

FÁTIMA DE LIMA PROCÓPIO DUARTE FIGUEIREDO

**DIFFMOBIL - UMA ARQUITETURA DE  
QUALIDADE DE SERVIÇO FIM-A-FIM EM  
REDES GPRS**

Belo Horizonte

14 de dezembro de 2004

FÁTIMA DE LIMA PROCÓPIO DUARTE FIGUEIREDO

**DIFFMOBIL - UMA ARQUITETURA DE  
QUALIDADE DE SERVIÇO FIM-A-FIM EM  
REDES GPRS**

Tese apresentada ao Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Doutor em Ciência da Computação.

Belo Horizonte



# Resumo

O crescente interesse pelo acesso à informação em qualquer lugar e a qualquer momento, associado ao desenvolvimento de aparelhos portáteis mais modernos, vem impulsionando a evolução das redes móveis. Esta evolução acontece de forma gradativa, através de diferentes gerações de tecnologias de telefonia celular: 1G, 2G, 2.XG, 3G, 4G. A cada geração são alcançadas capacidades mais altas de transmissão de voz, dados e imagens. O GPRS (*General Packet Radio Services*) é uma tecnologia 2.5G, que introduz comutação por pacotes em redes GSM. Aplicações multimídia e de tempo real necessitam de garantias de qualidade de serviço, tais como baixos tempo de resposta e confiabilidade na entrega de pacotes. Existem atualmente soluções que garantem qualidade de serviço em alguns segmentos GPRS, mas não existe uma solução que garanta níveis de QoS fim-a-fim. Algumas soluções de QoS da Internet, tais como Intserv, DiffServ e MPLS são adotadas como modelo e adaptadas às necessidades do núcleo GPRS. Este trabalho propõe uma arquitetura de QoS denominada DiffMobil, para redes GPRS. Para o núcleo IP da rede, a arquitetura DiffMobil adota o DiffServ. Módulos de negociação de QoS, diagnóstico de rede, controle de admissão e controle de fluxo são implementados para garantir QoS no enlace sem fio. Todos os módulos foram projetados a partir de medidas de desempenho e saturação de uma rede GPRS, obtidas através de simulações, dando prioridade de transmissão aos usuários e às aplicações de classes mais altas. Esta solução usa as mesmas classes da arquitetura DiffServ. Com isto, há no DiffMobil uma integração entre as soluções de QoS do núcleo e da inter-

face aérea, garantido qualidade de serviço fim-a-fim. Medidas de desempenho com e sem o uso de DiffMobil são mostradas para um ambiente GPRS simulado através da ferramenta ns-2 (*Network Simulator*). Os resultados mostram que o DiffMobil oferece às aplicações de classes de prioridades mais altas garantias de qualidade de serviço, quando a rede está congestionada, tanto na interface aérea quanto na rede de núcleo, sendo, portanto, uma arquitetura fim-a-fim.

# Abstract

The increasing interest on the possibility of accessing information anytime, anywhere, associated with the development of modern portable equipments, is stimulating the mobile network evolution. This is a gradual evolution, from the first mobile network generation, 1G, towards the fourth generation, 4G, passing through 2G, 2.XG, 3G. Higher video, voice and image transmission capacities are reached on each generation. GSM *Global System for Mobile Communications* is a 2G technology, while GPRS (*General Packet Radio Services*) is a 2.5G packet oriented technology. New demands regarding bandwidth and response time are generated through multimedia and real time applications. No end-to-end QoS (Quality of Service) guarantees are provided by GPRS operators nowadays. There are QoS solutions for some GPRS network segments, but there is no solution for the global environment. This work proposes an end-to-end GPRS QoS architecture, called DiffMobil. Some known Internet solutions like IntServ, DiffServ and MPLS are adopted as models and modified when necessary to the GPRS IP core. But this is a partial solution, since it doesn't consider the other network segments like the air interface. To the IP core, DiffMobil uses DiffServ. SLA negotiation, network diagnosis, admission control and congestion control modules were implemented to guarantee QoS in the air interface. Based on some GPRS network performance metrics obtained through simulations, and on the DiffServ QoS classes, these modules give priority to the conversational class of applications. So, there is an integration between the air interface and the core QoS solutions in the DiffMobil,

guaranting the end-to-end quality of service. Performance metrics obtained through simulations are shown when using DiffMobil and when using best-effort. DiffMobil offers better results for higher priority application classes when the network is congested either in the air interface or in the core.

*Dedico este trabalho ao meu grande amor, meu marido Luiz, e ao nosso bebê, Lúvia.*



# Agradecimentos

Agradeço ao Professor Antonio Alfredo pela orientação, pela compreensão em todos os momentos difíceis, pela oportunidade e condições de trabalho por ele proporcionadas. Conheço poucas pessoas tão humanas quanto meu orientador!

Agradeço à minha mãe por minha formação intelectual e moral. Este trabalho é mais um fruto do esforço da vida dela. Agradeço a ela pelo amor incondicional, pela dedicação integral, pela vida. Agradeço à minha irmã Ângela, por ser sempre minha amiga, pelo companheirismo e pela compreensão. Agradeço ao meu irmão Elias, pela orientação que sempre me deu, durante o doutorado e durante toda a minha vida, por ter aberto portas para mim, sendo sempre meu amigo e meu modelo.

Agradeço ao meu marido, Luiz, pelo incentivo, pelo apoio, pelo amor, por estar sempre ao meu lado. Agradeço ao Sr. Hilário, à D. Lúcia, à Maria Lúcia, à Maria Helena e a suas famílias pela acolhida carinhosa e pelo apoio.

Agradeço ao Vinícius Rosalen, ao Flávio Almeida, ao Augusto Pinho Gomes, ao Cláudio Márcio de Souza, ao Thiago H. de Paula pela colaboração efetiva, pelo trabalho, pela amizade. Ao Luiz Humberto, agradeço pelos debates técnicos.

Agradeço a todos os professores e funcionários do DCC pela convivência harmoniosa durante o período do meu doutorado.

Agradeço à Capes pelo apoio financeiro, sem o qual a realização deste trabalho seria impossível.

# Sumário

<b>Lista de Figuras</b>	<b>xii</b>
<b>Lista de Tabelas</b>	<b>xii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Descrição do Problema . . . . .	4
1.3 Trabalho Proposto . . . . .	6
1.4 Organização do texto . . . . .	9
<b>2 Fundamentos Teóricos de GPRS</b>	<b>10</b>
2.1 Características Gerais de uma Rede GPRS . . . . .	10
2.2 Gerenciamento de Sessão e Mobilidade no GPRS . . . . .	14
2.3 Pilha de Protocolos do GPRS . . . . .	16
2.4 Transmissão de Pacotes no Sentido <i>Downlink</i> . . . . .	20
2.4.1 Transmissão no SGSN . . . . .	21
2.4.2 Transmissão no BSS . . . . .	22
2.4.3 Transmissão na Interface Aérea . . . . .	22
2.5 A Interface Aérea do GPRS . . . . .	23
2.5.1 Canais Lógicos . . . . .	23
2.5.2 Mapeamento de Canais Lógicos em Canais Físicos . . . . .	24
2.5.3 Transferência de Quadros no Sentido <i>Uplink</i> . . . . .	25
2.5.4 Alocação Dinâmica ou Fixa de Blocos de Rádio . . . . .	25
2.5.5 Transferência de Blocos RLC . . . . .	26
2.5.6 Transferência de Quadros no Sentido <i>Downlink</i> . . . . .	26
<b>3 Fundamentos Teóricos de QoS e Revisão Bibliográfica</b>	<b>28</b>
3.1 Exemplos de Soluções de QoS . . . . .	30
3.1.1 IntServ . . . . .	30
3.1.2 DiffServ . . . . .	31
3.1.3 Especificação de QoS pelo 3GPP para Redes UMTS . . . . .	35
3.1.4 Soluções de QoS para GPRS . . . . .	37
3.2 Trabalhos Relacionados . . . . .	42
3.2.1 Solução Proposta por Stuckmann . . . . .	42

---

3.2.2	Solução Proposta por Oliver . . . . .	44
3.3	Revisão Bibliográfica Complementar . . . . .	49
3.4	Resumo do Capítulo . . . . .	54
<b>4</b>	<b>A Arquitetura DiffMobil</b>	<b>56</b>
4.1	Descrição da Arquitetura . . . . .	56
4.2	Política Implementada pelo DiffMobil no Núcleo IP . . . . .	63
4.3	Política Implementada pelo DiffMobil na Interface Aérea . . . . .	68
4.3.1	<i>Downlink</i> : Diferenciação de Serviços Fim-a-Fim . . . . .	68
4.3.2	<i>Uplink</i> . . . . .	71
<b>5</b>	<b>Simulações, Resultados e Avaliação da Arquitetura</b>	<b>77</b>
5.1	Modelo de Simulação . . . . .	78
5.2	DiffServ no Núcleo IP . . . . .	79
5.3	Controle de Admissão Baseado em Guarda e Diagnóstico da Rede . . .	82
5.4	CAC: Controle de Admissão Baseado em Guarda . . . . .	83
5.4.1	Atraso Médio no tempo, para diferentes tipos de tráfego . . . .	87
5.4.2	Variação do Atraso Médio no Tempo . . . . .	87
5.4.3	Vazão Média no tempo, para diferentes tipos de tráfego . . . . .	88
5.4.4	Utilização da Rede, para diferentes tipos de tráfego . . . . .	90
5.5	Módulo de Diagnóstico . . . . .	91
5.5.1	Atraso Médio no tempo, para diferentes tipos de tráfego . . . .	92
5.5.2	Variação do Atraso Médio no Tempo . . . . .	93
5.5.3	Vazão Média no tempo, para diferentes tipos de tráfego . . . . .	93
5.5.4	Utilização da Rede, para diferentes tipos de tráfego . . . . .	94
5.6	Módulo de Controle de Fluxo . . . . .	96
5.6.1	Atraso Médio no tempo, para diferentes tipos de tráfego . . . .	96
5.6.2	Variação do Atraso Médio no Tempo . . . . .	97
5.6.3	Vazão Média no tempo, para diferentes tipos de tráfego . . . . .	97
5.6.4	Utilização do Canal, para diferentes tipos de tráfego . . . . .	98
5.7	QoS na Interface Aérea e no Núcleo IP . . . . .	99
5.8	Conclusões . . . . .	100
<b>6</b>	<b>Conclusões</b>	<b>101</b>

# Lista de Figuras

2.1	Elementos do GPRS . . . . .	11
2.2	Estados de uma unidade móvel GPRS . . . . .	14
2.3	Pilha de Protocolos do GPRS . . . . .	20
3.1	Arquitetura de QoS para Redes UMTS [3] . . . . .	36
4.1	Arquitetura DiffMobil . . . . .	57
4.2	Diagrama UML de Caso de Uso DiffMobil . . . . .	58
4.3	Diagrama UML de Dependência entre os Módulos DiffMobil . . . . .	58
4.4	Vazão Média $\times$ Taxa de Geração de Tráfego . . . . .	61
4.5	Atraso Médio por pacote $\times$ Taxa de Geração de Tráfego . . . . .	61
4.6	Proposta de adaptação DiffServ ao núcleo IP GPRS . . . . .	65
4.7	Serviços do BSSGP . . . . .	70
5.1	Cenário GPRS Modelado . . . . .	79
5.2	Alocação de Banda $\times$ Variação % de Carga Background sem DiffServ . . . . .	80
5.3	Alocação de Banda $\times$ Variação % de Carga Background sem DiffServ . . . . .	81
5.4	Eventos da Simulação sem CAC . . . . .	85
5.5	Eventos da Simulação com CAC . . . . .	85
5.6	Atraso Médio sem CAC . . . . .	87
5.7	Atraso Médio com CAC . . . . .	88
5.8	Variação do Atraso Médio sem CAC . . . . .	89
5.9	Variação do Atraso Médio com CAC . . . . .	89
5.10	Vazão sem CAC . . . . .	90
5.11	Vazão com CAC . . . . .	90
5.12	Utilização de Canal sem CAC . . . . .	91
5.13	Utilização de Canal com CAC . . . . .	92
5.14	Atraso Médio com Diagnóstico . . . . .	93
5.15	Variação do Atraso Médio com Diagnóstico . . . . .	94
5.16	Vazão com Diagnóstico . . . . .	95
5.17	Utilização de Canal com Diagnóstico . . . . .	95
5.18	Atraso Médio com Controle de Fluxo . . . . .	96
5.19	Variação do Atraso Médio com Controle de Fluxo . . . . .	97
5.20	Vazão com Controle de Fluxo . . . . .	98
5.21	Utilização de Canal com Controle de Fluxo . . . . .	99

# Lista de Tabelas

3.1	PHB AF – quatro classes de tráfego independentes . . . . .	34
3.2	PHB Padrão – Best Effort . . . . .	34
3.3	PHB EF – Tráfego <i>premium</i> . . . . .	34
3.4	Mapeamento de Classes de QoS do UMTS em Classes PHB do Diffserv [67]	44
4.1	Mapeamento de Classes de QoS na Arquitetura DiffMobil . . . . .	63
5.1	Comparação entre a Arquitetura DiffMobil, Melhor Esforço e o AP-CAC de Oliver . . . . .	86
6.1	Mapeamento de Classes de QoS na Arquitetura DiffMobil . . . . .	104
6.2	Comparação entre a Arquitetura DiffMobil e as soluções de Oliver Stuckman . . . . .	104

# Capítulo 1

## Introdução

### 1.1 Motivação

Consultar informações pessoais em qualquer lugar ou trabalhar onde quer que se esteja são as vantagens proporcionadas pela computação móvel. O mundo da computação móvel inclui equipamentos portáteis (aparelhos celulares, *laptops*, *notebooks*, *palmtops* etc), *links* fixos de microondas, LANs sem fio, dados sobre redes de telefonia celular, WANs sem fio e canais de comunicação sem fio como enlaces de satélites, redes digitais, redes de *paging*, difusão infravermelha e comunicações baseadas em *laser*. É possível imaginar aplicações direcionadas a um mercado vertical, que abrange negócios específicos de determinadas empresas ou indústrias, e existem aplicações voltadas a um mercado horizontal, no qual um amplo espectro de consumidores e negócios utilizam comunicação móvel. O mercado horizontal cresce à medida que as tecnologias de redes sem fio tornam-se mais baratas, mais fáceis de usar e oferecem um maior número de aplicações práticas [64], com qualidade de serviço satisfatória.

