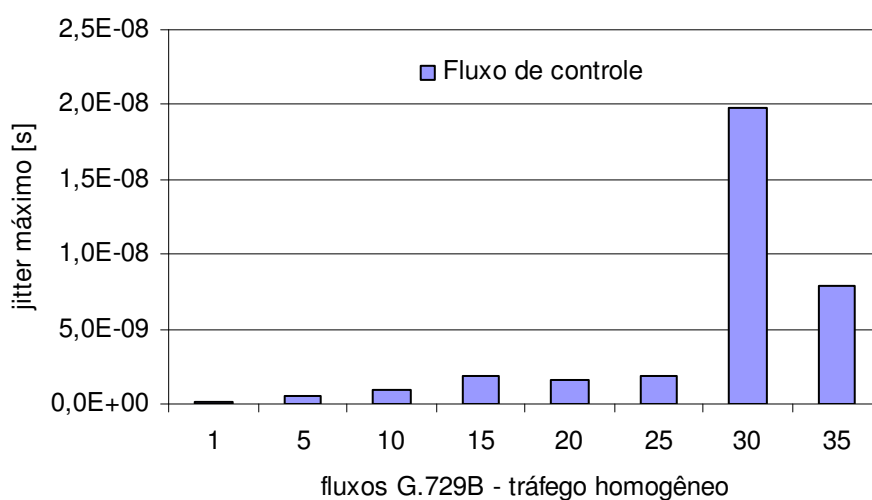


Figura A5.6: *Jitter* máximo no fluxo de controle G.729B em função do carregamento

A5.3 VARIAÇÃO DO RETARDO NORMALIZADO

Nas Figuras A5.7 e A5.8 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.729B em função do carregamento do agregado EF PHB (1 a 35 fluxos VoIP).

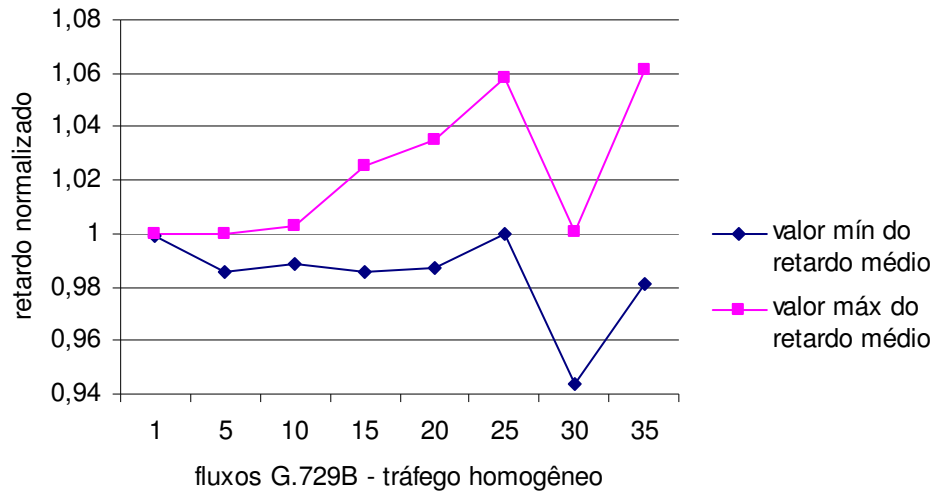


Figura A5.7: Variação do retardo médio normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.729B

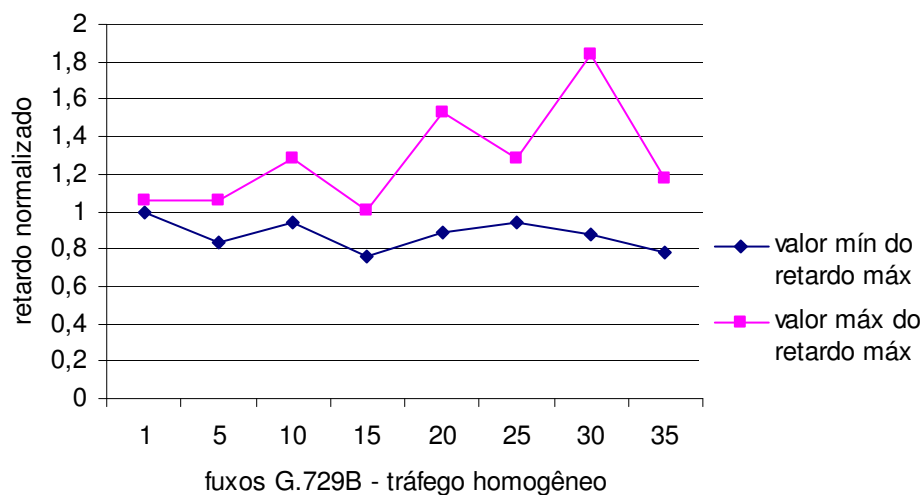


Figura A5.8: Variação do retardo máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.729B

A5.4 VARIAÇÃO DO JITTER NORMALIZADO

Nas Figuras A5.9 e A5.10 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.729B em função do carregamento do agregado EF PHB (1 a 35 fluxos VoIP).

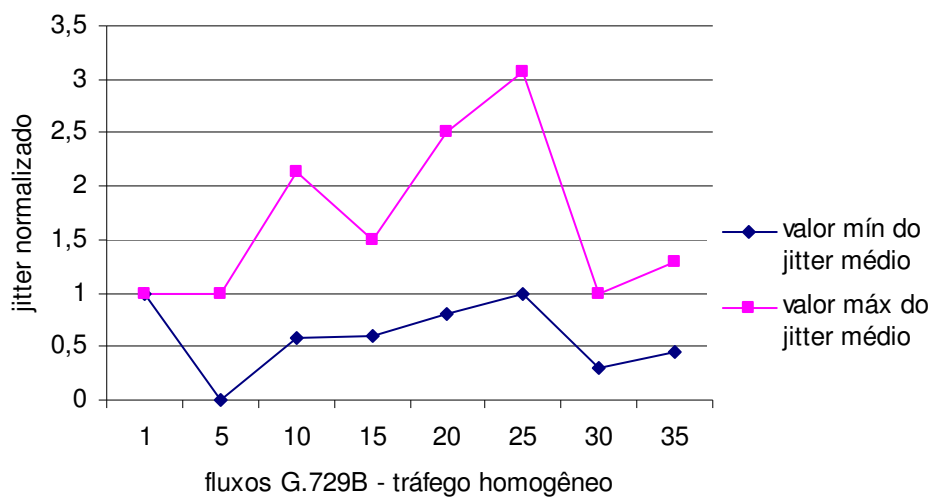


Figura A5.9: Variação do *jitter* médio normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.729B

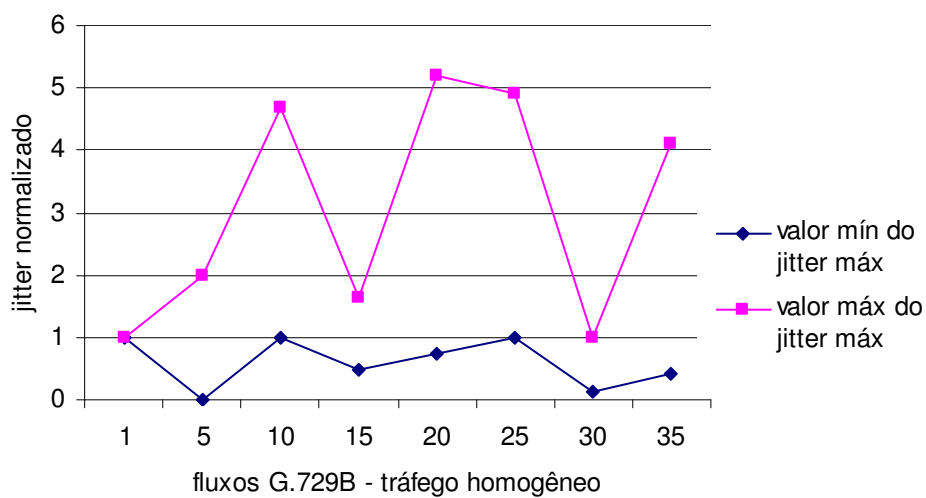


Figura A5.10: Variação do *jitter* máximo normalizado em função do carregamento do agregado com fluxos VoIP empregando *codec* G.729B

ANEXO 6

RESULTADOS TRÁFEGO HETEROGÊNEO G.711 E G.723.1

Esta condição de tráfego é formada por uma mistura de fluxos G.711 e G.723.1, sendo colocado um fluxo de controle para cada tipo de *codec*. O carregamento inicial foi de 15 fluxos G.711 e 1 fluxo G.723.1. Em cada passo da simulação foi retirado um fluxo G.711 e inserido o número de fluxos G.723.1 para completar a carga de 30% do E1. A condição final foi de 1 fluxo G.711 e 80 fluxos G.723.1.

Inicialmente são apresentados os resultados para o fluxo de controle G.711 e posteriormente para o fluxo de controle G.723.1.

A6.1 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE G.711

Nas Figuras A6.1 até A6.3 são apresentadas as distribuições do retardo no fluxo de controle G.711 em função do carregamento do agregado EF PHB.

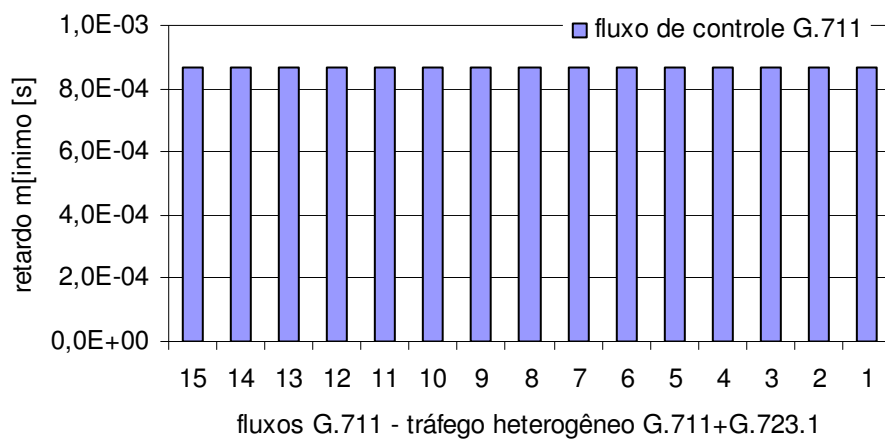


Figura A6.1: Retardo mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

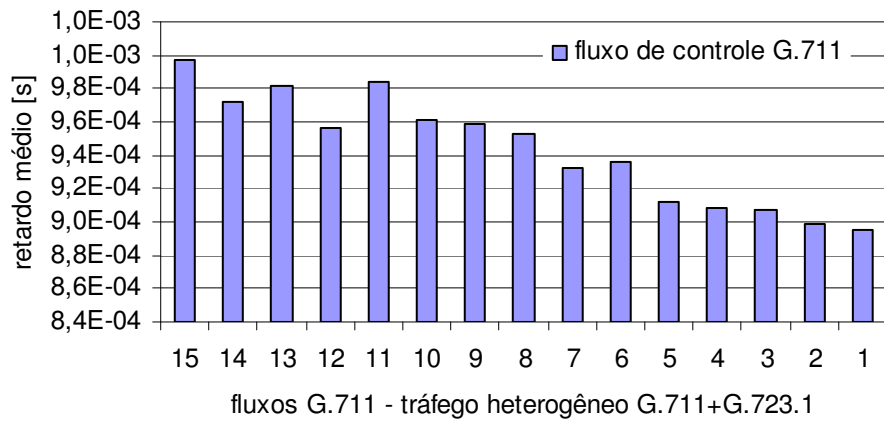


Figura A6.2: Retardo médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

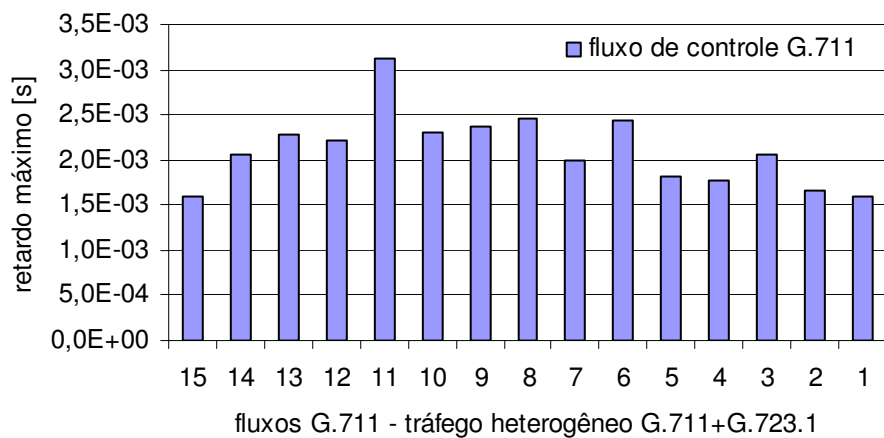


Figura A6.3: Retardo máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.2 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE G.711

Nas Figuras A6.4 até A6.6 são apresentadas as distribuições do *jitter* no fluxo de controle G.711 em função do carregamento do agregado EF PHB.

