UNIVERSIDADE FEDERAL DE UBERLÂNDIA

FACULDADE DE ENGENHARIA ELÉTRICA

PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

UMA PROPOSTA PARA A PROVISÃO DE QoS EM REDES *BACKBONE* UMTS

Paulo Dias de Alecrim

ABRIL

2006

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

FACULDADE DE ENGENHARIA ELÉTRICA

PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

UMA PROPOSTA PARA A PROVISÃO DE QoS EM REDES *BACKBONE* UMTS

Paulo Dias de Alecrim

Dissertação apresentada por Paulo Dias de Alecrim à Universidade Federal de Uberlândia para obtenção do título de Mestre em Ciências aprovada em 26 / 04 / 2006 pela Banca Examinadora:

Banca examinadora:

Professor: Paulo Roberto Guardieiro, Dr. (UFU) – Orientador

Professor: Solange da Silva, Dra. (UCG)

Professor: Jamil Salem Barbar, Dr. (UFU)

UMA PROPOSTA PARA A PROVISÃO DE QoS EM REDES *BACKBONE* UMTS

Paulo Dias de Alecrim

Dissertação apresentada à Universidade Federal de Uberlândia como parte dos requisitos para obtenção do título de Mestre em Ciências.

---	---
Prof. Paulo Roberto Guardieiro, Dr.	Prof. Darizon Alves de Andrade, Dr.
Orientador	Coordenador do Curso de Pós-Graduação

Dedicatória

A DEUS,

Aos meus queridos pais, José Alecrim e Francisca,

À minha querida esposa Ivane e a meus filhos, pelo apoio incessante,

pela compreensão, paciência e amor,

demonstrados nos momentos de dificuldade.

"Para ser sábio, é preciso primeiro temer a Deus, o SENHOR"

(Prov. 1:7-BLH)

Agradecimentos

Primeiramente, agradeço a Deus pelas incontáveis bênçãos recebidas ao longo de minha vida.

Ao meu orientador Prof. Dr. Paulo Roberto Guardieiro pela orientação competente e por todo apoio recebido.

À Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia pelos recursos oferecidos para a elaboração deste trabalho.

À minha eterna gratidão aos meus pais José Alecrim e Francisca Alecrim que abriram mão de muitos de seus recursos para proporcionar a mim os meus estudos.

À minha querida esposa Ivane Fonseca Alecrim, que sempre acreditou no meu potencial, agradeço por todo o amor, carinho e compreensão.

À minha filha Evelyn Alecrim e ao meu filho Ericssen Alecrim pela confiança e amor a mim dedicado.

Aos colegas e amigos do Laboratório de Redes de Computadores, pelo companheirismo, paciência e amizade.

A todos os meus amigos que, de alguma forma, contribuíram nesta caminhada, meus sinceros agradecimentos.

Resumo

Com o crescimento da demanda das aplicações multimídia na *Internet* e com os atuais *backbones* de alta capacidade, espera-se que as redes *backbone* UMTS suportem diferentes aplicações com requisitos de QoS, principalmente aquelas sensíveis a retardos, perdas e *jitter*. Entretanto, as aplicações de tempo real exigem mecanismos apropriados para garantir QoS de acordo com suas características específicas. Neste trabalho propõe-se a utilização de um mecanismo baseada no modelo *DiffServ,* que classificará e tratará cada pacote de acordo com as necessidades estabelecidas previamente por suas aplicações. Tal proposta provê um tratamento diferenciado ao tráfego de diversos fluxos na rede. Os resultados obtidos a partir de modelagem e simulação demonstraram que a utilização do mecanismo proposto trata as classes de tráfego UMTS de forma diferenciada, oferecendo um nível de QoS satisfatório mesmo em condições de mobilidade dos usuários.

Palavras-chave: Qualidade de Serviço (QoS), Serviços Diferenciados (*DiffServ*), *backbone* UMTS.

Abstract

With the growth of the demand of multimedia applications in the Internet and with the current backbones, it is expected that UMTS backbones support different applications with QoS requirements, mainly those which are sensitive to delays, packet losses and jitter. However, real time applications require proper mechanisms to guarantee QoS according to their specific characteristics. It is proposed in this work, the use of a mechanism based on the DiffServ model, which will classify and treat each packet in accordance to the needs which were previously established by its applications. Such proposal provides a differentiated treatment to the traffic of several flows in the network. The results reached from simulation and modeling showed that the use of the proposed mechanism treats the UMTS traffic classes in a differentiated way, offering a suitable level of QoS even under mobility conditions of the users.

Keywords: Quality of Service (QoS), Differentiated Services (DiffServ), UMTS backbone.

Sumário

1. Introdução..... 09

2. Redes Móveis sem fio: Visão Geral..... 15

2.1 Introdução..... 15

2.2 Componentes Básicos de uma Rede Móvel sem Fio..... 16

2.3 Evolução dos Sistemas Móveis sem Fio..... 18

2.3.1 Sistemas Analógicos de Primeira Geração (1G)..... 18

2.3.2 Sistemas Digitais de Segunda Geração (2G)..... 19

2.3.3 Evolução das Redes 2,5G em Direção aos Sistemas 3G..... 23

2.3.4 Sistema Universal de Telecomunicações Móveis - UMTS..... 25

2.3.5 Evolução das Redes 3G em Direção aos Sistemas 4G..... 31

2.4 Conclusão..... 34

3. Tecnologia para Comutação de Pacotes nas Redes Móveis sem Fio..... 36

3.1 Introdução..... 36

3.2 Arquitetura de Rede GPRS..... 37

3.3 Arquitetura EDGE..... 39

3.4 Arquitetura de Rede UMTS..... 42

3.4.1 A Pilha de Protocolos UMTS..... 46

3.4.2 A Interface Aérea do UMTS..... 47

3.4.3 QoS em Redes UMTS..... 50

3.4.4 Ativação do Contexto PDP..... 52

3.4.5 Mapeamento das Classes UMTS QoS..... 56

3.4.6 Parâmetros da Arquitetura de QoS para o UMTS..... 57

3.5 Conclusão..... 58

4. Qualidade de Serviço em Redes Móveis sem Fio..... 60

4.1 Introdução..... 60

4.2 Parâmetros que Influenciam na QoS..... 61

4.3. Mecanismo para Provimento de QoS na *Internet*..... 64

4.3.1. Arquitetura de Serviços Integrados *IntServ*..... 65

4.3.1.1. Classes de Serviço..... 66

4.3.2. Serviços Diferenciados *DiffServ*..... 67

4.3.2.1. Classes de Serviço..... 69

4.3.3. O Protocolo RSVP..... 72

4.3.4. Service Level Agreements SLA..... 73

4.3.5. Comutação Multi Protocolo (MPLS) 75

4.3.6. Engenharia de Tráfego..... 77

4.4. QoS em Redes Móveis..... 79

4.4.1. Controle de Admissão de Chamadas (CAC) 80

4.4.2. Alocação de Recursos de Rádio..... 82

4.4.3. Policiamento do Tráfego..... 82

4.4.4. Escalonamento..... 83

4.5. Disciplinas de Escalonamento de Filas..... 84

4.5.1. FIFO..... 85

4.5.2. PQ..... 86

4.5.3. FQ..... 87

4.5.4. WFQ..... 88

4.5.5. WRR..... 89

4.5.6. DWRR..... 91

4.6 Conclusão..... 92

5. Proposta para a Provisão de QoS em Redes *Backbone* UMTS..... 94

5.1 Introdução..... 94

5.2 QoS no Ambiente *Backbone* UMTS..... 95

5.3 Definição do Problema..... 97

5.4 Solução Proposta..... 98

5.5 Conclusão..... 101

6. Avaliação da Proposta para a Provisão de QoS em Redes *Backbone* UMTS..... 102

6.1 Introdução..... 102

6.2 Modelo de Simulação para o Ambiente *Backbone*..... 103

6.2.1 O Simulador para o Ambiente de Rede *Backbone*..... 104

6.3 Modelo de Simulação da Rede Móvel UMTS..... 105

6.4 Modelo de Simulação da Rede *Backbone* UMTS..... 108

6.4.1 Parâmetros para Simulação da Rede *Backbone* UMTS..... 109

6.4.2 Cenário das Simulações..... 111

6.5 Apresentação e Análise dos Resultados..... 112

6.5.1 Análise Comparativa entre Disciplinas de Escalonamento 113

6.6 Conclusão..... 124

7. Conclusões Gerais..... 127

Referências Bibliográficas..... 134

Lista de Figuras

2.1 Arquitetura de uma Rede Móvel sem Fio..... 17

3.1 Arquitetura de uma rede GPRS..... 37

3.2 Elementos básicos da arquitetura UMTS..... 43

3.3 Arquitetura UMTS..... 44

3.4 Arquitetura simplificada da pilha de protocolos UMTS..... 47

3.5 Estrutura de quadros no modo FDD do UMTS..... 48

3.6 Estrutura de quadro para o modo TDD do UMTS..... 49

3.7 Arquitetura de QoS no padrão UMTS..... 50

3.8 Ativação do contexto PDP pela MS..... 53

3.9 Ativação do contexto PDP pela rede..... 54

3.10 Desativação do contexto PDP pela MS..... 55

3.11 Desativação do contexto PDP pela rede..... 55

3.12 Componentes da arquitetura de QoS para UMTS..... 58

4.1 Layout do Campo DS..... 68

4.2 Classes do Serviço de Encaminhamento Assegurado..... 71

4.3 Funcionamento do Protocolo RSVP..... 73

4.4 Modelo lógico *DiffServ*..... 74

4.5 Encaminhamento de pacote em um domínio MPLS..... 76

4.6 Layout do cabeçalho do protocolo MPLS..... 76

4.7 Sistema de filas..... 84

4.8 FIFO..... 85

4.9 PQ..... 87

4.10 FQ..... 88

4.11 WFQ..... 89

4.12 WRR..... 90

4.13 DWRR..... 92

6.1 Modelo de Simulação da rede móvel UMTS 107

6.2 Modelo de Simulação da Rede *Backbone* UMTS..... 109

6.3 Atraso por disciplina de escalonamento..... 114

6.4 Atraso fim a fim para os fluxos F1 e F2..... 117

6.5 Atraso fim a fim para os fluxos F3 e F4..... 118

6.6 Banda alocada pelas classes F1, F2, F3 e F4 versus % carga de tráfego..... 119

6.7 Perda de Pacotes versus % carga de tráfego (F1, F2)..... 120

6.8 Perda de Pacotes versus % carga de tráfego (F3, F4)..... 121

6.9 Perda de Pacotes versus % carga de tráfego (F1, F2, F3, F4)..... 122

6.10 Variação de atraso (*Jitter*) versus % carga de tráfego (F1, F2)..... 124

Lista de Tabelas

2.1 Sistemas Analógicos de Primeira Geração (1G)..... 18

2.2 Características básicas da evolução dos sistemas móveis sem fio..... 25

2.3 Progresso de Implementação da UMTS..... 27

3.1 Mapeamento das Classes UMTS QoS em Classes DiffServ de PHB..... 56

3.2 Atributos de QoS para as Classes UMTS..... 57

4.1 Parâmetros de QoS definidos para um conjunto de aplicações..... 62

4.2 Vazão para diversas aplicações..... 63

4.3 Precedência IP..... 68

5.1 Exemplos de Aplicações para as Classes UMTS e suas Restrições Quanto à QoS 97

6.1 Parâmetros definidos para a simulação..... 110

6.2 Classes de precedência e atraso..... 113

6.3 Atraso em ms para os fluxos F1 e F2 com o aumento do tráfego..... 117

6.4 Alocação de Banda para os fluxos F1 e F2 com o aumento do tráfego..... 119

6.5 *Jitter* Médio em ms para os fluxos F1 e F2 com o aumento do tráfego..... 124

Lista de Abreviaturas

1G	*First Generation;*
2G	*Second Generation;*
2,5G	*Second and half Generation;*
3G	*Third Generation;*
3GPP	*Third Generation Partnership Project;*
AF	*Assured Forwarding;*
AMPS	*Advanced Mobile Phone System;*
ANATEL	Agência Nacional de Telecomunicações
APN	*Access Point Name;*
ATM	*Asynchronous Transfer Mode*;
AUC	*Authentication Center*;
BCS	*Block Check Sequence*;
BER	*Bit Error Rate*;
BS	*Base Station*;
BSC	*Base Station Controller*;
BSS	*Base Station SubSystem*;
BTS	*Base Transceiver Station*;

CAC	*Capacity-based Admission Control*;
CBI	*Constraint-Based Routing;*
CBQ	*Class-based queueing;*
CDMA	*Code Division Multiple Access*;
CDR	*Call Detail Record*;
CL	*Controlled Load Service;*
CN	*Core Network*;
CP	*Conditioning Paket;*
CQ	*Custom Queueing;*
CT	*Conditioning Traffic* ;
D-AMPS	*Digital* AMPS;
DCS-1800	*Digital Cellular System, 1,800 MHz;*
DiffServ	*Differentiated Services*;
DLCI	*Data Link Connection Identifier;*
DNS	*Domain Name Server;*
DS	*Direct Sequence*;
DSCP	*Differentiated Services Code Points;*
DSI	*Digital Speech Interpolation;*
DSL	*Digital Speech Link;*
DWDM	*Dense Wavelength Division Multiplexing;*
DWRR	*Deficit Weighted Round Robin;*
E-AMPS	*Extended* AMPS;
EDGE	*Enhanced Data Rates for Global Evolution;*
EF	*Expedited Forwarding;*

EIR	*Equipment Identity Register;*
ESN	*Eletronic Serial Number;*
ETSI	*European Telecommunications Standard Institute;*
FCFS	*First Come First Served;*
FDD	*Frequency Division Duplexing;*
FDMA	*Frequency Division Multiple Access;*
FH	*Frequency Hopping;*
FIFO	*First-In-First-Out;*
FM	*Frequency Modulation;*
FPLMTS	*Future Public Land Mobile Telecommunication System;*
FQ	*Fair Queuing;*
FSK	*Frequency Shift Keying;*
FTP	*File Transfer Protocol;*
FWA	*Fixed Wireless Access;*
GERAN	*GSM / EDGE Radio Access Network;*
GGSN	*Gateway GPRS Support Node;*
GMSC	*Gateway MSC;*
GMSK	*Gausssian Minimum Shift Keying;*
GP	*Guard Period;*
GPRS	*General Packet Radio System;*
GPS	*Generalized Processor Sharing;*
GR	*Resource Manager;*
GS	*Guaranteed Service;*
GSM	*Global System for Mobile Communication*s;

GSMK	*Gausssian Minimum Shift Keying*;
GTP	*GPRS Tunnel Protocol;*
HLR	*Home Location Register;*
HTTP	*Hiper Text Transfer Protocol;*
IETF	*Internet Engineering Task Force;*
IMSI	*International Mobile Subscriber Identity;*
IMT-2000	*International Mobile Telecommunications 2000;*
IntServ	*Integrated Services;*
IP	*Internet Protocol;*
IS-54	*Interim Standart - 54;*
IS-95	*Interim Standart - 95;*
IS-136	*Interim Standart - 136;*
IT	*Information Technology;*
ITU	*International Telecommunications Union;*
ITU-T	*International Telecommunications Union – Telecommunications;*
LL	*Layer Link;*
LSP	*Label-Switched Path;*
LSR	*Label-Switched Router;*
MAC	*Medium Access Control;*
MIN	*Mobile Identification Number;*
MMS	*Multimedia Messaging Service;*
MPLS	*MultiProtocol Label Switching;*
MS	*Mobile Stations;*
MSC	*Mobile Switching Center;*

NAM	*Network Animator;*
NMT	*Nordic Mobile Telephones;*
NOAH	*Non-Ad-Hoc routing agent;*
NS-2	*Network Simulator version 2;*
NTT	*Nippon Telephone and Telegraph;*
OTcl	*Object Tool command language;*
PCS-1900	*Personal Communication System, 1.900 MHz;*
PCU	*Packet Control Unit;*
PCUSN	PCU *Support Node;*
PDA	*Personal Digital Assistant;*
PDC	*Personal Digital Cellular;*
PDN	*Packet Data Networks;*
PDP	*Packet Data Protocol;*
PDU	*Packet Data Unit;*
PHB-EF	*Per Hop Behavior-EF;*
PLMN	*Public Land Mobile Network;*
PQ	*Priority Queueing;*
PSK	*Phase Shift Queueing;*
PSTN	*Public Switched Telephony Network;*
QoS	*Quality of Service;*
RA	*Routing Area;*
RAB	*Radio Access Bearer;*
RADIUS	*Remote Authentication Dial In User Service;*
RAN	*Radio Access Network;*

RDSI	Rede Digital de Serviço Integrado;
RLC	*Radio Link Control;*
RNC	*Radio Network Controller;*
RNS	*Radio Network Systems;*
RSVP	*Resource Reservation Protocol;*
RTPC	Rede Telefônica Pública Comutada;
RTT	*Radio Transmission Technologies;*
SAP	*Service Access Point;*
SDU	*Service Data Unit;*
SGSN	*Serving GPRS Support Node;*
SIM	*Subscriber Identification Module;*
SLA	*Service Level Agreement;*
SMS	*Short Message Service;*
SMTP	*Simple Mail Transfer Protocol;*
SNDCP	*Sub Network Dependent Convergence Protocol;*
TACS	*Total Access Communications System;*
TBF	*Temporary Block Flow;*
TCP	*Transmition Control Protocol;*
TDD	*Time Division Duplexing;*
TDMA	*Time Division Multiple Access;*
TID	*Tunnel Identifier;*
TLLI	*Temporary Logical Link Identifier;*
ToS	*Type-of-Service;*
TTL	*Time-to-Live;*

UDP	*User Datagram Protocol;*
UE	*User Equipment;*
UMTS	*Universal Mobile Telephone System;*
UTRAN	*UMTS Terrestrial Radio Access Network;*
VLR	*Visitor Location Register;*
VPN	*Virtual Private Network;*
VSELP	*Vector-Sum Excited Linear Prediction;*
WCDMA	*Wideband CDMA;*
WFQ	*Weighted Fair Queueing;*
WLAN	*Wireless LAN;*
WLL	*Wireless Local Loop;*
WRR	*Weighted Round Robin;*
WWW	*World Wide Web*

Capítulo 1

Introdução

O crescimento das redes sem fio com a integração de estações móveis (*MS-Mobile Station*) com alta capacidade de processamento tem despertado a necessidade de interligá-las à *Internet* e utilizá-las da mesma forma que os computadores convencionais, quebrando aquele conceito de que, para estar conectado deve-se estar à frente de um *desktop* em uma sala fechada. Para que estes dispositivos possam ter acesso às aplicações de tempo real, é necessário a implementação de mecanismos que ofereçam suporte de Qualidade de Serviço (*QoS-Quality of Service*) num ambiente com dispositivos em mobilidade [23], [55].

As estatísticas mostram que o número de redes móveis sem fio irá suplantar o de acessos fixos convencionais por volta do ano 2010 [77]. É surpreendente perceber que em apenas 20 anos os acessos sem fio terão alcançado a penetração de mercado atingida em cerca de 100 anos pelos serviços de acesso fixo convencional [54].

Relatórios da TELEBRÁS S/A hoje ANATEL (Agência Nacional de Telecomunicações) [78] informam que na década de 70 havia somente dois milhões de acessos fixos convencionais, mas com o advento das fibras ópticas houve o crescimento da digitalização dos sistemas telefônicos no Brasil. Assim, estes números foram acrescidos em 10 milhões de acessos fixos na década de 80, sendo que ao final da década de 90 já estavam em 22 milhões

de acessos fixos e 9 milhões de acessos móveis sem fio. Hoje, estes números são bem maiores, pois se têm na planta em funcionamento mais de 90 milhões de acessos móveis sem fio ao passo que os sistemas de acesso fixo decrescem a cada dia.

Anterior à década de 90 predominava a necessidade pelo chamado acesso básico (como voz e Fax). Após essa década os usuários passaram a requerer, nesse novo cenário, não somente dos serviços de voz e Fax, mas também *e-mail*, áudio em tempo real, TV móvel, acessíveis com qualidade a qualquer momento e sem depender de posições geográficas.

Desde sua criação, a *Internet* têm oferecido apenas o serviço de melhor esforço (*Best Effort*), caracterizado pela ausência de garantias de QoS às aplicações. A inadequação do serviço de melhor esforço aos requisitos das aplicações supracitadas motivou nos últimos anos o desenvolvimento de várias propostas para um novo modelo de serviços que agregasse o conceito de QoS na *Internet*, permitindo às aplicações requisitar serviços com diferentes características de qualidade [33], [45].

A possibilidade de acessar informações, aplicativos e serviços a qualquer momento e em qualquer lugar, está adequando uma nova sociedade da informação que vislumbra formas diferenciadas de acesso à *Internet* através de estações móveis, tais como PDAs (*Personal Digital Assistant*), *Notebooks*, *Laptops, Palms,* entre outras. [4]. Entretanto, a *Internet* não foi concebida para lidar com as características inerentes ao ambiente, onde vários usuários movimentam-se conectados à rede através de uma interface aérea com baixa largura de banda.

A infra-estrutura dos sistemas móveis sem fio, com estações móveis, está evoluindo com o objetivo de dar suporte às aplicações de usuários com soluções de QoS fim a fim baseadas no protocolo IP, capazes de prover qualidade de serviço aos seus usuários. Maior capacidade está sendo proporcionada às redes móveis sem fio, com mais largura de banda no enlace de rádio e a utilização de comutação por pacotes desde a interface aérea [69].

O mercado de telecomunicações apresenta atualmente duas principais demandas: usuários com uma crescente necessidade de acesso à informação com QoS; e usuários que necessitam acessar as informações a qualquer momento sem depender de posições geográficas. Atualmente, duas tecnologias têm atendido a estes dois requisitos através de diferentes alternativas, conhecidas como *Internet* e redes móveis sem fio, respectivamente.

A convergência de ambas as soluções é um dos principais objetivos dos sistemas móveis de terceira e quarta geração (3G/4G) denominada pelo ITU (*International Telecommunications Union*) de UMTS (*Universal Mobile Telephone System*) [56]. Estes sistemas permitem que as duas demandas sejam atendidas pela mesma plataforma e através das novas gerações de sistemas de acesso [3].

Sistemas UMTS abrem as portas para a comunicação e serviços de multimídia em banda larga com velocidade superior a 100 Mbps com estações móveis [7], [56]. Tais perspectivas prometem conduzir a conectividade e a capacidade de dados em um ambiente móvel em níveis não disponíveis atualmente. A comunicação móvel sem fio permitindo a troca de informações a altas taxas e com alta qualidade de serviço entre estações móveis que podem estar localizadas em qualquer parte do mundo, representa a fronteira a ser suplantada pelos sistemas UMTS [54], [77].

Estes fatores sem dúvida sinalizam que as soluções para implementação de QoS são imprescindíveis para o sucesso destes sistemas. São importantes tanto para que os usuários possam usufruir destas novas aplicações, quanto para que as operadoras e provedores de serviço possam ter controle sobre suas redes e conteúdos.

Objetivando alcançar um nível de qualidade de serviço satisfatório, as redes móveis sem fio deverão ser capazes de atender a requisitos de QoS de aplicações de tempo real [36]. O crescente interesse em torno de aplicações com estes perfis de QoS em redes móveis sem fio traz consigo desafios próprios que tornam a provisão de uma tarefa mais complexa. Tal

complexidade é resultado, dentre outros fatores, das limitações do ambiente e dos dispositivos envolvidos, como por exemplo: a qualidade da transmissão na interface aérea UTRAN (*UMTS Terrestrial Radio Access Network*) que é afetada pelo desvanecimento [55], [65]; os recursos escassos de largura de banda; a necessidade do controle de admissão dos pacotes nos pontos de ingresso ao *backbone* UMTS e a própria mobilidade das MSs, que implicará na necessidade do gerenciamento do congestionamento no núcleo deste *backbone*. Estes entre outros, são alguns dos problemas que dificultam o provimento de QoS neste tipo de rede.

Antes que uma MS possa efetivar a comunicação com outras PDNs (*Packet Data Networks*), o contexto PDP(*Packet Data Protocol*) deve ser ativado e nesta fase o perfil de QoS é negociado com o *backbone*. O *backbone* UMTS permite que sejam definidos perfis de tráfego para o qual o contexto pode se adequar tais como: (conversacional, *streaming*, interativa ou *background*), entre outros atributos que podem ser negociados [56]. Contudo, nenhum mecanismo de QoS é definido nesta fase que possa dar suporte a esses perfis e garantir o seu funcionamento.

Caminhando neste sentido, com o objetivo de prover QoS em redes *backbones* com tecnologia UMTS, propõe-se nesta dissertação uma proposta para a provisão QoS. Tal proposta emprega a arquitetura de serviços diferenciados juntamente com o algoritmo de escalonamento por prioridade PQ (*Priority Queueing*). O modelo de serviços diferenciados é denominado pelo IETF (*Internet Engineering Task Forc*e) de *DiffServ* e usa a classificação de pacotes como mecanismo para obtenção de QoS, sendo uma extensão dos trabalhos já consolidados [24], [57] e [63]. O mecanismo proposto é aplicada nos SGSN (*Serving GPRS Support Node*) e GGSN (*Gateway GPRS Support Node*) do *backbone* UMTS. Para tanto, foram avaliados parâmetros de QoS como atraso, perda de pacotes, *jitter* e alocação de largura de banda para as classes de tráfego UMTS [55].

A arquitetura *DiffServ* tem emergido como uma tecnologia de suporte a QoS em redes IP. Por esta razão, é redefinido o *layout* do *octeto* ToS (*Type of Service*) dos cabeçalhos dos protocolos IPv4 e do IPv6, que passa a ser chamado de campo DS (*Differentiated Services*) [5], [39], [57]. Assim, o tráfego que entra no *backbone* UMTS poderá ser tratado como as classes de tráfego UMTS, possuindo cada uma delas, diferentes prioridades por exigirem diferentes requisitos de Qualidade de Serviço. Desta forma, esta arquitetura pode ser usada como alternativa para fornecer QoS em redes *backbone* UMTS [52], [57], [58].

Mas o sucesso da implementação do modelo *DiffServ* depende também da escolha adequada do algoritmo de escalonamento. Portanto, um estudo comparativo será feito entre as seis disciplinas de escalonamento proposto pelo IETF com intuito de verificar quais delas atendem aos requisitos de atraso e *jitter* para aplicações sensíveis a estes atributos. O algoritmo PQ tem sido apontado como uma boa alternativa para o escalonamento de pacotes, capaz de minimizar o congestionamento no núcleo dos *backbones* UMTS através de um maior aproveitamento de seus recursos [38].

Espera-se com a implementação do mecanismo proposto, tornar viável o emprego das classes de tráfego UMTS propostas pelo grupo 3GPP (*Third Generation Partnership Project*) em *backbones* UMTS, concedendo um nível de QoS satisfatório a cada uma destas classes de forma diferenciada.

Esta dissertação está organizada da seguinte forma:

O Capítulo 2 apresenta uma visão geral dos sistemas móveis sem fio, aspectos de sua padronização, e os cenários de sua evolução considerando as redes 1G, 2G/2,5G e a evolução das redes 3G em direção aos sistemas 4G / UMTS.

O Capítulo 3 apresenta os principais conceitos relativos à tecnologia para comutação de pacotes nas redes móveis sem fio. Aspectos técnicos da arquitetura GPRS (*General Packet*

Radio System), conceitos básicos do contexto PDP e a arquitetura EDGE (*Enhanced Data Rates for Global Evolution*).

O Capítulo 4 descreve as arquiteturas que suportam qualidade de serviço em redes móveis sem fio, incluindo as arquiteturas IETF que têm sido propostas para a implementação de qualidade de serviço nas redes IP. Descreve também os principais algoritmos de escalonamento propostos pelo IETF para a *Internet*.

O Capítulo 5 apresenta a proposta para a provisão de QoS em redes *backbone* UMTS, definindo inicialmente o problema estudado e a solução proposta baseada na utilização do modelo *DiffServ* e a política de escalonamento PQ proposto pelo IETF para o provisionamento de QoS para estes sistemas.

O Capítulo 6 apresenta os procedimentos detalhados para a avaliação da proposta apresentada. Descreve-se os ambientes de simulação utilizados e os resultados obtidos utilizando-se o mecanismo de *DiffServ* e o algoritmo PQ em redes *backbone* UMTS. São analisados os itens tais como, atraso fim a fim, largura de banda alocada, perda de pacotes e *jitter*. São apresentadas as conclusões para cada resultado obtido na simulação.

Finalmente no Capítulo 7 são apresentadas as conclusões finais correspondentes aos pontos principais desta dissertação e sugestões para trabalhos futuros.

Capítulo 2

Redes Móveis sem Fio: Visão Geral

2.1 Introdução

A grande popularidade e disseminação que as redes sem fio vem atingindo nos últimos anos têm despertado o interesse de diversos seguimentos ligados à pesquisa e desenvolvimento, especialmente no que diz respeito à adequação das redes sem fio à *Internet*. Um sistema de comunicação móvel sem fio tem como característica a possibilidade de mobilidade relativa entre as partes como, por exemplo, a comunicação entre um usuário a partir de uma estação móvel e a estação rádio-base no sistema móvel. Redes móveis usam a tecnologia sem fio para possibilitar uma comunicação transparente enquanto a estação móvel se desloca [5], [6].

Nos anos 80 as estações móveis eram instaladas em veículos o que permitia apenas uma comunicação veicular, e não pessoal. A era da comunicação pessoal teve seu início efetivo no início dos anos 90, quando o usuário podia portar o aparelho embora suas dimensões iniciais fossem grandes.

Nestes últimos anos duas tecnologias têm crescido de uma maneira que excederam todas as expectativas: comunicações móveis sem fio e a *Internet*. O serviço móvel sem fio obteve um

crescimento muito grande de usuários juntamente com a redução drástica dos planos para os serviços de voz. Por outro lado, a Web proporciona uma grande possibilidade de acesso a informações para os usuários fixos. A possibilidade de acesso à informação e serviços a qualquer hora e em qualquer lugar do planeta poderá moldar uma nova sociedade de informação com os usuários acessando a informação na *Internet* através de diversos tipos de estações móveis, tais como telefones celulares, PDAs, *Laptops*, entre outros [5], [6].

Consideram-se, a seguir, os componentes básicos de uma rede móvel sem fio, bem como a evolução destes sistemas a partir dos padrões analógicos com tecnologia AMPS (*Advanced Mobile Phone System*). Os padrões 2G e suas características, e seus limitantes para uma tecnologia posterior 2,5G e a evolução da tecnologia 3G para 4G incluindo o padrão UMTS, são também apresentados.

2.2 Componentes Básicos de uma Rede Móvel sem Fio

Uma rede móvel sem fio é composta pelos seguintes elementos básicos, estações móveis (*MS - Mobile Station*), estações base (*BS - Base Station*), centro de comutação móvel (*MSC – Mobile Switching Center*), HLR (*Home Location Register*) e o VLR (*Visitor Location Register*) [12] [13].

Estações móveis comunicam-se com a estação base por meio de sinais de rádio. A estação base, que cobre uma determinada área denominada de célula, converte estes sinais e os transfere para um MSC via interface aérea ou mesmo via enlaces físicos. O MSC coordena e roteia as chamadas para outras estações móveis ou fixas conectando-as à rede pública de telefonia (*PSTN – Public Switched Telephone Network*). O HLR é uma base de dados de uma área de serviço que contém informações sobre as MSs cadastradas naquela área. Nos HLRs são armazenadas informações como, MIN (*Mobile Identification Number*), ESN (*Electronic Serial Number*) e perfil do usuário com dados sobre *billing*, *call forward-to number*, estado

(ativo ou inativo) e ponteiro para o último VLR onde a MS se registrou. O VLR funciona em conjunto com o HLR para suportar mobilidade das MS, ou seja, *roaming* automático. É um repositório local, temporário, com dados da MS que está fora de sua área de serviço. A Figura 2.1 esquematiza uma rede de comunicação móvel sem fio, com sua interligação à PSTN, HLR e VLR.

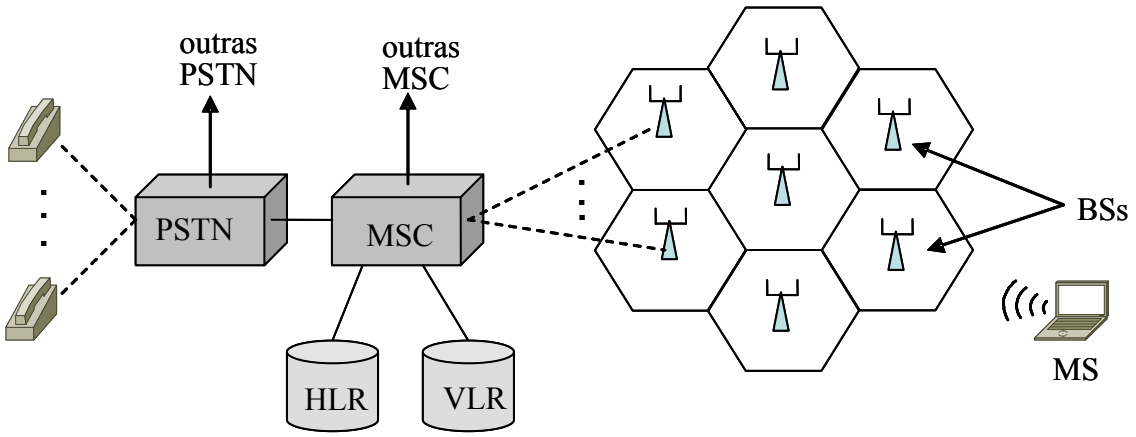

Figura 2.1: Arquitetura de uma Rede Móvel sem Fio.

O conceito introduzido pelos sistemas móveis sem fio permite que pequenas áreas de cobertura de rádio tornem-se parte de um grande sistema através da interconexão destas pequenas áreas por meio do MSC. Com o crescimento da demanda e o reduzido espectro de freqüências, as células têm uma tendência a reduzirem sua área de cobertura. Neste contexto, surgem os conceitos de macrocélula, microcélula e picocélula. As células reduzem, também, a potência de suas BSs. Um procedimento importante em sistemas móveis é denominado *handoff* e corresponde à mudança automática de chamada de uma célula para outra à medida que o usuário se desloca.

2.3 Evolução dos Sistemas Móveis sem Fio

Serão considerados a seguir, alguns detalhamentos básicos dos componentes de uma rede móvel sem fio, bem como a evolução destes sistemas a partir de padrões estabelecidos no âmbito do 3GPP.

2.3.1 Sistemas Analógicos de Primeira Geração (1G)

Os sistemas analógicos de primeira geração foram desenvolvidos para prover apenas serviços de voz. Dentre os principais sistemas desenvolvidos para prestar este tipo de serviço pode-se citar AMPS, NMT (*Nordic Mobile Telephones*), TACS (*Total Access Communications System*), C450 e NTT (*Nippon Telephone and Telegraph*). As características básicas destes sistemas são: modulação analógica em FM (*Frequency Modulation*), sinalização feita através da técnica de modulação FSK (*Frequency Shift Keying*), técnica de acesso FDMA (*Frequency Division Multiple Access*) com duplexação FDD (*Frequency Division Duplexing)*, tamanho de célula de 0,5 km a 10 km e potência de transmissão do terminal móvel de 1 a 8 Watts [5], [4]. A Tabela 2.1 resume as principais características dos sistemas analógicos de primeira geração utilizados em todo o mundo [5].

Tabela 2.1: Sistemas Analógicos de Primeira Geração (1G).