O sistema sem fio mais utilizado é o sistema móvel celular (SMC) [52]. A primeira geração (1G) de sistemas celulares analógicos utilizava tecnologia FDMA (*Frequency Division Multiplexing Access*). O TDMA (*Time Division Multiplexing Access*) e o

CDMA (*Code Division Multiplexing Access*) são as tecnologias que estão associadas à segunda geração de celulares (2G) [53]. O GSM (*Global System for Mobile Communications*), de tecnologia TDMA, é um sistema padronizado na Europa e utilizado hoje em todo o mundo. A limitação de uso exclusivo de um canal por um usuário, mesmo que ele não esteja transmitindo, tornou o GSM (*Global System for Mobile Communications*) e as demais tecnologias de segunda geração (2G) insatisfatórias sob o ponto de vista de capacidade e aproveitamento espectral [45]. Uma evolução, através de tecnologias denominadas genericamente 2.5G, possibilitou um melhor aproveitamento espectral, além de permitir uma integração com redes IP transparente ao usuário. As tecnologias HSCSD (*High Speed Circuit Switched Data*), GPRS (*General Packet Radio Service*) e EDGE (*Enhanced Data Rates for GSM Evolution*) representam um caminho evolutivo, cada uma com determinados avanços em relação às redes 2G, em direção às tecnologias 3G - UMTS (*Universal Mobile Telecommunication System*) e CDMA2000 (*Code Division Multiplexing Access 2000*) [59].

Sendo um sistema de comutação por circuito, o GSM disponibiliza taxas de transmissão de apenas 9.6 kbps, o que limita os tipos de serviços e aplicações suportados por ele. O GPRS é um sistema que introduz comutação por pacote, utilizando a mesma interface aérea do GSM. Permite taxas teóricas de até 171 kbps, sendo ideal para aplicações com tráfego em rajada, típicas da Internet. Com uma multiplexação de recursos de rádio mais eficaz que o GSM, o GPRS permite a utilização do sistema por um número maior de usuários. Uma eficiência espectral pode ser obtida no GPRS através da distribuição simultânea de pacotes para múltiplos usuários e mais de um *slot* pode ser utilizado por um usuário, para obtenção de velocidades mais elevadas. As requisições de pequenas rajadas de transmissão a altas velocidades (*download*) intercaladas com longos períodos de inatividade são mais indicadas para esta tecnologia sem conexão do que para um ambiente de comutação por circuito, no qual o precioso

tempo no ar é tarifado durante todo o período de uma conexão aberta [65].

QoS (*Quality of Service*) é um termo genérico relacionado ao desempenho de um sistema ou de uma rede [70, 74]. O desempenho é definido a partir de características tais como atraso, tempo de resposta, *throughput*, estabilidade, usabilidade, validade e precisão da informação, além de sua disponibilidade [29, 46]. A palavra “serviço”, do termo “qualidade de serviço”, refere-se a algo provido ao usuário e a qualidade está relacionada ao grau de satisfação deste usuário. Uma especificação de QoS é uma abstração usada para expressar características de qualidade exigidas pela aplicação, através de um conjunto de parâmetros. Está associada ao fluxo de informação do meio de comunicação. Por exemplo, uma especificação de QoS no nível de aplicação permite a seleção de um protocolo de transporte apropriado e, quando mapeado para o nível de rede, permite a reserva de recursos de rede adequados [32]. Sob o ponto de vista das aplicações, os parâmetros de QoS tratam, por exemplo, se a imagem em uma videoconferência deve ser em cores ou em preto e branco, se o som deve ter qualidade de telefone ou CD, se pode ou não haver interrupção da comunicação durante um *handoff*, entre outras coisas. Deste modo, deve existir um mecanismo de tradução destas exigências em parâmetros a serem negociados pelas entidades de rede, tais como taxa de transmissão ou atraso [66]. Logo, uma arquitetura de gerenciamento de QoS provê algumas funções, listadas abaixo, para obter e manter os requisitos de QoS especificados no contrato com a aplicação [37]. A aplicação de uma ou mais destas funções garante vários níveis de QoS (classes de serviço).

- Escolha dos parâmetros de QoS pela aplicação;
- Tradução em parâmetros que a rede possa “compreender”;
- Controle de admissão;
- Negociação com os elementos da rede;



- Reserva de recursos;
- Monitoração de QoS durante a comunicação e tomada de providências em casos de violações;
- Renegociações de parâmetros durante a comunicação, caso a aplicação necessite de novos níveis de QoS.

Para que um nível de QoS seja garantido às aplicações, devem ser reservados recursos da rede, tais como largura de banda, *buffers* nos roteadores, ciclos de CPU, entre outros. A racionalização no uso de recursos se faz necessária para que o maior número possível de usuários seja atendido simultaneamente. A arquitetura de gerenciamento de QoS e a aplicação estabelecem contratos sobre parâmetros de importância para a aplicação, exemplificados a seguir:

- Aplicações de áudio: perdas de pacotes devem ser suaves, de maneira que o entendimento do dado transmitido não seja prejudicado.
- Videoconferência: Para haver sincronia entre voz e imagem, os atrasos entre os pacotes não devem ser grandes e nem variáveis.
- Aplicações de ftp - *file transfer protocol*: perda de dados deve ser nula, não importando atrasos.

## 1.2 Descrição do Problema

Com a integração cada vez maior da Internet e dos sistemas móveis, usuários de redes móveis têm acesso a recursos de Telecomunicações e da Internet. Como historicamente as redes IP provêm serviços do tipo *best-effort* e não são orientadas a aplicações multimídia, QoS é um fator crítico no sucesso de redes GPRS. Para prover QoS perceptivo aos usuários finais, os recursos de rede devem ser utilizados de maneira otimizada.

Novas aplicações para Internet estão cada vez mais capazes de definir um nível de QoS desejado. As arquiteturas Intserv [66], Diffserv [20] e outras são as propostas para prover QoS fim-a-fim na Internet. Como adaptar estas soluções para ambientes móveis é uma questão que merece investigação e vem sendo estudada por vários pesquisadores. Para prover QoS fim-a-fim, de maneira consistente, os parâmetros de QoS escolhidos devem ser coerentes com as necessidades do fluxo de uma aplicação fim-a-fim.

Sistemas GPRS vêm sendo implantados em todo o mundo. Questões referentes à Qualidade de Serviço (*QoS*) ainda estão em aberto. As especificações do GPRS oferecem possibilidades de escolha de alguns parâmetros pelo fabricante ou pela operadora. Políticas de alocação e liberação de recursos, assim como a possibilidade de realizar operações com múltiplos ou único *slot* podem ser citados como exemplos. A escolha de parâmetros adequados afeta diretamente o desempenho da rede.

O modelo de QoS do GPRS foi projetado para ser independente de redes externas e seus mecanismos específicos de QoS. Para se conseguir garantias de QoS fim-a-fim em uma rede GPRS, é necessário que a mesma “entenda” mecanismos externos de QoS, tais como Diffserv, RSVP e MPLS [43], e seja compatível com aplicações que os utilize.

Nas especificações do GPRS, existem quatro classes de QoS propostas, além da negociação de parâmetros durante a ativação do contexto PDP. Para UMTS, o 3GPP definiu quatro classes de tráfego para as aplicações, dependendo da prioridade de transmissão. Uma arquitetura em camadas, para dar suporte a QoS no UMTS, foi proposta pelo 3GPP [3, 44].

Alguns trabalhos [9, 67] apresentam, na rede central IP fixa do GPRS, uma utilização do Diffserv, do IntServ, ou do MPLS. Alguns pesquisadores [9, 67, 76] sugerem mapeamentos das classes de tráfego e da arquitetura em camadas do UMTS para classes de QoS do Diffserv. Um tópico importante nesta linha de pesquisa é a derivação de mecanismos de encaminhamento de quadros da interface aérea, baseados em mar-

cação de pacotes da rede IP, que ainda, aparentemente, não foi proposto na literatura. Mecanismos eficientes de controle de admissão, reserva de recursos e controle de fluxo na interface aérea, integrados com mecanismos de garantia de qualidade de serviço no núcleo IP, ainda estão em aberto para implementações e avaliações.

### 1.3 Trabalho Proposto

Este trabalho propõe uma arquitetura de QoS fim-a-fim, que promove diferenciação de serviço homogênea, baseada nos mesmos critérios de prioridade, na interface aérea e no núcleo IP, em redes GPRS, denominada DiffMobil. O conceito fim-a-fim, no contexto do trabalho, refere-se à comunicação na rede GPRS, entre uma unidade móvel o *gateway* de comunicação com as redes de pacotes externas. A arquitetura DiffMobil é composta por DiffServ no núcleo IP da rede e por módulos de negociação de QoS, diagnóstico de rede, controle de admissão e controle de fluxo no enlace sem fio, desenvolvidos seguindo as seguintes etapas de trabalho:

1. Mapeamento entre classes de QoS do GPRS [5], do UMTS [3] e do DiffServ [26].
2. Simulação da interface aérea GPRS, através da ferramenta ns-2 [62] (*Network Simulator*) [60], que vem sendo amplamente utilizada na literatura em avaliações de diferentes protocolos de comunicação. Os resultados mostrados em [22] comparam o desempenho de uma rede GPRS com uma rede GSM.
3. Uma rede GPRS completa foi simulada através da ferramenta ns-2. A simulação do núcleo IP GPRS integrado à interface aérea foi mostrada em [23, 24], com resultados de análise de desempenho da rede, para diferentes cargas e diferentes aplicações.
4. O núcleo IP foi simulado com MIPv6, com o objetivo de se otimizar o roteamento, minimizando perda e atraso de pacotes após *handoffs*.

5. Uma adaptação do DiffServ foi feita para o núcleo IP, com resultados experimentais, obtidos através de simulação, mostrando que a diferenciação de serviço para classes prioritárias do GPRS garante largura de banda para as aplicações mais críticas de rede, o que não acontece quando se adota o modelo de melhor esforço.
6. Implementação, no ns-2, das seguintes funcionalidades da arquitetura de gerenciamento de QoS proposta :
  - Controle de admissão: Responsável pela inicialização de uma conexão sobre um enlace sem fio, alocação e liberação de canais de tráfego, com suporte a *handoff*; Trabalha com guarda de canais para classes prioritárias, tem como objetivo diminuir o número de bloqueios de novas chamadas de aplicações *conversational* e de *handoffs*;
  - Negociação de parâmetros de QoS: Módulo que permite que o usuário solicite à rede alguns parâmetros de QoS desejados: vazão e classe de serviço, inicialmente;
  - Diagnóstico da rede: Responsável por equilibrar o desempenho da rede, uma vez que o módulo de controle de admissão somente minimiza bloqueios. O diagnóstico mantém um controle do nível de utilização da rede e impede que novas chamadas sejam iniciadas em caso de congestionamento.
  - Controle de fluxo: Captura marcação DiffServ dos pacotes IP da rede fixa GPRS e transfere esta marcação para um mecanismo de escalonamento das estações rádio base da interface aérea. Ele está relacionado à priorização de escalonamento do fluxo na interface aérea.

Para cada uma das etapas acima, foram executados vários experimentos, com variação de parâmetros de entrada tais como tempo simulado, número de usuários,

tipos de aplicação por usuário e taxas de geração de tráfego. Como resultados, podem ser citados os seguintes parâmetros de desempenho:

- Atraso na transmissão - atraso médio por pacote, entre origem e destino. Esta medida se reflete no tempo de resposta percebido pelo usuário final da rede;
- Variação de atraso - diferença entre os atrasos nas chegadas de pacotes.
- Vazão - taxa efetiva de transmissão, dadas em bits por segundo (bps), obtida por uma aplicação;
- Utilização da rede por cada tipo de tráfego;
- Percentual de *handoffs* com perda de conexão, no caso do controle de admissão.