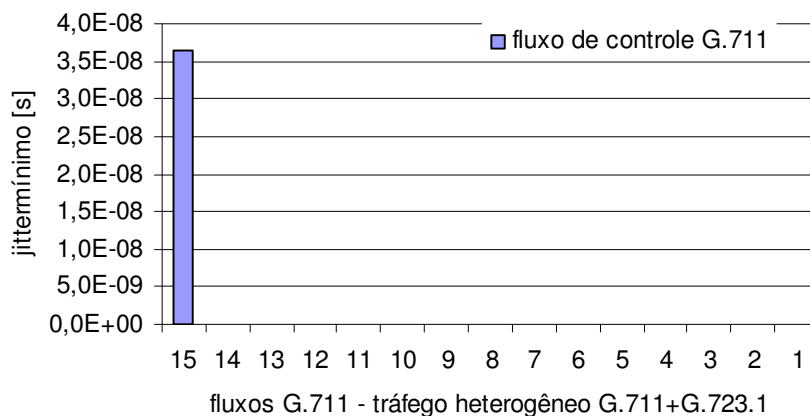


Figura A6.4: *Jitter* mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

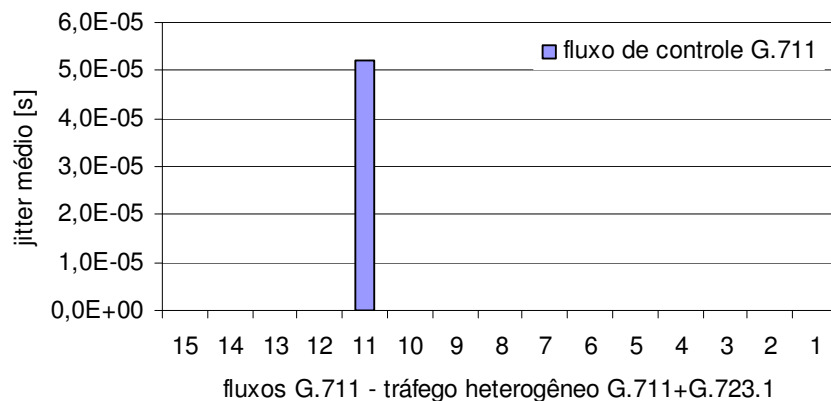


Figura A6.5: *Jitter* médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

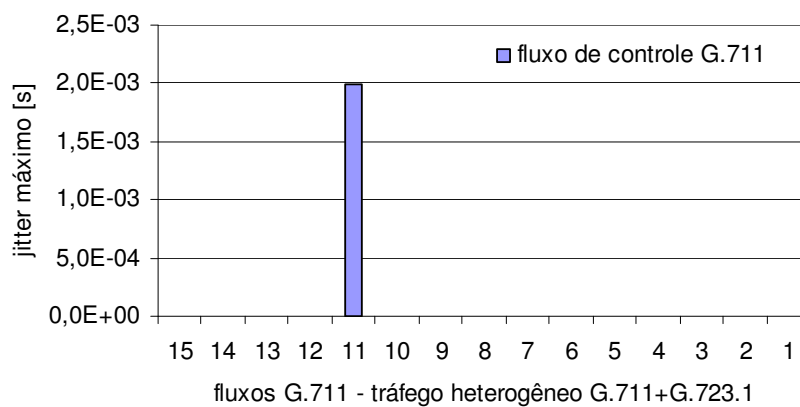


Figura A6.6: *Jitter* máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.3 VARIAÇÃO DO RETARDO NORMALIZADO NO FLUXO DE CONTROLE G.711

Nas Figuras A6.7 e A6.8 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.711 em função do carregamento do agregado EF PHB.

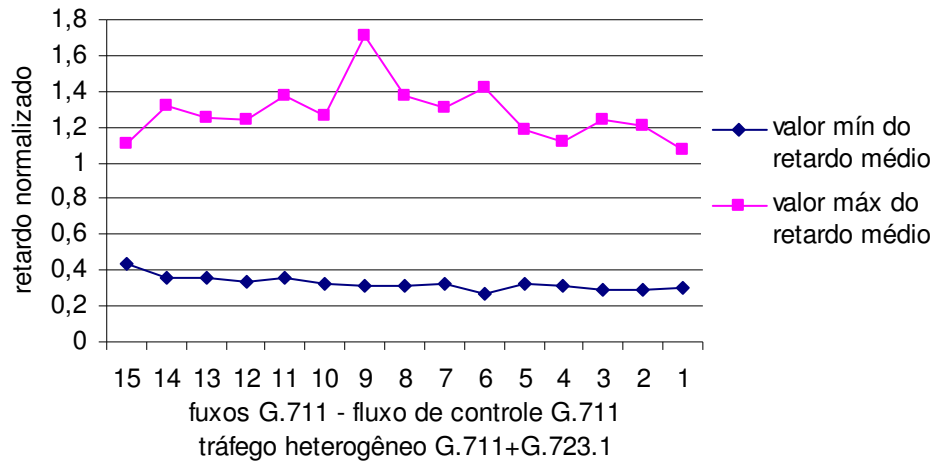


Figura A6.7: Variação do retardo médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

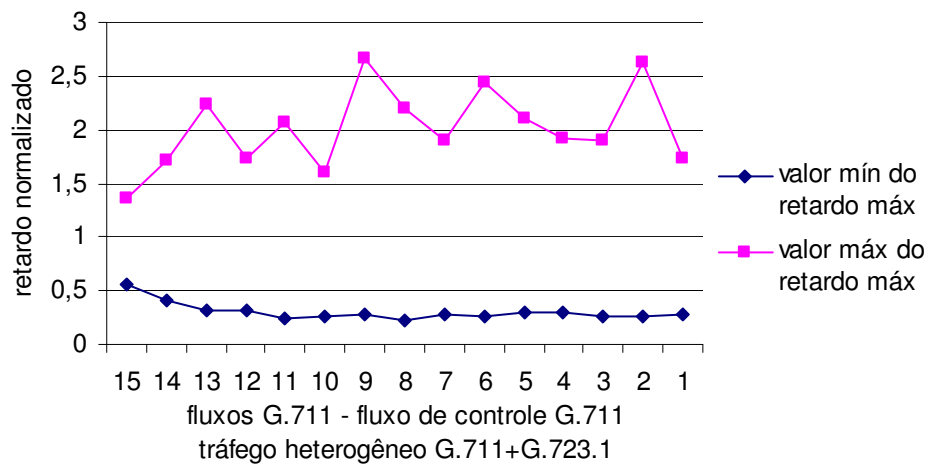


Figura A6.8: Variação do retardo máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.4 VARIAÇÃO DO JITTER NORMALIZADO NO FLUXO DE CONTROLE G.711

Nas Figuras A6.9 e A6.10 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.711 em função do carregamento do agregado EF PHB.

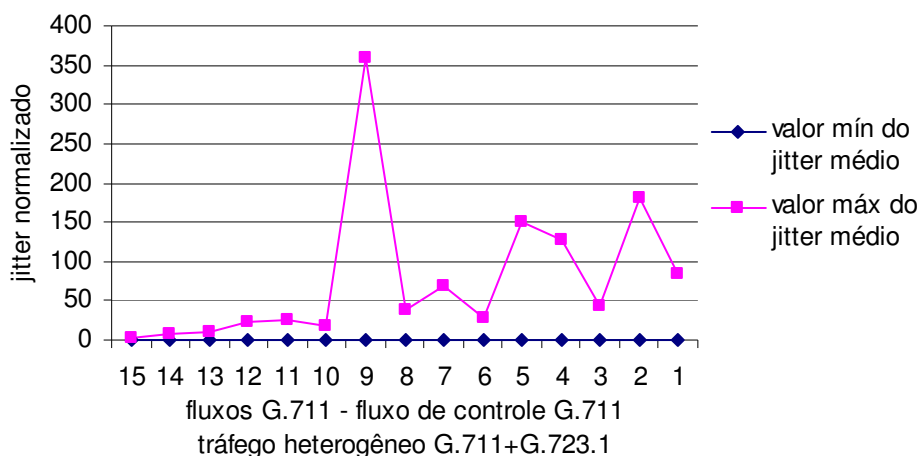


Figura A6.9: Variação do *jitter* médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

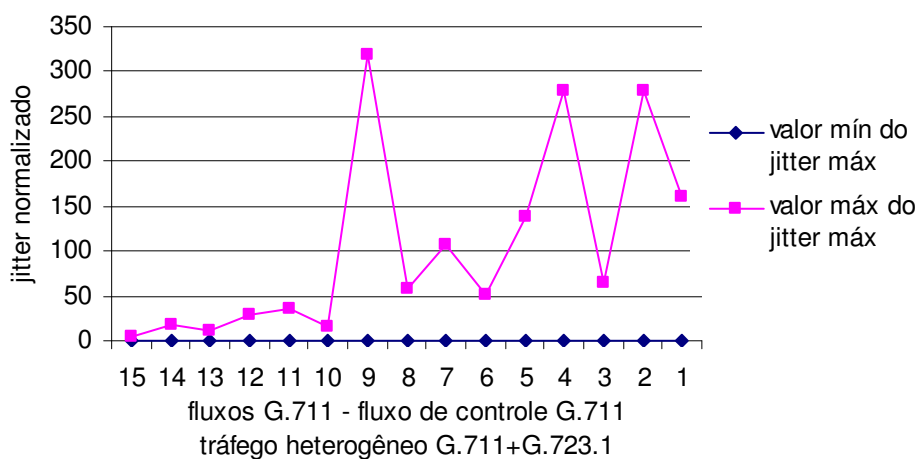


Figura A6.10: Variação do *jitter* máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.5 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A6.11 até A6.13 são apresentadas as distribuições do retardo no fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

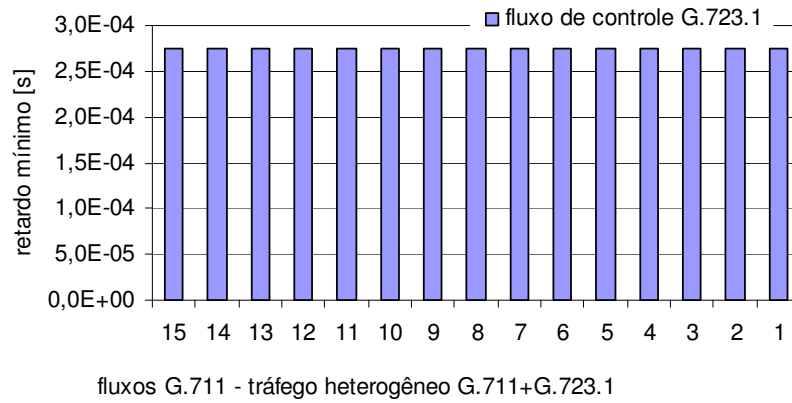


Figura A6.11: Retardo mínimo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

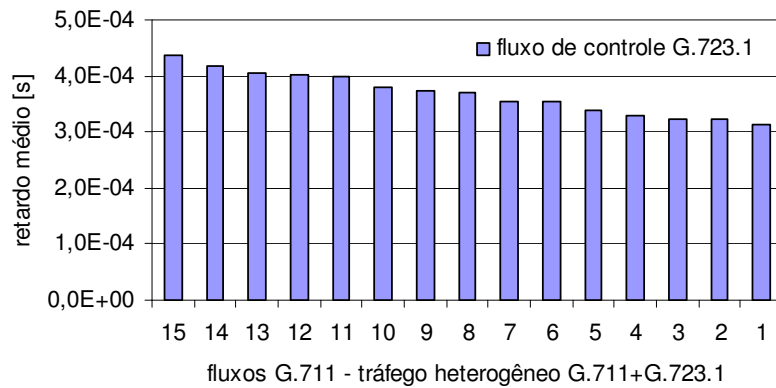


Figura A6.12: Retardo médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

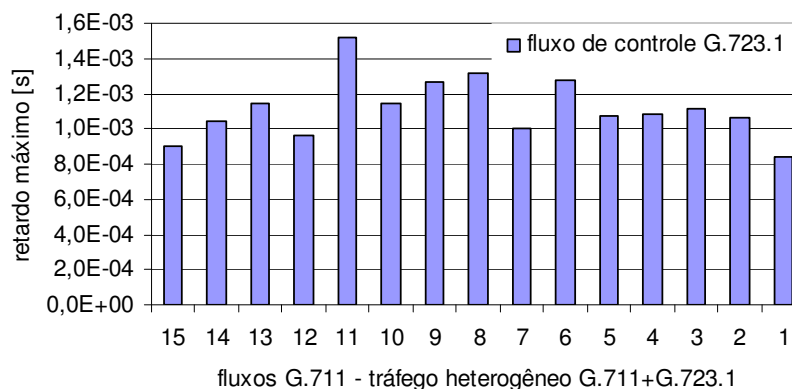


Figura A6.13: Retardo máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.6 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE G.723.1

Nas Figuras A6.14 e A6.15 são apresentadas as distribuições do *jitter* no fluxo de controle G.723.1 em função do carregamento do fluxo agregado VoIP.