SISTEMAS	**JAPÃO**	**VÁRIOS PAÍSES**	**VÁRIOS PAÍSES**	**VÁRIOS PAÍSES**	**ALEMANHA**
	NTT	**AMPS**	**TACS**	**NMT**	**C450**
Freqüência de Transmissão(MHz) Estação Rádio Base (BS) Estação Móvel (MS)	 870-885 955-940	 869-894 824-849	 917-950 872-905	 463-467.5 453-457.5	 461.3-465.74 451.3-455.74
Número de Canais	600	832	1320	180	222
Raio de Cobertura (Km)	5 - 15	2 – 20	2 – 20	1.8 – 40	5 – 30
Taxa na Sinalização (Kbps)	0.3	10	8	1.2	5.28

O sistema celular AMPS, proposto pela AT&T em 1971, divide a banda de freqüências em canais de rádio de 30 kHz. Quando em operação, cada canal de rádio serve a um único usuário (FDMA). Existem dois tipos de canais no sistema AMPS: canais de controle dedicados e canais de voz. Canais de controle dedicados enviam mensagens de *paging* e coordenam o acesso ao sistema. Os canais de controle utilizam modulação FSK que enviam mensagens digitais a uma taxa de 10kbps. Depois que o canal de controle tenha coordenado o acesso, um canal de rádio com modulação FM é atribuído ao terminal móvel para que seja efetuada a transferência do sinal de voz.

2.3.2 Sistemas Digitais de Segunda Geração (2G)

Os sistemas celulares de Segunda Geração possuem como característica comum o fato de empregarem esquemas de modulação digital também nos canais de voz, e não apenas nos canais de controle como já era feito nos sistemas de Primeira Geração. A digitalização permite o uso das técnicas TDMA (*Time Division Multiple Access)* e CDMA (*Code Division Multiple Access*) como alternativas à técnica FDMA. Dentre os sistemas de comunicação móvel celular digital podem-se destacar os padrões: GSM (*Global System for Mobile Communication*s), DCS-1800, PCS-1900, IS-54 D-AMPS e CDMA IS-95 [5].

Sistema GSM

O GSM foi desenvolvido com o objetivo de solucionar o problema da fragmentação existente nos primeiros sistemas celulares na Europa e de proporcionar uma série de serviços através da utilização de uma Rede Digital de Serviços Integrados (RDSI). Sua introdução na Europa foi em 1991 e desde então vários países da América do Sul, Ásia e Austrália têm adotado o GSM e sistemas equivalentes, tal como o DCS-1800. Dentre os principais serviços suportados pelo padrão GSM podem-se destacar:

- Serviços de telefonia, incluindo fax, videotexto e telex.
- Serviços de dados com possibilidade de comunicação de dados por pacotes a taxas de até 9600bps.
- Serviços de RDSI suplementares tais como desvio de chamada, identificação de assinante chamador e serviço de mensagem. Este último, chamado SMS (*Short Message Service*), permite a recepção de mensagens alfanuméricas mesmo durante uma conversação.

Outra grande inovação do sistema GSM é o módulo de identificação do usuário, o SIM (*Subscriber Identification Module*) que contém a identificação completa do usuário, chaves de código de privacidade e outras informações específicas sobre o usuário. O SIM pode apresentar-se sob a forma de um cartão de crédito ou de um *plug-in* que é conectado ao terminal GSM. Sem este módulo SIM o terminal fica inoperante [5].

A privacidade é outra característica marcante no padrão GSM. Isto é possível através da criptografia (embaralhamento) da seqüência de *bits* do usuário através de uma "chave" (regra de embaralhamento) que se altera freqüentemente [54], [70].

Resumo das principais características técnicas do Sistema GSM:

- Acesso FDMA/TDMA com canais de 200KHz e 8 *slots* temporais por canal (8 usuários por canal).
- Codificação de voz LPC-RPE (l3Kbps).
- Equalização adaptativa de canal.
- Taxa de transmissão no canal: 270Kbps.
- Taxa de transmissão por usuário (incluindo *bits* de codificação de canal): 22.8Kbps.
- Modulação GMSK com BT = 0,3.
- Máximo tamanho das células (limitado pelo tempo de guarda): 30Km.

- Potência de transmissão: equipamento de mão máx. 2W; equipamento portátil máx. 8W.
- Sinalização fora da faixa útil.

Os padrões DCS-1800 e PCS-1900 são baseados na arquitetura de rede do sistema GSM, sendo que o PCS-1900 é o equivalente nos Estados Unidos do padrão GSM [5].

IS-54 D-AMPS

O padrão IS-54, conhecido como D-AMPS (*Digital AMPS*) é o equivalente digital do sistema analógico AMPS. No sistema D-AMPS cada canal de 30KHz é compartilhado por três usuários em *slots* de tempo distintos. A capacidade é triplicada em relação ao AMPS. A principal motivação para o desenvolvimento desse sistema foi a possibilidade de transição suave da tecnologia analógica para a digital, ocupando a mesma faixa de freqüências. Uma recente variação do padrão IS-54, o IS-136 (anteriormente IS-54), tem as mesmas características do IS-54, diferenciando no tipo de modulação utilizada nos canais de controle ($\pi/4$ DQPSK), de forma a elevar a taxa de transmissão nesses canais e prover serviços adicionais tais como *paging* e envio de mensagens curtas entre usuários em grupos privados.

Uma outra evolução do padrão D-AMPS e que se encontra em fase de normalização é o E-AMPS (*Extended* AMPS). Nesse padrão é utilizada alocação dinâmica de canais baseada na atividade da voz, utilizando a técnica DSI (*Digital Speech Interpolation*). Essa técnica pode elevar a capacidade em 15 vezes em relação ao sistema analógico AMPS [9], [54].

Resumo das principais características técnicas do padrão IS-54 (D-AMPS):

- Método de acesso FDMA/TDMA, com canais de 30KHz e 3 *slots* de tempo por canal (3 usuários por canal).
- Taxa de transmissão no canal: 48,6Kbps.
- Codificação de voz VSELP (13Kbps).

- Triplica a capacidade do sistema analógico AMPS.
- Modulação π/4 DQPSK.
- Eficiência de largura de banda: 1,62bps/Hz.
- Complexo processo de gerenciamento temporal.

IS-95 (CDMA)

O padrão IS-95, baseado em tecnologia com espalhamento espectral (acesso múltiplo CDMA), foi desenvolvido também com o objetivo de possibilitar a transição suave da tecnologia analógica para a digital. No padrão IS-95 todos usuários utilizam a mesma faixa de freqüências, mesmo aqueles usuários em células adjacentes, sendo que cada canal IS-95 ocupa 1,25MHz (cerca de 10% da banda disponível para o padrão AMPS). Na prática as portadoras do sistema AMPS devem prover uma faixa de guarda de 270KHz (9 canais) em cada lado da faixa dedicada ao IS-95.

Os primeiros terminais *dual-mode* (IS-95/AMPS) foram disponibilizados pela Qualcomm em 1994.

No padrão IS-95 a taxa de transmissão no canal é variável em função dos requerimentos do terminal e do ciclo de atividade da voz. São possíveis 4 taxas de transmissão: 9600bps, 4800bps, 2400bps e 1200bps, sendo que a transmissão acontece sempre em rajadas de 9600bps em cada quadro (*frame*). A cada vez que a taxa é reduzida em função da atividade da voz são gerados intervalos durante os quais a potência de transmissão do terminal móvel é reduzida no mínimo em 20dB em relação ao período prévio de atividade. Esse procedimento reduz significativamente a quantidade de interferência no sistema, proporcionando um correspondente aumento da capacidade em termos do número médio de usuários ativos.

Resumo das principais características técnicas do padrão IS-95 (CDMA):

- Taxa de *chip* de 1.2288Mchips/s = 128 vezes 9600bps.

- No enlace direto (base para terminal móvel): Utilização de 64 códigos *Walsh-Hadamard* combinados para ortogonalidade entre os usuários; utilização de uma seqüência PN do tipo *m* [12] para um espalhamento efetivo e robustez contra os efeitos dos multipercursos
- No enlace reverso (terminal móvel para base): Utilização de seqüências PN tipo *m* melhores propriedades para Sincronismo.
- Largura de faixa de transmissão: 1,25MHz.
- Codificação de canal convolucional com taxa 1/2 no enlace direto e taxa 1/3 no enlace reverso.
- Uso de portadora piloto para sincronismo e detecção coerente nos terminais móveis.
- *Soft handoff* (mesma portadora para todo o sistema).
- Degradação suave na qualidade do enlace reverso quando há sobrecarga de usuários no sistema (*soft capacity*).
- Controle de potência do terminal móvel atualizado a cada 1 ms (ou conforme variação do desvanecimento) imperfeito devido a desvanecimentos muito profundos e/ou muito rápidos [9], [54].
- Utilização do receptor RAKE [13], [14], [17] ao invés de equalizador de canal.

2.3.3 Evolução das Redes 2,5G em Direção aos Sistemas 3G

A migração GSM, de tecnologia de segunda geração para a terceira geração, incorpora a otimização constante da capacidade e eficiência. O suporte à transmissão de dados por parte de tecnologias de segunda geração limita-se a algumas aplicações básicas, tais como serviço de mensagem e *e-mail* baseado em texto, não provendo conectividade adequada para acesso eficiente à *Internet.*

Um novo passo da evolução do padrão 2G foi sua integração com transmissão de pacotes de dados, devido muito à forte demanda de serviços de acesso à *Internet* para ambientes de redes sem fio. Esta nova capacidade recebeu o nome de sistemas 2,5G. O maior incremento que o 2,5G trouxe foi uma técnica avançada de modulação (comparado ao 2G), permitindo a comutação de pacotes ao invés de circuitos, a mesma técnica de transmissão adotada pelo IP da arquitetura TCP/IP. Diferentemente da comutação por circuito que aloca um circuito fim a fim durante a transmissão, a comutação de pacotes só utiliza o caminho quando de fato há dados para transmitir. Assim, a tecnologia 2,5G trouxe um uso mais eficiente do espectro de freqüência e da banda disponível, promovendo o meio de transporte mais apropriado para a navegação de aplicações na *Internet* a partir de dispositivos móveis sem fio, notadamente com o surgimento de aparelhos celulares com esta capacidade. Com 2,5G, pode-se atingir a taxa de transmissão de dados de até 144 Kbps. Na maioria dos casos, os sistemas 2,5G são implementados diretamente sobre as redes 2G existentes. Como resultado, um sistema 2,5G não é uma rede comutada a pacotes. Na verdade, pacotes de dados são transmitidos sobre redes de circuitos comutados. Redes móveis sem fio realmente comutadas a pacotes só serão mesmo disponibilizadas com o advento da geração 3G [8].

A próxima geração de redes de comunicação sem fio incluirá o desenvolvimento dos sistemas móveis em direção aos sistemas de terceira geração, também denominados de 3G (IMT-2000 para o ITU e UMTS para a Europa)[14], baseados na tecnologia WCDMA [01], [14]. Tais perspectivas prometem conduzir a conectividade e a capacidade de dados em um ambiente móvel em níveis não suportados atualmente. A Tabela 2.2 sumariza as gerações de sistemas móveis sem fio e suas características básicas.

Tabela 2.2: Características básicas da evolução dos sistemas móveis sem fio.

ARQUITETURA	**1G**	**2G**	**2,5G**	**3G**
Tipo do Sinal no Acesso	Analógico	Digital	Digital	Digital
Comutação da Rede	Circuito	Circuito	Pacote	Pacote
Aplicações disponíveis	Voz	Mensagem	*Internet*	Multimídia
Taxa de Transmissão (Kbps)	Não disponível	14	144	384 – 2000
Exemplo de Sistema	AMPS	TDMA, CDMA, GSM	GSM+GPRS	UMTS

No próximo Capítulo, serão mostrados os diferentes caminhos a serem adotados pelas diferentes tecnologias para alcançarem a terceira e quarta geração de telefonia móvel, 3G/4G.

2.3.4 Sistema Universal de Telecomunicações Móveis – UMTS

As redes móveis sem fio empregando tecnologia UMTS têm sido objeto de intensa pesquisa e discussão recentemente, sendo uma tecnologia de dados de alta velocidade, fazendo parte dos padrões sem fio de terceira geração (3G) da família IMT-2000 (*International Mobile Telecomunications 2000)* da ITU (*International Telecommunication Union*) [54]. O WCDMA (*Wideband* CDMA) é a tecnologia de rádio utilizada pela UMTS. Um resultado disso é o uso intercalado dos termos UMTS e WCDMA. Na Europa, Japão, Coréia e nos Estados Unidos, o WCDMA já está sendo atualmente padronizado [2], [71].

A UMTS é a principal tecnologia 3G empregada atualmente, oferecendo o potencial de cobertura mundial e possibilitando economias de escala, ou seja, *roaming* global e tecnologia prioritária para desenvolvimento de aplicativos e *software*. Esta nova tecnologia irá promover uma grande variedade de serviços, especialmente relacionados a multimídia com alta taxa de transmissão [54], [70] e [71].

Duas fortes discussões marcaram o início dos estudos sobre as redes de terceira geração: uma defendia a criação de um único padrão mundial; a outra defendia a evolução das redes e sistemas atuais de forma a atender aos requisitos definidos a partir da visão 3G [2]. Apesar de ambas as discussões possibilitarem economia de escala de fabricação para os equipamentos da rede, a segunda teve maior força, pois também permite que os grandes investimentos já realizados pelas operadoras na implantação das redes, pelos fabricantes em processos de fabricação e etapas de desenvolvimento de produtos em todo o mundo, fossem de certa forma protegidos.

Dessa maneira, o ITU elaborou um conjunto de requisitos de tal forma que pudessem ser apresentadas propostas para as tecnologias de transmissão via rádio RTT (*Radio Transmission Technologies*) candidatas a compor o conjunto de especificações para o futuro padrão mundial de sistema de comunicação móvel 3G [73], [74]. A esse sistema foi inicialmente dado o nome de FPLMTS (*Future Public Land Mobile Telecommunication System*), com o objetivo de atender tanto aos usuários fixos (*FWA–Fixed Wireless Access)*, ou WLL (*Wireless Local Loop*) quanto aos móveis, em redes públicas e privadas. O nome FPLMTS posteriormente foi mudado para IMT-2000 [75], nome este que é mantido e reconhecido até hoje.

UMTS é visto como o sucessor do GSM. Este padrão indica a mudança para a terceira geração de redes móveis. Além disso, UMTS também aparece para atender à crescente demanda de aplicações móveis e para *Internet*, dentro do meio das comunicações móveis.

A UMTS foi desenvolvida para prover serviços com altos níveis de consumo de banda, como multimídia *streaming*, transferência de arquivos grandes e videoconferências para uma grande variedade de aparelhos como telefones celulares, PDAs e laptops. A UMTS utiliza uma combinação das tecnologias CDMA e TDMA para um uso altamente eficiente do espectro. Uma análise conjunta desenvolvida pelos membros da *3G Americas*, e publicada por *Rysavy Research* em Novembro de 2002, descobriu que quando comparado com outras

tecnologias de terceira geração, a UMTS é a mais eficaz no uso de espectro para taxas de dados acima de 100 Kbps [72].

Esse sistema irá prover acesso, através de um ou mais enlaces de rádio, a uma ampla gama de serviços de telecomunicações suportados por redes fixas como a PSTN (*Public Switched Telephony Network*), RDSI (Rede digital de Serviços Integrado), ATM (*Asynchronous Transfer Mode*), redes IP (*Internet Protocol*), entre outras, e serviços específicos a usuários móveis. Deverão existir vários tipos de estações móveis com capacidade de acesso fixo ou móvel a redes baseadas em satélites e/ou redes terrestres. A Tabela 2.3 ilustra as etapas do processo de implementação da UMTS.

Tabela 2.3: Progresso de Implementação da UMTS.

ETAPA DE IMPLEMENTAÇÃO	**CARACTERÍSTICAS**
Implementação inicial de UMTS	Apenas uma parte da área de cobertura possui a UMTS A GSM / GPRS / EDGE oferece cobertura contínua A UMTS oferece recursos avançados e capacidade extra para a GSM
Interligação avançada da UMTS com a GSM / EDGE	Cobertura contínua da UMTS Cargas mais pesadas para a UMTS Usuários dirigidos para bandas dependendo da demanda por serviços e cargas
Capacidade plena de rede multi-rádio	Implementação de alta densidade de UMTS, incluindo micro-células Integração de equipamento GERAN e UTRAN Integração de qualidade de serviço Adicionamento de novas tecnologias de rádio como WLANs

A implementação da UMTS ocorre em várias etapas, iniciando com apenas uma parte da área de cobertura contemplando a UMTS, chegando à cobertura UMTS contínua, até chegar na operação altamente integrada de múltiplos rádios.

Com o tempo, as diversas redes de acesso GSM / EDGE (GERAN) e UMTS (UTRAN) devem sofrer uma consolidação. Essa etapa reduzirá o custo total da rede e melhorará a operação integrada das redes.

Os principais atributos do UMTS podem ser divididos em duas categorias: benefícios para o usuário e benefícios para a operadora.

Os principais atributos para o usuário incluem: [76]:

- A UMTS suporta taxas máximas de 2 Mbps quando a estação móvel está estacionária e de 384 kbps quando em mobilidade. A taxa média de transmissão é de 220 a 320 Kbps, suficiente para uma grande variedade de serviços de dados avançados, que incluem vídeo e áudio *streaming* de alta velocidade, acesso rápido à *Internet* e transmissão de arquivos grandes. Por exemplo, um clip MMS (*Multimedia Messaging Service*) de 100 Kbps leva 26,7 segundos para ser baixado no GPRS, mas apenas 6,8 segundos numa rede UMTS com taxas médias de 128 Kbps [23].

- A UMTS, como a DSL (*Digital Speech Link*) e banda larga via cabo, oferece uma conexão permanente com a *Internet*.

- A UMTS é compatível com a EDGE e a GPRS permitindo ao usuário sair de uma área de cobertura UMTS e ser automaticamente transferido para uma rede EDGE ou GPRS, dependendo de fatores como disponibilidade de rede e o consumo de banda do seu aplicativo.

- A UMTS suporta mecanismos sofisticados de QoS, assegurando a alocação da quantidade de espectro e recursos de infra-estrutura necessária para cada tipo de dado oferecido ao usuário. Por exemplo, em um aplicativo tipo vídeo *streaming* poderia ser feito uma alocação prévia de banda suficiente para oferecer uma imagem estável e de qualidade.

- A UMTS é a principal tecnologia 3G no mundo, disponível em redes licenciadas em mais de 50 países abrindo a oportunidade de *roaming* global e compatível com padrões anteriores do GPRS e/ou EDGE [71].

Os principais atributos para a operadora incluem [76]:

- A UMTS reutiliza investimentos anteriores, especialmente a infra-estrutura de rede de pacotes de dados implementada para a GPRS. Dependendo do fabricante, a atualização pode ser facilmente executada adicionando cartões de canais e *software* UMTS à infra-estrutura de rádio existente de GSM / GPRS / EDGE, que continua a prover serviços para clientes utilizando essas tecnologias.

- A UMTS opera numa variedade de bandas de espectro novas e atuais, incluindo a 1900 MHz. A UMTS também utiliza o espectro de forma altamente eficaz devido a uma combinação de tecnologias de acesso como, CDMA e TDMA.

- A UMTS suporta mecanismos sofisticados de qualidade de serviço (QoS) oferecendo mais qualidade às operadoras, assegurando que cada aplicativo receba a quantidade certa de banda. Os mecanismos de QoS da UMTS também suportam *Voice over Internet Protocol* (VoIP), a etapa final da visão UMTS, onde voz e dados trafegam a mesma infra-estrutura de pacotes.

O alto grau de compatibilidade entre tecnologias de um padrão mundial não só possibilitará grande economia de escala, mas facilitará a implementação do *roaming* global e impulsionará o investimento da indústria de Tecnologia da Informação (*IT-Information Technology*) em aplicações tais como serviços de multimídia que farão com que as redes de comunicações móveis possam ser vistas como uma extensão sem fio da *Internet* [77].

Para permitir cobertura e *roaming* global o sistema UMTS contará com a componente terrestre e a componente via satélite, atendendo aos usuários pico-celulares em interiores (*indoor* ou *in-building*), micro e macro-celulares em exteriores (*outdoor*) e em regiões remotas com cobertura global via satélite.

A natureza assimétrica dos dados que predominantemente trafegarão pelas redes UMTS é um outro ponto importante a considerar, necessitando de maior capacidade no enlace direto que no enlace reverso e, por conseqüência, atribuições assimétricas de banda.

Observando as faixas de freqüência escolhidas para o sistema UMTS pode-se notar que elas são significativamente mais altas que aquelas utilizadas pela maioria dos sistemas de segunda geração (abaixo de 1 GHz). Esse fato antecipa uma dificuldade maior no planejamento e implantação dos sistemas UMTS, pois a influência dos obstáculos entre transmissor e receptor torna-se mais significativa e leva à necessidade de utilização de ferramentas de predição de cobertura e análise de interferências mais precisas e bases de dados com maior resolução. E além da limitação por interferências inerentes aos sistemas com estrutura celular, principalmente devido ao reuso de freqüências, os sistemas UMTS terão também maior limitação de potência que os de segunda geração, devido às condições de propagação nessa faixa de freqüências serem mais severas que em faixas de freqüência mais baixa.

Uma topologia hierárquica clássica seria, então, uma arquitetura em camadas de células de rádio, onde haveria as mega-células para satélites, macro-células para cobertura de grandes

áreas, micro-células para cobertura urbana e pico-células para cobertura de ambientes internos. Haveria uma mistura de uso privado e público [51].

2.3.5 Evolução das Redes 3G em Direção aos Sistemas 4G

A evolução GSM de tecnologia 3G para 4G inclui basicamente a mudança da comutação por circuito para a comutação por pacote. Pode ser vista também como a adequação da atual rede *Internet* com fio para uma rede *Internet* sem fio. A visão em termos de serviço da 4G não mudou muito em relação à visão original da 3G. Ao menos por enquanto, a principal diferença entre as duas tecnologias é apenas um aumento das taxas de transmissão de dados, assim como ocorreu com o aumento nas taxas de transmissão entre as tecnologias 2G/2.5G e 3G. Contudo, muitos dos serviços prometidos para a 4G, como *streaming* de vídeo de alta qualidade, TV móvel, entre outros, já estão disponíveis em 3G [18] [46].

As velocidades de transmissão para uma MS utilizando a tecnologia 4G, alcançarão taxas da ordem de 10 Kbps a 100 Mbps operando em bandas de 5 a 60 GHz. Espera-se que a tecnologia 4G entregue versões mais avançadas das mesmas melhorias prometidas pela tecnologia 3G, tais como, multimídia de alta definição, *streaming* de vídeos com melhor definição, acesso universal e mobilidade absoluta para as MSs e portabilidade entre todos os tipos de MS. A tecnologia 4G poderá conectar qualquer um em qualquer lugar, ou seja, ser operada independentemente da posição geográfica da MS no planeta [18], [20].

Com a implementação dos sistemas GSM (2G/2,5G), a base para a terceira geração já estava consolidada na Europa. A UMTS foi concebido há aproximadamente uma década e, a partir de 2002, já existiam planos concretos para a difusão da tecnologia na Europa. O complemento entre as tecnologias GSM e UMTS acarretou em uma nova geração de redes móveis e sem fio, a tecnologia de quarta geração (4G) [19] [20].

A transição da tecnologia 3G para 4G será baseada não somente nas taxas de dados, mas deverá corresponder a UMTS Fase 2 a qual é chamada de IMT-2010 [71]. Esta transição não será norteada somente pela tecnologia, mas esta nova geração de rede é vista como um caminho para introduzir um novo método de acesso.

A nova geração de rede exige o desenvolvimento de uma estrutura que deverá proporcionar o crescimento do tráfego de dados, juntamente com a convergência de serviços, em uma rede multi-serviços. A próxima geração consistirá em uma rede totalmente orientada a pacotes, que será caracterizada principalmente pelos seguintes aspectos [56]:

- Estrutura de rede em camadas ou níveis;
- Uma camada de conectividade composta por uma rede de *backbone* comum;
- Arquitetura cliente-servidor entre servidores na camada de conectividade e de controle de níveis (*gateways* de mídia);
- Desenvolvimento de interfaces e padrões abertos, essenciais nesse tipo de arquitetura.

Analisando as condições técnicas para a tecnologia 4G de redes móveis sem fio pelos parâmetros desenvolvidos, pode-se dizer que consistirão em uma rede completa, se um conjunto destes parâmetros, conforme descrito a seguir, forem implementados [19], [56]:

- Acesso a serviços de voz e dados que serão oferecidos pelas redes móveis. Isto exige gerenciamento dos recursos, como por exemplo, a utilização de extensões *ad hoc* em sistemas sem fio;
- A rede móvel sem fio deverá proporcionar conectividade plena a outras PDNs;
- A segurança das redes virtuais privadas deverá ser garantida;
- A rede deverá ser capaz de se realinhar por si mesma, ou seja, ela deverá gerenciar diversos tipos de *backbones* e utilizar a melhor opção;
- O sistema deverá ser capaz de prover Qualidade de Serviço (QoS);
- Parâmetro de disponibilidade deverá estar próximo de 100%;

- Aplicações tais como: leitura de notícias através de multimídia, funções de agenda de bolso, gravação de voz, entre outras, deverão ser acessadas por MSs a qualquer momento e independentemente da posição geográfica;
- Uma interface universal de *software / hardware* deverá ser padronizada para facilitar o desenvolvimento de novos aplicativos e serviços.

A utilização da banda de 60 GHz requer um aumento no número de pontos de acesso, porém pode acarretar em um alto custo nos acessos com interface aérea. Assim, para que o sistema suporte altas taxas de dados a um custo baixo, antenas inteligentes estão sendo propostas para sistema de curto alcance do tipo WLAN (*Wireless Local Area Network*). Sistemas operando a 5 GHz oferecem uma faixa melhor e tem a vantagem de que muitos usuários podem compartilhar um único ponto de acesso, que oferece maior flexibilidade para a operadora. Os resultados das pesquisas apontam que a operação a 5 GHz é uma alternativa importante em sistema sem fio 4G. Adicionalmente, um estudo sobre a infra-estrutura é necessário para comparar a cobertura e a QoS versus o custo da infra-estrutura para os sistemas propostos. Para sistemas operando no Brasil, a ANATEL disponibilizou as faixas de freqüências de 3,5 GHZ e 10,5 GHz licenciadas e 5,8 GHz não licenciada [78]. Os requisitos para ocupação do espectro, taxa de transmissão, área de cobertura e área de serviço para a rede, estão sendo discutidos em função dos seguintes parâmetros [56]:

- Taxa de transmissão de *bits*: operação entre 20 e 40 Mbps (160 e 320 Mbps) em ambientes abertos e veiculares;
- Espectro: ainda não está definido, porém para suportar taxas de transmissão da ordem de 160 Mbps, recursos de freqüências suficientes devem ser considerados. Estima-se que freqüências a partir de 3 GHz serão utilizadas;
- Área de cobertura: uma característica forte na tecnologia 4G corresponde a possibilidade de se comunicar independente do tempo e da posição geográfica das

MSs. Além do mais, os sistemas 4G deverão operar em uma banda de alta freqüência, o que acarretará em uma perda da propagação do sinal sem fio maior do que os sistemas 2G e 3G;

- Área de serviços hierárquicos: apesar de esperar-se que todos os modelos de MS estejam conectados a uma rede através de enlace sem fio, haverá uma certa dificuldade para os modelos compactos estarem diretamente conectados aos sistemas 4G, em virtude do consumo de potência e do tamanho da antena. No entanto, dispositivos de pequeno porte serão capazes de trocar sinais sem fio em uma faixa estreita. Portanto, estes dispositivos compactos serão capazes de acessar a rede 4G através de mini estações-base. Esta configuração corresponde à implementação de áreas de serviços que consistirão em múltiplas células sobrepostas [18] [19] [20].

2.4 Conclusão

Neste Capítulo foram apresentados os componentes básicos de uma rede móvel sem fio e sua evolução. As principais características da evolução da implementação da família IMT-2000, isto é, 3G e UMTS e os tópicos básicos da evolução dos sistemas, iniciando pela primeira geração, ou seja, 1G, destacada como a tecnologia analógica para comunicações móveis. A segunda geração ou 2G / 2,5G, empregando tecnologias digitais, revolucionando a idéia de comunicação móvel sem fio através do oferecimento de diversos tipos de aplicativos e serviços. A terceira para a quarta geração, ou 4G, introduzindo uma nova forma de comunicação através da utilização de comutação plena por pacotes e sistemas totalmente digitais.

Os sistemas de terceira geração, ou seja, 3G, possibilita taxas de dados de até 2 Mbps, que são consideravelmente mais altas que as taxas de 9,6 Kbps dos sistemas 2G, como o GSM. Espera-se que os sistemas 3G e 4G sejam totalmente baseados na comutação de pacotes

utilizando o protocolo IP. A mudança dos sistemas 2G/2,5G, como GSM e TDMA/IS-54, que utilizam comutação por circuitos e os de quarta geração que utilizam comutação plena por pacotes, tem implicações importantes. Esta geração é, já na sua essência, uma rede totalmente digital de comutação de pacotes somente (e não de circuitos) para acesso móvel sem fio sem restrições. Com a nova geração de sistemas 4G vislumbra-se alcançar de forma plena e consistente a convergência da *Internet* com as redes móveis sem fio com mobilidade absoluta.

O próximo Capítulo abordará o tema tecnologia para comutação de pacotes nas redes móveis sem fio como uma alternativa de mecanismo para o acesso à *Internet*. Serão abordadas também, as vantagens das inovações tecnológicas destes mecanismos no sentido de elevar o nível de QoS nas redes *backbone* com tecnologia UMTS.

Capítulo 3

Tecnologia para Comutação de Pacotes nas Redes Móveis sem Fio

3.1 Introdução

Nos últimos anos, tem havido um contínuo crescimento tanto na área de comunicação móvel sem fio quanto na *Internet*, o que gerou a necessidade de acesso sem fio à *Internet*. No campo das redes de telefonia, numa evolução natural às atuais redes fixas de comutação de circuitos, surgiram as redes móveis que, além dos tradicionais serviços de voz, agregam novos recursos de comunicação móvel de dados. Tais tecnologias vêm ao encontro da crescente demanda do mercado de aplicativos e serviços que exigem não só altas velocidades, mas também Qualidade de Serviço (QoS), que são requisitos fundamentais. Neste cenário, novas tecnologias de comutação de pacotes em sistemas móveis sem fio usando tecnologias GSM e GPRS tem surgido em todo mundo, o que torna ainda mais oportuno o conhecimento e a análise destas arquiteturas e seus aspectos de QoS.

Consoante com esses objetivos, o GPRS surgiu para atender a essa demanda, estendendo uma das tecnologias mais utilizadas na área GSM [11]. Enquanto GSM utiliza comutação por circuitos, o GPRS utiliza comutação por pacotes para transmitir dados, permitindo

comunicação entre as estações móveis e redes externas de pacotes de dados, incluindo a *Internet*.

3.2 Arquitetura de Rede GPRS

A arquitetura GPRS é considerada como um serviço ou uma extensão do GSM, sendo que ambas utilizam a mesma infra-estrutura e compartilham os mesmos recursos de rádio. Isto é possível porque o GPRS utiliza a mesma banda de freqüência e técnicas de salto, a mesma estrutura do quadro TDMA e a mesma modulação que o GSM.

Para oferecer um serviço de comutação de pacotes, algumas modificações foram necessárias. Alguns dos nós já implementados nos sistemas atuais podem ser compartilhados entre o GPRS e os sistemas GSM. Apenas dois novos tipos de nós, o SGSN e o GGSN, foram adicionados à nova estrutura. A Figura 3.1 ilustra a arquitetura de uma rede GPRS com seus principais componentes e interfaces [21].

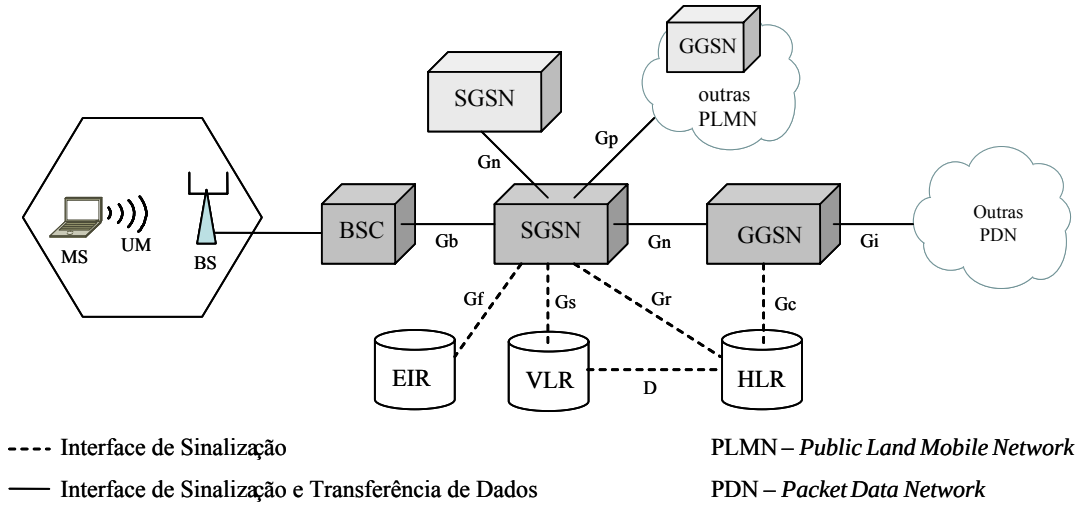

Figura 3.1: Arquitetura de uma rede GPRS.

O SGSN é um servidor GPRS responsável pela localização de uma MS dentro de sua área de serviço e fazer o roteamento de pacotes dentro da PLMN. O SGSN serve todos os

assinantes GPRS que estiverem fisicamente localizados dentro da área de serviço deste SGSN. O tráfego é roteado do SGSN para a BSC, BS e então MS. O SGSN também é responsável por fazer a tarifação e autenticação, gerenciamento de mobilidade, gerenciamento de enlace lógico para o MS, segurança e conexões para o HLR, MSC, BSC e GGSN.

O GGSN atua como um nó *gateway* entre uma rede de pacotes de dados externa e a rede *backbone* do GPRS. Para o caso de uma rede IP externa, o GGSN pode ser visto como um roteador IP convencional servindo a todos os endereços IP das estações móveis. Este nó pode incluir *firewalls* e mecanismo de filtragem de pacotes. Além disso, sua tarefa é atribuir o SGSN correto para a estação móvel, dependendo da localização geográfica da mesma.

Uma rede GPRS utiliza o mesmo BSS que os serviços de voz GSM, com uma atualização de *software* na BTS e no BSC e a instalação de um novo componente de *hardware* chamado PCU (*Packet Control Unit*), geralmente localizado no PCUSN (PCU *Support Node*) [10]. O PCUSN reside tipicamente entre o BSC e o SGSN sendo o responsável pela transferência de pacotes entre o BSS e a rede GPRS [21].

As entidades de armazenamento de dados VLR, HLR, EIR e AUC também fazem parte da arquitetura de uma rede GPRS. O AUC e o EIR são utilizados para controlar a segurança e autenticação dos assinantes móveis, o HLR mantém informações atualizadas de cada assinante registrado na rede GPRS (tais como: IMSI, informações de localização corrente, perfil de serviços contratados, endereço do SGSN corrente, um ou mais endereços PDPs (*Packet Data Protocol*), permissões de *roaming* e estado de atividade) e o VLR armazena informações sobre usuários ativos dentro da sua área de serviço, por exemplo, usuários em *roaming*.

A MSC é uma entidade que é usada em GPRS somente quando ambos os serviços de dados comutados a pacote e comutados a circuito são suportados. O tráfego de voz é enviado pelo

BSC para o MSC como no sistema GSM, enquanto o tráfego de dados é enviado para o SGSN, através da PCU, em uma interface *Frame Relay*.