As simulações foram executadas para o modelo de melhor esforço e para cada módulo adicional da arquitetura Diffmobil, tendo como objetivo avaliar sua eficácia. Assim, foram feitos experimentos separados, agregando módulos de cada etapa de desenvolvimento, da seguinte maneira: experimentos para melhor esforço, para DiffServ no núcleo IP, para controle de admissão, para controle de admissão com diagnóstico de rede e para controle de admissão com diagnóstico de rede, controle de fluxo e DiffServ no núcleo IP. Os resultados mostram que o DiffServ no núcleo IP oferece garantia de vazão a classes prioritárias, que o controle de admissão reduz a taxa de bloqueios de *handoffs* e de novas chamadas da classe *conversational*, mas piora o desempenho da rede. Para resolver este problema, o módulo de diagnóstico controla níveis de congestionamento da rede, bloqueando transmissões menos prioritárias cada vez que um nível pré-estabelecido de utilização é ultrapassado. O desempenho da rede é melhorado pelo módulo de diagnóstico, mas só é garantido nos níveis requisitados após a introdução do módulo de controle de fluxo. Este estende as marcações do Diffserv do núcleo IP, fazendo diferenciação de serviço na interface aérea e promovendo um escalonamento baseado em pesos de prioridades diferentes. Com a implementação dos módulos

propostos e a adoção do DiffServ no núcleo IP, o desempenho de classes prioritárias pode ser garantido fim-a-fim através da arquitetura Diffmobil, conforme mostram os resultados experimentais.

## 1.4 Organização do texto

O Capítulo 2 apresenta fundamentos teóricos do GPRS. O Capítulo 3 apresenta fundamentos teóricos de QoS e a revisão bibliográfica. O Capítulo 4 descreve a arquitetura proposta. O Capítulo 5 descreve o modelo de simulação do GPRS e da arquitetura DiffMobil, mostrando resultados experimentais. O Capítulo 6 conclui o trabalho.

# Capítulo 2

## Fundamentos Teóricos de GPRS

### 2.1 Características Gerais de uma Rede GPRS

O GPRS (*General Packet Radio Service*) é uma tecnologia que permite aos assinantes de operadoras GSM acesso a aplicações que envolvem transmissão de dados, tais como e-mail, comunicação com redes corporativas e com a Internet, através de seus telefones celulares. O GPRS está sendo implementado comercialmente, possibilitando a ligação entre provedores de serviços de celular e provedores de serviço de Internet. A maioria das operadoras GSM está seguindo esta tendência de evolução.

Conforme ilustra a Figura 2.1, novos nós são introduzidos ao GSM [25], [20], [71] e [72]. Estes nós são introduzidos para possibilitar a comutação por pacote, conforme explicado abaixo:

- **Nó Servidor de Suporte ao GPRS - SGSN** (*Serving GPRS Support Node*): Localiza unidades móveis e desempenha funções de segurança e controle de acesso. Recebe e envia dados de/para estações móveis, monitorando unidades presentes em sua área de abrangência. SGSNs enviam requisições aos HLR's (*Home Locator Register*) para obtenção de dados sobre o perfil de usuários e detectam a presença de novas estações móveis em uma determinada área. As suas funções

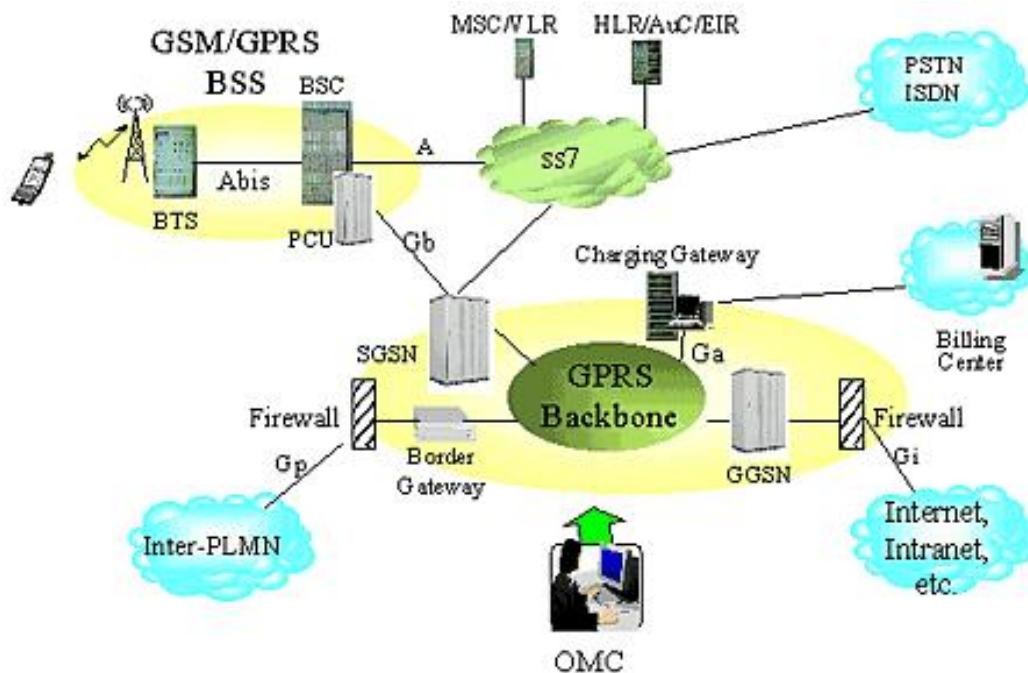


Figura 2.1: Elementos do GPRS

incluem gerenciamento de mobilidade e sessão, *handovers*, *paging* e verificação de registro de usuários.

- Nó *Gateway* de Suporte ao GPRS - **GGSN** (*Gateway GPRS Support Node*): GGSN é um nó de interface com redes de pacote externas (PDNs - *Packet Data Networks*), tais como Internet e X.25, para prover o roteamento de dados de/para unidades móveis. O GGSN mantém informações de roteamento necessárias para direcionar pacotes de dados aos SGSNs que servem determinadas estações móveis. Estão entre as funções deste nó a gerência de localização e autenticação, além da contabilidade do número de pacotes transmitidos, para prover tarifação precisa para um assinante. O GGSN é projetado para prover interligação com redes de pacote externas e é conectado ao SGSN através de uma rede GPRS central, baseada em IP.

Para permitir taxas de transmissão mais elevadas, o GPRS adota novos esquemas



de codificação de erros e múltiplos *timeslots*, de maneira similar ao HSCSD (*High-Speed Circuit-Switched Data*). Teoricamente, é possível atingir uma taxa máxima de transmissão de 171,2 kbps, utilizando oito *timeslots*. Entretanto, na prática, estes valores são bem inferiores devido à qualidade insatisfatória de interfaces aéreas e de telefones celulares. Como a comunicação por pacote permite alocação de recursos não contínua para um usuário, a tarifação não é baseada em tempo de conexão e sim em utilização de recursos.

Adicionalmente à inclusão de novos nós em redes GPRS, tornam-se necessárias algumas modificações na estrutura do GSM, descritas a seguir:

- Sistema de Estação Rádio Base -**BSS** (*Base Station System*): Precisa ser aperfeiçoado para reconhecer e enviar dados do usuário para o SGSN que está servindo à área. O BSS consiste em dois nós: um deles é o controlador da estação base (BSC - *Base Station Controller*), do qual faz parte a unidade de controle de pacote (PCU - *Packet Control Unit*). O BSS deve suportar todos os protocolos relevantes ao GPRS para comunicação sobre a interface aérea. A PCU deve converter pacotes em um formato que possa ser transmitido pela interface aérea, gerenciar recursos de rádio, implementar medições de QoS, configurar, supervisionar e desconectar chamadas comutadas por pacote, incluindo suporte para mudança de célula, configuração de recurso de rádio e atribuição de canal, sendo uma unidade extremamente importante para o desenvolvimento deste trabalho. O segundo nó é a estação base transceptora (BTS - *Base Transceiver Station*), que tem a função de encaminhamento, realizando modulação de frequências da portadora e demodulação de sinais.
- Repositórios de Informações de Localização de Usuários Locais e de Visitantes, **HLR** (*Home Locator Register*) e **VLR** (*Visitor Locator Register*). Precisam ser modificados com interfaces adicionais para registrar perfis de usuários GPRS e

responder a *queries* originadas em SGSNs referentes a estes perfis.

- Estação ou unidade móvel, UMI : Deve ser equipada com a pilha de protocolos GPRS, de maneira que o usuário possa utilizar esta tecnologia de rede. Três tipos de terminais são definidos no padrão GPRS: Classe A (suporta tráfego de comutação por pacote e por circuito simultaneamente), Classe B (suporta tráfego de comutação por pacote e por circuito, não simultaneamente), Classe C (suporta tráfego de comutação por pacote ou por circuito). Atualmente, a maioria dos fabricantes opta pela classe B.
- Novas interfaces físicas incluem a interface Gb, que conecta o SGSN à PCU e fica localizada no subsistema da estação base (BSS), a interface Gn, que conecta o GGSN ao SGSN e as interfaces Gc, Gr e Gs, que carregam sinais baseados em protocolos SS7, responsáveis pela comunicação entre nós GSM e GPRS.

Comutação por pacote indica que recursos de rádio somente são usados quando os usuários estão realmente enviando ou recebendo dados. Ao invés de um canal de rádio ficar dedicado para um usuário móvel durante um período fixo de tempo, vários usuários podem utilizá-lo concorrentemente. Isto significa que, com o GPRS, um maior número de usuários pode compartilhar a mesma largura de banda de uma célula. O número de usuários simultâneos depende das aplicações sendo utilizadas e de como os dados estão sendo transferidos. Em função disso, os telefones GPRS podem estar nos seguintes estados, conforme ilustra a figura 2.2:

- *Idle*: É o estado no qual um telefone móvel está em uma rede GSM. Ele recebe *paging* de comutação por circuito e se comporta como um telefone GSM.
- *Ready*: É o estado alcançado pelo equipamento móvel GPRS quando ele se conecta à rede, através de um procedimento denominado *attach*, que será descrito adiante. Neste estado, ele pode ativar um contexto PDP (*Packet Data Protocol*),

que permite que o telefone estabeleça uma sessão de transferência de pacotes com redes de dados externas. Quando um contexto PDP está ativo, blocos de recursos de rádio são alocados àquela sessão, até o encerramento de transferência de dados, quando o telefone passa para o estado *standby*.

- *Standby*: É o estado no qual o equipamento está conectado à rede GPRS, mas não há transmissão de dados. Se um pacote de dados chegar para o equipamento móvel, a rede envia um *paging* a ele que, então, ativa um contexto PDP, trazendo-o de volta ao estado *ready*.

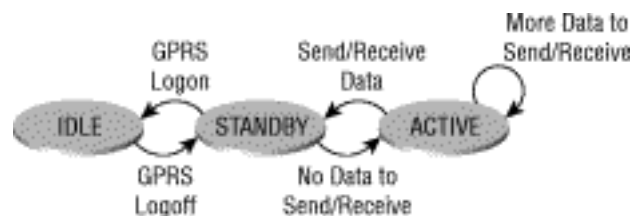


Figura 2.2: Estados de uma unidade móvel GPRS

## 2.2 Gerenciamento de Sessão e Mobilidade no GPRS

Três procedimentos - *attach*, ativação de contexto PDP (*Packet Data Protocol*) e localização de área de roteamento - são necessários para possibilitar a conexão de usuários GPRS com redes externas. A seguir, estes procedimentos são descritos:

- *Attach*: Este procedimento faz parte da gerência de mobilidade do GPRS (GMM). É sempre iniciado pelo usuário do móvel, seja manual ou automaticamente, quando ele liga o telefone. A requisição de um *attach* é feita diretamente ao SGSN, de maneira transparente à estação base. Inicialmente, a unidade móvel notifica o SGSN sobre sua identidade, como um IMSI (*International Mobile Subscriber Identity*) ou como um P-TMSI (*Packet Temporary Mobile Subscriber Identity*).

Logo após, ela envia sua antiga identificação de área de roteamento (RAI), classe (CKSN) e tipo de *attach* desejado. O tipo de *attach* indica se a unidade móvel deseja se conectar como um dispositivo GPRS, GSM, ou ambos. Ocorre uma verificação de autenticação entre o terminal e o HLR, no *attach*. O SGSN conecta a unidade móvel e informa ao HLR caso haja alteração de área de roteamento. Uma transferência de informações sobre o usuário é feita do HLR para a CCC/VLR. O usuário recebe notificação do SGSN, quando o processo termina. O objetivo do *attach*, é que a localização e as características de uma unidade móvel sejam conhecidas pela rede. Mas este procedimento não habilita a unidade móvel a transmitir e receber dados. Para isto, é necessário que a unidade móvel ative uma sessão de comunicação, utilizando um contexto PDP, descrito a seguir.