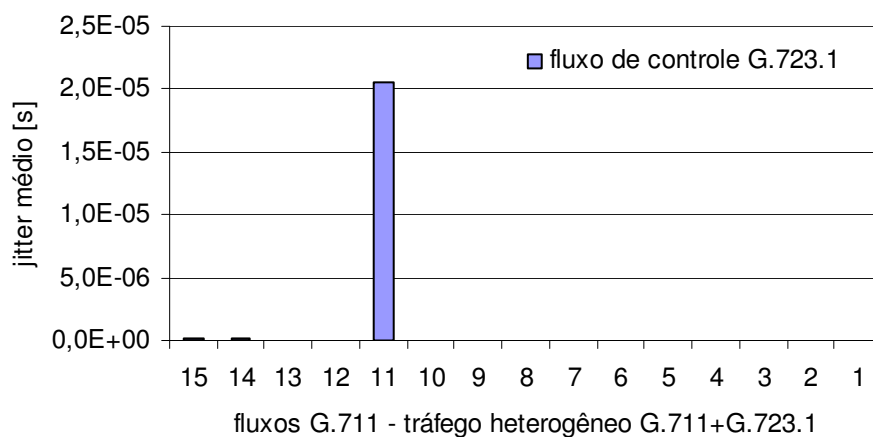


Figura A6.14: *Jitter* médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

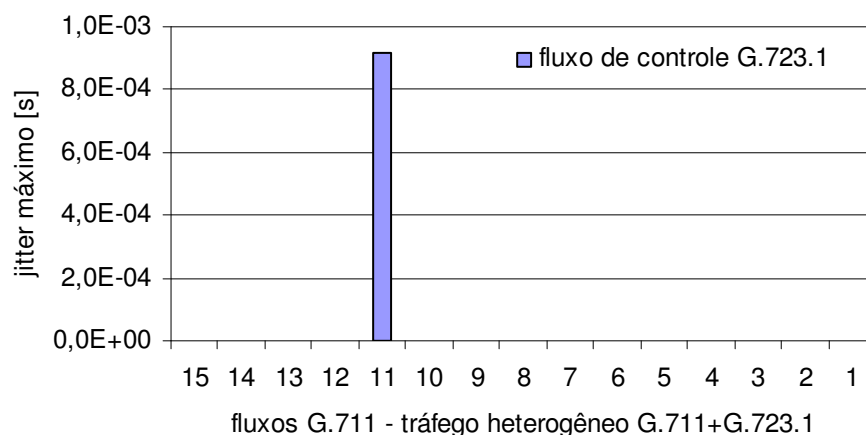


Figura A6.15: *Jitter* máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.7 VARIAÇÃO DO RETARDO NORMALIZADO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A6.16 e A6.17 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

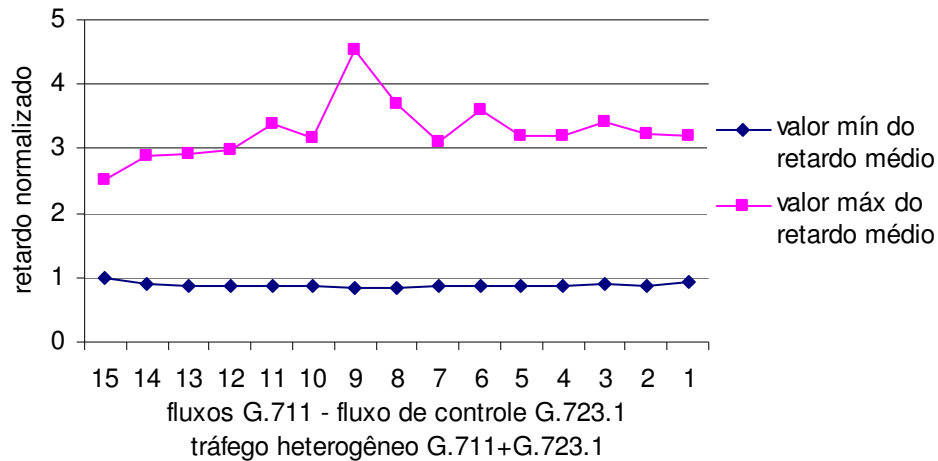


Figura A6.16: Variação do retardo médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

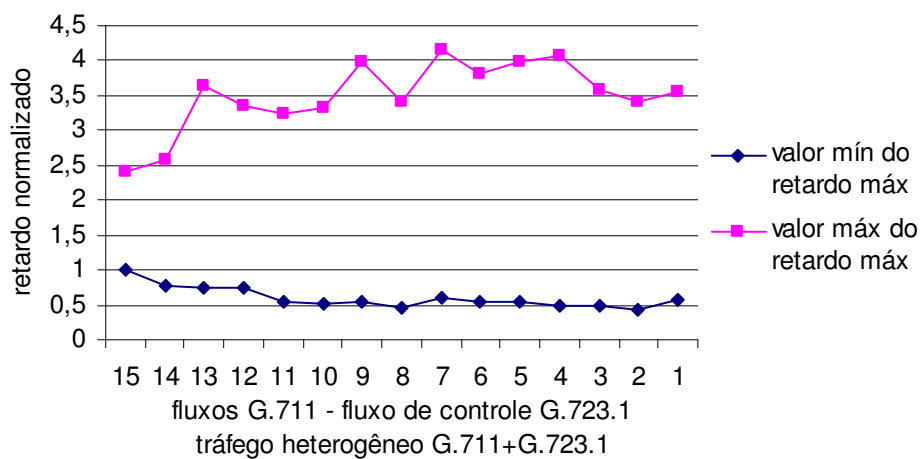


Figura A6.17: Variação do retardo máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

A6.8 VARIAÇÃO DO JITTER NORMALIZADO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A6.18 e A6.19 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

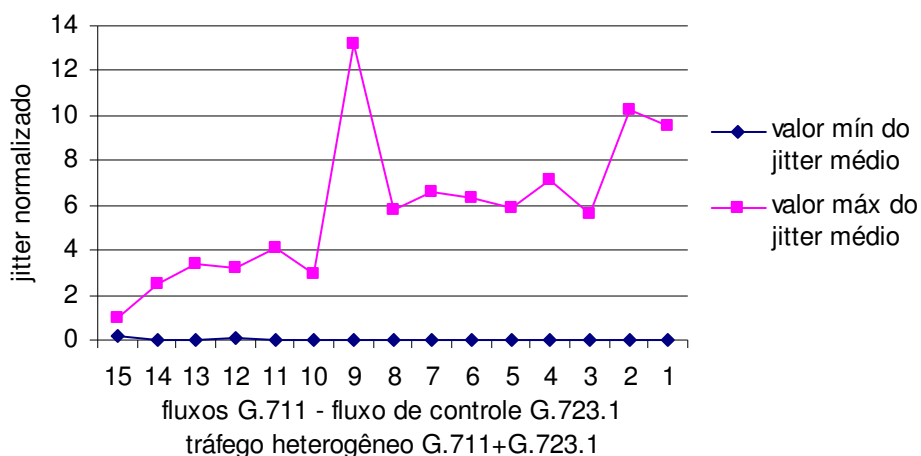


Figura A6.18: Variação do *jitter* médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

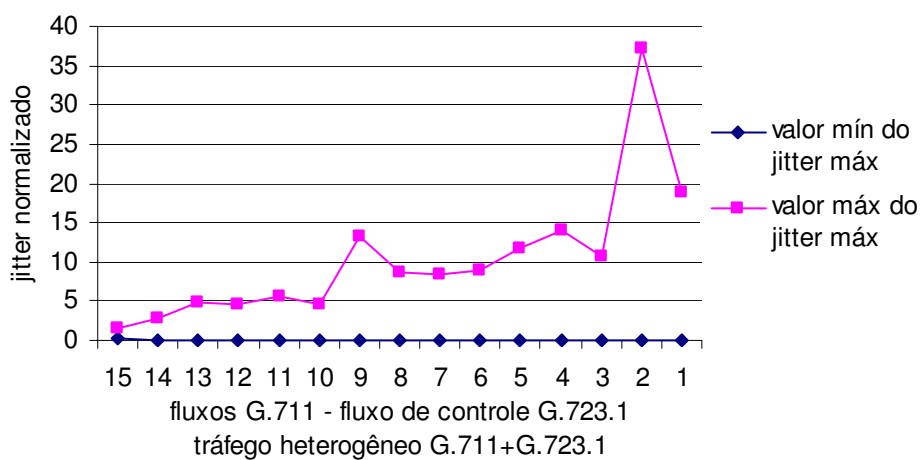


Figura A6.19: Variação do *jitter* máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.723.1

ANEXO 7

RESULTADOS TRÁFEGO HETEROGÊNEO G.711 E G.729B

Esta condição de tráfego é formada por uma mistura de fluxos G.711 e G.729B, sendo colocado um fluxo de controle para cada tipo de *codec*. O carregamento inicial foi de 15 fluxos G.711 e 1 fluxo G.729B. Em cada passo da simulação foi retirado um fluxo G.711 e inserido o número de fluxos G.729B para completar a carga de 30% do E1. A condição final foi de um fluxo G.711 e 30 fluxos G.729B.

Inicialmente são apresentados os resultados para o fluxo de controle G.711 e posteriormente para o fluxo de controle G.729B.

A7.1 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE G.711

Nas Figuras A7.1 até A7.3 são apresentadas as distribuições do retardo no fluxo de controle G.711 em função do carregamento do agregado EF PHB.

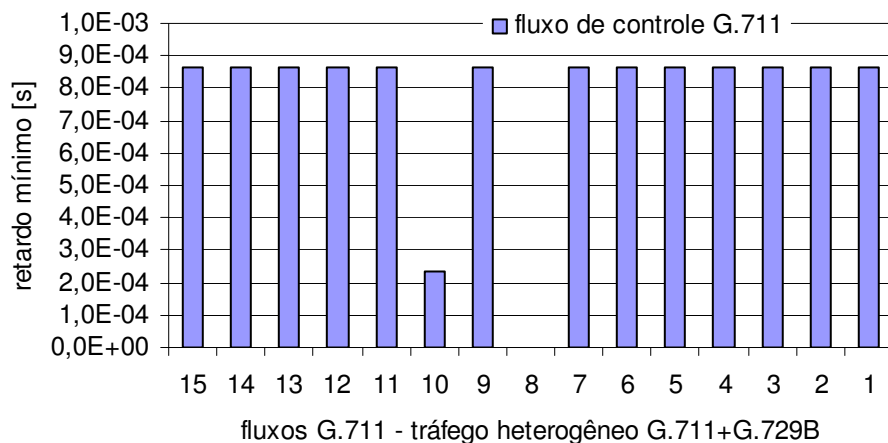


Figura A7.1: Retardo mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

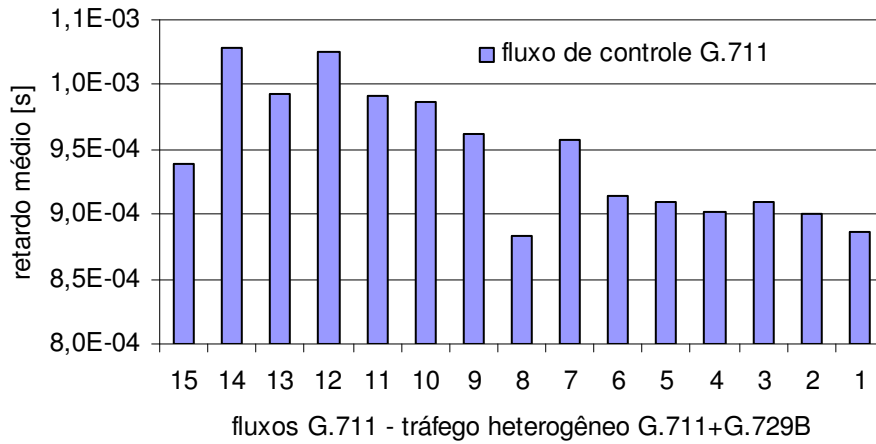


Figura A7.2: Retardo médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

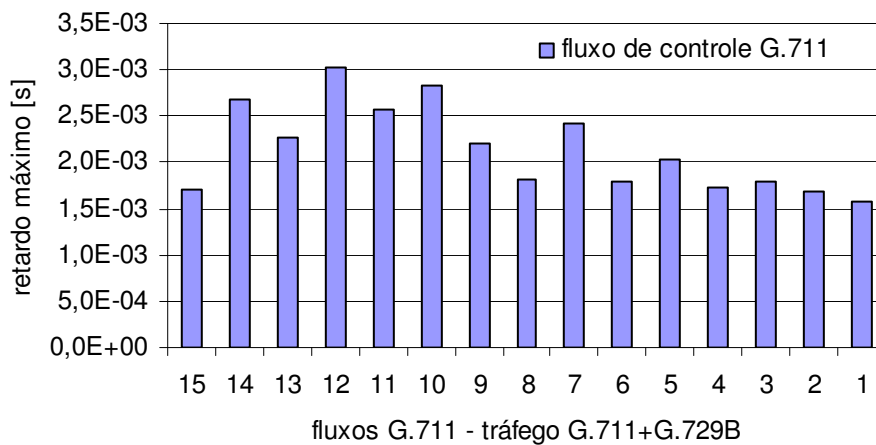


Figura A7.3: Retardo máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.2 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE G.711

Nas Figuras A7.4 até A7.6 são apresentadas as distribuições do *jitter* no fluxo de controle G.711 em função do carregamento do agregado EF PHB.