A BTS é responsável pela modulação, demodulação e transmissão e recepção no enlace de rádio. O BSC suporta os protocolos GPRS relevantes para a comunicação na interface aérea, gerencia as chamadas e executa a atribuição de recursos de rádio [10].

Foi introduzido um novo conjunto de interface para permitir a interação entre os elementos de uma rede GPRS, conforme mostrado na Figura 3.1. Dentre elas destacam-se: Interface Gb entre a PCU e o SGSN, usando *Frame Relay;* Interface Gr responsável pela troca de informações entre o SGSN e o HLR; Interface Gn utilizada entre o SGSN e o GGSN quando estes estão localizados na mesma PLMN; Interface Gp utilizada entre o SGSN e o GGSN caso estes estejam em PLMNs diferentes. Ambas as interfaces Gn e Gp transmitem dados do usuário e de sinalização entre o SGSN e o GGSN; Interface Gc possibilita a troca de informações entre o GGSN e o HLR; Interface Gf utilizada entre o SGSN e o EIR; Interface Gs utilizada entre o SGSN e o VLR; Interface Gi: utilizada entre o GGSN e PDNs externas.

3.3 Arquitetura EDGE

A tecnologia EDGE, é uma tecnologia de terceira geração (3G) de transmissão de dados e acesso à *Internet* de alta velocidade que transmite dados em velocidade de até 473 kbps. As taxas médias são rápidas o suficiente para permitir um amplo leque de serviços de dados avançados, incluindo *Streaming* de áudio e vídeo, acesso rápido à *Internet* e *download* de arquivos pesados [72]. A EDGE é conhecida como GPRS ampliada (E-GPRS; de *Enhanced* GPRS) porque aumenta em três ou quatro vezes a capacidade e o *throughput* de dados da GPRS. Isto é possível através da otimização da interface de rádio, permitindo que os outros componentes da rede, incluindo BSC, SGSN, GGSN e HLR, permaneçam inalterados. Assim

como a GPRS, a EDGE é um serviço baseado em pacotes que oferece aos clientes uma conexão permanente [22].

A tecnologia EDGE é totalmente compatível com a GPRS, e qualquer aplicação desenvolvida para o ambiente GPRS irá funcionar da mesma forma. A EDGE é uma tecnologia que emprega três técnicas avançadas na conexão de rádio, que permitem que atinja uma eficiência espectral mais alta. A primeira técnica é o acréscimo de um novo esquema de modulação por chaveamento de fase, conhecido como 8PSK. Esta técnica permite que o sinal de rádio transmita três *bits* de informação para cada símbolo de rádio. Em comparação, a modulação GSM/GPRS utiliza a modulação GSMK, que transmite um *bit* de informação por símbolo de rádio. A segunda técnica consiste na utilização de esquemas múltiplos de codificação, em que a rede pode ajustar o número de *bits* utilizados para o controle de erros de acordo com o ambiente do rádio. A tecnologia EDGE possui cinco esquemas de codificação disponíveis para o 8PSK e quatro esquemas para o GSMK, proporcionando até nove modulações e esquemas de codificação diferentes. O sistema seleciona dinamicamente a modulação e o esquema de codificação ótima para o ambiente de rádio disponível no momento. A terceira técnica consiste no emprego do mecanismo conhecido como redundância incremental. No caso de os blocos de dados serem recebidos com erro, é enviado um incremento de redundância de dados, para a correção destes, em cada retransmissão. Assim, cada retransmissão possui maior probabilidade de sucesso que a anterior [22] [72].

Além de melhorias no desempenho de rádio, a tecnologia EDGE apresenta outra função importante, que é a arquitetura de qualidade de serviço utilizada pelo UMTS. Esta arquitetura baseia-se nas especificações da versão 99, feitas pelo 3GPP. A seguir são descritas algumas vantagens apresentadas pela tecnologia EDGE em relação ao GPRS:

- **Velocidade:** a EDGE permite transmissão de dados em velocidade de até 473 kbps e taxa média entre 110 e 130 kbps. O *download* de um videoclipe de MMS de 100 Kb,

por exemplo, demora 26,7 segundos na GPRS, mas apenas 10 segundos na EDGE [23].

- **Conexão permanente:** assim como a conexão de banda larga por cabo e DSL, a EDGE oferece conexão permanente à *Internet*, eliminando a necessidade de se conectar cada vez que for usar. A EDGE também permite que os clientes mantenham uma transmissão de dados enquanto estão atendendo a uma ligação telefônica. Isto é totalmente único em tecnologias GSM [22].
- **Otimização:** a EDGE é um serviço baseado em pacotes que constitui uma forma mais eficiente para as operadoras oferecerem serviços. Transmissão por pacotes também significa que os clientes pagam apenas pelos dados que recebem e enviam, não pelo tempo gasto para iniciar uma conexão e esperar que o servidor responda.
- **Cobertura:** Considerando que a tecnologia EDGE é também compatível com a GPRS, quando os usuários saem da área de cobertura EDGE eles automaticamente passam a estar ligados a uma rede GPRS, que está disponível praticamente em todo o mundo.
- **Eficiência e flexibilidade quanto ao espectro:** a EDGE permite que as operadoras lancem a tecnologia 3G/UMTS agora usando o espectro já existente. Por ser uma tecnologia de banda estreita que utiliza canais de 200 KHz, a EDGE não exige um grande espectro, e pode ser implementada na maioria das bandas em uso atualmente: 850, 900, 1800 e 1900 MHz.
- **Facilidade para atualizar:** a EDGE utiliza a mesma estrutura de *frames* TDMA, canal lógico, e uma largura de banda de 200 KHz na portadora, como as redes GSM atuais. Por isso, implementar a EDGE não exige uma grande reengenharia dos planos de células. Se a infra-estrutura de rádio de uma operadora tem menos de cinco anos como é o caso da maioria das operadoras GSM, fazer o *Upgrade* para a EDGE em

geral implica ter novo S*oftware* e cartões de canal em cada local de célula. Portanto, atualizar todos os locais de célula numa grande cidade leva apenas algumas semanas, o que permite que as operadoras lancem serviços EDGE rapidamente. O custo relativamente baixo do *Upgrade* para a EDGE também implica que as operadoras podem fixar preços muito mais competitivos para seus serviços avançados do que se o processo exigisse a substituição dos principais elementos da infra-estrutura de rádio.

- **Maior controle:** a tecnologia EDGE inclui sofisticados mecanismos de controle da qualidade de serviço (QoS), que dão às operadoras mais controle, garantindo que cada aplicativo ou cliente tenha a largura de banda correta. A QoS é fundamental para um serviço voltado para clientes empresariais, e sistemas que requerem banda larga e altas velocidades [22] [72].

3.4 Arquitetura de Rede UMTS

As redes *backbone* UMTS serão totalmente baseadas no protocolo IP (*All IP*). Para a integração da rede *backbone* UMTS com a *Internet* é necessário um suporte eficiente do IP móvel, proposta da IETF para o suporte de mobilidade na camada de rede. A abordagem do 3GPP (*Third Generation Partnership Project*) [24], [55] para a integração com o IP Móvel é baseada em nós da rede *backbone* UMTS chamados de SGSN e GGSN [1], [2] , [3].

A arquitetura de uma rede UMTS consiste na interação de três domínios *backbone* UMTS, composta pela rede de núcleo CN (*Core Network*), *Terrestrial Radio Access Network (UTRAN)* e *User Equipment* (UE)/MS. A principal função do *backbone* UMTS é fazer comutação, roteamento e trânsito para tráfego e usuários. O *backbone* UMTS também contém o banco de dados e as funções de gerenciamento da rede. A Figura 3.2 ilustra os elementos básicos da arquitetura UMTS [6].

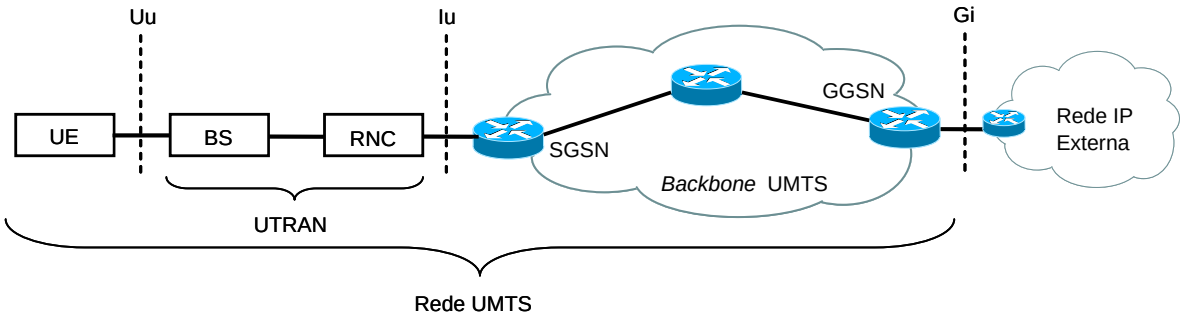

Figura 3.2: Elementos básicos da arquitetura UMTS.

A arquitetura básica do *backbone* UMTS é baseada na rede GSM com GPRS. Todos os equipamentos têm de ser modificados para operações e serviços UMTS. A UTRAN fornece para o UE/MS, o método de acesso à interface aérea. A Estação Base é conhecida como Node-B e o equipamento que a controla é chamado de *Radio Network Controller* (RNC). O tráfego de pacotes está concentrado em um novo elemento comutador, o SGGN. Na borda da rede *backbone* UMTS está o GGSN comutador de pacotes que é um *gateway* que se conecta à rede IP externa ou às *intranets* corporativas. A interoperabilidade entre diferentes nós da arquitetura de rede UMTS é definida através de interfaces padrão, sendo algumas abertas e outras serão desenvolvidas através de soluções proprietárias. A Figura 3.3 ilustra uma arquitetura UMTS simplificada composta de uma rede *backbone* UMTS conectada à rede de acesso UTRAN através da interface Iu [6].

A arquitetura UMTS compreende uma rede *backbone* UMTS conectada à rede de acesso UTRAN através da interface Iu. A UTRAN é conectada ao equipamento do usuário UE com a interface Uu. A rede *backbone* UMTS é conectada à rede IP externa através da interface Gi. A funcionalidade dos principais componentes e interfaces são descritos a seguir:

A Estação Base (BS) (Nó B): este elemento é responsável pela conversão, transmissão e recepção de dados na interface aérea (Uu) oriundos dos nós móveis e destinados às MS. É

responsável pela correção de erros, adaptação de taxa e modulação. Cada nó B pode ter um número de transmissores de rádio e cobre um número de células e também envia relatórios de medições para o RNC.

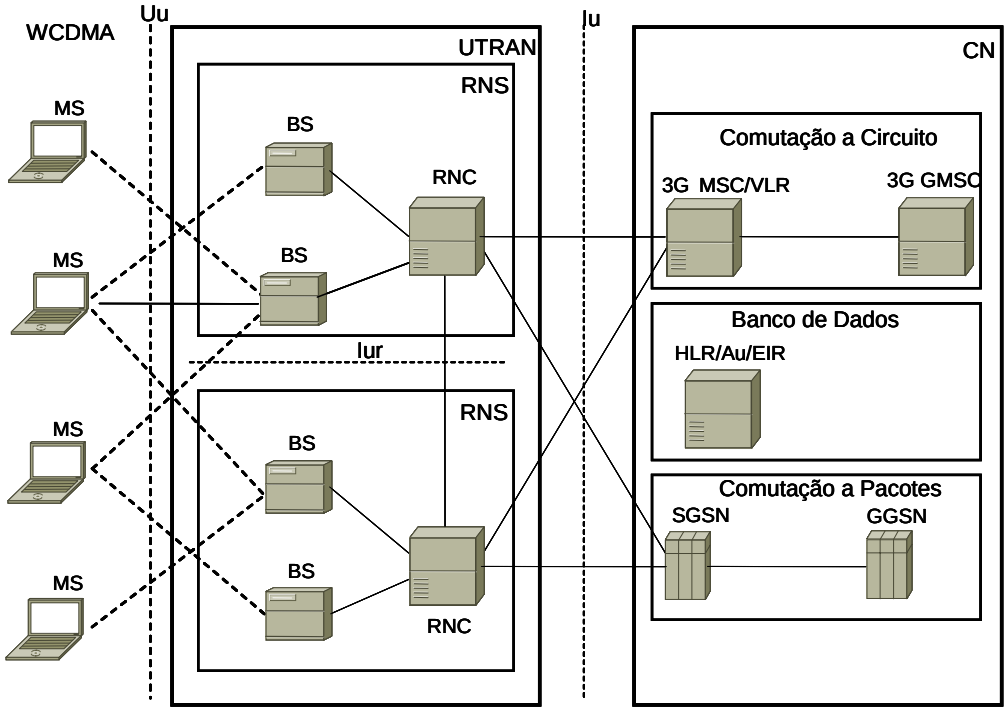

Figura 3.3: Arquitetura UMTS.

O Controlador de Rede de Rádio (RNC): é um comutador ATM que pode multiplexar e demultiplexar dados provenientes da comutação de circuitos e da comutação pacotes juntos. Ao contrário do GSM, os RNCs estão conectados juntos através da interface Iu e, portanto, podem manipular todas as questões de recursos de rádio de maneira autônoma. Cada RNC controla um número de nós BSs, sendo o conjunto conhecido como Sistema de Rede de Rádio (RNS). O RNC controla congestionamento e *handoff* suave (envolvendo deslocamento entre diferentes nós Bs), assim como é responsável pela operação e manutenção dentro do RNS.

O servidor GPRS (SGSN): é responsável pelo gerenciamento de sessão, produzindo informação de tarifação, e também faz o encaminhamento de pacotes até o correto RNC.

Funções como estabelecimento de sessão e os caminhos de QoS também são manipulados pelo SGSN, que também controla a localização dos usuários móveis, além de outras informações relativas ao assinante e seu equipamento móvel. Existe um SGSN para cada área de serviço, sendo que este é normalmente colocado junto ao MSC;

O *gateway* IP (GGSN): é um roteador de borda para interconexões com a rede IP externa. O GGSN conecta a rede núcleo GPRS a uma ou mais redes de comutação de pacotes de dados PDNs (*Packet Data Network*). Entre suas funções está a de converter os pacotes entrantes para o protocolo apropriado, a fim de encaminhá-los à PDN. Além disso, o GGSN é responsável pelo gerenciamento de sessão e pela correta atribuição de um certo SGSN a uma MS, dependendo da sua localização. O GGSN também contribui na coleta de informação para fins de bilhetagem, feita pelo subsistema de tarifação GPRS.

Centro de Comutação Móvel / Registro de Posição de visitantes (MSC/VLR): este elemento manipula as responsabilidades da comutação de circuitos orientados à conexão, incluindo gerenciamento de conexão e funções de gerenciamento de mobilidade. É também responsável por algumas das questões de segurança e pela geração do registro de detalhe da chamada CDR (*Call Detail Record)* para fins de bilhetagem;

O *gateway* MSC (GMSC): responsável pelas conexões da rede móvel com outras redes fixas externas, tais como, PSTN e ISDN para o tráfego comutado por circuito.

Registro de Posições dos usuários Locais (HLR): este elemento executa as mesmas funções das redes GSM já citadas em seções anteriores. Trata-se de um grande banco de dados com informações sobre usuários e seus serviços (por exemplo, se são pré ou pós-pagos, se estão com a função *roaming* ativada e as classes de QoS as quais assinam). Novas classes foram adicionadas para o UMTS – especialmente relacionadas aos serviços de dados.

A UTRAN compreende de um conjunto de RNSs (*Radio Network Subsystems*) conectado à rede *backbone* UMTS através da interface *Iu*. Um RNS consiste de um RNC e um ou mais

nós do tipo *B* (BS). Um nó *B* é conectado a um RNC através da interface *Iub*. O RNC é responsável pela decisão de *handoff* que requer sinalização até o UE. Na UTRAN, os RNCs dos RNSs podem ser interconectados através da interface *Iur*. Cada RNS é responsável pelos recursos do seu conjunto de células e cada nó *B* possui uma ou mais células.

Interface Iu: responsável pela conexão da UTRAN com a rede de núcleo (CN). Trata-se de uma interface aberta que possibilita às operadoras adquirir UTRAN e rede CN de diferentes fornecedores. Desempenha, também, outras funções, como *handoff* entre RNSs, acesso à portadora de rádio, gerenciamento de sobrecarga, informação referente a dados com erro de transmissão entre outros.

Interface Iur: protocolo que possibilita a "comunicação" entre RNCs diferentes. As funções desempenhadas por essa interface incluem suporte ao *handoff*, suporte a registros de localização de estações móveis entre RNCs, transferência de medições da ERB entre dois RNCs etc;

Interface Iub: esta interface permite comunicação entre RNC e Nó B, e desempenha funções como estabelecimento do "canal" ou "enlace" de rádio para atender a uma estação móvel, tratamento dos canais de controle, troca de informações sobre medições da ERB e gerenciamento de falhas.

O gerenciamento de mobilidade em um sistema *backbone* UMTS geralmente ocorre dentro da RAN. Um dos objetivos do planejamento e projeto para implantação da RAN é esconder da rede *backbone* UMTS, o quanto for possível, as conseqüências da mobilidade do usuário quando ele se move entre as estações base.

3.4.1 A Pilha de Protocolos UMTS

A Figura 3.4 mostra uma estrutura simplificada da pilha de protocolos para o domínio UMTS. Todos os dados dentro do *backbone*, isto é, entre SGSN e GGSN, são transferidos

usando o GTP (*GPRS Tunnelling Protocol*) que encapsula as PDUs que passam pelo *backbone* GPRS, adicionando informação de roteamento [15]. O protocolo IPv6 é utilizado como protocolo da camada de rede no *backbone* GPRS [23], [79]. Os protocolos das pilhas Ethernet, ISDN ou ATM podem ser utilizados abaixo do IPv6.

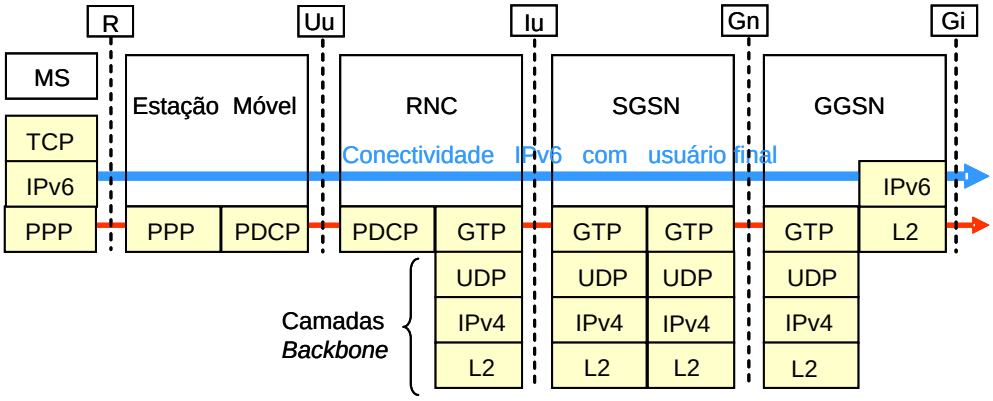

Figura 3.4: Arquitetura simplificada da pilha de protocolos UMTS [23].

3.4.2 A Interface Aérea do UMTS

O padrão WCDMA ou CDMA de banda larga é uma interface aérea do UMTS, como é denominada a IMT-2000 para o ITU, que usa técnicas de acesso baseadas em tecnologia CDMA que é a tecnologia recomendada pela ITU. O WCDMA suporta a tecnologia UMTS para a transmissão de sinais multimídia pela sua eficiência na utilização do espectro disponível, possibilita o aumento na capacidade de rede através da técnica de estruturação hierárquica de células e demodulação bidirecional. A seguir serão descritos os principais modos de acesso:

Modo FDD (*Frequency Division Duplexing*): Esta técnica utiliza o WCDMA com *direct sequence spreading*, com o *uplink* e *downlink* utilizando freqüências distintas. A estação móvel faz a transmissão via *uplink* utilizando uma portadora entre 1920 e 1980 MHz, a estação base usa a faixa de 2110 a 2170 MHz para *downlink*. Este modo proporciona

aproximadamente 250 canais para o tráfego do usuário. Cada canal físico corresponde a uma freqüência da portadora, um certo código de espalhamento DS (*Direct Sequence*) e uma fase relativa (apenas no *uplink*) e proporciona taxas de dados de 2 Mbps. O WCDMA no modo FDD suporta vários canais físicos que transmitem os dados nos canais de transporte.[24]

As estruturas de quadros UMTS no modo FDD são apresentadas na Figura 3.5. Um *superframe* com a duração de 720 ms consiste de 72 quadros (*frames*) que contêm, por sua vez, 16 *slots* com duração de 10 ms para cada *slot*. O exemplo ilustra três canais diferentes que podem utilizar um *slot*. Nos canais de controle, o *pilot*, é usado para obter estimativas do canal. O modo FDD do UMTS suporta 4,096 Mchip/s e pode ser estendido para 8,192 e 16,384 Mchip/s em futuras aplicações. [5]

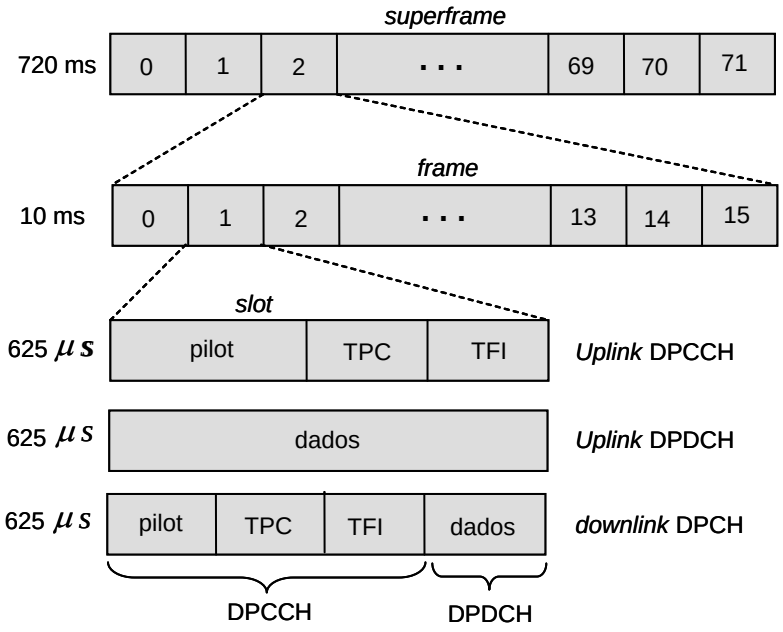

Figura 3.5: Estrutura de quadros no modo FDD do UMTS.

Modo TDD (*Time Division Duplexing*): Esta técnica utiliza o *Wideband TDMA/CDMA* para o acesso ao meio e, a mesma freqüência é utilizada tanto pelo *uplink* quanto pelo

downlink. Este modo oferece 2 Mbit/s para, aproximadamente, 120 canais ortogonais para o tráfego de usuário. Isto significa que esta técnica utiliza aproximadamente metade da capacidade do modo FDD, entretanto, requerendo apenas metade da largura de banda. Canais físicos são representados por um *slot* e um código de espalhamento (*direct sequence*). Tipicamente, o sistema pode transmitir oito rajadas de canais diferentes no mesmo *time slot*. A Figura 3.6 mostra uma estrutura de rajada típica para um *slot* com um campo para dados do usuário, um para estimativa (*midample*) do canal e um campo GP (*Guard Period)* para evitar interferências entre diferentes slots (23,4 ms). Cada *slot* tem duração de 625 ms e contém 2.560 *chips*. Um quadro é formado por 16 *slots*, cada *slot* com duração de 10 ms.

A estrutura de modulação no modo TDD é o QPSK. O controle de potência pode ser menor comparado ao do FDD (são necessários apenas de 100 a 800 ciclos de controle de potência por segundo).

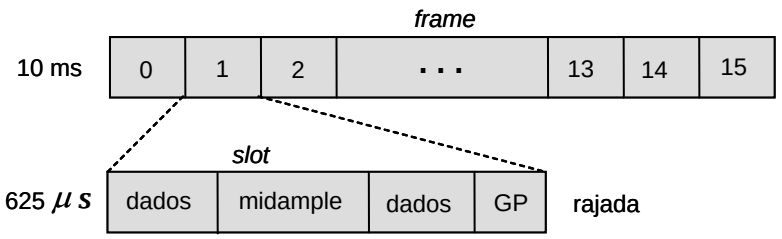

Figura 3.6: Estrutura de quadro para o modo TDD do UMTS.

O WCDMA por sua vez, oferece uma oportunidade para criar um padrão mundial para serviços UMTS oferecendo *roaming* entre todos os paises, pois apresenta uma facilidade maior de migração dos atuais sistemas GSM, que representam aproximadamente 80% de todos os sistemas em funcionamento no planeta, para sistemas 3G/4G com tecnologia UMTS.

3.4.3 QoS em Redes UMTS

Um elemento chave no transporte de serviços multimídia em redes UMTS será a Qualidade de Serviço (QoS) oferecida às aplicações, devido à natureza da rede, baseada em pacotes (*All IP*), e aos recursos limitados da interface aérea. A mobilidade das MSs entre redes (*roaming*) terá um impacto nas soluções de QoS, que necessitam garantir o nível de serviço quando os usuários finais cruzarem os domínios. A arquitetura para QoS das redes *backbone* UMTS é uma evolução da arquitetura de QoS para o GPRS. Nesta arquitetura um *bearer service*, define as características e funcionalidades entre os pontos finais da comunicação com o objetivo de realizar um suporte consistente de QoS para serviços fim a fim. O *end-to-end bearer service* é realizado através de uma arquitetura em camadas, mostrada na Figura 3.7.

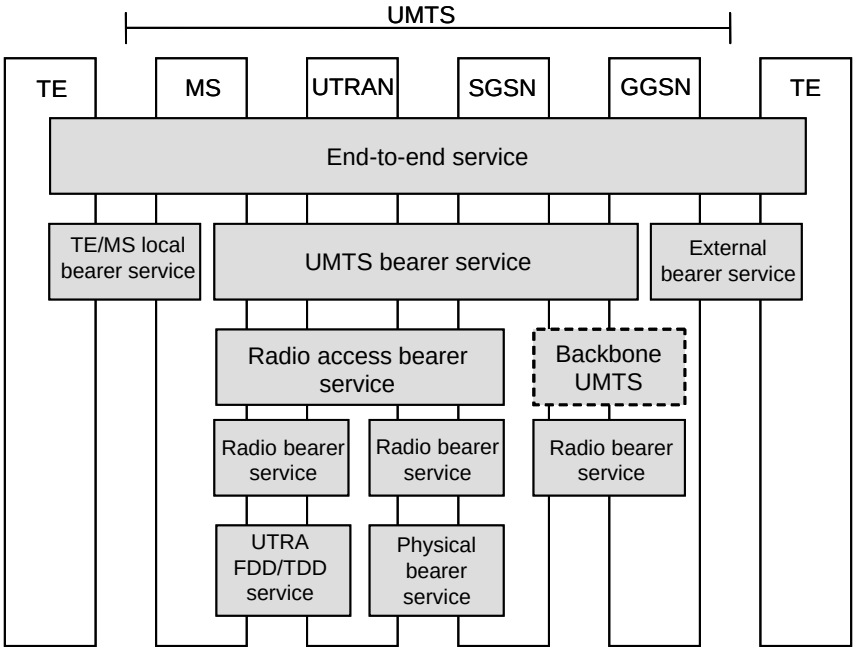

Figura 3.7: Arquitetura de QoS no padrão UMTS [55].

O *end-to-end bearer service* é a composição dos serviços das diferentes redes ao longo do caminho fim a fim quando várias redes estão envolvidas. Na arquitetura UMTS, o *end-to-end bearer service* pode ser compreendida em três componentes básicos [6], [19]:

- O UE/MS *local bearer service*, que permite a comunicação entre os diferentes componentes de uma estação móvel. Estes componentes correspondem às MSs, que é responsável pela conexão física com a UTRAN por meio da interface aérea, e um ou mais dispositivos de usuário, tais como, *laptops* , PDAs e telefones móveis tradicionais.
- O UMTS *bearer service*, que é oferecido pela operadora. Provê QoS dentro da rede UMTS e realiza as funções necessárias para o interfuncionamento com as redes externas. Este serviço ainda é composto do RAB (*Radio Access Bearer) Service* e do *backbone* UMTS *bearer service*.
- O *external bearer service* conecta o *backbone* UMTS e o nó de destino localizado em uma rede externa. Este serviço pode utilizar o transporte IP ou outras alternativas.

As funções de gerenciamento de QoS em redes *backbone* UMTS para aplicações multimídias incluem um mecanismo que efetue o controle de admissão, classificação, mapeamento e gerenciamento da mobilidade das MSs, bem como, a administração dos recursos nos pontos de ingresso à rede *backbone* UMTS. A função de gerenciamento de recursos é realizada através de mecanismos de escalonamento, gerenciamento de filas e controle dos recursos de rádio, a fim de distribuir os recursos disponíveis entre os serviços.

O estabelecimento de QoS dentro da rede *backbone* UMTS se dá através do procedimento de ativação de um contexto PDP, que é ativado através de um pedido feito pela estação móvel ao SGSN, onde estão localizados o perfil de QoS e outros parâmetros necessários à ativação deste.

3.4.4 Ativação do Contexto PDP

O contexto PDP deve ser estabelecido antes que uma estação móvel possa efetivar a comunicação com a rede de dados externa e após o seu registro com um SGSN da sua rede *backbone* UMTS. O contexto PDP é um protocolo de sinalização usado para estabelecer diferentes comunicações possibilitando a interoperabilidade com PDNs externas. O contexto PDP descreve as características da conexão com a rede de dados externa, como: tipo de rede, endereço de rede, APN (*Access Point Name*), QoS, etc. Para que isso ocorra a MS requisita uma ativação de contexto PDP. O SGSN valida a requisição baseando-se na informação de inscrição recebida do HLR durante o processamento do GTP [16], [58].

O APN é enviado para o DNS (*Domain Name Server*) no SGSN para que se encontre o endereço IP do GGSN relevante. É criada uma conexão lógica entre o SGSN e o GGSN / GTP. O GGSN atribui à estação móvel um endereço IP dinâmico da faixa de endereços alocados à PLMN ou a partir de um servidor RADIUS (*Remote Authentication Dial In User Service*).

Após a estação móvel ter se registrado com um SGSN do seu *backbone* UMTS, ela precisa alocar um ou mais endereços da camada de rede, chamados de endereços PDP, de modo a enviar ou receber pacotes de uma estação móvel residindo em uma PDN externa. Essa operação estabelece uma associação entre o SGSN com a qual a MS se registrou e o GGSN que mantém o(s) endereço(s) PDP atribuído(s) à MS, possibilitando a interoperabilidade com PDNs externas.

O registro da associação entre um endereço PDP individual que identifica a aplicação de um usuário móvel, um SGSN e um GGSN, é chamado de contexto PDP. Os seguintes parâmetros podem ser definidos em um contexto PDP:

- Endereço PDP (PDN) e tipo PDP (IPv4, IPv6, PDN) atribuído para uma MS;
- Endereço de um GGSN que serve como ponto de acesso para uma PDN externa;

- Perfil de QoS: define a QoS esperada por um assinante móvel em termos dos seguintes parâmetros: atraso, confiabilidade, precedência de serviço e *throughput*.

A ativação de um contexto PDP pode ser executada por uma MS ou pela rede (GGSN). Os procedimentos utilizados para a ativação de um contexto PDP iniciado por uma MS estão ilustrados na Figura 3.8 [16]:

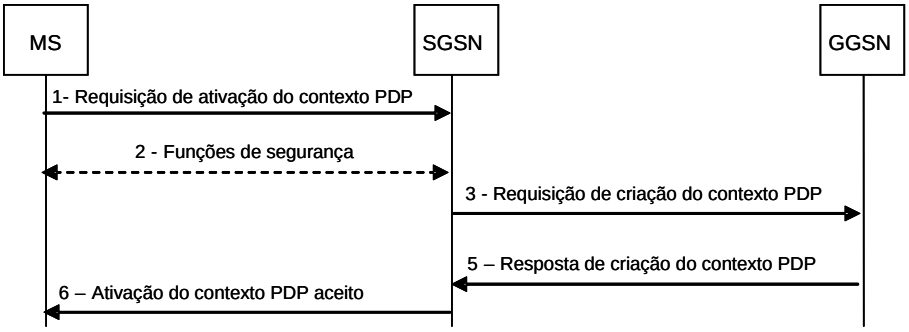

Figura 3.8: Ativação do contexto PDP pela MS.

- Uma MS envia uma mensagem de "Pedido de Ativação de Contexto PDP" para o SGSN contendo as seguintes informações: um endereço PDP, um tipo PDP, um APN opcional que identifica um GGSN através de um endereço IP ou um nome lógico, um NSAPI e um perfil de QoS;
- Para aceitar o perfil de QoS solicitado, o SGSN precisa autenticar a MS e verificar o contrato de serviço assinado com a operadora de rede em termos dos níveis de QoS; A qualquer momento, dependendo das condições de carga de uma PLMN GPRS e dos recursos disponíveis, um perfil de QoS pode ser modificado pelo SGSN;
- O SGSN envia o APN recebido a um DNS através de um pedido de busca. O DNS responde ao pedido com uma lista, em ordem de preferência, dos GGSNs disponíveis para uso. O SGSN seleciona o GGSN adequado e envia a este uma mensagem de "Pedido de Criação de Contexto PDP". Para enviar essa mensagem, um túnel GTP

(caminho bi-direcional ponto-a-ponto) é aberto entre o SGSN e o GGSN, caso este ainda não exista. Esse túnel é identificado por um TID (*Tunnel Identifier*);

- O GGSN responde ao pedido retornando uma mensagem de confirmação "Resposta de Criação de Contexto PDP" para o SGSN contendo o endereço PDP para a MS e o perfil de QoS negociado. Além disso, o GGSN atualiza uma das entradas de sua tabela onde ele mapeia o TID e o endereço IP do SGSN com a MS associada a eles;
- O SGSN envia uma mensagem para a MS indicando que o contexto PDP solicitado foi ativado corretamente. Finalmente, o SGSN também atualiza uma entrada de sua tabela mapeando a identidade móvel com o TID e o endereço IP do GGSN com o qual ele estabeleceu o túnel.

Para o caso em que o pacote destinado a um endereço PDP específico for recebido de uma PDN e não existir um contexto PDP ativo para esse endereço, o GGSN poderá simplesmente descartar o pacote ou tentar ativar um contexto PDP com a MS correspondente. A Figura 3.9 ilustra a ativação do contexto PDP pela rede.

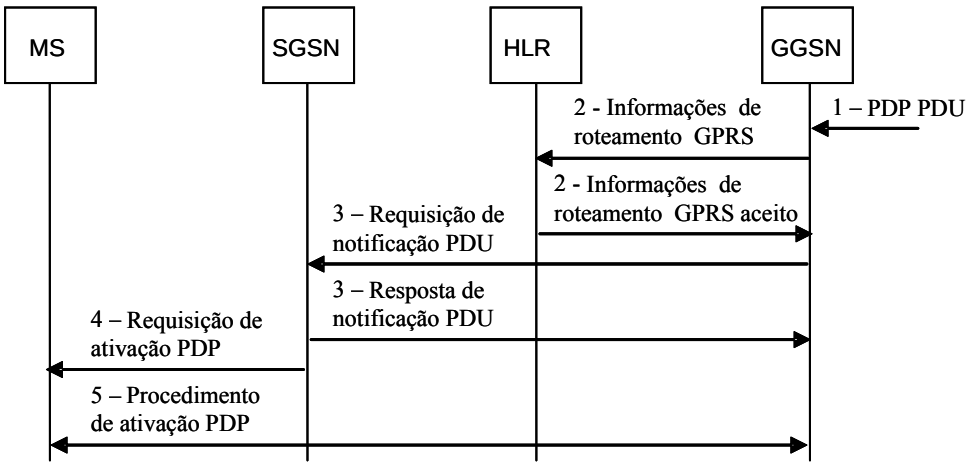

Figura 3.9: Ativação do contexto PDP pela rede.