- *Ativação de Contexto PDP (Packet Data Protocol) :*

Um contexto PDP ativa uma sessão de comunicação de uma UM com um SGSN. Durante o procedimento de ativação, a unidade móvel provê um endereço IP estático ou requisita um temporário à rede. Ela também especifica um APN (*Access Point Name*) com o qual deseja se comunicar como, por exemplo, um endereço da Internet. A UM requisita um nível desejável de QoS e um identificador de ponto de acesso ao serviço (NSAPI). O NSAPI é utilizado para identificar pacotes de uma aplicação específica, uma vez que uma UM pode estabelecer múltiplas sessões de contexto PDP para diferentes aplicações. Após receber informações da UM, o SGSN determina qual GGSN está conectado àquele APN e encaminha a requisição. Se a UM possui um endereço IP estático, o GGSN a conecta diretamente ao ponto de acesso desejado. Senão, ele obtém um endereço temporário do APN. O GGSN também provê alguns identificadores para comunicações de dados entre o SGSN e o GGSN. O SGSN é a entidade do GPRS que provê QoS, baseado no registro do assinante da UM e na disponibilidade de

recursos. Quando o procedimento de ativação termina, a informação para transferência de dados é encaminhada à UM. Os passos de uma ativação de contexto PDP são descritos a seguir:

1. A ativação de contexto PDP é requisitada pelo usuário, quando este envia o primeiro pacote através de seu equipamento portátil, provocando a criação de um circuito full-duplex entre a UM e a interface Gi.
  2. O SGSN, baseado em informação recebida do HLR, durante o *attach*, valida a requisição.
  3. O APN é enviado para o DNS (*Domain Name System*), para que se encontre o endereço IP do GGSN relevante.
  4. Uma conexão lógica é criada entre o SGSN e o GGSN.
  5. O GGSN atribui ao terminal móvel um endereço IP dinâmico, da faixa de endereços alocados à PLMN (*Public Land Mobile Network*), ou a partir de um servidor RADIUS (*Remote Authentication Dial In User Service*).
- **Localização de área de roteamento:** Uma área de roteamento é composta por um conjunto de células definido pela operadora. Quando uma UM muda de área de roteamento, os gateways GPRS e a informação de contexto PDP precisam ser atualizados.

## 2.3 Pilha de Protocolos do GPRS

Os dados que passam no *backbone* IP, entre GGSNs e SGSNs, são transmitidos usando o protocolo GTP (*GPRS Tunneling Protocol*), que encapsula pacotes de dados, adicionando informações de roteamento. O GTP usa TCP ou UDP para transporte de pacotes. O IP é o protocolo da camada de rede em um *backbone* GPRS, podendo haver

Ethernet, ISDN ou ATM abaixo dele. Como uma rede pode possuir diversas interfaces GGSN e SGSN (Gn), o GTP provê um identificador de túnel (TID) para cada pacote. Este TID identifica o destino e a transação à qual o pacote pertence. As transações têm identificadores lógicos ou IMSI's (*Internet Mobile Subscriber Identity*).

O SNDCP (*Subnetwork Dependent Convergent Protocol*) é usado entre um SGSN e a UM. Este protocolo converte pacotes da camada de rede da interface Gn em um formato inteligível pela arquitetura de rede GPRS. As funções do SNDCP são: multiplexação de N-PDUs de uma ou mais entidades de rede em uma conexão LLC, armazenamento de N-PDUs para serviço confirmado, gerência da seqüência de entrega para cada NSAPI, compressão e descompressão de informação do protocolo e de dados do usuário, segmentação e recomposição de dados comprimidos no tamanho máximo PDUs-LLC, negociação de parâmetros XID entre entidades SNDCP.

O LLC (*Logical Link Protocol*) provê um enlace lógico confiável entre a UM e seu SGSN. O LLC usa modos com e sem confirmação para transmissão de quadros, dependendo do QoS negociado com o usuário. Este protocolo também gerencia retransmissões de quadros, *buffers* e tamanho da informação baseado na classe de atraso negociada como parâmetro de QoS.

O BSSGP (*Base Station System GPRS Protocol*) encaminha informações entre o SGSN e o BSS. Este protocolo tem informação de QoS, mas não possui nenhuma forma de correção de erro. Sua principal função é prover informação relacionada a rádio para os protocolos RLC e MAC da interface aérea. O LLC usa serviços do BSSGP para transferência de dados. A função “*relay*” do BSS transfere quadros LLC entre a camada RLC/MAC e a camada BSSGP. O BSSGP envia informação para a camada de serviços de rede para determinar o destino da transferência:

- *BVCI (BSSGP Virtual Connection Identifier)*: É enviado para camadas de serviços de rede para sinalização de roteamento e transferência de dados entre as entidades

par corretas. Cada BVCI entre duas entidades par é único.

- *LSP (Link Selection Parameter)*: É usado em conjunto com o BVCI para auxiliar na seleção de um enlace físico.
- *NSEI (Network Service Entity Identifier)*: Utilizado no BSS e no SGSN para prover a funcionalidade de gerência de rede necessária na operação da interface Gb. O NSEI junto com o BVCI identificam unicamente uma conexão virtual BSSGP.

O NS (*Network Service*) é uma conexão ponto-a-ponto entre o SGSN e o BSS. A camada NS usa uma tabela DLCI (*Data Link Connection Identifier*) para indicar a melhor rota entre o SGSN e o BSS. O valor inicial do DLCI é derivado do BVCI, do NSEI e do LSP providos pelo BSSGP. Este valor vai mudando à medida que os quadros vão passando pela rede *frame relay* e atingindo seu destino final.

A comunicação via rádio entre a UM e a rede GPRS abrange funções das camadas de enlace e física, representadas na figura 2.3 pela área tracejada.

O RLC (*Radio Link Control*) localiza-se na PCU e é responsável pelas seguintes funções:

- Transferência de PDUs LLC entre as camadas LLC e MAC;
- Segmentação de PDUs LLC em blocos de dados RLC e recomposição de blocos de dados RLC em blocos de quadros TDMA;
- Segmentação e recomposição de mensagens de controle RLC/MAC em blocos de controle RLC/MAC;
- Correção de erros em transmissões seletivas de blocos de dados RLC.

A função de segmentação do RLC corresponde a um processo de pegar uma ou mais PDUs LLC e dividi-las em blocos RLC menores. As PDUs LLC são agrupadas em uma

estrutura denominada TBF (*Temporary Block Flow*) e recebem a alocação de um ou mais canais PDCH (*Packet Data Channel*). O TBF é temporário, mantido somente durante a transferência e recebe uma identificação da rede: TFI (*Temporary Flow Identity*), que identifica a UM na camada RLC/MAC. O TBF pode ser entendido como uma conexão física usada entre duas entidades de rádio para suportar a transferência de PDUs LLC em canais físicos. Todos os quadros LLC que foram segmentados formam um TBF no enlace lógico da interface aérea. Blocos de dados RLC consistem em um cabeçalho, uma unidade de dados e bits sobressalentes. O bloco RLC, juntamente com um cabeçalho MAC, deve ser codificado usando-se um dos quatro tipos de codificação (CS1, CS2, CS3, ou CS4) definidos para o GPRS. O tipo de codificação, explicado adiante, influencia diretamente o processo de segmentação.

O protocolo MAC (*Medium Access Control*) [6] controla a sinalização de acesso à interface aérea, incluindo a gerência de recursos compartilhados (atribuição de blocos de rádio a múltiplos usuários no mesmo *timeslot*). O MAC obtém estas funcionalidade colocando um cabeçalho em blocos de dados e de controle RLC/MAC. O cabeçalho MAC contém vários elementos, alguns deles são específicos da direção, *downlink* ou *uplink*. Os parâmetros chave do cabeçalho MAC são:

- USF (*Uplink Status Flag*): É enviado em todos os blocos *downlink* RLC/MAC e indica o dono ou uso do próximo bloco de rádio *uplink* no mesmo *timeslot*. Como múltiplos usuários podem ser multiplexados no mesmo *timeslot*, eles só podem transmitir quando o USF indica a vez de cada um.
- RRBP (*Relative Reserved Block Period*): Identifica um bloco único *uplink* no qual a UM transmite informação de controle.
- PT (*Payload Type*): Tipo de informação (dados ou controle) que segue no restante do bloco RLC/MAC.



- CV (*Countdown Value*): É enviado pela UM para permitir que a rede calcule o número de blocos RLC que restam no TBF *uplink* corrente.

O MAC é o responsável pela multiplexação de um canal físico para várias UMs e permite que uma UM utilize vários canais físicos em paralelo, através da alocação de múltiplos *slots* de tempo. A camada RF (*Radio Frequency*) transmite e recebe ondas moduladas na frequência da portadora.

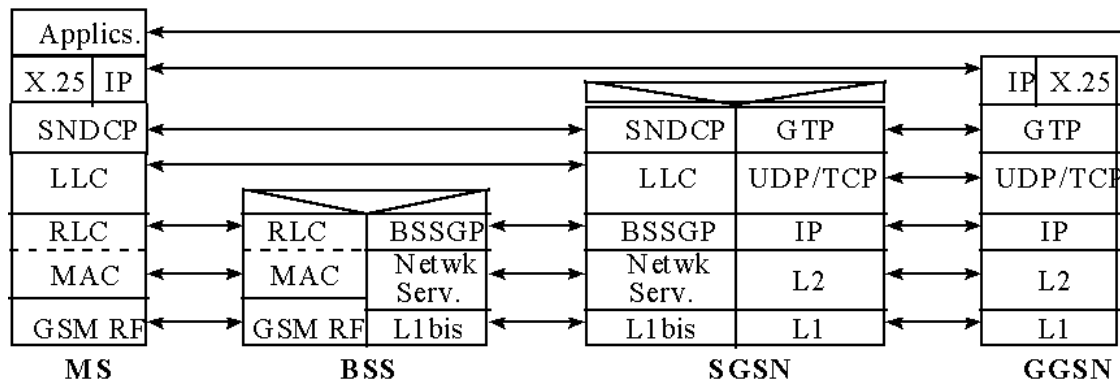


Figura 2.3: Pilha de Protocolos do GPRS

## 2.4 Transmissão de Pacotes no Sentido *Downlink*

Quando um datagrama IP, enviado de uma rede externa ao GPRS, chega a um GGSN, ele é chamado de N-PDU (*Network Data Packet Unit*). Este pacote vai estar endereçado a um IMSI ou a um endereço IP. À medida que o pacote vai descendo na pilha de protocolos do GGSN, cabeçalhos vão sendo adicionados a ele. Depois de passar pelo encapsulamento GTP, a N-PDU passa pelo TCP/UDP, que coloca seu próprio

cabeçalho, portos fonte e destino, informação de roteamento e controle de fluxo. Na camada IP, o cabeçalho recebe o endereço destino do SGSN. Dependendo do tamanho da N-PDU e do tamanho máximo de unidade de transmissão da camada IP, uma fragmentação de PDU pode ser necessária. Finalmente, a N-PDU com todos os cabeçalhos é transportada no nível físico da interface Gn até o SGSN.

### 2.4.1 Transmissão no SGSN

No SGSN, os cabeçalhos são retirados e a N-PDU é entregue ao protocolo SNDCP. O SNDCP vai segmentar a N-PDU, de maneira a satisfazer o tamanho máximo de transmissão de 1520 octetos sobre a camada NS da interface Gb. Pode haver uma compressão opcional também. O SNDCP vai classificar a N-PDU como orientada a conexão (SN-DATA PDU) ou não (SN-UNIDATA PDU). Finalmente, ele adiciona o cabeçalho SNDCP, contendo informações sobre compressão e segmentação e envia a PDU para a camada LLC inferior.

Conforme mencionado anteriormente, a principal função do LLC é o provimento de uma conexão lógica entre o SGSN e a unidade móvel. Ele encapsula a PDU SN-DATA ou SN-UNIDATA em um quadro LLC com seu próprio cabeçalho. O quadro LLC contendo a SN-PDU é chamado de bloco LLC. O cabeçalho LLC adiciona informação de controle (usada em modo de confirmação), seqüência de controle de quadro e valores SAPI. O SAPI neste caso se refere ao serviço associado com o quadro LLC nesta sessão PDP. Os serviços para este quadro podem ser gerência de mobilidade ou dados de usuário com níveis de parâmetros de QoS (1 a 5). Estes níveis de serviço são decididos no processo de negociação de QoS.

A camada BSSGP, abaixo do LLC, provê informação de roteamento ao nível NS para rotar o bloco LLC sobre a camada física *frame relay*. O BSSGP também adiciona um cabeçalho ao bloco LLC, contendo informações essenciais às camadas RLC/MAC da

interface aérea a respeito da transmissão de blocos, incluindo parâmetros de prioridade, TLLI (*Temporary Logical Link Identifier*) etc.