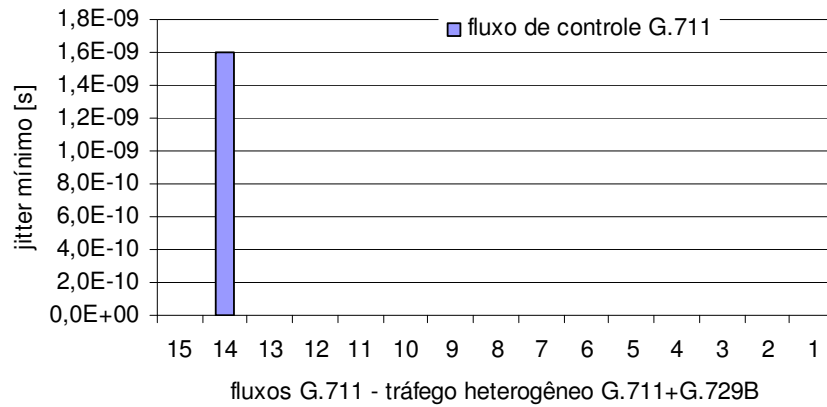


Figura A7.4: *Jitter* mínimo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

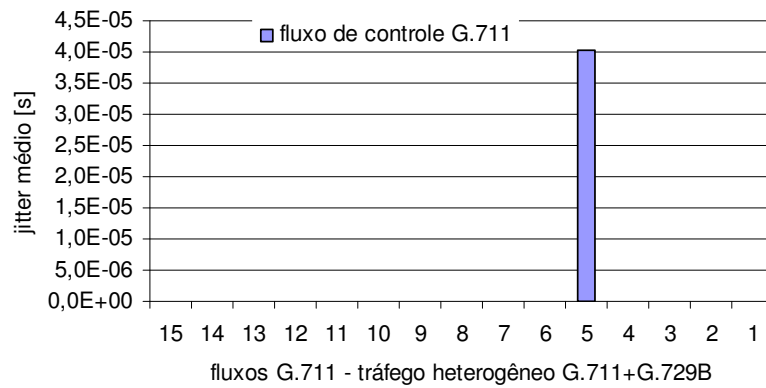


Figura A7.5: *Jitter* médio no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

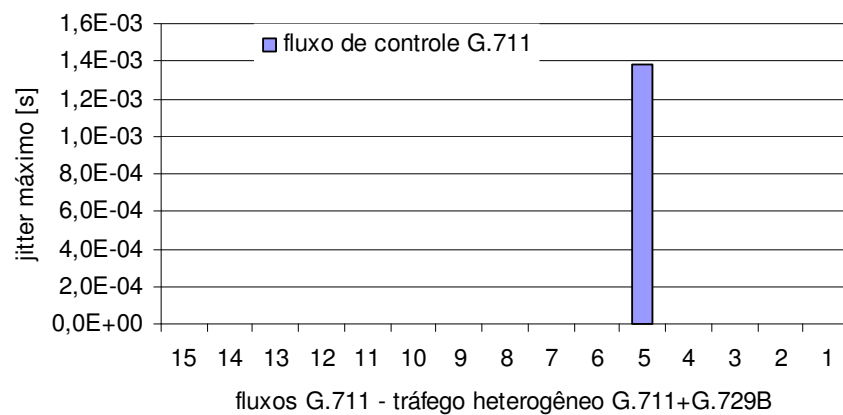


Figura A7.6: *Jitter* máximo no fluxo de controle G.711 em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.3 VARIAÇÃO DO RETARDO NORMALIZADO NO FLUXO DE CONTROLE G.711

Nas Figuras A7.7 e A7.8 são apresentadas as variações da métrica retardo normalizado em relação ao fluxo de controle G.711 em função do carregamento do agregado EF PHB.

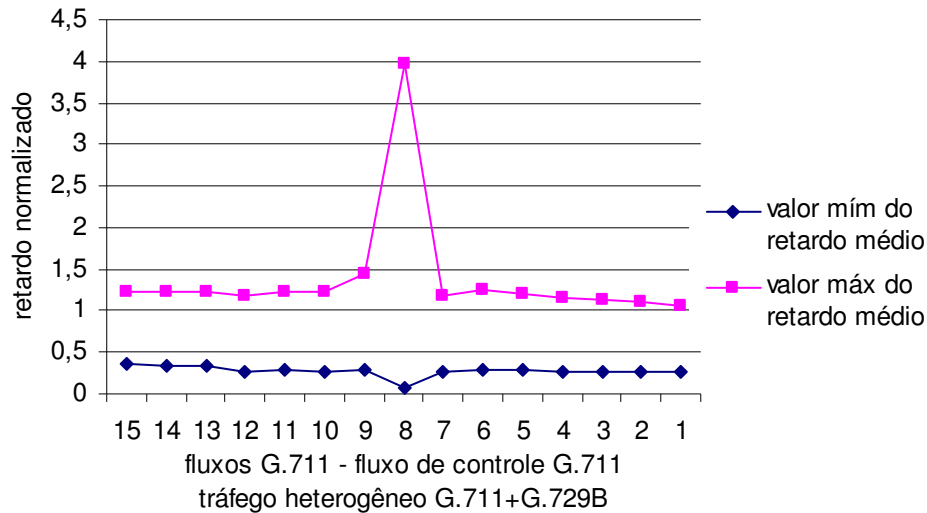


Figura A7.7: Variação do retardo médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

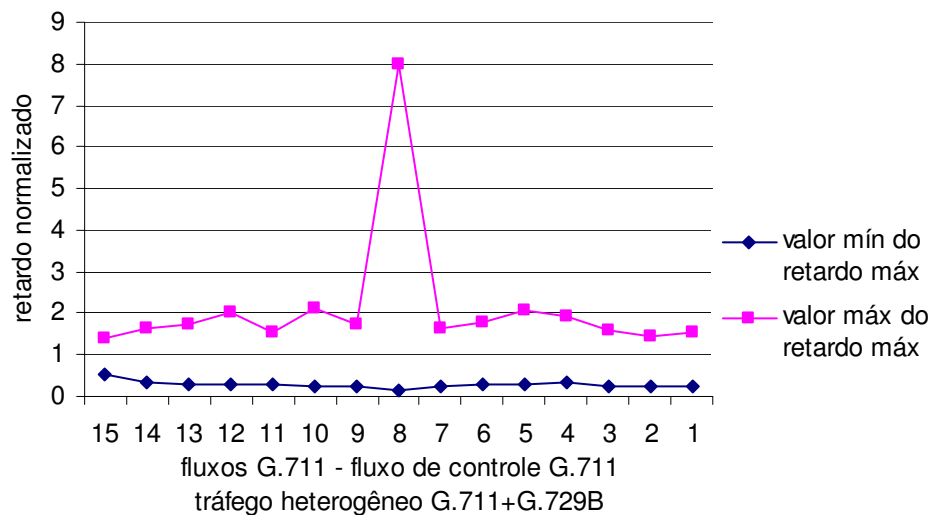


Figura A7.8: Variação do retardo máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.4 VARIAÇÃO DO JITTER NORMALIZADO NO FLUXO DE CONTROLE G.711

Nas Figuras A7.9 e A7.10 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.711 em função do carregamento do agregado EF PHB.

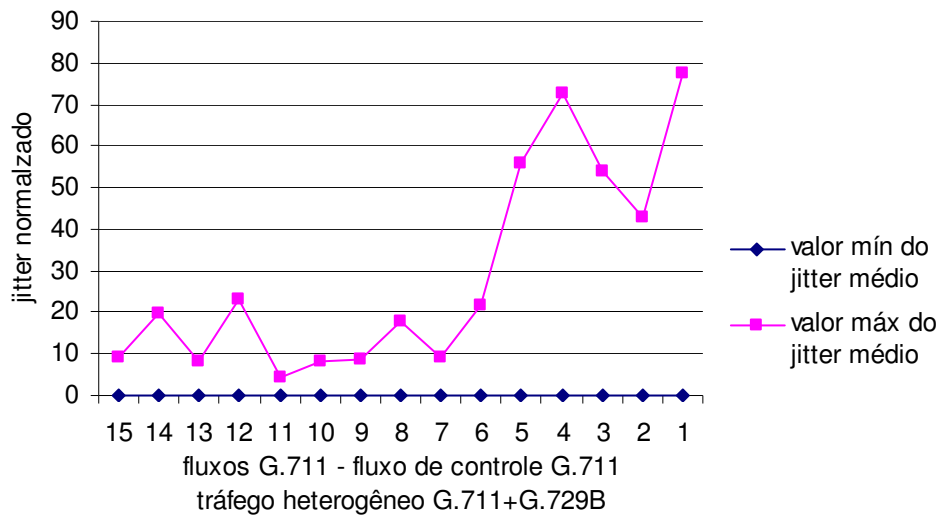


Figura A7.9: Variação do *jitter* médio normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

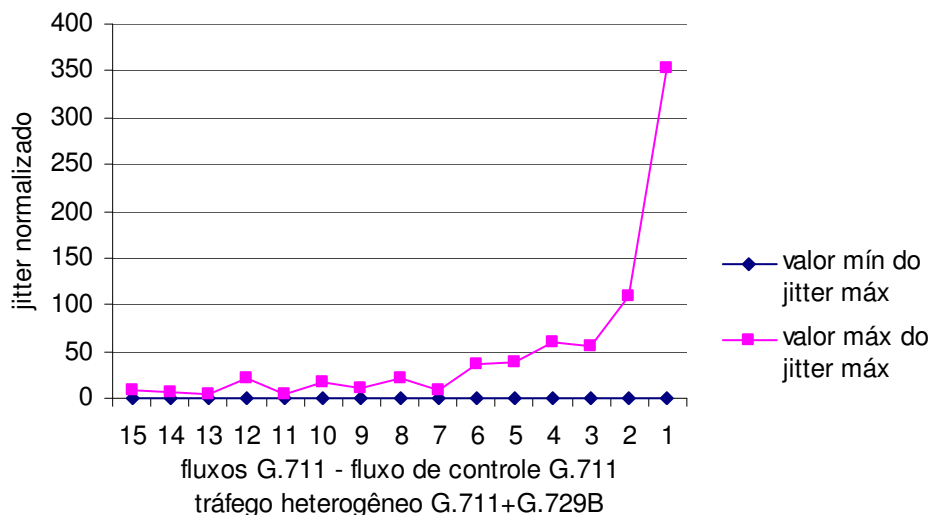


Figura A7.10: Variação do *jitter* máximo normalizado em relação ao fluxo de controle G.711, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.5 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE G.729B

Nas Figuras A7.11 até A7.13 são apresentadas as distribuições do retardo no fluxo de controle G.729B em função do carregamento do agregado EF PHB.

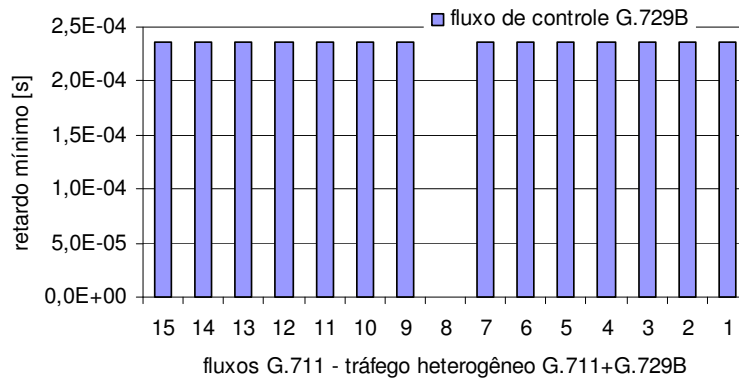


Figura A7.11: Retardo mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

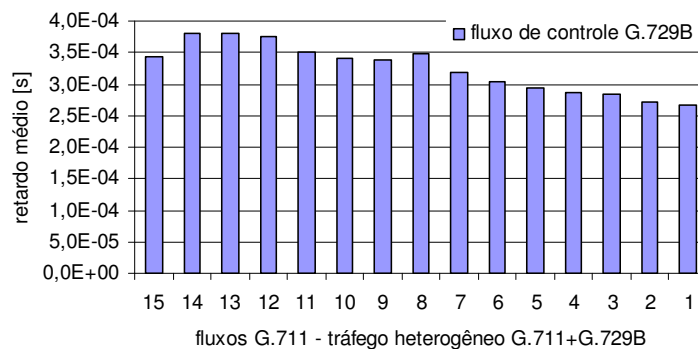


Figura A7.12: Retardo médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

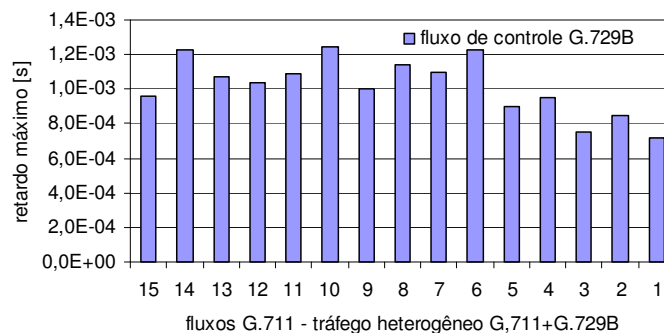


Figura A7.13: Retardo máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.6 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE G.729B

Nas Figuras A7.14 e A7.15 são apresentadas as distribuições do *jitter* no fluxo de controle G.729B em função do carregamento do fluxo agregado VoIP.