Se necessário, o GGSN, pede ao HLR informações de roteamento necessárias para o endereço PDP. Se o HLR possuir essas informações, ele enviará uma mensagem de

confirmação contendo a IMSI e o endereço IP do SGSN servindo a MS. Caso contrário, o HLR informará a causa indicando a razão para a resposta negativa. Se uma resposta positiva for recebida, o GGSN pedirá ao SGSN para enviar uma mensagem a MS (contendo o endereço PDP e o tipo PDP) solicitando a ativação de um contexto PDP. As Figuras 3.10 e 3.11 ilustram a desativação do contexto PDP.

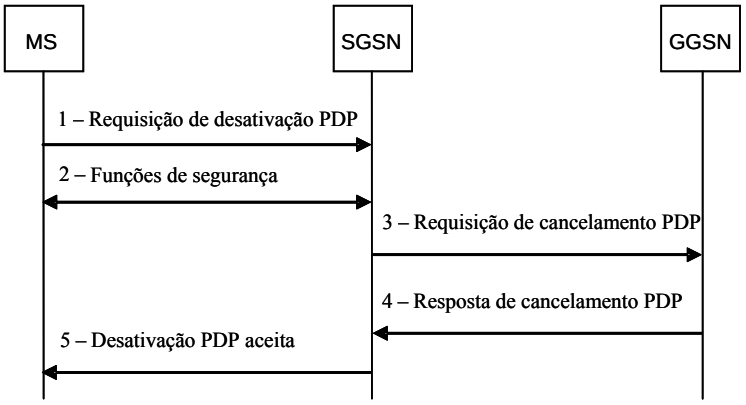

Figura 3.10: Desativação do contexto PDP pela MS.

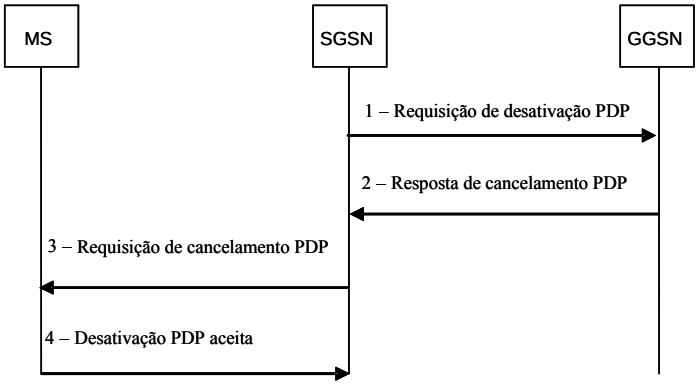

Figura 3.11: Desativação do contexto PDP pela rede.

A desativação do contexto PDP pode ocorrer a partir da MS ou a partir de outras redes de dados externa (PDNs).

3.4.5 Mapeamento das classes UMTS QoS

Os tipos de tráfegos para a arquitetura de QoS UMTS definidos pelo 3GPP, são divididos em quatro classes. Esta classificação é baseada nos requisitos de atraso, taxa de *bit*, taxa de erro e prioridade no tratamento do tráfego, conforme descrito a seguir [3], [24]:

- Classe Conversacional: Essa classe representa os serviços de tempo real que possuem restrições quanto ao atraso, e *jitter*. Aplicações de voz são exemplos para essa classe.
- Classe *Streaming* : Para as aplicações enquadradas nesta classe, uma certa tolerância de atraso e *jitter* é permitida. Como exemplo para esta classe, pode-se citar: vídeo de tempo real, aplicações de fluxos contínuos.
- Classe Interativa: é aplicável a serviços que requerem uma vazão garantida. Alguns exemplos desta classe, incluem: *e-comerce*, navegação Web interativa.
- Classe B*ackground*: É utilizada para o tráfego *best-effort* tradicional, como transferência de arquivos e *e-mail*. Esse tráfego possue a menor prioridade dentre todas as classes.

As Tabelas 3.1 e 3.2 apresentam o mapeamento das classes de QoS UMTS em classes *DiffServ* de PHB e os atributos de QoS para as quatro classes de tráfego UMTS [55].

Tabela 3.1: Mapeamento das Classes UMTS QoS em Classes *DiffServ* de PHB.

Classes de Tráfego UMTS	***DiffServ* com PHB**
Conversacional	EF
Streaming	AF1
Interativa	AF2
Background	AF3 ou *Best Effort*

O processo de marcar pacotes com DSCPs apropriados nos roteadores de borda também é chamado "classificação de pacote" Os roteadores intermediários da rede *backbone* UMTS

usam este DSCPs para o escalonamento e o encaminhamento apropriado dos datagramas IP [5], [34].

Tabela 3.2: Atributos de QoS para as Classes UMTS [55].

Classes de Tráfego	**Conversacional**	***Streaming***	**Interativa**	***Background***
Taxa máxima de *bits* (bps)	$\leq$ 16000	$\leq$ 16000	$\leq$ 16000	$\leq$ 16000
Tamanho máximo SDU (*byte*)	$\leq$ 1500 ou 1502	$\leq$ 1500 ou 1502	$\leq$ 1500 ou 1502	$\leq$ 1500 ou 1502
BER residual	5×10^{-2}, 10^{-2} ,5×10^{-3}, 10^{-3} ,10^{-4}, 10^{-5} ,10^{-6}	5×10^{-2}, 10^{-2} ,5×10^{-3}, 10^{-3} ,10^{-4}, 10^{-5} ,10^{-6}	4×10^{-3}, 10^{-5} ,6×10^{-8}	4×10^{-3}, 10^{-5} ,6×10^{-8}
Taxa de erro SDU	10^{-2}, 7×10^{-3}, 10^{-3} ,10^{-4}, 10^{-5}	10^{-1}, 10^{-2}, 7×10^{-3}, 10^{-3} ,10^{-4}, 10^{-5}	10^{-3}, 10^{-4} ,10^{-8}	10^{-3}, 10^{-4} ,10^{-8}
Atraso de transmissão (ms)	80 (max)	300 (Max)	-	-
Taxa garantida de *bit* (Kbps)	$\leq$ 16000	$\leq$ 16000	-	-
Prioridade na alocação / retenção	1,2,3	1,2,3	1,2,3	1,2,3

3.4.6 Parâmetros da Arquitetura de QoS para o UMTS

As redes UMTS são suportadas por uma arquitetura de QoS que é uma evolução da arquitetura de QoS GPRS. É incluído em todos os nós um dispositivo responsável pelo controle de admissão a fim de determinar se os recursos requisitados pelos *bearer services* estão disponíveis e reservá-los em caso positivo. A Figura 3.12 ilustra os componentes de QoS para UMTS. O gerente de recurso (GR) é responsável por gerenciar o acesso aos recursos de acordo com o *bearer service* particular [53]. A coordenação da sinalização do plano de controle e a tradução dos atributos de QoS nos atributos suportados por um determinado *bearer service*, são feitas pelo *bearer service manager*. O policiamento do tráfego é feito pelo condicionador de tráfego (CT) de forma que este continue com o comportamento que foi especificado e respeite o acordo na manutenção do contrato estabelecido na fase de admissão e alocação de recursos. O classificador de pacotes (CP)

classifica os pacotes de acordo com as classes de tráfego com base no campo DSCP do *DiffServ* [57], [58].

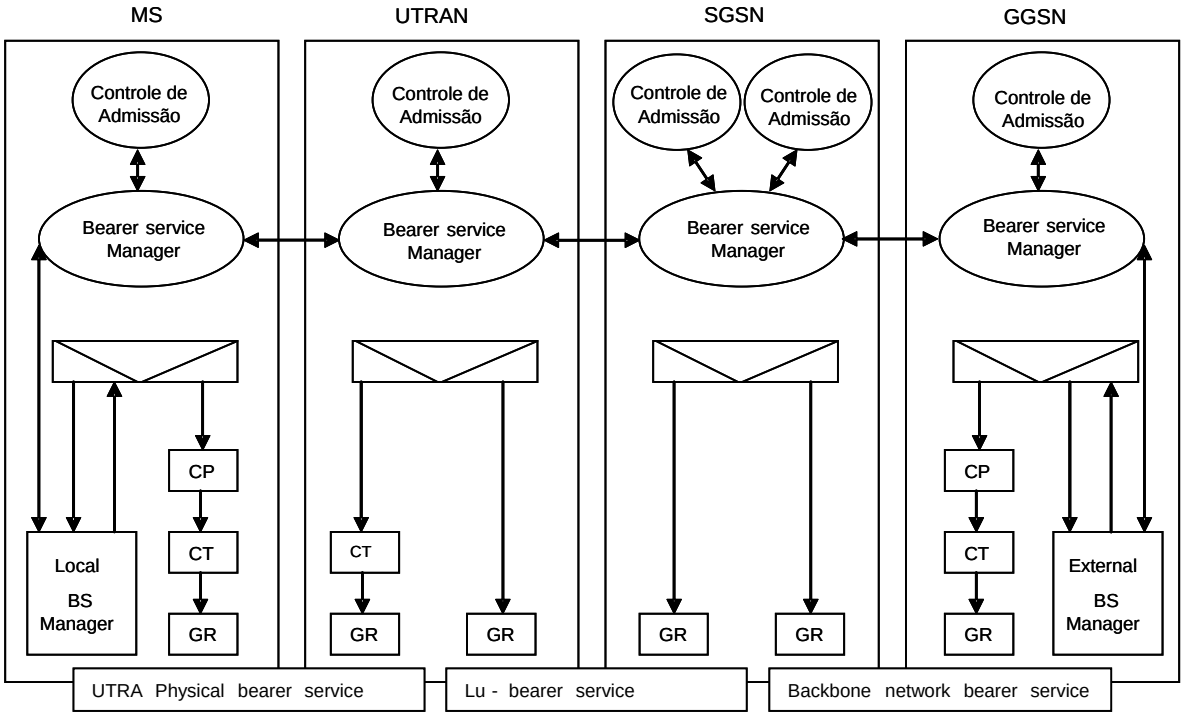

Figura 3.12: Componentes da arquitetura de QoS para UMTS [69].

3.5 Conclusão

A tecnologia para comutação de pacotes nas redes móveis sem fio, tal como o GPRS, inclui recursos de comutação de pacotes ao GSM que abrem as portas a novos serviços baseados em redes IP e outras aplicações de alta velocidade. Entretanto, o GPRS também inclui novos protocolos e complexidade à rede. Compreender a tecnologia e as mudanças que ela traz será sem dúvida o ponto chave para uma implementação bem sucedida do GPRS e a realização plena dos benefícios que ela proporcionará às redes móveis sem fio.

Redes GPRS otimizam a utilização de recursos escassos e fornecem tempos de acessos mais rápidos. Na interface aérea de uma rede GPRS, os recursos de rádio e da rede são

alocados temporariamente quando usuários móveis estiverem enviando ou recebendo dados, ao contrário, por exemplo, de uma rede GSM comutada a circuito onde os recursos são alocados para um usuário por toda a duração de uma conexão. Com isso, a tarifação em uma rede GPRS pode ser baseada no volume de dados transferidos, proporcionando serviços com custos relativamente mais baixos.

No caminho da evolução das comunicações móveis sem fio, os sistemas móveis de terceira e quarta geração (3G/4G), incluindo as redes *backbone* UMTS, estão adotando cada vez mais o protocolo IP. O cenário atual das redes móveis sem fio, sinaliza que seja possível aplicar tecnologia IP fim a fim também no ambiente móvel sem fio, produzindo uma rede *All-IP*. O uso de pacotes de dados IP, tanto para o transporte quanto para a sinalização, traça o caminho para uma convergência natural entre redes fixas e móvel. Proporcionar um nível de QoS satisfatório fim a fim é um dos maiores desafios no projeto de sistemas *All-IP*. Serão necessárias uma interconexão plena das redes *backbone* UMTS e as redes IP de próxima geração.

Conforme apresentado neste Capítulo, o 3GPP definiu uma arquitetura para o controle de QoS fim a fim na rede *backbone* UMTS, bem como as classes de tráfego para representarem as aplicações oferecidas pela rede UMTS, e por fim foi demonstrado a necessidade que existe em se prover QoS à rede *backbone* UMTS.

Neste contexto, será abordado no próximo Capítulo o tema Qualidade de Serviço nas redes móveis sem fio, discorrendo sobre seu significado, quais são as exigências dos diversos tipos de aplicações e os vários mecanismos existentes para a implementação de QoS.

Capítulo 4

Qualidade de Serviço em Redes Móveis sem Fio

4.1 Introdução

Idealmente, as redes móveis sem fio deveriam ser capazes de garantir a entrega de aplicações sensíveis a atrasos, perdas de pacotes e *jitter* com Qualidade de Serviço (QoS) requisitada pelos usuários. Mas a crescente demanda por aplicações multimídia em redes móveis sem fio traz consigo desafios próprios que tornaram a provisão de QoS uma tarefa mais complexa. Tal complexidade é resultado, dentre outros fatores, das limitações do ambiente e dos parâmetros envolvidos, tais como: a qualidade da transmissão na interface aérea; restrição quanto à fonte de energia das MSs, pois modelos compactos utilizam baterias de energia limitada e, portanto, a utilização controlada de energia é um parâmetro importante; os recursos escassos de largura de banda; a grande variabilidade da qualidade do enlace devido ao desvanecimento e a própria mobilidade das MSs (*roaming*), que implicará na necessidade do gerenciamento de suas localizações.

No âmbito de redes móveis sem fio baseadas em IP, um dos objetivos para alcançar um certo nível de QoS, é criar uma solução de rede IP multiserviço onde o tráfego com menor

prioridade possa compartilhar a mesma infra-estrutura que os tráfegos que possuam requisitos mais rigorosos de Qualidade de Serviço [62].

Aplicações de tempo real como, por exemplo, os serviços de voz, são sensíveis a atraso. Este tipo de serviço impõe limites rígidos para o tempo de transferência de pacotes fim a fim na rede e a variação no atraso entre os pacotes consecutivos deve ser compensada. Outros tipos de aplicações, tais como aplicações de transferência de arquivos, caracterizadas por fluxos intensos de dados intercalados por períodos de inatividade, não são sensíveis a atrasos embora exijam uma recepção confiável dos dados transmitidos.

Com a evolução das redes 3G em direção aos sistemas 4G, as redes móveis sem fio passaram a utilizar plenamente o protocolo IP. Assim, tornou-se necessário dar maior ênfase à Qualidade de Serviço, pois a principal aplicação oferecida por essas redes é a transmissão de voz, que é uma aplicação de tempo real.

4.2. Parâmetros que Influenciam na QoS

As redes 3G/4G têm como um dos seus objetivos a transmissão de aplicações multimídia. Desta forma, um grande desafio a ser suplantado por essas redes com relação à transmissão de aplicações em tempo real, como o tráfego de voz, é o provisionamento de QoS [19]. Isto porque, a mobilidade que deverá ser garantida às MSs dificultará um gerenciamento de recursos adequado na rede, dificultando a oferta de aplicações com QoS ao usuário. Isto justifica a necessidade de mecanismos eficientes de controle de tráfego ainda maior.

O atraso, a perda de pacotes e a capacidade de transmissão da rede são parâmetros importantes para as aplicações de tempo real, pois, embora elas possam tolerar pequenas perdas de pacotes, em contrapartida impõem restrições severas de temporização ou de capacidade mínima de transmissão para garantir sua entrega ao usuário final.

Um conjunto de parâmetros define o nível de QoS desejável que caracteriza um serviço de comunicação. Esses parâmetros incluem alocação de banda, limite de atraso fim a fim permitido, variação do atraso (jitter) e taxa de perda de pacotes. Os parâmetros de QoS são normalmente utilizados na contratação de um serviço e, por isso, devem ser bem definidos. A Tabela 4.1 ilustra os parâmetros de QoS definidos para algumas aplicações comumente aceitos por um usuário final [26]. Esses parâmetros podem variar de aplicação para aplicação, de acordo com as características desejadas para o serviço oferecido pela rede [62].

Tabela 4.1: Parâmetros de QoS definidos para um conjunto de aplicações.

APLICAÇÃO	**VAZÃO (Kbps)**	**ATRASO**	***JITTER***	**PERDA DE PACOTES**
Acesso Remoto (*Telnet*)	< 1	< 250ms	-	Zero
Áudio *Streaming* (Música)	5 – 128	< 10s	< 2s	< 1%
Correio Eletrônico (*e-mail*)	-	*Best Effort*	-	Zero
Mensagem de Voz	4-13	< 1s	< 1ms	< 3%
Transferência de Arquivo (FTP)	< 384	< 10s	-	Zero
Vídeo *One-way*	20-384	< 10s	< 2s	< 2%

- **Vazão:** o parâmetro vazão, ou a taxa de transmissão, é o número de bytes por segundo recebidos pelo receptor em uma determinada conexão. É uma grandeza finita, normalmente limitada pelas características físicas do meio de comunicação, capacidade do equipamento ou limitações impostas por acordo comercial. A vazão de uma determinada aplicação está diretamente ligada ao nível de congestionamento na rede. A Tabela 4.2 apresenta a vazão requerida por diversas aplicações. Este fator é devido ao aumento do atraso tendo como conseqüência um maior atraso da chegada dos pacotes ao seu destino, e devido também às perdas de pacotes que aumentam quando a rede se encontra em situações de congestionamento, fazendo com que o número de pacotes que chegam no destino seja reduzido, diminuindo conseqüentemente a vazão [34].

Tabela 4.2: Vazão requerida por diversas aplicações.

APLICAÇÕES	VAZÃO
Aplicações Transacionais	1 Kbps a 50 Kbps
Quadro Branco (*Whiteboard*)	10 Kbps a 100 Kbps
Voz	10 Kbps a 120 Kbps
Aplicações Web (www)	10 Kbps a 500 Kbps
Transferência de Arquivos (FTP)	10 Kbps a 1 Mbps
Vídeo (*Streaming*)	100 Kbps a 1 Mbps
Aplicação Conferência	500 Kbps a 1 Mbps
Vídeo MPEG	1 Mbps a 10 Mbps
Aplicação Imagens Médicas	10 Mbps a 100 Mbps
Aplicação Realidade Virtual	80 Mbps a 150 Mbps

- **Atraso:** o parâmetro atraso corresponde ao tempo gasto desde a geração de um pacote na origem até seu recebimento pelo usuário final. Ele depende, basicamente, do meio de transmissão, da distância, do tempo de processamento nos nós e, principalmente, do nível de ocupação da rede. O atraso é causado por fatores fixos e variáveis. Os fatores fixos são aqueles que, independentemente da situação da rede, não sofrem qualquer variação. Estes fatores são normalmente aqueles inerentes ao meio de comunicação, como o atraso de propagação, mas também podem estar associados aos mecanismos de codificação e decodificação de sinais. Os fatores variáveis dependem da situação atual da rede, e normalmente são causados por situações de congestionamento. Os principais tipos de atraso em uma transmissão de dados são:

    1. Atraso de processamento**:** é o tempo necessário para processar o cabeçalho do pacote e determinar o seu destino final.
    2. Atraso de enfileiramento: é função do tempo de armazenamento de um pacote em uma fila antes do seu efetivo encaminhamento devido aos recursos finitos da rede. Este atraso é o principal responsável pela aleatoriedade no atraso do pacote, podendo assumir valores críticos em situações de congestionamentos na rede.

3. Atraso de transmissão: é o tempo necessário para a transmissão de todos os *bits* do pacote para o enlace.
4. Atraso de propagação: é o tempo necessário para que o pacote se propague desde o início do enlace até o próximo nó da rede.

- ***Jitter*:** é uma distorção ocorrida nos tempos de chegada entre pacotes comparados aos tempos originais de transmissão entre pacotes. É uma distorção que acontece, por exemplo, quando um determinado fluxo é transmitido em uma rede e os pacotes não chegam no seu destino dentro da ordem sucessiva ou em uma determinada cadência. Ou seja, eles variam em termos de tempo de atraso [47].
- **Perda de pacotes:** a rede IP não fornece um serviço de entrega de pacotes confiável, ou seja, o protocolo IP não verifica se um pacote foi recebido corretamente, deixando essa responsabilidade para os protocolos das camadas superiores. É importante destacar que diversos fatores podem ocasionar perdas, incluindo erros no meio físico de transmissão, *buffer overflow*, colisões no caso de protocolos de acesso aleatório e pacotes descartados devido a um atraso excessivo (no caso de aplicações multimídia com restrição de atraso).

4.3. Mecanismos para Provimento de QoS na *Internet*

Pode-se afirmar que o objetivo da introdução de QoS nas atuais redes IP é fornecer algum nível de gerenciamento dos parâmetros de qualidade e controle, além do serviço de melhor esforço da *Internet*, pois somente assim será possível atender às crescentes demandas das aplicações atuais e futuras [62]. O maior desafio da *Internet* não é somente o aumento da velocidade de processamento nos nós comutadores, mas sim suportar diversas aplicações, garantindo certo nível de QoS para o correto funcionamento mesmo em condições de congestionamento. Portanto, pacotes IP devem se comportar de uma forma aceitável para que

características funcionais da aplicação sejam mantidas, independentes de congestionamento. Pode-se citar como exemplo, o parâmetro de variação de atraso de pacotes (*jitter*) em uma aplicação de voz sobre IP. Uma variação de atraso superior aos limites admissíveis poderá deixar uma aplicação de voz ininteligível, apesar de, até mesmo, existir uma taxa nominal de transmissão adequada para a aplicação.

O tema de QoS na *Internet* tem sido objeto de intensas pesquisas nos últimos anos e várias abordagens foram propostas nesse sentido, sempre tendo em mente que o sucesso da *Internet* se explica, em grande parte, pela simplicidade dos protocolos que funcionam sobre a rede, portanto isto sinaliza a sua continuidade [5] [32].

Neste sentido, a IETF [27] possui alguns grupos de trabalho que objetivam estender a arquitetura atual da *Internet* para prover diferentes níveis de qualidade de serviço para os usuários, e dentre eles destacam-se os grupos de Serviços Integrados [28], MPLS (*MultiProtocol Label Switching*) [29], Serviços Diferenciados [30], RSVP [31], Engenharia de Tráfego e Roteamento Baseado em Restrição [32]. Portanto, serão apresentados a seguir os princípios básicos destas arquiteturas.

4.3.1. Arquitetura de Serviços Integrados *IntServ*

A arquitetura de Serviços Integrados (*IntServ*) descreve as primeiras funções de provisão de QoS que devem estar presentes em todas as estações e roteadores, em especial o controle de admissão, a alocação de recursos e a classificação e escalonamento de pacotes. Foi a primeira das soluções propostas para dar suporte a QoS na *Internet*, no âmbito da IETF [33]. Tinha como objetivo inicial atender certas garantias exigidas por determinados tipos de tráfego, como transmissão de áudio e vídeo. A arquitetura *IntServ* visa fornecer, em uma rede comutada por pacotes, como a *Internet*, um serviço cuja QoS é determinada por uma das

classes de Serviço Garantido (GS-*Guaranteed Service*), e Serviço de Carga Controlada (CL-*Controlled Load Service*), [34].

A idéia principal do modelo *IntServ* é a de reserva de recursos. Antes de iniciar a transmissão dos dados, as aplicações precisam encontrar um caminho até o receptor que satisfaça suas demandas de QoS, reservando, ao longo do mesmo, os recursos necessários. Objetivando alcançar uma determinada QoS, as aplicações fazem uso do protocolo RSVP, que é um protocolo de controle e sinalização que atua na camada de rede, sendo responsável por reservar caminhos e recursos na sub-rede de comunicação, conforme será apresentado na seção 4.4.3.

4.3.1.1. Classes de Serviço

A arquitetura de Serviços Integrados define duas classes de serviço, além do modelo de melhor esforço, tradicionalmente encontrado nas redes IP. São elas:

- **Serviço Garantido** (GS)

  Serviço Garantido, especificado em [28], fornece um limite superior rígido para o atraso fim a fim, além de garantir a disponibilidade de largura de banda. Este serviço se destina a aplicações que possuem requisitos estritos de tempo real para funcionar, como mídias contínuas que não toleram atrasos, *jitter* e perdas, que precisam de uma garantia firme de que um pacote não irá chegar no destino depois de um tempo maior que um limite especificado. Esta classe de serviço apresenta um alto nível de QoS na *Internet* [35].

- **Serviço de Carga Controlada** (CL)

  O Serviço de Carga Controlada, especificado em [36], não oferece garantias quantitativas rígidas, como o Serviço Garantido, porém fornece um serviço equivalente ao modelo de melhor esforço em uma rede pouco utilizada com quase

nenhuma perda ou atraso. Em situações de sobrecarga, esta abordagem será capaz de compartilhar a largura de banda entre múltiplos fluxos, de uma maneira controlada, garantindo um serviço melhor que o usual. Entretanto, este modelo não oferece garantias de atraso máximo, apenas um limiar probabilístico, assim como também não pode assegurar que pacotes não serão perdidos.

4.3.2. Serviços Diferenciados (*DiffServ)*

Em face das limitações encontradas no modelo *IntServ*, foi proposta no âmbito da IETF a Arquitetura de Serviços Diferenciados [37] que oferece *QoS* na *Internet* com escalabilidade, sem estado para cada fluxo e sinalização a cada roteador, ou seja, a cada nó.

Ao passo que o *IntServ* realiza as reservas de recursos em nível de fluxo, exigindo dos roteadores a manutenção de informações de estado para cada fluxo, fim a fim, a arquitetura *DiffServ* baseia-se na idéia de agregação de fluxos em umas poucas classes de serviço. Dessa forma, o *DiffServ* fornece diferenciação de serviços local para grandes agregados de tráfego, enquanto a arquitetura *IntServ* dá garantias de desempenho fim a fim para fluxos individuais. [38]

O modelo de Serviços Diferenciados usa a classificação de pacotes como mecanismo para o provisionamento de QoS. Por esta razão, é redefinido o *layout* do *octeto* ToS (*Type-of-Service)* do cabeçalho do protocolo IPv4 (ou o campo *Traffic Class* do IPv6), que passa a ser chamado de campo DS (*Differentiated Services*) [39]. Até o momento, somente os seis primeiros *bits* do campo DS são usados, recebendo a denominação de campo DSCP (*Differentiated Services Codepoint*), como mostrado na Figura 4.1. Seu objetivo é especificar o tratamento dado ao encaminhamento de pacotes em cada roteador, o chamado PHB (*Per-hop Behavior*). Os PHBs são mecanismos de priorização que permitem a agregação de fluxos gerados por diferentes aplicações, definindo uma classe de serviço [34].

Como mostrado na Figura 4.1, os *bits* 0 a 2 são usados para precedência. Estes *bits* podem assumir um dos valores especificados na Tabela 4.3. O *bit* 3 (D) é usado para indicar atraso. Quando D = 0, indica atraso normal; D = 1 minimiza atraso. O *bit* 4 (T) é usado para *throughput*. Quando T = 0 indica vazão normal; T = 1 maximiza vazão. O *bit* 5 (R) é usado para indicar a confiabilidade dos pacotes recebidos. Quando R = 0 indica confiabilidade normal; R = 1 indica alta confiabilidade. Os *bits* 6 e 7 são reservados para uso futuro.

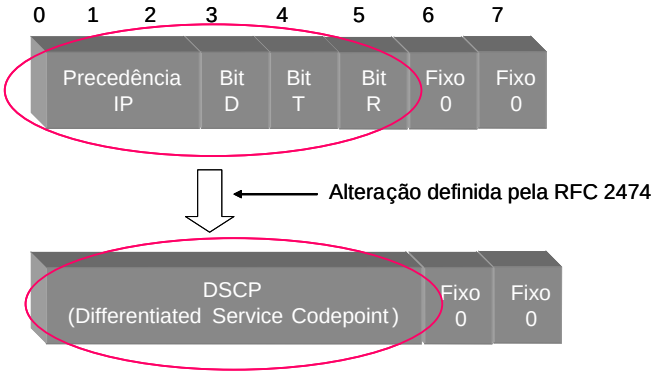

Figura 4.1: Layout do Campo DS.

Portanto, a arquitetura *DiffServ* funciona através da marcação de pacotes de forma distinta, o que permite que os mesmos sejam tratados internamente à rede de maneira diferenciada, segundo a classe de serviço à qual pertencem. A abordagem empregada pelo *DiffServ* baseia-se em um esquema de prioridades relativas, ou seja, ele garante que o tráfego gerado por uma aplicação, com um certo nível de prioridade, receberá um tratamento melhor que o gerado por qualquer outra que possua uma prioridade inferior [40].

Tabela 4.3: Precedência IP.

BITS	**PRECEDÊNCIA**
111	Network Control / Controle e gerenciamento de Rede
110	*Internetwork* Control / Controle e gerenciamento de Inter-Rede
101	*Critical*
100	*Flash Override*
011	Flash
110	Imediato
001	Prioridade
000	Comum

A marcação dos pacotes se dá nos pontos de ingresso na rede, como os *hosts* finais e os roteadores de borda. Dessa forma, retém-se um dos princípios básicos do projeto da *Internet*, que é colocar a complexidade na fronteira da rede, ao contrário da abordagem *IntServ*, que exige complexidade fim a fim. O modelo *DiffServ* não necessita de um protocolo próprio (como no caso do RSVP, usado no *IntServ*), pois aqui se utiliza um campo do próprio datagrama IP. Tudo que um roteador precisa fazer é examinar o campo DSCP de cada pacote para determinar qual o tratamento a ser dado ao mesmo. Assim, pacotes marcados da mesma forma recebem tratamento igual. Não é necessário manter nenhum tipo de informação relacionada a fluxos, pois os roteadores só precisam ser capazes de distinguir entre um certo número de classes de serviço pré-definidas.

4.3.2.1. Classes de Serviço

Embora outras alternativas sejam possíveis, existem, atualmente padronizados pelo IETF, duas classes de serviço principais definidas no modelo *DiffServ*, que são:

- Encaminhamento Expresso EF (*Expedited Forwarding*);
- Encaminhamento Assegurado AF (*Assured Forwarding*).

Encaminhamento Expresso

A classe de serviço chamada de Encaminhamento Expresso permite a adaptação do modelo de serviço garantido da arquitetura *IntServ* à arquitetura de Serviços Diferenciados. Ele oferece garantias de QoS absoluta, com baixos valores de perda, atraso e *jitter*, fornecendo o equivalente a uma linha privada virtual com largura de banda fixa entre dois *hosts*. É indicado para aplicações de tempo real como voz e vídeo, e para a criação de linhas dedicadas em redes privadas virtuais VPNs (*Virtual Private Network*) [41].

A principal vantagem sobre o serviço equivalente na arquitetura *IntServ* está na simplicidade de implementação, pois não é necessário manter nos roteadores nenhuma

informação relativa a fluxos. Os equipamentos que implementam este comportamento PHB (*Per Hop Behavior*) simplesmente devem escalonar o tráfego de maneira a manter descongestionadas as filas de saída, a fim de que o tráfego passe o menor tempo possível nos nós da rede. Em geral, os pacotes pertencentes a esta classe são colocados em uma fila de maior prioridade que a do tráfego de melhor esforço e são os primeiros a serem encaminhados em qualquer situação. Em relação à política de descarte, este serviço evita a todo custo descartar os pacotes para o tráfego em conformidade com o perfil contratado. Já para o tráfego sem conformidade, ele é implacável, os pacotes simplesmente são descartados. Uma conseqüência disso é que os usuários não podem exceder a taxa de pico solicitada, caso contrário, o tráfego em excesso será descartado. Em contrapartida, o modelo *DiffServ* garante que a largura de banda contratada estará disponível quando o tráfego for enviado [41].

Encaminhamento Assegurado

A classe de serviço AF destina-se a aplicações que demandem da rede um serviço mais confiável que aquele de melhor esforço, mas sem todas as garantias de QoS dadas pelo EF. Este serviço não oferece limites superiores para o atraso e *jitter*, mas garante um tratamento preferencial ao tráfego que dele se utilize. Ele é indicado quando se deseja obter da rede um serviço de entrega de pacotes mais consistente, por exemplo, a fim de oferecer uma melhor QoS a agregados de tráfego consistindo de rajadas de curta duração com destinos diferentes [42].

O princípio aqui é a divisão do tráfego em N classes, cada uma com alguns níveis de precedência de descarte (M). A especificação atual define $N = 4$ e $M = 3$ (apresentada na Figura 4.2), embora, em uma situação real, nem todas elas possam vir a ser necessária. O serviço fornecido por uma certa classe independe do serviço das demais classes, sendo função apenas dos recursos alocados para cada classe pelo sistema.

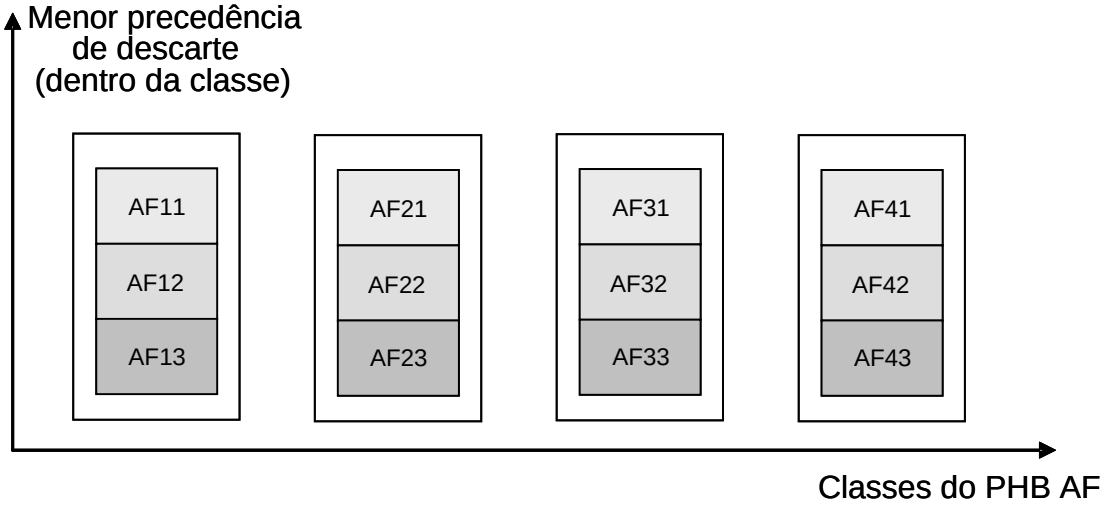

Figura 4.2: Classes do Serviço de Encaminhamento Assegurado.

Um usuário pode contratar de um provedor um dos quatro serviços de encaminhamento diferenciado, cada um com três níveis de prioridade de descarte. Em situações de sobrecarga, o tráfego de uma classe superior tem menor probabilidade de sofrer congestionamento que o tráfego de uma classe inferior. O mecanismo de diferenciação aqui utilizado se baseia na prioridade de descarte quando este for inevitável, primeiro descartam-se os pacotes pertencentes ao serviço de melhor esforço e, só então, passa-se para os pacotes associados ao serviço de encaminhamento assegurado, segundo sua classe e nível de precedência na classe. Portanto, os pacotes pertencentes ao serviço AF são os últimos a serem descartados em situações de congestionamento.