### 2.4.2 Transmissão no BSS

O dado é enviado sobre a conexão física entre o SGSN e o BSS. O BSSGP no BSS envia, então, toda a informação ao RLC. A função mais importante do RLC é a segmentação de blocos LLC em blocos RLC menores. Um grupo de blocos LLC, que foram fragmentados em blocos menores, é conhecido como TBF (*Temporary Block Flow*), conforme mencionado anteriormente. Cada bloco TBF recebe a alocação de um ou mais canais do tipo PDTCH (*Packet Data Channel*). O TBF é mantido somente durante o tempo de transmissão. O RLC adiciona um cabeçalho aos blocos de dados contendo o TFI, TLLI, número de seqüência do bloco RLC, indicação do último bloco e informação sobre a direção (*uplink* ou *downlink*). O tamanho do bloco LLC dentro de um bloco RLC depende do tipo de codificação escolhido para ser usado na interface aérea. Existem quatro esquemas de codificação definidos para GPRS: CS1, CS2, CS3 e CS4, que contém no máximo 22, 32, 38 e 52 octetos, respectivamente. A escolha da codificação depende do compromisso entre vazão e confiabilidade do sistema.

### 2.4.3 Transmissão na Interface Aérea

Após a segmentação RLC e a inserção de cabeçalhos, os blocos RLC são transmitidos sobre a interface aérea. O protocolo MAC, responsável por controlar sinalização de acesso, inclui alocação de blocos *uplink* e *downlink*. Ele adiciona seu próprio cabeçalho que fica sendo monitorado pela UM. O dado é transmitido sobre a interface aérea até a UM através da camada física (GSM RF). O dado se move na pilha de protocolos da UM, com os cabeçalhos sendo retirados em cada camada, até a informação chegar à camada de aplicação.

## 2.5 A Interface Aérea do GPRS

Os canais físicos do GPRS são formados utilizando a mesma estrutura do GSM, compartilhando *slots* de tempo, quadros TDMA e bandas de frequência (900, 1800 e 1900 MHz). Um *slot* de tempo pode ser usado tanto para transmissão de dados de comutação por pacote do GPRS, assim como para dados de comutação por circuito do GSM. O GPRS consegue uma capacidade maior através da multiplexação de múltiplos usuários nos mesmos canais físicos (*timeslots*).

Os quadros GSM ou GPRS consistem em oito *timeslots*. O *timeslot* 0 é reservado para *broadcast*. **PDCH** (*Packet Data Channel*) é o nome dado a um canal físico do GPRS. **Bloco de Rádio** é o nome que se dá à estrutura básica de transmissão do GPRS. Seu tamanho é de 456 *bytes*, sendo apenas uma parte dedicada à transmissão de dados propriamente dita (*payload*), devido à técnica de correção de erros FEC (*Forward Error Detection*). São necessários quatro *slots* de tempo em quatro quadros TDMA consecutivos, para transmissão de um bloco de rádio.

### 2.5.1 Canais Lógicos

Os tipos de canais lógicos transmitem informações de usuários ou de controle do GPRS. Eles são descritos a seguir:

- **PRACH** (*Packet Random Access Channel*): É um canal de controle utilizado pela UM para requisitar acesso para transferência no *uplink*. Consiste em dois canais multiplexados no tempo: S-PRACH (sinalização) e U-PRACH (dados de usuário). O S-PRACH é acessado por novas chamadas e por requisições de *handoffs*, através do protocolo de acesso randômico *Slotted Aloha*, para obter acesso aos canais PDTCH. O U-PRACH é utilizado para reserva de recursos PDTCH sob demanda, também através do *Slotted Aloha*.

- PPCH (*Packet Paging Channel*): Através dele, uma estação base faz um *paging* em uma UM, antes da transmissão no *downlink*.
- PAGCH (*Packet Access Grant Channel*): Canal de controle, usado para enviar atribuições de recursos nas direções *uplink* e *downlink*, antes de iniciar a transferência de pacotes.
- PNCH (*Packet Notification Channel*): Canal de controle usado pela estação base para transmissão de uma notificação para um grupo de unidades móveis de uma célula, antes de uma transferência de pacotes *multicast*.
- PBCCH (*Packet Broadcast Control Channel*): Canal de controle usado pela estação base para transmissão de uma informação para todas as unidades móveis de uma célula.
- PDTCH (*Packet Data Transfer Channel*): Canal de tráfego de pacotes, usado para transmissão de pacotes de dados para *uplink* e *downlink*.
- PACCH (*Packet Associated Control Channel*): Usado para transmissão de informação sobre uma UM e seus PDTCHs como, por exemplo, controle de potência, confirmações de pacotes ou realocação de recursos.
- PTCCH (*Packet Timing Advance Control Channel*): Usado no *uplink* para transmissão de rajadas de acesso randômico. Ele permite uma estimativa do avanço de tempo necessário pela UM no modo de transferência de pacotes. No *downlink* pode ser usado para atualizar avanço de tempo para múltiplas UMs.

## 2.5.2 Mapeamento de Canais Lógicos em Canais Físicos

Os canais lógicos definidos na seção anterior não requerem recursos físicos separados. Eles podem ser mapeados para os canais físicos através de uma técnica denominada

*multiframeing*. Um quadro múltiplo é um conjunto de quadros TDMA que recebem a mesma funcionalidade. No GPRS, um quadro múltiplo é constituído de 52 quadros TDMA. Cada quadro múltiplo é dividido em 12 blocos, cada um com 4 quadros TDMA. Estes 12 blocos são os blocos de rádio, que recebem funções dos canais lógicos. Os blocos de rádio totalizam 48 quadros TDMA. Dos 4 quadros restantes, dois são usados para temporização e dois são deixados livres para serem usados para medidas de interferência com células vizinhas e controle de potência.

### 2.5.3 Transferência de Quadros no Sentido *Uplink*

Uma unidade móvel deve requisitar recursos à estação base. Ela inicia uma transferência de dados através de uma requisição de um canal, usando o PRACH. A rede responde, no canal PAGCH, com atribuição e reserva de recursos em um ou mais PDCHs para a transferência de blocos de rádio *uplink*. Usando PRACH na requisição do canal, a UM envia informação suficiente para a rede poder atribuir e reservar recursos à transferência. Estas informações incluem classe de múltiplos *slots*, número de blocos de rádio, método de acesso ao meio para aquele TBF (capacidade de acesso a rádio, nível de interferência em cada *timeslot*, motivo da requisição, nível de recepção de sinais etc). Se a rede não responde, a UM retransmite a requisição após um tempo de espera randômico.

### 2.5.4 Alocação Dinâmica ou Fixa de Blocos de Rádio

A alocação de blocos de rádio pode ser dinâmica ou fixa. O processo que informa à UM sobre a disponibilidade de recursos para transmissão bloco a bloco é conhecido como alocação dinâmica. A rede envia uma mensagem de atribuição à UM, com uma lista de PDCHs *timeslot* e portadora, com USFs correspondentes. Um TFI único é alocado e incluído em cada bloco de controle e de dados relacionado àquele TBF. A UM monitora

os USFs dos PDCHs alocados e transmite blocos de rádio no canais que possuem o USF destinado a ela. O USF é um campo de 3 bits do cabeçalho MAC. Ele atribui 1 ou 4 blocos de rádio, em um momento, para a UM. Logo, uma UM, quando detecta o seu USF em uma transmissão *downlink*, envia um ou quatro blocos de rádio e também o número de blocos de rádio restantes no TBF. A rede continua alocando canais para aquela UM até que ela informe que não há mais blocos de rádio a transmitir.

Na alocação fixa, a rede envia na mensagem de resposta a atribuição de canal para todos os blocos de rádio. Nesta atribuição estão incluídos o parâmetro de frequência, *timeslots*, um *bitmap* de 127 bits e o número do quadro para o início da transmissão. Quando a UM recebe esta informação, ela começa a transmitir no blocos correspondentes, a partir do quadro inicial, usando o *bitmap* da seguinte maneira: 0 é um bloco fora da atribuição dada a ela e 1 é um bloco atribuído a ela.

### 2.5.5 Transferência de Blocos RLC

O processo de transferência de blocos tem um mecanismo de confirmação e de numeração de blocos de um TBF. Ao final de cada janela de transmissão, o receptor envia uma mensagem de confirmação. A mensagem de confirmação contém um *bitmap* dos números dos blocos RLC, com o bit 1 representando um bloco recebido corretamente (*ack*) e com o bit 0 representando um bloco não recebido corretamente (*nack*). Quando a UM envia o último bloco de um TBF e informa à rede que não tem mais blocos a transmitir, um temporizador de 5 segundos é iniciado. Quando ele expira, a UM passa a considerar a atribuição de recursos que recebeu inválida.

### 2.5.6 Transferência de Quadros no Sentido *Downlink*

A rede inicia a transmissão de pacotes para uma UM no estado *ready* usando uma mensagem de atribuição de *downlink*. Se há uma transferência *uplink* em progresso

para aquela UM, a mensagem de atribuição pode ser enviada em um PACCH. Senão, pode ser enviada em um PPCH. Esta mensagem informa à UM o número de *timeslots*, parâmetros sobre frequência, avanço de tempo, controle de potência, TFI e o número de quadro TDMA inicial.

A rede envia blocos RLC/MAC pertencentes a um TBF nos canais atribuídos àquela transferência de *downlink*. Normalmente, mais de uma UM é multiplexada nos PDCHs ao mesmo tempo, sendo que cada uma é identificada pelo TFI, conforme mencionado anteriormente.

Em *downlink*, o cabeçalho MAC dos blocos RLC contém um bit para fazer *polling* na unidade móvel. São informados à UM: número relativo do quadro após o qual a UM deve enviar um ack/nack e um indicador do último bloco transmitido. Após receber o ack final, a rede inicia um temporizador. Quando ele expira, o TFI e os canais alocados são liberados.



# Capítulo 3

## Fundamentos Teóricos de QoS e Revisão Bibliográfica

O conceito de Qualidade de Serviço, ou QoS, está relacionado ao grau de satisfação de usuários de uma rede. Níveis diferenciados de QoS são garantidos a tipos diferentes de aplicações de uma rede. Quando se garante QoS a uma aplicação, está se garantindo um desempenho requisitado pelo usuário da aplicação. A escolha de parâmetros de QoS pode ser feita pela seleção de classes de serviço, que assumem valores a serem negociados com a camada de transporte. Os seguintes parâmetros de tráfego são, geralmente, negociados entre a aplicação e as camadas inferiores:

- *Vazão (throughput)*: Taxa efetiva de transmissão, normalmente dada em bits por segundo, obtida por uma aplicação. Podem ser especificadas a vazão de pico e a vazão média requisitadas em uma negociação de nível de serviço. Aplicações de tempo real e de transmissão de voz sobre IP (*VoIP*) exigem alta vazão.
- *Atraso de transmissão (delay)*: Tempo que o dado gasta para ir do transmissor ao receptor, sendo causado pela latência do meio e em *buffers* intermediários. Aplicações de transmissão de voz requisitam baixos atrasos.

- Variação do atraso (*jitter*): Diferença entre os atrasos de pacotes consecutivos de uma determinada aplicação. Aplicações multimídia e *VoIP* não suportam variações de atraso, por exemplo.
- Confiabilidade da transmissão: Tolerância à perda de pacotes devido à falta de espaço em *buffers* ou erros de transmissão. Como exemplo, aplicações de videoconferência podem ter certa tolerância à perda de pacotes, mas ftp ou correio eletrônico não.

Em redes sem fio, a mobilidade de um elemento ocasiona problemas para se garantir QoS à aplicação. Há necessidade de se fazer reserva de recursos entre diversas células, além de se garantir um roteamento correto. Além da mobilidade no ambiente sem fio, novos problemas devem ser tratados: restrição no consumo de energia, alta taxa de erros, variabilidade na qualidade do enlace que provoca desconexões. Estes novos problemas implicam em especificações de novos parâmetros de QoS, assim como tempo máximo da interrupção da conexão, frequência máxima de interrupção de conexão, *handoff* suave, prioridades na manutenção de conexões em caso de escassez de recursos, quando ocorre a chegada de um novo usuário em uma célula.

A negociação de parâmetros de QoS deve ser a mesma tanto na parte fixa quanto na parte móvel, quando ocorre interconexão dos dois tipos de rede. Um dos principais problemas é o suporte à mobilidade e roteamento correto de dados. Uma das soluções propostas para resolver este problema é o IP Móvel [50]. Uma alternativa de roteamento, neste caso, envia pacotes à unidade móvel de forma transparente à rede fixa, que deve simplesmente conhecer um endereço no qual os pacotes são coletados e entregues à unidade móvel, independentemente de sua localização. Para a parte móvel, deve-se minimizar o volume de dados transmitidos em função da velocidade ser mais baixa que na rede fixa, da qualidade da transmissão ser variável e da restrição no consumo de energia.

O restante deste capítulo está organizado da seguinte maneira: a seção 3.1 faz um levantamento das principais soluções de QoS encontradas na literatura. A seção 3.2 apresenta uma revisão da literatura relacionada ao trabalho. A seção 3.3 apresenta uma revisão da literatura complementar.

## 3.1 Exemplos de Soluções de QoS

Nesta seção são apresentadas as principais soluções de QoS para redes IP. MPLS (*Multiprotocol Label Switching*) é uma técnica de encaminhamento de pacotes padronizada pela IETF [33]. Utiliza *labels* para tomar decisões sobre o encaminhamento em nós da rede, em contraste ao encaminhamento tradicional *hop-by-hop* das redes IP. A arquitetura de Serviços Integrados (*IntServ*) e a arquitetura de Serviços Diferenciados *DiffServ* foram desenvolvidas tendo como objetivo proporcionar QoS em redes IP.