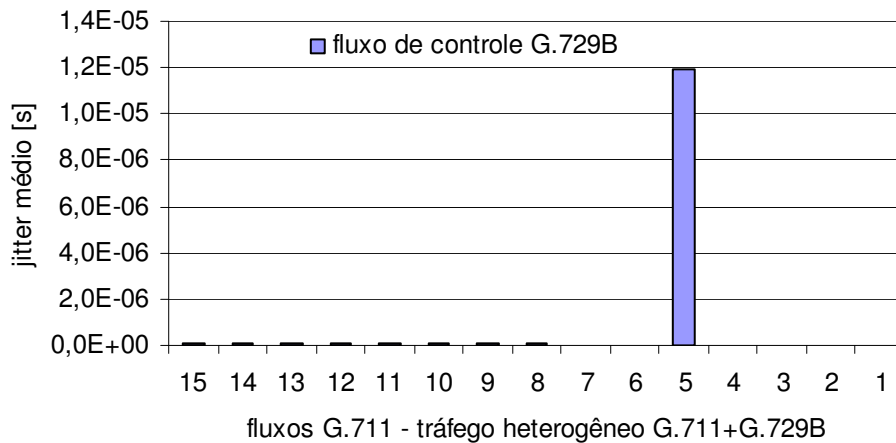


Figura A7.14: *Jitter* médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

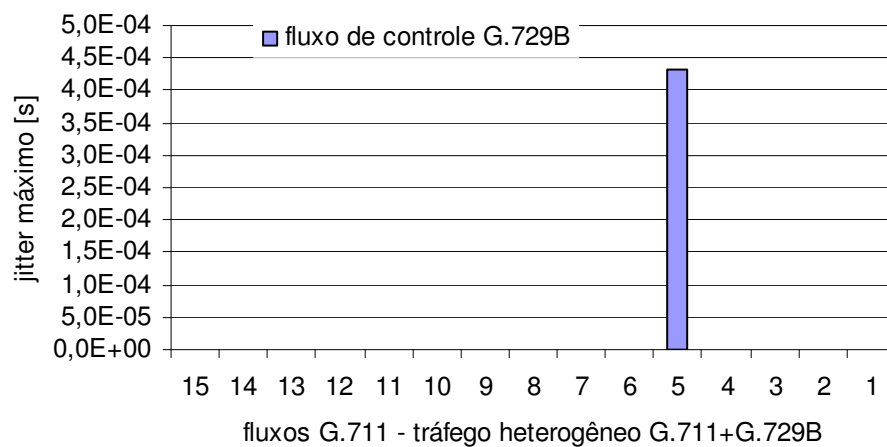


Figura A7.15: *Jitter* máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.7 VARIAÇÃO DO RETARDO NORMALIZADO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A7.16 e A7.17 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

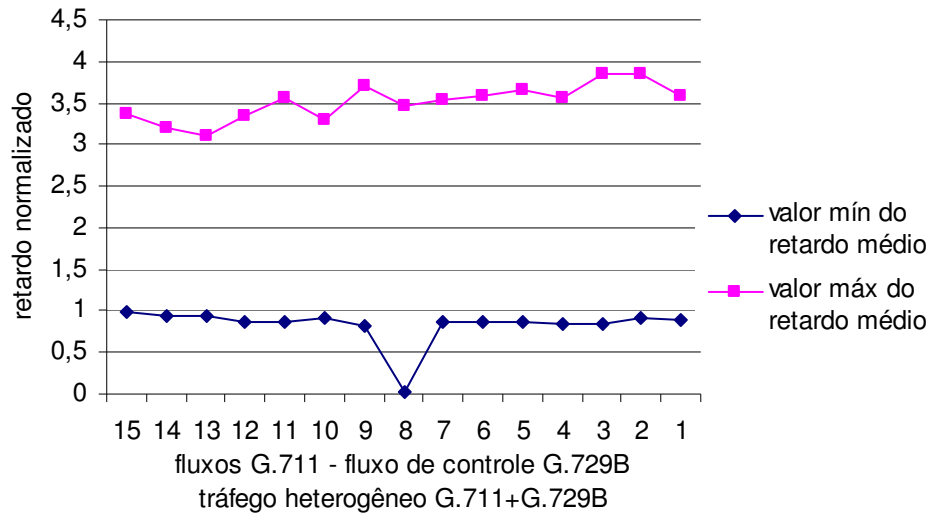


Figura A7.16: Variação do retardo médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

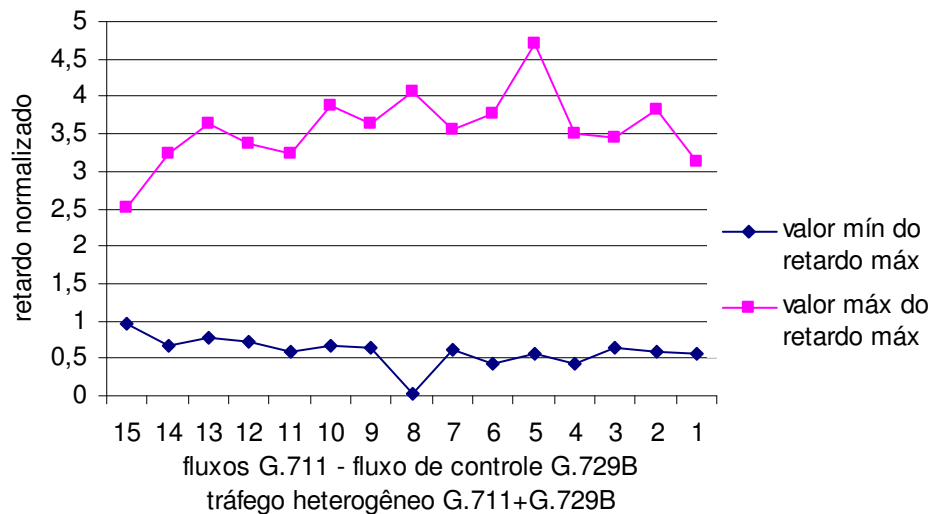


Figura A7.17: Variação do retardo máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

A7.8 VARIAÇÃO DO JITTER NORMALIZADO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A7.18 e A7.19 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

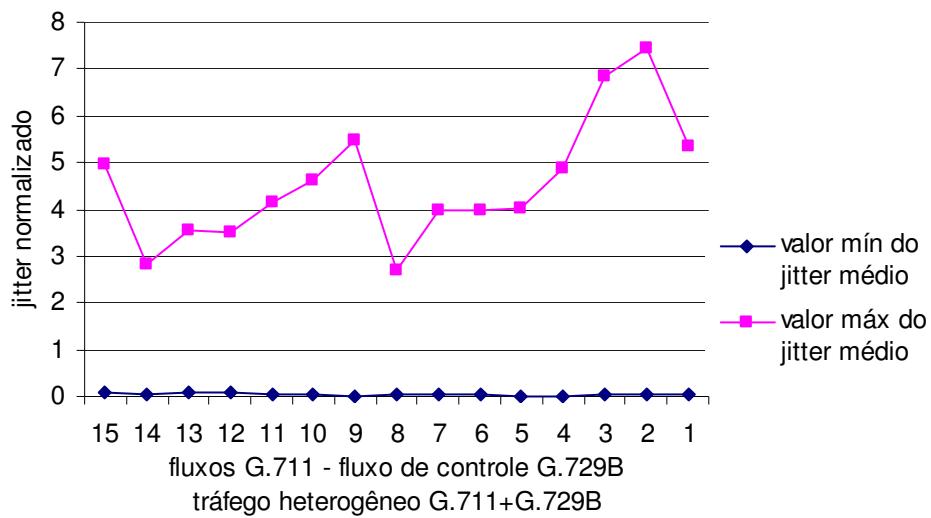


Figura A7.18: Variação do *jitter* médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

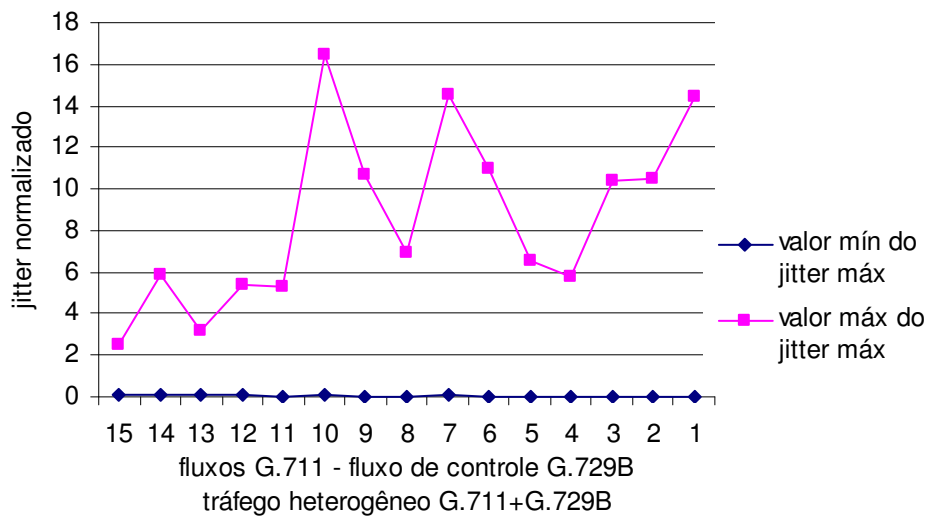


Figura A7.19: Variação do *jitter* máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.711 e G.729B

ANEXO 8

RESULTADOS TRÁFEGO HETEROGÊNEO G.729B E G.723.1

Esta condição de tráfego é formada por uma mistura de fluxos G.729B e G.723.1, sendo colocado um fluxo de controle para cada tipo de *codec*. O carregamento inicial foi de 34 fluxos G.729B e 1 fluxo G.723.1. Em cada passo da simulação foi retirado um fluxo G.729B e inserido o número de fluxos G.723.1 para completar a carga de 30% do E1. A condição final foi de 1 fluxo G.729B e 48 fluxos G.723.1.

Inicialmente são apresentados os resultados para o fluxo de controle G.729B e posteriormente para o fluxo de controle G.723.1.

A8.1 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE G.729B

Nas Figuras A8.1 até A8.3 são apresentadas as distribuições do retardo no fluxo de controle G.729B em função do carregamento do agregado EF PHB.