É possível estabelecer uma relação entre os serviços das arquiteturas de serviços integrados e diferenciados. Em geral, pode-se mapear o Serviço Garantido (*IntServ*) para o Serviço de Encaminhamento Expresso (*DiffServ*) e o Serviço de Carga Controlada (*IntServ*) para o Serviço de Encaminhamento Assegurado (*DiffServ*). Esta possibilidade de mapeamento é especialmente interessante em situações em que se utilizam as duas arquiteturas em conjunto para a provisão de uma QoS fim a fim.

4.3.3. O Protocolo RSVP

O RSVP (*Resource Reservation Protocol*) [31] é um protocolo de sinalização desenvolvido para realizar reserva de recursos ao longo de uma rede que utiliza a arquitetura *IntServ*. O RSVP é utilizado por sistemas finais para requisitar à rede níveis específicos de QoS para as aplicações [25]. As duas mensagens mais importantes do protocolo RSVP são *PATH*, que é originado no transmissor e RESV, que é originado no receptor.

A Figura 4.3 ilustra o funcionamento do protocolo RSVP: o emissor envia uma mensagem do tipo *PATH* (setas vermelhas) a um ou mais receptores informando as características do tráfego a ser gerado. À medida que a mensagem *PATH* se propaga pela rede, cada roteador ao longo do enlace grava informações sobre o caminho na mensagem. O receptor, ao receber a mensagem *PATH,* responde com uma mensagem do tipo RESV (setas verdes), requisitando recursos ao longo do caminho gravado na mensagem *PATH,* em ordem inversa, do receptor para o emissor. Quando um roteador intermediário recebe a mensagem RESV, ele verifica se pode atender a solicitação de recursos. Caso seja possível, ele faz a reserva de largura de banda e espaço em *buffers* e passa a mensagem RESV adiante, caso contrário, envia uma mensagem de erro ao receptor. Se o emissor recebe a mensagem RESV, isto indica que a solicitação de reserva de recursos foi bem sucedida e é possível iniciar a transmissão dos pacotes.[34]

Através da utilização do protocolo RSVP, é possível conseguir um alto nível de *QoS* na *Internet*, pois ele permite fazer o gerenciamento com menor granularidade, em nível de fluxo, conseguindo dar excelentes garantias de qualidade às aplicações. Entretanto, há um preço a se pagar por isso: o RSVP possui sérios problemas de gerenciamento e escalabilidade. Cada roteador ao longo do caminho precisa dar suporte à RSVP para que se possa assegurar a QoS, sendo necessário manter as informações de estado e fazer o escalonamento e enfileiramento dos pacotes para cada fluxo.

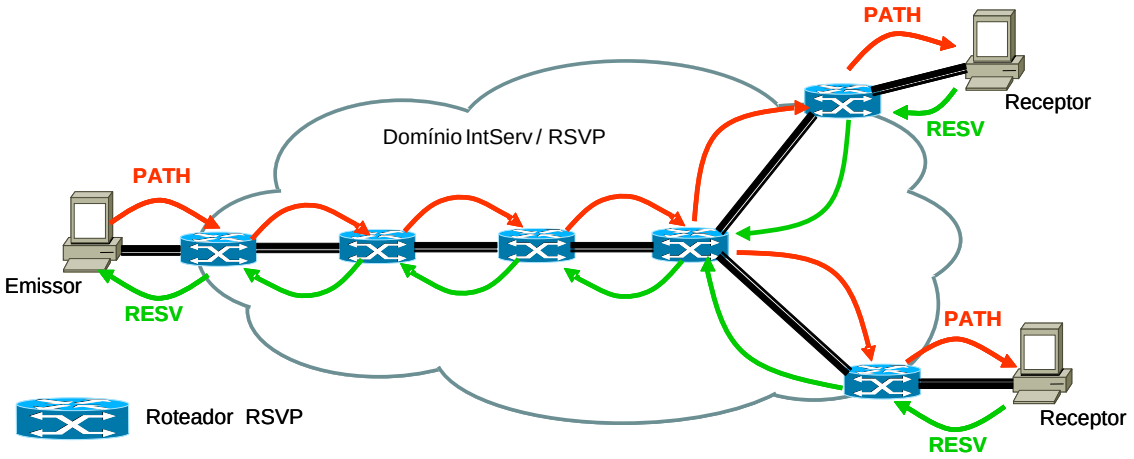

Figura 4.3: Funcionamento do Protocolo RSVP.

4.3.4. Service Level Agreements (SLA)

A fim de que um determinado usuário possa receber serviços diferenciados de um provedor, os domínios DS negociam entre si "acordos de serviço" denominado SLAs (*Service Level Agreements*) que visam o oferecimento de garantias mínimas de QoS para as aplicações dos usuários. Em outras palavras, define as classes de serviço contratadas e a quantidade de tráfego permitida em cada classe. Todos os pacotes que fluem de um domínio para outro são fiscalizados (policiados) nos roteadores de borda para verificar sua conformidade com os contratos. No centro da rede, os roteadores simplesmente encaminham os pacotes para os seus destinos, oferecendo algumas garantias de QoS a determinados pacotes.

Ao ser enviado um tráfego, o domínio de origem tem a responsabilidade de policiá-lo e suavizá-lo nos pontos de saída (*egress points*), pois um tráfego fora do perfil contratado não receberá nenhum tipo de garantia de QoS ao chegar ao próximo ponto de ingresso (*ingress point*), em outro domínio. Quando um pacote deixa um domínio e segue para outro, às vezes pode ser necessário remarcar o seu campo DS, como resultado de um SLA estabelecido entre

os dois domínios, embora o ideal seja que o tráfego experimente o mesmo nível de QoS ao longo de todo o percurso, da origem até o destino, o que nem sempre é possível.

Os critérios empregados para a aplicação das políticas de QoS podem ser data e hora, endereços de origem e destino, números de portas ou qualquer outra informação que possa ser extraída do conteúdo do tráfego, inclusive a contida nos cabeçalhos. Além desses aspectos, um SLA também pode especificar procedimentos de tarifação e cobrança, serviços de criptografia e autenticação, procedimentos de renegociação dos parâmetros do SLA, ações a serem tomadas para o tráfego fora de perfil, entre outros.

O oferecimento de um serviço fim a fim é realizado através da concatenação de vários domínios DS, onde os SLAs são negociados em cada uma das bordas entre os domínios existentes. O modelo lógico *DiffServ*, com os vários domínios e SLAs nas bordas é mostrada na Figura 4.4. Um domínio usuário de um serviço não estabelece um SLA direto com o domínio final do serviço, a não ser que haja uma ligação direta entre eles. Caso contrário, ele negocia com o próximo domínio DS no caminho e assim por diante até o domínio final.

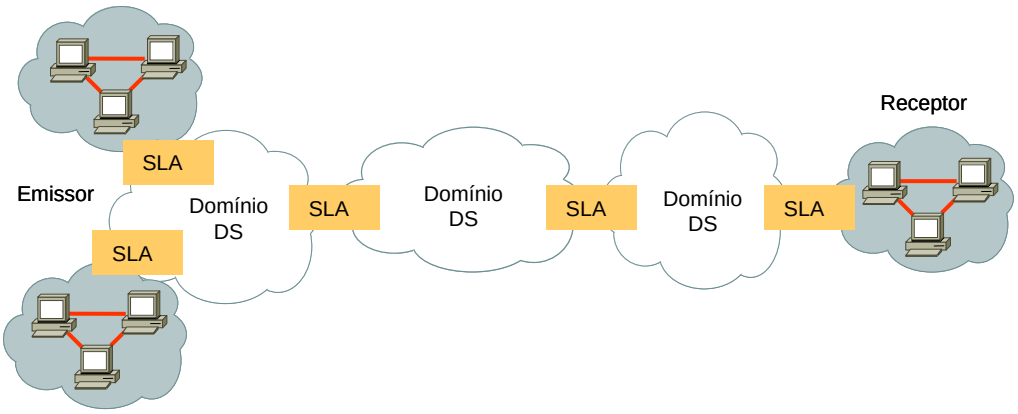

Figura 4.4: Modelo lógico *DiffServ*.

A especificação de condicionamento de tráfego é uma parte importante do SLA, que define entre outras coisas alguns parâmetros detalhados do desempenho do serviço, como níveis esperados de vazão, atraso e perda de pacotes. Além disso, define perfis de tráfego, através de

parâmetros de balde de fichas, que definem as características do tráfego como também as ações que podem ser tomadas caso o usuário não cumpra as especificações.

4.3.5. Comutação Multi Protocolo (MPLS)

Em redes IP, quando um roteador recebe um pacote, ele faz uma busca na sua tabela de roteamento e então, baseado no endereço IP do pacote, decide para onde enviá-lo. Essa procura pode levar bastante tempo, dependendo do tamanho da tabela de cada roteador. O modelo MPLS [29] propõe uma solução para esse problema, usando um rótulo de tamanho fixo a partir do qual o roteador decide por onde enviar os pacotes. Assim como o *DiffServ*, o MPLS inclui um rótulo nos pacotes nos pontos de ingresso na rede e os remove nos pontos de saída, contudo, ao contrário do *DiffServ*, em que a inclusão do rótulo é apenas uma forma de atribuir prioridade aos pacotes dentro dos roteadores, o MPLS a utiliza efetivamente para fazer o roteamento, determinando qual o próximo passo (*hop)* a ser dado conforme o rótulo colocado no pacote. O protocolo MPLS funciona integralmente nos roteadores, sem nenhum componente nos sistemas finais.[43].

Um rótulo identifica o caminho que o pacote poderá percorrer na rede, permitindo que sejam estabelecidos canais com largura de banda fixa, nos moldes dos circuitos virtuais de uma rede ATM ou *Frame Relay*. Como este protocolo opera entre as camadas de rede e enlace, MPLS é neutra quanto à tecnologia de rede, podendo ser implementada sobre ATM, DWDM, *Frame Relay, Ethernet* e outros. [44]

A Figura 4.5 ilustra a operação básica da rede MPLS. Neste exemplo, o *host* **A** emissor envia um pacote para o *host* **B** receptor.

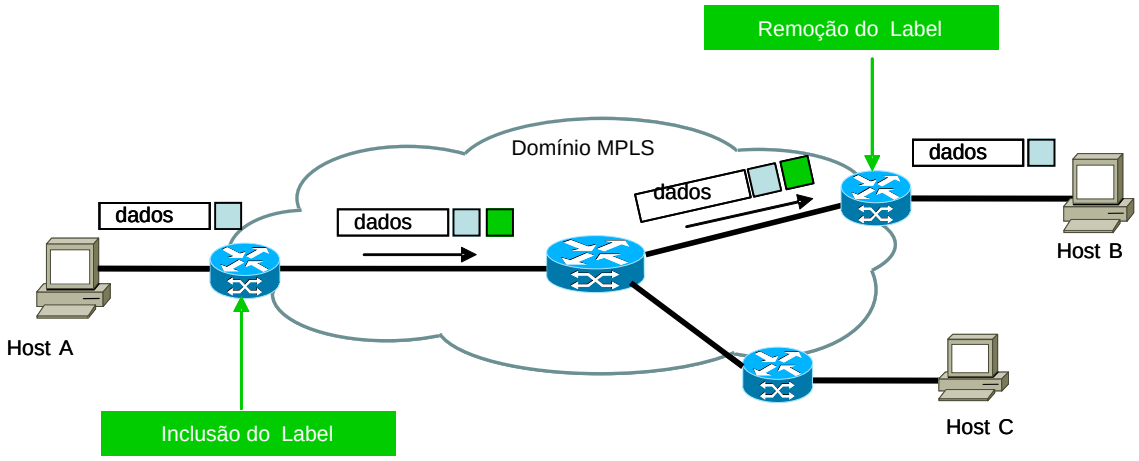

Figura 4.5: Encaminhamento de pacote em um domínio MPLS.

O formato do cabeçalho MPLS está ilustrado na Figura 4.6, o qual possui um rótulo de 20 *bits*. É este rótulo que será processado pelos roteadores habilitados para MPLS - LSRs, funcionando como um índice para uma tabela que especifica o próximo *hop* a ser dado e um novo rótulo para o pacote. Há também um campo de três *bits* reservado para fins experimentais, um *bit* usado para permitir o alinhamento de rotas MPLS (*Stack Flag*) e um campo de 8 *bits* para estabelecer um tempo de vida para o pacote (*Time-to-Live* - TTL).

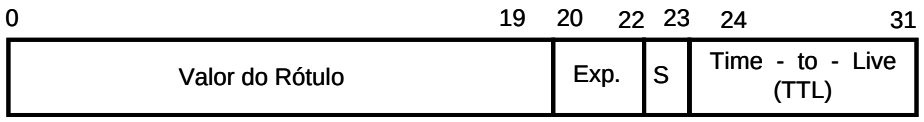

Figura 4.6: Layout do cabeçalho do protocolo MPLS.

Um aspecto importante na implantação do MPLS está na forma de gerenciamento e distribuição dos rótulos entre os roteadores MPLS, a fim de garantir um significado comum para todos eles.

4.3.6. Engenharia de Tráfego

São dois os fatores que podem causar congestionamento na rede, a falta de recursos ou a distribuição irregular do tráfego. Analisando o primeiro caso, todos os roteadores e enlaces estão sobrecarregados. Para o segundo caso, partes da rede estão sobrecarregada enquanto que outras estão subutilizadas. A distribuição desordenada do tráfego pode ser causada pelos protocolos de roteamento dinâmicos que utilizam sempre o caminho mais curto para encaminhar os pacotes. Devem-se utilizar outros protocolos para contornar este caso, que levem em consideração outros fatores. Para pequenas redes, é possível configurar manualmente o custo de cada enlace, de modo que o tráfego possa ser distribuído de maneira uniforme, mas para redes complexas, isto se torna impossível.

A Engenharia de Tráfego é o processo que tenta distribuir os fluxos de tráfegos através da rede, de maneira que este tipo de congestionamento possa ser evitado. Esta tecnologia é direcionada à otimização de desempenho de redes operacionais. Em geral, ela engloba a aplicação de princípios tecnológicos e científicos para medir, modelar, caracterizar e controlar o tráfego na *Internet* e a aplicação dessas técnicas e conhecimentos para atingir determinados objetivos de desempenho [43]. Um dos pontos relevantes da Engenharia de Tráfego na *Internet* é facilitar a operação eficiente e confiável da rede enquanto que ao mesmo tempo otimiza a sua utilização e desempenho. A Engenharia de Tráfego já é atualmente uma função indispensável em grandes redes por causa do custo alto dos equipamentos e da natureza comercial e competitiva da *Internet*. Como altera o fluxo normal dos pacotes, ela pode ser utilizada para atender a requisitos de QoS de determinados fluxos de dados. Os objetivos de desempenho da Engenharia de Tráfego podem ser classificados como:

- Orientados a tráfego: incluem os aspectos relacionados à manutenção das garantias de QoS dos fluxos de dados (ou agregações de fluxos).

- Orientados a recursos: estão relacionados à otimização dos recursos da rede, como impedir que certas partes da rede se tornem congestionadas, enquanto outras permaneçam com recursos ociosos.

Esses objetivos evidentemente não podem ser alcançados por meio do roteamento IP baseado em destino. Obviamente há pouca informação disponível no roteamento IP para possibilitar tal otimização. Na Engenharia de Tráfego, técnicas avançadas de seleção de rotas, geralmente chamadas de roteamento baseado em restrições CBR (*Constraint-Based Routing*), são usadas para se calcular troncos de tráfego baseados nos objetivos de otimização. Portanto, o sistema de Engenharia de Tráfego geralmente precisa de informações sobre a topologia e as demandas de tráfego de toda a rede. Com isso, a Engenharia de Tráfego é tipicamente confinada a um único domínio administrativo.

As fontes CBR produzirão rotas que serão, evidentemente, diferentes daquelas no roteamento baseado em destino. Por esta razão, essas rotas baseadas em restrições não podem ser implementadas pelo encaminhamento baseado em destino. Em modelos anteriores, muitos provedores de serviço usavam ATM nos *backbones* para dar suporte ao CBR. A rede IP era sobreposta aos circuitos da rede ATM, num modelo conhecido como *overlay*. A tecnologia MPLS oferece uma melhor alternativa, uma vez que oferece funções similares e ainda pode ser fortemente integrado às redes IP, e a um preço menor. Entretanto, o MPLS oferece a possibilidade de se automatizar aspectos da função de Engenharia de Tráfego.

Redes de grande porte poderão suportar a Engenharia de Tráfego sobre MPLS, se estas derem suporte aos seguintes atributos:

- Troncos de tráfego: é uma agregação de fluxos de tráfego da mesma classe que são colocadas dentro de um LSP. Essencialmente, um tronco de tráfego é uma representação abstrata do tráfego à qual característica específica pode ser associada.

Isso é útil para se ver os troncos de tráfego como objetos que podem ser roteados, ou seja, o caminho pelo qual o tronco trafega pode ser mudado.

- Atributos de recursos: são atributos associados aos recursos que restringem a colocação dos troncos de tráfego.
- Roteamento baseado em restrições (CBR): é uma das ferramentas utilizadas para automatizar o processo de engenharia de Tráfego. Dessa forma, é uma técnica que auxilia a otimização da operação dos protocolos que habilitam QoS, onde a forma de operação visa, através de mecanismo de policiamento, computar rotas e distribuir cargas de forma a ir ao encontro dos requisitos de QoS de determinado fluxo ou agregado de fluxos.

Parâmetros como: atraso, variação de atraso (*jitter*), custo, contagem de nós e largura de banda, podem ser usados como métrica na computação de rotas. O CBR utiliza um algoritmo chamado NP-*complete,* que usa duas ou mais métricas para computar rotas ótimas de forma a atender a QoS solicitada por determinados fluxos [45].

4.4. QoS em Redes Móveis

Para aplicações típicas em redes móveis sem fio, os parâmetros de QoS incluem limites tais como: banda requerida, atraso de pacote, taxa de perda de pacotes e variação de atraso. Por outro lado, a provisão e a manutenção de QoS no ambiente móvel sem fio possuem características únicas. [46], [53]. Diferente das redes fixas, que assumem uma baixa taxa de erro e usuários fixos, existem alguns fatores que dificultam a provisão de garantias de QoS para redes móveis sem fio. Atender aos requisitos de desempenho de diferentes aplicações, em termos da disponibilidade da rede, largura de banda, atraso, variação de atraso e taxas de confiabilidade na transmissão de dados, é uma das principais metas de pesquisa na área de QoS.

Várias aplicações são sensíveis a atrasos como, por exemplo, os serviços de voz. Essas aplicações impõem limites rígidos para o tempo de transferência de pacotes do ponto de origem ao ponto de destino da rede e a variação no atraso entre os pacotes consecutivos deve ser compensada. Aplicações, tais como transferência de arquivos, caracterizadas por fluxos intensos de dados intercalados por períodos de inatividade, não são sensíveis a atrasos embora exijam uma recepção confiável dos dados transmitidos.

Um dos problemas das redes móveis sem fio em fornecer um nível adequado de qualidade para a correta operação de diferentes aplicações é gerenciar a escassa largura de banda. Devido às características de mobilidade, restrições de energia e alta variabilidade da qualidade do enlace de comunicação sem fio, a utilização adequada desse recurso de rádio se torna um desafio para a rede, principalmente em condições de congestionamento.

4.4.1. Controle de Admissão de Chamadas (CAC)

Para que se possa oferecer o nível de QoS requerido pelas aplicações em uma rede móvel sem fio, um dos fatores fundamentais é que os recursos disponíveis sejam suficientes para garantir os parâmetros exigidos por uma aplicação durante todo o tempo da conexão. Isso requer um mecanismo capaz de estimar se há recursos, como por exemplo, largura de banda suficiente para que uma nova sessão de usuário possa ser admitida ou não, de acordo com o nível de QoS requisitado. Este mecanismo é o controle de admissão de chamadas (CAC).

Quando a ativação de um contexto PDP é solicitada, o processamento do CAC em uma rede *backbone* UMTS é realizado no SGSN e GGSN. Quando uma estação móvel desejar dar início a uma transferência de dados, ela deve alocar um endereço PDP através do envio de uma mensagem de pedido de ativação de um contexto PDP. Nessa mensagem a unidade móvel requisita um perfil de QoS contendo as classes de serviços: atraso, confiabilidade, precedência de serviço e vazão que melhor atendam as exigências de sua aplicação corrente.

Baseado na comparação do perfil de QoS solicitado e da disponibilidade de recursos e carga de uma célula em particular, um algoritmo de CAC é executado para decidir se uma determinada aplicação pode ou não ser admitida. No controle de admissão para o tráfego *downlink*, o SGSN utiliza os parâmetros de controle de fluxo ajustados pelo BSS, tais como: a quantidade de espaço em *buffer* para um dado tunelamento (GTP) entre o BSS e o SGSN e a quantidade média de tráfego de dados que o BSS pode transmitir nesse circuito [28].

O GGSN participa do CAC determinando se ele tem capacidade disponível em sua interface com o SGSN e com outras redes de dados externas. Os atributos de vazão médio e máximo, especificados no perfil de QoS, podem ser analisados para verificar se há largura de banda disponível na interface aérea, uma vez que esta representa o ponto de "gargalo" da rede.

Em relação aos recursos de rádio de uma rede sem fio, o objetivo de um controle de admissão é limitar o número de conexões de tal forma que, uma vez que uma conexão é admitida, a probabilidade dela encontrar a rede sob congestionamento seja considerada baixa, provendo a qualidade de serviço exigida. Por essa razão, a probabilidade de bloqueio de uma conexão ativa em uma rede GPRS deve ser utilizada como uma métrica fundamental de QoS para todos os tipos de fontes de tráfego. Após as exigências terem sido aceitas, o contexto PDP pode ser ativado e a rede *backbone* UMTS deve ser capaz de fornecer os recursos requisitados durante o período de atividade de uma conexão. Caso contrário, se a rede não for capaz de suprir as exigências solicitadas, um novo perfil de QoS deverá ser renegociado.

Em pesquisas realizadas atualmente, grande parte dos algoritmos de CAC propostos tem sido aplicada a redes móveis sem fio. Com base nos requisitos de QoS fornecidos pelos usuários móveis, é proposto um esquema de CAC e reserva de recursos adaptativa para garantir a qualidade exigida por diferentes tráfegos multimídia transportados em redes sem fio de alta velocidade [48]. Para isso os autores classificam um tráfego multimídia em duas

classes distintas: classe de tráfego de tempo real e classe de tráfego de tempo não real. Para o tráfego em tempo real, caracterizado pelas aplicações sensíveis a atrasos, a conexão é aceita somente se a largura de banda requisitada puder ser reservada na célula onde o pedido é originado e nas suas células adjacentes. Para o tráfego de tempo não real, caracterizado pelos serviços de melhor esforço, a largura de banda solicitada é alocada apenas na célula corrente.

4.4.2. Alocação de Recursos de Rádio

O BSS executa a alocação de recursos de rádio durante o estabelecimento de um TBF (*Temporary Block Flow*) de acordo com o contrato de QoS negociado entre uma estação móvel e a rede *backbone* UMTS. Com base na classe de vazão, o BSS pode determinar a porção da largura de banda que deve ser alocada, dinamicamente ou estaticamente, para uma determinada aplicação de modo que esta transmita seus pacotes de acordo com a taxa de dados negociada.

Com relação a uma alocação estática (ou fixa), cada canal é atribuído durante o período inteiro de uma conexão. Conseqüentemente, esse tipo de alocação contribui para o desperdício da largura de banda total disponível no sistema, quando o canal fica ocioso, isto é, não ocorre transmissão de dados, durante o intervalo de tempo no qual a porção da largura de banda foi alocada para a conexão. Por esse motivo, a maioria dos sistemas opta por alocar seus canais dinamicamente à medida que os mesmos são necessários.

4.4.3. Policiamento do Tráfego

O policiamento do tráfego em redes *backbone* UMTS é executado durante a transmissão de dados para regular o tráfego e assegurar o perfil de QoS negociado. O policiamento pode ser implementado no SGSN e no GGSN para policiar os pacotes de entrada e manter o tráfego abaixo da taxa máxima admitida. Este processo ocorre após a conexão ser aceita e os recursos serem alocados. Uma técnica bastante utilizada para a função de policiamento e conformação

de tráfego é o *token bucket* (balde de fichas). O uso dessa técnica permite um controle preciso do intervalo de tempo entre uma seqüência de pacotes enviados. O mecanismo *token bucket* é implementado baseado nos principais componentes básicos, comprimento de rajada, taxa média e intervalo de tempo [25].

4.4.4. Escalonamento

As interfaces aéreas em geral, bem como as redes *backbones* UMTS, apresentam problemas que estão relacionados com a alocação de uma quantidade limitada de recursos de rádio para um número cada vez maior de aplicações. Quando há contenção de recursos, é fundamental que os recursos disponíveis sejam alocados e escalonados de forma eficiente e justa. Para isso ser possível é necessário o emprego de sistemas de filas. As filas são utilizadas para armazenar eventos, tal como uma fila de pacotes, ocorridos em tempos randômicos, organizá-los em seqüência e serví-los de acordo com uma disciplina de escalonamento de fila.

Em *backbones* UMTS, uma disciplina de escalonamento de filas é implementada por um escalonador localizado no BSS [61]. No *downlink* o escalonamento pode ocorrer na chegada de um bloco de dados RLC na camada MAC do BSS enquanto no *uplink* o escalonamento ocorre quando chega um pedido de recurso, contido em um bloco de controle RLC/MAC, para dar início a uma transferência de dados. Um cabeçalho MAC é adicionado a esses blocos, formando um bloco de rádio. A camada MAC do BSS é, então, responsável pelo controle dos blocos de rádio pela interface aérea em ambas as direções, com o objetivo de satisfazer os contratos de serviço de todos os assinantes móveis [61], [67].

Baseado no perfil de QoS contido em um contexto PDP associado a um TBF, o escalonador toma a decisão de quais blocos de rádio devem ser servidos. Os blocos de rádio que chegam no BSS podem ser distribuídos em uma ou mais filas distintas de acordo, por exemplo, com sua classe de precedência de serviço. Dentro das filas, esses blocos são escalonados seguindo

uma disciplina de escalonamento de filas. A próxima seção descreverá os conceitos fundamentais relacionados a um sistema de filas e introduzir alguns conceitos sobre disciplinas tradicionais de escalonamento de filas.

4.5. Disciplinas de Escalonamento de Filas

Minimizar o congestionamento na maioria das redes de comunicação é um aspecto fundamental para garantir o nível de qualidade exigido por diferentes fluxos de dados, principalmente àqueles sensíveis a atraso e *jitter*. Isso se faz necessário, pois uma rede congestionada reduz a sua vazão, em termos da quantidade de dados livres de erros que alcançam o destino, aumenta o atraso fim a fim na transmissão dos pacotes e, pode ocasionar perdas de pacotes caso não haja espaço suficiente em uma área de armazenamento (chamada "*buffer*" ou "fila") para manter os pacotes que aguardam para serem transmitidos [25].

De modo a controlar o tráfego total transportado em uma rede, esta pode ser representada como um sistema de filas [34]. Um sistema de filas é composto por pacotes e por um ou mais servidores que representam o meio de transmissão entre os elementos da rede [50].

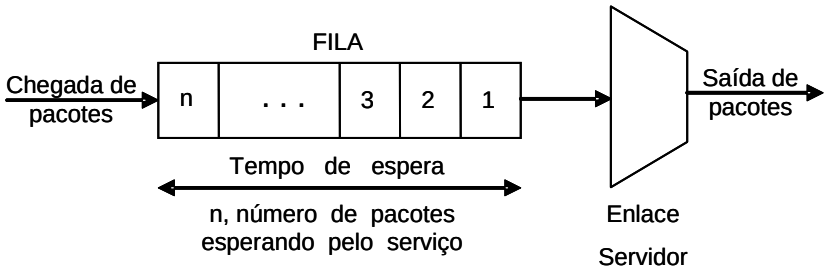

Figura 4.7: Sistema de filas [50].

A Figura 4.7 mostra a condição em que os pacotes chegam a um sistema de filas requisitando um certo serviço como, transmissão de pacotes, por exemplo, esperam em uma fila até serem atendidos caso o serviço não esteja disponível imediatamente e eventualmente partem do sistema após receberem o serviço desejado. É importante lembrar que os pacotes podem ser descartados sem serem atendidos.

4.5.1. FIFO

A disciplina de escalonamento mais simples e comumente utilizada é a FIFO (*First-In-First-Out*). Todos os pacotes são tratados igualmente, colocados em uma única fila e servidos na mesma ordem que chegam na fila como mostrado na Figura 4.8. Esta disciplina também é conhecida como FCFS [38]. As seguintes vantagens são oferecidas por esta disciplina [50]:

1. Carga computacional extremamente baixa se comparada com outras disciplinas.
2. Comportamento previsível. Atraso é determinado pelo tamanho da fila.

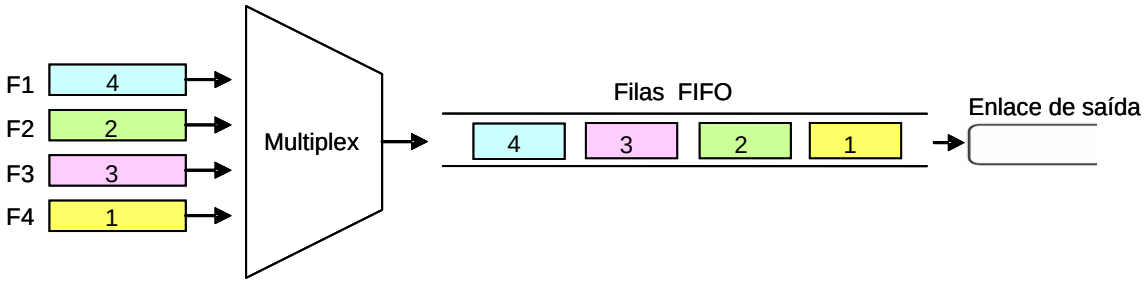

Figura 4.8: FIFO.

A seguir são apresentadas as principais limitações da FIFO:

1. Uma única fila FIFO não permite reordenar/organizar os pacotes armazenados sendo incapaz de priorizar classes de tráfego.
2. Uma única fila FIFO trata igualmente todos os fluxos. Assim, o atraso aumenta igualmente para todos os fluxos à medida que ocorre um congestionamento. Como resultado, FIFO pode significar atraso, *jitter* e perda para aplicação de tempo real.
3. Durante um congestionamento, a FIFO beneficia fluxos de UDP sobre fluxos de TCP. Quando ocorre perda de pacote durante o congestionamento, as aplicações baseadas em TCP reduzem suas taxas de transmissão, mas aplicações baseadas em UDP permanecem inalteradas. Assim, o uso da FIFO pode resultar em aumento de atraso, *jitter* e redução da largura de banda consumida pelas aplicações TCP.
4. Uma rajada pode consumir toda a fila e prejudicar os outros fluxos.

4.5.2. PQ

A disciplina de escalonamento por prioridade PQ (*Priority Queuing*) permite que sejam atribuídos até quatro níveis de prioridades aos diferentes tráfegos gerados na rede: alta, média, baixa e normal. Os pacotes que não são marcados são tratados sem nenhuma prioridade, isto é, sua classificação será de acordo com a prioridade normal, ou seja, equivalente ao *Best Effort* da *Internet.*

Neste modelo de escalonamento o tráfego classificado e marcado como prioritário tem preferência absoluta em relação aos outros fluxos. Esta é uma das desvantagens do PQ quando implementado em redes de baixa velocidade, pois isso implica no aumento de atraso e *jitter* em aplicações de menor prioridade. Mas este mecanismo é não-preemptivo, isto é, a transmissão de um pacote de uma fila de baixa prioridade não é interrompida se um pacote de alta prioridade chegar em sua interface. Dessa forma, um pequeno atraso é adicionado a qualquer pacote que esteja esperando para ser servido, enquanto pacotes de baixa prioridade são transmitidos. Este esquema de escalonamento apresenta alta eficiência quando implementado em redes tipo *backbones*, por ser esta de alta velocidade.

A implementação da PQ em redes tipo *backbone* é bastante útil quando se quer garantir que um tráfego sensível a atraso e *jitter* seja transmitido pela rede e tenha tratamento diferenciado em cada nó ao longo da rede. Por esta razão esta disciplina é a recomendada para servir o trafego das classes sensíveis a estes atributos. Estas classes representam aplicações de tempo real, e estas devem ser tratadas com prioridade em cada nó da rede, em detrimento das aplicações de menor prioridade.

PQ é a base para uma disciplina de escalonamento que suporte classes de serviço diferenciadas. Na PQ clássica, os pacotes são classificados primeiro pelo classificador e então colocados em filas (FIFO) de prioridade diferentes (Figura 4.9). Os pacotes apenas são retirados de uma determinada fila se as filas de prioridades mais altas estiverem vazias [38].

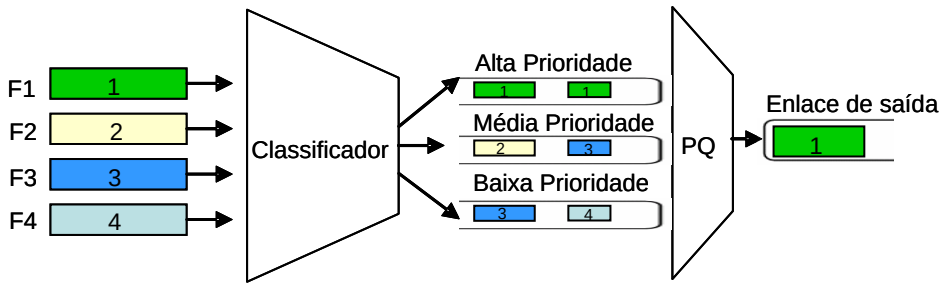

Figura 4.9: PQ.

A seguir são listadas as principais vantagens oferecidas pela disciplina PQ:

1. Carga computacional relativamente baixa se comparada com outras disciplinas.
2. Permite organizar as filas e após a classificação, o tráfego é associado à fila com o nível de prioridade correspondente.

As principais limitações são:

1. Se ocorrer uma grande quantidade de tráfego de alta prioridade, o tráfego de baixa prioridade pode sofrer atrasos excessivos ou ser descartado caso a fila preferida esteja recebendo tráfego constantemente.
2. Pode ocorrer estouro de capacidade das filas de fluxo de baixa prioridade.
3. Um fluxo de alta prioridade mal comportado pode causar atraso e *jitter* em outro fluxo dentro da mesma fila.
4. PQ não é uma solução para atender a limitação da fila FIFO com relação a favorecer o tráfego UDP.

4.5.3. FQ

A disciplina FQ (*Fair Queuing*) possui uma única fila. Assim, nesta disciplina a banda é dividida igualmente pelas N filas, atribuindo a cada fluxo uma fração 1/N da banda do enlace. Essa disciplina classifica os pacotes recebidos pela rede em fluxos e para cada fluxo há uma fila dedicada unicamente a ele. Isolando cada fluxo em sua fila, previne-se que um fluxo mal comportado degrade o desempenho de outros fluxos. O escalonador visita as filas não vazias

de modo circular, servindo um pacote de cada fila por vez. O objetivo da FQ é alocar a mesma quantidade de largura de banda para todos os fluxos: ela não foi projetada para dar suporte a fluxos com necessidades diferentes de largura de banda. A Figura 4.10 ilustra a disciplina FQ [50].

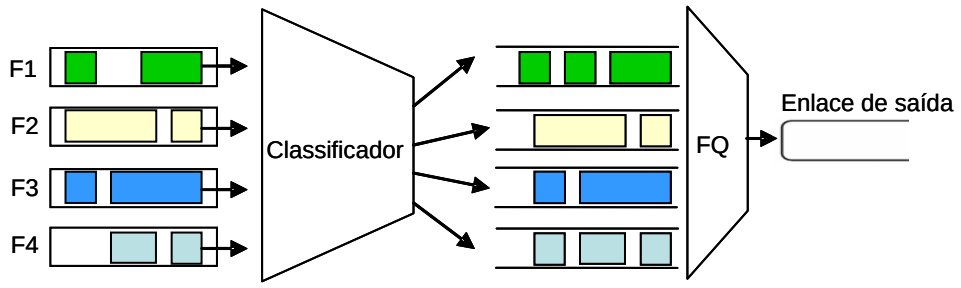

Figura 4.10: FQ.