### 3.1.1 IntServ

A arquitetura IntServ é composta por quatro elementos, listados a seguir.

1. RSVP (*Resource ReSerVation Protocol*): Efetua reserva de recursos ao longo do caminho da rede que o pacote percorre [27].
2. Controle de Fluxo: Através de escalonamento que utiliza uma disciplina de filas, faz um controle de fluxo de pacotes.
3. Controle de Admissão: É responsável por verificar se um novo fluxo que chega pode ter seus requisitos de QoS garantidos sem prejudicar garantias feitas a fluxos anteriores.
4. Policiamento: Verifica se as condições do fluxo estão de acordo com o contrato estabelecido na conexão, podendo descartar pacotes em casos de violações.

### 3.1.2 DiffServ

Esta subseção apresenta os principais conceitos da arquitetura de QoS DiffServ, que é utilizada como parte da arquitetura DiffMobil.

O DiffServ é uma solução de QoS que controla o tráfego IP, identificando pacotes que devem receber prioridade de transferência. A RFC 2475 [12] define o DiffServ como uma arquitetura para implementação de diferenciação escalável de serviços na Internet. Características de um serviço podem ser expressas em termos quantitativos tais como vazão, atraso, variação de atraso e perda, além de algumas prioridades relativas ao acesso aos recursos de rede. A diferenciação de serviços é desejável para atender a requisições de aplicações heterogêneas, a expectativas dos usuários em relação ao desempenho da rede e para permitir diferenciação de preços em serviços da Internet [38].

A arquitetura DiffServ é composta por mecanismos implementados em nós da rede, incluindo encaminhamento *per-hop*, funções de classificação de pacotes e funções de condicionamento de tráfego, tais como medição, marcação, formatação e policiamento.

Decisões de encaminhamento são tomadas pacote a pacote, de acordo com políticas estabelecidas pelo operador da rede. Pacotes de alta prioridade e baixa latência devem ser enviados a uma fila de alta prioridade, sendo transmitidos mais rapidamente que pacotes de baixa prioridade [70]. A arquitetura DiffServ propõe um tratamento agregado ao tráfego que passa pelos roteadores, de maneira a melhorar escalabilidade da rede e a diminuir a quantidade de informação de fluxo existente na arquitetura IntServ. O DiffServ categoriza o tráfego em diferentes classes. Os pacotes recebem uma marcação no *byte* de tipo de serviço do cabeçalho IP, de acordo com as classes às quais pertencem. Um padrão de 6 bits, denominado DSCP (*Differentiated Service Code Point*) [13], é atribuído ao campo TOS (*Type of Service*) ou TCO (*Traffic Class Octet*) do cabeçalho do IPv4 ou do IPv6, respectivamente [69]. A rede é dividida em

domínios e apenas os roteadores de borda assumem funções de classificação e marcação de pacotes. Após a classificação de pacotes nas bordas de um domínio, procedimentos específicos de encaminhamento de pacotes, chamados de PHB (*Per-Hop Behavior*), são aplicados em cada elemento da rede de maneira a prover parâmetros de QoS adequados a cada classe de serviço. A quantidade de informação a ser mantida em cada elemento, assim como a quantidade de informação de QoS a ser transmitida entre os elementos da rede DiffServ, é drasticamente reduzida, resultando em uma solução de QoS escalável.

Na arquitetura DiffServ, o tráfego que entra em um domínio da rede é classificado e condicionado nas fronteiras do domínio. Um SLA (*Service Level Agreement*) é estabelecido entre a rede e o domínio, para que o tráfego receba diferentes tipos de comportamento agregado (PHB) em cada domínio.

Os seguintes componentes são utilizados na classificação e no condicionamento de tráfego:

- **Classificador:** Seleciona pacotes do tráfego de entrada, baseado no conteúdo de uma parte do cabeçalho IP. Existem dois tipos: BA (*Behavior Aggregate*), que classifica os pacotes baseado somente no DSCP; MF (*Multi-Field*), que seleciona pacotes através de uma combinação de um ou mais campos do cabeçalho, tais como endereços de origem e de destino, números de porto de origem e de destino e interface de entrada.
- **Perfil de Tráfego:** Um perfil de tráfego especifica propriedades temporais de um fluxo de tráfego selecionado pelo classificador e provê regras para determinar se um pacote está de acordo ou não com o perfil. Este perfil indica que um pacote marcado com DS codepoint X, deve ser medido por um medidor token-bucket com taxa  $r$  e tamanho de rajada  $b$ .
- **Condicionador:** deve conter os seguintes elementos: medidor, marcador, formata-dor e “conta-gotas” Um fluxo de tráfego é selecionado por um classificador, que

direciona os pacotes até uma instância lógica de um condicionador de tráfego. O medidor verifica propriedades temporais de acordo com o perfil de tráfego do fluxo de pacotes e passa informações para os próximos elementos do condicionador, dizendo se o fluxo está ou não de acordo com o perfil. Um marcador estabelece um valor de *codepoint*, adicionando o mesmo a um comportamento agregado. Um formatador atrasa um ou mais pacotes de um fluxo de tráfego, para manter o fluxo de acordo com o perfil estabelecido. Um formatador tem um *buffer* de tamanho finito. Caso não haja espaço neste *buffer*, pacotes poderão ser descartados. Um “conta-gotas” descarta alguns dos pacotes do fluxo, de acordo com o perfil de tráfego. Este processo é conhecido como policiamento do fluxo. Um “conta-gotas” pode ser implementado como um caso especial de formatador de *buffer* de tamanho zero.

Classificadores e condicionadores estão normalmente localizados em nós de entrada e saída do domínio.

Cada comportamento agregado é identificado por um único *code point*. O Diff-serv propõe uma diferenciação no enfileiramento e encaminhamento de pacotes nos roteadores do domínio, tendo como base o *code point* colocado no campo DS do cabeçalho, no momento da entrada no domínio.

Dois grupos de comportamento agregado foram padronizados pelo IETF. Um deles é o EF (*Expedited Forwarding*) [35], com uma classe, e o outro é o AF (*Assured Forwarding*) [30], com quatro classes, cada uma contendo três níveis de precedência de descarte. As políticas atualmente usadas para marcar, enfileirar e encaminhar pacotes em roteadores de um domínio Diff-serv estão relacionadas a implementações específicas. O grupo PHB EF foi definido com a intenção de prover serviços prioritários, utilizando Diff-serv. O PHB EF é também conhecido como serviço *premium* ou de canal dedicado. É geralmente utilizado por tráfego com requisitos de baixa perda, baixo atraso e baixo

*jitter* (variação de atraso) e alta largura de banda. Por ser prioritária, a implementação do EF deve prover mecanismos de limitar o dano desse tipo de tráfego sobre os demais, enquanto os nós de borda têm a função de policiar e descartar o excesso.

O PHB AF tem por objetivo garantir a largura de banda mas não oferece garantias quanto ao atraso. Cada uma das quatro classes possui três níveis de precedência de descarte, determinando a importância do pacote. Isso permite a um nó congestionado, descartar pacotes de uma mesma classe de maneira diferenciada.

Os três primeiros bits do DSCP estão relacionados com a classe de transmissão. Os três últimos com a precedência de descarte - 010 para precedência mais baixa (último a ser descartado), 100 para média e 110 para a mais alta precedência (primeiro a ser descartado). As Tabelas 3.1.2, 3.1.2 e 3.1.2 a seguir apresentam os DSCPs recomendados para as classes de tráfego BE, EF e AF.

Precedência de descarte	Classe AF1		Classe AF2		Classe AF3		Classe AF4	
	Binário	Dec.	Binário	Dec.	Binário	Dec.	Binário	Dec.
Baixa	001010	11	010010	21	011010	31	100010	41
Média	001100	12	010100	22	011100	32	100100	42
Alta	001110	13	010110	23	011110	33	100110	43

Tabela 3.1: PHB AF – quatro classes de tráfego independentes

Precedência de descarte	Classe BE	
	Binário	Decimal
N/A	000000	0

Tabela 3.2: PHB Padrão – Best Effort

Precedência de descarte	Classe EF	
	Binário	Decimal
N/A	101110	46

Tabela 3.3: PHB EF – Tráfego *premium*

Na marcação AF, a partir do momento em que o pacote excede a taxa acordada

ou contratada CIR (*Committed Information Rate*), ele é marcado como verde, amarelo ou vermelho, dependendo em quanto o CIR for excedido. As cores determinam a prioridade de descarte, sendo a verde a menor e a vermelha a maior. Esse enfoque tem o objetivo de prover condições para o funcionamento de algum mecanismo de descarte como o RED (*Random Early Detection*) [34].

O mecanismo de controle de congestionamento deve manter a relação entre os diferentes níveis de precedência [16].

### 3.1.3 Especificação de QoS pelo 3GPP para Redes UMTS

Esta subseção apresenta os principais conceitos da arquitetura de QoS do 3GPP, cujo modelo inspirou a organização da arquitetura DiffMobil.

Serviços de rede são considerados fim-a-fim, ou seja entre entidades pares, que podem depender de vários sub-sistemas intermediários. Um serviço fim-a-fim pode ter uma certa QoS provida a um usuário da rede UMTS. Um serviço portador inclui todos os aspectos para permitir o provisionamento de QoS contratada. Alguns destes aspectos incluem sinalização de controle, transporte no plano de usuário e funcionalidades de gerência de QoS. Uma arquitetura em camadas, ilustrada pela Figura 3.1, é proposta pelo 3GPP para redes UMTS [1, 2, 3]. Nesta arquitetura, cada serviço portador de uma determinada camada oferece o seu serviço utilizando serviços das camadas inferiores [44].

Das camadas especificadas pelo 3GPP, pode-se ressaltar as seguintes funcionalidades:

- A operadora da rede UMTS provê QoS através da camada *UMTS Bearer Service*.
- A camada *Radio Access Bearer Service* provê o transporte confiável, com QoS, de dados de usuário e de sinalização entre a UM e o nó de borda da interface aérea



## Arquitetura de QoS no UMTS

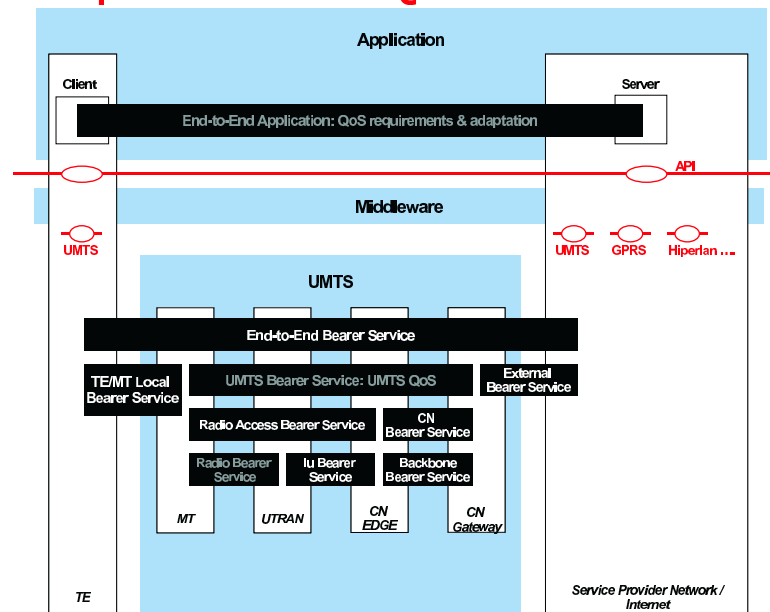


Figura 3.1: Arquitetura de QoS para Redes UMTS [3]

correspondente. Este serviço é baseado em características da interface aérea e é mantido para uma UM (unidade móvel) em movimento.

- Proteção a erros é um serviço provido pela camada inferior *Radio Bearer Service*.
- A camada *Core Network Bearer Service* controla e utiliza o *backbone* da rede para prover o serviço UMTS contratado. A rede central de pacotes do UMTS pode suportar diferentes serviços e uma variedade de níveis de QoS no *backbone*.
- As camadas *Radio Bearer Service* e *IU Bearer Service* são responsáveis pelo transporte de informações na interface aérea, provendo níveis diferenciados de QoS e utilizando as camadas inferiores.
- A camada *Backbone Bearer Service* não é específica do UMTS e, assim, pode usar um padrão existente. Esta camada corresponde aos níveis 1 e 2 da rede e é selecionada pelo operador para prover QoS da camada *Core Network Bearer Service*.