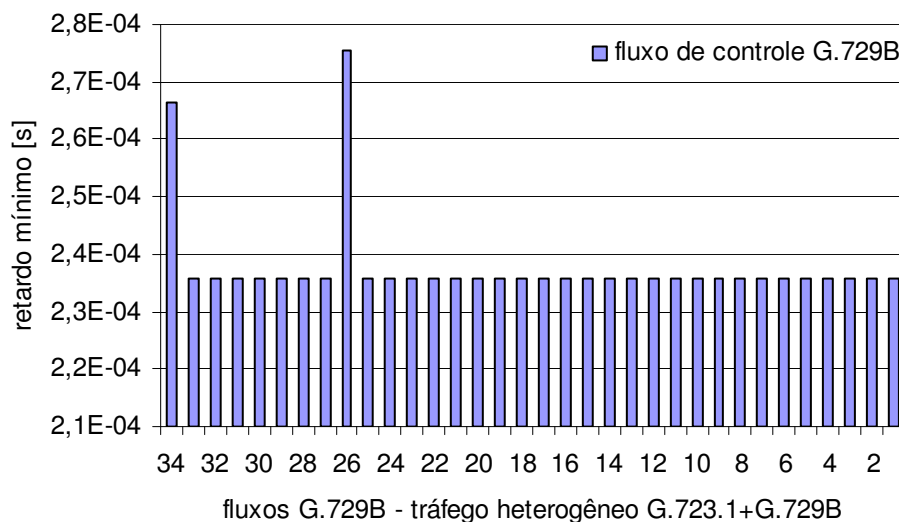


Figura A8.1: Retardo mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

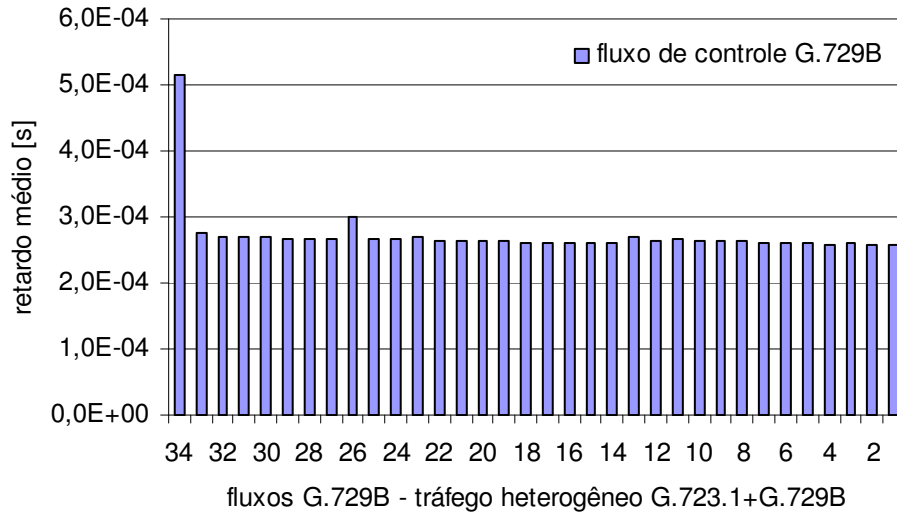


Figura A8.2: Retardo médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

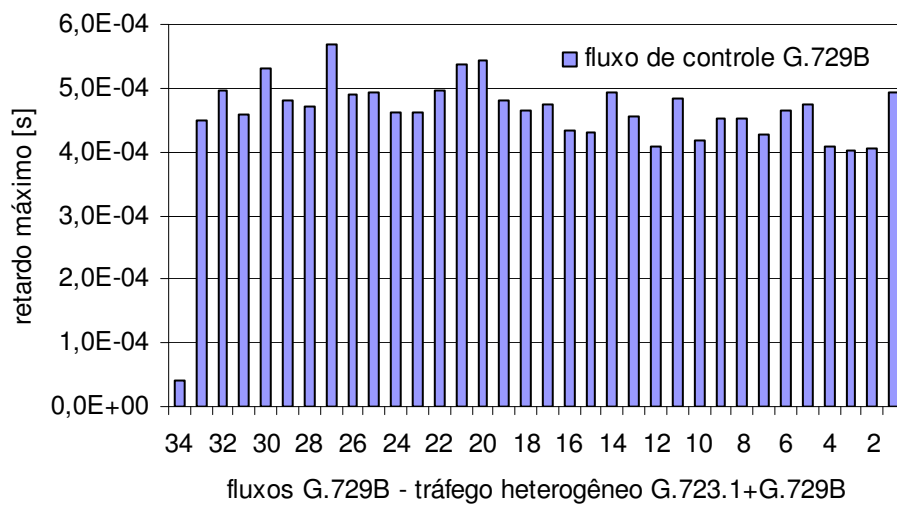


Figura A8.3: Retardo máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.2 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE G.729B

Nas Figuras A8.4 até A8.6 são apresentadas as distribuições do *jitter* no fluxo de controle G.729B em função do carregamento do agregado EF PHB.

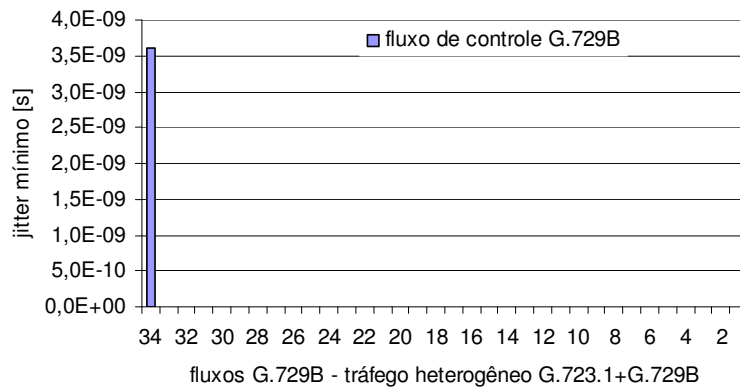


Figura A8.4: *Jitter* mínimo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

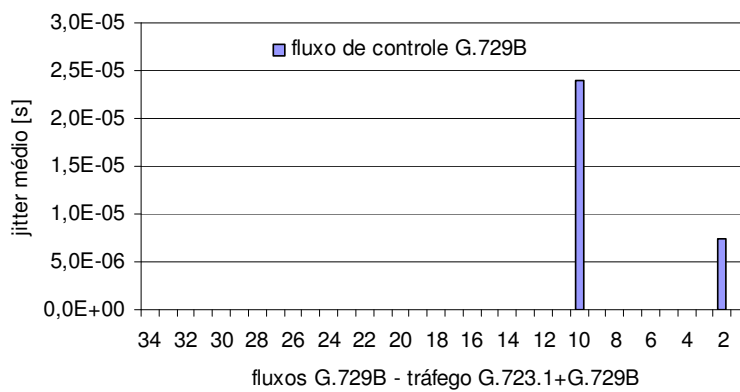


Figura A8.5: *Jitter* médio no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

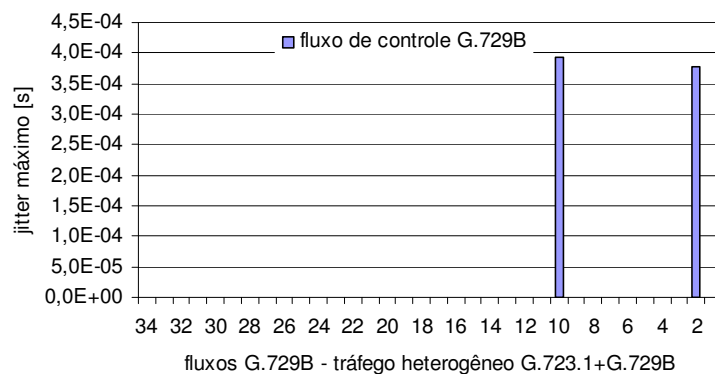


Figura A8.6: *Jitter* máximo no fluxo de controle G.729B em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.3 VARIAÇÃO DO RETARDO NORMALIZADO NO FLUXO DE CONTROLE G.729B

Nas Figuras A8.7 e A8.8 são apresentadas as variações da métrica retardo normalizado em relação ao fluxo de controle G.729B em função do carregamento do agregado EF PHB.

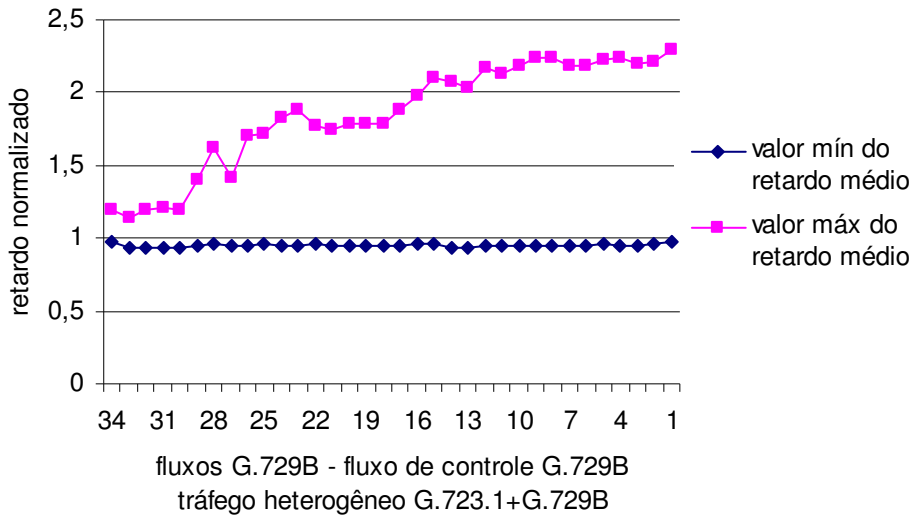


Figura A8.7: Variação do retardo médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

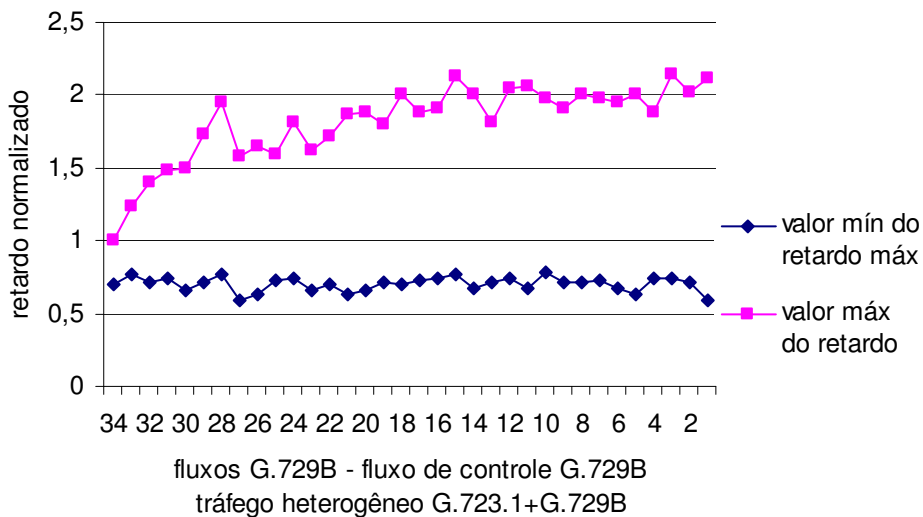


Figura A8.8: Variação do retardo máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.4 VARIAÇÃO DO JITTER NORMALIZADO NO FLUXO DE CONTROLE G.729B

Nas Figuras A8.9 e A8.10 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.729B em função do carregamento do agregado EF PHB.

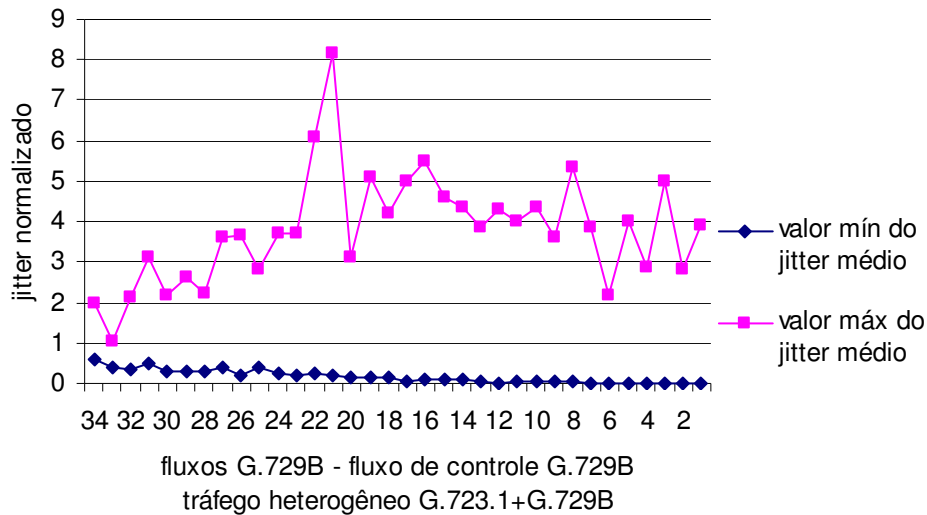


Figura A8.9: Variação do *jitter* médio normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

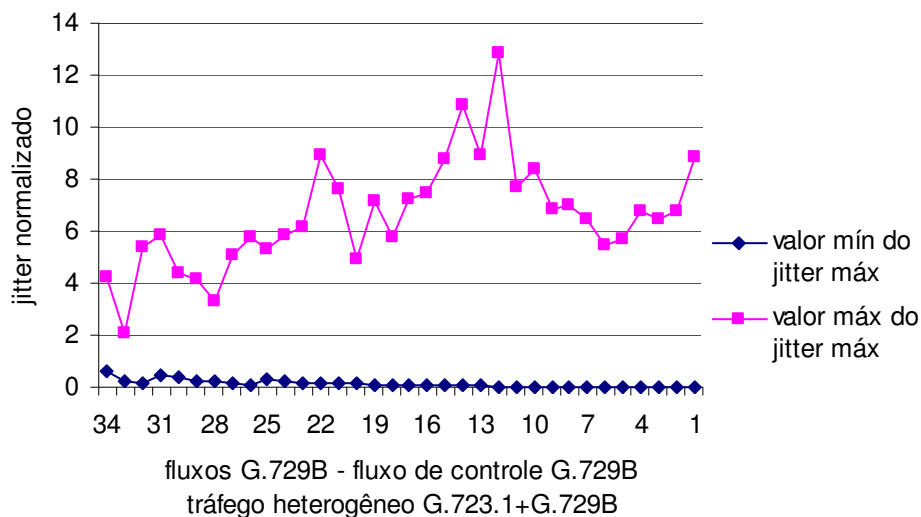


Figura A8.10: Variação do *jitter* máximo normalizado em relação ao fluxo de controle G.729B, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.5 DISTRIBUIÇÃO DO RETARDO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A8.11 até A8.13 são apresentadas as distribuições do retardo no fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