4.5.4. WFQ

A disciplina de escalonamento WFQ (*Weighted Fair Queueing)* foi desenvolvida para suprir limitações da disciplina FQ abordada na seção 4.5.3, oferecendo, no entanto, a possibilidade de se diferenciar os fluxos, através de um peso atribuído a cada um. Esse peso é utilizado para calcular o tempo associado a cada pacote. A WFQ também suporta pacotes de tamanhos variáveis, e assim fluxos com pacotes maiores não locam mais banda que os fluxos de pacotes menores, o que a torna bastante interessante.

A WFQ possui um classificador que classifica os fluxos de chegada e os insere em filas (FIFO) que podem possuir outras larguras de bandas. O classificador também calcula e etiqueta cada pacote com um número diferente denominado "*Finish Time*". O *Finish Time* representa a ordem com que os pacotes devem ser retirados das filas e encaminhados. Este é calculado a partir da taxa de *bit* de saída, do número de filas ativas, do peso relativo de cada fila e do tamanho do pacote. O escalonador verifica nas cabeças de cada fila qual delas possui o pacote com menor *Finish Time.* O pacote que tiver o menor número é retirado da fila e encaminhado para a saída, conforme mostrado na Figura 4.11 [50].

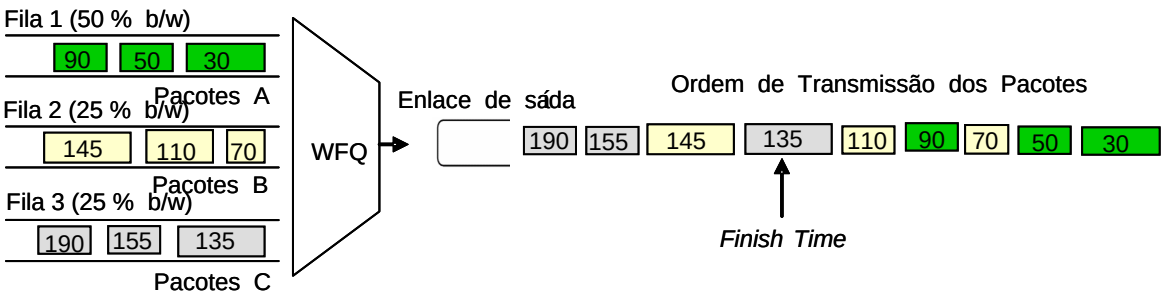

Figura 4.11: WFQ.

O *Finish Time* pode ser calculado pela fórmula a seguir:

$$F_i(k,t) = \max \{F_i(k-1,t), R(t)\} + P(k,t) / \phi(i) \quad (4.1)$$

Onde: $F_i(k,t)$ é o *Finish Time* para o pacote k, da conexão i e no tempo t. O termo $F_i(k-1,t)$ é o *Finish Time* do pacote anterior, R(t) representa o número do *round* no instante t, P(k,t) é o tamanho do pacote k da conexão i no instante t e $\phi(i)$ é o peso da conexão i [49].

A disciplina WFQ apresenta a seguinte vantagem:

1. A WFQ provê proteção para cada classe de serviço assegurando um mínimo de largura de banda independente do comportamento de cada classe.

As principais limitações são:

1. A implementação é feita apenas em *software* e não em *hardware*.
2. Fluxo mal comportado pode impactar no desempenho de outros fluxos.
3. Algoritmo complexo.

4.5.5. WRR

Na disciplina WRR (*Weighted Round Robin*) um peso é associado a cada classe de tráfego. Este peso é utilizado para determinar qual o percentual de banda do enlace deve ser utilizado por cada classe de tráfego associada a uma fila. O percentual destinado a cada tipo de tráfego pode ser representado pela fórmula a seguir:

$$P_i = (W_i / S) \times B. \quad (4.2)$$

Onde P_i representa a taxa em bps alocado para o tráfego da classe i, W_i representa o peso associado à classe i, e S é o somatório dos pesos atribuídos a todas as classes.

As filas são atendidas em ordem decrescente de peso, é um mecanismo semelhante ao CQ (*Custom Queueing*), porém uma fila de maior prioridade deixa de ser atendida quando não há mais pacotes para serem encaminhados ou se o percentual de banda tiver sido utilizado dentro de um mesmo ciclo.

A disciplina de escalonamento WRR também conhecida por CBQ foi desenvolvida para suprir limitações da disciplina FQ e PQ. O WRR supre as limitações do FQ suportando fluxos com diferentes requisitos de largura de banda. O WRR supre as limitações do FQ assegurando que o tráfego de baixa prioridade seja prejudicado. Como o WRR, no mínimo um pacote é removido de cada fila a cada *round.*

O WRR primeiro classifica os pacotes que chegam em várias classes de serviço e então insere os pacotes nas filas dedicadas a cada classe de serviço. Cada fila é servida por um ordenador *Round Robin*. A Figura 4.12 ilustra este algoritmo. [50]

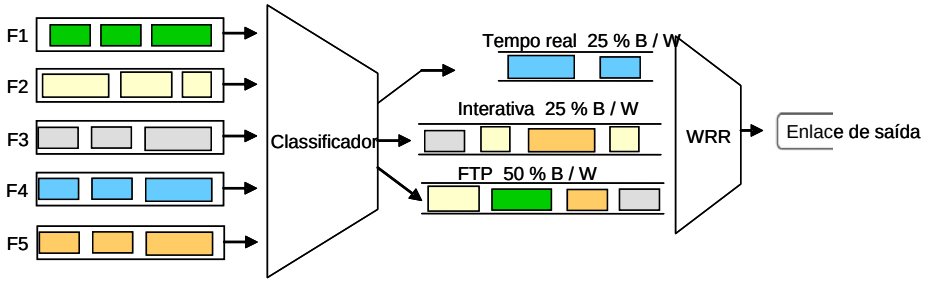

Figura 4.12: WRR.

As seguintes vantagens são oferecidas pela disciplina WRR:

1. Pode ser implementada em *hardware.*
2. Provê um controle sobre cada porcentagem de largura de banda de saída.
3. Assegura que as classes de serviço tenham acesso a uma largura de banda de saída configurada.

A principal limitação é:

- Provê uma correta porcentagem de largura de banda de saída apenas se os pacotes inseridos em todas as filas forem do mesmo tamanho.

4.5.6. DWRR

As limitações dos algoritmos de escalonamento WRR e WFQ são suplantadas pelo desenvolvimento do DWRR (*Deficit Weighted Round Robin*). O DWRR supre as limitações do WRR através de uma distribuição de largura de banda de saída justa, para filas que contenham pacotes de tamanhos variáveis. O DWRR supre as limitações do WFQ reduzindo a complexidade computacional do algoritmo e também por permitir a sua implementação em *hardware*, o que possibilita o uso de interface de alta velocidade [50].

No DWRR cada fila é configurada com vários parâmetros:

1. Um *Weight* define a porcentagem de largura de banda de saída alocada para cada fila. Um *Deficit Counter* especifica o número total de bytes ou *bits* que uma fila pode transmitir cada vez que é visitada pelo escalonador. Se o tamanho do pacote for maior do que o *Deficit Counter,* este último é creditado no saldo da fila para o próximo *round,* onde é feita uma nova comparação.
2. Um *Quantum* de serviço que é proporcional ao peso da fila e expresso em bytes. O *Deficit Counter* é incrementado de um *Quantum* cada vez que a fila é visitada pelo escalonador.

No algoritmo DWRR clássico, o escalonador visita cada fila não vazia e determina o número de bytes do pacote da cabeça da fila. A variável *Deficit Counter* é incrementada do valor *Quantum*. Se o valor do pacote da cabeça da fila é maior que a variável *Deficit Counter* então o escalonador se move para a próxima fila. Se o tamanho do pacote da cabeça da fila é menos ou igual a variável *Deficit Counter,* então esta é decrementada do tamanho do pacote e

o pacote é transmitido pela porta de saída. O escalonador continua retirando os pacotes e decrementando a variável *Deficit Counter* do tamanho do pacote transmitido até o tamanho do pacote da cabeça da fila seja maior que a variável *Deficit Counter* ou que a fila esteja vazia. Se a fila está vazia o valor *Deficit Counter* recebe zero. Quando isso ocorre o escalonador se move para a próxima fila não vazia. A Figura 4.13 mostra a operação deste algoritmo.

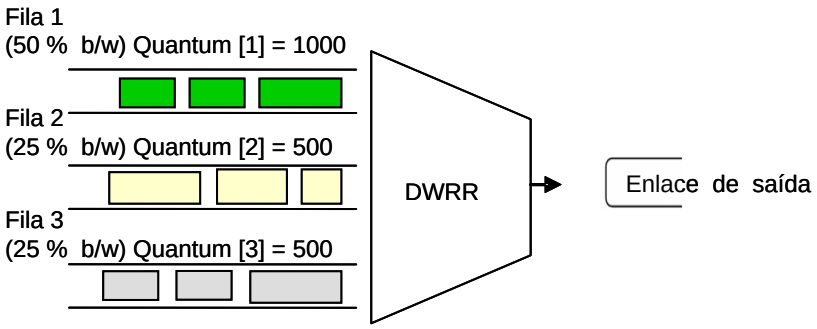

Figura 4.13: DWRR.

As seguintes vantagens são oferecidas pelo DWRR:

1. Supera as limitações do WRR provendo um controle preciso da porcentagem de largura de banda de saída alocada para cada classe de serviços quando se encaminham pacotes de tamanhos variáveis.
2. Supera as limitações do PQ assegurando que todas as classes de serviços terão acesso a uma quantidade de largura de banda de saída configurada.
3. A implementação do algoritmo é relativamente simples.

A principal limitação é:

- Fluxos mal comportados podem impactar no desempenho de outros fluxos.

4.6 Conclusão

Neste Capítulo foram apresentados os conceitos básicos referentes a Qualidade de Serviço, parâmetros que influenciam na QoS, alternativas de mecanismos para provimento de QoS na

Internet e as disciplinas de escalonamento de filas. As diversas alternativas de mecanismos e parâmetros apresentados atendem a certa necessidade de controle sobre o tráfego e possuem diversas implementações possíveis. Tais mecanismos serão implementados nas mais diversas combinações, relacionando as técnicas existentes para prover QoS em redes *backbones* UMTS, que serão descritas no próximo Capítulo.

A interligação de estações móveis sem fio à *Internet* é um caminho sem volta. Assim, a qualidade de serviço nestas redes tem o intuito de atender os requisitos dos diversos tipos de aplicações não só no ambiente aéreo limitado e com problemas inerentes, mas também na rede *backbone* UMTS.

Os mecanismos de provimento de qualidade de serviço, tais como controle de admissão, classificação de pacotes, policiamento e condicionador de tráfego, desenvolvidos e aplicados nos pontos de ingresso à rede *backbone* UMTS podem ser uma alternativa eficiente para atender os requisitos de QoS.

Embora existam diversas pesquisas sobre provimento de QoS através dos referidos mecanismos, o fornecimento de QoS fim a fim ainda continua sendo uma vasta área de pesquisa em aberto. Como pôde ser visto na seção 4.6, diferentes mecanismos e disciplinas de escalonamento foram propostos com o intuito de prover um nível de QoS aceitável em redes *backbone* UMTS.

Foram apresentados neste Capítulo os conceitos referentes à qualidade de serviço em redes *backbone* com tecnologia UMTS, as aplicações que necessitam de garantias de QoS e os diversos mecanismos utilizados para implementá-la.

A descrição detalhada para uma proposta de provisão de QoS em redes *backbone* UMTS que efetue a classificação de pacotes e gerenciamento das estações móveis, bem como, a administração dos recursos nos pontos de ingresso à rede *backbone* UMTS será descrita no próximo Capítulo.

Capítulo 5

Proposta para a Provisão de QoS em Redes *Backbone* UMTS

5.1 Introdução

Uma rede *backbone* com tecnologia UMTS é composta pelos servidores SGSN, GGSN e os reteadores de núcleo. Assim, os pontos de ingresso e egresso a esta rede, são determinados por estes servidores. Como descrito no Capítulo 3, seção 3.4.4, antes que uma MS possa efetivar a comunicação com a rede de dados externa, o contexto PDP deve ser ativado. Nesta fase são descritas as características da conexão com a rede de dados externa, tais como, tipo de rede, endereço de rede, APN (*Access Point Name*), QoS, entre outros. Assim, a rede poderá garantir um nível aceitável da qualidade de serviço na requisição da ativação do contexto PDP. O SGSN que estabeleceu a conexão com a MS solicitante valida a requisição baseando-se na informação de inscrição recebida do HLR durante o processamento do GTP.

Aplicações típicas requerem parâmetros de QoS que incluem limites para largura de banda, atraso fim a fim, taxa de perda de pacotes e variação de atraso. Por outro lado, a provisão e a

manutenção de QoS na interface aérea possuem características únicas. Enquanto que nas redes fixas tipo *backbone* a perda de pacotes geralmente é causada pelo congestionamento da rede.

Os *backbones* UMTS deverão dar suporte às aplicações multimídia, respeitando sempre os requisitos de QoS para cada tipo de aplicação, incluindo serviço de tempo real com garantias. Para tanto, propõe-se neste trabalho um mecanismo para o provimento de qualidade de serviço a ser implementado nas redes *backbone* UMTS, com base na diferenciação de serviços e políticas de escalonamento adequadas para esse tipo de rede.

As aplicações que envolvem fluxos de tempo real ou interatividade necessitam que as redes com tais tecnologias, apresentem melhores e previsíveis tempos de resposta, que possa atender aos perfis de QoS definidos pelo 3GPP além de garantias de confiabilidade e de largura de banda [6], [55].

5.2 QoS no Ambiente *Backbone* UMTS

Devido às características da rede baseada em pacotes e aos limitados recursos da interface aérea, a QoS será um componente chave na transmissão de dados no ambiente *backbone* UMTS. O deslocamento entre redes (*roaming*) terá um impacto nas soluções de QoS, que necessitam garantir o nível de serviço quando os usuários finais cruzarem os domínios. A QoS pode também ser influenciada pela a tarifação diferenciada, onde o mesmo serviço poderia ser oferecido com diferentes níveis de qualidade e, conseqüentemente, com preços distintos [52].

Os requisitos específicos para a interface aérea e para os *backbones* UMTS dentro do padrão 3G/4G estão sob desenvolvimento no grupo 3GPP. São necessários também parâmetros para a QoS entre redes, de maneira a garantir a portabilidade do serviço e a interoperabilidade. O principal objetivo da padronização de QoS do serviço de pacotes UMTS é permitir a entrega de dados com garantias de QoS apropriadas. Para tanto, foi proposta uma arquitetura em camadas com os seguintes elementos chave [53]:

1. mapeamento do serviço fim a fim em serviços providos pelo UE, UTRAN, rede *backbone* UMTS e redes IP externas;
2. classes de tráfego e parâmetros de QoS associados;
3. localização das funções de QoS;
4. negociação de QoS;
5. um modelo de entrega de dados fim a fim.

Conforme apresentado no Capítulo 3 seção 3.4.5, em geral, a classe conversacional possui requisitos mais rígidos de QoS, enquanto que a classe *background* apresenta os requisitos mais baixos de QoS, podendo ser vista como uma classe *Best Effort*. O principal fator de diferença entre estas classes é a sensibilidade ao atraso. A classe conversacional é a mais sensível, enquanto que a *background* é a menos sensível. As classes conversacional e *streaming* são destinadas ao tráfego de tempo real. Ambas preservam a relação entre os elementos de informação do fluxo. Exemplos de aplicações são: telefonia IP (VoIP) para a classe conversacional, e seqüência de vídeo para a classe *streaming* [26].

Os atributos de QoS foram definidos pelo grupo 3GPP que caracterizam as quatro classes de tráfego: taxa máxima de *bits*, taxa garantida de *bits*, ordem de entrega, tamanho máximo da Unidade de Serviço de Dados (SDU), informação do formato SDU, média de erro SDU, média de erro residual de *bit*, entrega de SDUs incorretos, atraso de transferência, prioridade na manipulação do tráfego, prioridade de alocação e retenção [24].

Exemplos de aplicações para as classes UMTS QoS para os serviços oferecidos pelos *backbones* UMTS, que incluem a interoperabilidade com as soluções baseadas em pacote da segunda geração são mostrados na Tabela 5.1 [54], [55]:

Tabela 5.1: Exemplos de Aplicações para as Classes UMTS
e suas Restrições Quanto à QoS.

TIPO	SERVIÇO	TAXA DE DADOS Kbps	ATRASO	*JITTER*	CONFIABILIDADE
Conversacional (tempo real)	Voz (conversacional)	4 -25	< 150 ms	< 1ms	< 3 % Frame Error Rate (FER)
	Videofone	32 - 384	< 150 ms	-	< 1 % FER
	Telemetria (controle)	< 28.8	< 250 ms	-	~ 0 % FER
	Jogos	< 1	< 250 ms	-	< 3 % FER
Interativo	Mensagem de voz	4 -13	< 1 s	< 1 ms	< 3 % FER
	Navegação Web	-	4s / pag		
	e-comerce	-	4 s		~ 0 % FER
Fluxo contínuo (*streaming*)	Fluxo contínuo de áudio	32 - 384	< 10 s	< 1 ms	< 1 % FER
	Vídeo	32 - 384	< 10 s	-	< 1 % FER
	Telemetria	< 28.8	< 10 s	-	~ 0 % FER

A arquitetura de QoS disponível, tanto para as redes de acesso de rádio EDGE quanto WCDMA, envolve a negociação e priorização do tráfego na UTRAN, no *backbone* UMTS e nas interfaces com as redes externas, como no caso da *Internet*. Conseqüentemente, as aplicações podem negociar os parâmetros de qualidade de serviço do início ao fim [6], [55].

5.3 Definição do Problema

A demanda por aplicações de tempo real que possam ser acessadas a partir de estações móveis suportadas pelas redes IP tem tido crescimento rápido nos últimos anos. Entretanto, ainda existem obstáculos para que tais aplicações sejam providas com QoS não só na rede de acesso sem fio, mas também no *backbone* UMTS [55], [65]. Os principais obstáculos são: recursos limitados na interface aérea (UTRAN); o deslocamento entre redes (*roaming*) que provoca grandes congestionamentos à entrada do *backbone*; garantias dos níveis de QoS fim a fim para os usuários de outras PDNs.

Antes que uma MS possa efetivar a comunicação com a rede de dados externa, o contexto PDP deve ser ativado e nesta fase o perfil de QoS é negociado com o *backbone* UMTS. Quando uma estação móvel solicita uma conexão com a rede, um tunelamento é estabelecido

entre uma MS e o GGSN. Este tunelamento, conhecido na UMTS como contexto PDP, provê o UMTS *Bearer Service*. O *backbone* UMTS permite que sejam definidos perfis de tráfego para o qual o contexto pode se adequar (conversacional, *streaming*, interativa ou *background*), entre outros atributos que podem ser negociados [56]. Contudo, nenhuma proposta para prover QoS é definida nesta fase que possa dar suporte a esses perfis e garantir o seu funcionamento. Todavia, a QoS é considerada um dos atributos importantes de um contexto PDP e pode ser negociada durante a sua ativação.

5.4 Solução Proposta

Para que um *backbone* com tecnologia UMTS possa oferecer um nível de QoS exigido pelas aplicações multimídia mesmo em condições de mobilidade das MSs (*roaming*), propõe-se um mecanismo para o provimento de QoS a ser implementada nos SGSN e GGSN. O mecanismo proposto é capaz de realizar a diferenciação do tráfego entrante e o controle de admissão para a rede *backbone* UMTS empregando Serviços Diferenciados como uma alternativa para o provimento de QoS. A arquitetura de Serviços Diferenciados é a tecnologia recomendada para fornecer diferentes níveis de serviço em *backbones* UMTS [24], [57]. Esse mecanismo utiliza uma arquitetura denominada pelo IETF de *DiffServ* e usa a classificação de pacotes como método para obtenção de QoS [27], [37]. Tal arquitetura é de simples implementação nos roteadores atuais e, portanto, tem sido apontada como um bom padrão para implementar QoS em *backbones* UMTS [58]. Além disso, o mecanismo proposto emprega um algoritmo composto por policiadores de tráfego, gerenciadores de filas e escalonadores de tráfego.

A literatura define quatro classes de tráfego para representar as aplicações oferecidas pela tecnologia UMTS [55]. Essas classes são: conversacional, *streaming*, interativa e *background*. A classe conversacional representa as aplicações de tempo real, como voz, a classe *streaming*

representa aplicações de fluxo contínuo de vídeo e áudio, interativa representa aplicações como W.W.W. (*World Wide Web*), e a classe *background* representa aplicações como FTP, e *e-mail*.

Dentro do *backbone* UMTS, os pacotes de dados são encapsulados em pacotes IP e transmitidos através do GTP, sendo os parâmetros de QoS UMTS negociados por meio do contexto PDP via GTP.

Conforme apresentado na seção 4.3.2 do Capítulo 4, um classificador de pacote é responsável por marcar cada pacote, nos pontos de ingresso ao *backbone* UMTS, com o DSCP apropriado. O pacote entrante que chega a um servidor SGSN tem o seu QoS mapeado conforme o seu próprio PHB e então é encapsulado com o DSCP adequado em seu cabeçalho. Em seguida, o pacote aceito é enfileirado.

Para alcançar os níveis de serviços específicos das quatro classes de tráfego QoS UMTS definidas pelo 3GPP, como baixo atraso e *jitter*, é necessária a escolha adequada de um algoritmo de escalonamento de pacotes. A partir de um estudo comparativo de desempenho entre as disciplinas apresentadas, o algoritmo PQ foi a opção adotada, por atender aos requisitos de atraso e *jitter* para aplicações de tempo real, tal como a classe conversacional que apresenta maior sensibilidade a estes atributos. Conforme apresentado no Capítulo 4 seção 4.5.2, o escalonamento PQ controla a taxa média permitida para esses tráfegos ao longo do *backbone* UMTS [50]. Esse controle realiza um tipo de alocação de largura de banda, pois o tráfego admitido na rede estará dentro dos limites permitidos, ocupando apenas este limite do enlace. O tráfego correspondentes as quatro classes de QoS UMTS obedecerão a uma ordem de prioridade conforme Tabela 3.1 [34]. Dessa forma, a fila da classe conversacional terá maior porcentagem de largura de banda do enlace de saída por apresentar maiores requisitos do que a classe *streaming*, que é tratada com menor prioridade. Assim, as filas das classes interativa e *background* serão escalonadas pela disciplina PQ, que transmite o tráfego

dessas filas de acordo com a porcentagem do enlace de saída especificada para cada uma delas. Isso faz com que cada classe utilize somente essa porcentagem, realizando alocação de largura de banda a cada uma delas.

Nos Capítulos subseqüentes serão apresentados os resultados da avaliação de desempenho da aplicação da alternativa proposta, mensurado através de modelagem e simulação dos parâmetros de QoS, como atraso fim a fim, perdas de pacotes, *Jitter* e alocação de largura de banda para as classes de tráfego QoS UMTS.

Trabalhos Relacionados

A necessidade de provisionamento de QoS em redes *backbones* UMTS, foi objeto de pesquisa em [5], [24], [57] e [63]. Os autores em [5] propõem mecanismos para dar suporte às classes de tráfego UMTS baseado na mobilidade das estações móveis (*roaming*). O autor em [24] propõe o uso do modelo *DiffServ* para prover QoS no *backbone* UMTS. O autor em [57] propõem algoritmos específicos de escalonamento do tráfego com o objetivo de minimizar o congestionamento no *backbone* e controle de admissão e classificação de pacotes nos pontos de ingresso e egresso ao *backbone*, com base no conceito da arquitetura *DiffServ*. Em [63] o autor propõem um gerenciamento de QoS baseado em políticas para uma rede UMTS através de um mecanismo para o provimento de QoS baseado em *DiffServ*.

Contudo, estes trabalhos não propuseram nenhum mecanismo para a provisão de QoS em redes *backbones* com tecnologia UMTS que pudesse ser implementado durante a fase da ativação do contesto PDP. Um mecanismo que realizasse a classificação de pacotes nos pontos de ingresso ao *backbone* e uma política de escalonamento para cada uma das classes de QoS UMTS de forma individual, conforme foram apresentados nos Capítulos 4 e 5 desta dissertação. Assim sendo, pretende-se com a referida proposta, implementar durante a fase da ativação do contexto PDP um mecanismo que efetue a classificação de pacotes e alocação de

largura de banda para cada classe de tráfego de QoS UMTS de forma a oferecer um nível de QoS aceitável a cada uma delas mesmo em condições de congestionamento, além de distribuir a banda passante não utilizada entre as demais classes, caso uma delas não utilize toda a banda a ela alocada.

5.5 Conclusão

Neste Capítulo foi apresentada uma alternativa de mecanismo de QoS para a definição do problema encontrado em redes *backbone* com tecnologia UMTS bem como uma visão global da arquitetura de QoS proposta para estes sistemas, com exemplos de soluções para as diversas partes da arquitetura.

A proposta para provisão de QoS em redes *backbone* UMTS é baseada nos modelos *DiffServ* e o algoritmo PQ. Essa proposta deverá gerenciar a alocação de recursos desta rede para que seja viável o suporte de aplicações sensíveis a atrasos como as de tempo real com QoS e o suporte ao alto volume de tráfego nos pontos de ingresso à rede devido a mobilidade das MSs.

A referida proposta para o provimento de QoS será implementada com o objetivo de realizar o controle de admissão, classificação de pacotes, policiamento e condicionamento de tráfego para cada classe de QoS UMTS. Com tais reservas de recursos o tráfego dessas classes é protegido nos momentos de congestionamento, de forma que as classes, cujo tráfego excede aos parâmetros contratados e tende a ocupar o máximo possível da capacidade do enlace, são penalizadas com descarte de pacotes para que mantenha o tráfego transmitido dentro da banda a elas alocada.

No próximo Capítulo, os mecanismos *DiffServ* e o algoritmo de escalonamento por prioridade PQ serão avaliados como uma alternativa capaz de prover qualidade de serviço quando implantados em redes *backbone* empregando tecnologia UMTS.

Capítulo 6

Avaliação da Proposta para a Provisão de QoS em Redes *Backbone* UMTS

6.1 Introdução

Os Capítulos anteriores apresentaram uma visão geral das redes móveis sem fio e seus esquemas para comutação de pacotes e provimento de QoS. Foi apresentado também uma descrição dos mecanismos que podem ser aplicados nos elementos da rede para prover um tratamento diferenciado a diferentes tipos de mídias e suprir os requisitos de QoS expostos por tais aplicações. Verificou-se também, que a implementação de QoS ao longo do *backbone* UMTS é possível através do procedimento de ativação de um contexto PDP, que é ativado através de uma solicitação feita por uma MS ao SGSN e nesta fase o perfil de QoS é negociado com o *backbone* UMTS.

O Capítulos 5 apresentou a proposta de mecanismo para o provimento de QoS a ser implementado nos pontos de ingresso e egresso à rede *backbone* UMTS. Apresentou-se também, que tal mecanismo é capaz de realizar a diferenciação do tráfego entrante e a classificação de pacotes para esta rede. Esta proposta se baseia na arquitetura *DiffServ* e no algoritmo de escalonamento por prioridade PQ implementado no núcleo do *backbone* UMTS.

Neste Capítulo, através de um estudo baseado em modelagem e simulação avalia-se o desempenho do mecanismo de QoS proposto quando implementados em uma rede *backbone* com tecnologia UMTS. Para tanto, serão analisados os parâmetros de QoS como atraso, alocação de largura de banda, perda de pacotes e *jitter* para as quatro classes de tráfego UMTS, a fim de verificar a ocorrência de diferenciação de serviço entre elas e avaliar o comportamento da rede sob condições de alto tráfego devido ao *roaming* das MS.

O *backbone* UMTS através do servidor SGSN e o *gateway* GGSN recebe o tráfego gerado por uma quantidade aleatória de MSs ou PDNs, sendo que o SGSN atua como o servidor de ingresso a esta rede, implementando funções de classificação de pacotes, conformação do tráfego e marcando os pacotes com os DSCP apropriados. O núcleo do *backbone* apenas fará o escalonamento dos pacotes e gerenciamento do tráfego baseado em sua marcação [55], [65].

Na seção subseqüente, será descrito como o ambiente de teste foi construído e, em seguida, serão apresentadas as análises de resultados e as conclusões.

6.2 Modelo de Simulação para o Ambiente *Backbone*

O modelo de simulação descrito nesta seção será desenvolvido com objetivo de fornecer condições para uma avaliação comparativa e consistente da proposta de uma das alternativas para o provimento de QoS nos *backbones* UMTS. O estudo de desempenho do mecanismo de QoS proposto foi a partir da implementação de algoritmos que efetuem a classificação de pacotes nos pontos de ingresso e egresso ao *backbone* UMTS e o gerenciamento do tráfego e o escalonamento dos pacotes ao longo deste *backbone*.

O comportamento destes *backbones* pode ser estudado basicamente por duas formas que possam avaliar os parâmetros de QoS. Estes dois modelos podem ser aplicados para representar e medir o desempenho de um sistema em análise, isto é, o modelo analítico e o modelo de simulação. Os modelos analíticos se baseiam em fórmulas matemáticas como, por

exemplo, equações que representam um modelo de filas de um sistema, de modo a gerar medidas de desempenho a partir de parâmetros fornecidos.

Após a identificação dos cenários para uma determinada topologia de uma rede em análise, um modelo de simulação pode ser implementado em *software* para simular o comportamento do sistema. Os modelos de simulação são amplamente utilizados em várias áreas da engenharia para aplicações bastante variadas. Uma vez identificados os cenários do sistema a ser modelado, uma rede móvel sem fio, por exemplo, pode ser modelada através de um simulador baseado em eventos. Através de um simulador é possível efetuar diversas análises sobre o funcionamento de diversos protocolos e elementos de uma rede. Um dos objetivos de uma simulação é permitir que um usuário configure topologias e cenários dificilmente implementáveis na prática de laboratório.

6.2.1 O Simulador para o Ambiente de Rede *Backbone*

O simulador de rede NS (*Network Simulator*) versão 2 [59] foi a opção adotada para o estudo do comportamento dos *backbones* UMTS e posteriormente avaliar a alternativa proposta para atender requisitos de QoS de diferentes aplicações. O NS é um simulador baseado em eventos e de domínio público desenvolvido para simular redes de comunicação fixas e móveis. Sua escolha se deve ao fato da implementação do NS ser um código aberto onde novos módulos podem ser facilmente adicionados e módulos existentes modificados. Vários pesquisadores têm colaborado com módulos adicionais sendo que dois destes módulos (seção 6.3) foram utilizados na implementação do simulador usado nesta dissertação.

O NS é uma ferramenta orientada a objetos, sendo que toda a sua hierarquia de classes é construída em C++, e utiliza a linguagem *script* OTcl como interface de comando e de configuração. Cada classe em NS é composta por parâmetros de configuração, variáveis de

estado e métodos. Por ser uma arquitetura aberta, este simulador permite ao usuário adicionar novos módulos e funcionalidades a partir da extensão da hierarquia de classes existentes.

Um modelo de simulação consiste em uma topologia e eventos de escalonamento como, por exemplo, início e término de uma fonte de tráfego durante a simulação. Um cenário é construído a partir dos modelos de tráfego gerados na rede, dos elementos (ou nós) da rede e suas interconexões. Para interconectar os diferentes nós de uma rede, são usados enlaces configurados com uma certa largura de banda, atraso e disciplina de escalonamento de filas.

As próximas seções descrevem o modelo de simulação para uma rede móvel UMTS, a qual alimentará o *backbone* UMTS, em que será aplicado o mecanismo proposto para a provisão de QoS. Desta forma, utilizando a ferramenta de simulação NS-2, obtém-se o tráfego para esta rede, cujos resultados não serão objeto de análise, sendo que, o intuito deste estudo baseado em simulações é avaliar o comportamento do *backbone* UMTS mediante a aplicação do mecanismo proposto na seção 5.4 desta dissertação para o provimento de QoS.

6.3 Modelo de Simulação da Rede Móvel UMTS

O simulador NS em sua implementação original não contempla o módulo para interface aérea UMTS. Geralmente, os módulos para simulação de redes móveis fornecidos em NS, (versão 2.1b9a) são configurados como módulos "*ad-hoc*" podendo se comunicar diretamente com qualquer outra estação móvel dentro de seu alcance. Considerando uma topologia UMTS, cada estação móvel (MS) é criada para se comunicar apenas com o BSS dentro de sua área de serviço, isto é, ela não deve se comunicar com outras MSs. A fim de solucionar esta limitação, o agente de roteamento NOAH (*Non-Ad-hoc*) desenvolvido por Alfredo Todini [61] foi utilizado. Um agente NOAH roteia todos os pacotes enviados para ou de uma MS através de um BSS, inviabilizando a comunicação entre MSs.

Foi necessário adicionar outras adequações a esse módulo e aplicada à camada RLC (*Radio Link Control*). Esta camada substitui a camada de enlace na arquitetura de protocolos de uma rede móvel sem fio. Em [61] a camada RLC é parcialmente baseada na classe GPRS RLC. Foi adicionado o suporte a múltiplas entidades RLC em cada nó, além de modificações extensivas aos protocolos para modelar precisamente as características da camada RLC UMTS. Outras alterações ocorreram na camada MAC (*Medium Access Control*). A camada MAC UMTS substitui a camada MAC 802.11 na arquitetura de protocolos de rede móvel sem fio. Desta forma, o *backbone* UMTS pode ser simulada como uma rede toda IP (*All-IP*) onde todas as transações são feitas utilizando o protocolo IP [67], [68].

Está prevista a existência de somente uma única célula, sendo que o número de estações móveis está limitado mais por restrições de memória da máquina do que pelo módulo em questão. As estações móveis representam o módulo da interface aérea UMTS, sendo que a utilização deste módulo foi apenas para gerar o tráfego de entrada para o *backbone* UMTS e fez surgir duas restrições. A primeira é que ele está disponível somente para a versão 2.1b9a do simulador, ao passo que o módulo *DiffServ* está disponível apenas para a versão 2.1b8a. A segunda é que ele suporta somente uma célula. Para suplantar a primeira restrição foi necessário compatibilizar duas simulações em diferentes versões do NS. Sendo a simulação da rede móvel UMTS na versão 2.1b9a do NS e a simulação do *backbone* UMTS na versão 2.1b8a, respectivamente.

O modelo de simulação para uma rede móvel UMTS *uplink,* ou seja, da MS para a BS, está ilustrado na Figura 6.1. O modelo abrange um *gateway* GGSN, um servidor SGSN, um BSC, uma BS e conexões concorrentes para várias MSs. O GGSN é o elemento de origem dos fluxos de tráfego, o servidor SGSN e o BSS são os elementos intermediários da rede e as MSs são os elementos destino das conexões. O BSS tem como função alocar recursos de rádio na célula corrente da MS e reservar recursos nas células adjacentes caso seja necessário.

Neste modelo de simulação considera-se a presença de 16 estações móveis, sendo que os tipos de tráfegos considerados foram os citados na Tabela 6.1. Para cada experimento foi considerado o número máximo de MS e os seguintes parâmetros:

- número de estações móveis como carga de tráfego, variando entre (1 a 16 MSs);
- tempos iniciais e finais de cada tipo de tráfego;
- ocorrência ou não de cada tipo de tráfego;
- mobilidade ou não das MSs;
- número de estações base.