As funções de gerência de QoS de todas as entidades UMTS garantem o nível de QoS negociado para um determinado serviço. A seguir, serão descritas as principais funções de gerência de QoS da camada *UMTS Bearer Service*, no plano de controle:

- **Gerência de Serviço:** Coordena funções do plano de controle para estabelecer, modificar e manter o serviço pelo qual é responsável. Provê funções de gerência de QoS para o plano de usuário. Oferece serviços a outras instâncias e traduz atributos para requisitar serviços de camadas inferiores.
- **Função de Tradução:** Converte atributos UMTS em QoS de redes externas e vice-versa.
- **Controle de Admissão/Capacidade:** Mantém informações sobre todos os recursos disponíveis de uma entidade de rede e sobre todos os recursos alocados para serviços UMTS. Determina, para cada serviço requisitado, se há possibilidade de provisão de recursos pela entidade e reserva estes recursos em caso positivo. Verifica a capacidade das entidades para prover o serviço requisitado, ou seja, se o serviço requisitado é implementado e não está bloqueado por questões administrativas. Há também a possibilidade de se trabalhar com retenção de serviço através do controle de admissão.
- **Controle de Inscrição:** Verifica direitos administrativos do usuário da rede UMTS, para utilizar o serviço requisitado com os atributos de QoS específicos.

### 3.1.4 Soluções de QoS para GPRS

Esta subseção apresenta as principais características de QoS de redes GPRS.

O GPRS define vários perfis de QoS que podem ser requisitados por um usuário móvel. Estes perfis permitem que a rede ajuste o tráfego, objetivando alcançar o nível de serviço desejado. Garantir uma qualidade de serviço específica é difícil, pois cada

reserva de recurso para um determinado usuário depende da disponibilidade de recursos de rádio em uma rede móvel. O perfil de QoS de uma aplicação móvel consiste de vários componentes, listados a seguir [55]:

- **Classes de Precedência:** Aplicações recebem uma precedência de 1, 2, ou 3, sendo a precedência 1 a maior. Se uma aplicação possui uma precedência maior que a outra, seu tráfego receberá uma prioridade mais alta.
- **Classes de Atraso:** Aplicações podem requisitar classes de atraso que garantem atrasos médios. Estes atrasos incluem escalonamento e acesso a recursos de rádio.
- **Classes de Confiabilidade:** A aplicação pode requisitar diferentes classes de confiabilidade, dependendo de sua habilidade em lidar com dados corrompidos ou duplicados.
- **Classes de Vazão (*Throughput*):** A aplicação pode requisitar diferentes valores para vazão média e máxima. A vazão máxima corresponde a um intervalo de tempo curto, no qual a taxa de transmissão é a máxima possível. A vazão média corresponde à taxa de transferência durante um período de tempo maior, durante o qual pode haver ociosidade em determinados intervalos. A combinação dos dois permite que a aplicação determine a vazão desejada ao tráfego IP em rajadas.

As estações móveis requisitam os perfis de QoS desejados durante a fase de ativação de contexto PDP, na mensagem *Activate PDP Req.* A rede responde com a qualidade de serviço que pode prover. O SGSN informa à UM a QoS alocada, na mensagem *Activate PDP Accept Message.* A rede pode modificar a QoS durante toda a existência de um contexto PDP. Isto permite que o SGSN ajuste a largura de banda reservada a um determinado usuário, na medida em que ocorrem trocas no nível de recursos disponíveis. Quando não há tráfego por um determinado período de tempo, a conexão é encerrada, permitindo que os recursos de rádio sejam alocados a outros usuários.

### Medidas de QoS para a Camada de Rede

A forma mais precisa de se quantificar o comportamento de uma rede GPRS é através de uma análise da camada IP. Espera-se que a qualidade de serviço no GPRS seja correspondente ao desempenho da camada IP, o que permite a utilização de métricas da RFC 1242 [28] como base das medidas de QoS fim-a-fim para o GPRS. Estas medidas são listadas a seguir:

- **Vazão:** Taxa na qual uma seqüência de pacotes, ou bits, podem ser transferidos entre uma UM e uma interface Gi, em um dado intervalo de tempo. Normalmente é dada em bits por segundo.
- **Perda:** Razão entre pacotes perdidos durante a transmissão entre a UM e a Gi e o total de pacotes transmitidos. Determina a confiabilidade da transmissão dos dados pela rede.
- **Atraso:** O atraso de propagação de um pacote entre a UM e a Gi e entre a Gi e a UM. Este parâmetro tem grande relevância na análise de desempenho de protocolos IP com confirmação, como o TCP.
- **Jitter** ou variação de atraso: A variação estatística de medidas de latência sobre um número de transmissões de pacotes. Descreve o quão previsível é a latência de uma rede GPRS.
- **Seqüenciamento:** Comparação entre o número de pacotes que chegam em ordem e o número de pacotes que chegam fora de ordem. Um seqüenciamento fraco de pacotes pode levar a retransmissões no TCP, o que afeta, negativamente, a vazão do TCP.

As medidas de desempenho da camada de rede são indicadas para avaliação de desempenho da infra-estrutura da rede. Sua coleta deve ser realizada em situações

durante as quais o equipamento sendo testado não esteja sendo utilizado, pois os testes geram um grande tráfego [55].

### Fatores que afetam a QoS no GPRS

Existem diferentes fatores que definem a qualidade do serviço que um usuário recebe. O mais óbvio é a classe de serviço garantida a uma estação móvel, na ativação de contexto PDP, mencionada anteriormente. Este mecanismo permite que o GPRS e a estação móvel priorizem o tráfego. Apesar dos esforços da estação móvel e do GPRS, vários fatores podem degradar a qualidade de serviço do usuário, de maneira que ele fique insatisfeito. Estes fatores, citados a seguir, constituem o foco de otimizações em redes GSM e GPRS.

- **Qualidade do Sinal de Rádio:** O GPRS necessita mais de boa qualidade de rádio que o GSM. Quando um sinal de voz é digitalizado para transmissão sobre o GSM tradicional, pequenos erros e perdas de quadros sobre o enlace RLC não impedem o entendimento da conversa. O GPRS, sendo uma tecnologia de comutação por pacote, é direcionado ao uso de serviços IP tradicionais, cuja tolerância a perdas e a erros é bastante pequena.

Como mencionado anteriormente, o GPRS define diversos esquemas de codificação em interfaces de rádio. Isto permite que o usuário móvel tire vantagem da qualidade do sinal e envie dados utilizando esquemas de codificação eficazes, obtendo taxas de vazão mais altas. A rede pode efetuar trocas para esquemas de codificação mais redundantes, em casos de ocorrência de degradação do sinal de rádio. Nestes casos, a vazão pode sofrer grande impacto. No pior caso, o usuário móvel pode ter que migrar do CS-4 para o CS-1, o que significa que a vazão cai para menos da metade. Mesmo que o usuário móvel esteja usando CS-1 - o esquema de codificação mais redundante - a taxa de erro pode atingir níveis

elevados, resultando em erros incorrigíveis em quadros RLC. Neste momento, retransmissões podem piorar, ainda mais, a vazão, o atraso e o *jitter*.

- **Congestionamento da Rede:** Transações no GPRS, tais como ativação e desativação de contexto PDP, causam maior carga nos componentes de sinalização que transações no GSM tradicional. O HLR e o VLR devem atualizar a informação de estado do contexto PDP, que pode sofrer várias alterações em um único minuto, o que não ocorre à sinalização associada às chamadas de voz tradicionais no GSM. Quando a operadora da rede resolve cobrar por utilização da rede, ocorre uma outra sobrecarga na rede, devido ao número de registros sendo gerados. O crescimento no tráfego da rede terrestre GPRS acarreta alguns problemas. Projetistas de rádio devem disponibilizar mais antenas para contornar a necessidade de maior largura de banda na interface de rádio. Projetistas da rede podem precisar aumentar o número de enlaces físicos entre o SGSN e os BSSs. GSNs adicionais podem ser necessários para contornar o aumento do tráfego IP. A interface Gi - o ponto de acesso à Internet - deverá ser atualizada. Limitações de desempenho da rede terrestre poderão afetar a QoS no GPRS.
- **Congestionamento na Internet:** Um congestionamento em um roteador ou um servidor lento que provoquem dificuldades no *download* de tráfego de aplicações podem levar os usuários a reclamar do GPRS, mesmo que o problema não esteja nele.
- **Mudança de Área:** O GPRS é uma tecnologia orientada a pacotes, com roteamento dinâmico. Manter um roteamento dinâmico para protocolos orientados a pacotes é mais complicado que para conexões comutadas por circuito, como chamadas de voz GSM. Com a comutação por circuito, há pouco enfileiramento de informação, além de ser pouco provável a ocorrência de perdas durante tro-

cas de informações sobre roteamento. Em uma rede IP fixa tradicional, o que acarreta atraso de pacotes transferidos de um ponto a outro é o enfileiramento em pontos intermediários, tais como roteadores e *switches*. Como o GPRS usa roteamento dinâmico, o caminho que os pacotes tomam pela rede muda mais que em uma rede fixa. Cada vez que o caminho muda, também muda o atraso médio dos pacotes, introduzindo uma variação de atraso no desempenho das aplicações. Quando o atraso dos pacotes é altamente variável, existe uma chance dos pacotes chegarem fora de ordem. Do ponto de vista prático, manter conexões após *handoffs*, sem perda de dados, assegurando níveis de QoS satisfatórios é um dos principais desafios das operadoras de redes móveis.

## 3.2 Trabalhos Relacionados

Esta seção apresenta dois trabalhos da literatura que propõem soluções de QoS fim-a-fim em redes GPRS, que serão comparadas com a arquitetura DiffMobil.

### 3.2.1 Solução Proposta por Stuckmann

Stuckmann analisa ganhos de desempenho com QoS e o comportamento do Diffserv em uma rede IP central GPRS, utilizando o simulador GPRSim [9, 68]. Uma política de controle de admissão e uma estratégia de escalonamento são implementadas e analisadas por Stuckmann em [67]. Ele propõe em sua política de controle de admissão que requisições de ativação de contexto PDP sejam diferenciadas em dois níveis: *Premium*, *Standard*, *Best Effort* e *Conversational*, *Streaming*, *Interactive*, *Background*, para a classe *standard*. Em sua implementação, os recursos da rede ficam disponíveis para qualquer tipo de tráfego. Quando a carga da rede fica alta, os fluxos de aplicações que demandam mais QoS podem descartar fluxos de baixos requisitos de QoS, mas até um certo limite, onde P e I representam os limites apropriados. Quando os limites P e I

são atingidos, a QoS requisitada não é atendida, mas rebaixada para uma classe de prioridade inferior.

O escalonamento de blocos de rádio é feito em três estágios:

1. Blocos de rádio são distribuídos de acordo com a inscrição do assinante associado ao tráfego em três filas: *Premium*, *Standard* e *Best Effort*.
2. O segundo estágio só é válido para tráfego de assinantes do tipo *Standard*. Dependendo do perfil de QoS da aplicação de um pacote, uma fila adequada à classe de tráfego (*Conversational*, *Streaming*, *Interactive*, *Background*) é selecionada. Nas filas de pacotes, o escalonamento ocorre de acordo com um TBF (*Token Bucket Filter*) e um algoritmo *Round Robin*.
3. O terceiro estágio é construído por um mecanismo simples de prioridade, servindo às filas de classes de tráfego em ordem da mais alta prioridade (*Premium*) para a mais baixa (*Best Effort*).

A vazão, a utilização de PDCHs e o atraso de 95% dos pacotes foram as medidas de desempenho utilizadas por Stuckmann, em função do número de unidades móveis. Ele as considerou como melhores métricas de QoS para usuários de aplicações de e-mail (70%) e WEB (30%), simuladas em [67], com intervalo de 12 segundos entre elas. Utilizando CS-2, as taxas de transmissão variaram entre 5 e 25 kbps na simulação, onde 10% dos usuários foram modelados como *premium* e 90% como *standard*. Com a utilização do CAC proposto, 20% das aplicações *background* são encerradas quando o tráfego está mais intenso, sendo que a utilização dos canais não ultrapassa 75%.

O Diffserv pode ser usado no núcleo IP GPRS, suportando classes de serviço definidas para UMTS em [3]. Para isto, torna-se necessário fazer um mapeamento de classes de QoS no UMTS para classes PHB do Diffserv. Na literatura, vários mapeamentos são propostos. A Tabela 3.4 mostra o mapeamento proposto por Stuck-



mann [67].

Classe de Tráfego UMTS	Conformidade, Policiamento e Prioridade de Tratamento de Tráfego
Conversational	EF
Streaming	AF4x
Interactive	AF3x
Background	AF2x, AF1x, ou Best Effort

Tabela 3.4: Mapeamento de Classes de QoS do UMTS em Classes PHB do Diffserv [67]

A classe *conversational* exige prioridades mais altas e é normalmente associada a aplicações de tempo real. A classe *streaming* também requer alta prioridade sobre as demais, sendo menos exigente em relação à conversacional. As classes *interactive* e *background* são normalmente utilizadas por aplicações de transferência de arquivos e e-mails e podem ter menos recursos.

A classificação de pacotes deve ser feita nos nodos de borda da rede central (SGSN e GGSN), no momento da ativação ou modificação de contexto PDP. Isto requer algumas mudanças nas mensagens PDP Activate e PDP Modify, que devem incluir informações sobre o cabeçalho IP, usadas no processo de classificação dos pacotes.