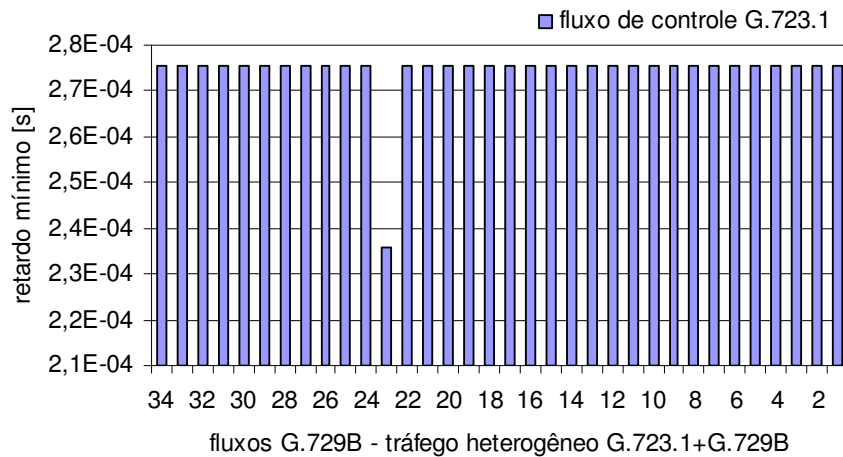


Figura A8.11: Retardo mínimo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

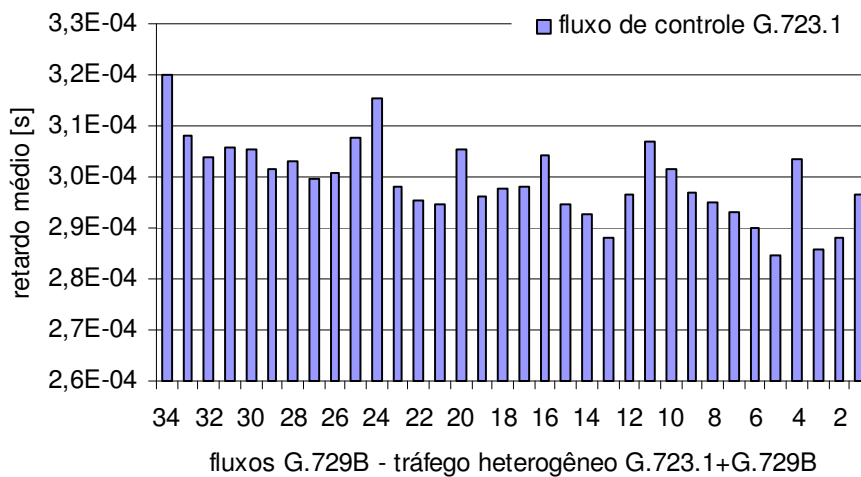


Figura A8.12: Retardo médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

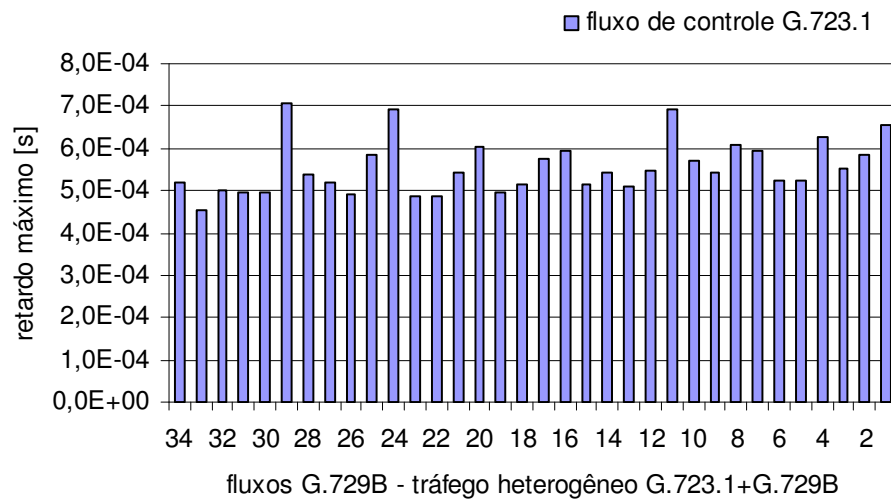


Figura A8.13: Retardo máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.6 DISTRIBUIÇÃO DO JITTER NO FLUXO DE CONTROLE G.723.1

Nas Figuras A8.14 e A8.15 são apresentadas as distribuições do *jitter* no fluxo de controle G.723.1 em função do carregamento do fluxo agregado VoIP.

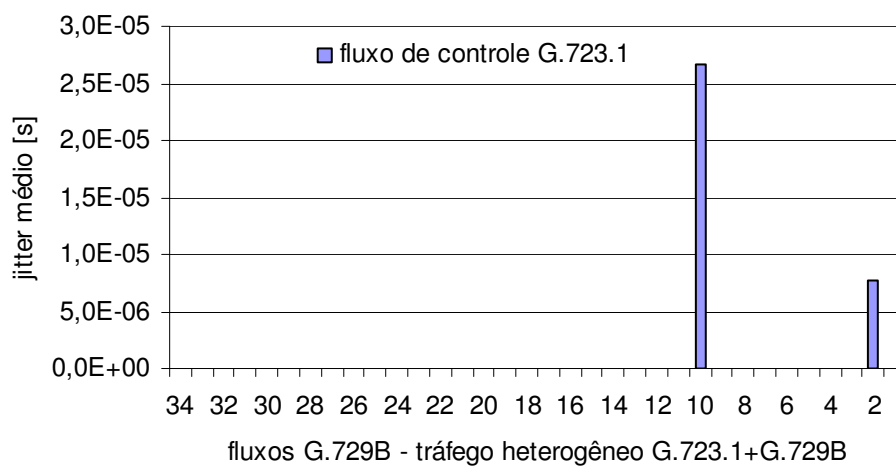


Figura A8.14: *Jitter* médio no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

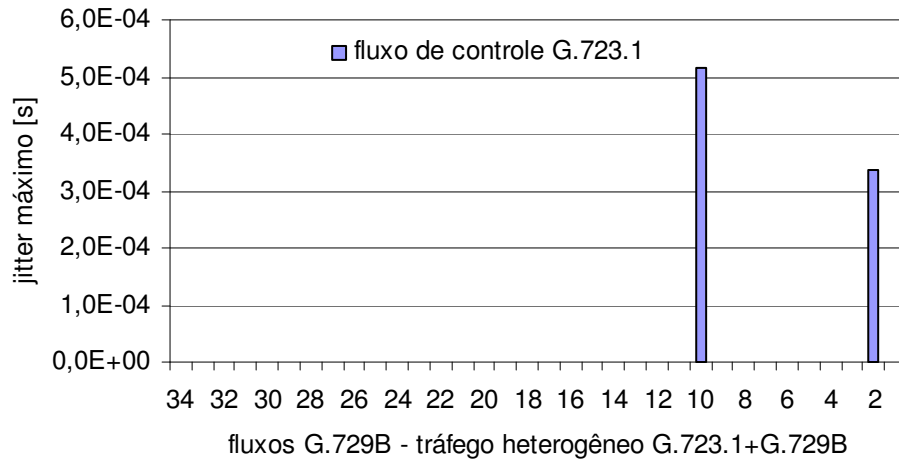


Figura A8.15: Jitter máximo no fluxo de controle G.723.1 em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.7 VARIAÇÃO DO RETARDO NORMALIZADO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A8.16 e A8.17 são apresentadas as variações da métrica retardo normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

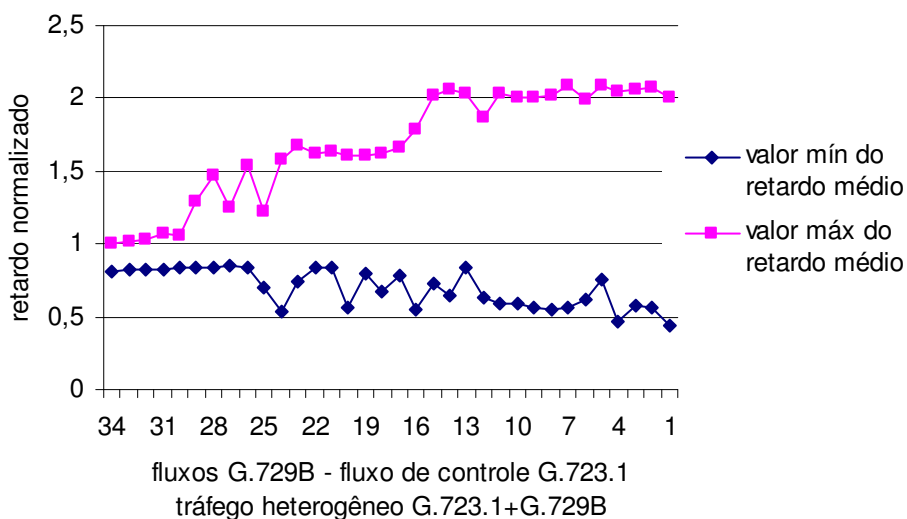


Figura A8.16: Variação do retardo médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

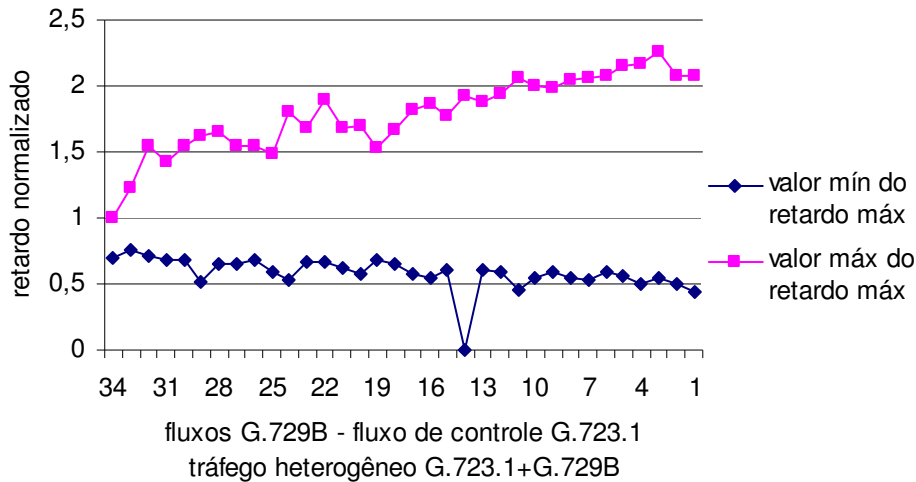


Figura A8.17: Variação do retardo máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

A8.8 VARIAÇÃO DO JITTER NORMALIZADO NO FLUXO DE CONTROLE G.723.1

Nas Figuras A8.18 e A8.19 são apresentadas as variações da métrica *jitter* normalizada em relação ao fluxo de controle G.723.1 em função do carregamento do agregado EF PHB.

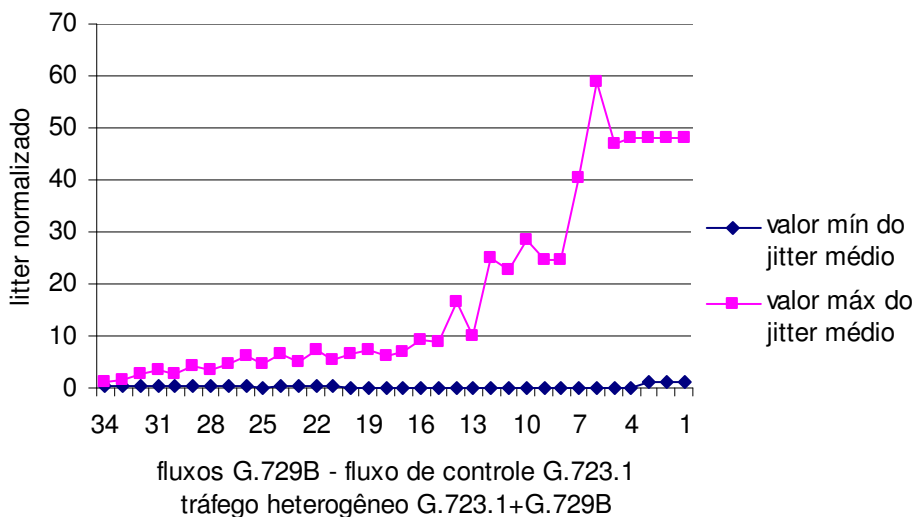


Figura A8.18: Variação do *jitter* médio normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