Estes parâmetros são padronizados para este módulo, os demais basicamente definem o ambiente dentro da célula, informando quantas MSs existirão, se a simulação é no modo GSM, GPRS ou UMTS, o número de freqüência utilizadas, entre outros. Os parâmetros escolhidos definem apenas MSs que transmitem somente tráfego de pacotes para o *backbone* UMTS iniciando e sendo finalizadas em 30 minutos por simulação.

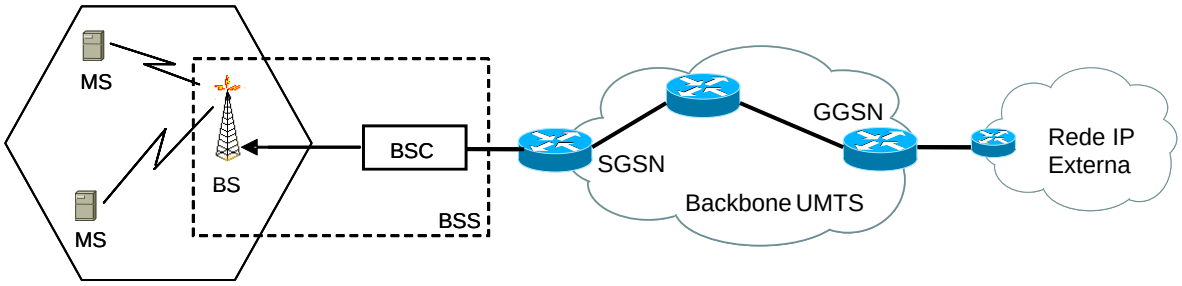

Figura 6.1: Modelo de Simulação da Rede Móvel UMTS.

6.4 Modelo de Simulação da Rede *Backbone* UMTS

Com a definição do modelo de simulação de rede na seção 6.2.1 e dos parâmetros de simulação da rede móvel na seção 6.3, foi considerada a interconexão desta rede com a rede *backbone* UMTS. Desta forma, foi possível gerar o tráfego para o ambiente *backbone*, utilizando as duas versões do simulador de acordo com [63] e com parâmetros definidos em [55]. Para tanto, foi utilizada uma funcionalidade do simulador, que permite tomar um arquivo de registro (*trace*) gerado por uma simulação, transformá-lo em um arquivo contendo atraso de pacotes, alocação de banda, perdas de pacotes e *jitter*, e em seguida converter este arquivo obtido em formato binário.

Como definido nesta mesma seção, os resultados obtidos deste arquivo resultante podem ser utilizados como tráfego de entrada em outra simulação. Cada simulação representa uma agregação de tráfego interferente, sendo que o tráfego resultante de todas as simulações foi combinado como entrada no *backbone* UMTS na segunda simulação. O aumento do número de células e MSs será caracterizado como o acréscimo do percentual da carga de tráfego interferente.

Para isso, a realização dos testes foi com base na versão 2.1b8a do simulador em questão. Nesta versão existe, além do módulo que simula o modelo *DiffServ* já previsto na implementação original do simulador, um módulo adicional que foi desenvolvido em [60], o qual amplia e facilita a programação deste método. Este módulo basicamente cria novos escalonadores, novas políticas de descarte e permite uma nova alternativa de mecanismo de QoS para marcação de pacotes baseados em fatores como origem, destino, protocolos de transporte e aplicação.

A Figura 6.2 ilustra o modelo utilizado nas simulações para um *backbone* UMTS, considerando-se o nó 1 como um SGSN representando o roteador de ingresso à esta rede e o nó 2 como um GGSN que representa o *gateway* de interconexão com outras PDNs.

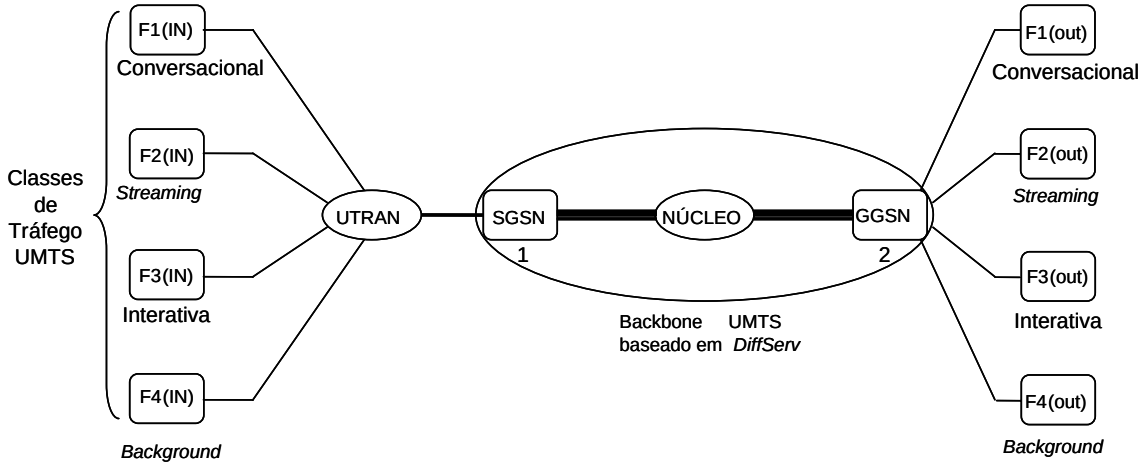

Figura 6.2: Modelo de Simulação da Rede *Backbone* UMTS.

6.4.1 Parâmetros de Simulação da Rede *Backbone* UMTS

Nesta seção avalia-se o comportamento da rede *backbone* UMTS mediante a aplicação do mecanismo para o provimento de QoS, conforme os requisitos do perfil de QoS, tais como: perdas de pacotes, atraso, alocação de banda e *jitter* para as classes de tráfego UMTS. Considerando o crescimento da sobrecarga de tráfego na rede, a avaliação terá como base o nível de diferenciação das classes de tráfego UMTS.

Conforme definido nas seções 6.2 e 6.3 a rede móvel sem fio é composta por quatro nós de entrada, os quais representam as fontes das classes de tráfego UMTS. O *backbone* UMTS compõe-se basicamente dos seguintes elementos: o nó 1 de ingresso a esta rede representado pelo SGSN, rede de núcleo que é formada por vários nós e o nó 2 de egresso ao *backbone* representado pelo GGSN [65].

O congestionamento do *backbone* UMTS é modelado através de um enlace situado entre o roteador SGSN e o *gateway* GGSN conforme modelo de simulação da rede móvel UMTS e o *backbone* UMTS ilustrado na Figura 6.2. Este enlace possui uma capacidade de escoamento do tráfego de 0,8 Mbps que será compartilhado com as classes de tráfego UMTS definidas na Tabela 6.1. Para cada simulação, foi provocado um aumento da carga de tráfego, onde a

intensidade de tráfego variou entre 140% e 200% de sua taxa inicial, totalizando assim 11 simulações para cada experimento. Os diferentes tipos de tráfego seguem a distribuição de Poisson, pois o intervalo de tempo entre chegadas de novos fluxos de tráfego à BSS é aleatório seguindo uma distribuição exponencial. Todas as simulações foram realizadas com um tempo igual a 30 minutos. A Tabela 6.1 apresenta os parâmetros definidos para as quatro classes de tráfego UMTS.

Tabela 6.1: Parâmetros definidos para a simulação.

Classes de Tráfego UMTS	Taxa	Padrão *DiffServ*	Algoritmo de Escalonamento
Conversacional	32 Kbps	EF	*Priority Queueing*
Streaming	64 Kbps	AF1	*Priority Queueing*
Interativa	96 Kbps	AF2	*Priority Queueing*
Background	128 Kbps	*Best Effort*	*Priority Queueing*
Fluxo FTP (máx.)	1.100 Kbps		

Considerando estes parâmetros, pode-se chegar a um modelo com os seus componentes, que implementados em uma rede *backbone* UMTS possa prover uma qualidade de serviço satisfatória, sendo que o tráfego entrante nesta rede é representado pelas quatro classes de tráfego UMTS [55], [65]. As aplicações são do tipo conversacional, *streaming*, interativa e *background*, utilizando os protocolos de transporte UDP e TCP. Estas aplicações foram escolhidas por apresentar diferentes requisitos de QoS e por utilizarem protocolos de transporte com comportamentos distintos.

Conforme apresentado na seção 6.2, o objetivo das simulações é avaliar o desempenho do *backbone* UMTS sob a aplicação do mecanismo proposto para provimento de QoS nesta rede. Como apresentado na seção 5.4, a proposta para a provisão de QoS consiste na implementação da arquitetura *DiffServ* neste *backbone*. Para tanto, um plano de mapeamento das classes UMTS para PHBs *DiffServ* é proposto como um modelo capaz de implementar um

nível de qualidade de serviço satisfatório. Antes que uma MS estabeleça uma conexão com uma rede de dados externa, um contexto PDP deve ser ativado. Este *backbone* permite que sejam definidos perfis de tráfego para o qual o contexto pode se adequar (conversacional, *streaming*, interativa ou *background*), entre outros atributos que podem ser negociados. Entretanto nenhum método para prover QoS é definido nesta fase que possa dar suporte a esses perfis e garantir o seu funcionamento.

O *backbone* UMTS recebe o tráfego gerado pela rede móvel UMTS, sendo que o nó de ingresso atua como o roteador de borda, implementando funções de controle de admissão, conformação do tráfego e marcação de pacotes com os DSCP apropriados. Conforme apresentado nas seções 6.3 e 6.4 os elementos SGSN e GGSN desta rede, são os responsáveis por estas tarefas. Os nós seguintes compõem a rede de núcleo e são caracterizados como roteadores de núcleo, que simplesmente usarão o DSCP para o escalonamento e o encaminhamento apropriado dos datagramas IP [5], conforme apresentado nas seções 3.4.4 e 4.4.2 desta dissertação.

6.4.2 Cenário das Simulações

Por meio de modelagem e simulação, utilizando o software NS-2 [60], foi realizado um estudo de desempenho da aplicação do mecanismo de QoS proposto na seção 5.4. A Figura 6.2 na seção 6.4 ilustra o modelo seguido nas implementações dos cenários utilizados nas simulações.

As fontes de tráfego escolhidas para este experimento são as classes de tráfego UMTS combinadas entre si, ou seja, conversacional, *streaming*, interativa ou mesmo a *background*. Tomando este modelo de tráfego e a estrutura *DiffServ*, o PHB EF foi a opção escolhida para acomodar a classe conversacional, porque garante parâmetros mais rígidos. As outras três classes de tráfego UMTS podem ser acomodadas pelos PHBs AF, que são mais flexíveis. A

classe *streaming* pode corresponder à classe AF1, a classe interativa à classe AF2 e a classe *background* corresponderia à classe de melhor-esforço (*Best Effort*) [34], [38].

O tráfego denominado de fluxo F1 é gerado a partir de fontes de tráfego que possuem a exigência de atraso e *jitter* semelhantes aos da classe de tráfego conversacional. O fluxo F2 representa o tráfego *streaming*. O tráfego interativo F3 vem da agregação de tráfegos do tipo HTTP. Finalmente, o fluxo *background* F4, que modela o tráfego, tipicamente de e-mails e de fluxos de FTP, este foi utilizado para congestionar a rede, ou seja, o tráfego interferente. F1 possui alta sensibilidade a atrasos e *jitter*. F2 tem uma alta sensibilidade a *jitter* e F3 tem uma alta sensibilidade a atrasos. F4 não é sensível a atrasos ou *jitter*.

6.5 Apresentação e Análise dos Resultados

Nesta seção apresentam-se os resultados de simulação referentes aos modelos apresentados nas seções 6.3 e 6.4, bem como uma análise destes resultados. Para isso, foram avaliados parâmetros de QoS fim a fim no *backbone* como atraso de pacotes, alocação de banda, perdas de pacotes e *Jitter* para as classes de tráfego UMTS. Uma vez obtidos e analisados os resultados da simulação, é possível avaliar também a performance do *backbone* UMTS ao empregar o mecanismo baseado em *DiffServ* como alternativa para o provimento de QoS.

Com o objetivo de avaliar cada alternativa do mecanismo proposto, ferramentas gráficas e arquivos de análise numérica disponível pelo *software* NS foram utilizados. Animações dos elementos do *backbone* UMTS empregando o mecanismo para provisão de QoS puderam ser visualizadas através de uma ilustração gráfica da topologia de uma rede simulada através de uma ferramenta gráfica chamada NAM (*Network Animator*) [64].

Todas as classes de tráfego UMTS apresentaram uma taxa de ocupação inicial de 140% incluindo o fluxo FTP. Assim sendo, foi provocado um congestionamento na rede com o

aumento do tráfego, através dos fluxos F2 e o FTP como tráfego interferente até um limite de 200%.

6.5.1 Análise Comparativa entre Disciplinas de Escalonamento.

O estudo apresentado nesta seção tem como objetivo principal obter resultados para uma análise comparativa entre as seis disciplinas de escalonamento apresentadas na seção 4.6 (FIFO, PQ, FQ, WFQ, WRR, e DWRR). Verificar também quais delas atendem aos requisitos de atraso para a aplicação conversacional, em função de esta ser mais sensível a atraso. Para uma avaliação mais consistente dessas disciplinas, as mesmas serão avaliadas sob condições de congestionamento na rede provocado pelo aumento do tráfego através de fluxos FTP até um limite de 200%. A avaliação do atraso fim a fim no *backbone* será caracterizada com o aumento da intensidade de tráfego entre 140% e 200%. A Figura 6.3 mostra os resultados obtidos desta simulação para as seis disciplinas de escalonamento.

Para esta simulação com base nas classes de precedência e o atraso atribuído às aplicações o qual permite que as mesmas sejam tratadas de forma diferenciada, o 3GPP em [55] determina as classes de precedência e as classes de atraso para as SDUs conforme Tabela 6.2 a seguir e a Tabela 3.2 da seção 3.4.4.

Tabela 6.2: Classes de precedência e atraso [55]

Aplicações	Classe de precedência	Classe de atraso	Taxa garantida máxima [bps]	Atraso máximo [ms]
Background	Baixa	4	-	-
Interativa	Baixa	3	-	-
Streaming	Média	2	16000	300
Conversacional	Alta	1	16000	80

Esta simulação permite uma avaliação das disciplinas de escalonamento sob o aspecto de atraso de pacotes compreendendo os pontos de ingresso e egresso do *backbone* para a classe de tráfego UMTS conversacional.

Observando os resultados da Figura 6.3 verifica-se que todos os algoritmos atenderam às classes de atraso até a um congestionamento de 165% da rede. Observa-se também que a disciplina FIFO apresentou atraso compatível com outros algoritmos até 170%, isso ocorre devido a sua baixa carga computacional e a variação de atraso dos pacotes. Contudo a disciplina PQ apresentou o menor atraso para a aplicação conversacional (mais sensível ao atraso), por atribuir prioridade absoluta a essa aplicação. Apesar dos atrasos serem baixos, pode ser que os mesmos não sejam suficientes para atender às exigências desta classe de tráfego, tal como uma videoconferência de tempo real. Assim, em uma transmissão de tempo real seria negociada uma maior capacidade de transmissão durante a ativação do contexto PDP para prover atrasos máximos inferiores a 80 ms conforme Tabela 6.2 e [55]. Entretanto, esta simulação tem o objetivo de avaliar quais disciplinas de escalonamento atendem melhor às classes de tráfego para uma aplicação de tempo real. Esta aplicação será mantida como referência de classe de precedência alta ou classe de atraso 1 no *backbone* UMTS.

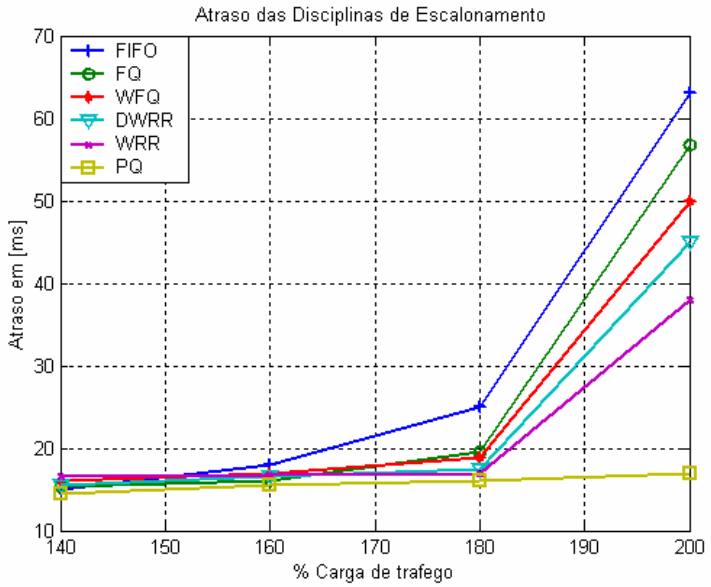

Figura 6.3: Atraso por disciplina de escalonamento.

A Figura 6.3 mostra que ocorreu um aumento considerável dos atrasos a partir de 180% da carga de tráfego imposta por F2 (*streaming*) e o FTP. Observa-se também que a disciplina

FIFO apresenta um aumento dos atrasos a partir de 160% de utilização da rede e intensificando a partir dos 180%. Este resultado já era esperado, uma vez que a disciplina FIFO não possui métodos de classificação e escalonamento de pacotes.

Quando a rede atinge um nível de utilização de 180%, isto é, um congestionamento causado pela alta intensidade de tráfego, este cenário representa um esforço maior para a rede manter os atrasos dentro de níveis aceitáveis. Observa-se que 180% de utilização da rede representa o ponto limite para o restante das disciplinas, uma vez que todas deixaram de atender ao parâmetro atraso como já era esperado. Este cenário apresenta uma condição extremamente crítica não só para a classe conversacional, mas também para todas as classes.

A disciplina FIFO, por tratar todos pacotes da mesma maneira, obteve atrasos bastante elevados a partir dos 160% de congestionamento da rede. Nota-se que até o limite de 180% de utilização da rede os atrasos com a disciplina WFQ e FQ se mantiveram na média. Os resultados obtidos pelas disciplinas WRR e DWRR foram bastante próximos durante todo o tempo de observação, entretanto apesar do bom desempenho não puderam atender os requisitos de atraso dentro dos limites aceitáveis.

Observando a Figura 6.3 quando a carga de tráfego varia entre 140% e 180%, onde as disciplinas de escalonamento de pacotes se mantiveram de forma estável, é importante destacar os benefícios de se trabalhar com a configuração dos pesos das disciplinas, WFQ, WRR e DWRR. Para esta fase os pesos podem ser atribuídos diferentemente de forma a manter os atrasos dentro de condições desejadas. Tais recursos podem ser configurados de forma a adequar o tráfego a uma situação desejada e com uma boa precisão.

Analisando a disciplina PQ comparada a outras disciplinas observa-se pela Figura 6.3 que ela obteve menores atrasos de forma estabilizada durante todo o período observado mesmo com o aumento da taxa de congestionamento na rede.

Diante dos resultados obtidos, apresenta-se a justificativa para a escolha do algoritmo PQ para a realização das simulações empregando o mecanismo proposto baseado na arquitetura *DiffServ* e neste algoritmo, em função do mesmo apresentar menor atraso para aplicações de tempo real.

Como descrito na seção 4.6.2 um classificador de pacotes é responsável por marcar cada pacote, nas bordas da rede *backbone* UMTS (SGSN e GGSN), com o DSCP apropriado. Os pacotes que chegam no SGSN (ponto de ingresso do *backbone*) tem o seu QoS mapeado conforme o seu próprio PHB. Dentro desta rede, os pacotes de dados são encapsulados em pacotes IP e transmitidos através do GTP, sendo que os parâmetros de QoS UMTS são previamente negociados por meio do contexto PDP via GTP.

a) Experimento 1: Atraso de pacotes para os fluxos F1, F2, F3 e F4.

A avaliação do atraso para os fluxos F1 (Conversacional) e fluxo F2 (*Streaming*) nos pontos de ingresso e egresso à rede foi caracterizada quando a intensidade de tráfego oferecida variou entre 140% e 200% de sua taxa inicial. Em relação aos requisitos apresentados pelos fluxos quanto aos itens de atraso e largura de banda, verifica-se que o atraso fim a fim sofrido pelo fluxo F1 é menor que os valores encontrados pelo fluxo F2. Esta característica está de acordo com o esperado, pois no escalonamento realizado por PQ a maior prioridade é dada ao tráfego de classe conversacional, como faz para o EF PHB (resultados obtidos na seção 6.5.1, Figura 6.3). Desta forma, assim que houver pacotes dessa classe a serem transmitidos, sua fila será atendida em prioridade à fila do fluxo F2, sem contudo interromper a transmissão de pacotes do fluxo F2, caso os pacotes de fluxo F1 tenham chegado durante a sua transmissão. A Figura 6.4 mostra o atraso desde o ponto de ingresso até o egresso ao *backbone* sofrido pelas aplicações F1 e F2.

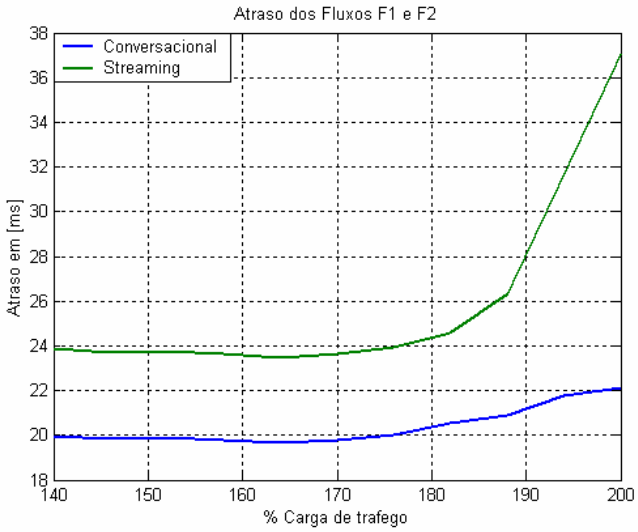

Figura 6.4: Atraso fim a fim para os fluxos F1 e F2.

Analisando a Figura 6.4 e a Tabela 6.3, com a aplicação do mecanismo baseado em *DiffServ* observa-se que o fluxo F1 obteve menores atrasos do que F2 de forma estabilizada durante todo o período observado mesmo com o aumento da taxa de congestionamento na rede. Os atrasos dos fluxos F2 (Fig. 6.4), F3 e F4 (Fig. 6.5) se justificam em função da alta prioridade concedida ao fluxo F1, que pode ser explicado pelas longas filas escolhidas em face da ação do algoritmo PQ. Um tamanho menor de fila teria resultado em atrasos menores. A Tabela 6.3 mostra os valores correspondentes ao intervalo de confiança para o atraso fim a fim, ou seja, nos pontos de ingresso e egresso ao *backbone* considerando um nível de confiança de 195%.

Tabela 6.3: Atraso em ms para os fluxos F1 e F2 com o aumento do tráfego.

Classes de Tráfego	**Atraso Médio [ms]**	**Intervalo de Confiança**
Conversacional (F1)	20,39036	19,8826 / 20,8981
Streaming (F2)	25,97934	23,3774 / 28,5806

Os atrasos dos fluxos F2, F3 e F4 foram superiores ao F1 durante todos os períodos observados, apresentando seu pico no ponto de maior congestionamento da rede, isto é, acima de 190%. Os pacotes dos fluxos F3 e F4 foram atrasados e sua entrega continua depois de finalizada a entrega dos pacotes dos outros fluxos. Durante este período final seu atraso foi reduzido, pois não havia tráfego na rede. A Figura 6.5 a seguir mostra os atrasos sofridos pelos fluxos F3 (Interativa) e fluxo F4 (*background*), a qual confirma os resultados obtidos na seção 6.5.1 e mostrados na Figura 6.3, onde o algoritmo PQ concede alta prioridade ao fluxo F1 e média ao fluxo F3 e baixa prioridade ao F4.

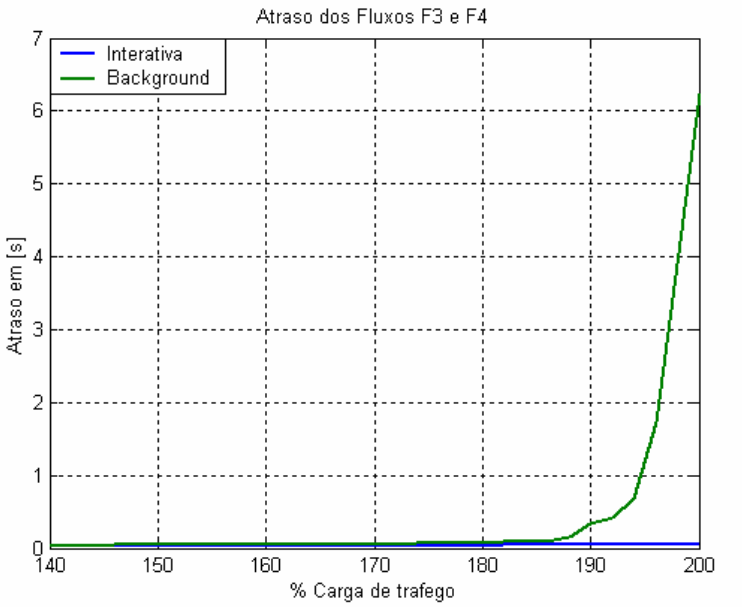

Figura 6.5: Atraso fim a fim para os fluxos F3 e F4.

b) Experimento 2: Banda utilizada para os fluxos de tráfego F1, F2, F3 e F4.

A Figura 6.6 mostra o resultado da banda utilizada para cada fluxo quando a carga de tráfego varia de 140% a 200% e a Tabela 6.4 o valor médio e os intervalos de confiança para os fluxos F1 e F2 considerando um nível de confiança de 95%. Verifica-se que o mecanismo proposto foi capaz de alocar a banda requerida por cada uma das classes de tráfego UMTS, isto é, fluxos F1, F2, de forma que estes fluxos foram protegidos durante todo o tempo mesmo

nos pontos de alto tráfego. Verifica-se também, que nos pontos onde a carga de tráfego variou-se entre 150% e 190 % da sua taxa inicial, os fluxos, F3 e F4 tiveram suas vazões diminuídas, ou seja, a capacidade de banda reservada para essas classes foi alocada pelos fluxos F1 e F2. Pelo gráfico da Figura 6.6 observa-se que a alocação de banda se manteve estabilizada para o fluxo F1 durante todo o tempo de observação, ao passo que o fluxo F4, por exemplo, quando o aumento do tráfego atingiu 190%, teve a sua alocação de banda reduzida drasticamente. Este resultado já era esperado, pois em função da política do algoritmo PQ ocorreu descarte dos pacotes com menor prioridade, tendo o seu valor de pico nos pontos de maior congestionamento da rede, isto é, acima de 190%.

Tabela 6.4: Banda Alocada para os fluxos F1 e F2 com o aumento do tráfego.

Classes de Tráfego	**Largura de Banda Média em [Kbps]**	**Intervalo de Confiança**
Conversacional (F1)	14,0386	12,3778 / 15,6994
Streaming (F2)	28,0207	24,8155 / 31,2259

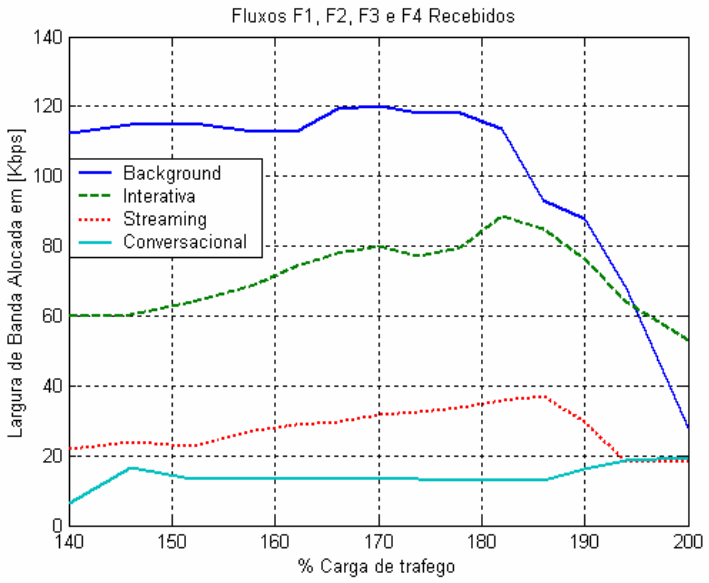

Figura 6.6: Banda alocada pelas classes F1, F2, F3 e F4 versus % carga de tráfego.

c) Experimento 3: Perda de pacotes para os fluxos F1, F2, F3 e F4.

As Figuras 6.7, 6.8 e 6.9 mostram a perda de pacotes apresentada por cada aplicação quando a carga de tráfego varia de 140% a 200%. A avaliação destas perdas foi caracterizada quando a carga de tráfego variou entre 140% e 200% além de seu valor previamente estabelecido.

A disciplina PQ faz o escalonamento de pacotes de acordo com a classe de precedência, onde a escolha de prioridade é feita conforme Tabela 6.2, ou seja, alta, média e baixa prioridade. O pacote é inserido na fila de acordo com sua classe de prioridade. Assim, os pacotes com classe mais baixa como *background* e interativa sofrerão maiores perdas para que se possa alocar mais banda às classes de maior precedência (ex: conversacional e *streaming*).

Como conseqüência disso, tem-se para estas classes uma perda muito pequena e praticamente constante principalmente nos pontos de alta intensidade da carga de tráfego na rede, ou seja, após 190%.

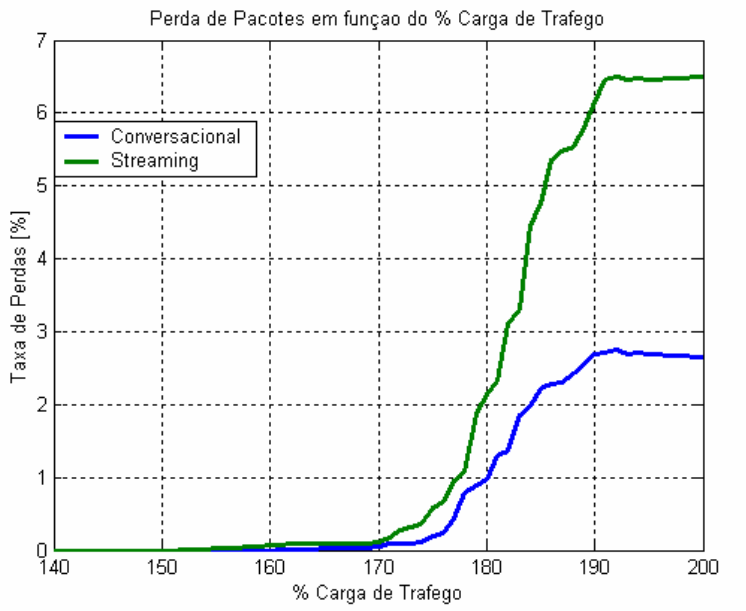

Figura 6.7: Perda de Pacotes versus % carga de tráfego (F1, F2).

Observa-se pela Figura 6.7 que as classes conversacional e *streaming* mantiveram praticamente constante a taxa de perda de pacotes nos pontos de alto tráfego, isto é, entre 193% e 200%, sendo que a taxa de perda para a classe *streaming* foi pior neste período apenas 4,23%. Contudo, para outras aplicações a perda de pacotes será tão alta quanto maior for a intensidade de tráfego. Este resultado se justifica em função da eficiência do mecanismo empregado baseado na arquitetura *DiffServ* e no algoritmo PQ implementados nos roteadores SGSN de ingresso ao *backbone* UMTS.

Pelas Figuras 6.7, 6.8 e 6.9 pode-se concluir que o mecanismo utilizado como alternativa para o provimento de QoS e pelos resultados obtidos por esta simulação, têm-se que o mesmo é ideal para aplicações de tempo real, conforme apresentado em seções anteriores.

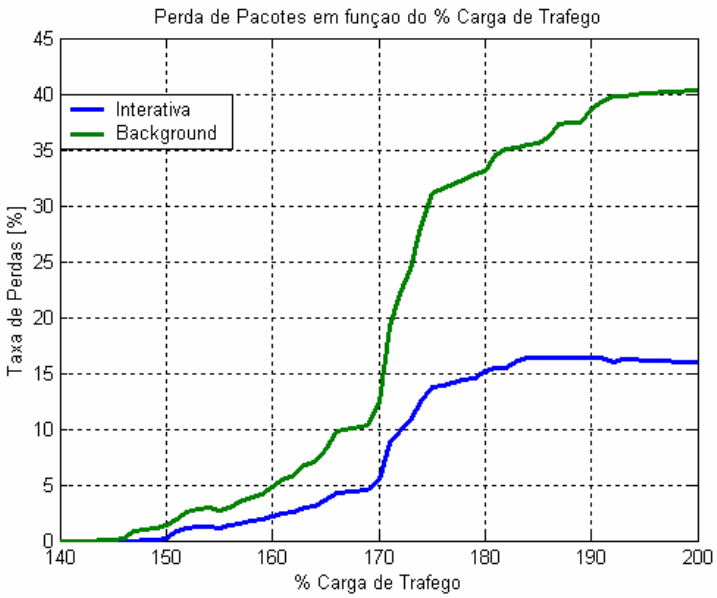

Figura 6.8: Perda de Pacotes versus % carga de tráfego (F3, F4).

Pela Figura 6.9 pode-se confirmar que, com o aumento da carga de tráfego ocorrem perdas de pacotes para todos os fluxos. Porém, para os fluxos cuja classe de precedência é baixa, as perdas são maiores. Caso da classe *background*, que possui classe de precedência semelhante

à *Best Effort*. Isto ocorre devido o descarte de pacotes excedente ser uma ação do algoritmo PQ ao longo do *backbone*, que faz o escalonamento destes pacotes em função do controle de admissão e classificação de pacotes implementada nos pontos de ingresso deste tipo de rede, o que pode ser observado pela Figura 6.9.

Pode-se verificar também pela Figura 6.9 que o mecanismo implementado através da arquitetura *DiffServ* e o algoritmo PQ no *backbone* UMTS protege todos os tráfegos (F1, F2, F3 e F4) de forma consistente e estabilizada durante todo o período observado, mesmo com o aumento da taxa de congestionamento na rede, isto é, entre 190% e 200%. Verifica-se também que as perdas de pacotes só ocorreram para intensidade de tráfego superior a 170%.

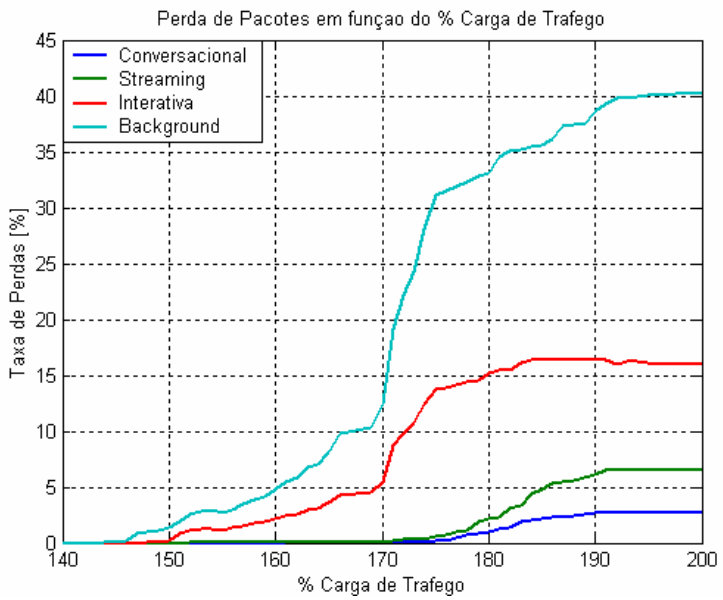

Figura 6.9: Perda de Pacotes versus % carga de tráfego (F1, F2, F3, F4).

d) Experimento 4: Variação de atraso (*Jitter*) para os fluxos F1, F2.