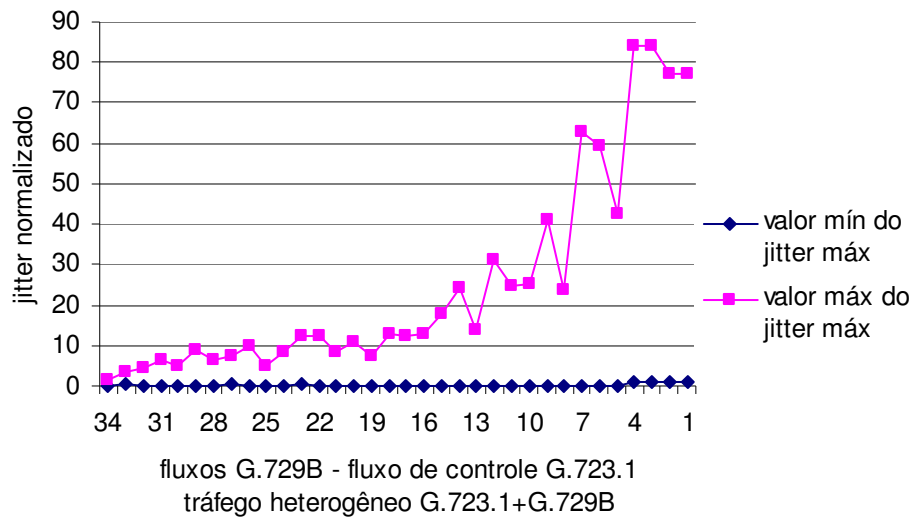


Figura A8.19: Variação do *jitter* máximo normalizado em relação ao fluxo de controle G.723.1, em função do carregamento com fluxos VoIP empregando *codecs* G.729B e G.723.1

REFERÊNCIAS BIBLIOGRÁFICAS

- [AA1997] Adas, A. **Traffic Models in Broadband Networks**. IEEE Communications Magazine, Vol. 35, no. 7, July 1997.
- [AA2003] Asgari, A.; et al. **Building Quality-of-Service Monitoring Systems for Traffic Engineering and Service Management**. Journal of Network and Systems Management, Vol. 11, No 4, December 2003.
- [AG2002] Armitage, G.; *et al.* A Delay Bound Alternative Revision of RFC 2598. **IETF RFC 3248**, March 2002. (
- [AR2004] Ali, R. B.; Pierre, S.; Lemieux, Y. **DiffServ QoS Evaluation of Multimedia Telephony**. CCECE 2004-CCGEI 2004, Niagara Falls, 2004. IEEE, 2004.
- [BB2002] Bellalta, B.; Oliver, M.; Rincón, D. **Capacity and Traffic Analysis of VoIP Services over GPRS Mobile Networks**. Report TIC2001-0956-C04-04, 2001.
- [BJ1995] Beran, J.; Sherman, R.; Taqq, M. S.; Willinger, W.. **Long-Range Depedence in Variable-Bit-Rate Video Traffic**. IEEE Transactions on Communications, Vol. 43, no. 2/3/4, February/March/April 1995.
- [BP1965] Brady, P. T. **A technique for investigation on-off patterns of speech**. BSTJ, 44 (1): 1-22, January 1965.
- [BP1968] Brady, P. T. **A statistical analysis of on-off patterns in 16 conversations**. BSTJ, 47 (1): 73-91, January 1968.
- [BP1969] Brady, P. T. **A model for generating on-off speech patterns in two-way conversation**. BSTJ, 48 (9): 2445-2472, September 1969.
- [BR1994] Braden, R.; Clark, D.; Shenker, S. Integrated Services in the Internet Architecture: an Overview. **IETF RFC 1633**, June 1994.
- [BS1998] Blake, S.; et al. An Architecture for Differentiated Service. **IETF RFC 2475**, December 1998. (*Updated by RFC3260*).
- [CC2000] Chuah, C. N.; Katz, R. H. **Characterizing Packet Audio Streams from Internet Multimedia Applications**. University of California, Berkeley, USA, 2000.
- [CE1998] Crawley, E.; et al. A Framework for QoS-based Routing in the Internet. **IETF RFC 2386**, August, 1998.

- [CE2002] Casilari, E.; Montes, H.; Sandoval, F. **Modeling of Voice Traffic over IP Networks**. CSNDSP 2002, Network Communications K1.5, July 2002.
- [CM1997] Crovella, M.; Bestavros, A. **Self-similarity in World Wide Web traffic: Evidence and possible causes**. IEEE/ACM Transactions on Networking, vol. 5, no. 6, pp. 835-846, December 1997.
- [CR1997] Cox, R. V. **Three new speech coders from the ITU cover a range of applications**. IEEE Communications Magazine, Vol. 35, no.9, September 1997.
- [DT2003] Dang, T. D.; Sonkoly, B.; Molnár, S. **Fractal Analysis and Modeling of VoIP Traffic**. Report to Ministry of Education, IKTA-0092/2002, Hungary, 2003.
- [EA2003] Estepa, A.; Estepa, R.; Vozmediano J. **Paquetization and silence influence on VoIP traffic profile**. Lecture Notes on Computer Science 2899 – Multimedia Interactive Protocols and Systems MIPS 2003, Naples, Italy, 2003.
- [EA2004] Estepa, A.; Estepa, R.; Vozmediano, J. **A New Approach for VoIP Traffic Characterization**. IEEE Communications Letters, Vol. 8, No. 10, October 2004.
- [ET2003] Engel, T.; et al. **AQUILA: Adaptive Resource Control for QoS using an IP-based Layered Architecture**. IEEE Communications Magazine. January 2003. pg. 46-53.
- [FV1994] Frost, V.; Melamed, B. **Traffic Modeling for Telecommunications Networks**. IEEE Communications Magazine, pp. 70-81, March 1994.
- [GD2002] Grossman, D. New terminology and Classifications for DiffServ. **IETF RFC 3260**. April, 2002. (*Updates RFC 2475 and RFC 2597*)
- [GJ2003] Gozdecki, J.; Jajszczyk, A.; Stankiewicz, R. **Quality of Service Terminology in IP Networks**. IEEE Communications Magazine, March 2003, pg. 153-159.
- [HJ1999] Heinanen, J.; et al. Assured Forwarding PHB Group. **IETF RFC 2597**, June 1999. (*Updated by RFC 3260*)
- [HO2002] Hersent, O.; Guide, D.; Petit, J. P. **Telefonia IP: comunicação multimídia baseada em pacotes**. Prentice Hall, 2002, São Paulo, Brasil. ISBN 85-88639-02-5, pp.77-79

- [HO2003] Holland, O. **Over-the-air Downloads for Mobile Terminal Reconfiguration**. Technical Report. Centre for Telecommunications Research, King's College London, 2003.
- [HW2001] Hardy, W. C. **QoS Measurement and Evaluation of Telecommunications Quality of Service**. Wiley, 2001.
- [JV1999] Jacobson, V; Nichols, K.; Poduri, K.. An Expedited Forwarding PHB. **IETF RFC 2598**, June 1999. (*Obsoleted by RFC 3246*)
- [JR1991] Jain, R. **The Art of Computer Systems Performance Analysis: techniques for experimental design, measurement, simulation and modeling**. John Wiley & Sons, New York, 1991. ISBN 0-471-50336-3.
- [JW2000] Jiang, W.; Schulzrinne, H. **Analysis of On-Off Patterns in VoIP and their Effect on Voice Traffic Aggregation**. Proceedings of 9th IEEE International Conference on Computer Communication Networks, pp.82-7, Las Vegas, NV, October 2000.
- [KT1998] Kostas T. J.; Borella, M. S.; Sidhu, I.; Schuster, G. M.; Gabrieic, J.; Mahler, J. **Real-time voice over packet-switched networks**. IEEE Network, Vol. 12, no. 1, pp.18-27, Jan/Feb. 1998.
- [LH2003] Lu, Hui-Lan; Faynberg, I. **An Arquitectural Framework for Support of Quality of Service in Packet Network**. IEEE Communications Magazine, June 2003, pg. 98-105.
- [LW1994] Leland, W. E.; Taquq, M. S.; Willinger, W.; Wilson, D. V. **On the Self-Similar Nature of Ethernet Traffic** (Extended Version). IEEE/ACM Transaction on Networking, 2(1): 1-15, February 1994.
- [LX2003] Liu, X.; Baras, J. S. **Aggregation of Heavy-Tailed On-Off Flows is Multifractal**. University of Maryland, College Park, USA, 2003.
- [ME2003] Mykoniati, E.; et al. **Admission Control for Providing QoS in DiffServ IP Networks: The TEQUILA Approach**. IEEE Communications Magazine, January 2003. pg. 38-44.
- [MF2005] Michaut, F.; Lepage, F. **Application-oriented Network Metrology: Metrics and Active Measurements Tools**. IEEE Communications Surveys and Tutorials, 2nd Quarter 2005, Vol. 7, no. 2.
- [NE2001] Noel, E.; Tang, K. W. **Performance Analysis of a VoIP Access Architecture**. AT&T Laboratories, Middletown, NJ, USA, 2000.
- [OPNET] **OPNET Modeller 10.0 Manuals**, <http://www.opnet.com>

- [PK1996] Park, K.; Kim, G.; Crovella, M. **On the Relationship Between File Sizes, Transport Protocols, and Self-similar Network Traffic**. Proceedings of International Conference on Network Protocols, Columbus, OH, USA, pp. 171-180, October 1996.
- [PK2000] Park, K.; Willinger, W.. **Self-Similar Network Traffic and Performance Evaluation**. John Wiley & Sons, New York, 2000. ISBN 0-471-31974-0
- [PV1995] Paxson, V.; Floyd, S. **Wide-area traffic: the failure of Poisson modeling**. IEEE/ACM Transactions on Networking, 3 (3): 226-244, June 1995.
- [PV1998] Paxson, V.; Almes, G.; Mahdavi, J.; Mathis, M. Framework for IP Performance Metrics. **IETF RFC 2330**, May 1998.
- [Rec. E.800] ITU-T. **Recommendation E.800**: Terms and Definitions Related to Quality of Service and Network Performance Including Dependability. ITU. August, 1993.
- [Rec. G.111] ITU-T. **Recommendation G.111**: Loudness Ratings (LRs) in an International Connection. ITU. March, 1993.
- [Rec. G.114] ITU-T. **Recommendation G.114**: One-way transmission time. ITU. May, 2003.
- [Rec. G.122] ITU-T. **Recommendation G.122**: Influence of National Systems on Stability and Talker Echo in International Connections. ITU. March, 1993.
- [Rec. G.131] ITU-T. **Recommendation G.131**: Talker Echo and Its Control. ITU. November, 2003.
- [Rec. G.711] ITU-T. **Recommendation G.711**: Pulse Code Modulation (PCM) of Voice Frequency. ITU. November, 1988.
- [Rec. G.723] ITU-T. **Recommendation G.723.1**: Dual Rate Speech Coder for Multimedia Communications Transmitting at 5,3 Kb/s and 6,3 Kb/s. ITU. March, 1996.
- [Rec. G.729] ITU-T. **Recommendation G.729**: Coding of Speech at 8 Kb/s using Conjugate-Structure Algebraic-Code-Excited Linear Prediction. ITU. March, 1996.
- [Rec. G.1000] ITU-T. **Recommendation G.1000**: Communications Quality of Service – A framework and definitions. ITU. November, 2001.
- [Rec. G.1010] ITU-T. **Recommendation G.1010**: Quality of service and performance – End-user multimedia QoS categories. ITU. November, 2001.

- [Rec. I.350] ITU-T. **Recommendation I.350**: General Aspects of Quality of Service and Network Desempenho in Digital Networks, Including ISDN. ITU. March, 1993.
- [Rec. Y.1241] ITU-T. **Recommendation Y.1241**: Support of IP-based Services Using IP Transfer Capabilities. ITU, March 2001.
- [Rec. Y.1541] ITU-T. **Recommendation Y.1541**: Network performance objectives for IP-based services. ITU. May, 2002.
- [RT1996] Rappaport, T. S.. **Wireless Communications: Principles & Practice**. Prentice Hall, New Jersey 1999 (2nd Edition). ISBN 0-13-375536-3.
- [SJ2003] Serger, J. **Modeling Approach for VoIP Traffic Aggregations for Transferring Tele-traffic Trunks in a QoS enabled IP-Backbone Environment**. International Workshop on Inter-domain Desempenho and Simulation, Salzburg, Austria, February 2003.
- [SK1986] Sriram, K.; Whitt, W.. **Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data**. IEEE Journal on Selected Areas in Communications. Vol. SAC-4, no. 6, September 1986, pp 833-846
- [SM1977] Spiegel, M. R. **Probabilidade e Estatística**. McGraw Hill do Brasil. Coleção Schaun. Rio de Janeiro, 1977.
- [SN2003] Seitz, N. **ITU-T QoS Standards for IP-based Networks**. IEEE Communications Magazine. June 2003, pg. 82-89.
- [TA2000] Tyagi, A.; Muppala, J. K.; de Meer, H.; Street, G. **VoIP support on Differentiated Services using Expedited Forwarding**. Proceedings of the IEEE International Conference on Performance, Computing and Communications, 2000 – IPCC'00, pp. 574-580, 20-22 February, 2000.
- [TM1997] Taqq, M. S.; Willinger, W.; Sherman, R. **Proof of a fundamental result in self-similar traffic modeling**. ACM/SIGCOMM Computer Communications Review, 27 (2): 5-23, 1997.
- [WW1997] Willinger, W.; Taqq, M. S.; Shermam, R.; Wilson, D. V.. **Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level**. IEEE/ACM Transactions on Networking, Vol. 5, no. 1, February 1997.
- [WZ2001] Wang, Zheng. **Internet QoS – Architectures and Mechanisms for Quality of Service**. Morgan Kaufmann Publishers. San Francisco, 2001. 239 p. ISBN 1-55860-608-4.