Este cenário apresenta variação de atraso dos pacotes, desde os pontos de ingresso à rede *backbone* UMTS (servidor SGSN) até sua entrega no final desta. Em relação aos resultados obtidos, verifica-se pela Figura 6.10 que o *jitter* , apesar de em termos relativos comportar-se

como esperado, apresentou resultados altos para o fluxo F2 nos pontos de alto tráfego, conforme apresentado em [66], o *jitter* não deve exceder a 10% do atraso de transmissão fim a fim. Entretanto para o fluxo F1 (conversacional) foi plenamente atendido durante todo o tempo observado mesmo com alta intensidade de tráfego.

Como o atraso de transmissão fim a fim no *backbone* UMTS, para cada tipo de tráfego, apresentou comportamento de acordo com a Tabela 6.2 e Figuras 6.4 e 6.5, o atraso total sofrido pelos pacotes dependeu apenas do tempo de espera na fila. Conclui-se que o resultado obtido e mostrado na Figura 6.10 comprova a eficiência do mecanismo baseado na arquitetura *DiffServ* e no algoritmo PQ quando implementados nos pontos de ingresso deste *backbone* o qual protege principalmente a classe conversacional.

Uma investigação possível seria a implementação de políticas de escalonamento que se adequasse aos fluxos que tiveram valores de *jitter* extrapolados, caso do fluxo F2 (Fig. 6.10 e Tabela 6.5) que teve seus valores fora do especificado apenas quando a intensidade de tráfego superou 180%.

Com relação aos resultados obtidos, pode-se observar, através da Figura 6.10, que o uso de *DiffServ* diminuiu o *jitter* para os tráfegos com maior precedência. Verifica-se que os fluxos F1 e F2 tiveram valores dentro do especificado até a intensidade de tráfego atingir 180%. Como o fluxo F1 estava protegido contra grandes atrasos em função das ações do mecanismo baseado no modelo *DiffServ* e o algoritmo PQ, observa-se que seus valores permaneceram abaixo de 1 ms durante todo o tempo de observação, salvo quando a intensidade de tráfego superou os 180%. Isto significa que o intervalo entre as chegadas dos pacotes tornou-se mais constante, refletindo que as políticas de conformação do tráfego adotadas foram coerentes quando aplicadas em redes *backbone* com tecnologia UMTS (Fig. 6.1). A Tabela 6.5 mostra os valores médio do *jitter* em ms para os fluxos F1 e F2 e o intervalo de confiança considerando um nível de confiança de 95%.

Tabela 6.5: *Jitter* Médio em ms para os fluxos F1 e F2 com o aumento do tráfego.

Classes de Tráfego	***Jitter* Médio [ms]**	**Intervalo de Confiança**
Conversacional (F1)	0,8416363	0,6551361 / 1,0281365
Streaming (F2)	3,3518181	2,6038827 / 4,0997535

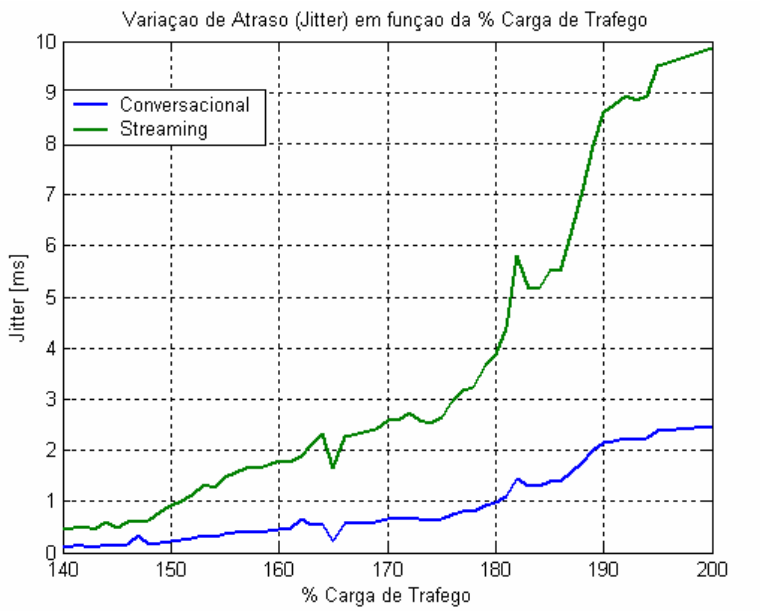

Figura 6.10: Variação de atraso (*Jitter*) versus % carga de tráfego (F1, F2).

6.6 Conclusões

Neste Capítulo realizou-se a apresentação e análise de resultados obtidos a partir da modelagem e simulação do mecanismo proposto para o provimento de QoS em redes *backbone* UMTS, a qual é baseada na arquitetura *DiffServ* e no algoritmo PQ.

Com este objetivo foram feitas simulações considerando os modelos apresentados em seções anteriores, avaliando parâmetros como perda de pacotes, atrasos, alocação de banda e

jitter, em modelos onde havia métodos que implementasse controle sobre o tráfego entrante e, o seu escalonamento ao longo do *backbone*.

Os resultados obtidos através dos experimentos, mostraram também que a solução proposta, foi capaz de prover uma qualidade de serviço satisfatória para as aplicações consideradas.

A arquitetura *DiffServ* se mostrou como uma boa alternativa para garantir um nível de QoS aceitável para redes *backbones* com tecnologia UMTS. Com base nas simulações realizadas verifica-se porém, que o sucesso de sua implementação depende também da escolha adequada do algoritmo de escalonamento e priorizar adequadamente o tráfego. Isto é, classificar o tráfego de acordo com as características particulares de cada tipo de aplicação como apresentado na seção 4.5 e Figura 6.3. Uma classificação inadequada pode priorizar excessivamente determinado tráfego afetando pontos cruciais dos demais.

Comparando os resultados obtidos e apresentados nos gráficos das Figuras 6.4, 6.7 e 6.10 e Tabelas 6.3, 6.4 e 6.5, verifica-se que estes resultados demonstram que a classe conversacional obteve melhor desempenho do que a classe *streaming*. Este resultado foi possível através da utilização da disciplina PQ no escalonamento das filas referentes a essas classes, dando maior prioridade à fila da classe conversacional devido a seus requisitos de QoS mais exigentes do que a classe *streaming*, como menor atraso, alocação de banda estabilizada, menor perda de pacotes e menor *jitter*. A classe interativa apresentou melhor desempenho do que a classe *background*. Este resultado foi possível através da eficiência da disciplina PQ no escalonamento das filas das quatro classes QoS, onde é dado maior prioridade à fila da classe interativa, por ser a classe referente ao tráfego que tem requisitos mais estritos de atraso, perda de pacotes e *jitter*. As classes interativa e *background* possuem filas com média e baixa prioridade, respectivamente, comparadas às classes *streaming* e conversacional, por apresentar requisitos mais flexíveis. No escalonamento PQ a classe

background recebe menor prioridade, pois representa as aplicações conhecidas como serviço de melhor esforço (*Best Effort*), e por esta razão não possui requisitos estritos quanto a atraso, banda utilizada, perda de pacotes e *jitter*, conforme apresentado nas Tabelas 6.1 e 6.2.

Por meio dos resultados obtidos pôde-se demonstrar que o mecanismo proposto tornou possível o uso de vários serviços em um *backbone* UMTS de acordo com seus requisitos de QoS, o que em muitos casos não seria possível sem um gerenciamento de QoS. As Tabelas e os gráficos mostrados nos quatro experimentos apresentando os resultados obtidos, demonstram que o uso do *DiffServ* com a configuração utilizada proporcionou melhor aproveitamento no escoamento do tráfego apresentado à rede, priorizando a entrega dos fluxos prioritários, garantindo menor taxa de perda de pacotes e menor atraso.

Observou-se, pelos resultados obtidos, que o comportamento do *backbone* UMTS manteve-se coerente durante todo o período de aumento da carga de tráfego. Isto indica que com a implementação do mecanismo baseado no modelo *DiffServ* e no algoritmo PQ, através do uso de escalonadores apropriados, métodos de marcação e conformação de tráfego e de políticas de descarte, pode se ter uma eficiente alternativa para o provimento de qualidade de serviço à diferentes classes de tráfego.

Capítulo 7

Conclusões Gerais

As redes móveis sem fio surgiram como uma nova revolução na sociedade da informação, permitindo atender a clientes que exigem cada vez mais um acesso o mais eficiente possível e com mobilidade. Atualmente, estes requisitos têm sido atendidos por diferentes soluções, conhecidas como redes IP e redes móveis sem fio. A convergência de ambas as soluções é um dos objetivos dos sistemas móveis de terceira e quarta geração (3G/4G) denominado pela 3GPP por UMTS, permitindo que as duas necessidades sejam atendidas pela mesma plataforma e através das novas gerações de dispositivos de acesso.

Um dos desafios das redes móveis sem fio é permitir que, cada vez mais, informações e serviços possam estar disponíveis aos usuários sem as barreiras impostas pela necessidade de conexões físicas com a rede e, sobretudo, permitindo acesso às informações a qualquer momento e em qualquer lugar do planeta com altas taxas de transmissão.

As redes 3G com tecnologia UMTS possibilitam taxas de dados de até 2 Mbps, que são consideravelmente mais altas que as taxas suportadas pelos sistemas 2G/2,5G, como o GSM. A evolução dos atuais sistemas 2,5G/3G e TDMA/IS-136, que compartilham comutação por pacotes e circuitos, para sistemas 4G que serão totalmente digitais e somente IP (*ALL IP*), atingirão taxas superiores a 100 Mbps.

Este trabalho abordou vários conceitos concernentes à evolução e objetivos dos sistemas móveis sem fio empregando a tecnologia proposta pela 3GPP para redes *backbone* UMTS. Além disso, destacou a integração da plataforma *Internet* com a rede móvel sem fio do ponto de vista da qualidade de serviço, que é um dos objetivos dos sistemas UMTS.

Um dos grandes desafios das operadoras de rede *backbone* é manter ao longo da rede um nível de Qualidade de Serviço (QoS) satisfatório. Contudo, considerando sistemas móveis sem fio, o desafio é maior ainda, devido à escassez do espectro de freqüência, diversidade dos requisitos de desempenho dos serviços e características dos enlaces sem fio quanto ao desvanecimento.

O problema para o provimento de QoS em redes *backbone* com tecnologia UMTS pode estar relacionado às seguintes restrinções: recursos limitados na interface aérea (UTRAN); o deslocamento entre redes (*roaming*), que pode acarretar grandes congestionamentos à entrada do *backbone*; garantias dos níveis de QoS desde os pontos de ingresso e egresso do *backbone* para os usuários de outras PDNs.

Antes que uma MS possa efetivar a comunicação com a rede de dados externa, o contexto PDP deve ser ativado e nesta fase o perfil de QoS é negociado com o *backbone* UMTS. Este *backbone* permite que sejam definidos perfis de tráfego para o qual o contexto pode se adequar (conversacional, *streaming*, interativa ou *background*), entre outros atributos que podem ser negociados. Contudo, nenhum mecanismo para prover QoS é definido nesta fase que possa dar suporte a esses perfis e garantir o seu funcionamento. QoS é considerado um dos atributos importantes de um contexto PDP e é negociado durante a sua ativação.

Uma das contribuições deste trabalho é uma proposta de mecanismo para o provimento de QoS em redes tipo *backbone* UMTS baseada no modelo de serviços diferenciados. Neste ambiente foi implementado o modelo *DiffServ*, que é uma das arquiteturas que têm sido propostas para prover qualidade de serviço nestes *backbones* baseadas em IP.

Através de modelagem e simulação de vários cenários com condições de tráfego e configurações diferentes, obtiveram-se resultados consistentes sobre o desempenho destas redes em função da implementação do referido mecanismo. Os resultados obtidos através de modelagem e simulação confirmaram que o mecanismo baseado na arquitetura *DiffServ* e no algoritmo PQ pôde diferenciar classes de tráfegos dentro do *backbone*, garantindo um nível de QoS satisfatório a cada uma das quatro classes de tráfego definidas pelo 3GPP.

A fim de comparar as seis disciplinas de escalonamento no ambiente de rede considerado, vários experimentos foram efetuados. Os resultados obtidos deram suporte para identificar qual algoritmo que mais se adequa à classe de tráfego conversacional, por ser esta mais sensível ao atraso. Para uma avaliação mais consistente dessas disciplinas, as mesmas foram avaliadas sob condições de congestionamento na rede provocado pelo aumento do tráfego através dos fluxos F2 e FTP até um limite de 200% da carga de tráfego. A avaliação do atraso fim a fim no *backbone* foi caracterizada quando a intensidade de tráfego variou entre 140% e 200%.

Os resultados obtidos por meio de simulação mostraram que a disciplina PQ oferece menores atrasos de forma consistente e estabilizada durante todo o período observado, mesmo com o aumento da taxa de congestionamento na rede.

Com exceção da disciplina PQ, observou-se que ocorreu um aumento considerável dos atrasos a partir de 175% de utilização da rede. Verificou-se também que a disciplina FIFO apresentou um aumento dos atrasos a partir de 160% de utilização da rede e intensificando a partir dos 180%. Este resultado já era esperado, uma vez que esta disciplina não possui nenhuma técnica de classificação e escalonamento de pacotes. Diante dos resultados obtidos, obteve-se a justificativa para a escolha do algoritmo PQ para a realização das simulações empregando a arquitetura *DiffServ* no *backbone* UMTS.

As fontes de tráfego escolhidas para a avaliação do mecanismo proposto foram uma combinação de todas as classes de tráfego UMTS, tais como: conversacional, *streaming*, interativa e *background*. Foram escolhidos estes modelos de tráfego e a estrutura *DiffServ,* sendo o PHB EF a opção adotada para acomodar a classe conversacional, por garantir parâmetros mais rígidos. As outras três classes de tráfego UMTS foram acomodadas pelos PHBs AF, que são mais flexíveis. A classe *streaming* correspondeu à classe AF1; a classe interativa à classe AF2, e a classe *background* correspondeu à classe (*Best Effort*).

Os resultados referentes ao atraso fim a fim no *backbone* para os fluxos F1 (Conversacional) e fluxo F2 (*Streaming*) foram obtidos quando a intensidade de tráfego e o fluxo FTP oferecida variaram entre 140% e 200% de sua taxa inicial. Em relação aos requisitos apresentados pelos fluxos quanto aos itens de atraso e largura de banda, verificou-se que o atraso fim a fim no *backbone* sofrido pelo tráfego gerado pelo fluxo F1 foi menor que os valores encontrados pelas fontes do fluxo F2. Esta característica está de acordo com o esperado, pois no escalonamento realizado por PQ a maior prioridade foi dada ao tráfego de classe conversacional, conforme atribuição adotada para esta classe.

Os resultados referentes a alocação de largura de banda para os fluxos de tráfego F1, F2, F3 e F4 foram obtidos quando a carga de tráfego e fluxo FTP no *backbone* UMTS variaram entre 140% e 200%. Verificou-se pelos gráficos apresentados que o modelo proposto foi capaz de alocar a banda requerida por cada uma das classes de tráfego UMTS, isto é, os fluxos F1 e F2 foram protegidos durante todo o tempo, mesmo nos pontos de alto tráfego. Verificou-se também, que nos pontos onde a carga de tráfego e o fluxo FTP variaram entre 150% e 190 % da sua taxa inicial, os fluxos, F3 e F4 tiveram a alocação de banda diminuída, ou seja, a capacidade da banda reservada para essas classes foi alocada pelos fluxos F1 e F2. Foi possível observar pelos gráficos apresentados neste cenário que a alocação de banda se manteve estabilizada para o fluxo F1 durante todo o tempo de observação, ao passo que o

fluxo F4, por exemplo, quando o aumento do tráfego atingiu 190%, teve a sua banda reduzida drasticamente. Este resultado já era esperado, pois em função da política do algoritmo PQ ocorreu descarte dos pacotes com menor prioridade, tendo o seu valor de pico nos pontos de maior congestionamento da rede, isto é, acima de 190%.

A avaliação das perdas de pacotes para os fluxos F1, F2, F3 e F4 foi caracterizada quando a intensidade de tráfego e o fluxo FTP oferecida variaram entre 140% e 200% de sua taxa previamente estabelecida.

Os gráficos referentes às perdas de pacotes mostraram que as classes conversacional e *streaming* mantiveram praticamente constante a taxa de perda de pacotes nos pontos de alto tráfego, isto é, entre 193% e 200%, sendo que a taxa de perda para a classe *streaming* foi pior neste período apenas 4,23%. Contudo, para outras aplicações, a perda de pacotes aumentou com o aumento da intensidade de tráfego interferente. Este resultado comprovou a eficiência do mecanismo proposto quando da utilização da arquitetura *DiffServ* e o algoritmo PQ nos roteadores SGSN de ingresso ao *backbone* UMTS.

Foi possível verificar por este experimento que o mecanismo utilizado baseado no modelo *DiffServ* implementado no *backbone* UMTS protegeu todos os tráfegos de forma estabilizada durante todo o período observado, mesmo com o aumento da taxa de congestionamento na rede, isto é, entre 190% e 200%. Conclui-se que a implementação do mecanismo proposto para o provimento de QoS e pelos resultados obtidos por esta simulação, que este mecanismo é ideal para aplicações de tempo real, conforme descrito em seções anteriores.

Os gráficos que mostraram os resultados da variação de atraso (*jitter*) dos pacotes, desde os pontos de ingresso ao *backbone* UMTS até sua entrega no final desta, demonstraram que o mecanismo proposto foi eficiente quando implementada em redes tipo *backbone* UMTS. Em relação aos resultados obtidos, verificou-se também que o *jitter*, apesar de em termos relativos comportar-se como esperado, apresentou resultados altos para o fluxo F2 nos pontos de alto

tráfego. Entretanto para o fluxo F1 os resultados foram satisfatórios durante todo o tempo observado mesmo com alta intensidade do tráfego.

Através dos resultados obtidos neste experimento, observou-se, pelos gráficos, que o uso de *DiffServ* diminuiu o *jitter* para os tráfegos com maior precedência. Verificou-se também que os aplicativos F1 e F2 tiveram valores dentro do especificado até a intensidade de tráfego atingir 180%. Como o fluxo F1 foi protegido contra grandes atrasos através deste método, nota-se que seus valores permaneceram abaixo de 1 ms durante todo o tempo de observação, salvo quando a intensidade de tráfego superou os 180%. Isto significa que o intervalo entre as chegadas dos pacotes tornou-se mais constante, refletindo que as políticas de conformação do tráfego adotadas foram coerentes para o modelo de simulação adotado.

Assim, conclui-se que o mecanismo proposto atendeu efetivamente o objetivo inicial desta dissertação, a qual se refere ao desenvolvimento de um trabalho para ampliar e enriquecer as condições de análise das alternativas para provimento de QoS dentro do ambiente de *backbone* UMTS.

Como contribuição deste trabalho, propõe-se o referido mecanismo como alternativa para se obter QoS em redes *backbone* com tecnologia UMTS. O mecanismo proposto emprega o modelo *DiffServ* juntamente com o algoritmo de escalonamento por prioridade PQ, o qual foi demonstrado através de simulação a sua eficiência em proteger aplicativos sensíveis a atraso e *jitter*. Desta forma, com este mecanismo torna-se viável o emprego de todas as classes de tráfego UMTS proposta pelo grupo 3GPP em *backbones* UMTS, concedendo a cada uma delas a QoS requerida.

Como sugestão para trabalhos futuros, visando a continuidade deste trabalho, propõe-se ampliar o mecanismo proposto, incluindo o controle de admissão no *backbone* UMTS que envolve a autorização, contrato de serviço (SLA) e políticas do operador. Este novo mecanismo seria baseado em diferenciação de serviços e políticas de escalonamento que

permitissem a associação de pesos a cada classe de tráfego. Desta forma, haveria a possibilidade de se diferenciar os fluxos, através de um peso atribuído a cada um. Um esquema baseado em pesos poderia ser do tipo: WFQ, WRR e DWRR. Assim, seria possível avaliar não só o controle de admissão dos pacotes, mas também o escalonamento e escoamento dos mesmos ao longo deste *backbone*.

Outra sugestão seria implementação de disciplina de escalonamento de pacotes tal como, WFQ ou WRR, onde se pode adequar os seus pesos conforme aplicação e implementar a arquitetura *DiffServ* juntamente com MPLS no *backbone* UMTS.

Sugere-se ainda aumentar o número de estações móveis (MS) em várias células e avaliar o desempenho do *backbone* UMTS como um todo frente à mobilidade e o processo de *roaming* empregando o protocolo MIPv6.

Por fim, o trabalho apresentado alcançou os objetivos propostos nesta dissertação e poderá servir de base para o desenvolvimento de futuras investigações.

Referências Bibliográficas

[1] 3GPP - http://www.3gpp.org/specs/specs.htm. Acessado em Novembro de 2005.

[2] 3GPP2 - http://www.3gpp2.org/Public_html/Misc/v&vindex.cfm. Acessado em Novembro de 2005.

[3] CLAPTON, A. and GROVES, I., "Third Generation Mobile Systems", BT Tecnology Journal, July 1996, Vol. 14, No. 3, pg. 115-122.

[4] ERICSSON, Redes GSM / GPRS / UMTS, disponível em: http://www.3gamericas.org/portuguese/Technology_Center/WhitePapers/. Acessado em Dezembro de 2005.

[5] DIAS, K. L. e SADOK, D. F. H., *Internet* Móvel: Tecnologias, Aplicações e QoS. XIX Livro Texto dos Minicursos – SBRC'2001, Maio, 2001.

[6] IETF NEMO WG, http://www.ietf.org/html.charters/nemo-charter.html, acessado em Dezembro de 2005.

[7] CLARK, M. P., "Wireless Access Network – Fixed Wireless Access and WLL Networks", John Wiley and Sons, New York – USA, 2000.

[8] d'ÁVILA, C. K., "Telefonia Móvel Celular", CEDETEC, INATEL, Apostila Vol. 2, 1998.

[9] HARTE, L.; SMITH, A. and JACOBS, C., IS-136 TDMA Technology , Economics and Services Artech House, 1998.

[10] BATES, R. J. (BUD), *GPRS General Packet Radio Service*. USA : McGraw- Hill, 2002.

[11] ETSI TS 144 065 V5.0.0 (2002-06), "Digital cellular telecommunications system (Phase 2+); Mobile Station (MS) - Serving (GPRS); Support Node (SGSN); Subnetwork Dependent Convergence Protocol (SNDCP)" (3GPP TS 44.065 version 5.0.0 Release 5).

[12] ETSI EN 301 113 V6.3.1 (2000-11), "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description; Stage 1" (GSM 02.60 version 6.3.1 Release 1997)

[13] ETSI TS 144 064 V5.1.0 (2002-03), "Digital cellular telecommunications system (Phase 2+); Mobile Station (MS) - Serving (GPRS); Support Node (MS-SGSN); Logical Link Control (LLC) Layer Specification ".(3GPP TS 44.064 version 5.1.0 Release 5).

[14] DIXIT, S.; GUO, Y. and ANTONIOU, Z., "Resource Management and Quality of Service in Third-Generation Wireless Networks", *IEEE Communications Magazine,* Frebuary 2002, vol 39, no. 2, pp. 125-33.

[15] ETSI EN 301 344 V6.7.1 (2000-09), "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description; Stage 2" (GSM 03.60 version 6.7.1 Release 1997).

[16] ETSI TS 101 347 V7.10.0 (2002-12), "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); GPRS Tunnelling Protocol (GTP) across the GN and Interface". (3GPP TS 09.60 version 7.10.0 Release 1998).

[17] MEHROTRA, A., *GSM System Engineering*. Mobile Communications Series. London : Artech House Publishers, 1986.

[18] CUEVAS, A.; SERRANO, P.; BERNARDOS, C. J.; MORENO, J. I.; JAEHNERT, J.; HYUNG-WOO, K.; ZHOU, J.; GOMES, D.; GONÇALVES, P. and AGUIAR, R., Field Evaluation of a 4G True-IP network. In *IST Mobile Summit 2004*, Lyon, France, 2004.

[19] MARQUES, V. and AGUIAR, R., ET AL., An IP-Based QoS Architecture for 4G Operator Scenarios. IEEE Wireless Communications Magazine, vol. 41, issue 3, March, 2003, pp. 120-124.

[20] VIDALES, P., CHAKRAVORTY, R. and POLICRONIADES, C. P., A Policy-based Solution for Future 4G devices. In *Proceedings of 5th International Workshop on Policies for Distributed Systems and Networks (IEEE POLICY 2004)*, New York, United States, June 2004.

[21] KARAGIANNIS, G. and HEIJENK, G., *QoS in GPRS*. CTIT Technical Report Series, n. 00-29, ISSN 1381-3625. December, 2000.

[22] RYSAVY, P., Aumento da Capacidade de Transmissão de Dados para a Evolução da Tecnologia GSM para a UMTS, Livro Branco para as Américas, 19/11/2002.

[23] NOKIA, Introducing Mobile IPv6 in 2G and 3G Mobile Networks, September 2001.

[24] AGHAREBPARAST, F. and LEUNG, V., QoS Support in the UMTS/GPRS Backbone Network Using DiffServ, Dept. of Electrical and Comp. Eng., Univ. of British Columbia, Canada, 2002.

[25] KUROSE, J. F. e ROSS, K. W., Redes de Computadores e a *Internet*: Uma Nova Abordagem, Addison Wesley, 1 ed., São Paulo, 2003.

[26] *Universal Mobile Telecommunications System (UMTS); Services & service capabilities*. ETSI TS 122 105 (2002-06). 3GPP TS 22.105, v. 5.2.0, Release 5.

[27] Internet Engineering Task Force, http://www.ietf.org. Acessado em: Dezembro de 2005.

[28] BRADEN, R.; CLARCK, D. and SHENKER, S., Integrated Services in the Internet Architecture: an Overview, RFC 1633, June 1994.

[29] ROSEN, E.; VISWANATHAN, A. and CALLON, R., (2001). *Multiprotocol Label Switching Architecture*. RFC 3031, IETF.

[30] BLAKE, S.; BLACK, D.; CARLSON, M.; DAVIES, E.; WANG, Z. and WEISS, W., An Architecture for Differentiated Services, RFC 2475, December 1998.

[31] BRADEN, R. Et AL., Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification, RFC 2205, September 1997.

[32] CRAWLEY, E.; NAIR, R. and RAJAGOPALAN, B., A Framework for QoS-Based Routing in the Internet, RFC 2386, August 1998.

[33] XIAO, X. and NI, L. M., Internet QoS: A Big Picture, IEEE Network, March/April 1999.

[34] EL-GENDY, M. A.; BOSE, A. and SHIN, K. G., Evolution of the Internet and support for soft real-time applications. Proceedings of the IEEE 91(7): pp. 1086-1104, June, 2003.

[35] KAMIENSKI, C. A. e SADOK, D., Qualidade de Serviço na Internet, 18º SBRC (Simpósio Brasileiro de Redes de Computadores), Belo Horizonte/ MG, maio de 2000. (http://www.cin.ufpe.br/~cak/publications). Acessado em Setembro de 2005.

[36] WROCLAWSKI, J., Specification of the Controlled-Load Network Element Service, RFC 2211, September, 1997.

[37] RFC 3290, An Informal Management Model for Diffserv Routers, Y. Bernet, S. Blake, D. Grossman, A. Smith, May 2002.

[38] CARPENTER, B. E. and NICHOLS, K., Differentiated Services in the Internet, IEEE Proceedings, volume 90, (no 9), pages 1479-94 in 2002.

[39] NICHOLS, K.; BLAKE, S.; BAKER, F. and BLACK, D., (1998). *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*. RFC 2474, IETF.

[40] KAMIENSKI, C. A., An Architecture for Providing End-to-End -based Advanced Services in the Internet, PhD. Thesis, February 2003.(disponível em: http://www.cin.ufpe.br/~cak/publications). Acessado em Setembro de 2005.

[41] JACOBSON, V.; NICHOLS, K. and PODURI, K., An Expedited Forwarding PHB, RFC 2598, June 1999.

[42] HEINAMEN, J. ET. AL., Assured Forwarding PHB Group, RFC 2597, June 1999.

[43] RFC 3496 Protocol Extension for Support of Asynchronous Transfer Mode(ATM)Service Class-aware Multiprotocol Label Switching (MPLS) Traffic Engineering, A. G. Malis, T. Hsiao, March 2003.

[44] KAMIENSKI, C. A. and SADOK, D., Engenharia de Tráfego em uma Rede de Serviços Diferenciados, 18o SBRC (Simpósio Brasileiro de Redes de Computadores), Belo Horizonte/MG, Maio/2000. (disponível em: http://www.cin.ufpe.br/~cak/publications). Acessado em Setembro de 2005.

[45] WANG, Z., "Internet QoS: Architecture and Mechanisms for Quality of Service", Morgan Kaufmann publishers, New York, 2001.

[46] SIEMENS MOBILE, Wireless LAN, disponível em: www.siemens-mobile.com, acessado em Dezembro, 2005;

[47] ARMITAGE, G., Quality of Service in IP Networks – Foundations for a Multi-Service Internet, MacMillan Technical Publishing, USA, 2000.

[48] OLIVEIRA, C.; KIM, J. B. and SUDA, T., An Adaptive Bandwidth Reservation Scheme for High Speed Multimedia Wireless Networks. *IEEE Journal on Selected Areas in Communications*, USA, v. 16, p. 858-874, August, 1998.

[49] KESHAV, S., An Engineering Approach to Computer Networking, ed. 1997.

[50] CHUCK, S., Supporting Differentiated Service Classes: Queue Scheduling Disciplines, Juniper Networks, USA, 2001.

[51] WISELY, D.; EARDLEY P.; BURNESS L., *IP for 3G*. England: John Wiley & Sons, 2002.

[52] "Enabling UMTS/Third Generation Services and Application", No. 11 Report from UMTS Forum, October 2000.

[53] GRONBAEK, I., IP QoS Bearer Service Elements for the Converging Fixed and Mobile 3G.IP Network, *IEEE Communications Magazine,* 2000.

[54] ITU - http://www.itu.int/home/imt.html . Acessado em Dezembro de 2005.

[55] 3GPP, End-to-End Quality of Service (QoS) Concept and Architecture, TS 23.207 V6.3.0 (2004-06).

[56] B3G/4G, Applications and Services. IEEE Wireless Communications, Outubro, 2004,Vol. 11 Nº 5.

[57] PUUSKARI, M., Quality of Service Framework in GPRS and Evolution towards UMTS, 3rd European Personal Mobile Communication Conference, March, 1999.

[58] GSM 03 60, ETSI Technical Specification: General Packet Radio Service (GPRS); Service Description, 1997.

[59] FALL, K. and VARADHAN K., Network Simulator – NS (version 2) notes and documentation, tech. Rep. The VINT Project, July 2002. URL: http://www.ash.cs.berkeley.edu/ns/. Acessado em: Agosto de 2005.

[60] ANDREOZZI, S., DiffServ Simulations using the Network Simulator: requirements, issues and solutions, Universitá degli Studi di Pisa, Italia, 2001.

[61] TODINI, A. and VACIRCA, F., UMTS Manual, disponível em http://www.net.infocom.uniroma1.it/reti_files/reti_downloads1.php, November, 2003. Acessado em Agosto de 2005.

[62] VEGESNA, S., IP Quality of Service, (Cisco Networking Fundamentals), January 2001

[63] KIM, M. K. and LEE, H. S., A case study of policy-based QoS management in 3Gnetworks, VTC2003 Spring, Session 11G, April, 2003.

[64] The VINT Project. *Nam: Network Animator*. UC Berkeley. Disponível em: http://www.isi.edu/nsnam/nam. Acessado em Agosto 2005.

[65] MANIATIS, S.; GRECAS C. and VERIERIS, I., End-to-end quality of service issues over next generation mobile internet, *Symp. Commun. & Vehricular Technology*, pp. 150-154, Leuven, Belgium, Octuber, 2000.

[66] PIEPRZYCKI, A. and JAJSZCZYK, A., Some Useful Formulas for UMTS Radio Network Planning and Optimisation, Dept. of Computer Science, TU PWSZ Tarnów, Dept. of Telecommunications, AGH University of Technology, Kraków, Poland, 2002.

[67] ETSI TS 101 349 V8.19.0 (2003-07), "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Mobile Station (MS) – Base Station System (BSS) interface; Radio Link Control/Medium Access Control (RLC/MAC) protocol", (3GPP TS 04.60 version 8.19.0 Release 1999).

[68] ETSI TS 148 016 V5.2.0 (2003-07), "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); (BSS) – Serving GPRS Support Node (SGSN) Interface; Network Service". (3GPP TS 148 016 version 5.2.0 Release 5).

[69] J3GPP, IP Transport in UTRAN Work Task Technical Report, TR 25.933, R-00, RAN WG3.

[70] 3GPP - http://www.3gpp.org/specs/numbering.htm. Acessado em Dezembro de 2005.

[71] UMTS - Fórum - http://www.umts-forum.org/servlet/dycon/ztumts/umts/Live/en/umts/Resources_Papers_index. Acessado em Dezembro de 2005.

[72] 3G Américas - http://www.3gamericas.org/English/Technology_Center/Articles. Acessado em Dezembro de 2005.

[73] Rec. ITU-R M.1225 "Guidelines for Evaluation of Radio Transmission Technologies (RTT) for IMT-2000", 1997.

[74] ITU-R, Circular Letter: Request for Submission of Candidate Radio Transmission Technologies (RTTs) for the IMT-2000/FPLMTS Radio Interface, Circular Letter 8/LCCE/47.

[75] Rec. ITU-R M.687-2, International Móbile Telecommunications-2000 (IMT-2000), 1997.

[76] Radiocommunication Study Group 8, Draft New Recommendation ITU-R [IMT.RKEY] Key Characteristics for the IMT-2000 Radio Interfaces, May 2000.

[77] SINGH, M. and SINGH, M., 3G Wireless With Respect to IMT-2000 and Beyond, *ITU documentation on-line* (http://www.itu.int), February 2000.

[78] Agência Nacional de Telecomunicações - ANATEL, http://www.anatel.gov.br, acessado em Dezembro de 2005.

[79] JOHNSON, D. B.; PERKINS C. E. and ARKKO, J., Mobility Support in IPv6 (RFC 3775) http://www.ieft.org, June 2004, acessado em Dezembro de 2005.

Trabalhos Aceitos para Publicação pelo Autor:

[80] ALECRIM, P. D. e GUARDIEIR0, P. R. Uma Proposta de Mecanismo para Controle de Admissão para a Rede Backbone UMTS empregando Serviços Diferênciados. Aceito para publicação na III Escola Regional de Redes de Computadores (ERRC 2005), Santa Cruz do Sul- RS, Agosto 2005.

[81] ALECRIM, P. D. and GUARDIEIR0, P. R. A Mechanism for Admission Control of the Input Traffic to the UMTS Backbone Network. Accepted for publication in IEEE Malaysia International Conference on Communications and IEEE International Conference on Networks (MICC & ICON 2005). Malaysia, november, 2005.

[82] ALECRIM, P. D. and GUARDIEIR0, P. R. A Mechanism for Admission Control of the Input Traffic to the UMTS Backbone Network. Accepted for publication in 4th International Information and Telecommunication Technologies Symposium (I2TS – Federal University of Santa Catarina – UFSC, Florianópolis, SC - Brazil), 2005.

[83] ALECRIM, P. D. and GUARDIEIR0, P. R. A Mechanism for Admission Control of the Input Traffic to the UMTS Backbone Network. Accepted for publication in Advanced International Conference on Telecommunications (AICT 2006) in Guadeloupe, French Caribbean, Frebuary 2006.