### 3.2.2 Solução Proposta por Oliver

Os protocolos GQ-MAC (*Guaranteed QoS Medium Access Control*) e AP-CAC (*Adaptive Prioritized-Handoff Call Admission Control*) foram propostos em [76], com o objetivo de garantir QoS de maneira adaptativa. O protocolo GQ-MAC suporta reserva dedicada por sessão para a classe de tráfego *streaming* e reserva prioritária sob demanda para classes de tráfego conversacional e interativa. Alocação dinâmica de largura de banda, adaptada ao tamanho corrente da fila, é uma proposta para amenizar o impacto do tráfego em rajadas. O AP-CAC provê admissão dinâmica adaptativa por prioridade, através de diferenciação de requisições de *handoff*, que têm prioridades de

admissão mais altas que novas chamadas. O algoritmo utiliza, para isto, um esquema de proteção dinâmica de múltiplos canais. O algoritmo também adapta a capacidade reservada para tratar requisições de *handoff* baseado em suas condições atuais de tráfego de células vizinhas. Oliver [76] adota no núcleo IP o IntServ sobre uma arquitetura MPLS. A QoS fim-a-fim é provida através de um mapeamento entre IntServ e GPRS e, através de uma extensão do AP-CAC para o núcleo, para prover um CAC fim-a-fim unificado, com prioridades de admissão adaptativas e dinâmicas. No GQ-MAC, o canal S-PRACH é acessado por novas chamadas e *handoffs*, através do protocolo de acesso randômico *Slotted Aloha*, sendo que *handoffs* têm prioridade mais alta. Tráfego do tipo *streaming* recebe, neste algoritmo, reserva dedicada por sessão de canais do tipo PDTCH. Tráfego do tipo conversacional ou interativo deve reservar canais PDTCH sob demanda, usando o U-PRACH através do *Slotted Aloha*. Fontes de tráfego da classe *background* utilizam os PDTCHs não utilizados, através do algoritmo *Round-Robin*. O protocolo *Slotted Aloha* é usado para transmissão de sinais de requisição de acesso ao S-PRACH por novas chamadas e por *handoffs*. Estações móveis com pacotes de sinalização transmitem imediatamente no primeiro *slot* disponível. Se ocorre colisão, elas transmitem com probabilidade  $P_a$  no próximo *slot*. Requisições de *handoff* têm maior prioridade através de um  $P_a$  maior. O protocolo *tree limited-control* é usado em requisições de acesso a canal por classes de tráfego conversacional, porque provê um tempo de acesso determinístico. Inicialmente, ele permite que todas as unidades móveis de uma célula acessem um canal. Se ocorre uma colisão em um *slot*  $k$ , a estação base inicia um ciclo de contenção, que reserva o *slot*  $k+1$  para a UM que tiver o MSB (*Most Significant Bit*) de seus identificadores igual a "0". Se ocorre uma colisão novamente no *slot*  $k+1$ , o *slot*  $k+2$  é reservado para a UM que tiver os dois primeiros bits de seu MSB iguais a "00". Por outro lado, se uma requisição foi transmitida com sucesso no *slot*  $k+1$ , ou se o *slot* estava livre, o *slot*  $k+2$  será reservado para UMs cujos

bits MSB sejam iguais a “1XX”. Por permitir somente tráfego conversacional no ciclo de resolução de contenção, o protocolo garante um limite de 20 ms de atraso no acesso ao canal, considerando um quadro TDMA de 4,615 ms.

O algoritmo *Modified Slotted Aloha* é usado por fontes de tráfego interativo, para acesso ao canal U-PRACH. É semelhante ao *Slotted Aloha*, com as seguintes modificações: (1) Se ocorre uma colisão, um algoritmo de *backoff* exponencial binário é utilizado para reduzir a probabilidade de transmissão,  $P(x)$ , do próximo *slot* em 0,5. A probabilidade  $P(x)$  é calculada como uma função do tamanho da fila de saída e a fração de QoS que foi satisfeita. (2) Quando um ciclo de resolução de contenção está em progresso,  $P(x)$  é reduzido a 0, ou seja, UMs com tráfego interativo não participam do ciclo de contenção.

## O AP-CAC

O suporte a *handoff* exige re-roteamento de tráfego para uma nova estação base, através de um enlace sem fio diferente. Dependendo do esquema de *handoff* adotado, chamadas em curso podem sofrer vários níveis de degradação, tais como perda de dados, variação de atraso e erros na seqüência dos dados. Requisições de *handoff* competem com requisições de novas chamadas, para ganhar acesso à célula. Em geral, um protocolo de controle de admissão deve maximizar o número de fontes de tráfego admitidas, garantindo seus requisitos de QoS. Para protocolos MAC que suportam reserva dedicada por sessão, o protocolo de controle de admissão deve dar admissão prioritária a requisições de *handoff*, pois a interrupção dos serviços já iniciados é mais questionável pelo usuário do que o bloqueio de novas chamadas. Para o protocolo GQ-MAC, que suporta múltiplas classes de tráfego, o artigo propõe um AP-CAC que diferencia requisições de *handoff* com prioridades de admissão mais altas diferentes que as de novas chamadas, através de um esquema de guarda de canal. O protocolo proposto define que requisições de *handoff* associadas a reserva dedicada por sessão (tráfego *streaming*)

têm prioridade de admissão maior que as reservas sob demanda (tráfego conversacional, interativo e *background*). Um número fixo de canais é reservado em AP-CAC. Estes canais se dividem em dois grupos, ficando NG1 canais reservados para um grupo e NG2 para outro grupo. Um esquema de admissão de três níveis de prioridade é adotado: dois níveis de prioridade *premium* para *handoffs* e o nível básico para novas chamadas. As três classes de prioridade em ordem ascendente são: (1) Classe N, associada a novas chamadas, que só são admitidas se o número de canais livres excede  $N_{G1}$ ; (2) Classe H1, associada a requisições de *handoff* com tráfego conversacional, interativo ou *background*. São admitidas se o número de canais livres excede  $N_{G2}$ ; (3) Classe H2, associada a requisições de *handoff* com tráfego *streaming*. Admitidos sempre que houver canais livres. Cada nível de canais de reserva é continuamente adaptado, de acordo com a estimativa instantânea da taxa de chegada de requisições de *handoffs*, que depende do número de UMs ativas com chamadas em curso nas células vizinhas, do padrão de mobilidade em termos de velocidade e direção das UMs ativas, do tamanho das células onde as UMs estão e do tempo restante das chamadas em curso.

Para um número de canais  $N_T$ , os intervalos de variação de  $N_{G1}$  e  $N_{G2}$  no tempo são dados por:  $0 \leq N_{G2}(t) \leq N_{G2max} = \beta_2 N_T$ ,  $N_{G2}(t) \leq N_{G1}(t) \leq N_{G1max} = \beta_1 N_T$ , onde  $0 < \beta_2 \leq \beta_1 \leq 1$ . Na derivação de probabilidade de bloqueio, baseada em resultados apresentados por Rappaport em [31], para cada uma das três classes de prioridade, as chegadas de *handoffs* H2, de *handoffs* H1 e de novas chamadas seguem uma distribuição de Poisson, com taxas  $\lambda_{H1}(t)$ ,  $\lambda_{H2}(t)$  e  $\lambda_N(t)$ , respectivamente. As saídas também seguem uma distribuição de Poisson, com taxa constante igual a  $\mu$ . As probabilidades de bloqueio  $B_{H2}(t)$ ,  $B_{H1}(t)$  e  $B_N(t)$  são dadas por:

$$B_N(t) = \sum_{j=N_T-N_{G1}(t)}^{N_T} P_j(t), \quad (3.1)$$

$$B_{H1}(t) = \sum_{j=N_T-N_{G2}(t)}^{N_T} P_j(t), \quad (3.2)$$

$$B_{H2}(t) = P_{N_T}(t). \quad (3.3)$$

Com:

$$\lambda_{\Delta}(t) = \lambda_{H1}(t) + \lambda_{H2}(t) + \lambda_N(t), \quad (3.4)$$

$$\lambda_{\Omega}(t) = \lambda_{H1}(t) + \lambda_{H2}(t), \quad (3.5)$$

$$N_{\Delta}(t) = N_T(t) - N_{G1}(t), \quad (3.6)$$

$$N_{\Omega}(t) = N_T(t) - N_{G2}(t), \quad (3.7)$$

Oliver [76] define que:

$$P_j(t) = \begin{cases} \frac{\lambda_{\Delta}(t)^j P_0(t)}{j! \mu^j} & \forall j \in 1 a N_{\Delta}(t) = N_T(t) - N_{G1}(t), \\ \frac{\lambda_{\Delta}(t)^{N_{\Delta}(t)} \lambda_{\Omega}(t)^{j-N_{\Delta}(t)} P_0(t)}{j! \mu^j} & \forall j \in N_{\Delta}(t) + 1 a N_{\Omega}(t), \\ \frac{\lambda_{\Delta}(t)^{N_{\Delta}(t)} \lambda_{\Omega}(t)^{N_{G1}(t)-N_{G2}(t)} \lambda_{H2}(t)^{j-N_{\Omega}(t)} P_0(t)}{j! \mu^j} & \forall j \in N_{\Omega}(t) + 1 a N_T(t), \end{cases}$$

com

$$P_0(t) = \left[ \begin{array}{l} \sum_{i=0}^{N_{\Delta}(t)} \frac{\lambda_{\Delta}(t)^i}{i! \mu^i} + \sum_{i=N_{\Delta}(t)+1}^{N_{\Omega}(t)} \frac{\lambda_{\Delta}(t)^{N_{\Delta}(t)} \lambda_{\Omega}(t)^{i-N_{\Delta}(t)}}{i! \mu^i} \\ + \sum_{i=N_{\Omega}(t)+1}^{N_T(t)} \frac{\lambda_{\Delta}(t)^{N_{\Delta}(t)} \lambda_{\Omega}(t)^{N_{G1}(t)-N_{G2}(t)} \lambda_{H2}(t)^{i-N_{\Omega}(t)}}{i! \mu^i} \end{array} \right]^{-1}$$

### Resultados Obtidos por Oliver

As probabilidades assumidas para as classes  $H1$  e  $H2$  foram de 0.8 e 0.2, respectivamente. Com distribuição exponencial, Oliver assumiu duração de 150 s e 300 s para as respectivas classes não impedidas e tempos de 120 s e 240 s distribuídos exponencialmente para os tempos que mantinham um canal preso. Com taxas de chegada de 0.475, em média, considerando-se um sistema estável, os objetivos eram taxas de bloqueio  $B_{H2} = 0.003$ ,  $B_{H1} = 0.05$  e  $B_N = 0.1$ . Utilizando  $N_{G1} = 2$  e  $N_{G2} = 1$ , os resultados mostraram que  $B_N = 0.17$  se desviou em 70% do objetivo, enquanto  $B_{H2}$  e  $B_{H1}$  se mantiveram dentro do objetivo. Os resultados de Oliver, se comparados aos de [75], que usa guarda estática de canais, apresenta ganhos em relação a *handoffs* das classes  $H1$  e  $H2$ , que se desviaram em 25% e 300%, com  $B_{H1} = 0.06$  e  $B_{H2} = 0.05$ , respectivamente, mas apresenta perda em relação a bloqueios de novas chamadas, que ficaram 40% abaixo do esperado em [75].

## 3.3 Revisão Bibliográfica Complementar

Em sua tese de doutorado, Choi [18] analisa QoS em redes locais sem fio em dois níveis: (1) atraso, vazão e erros de pacotes e (2) estabelecimento e gerência de conexões. Para prover QoS em nível de pacote, ele desenvolveu uma arquitetura para LANs sem fio, definindo um protocolo MAC baseado em *polling*, para acessos *uplink*.

Uma arquitetura, para provisão de QoS em uma rede móvel baseada em IP (Projeto BRAIN), é a proposta apresentada em [15]. A arquitetura de QoS Brain, seguindo especificações do ISSLL (*IETF Integrated Services over Specific Link Layers*) *working group*, implementa um controle de tráfego pelo protocolo RSVP, em uma rede de serviços diferenciados.

Um paradigma de controle de admissão orientado a prioridade com renegociação

de QoS para serviços multimídia em UMTS é a idéia do trabalho apresentado em [47]. Nele, Li assume que classes de tráfego de baixa prioridade podem liberar largura de banda para uma classe de prioridade mais alta, possivelmente sacrificando o atraso estimado, quando necessário. Os pacotes do tráfego de mais baixa prioridade são armazenados em um *buffer* para transmissão posterior. Os pacotes são descartados quando o atraso excede o limite estimado.

O controle de admissão proposto por [78] utiliza a informação sobre mobilidade do usuário para reservar uma quantidade apropriada de recursos a *handoffs* potenciais, com a intenção de obter baixa taxa de descarte de conexões provenientes de *handoffs*. O controle de acesso múltiplo é tratado através de um esquema baseado em *polling* com prioridades a serviços heterogêneos, incluindo CBR (*Constant Bit Rate*), VBR (*Variable Bit Rate*) e ABR (*Available Bit Rate*). Limites de variação de atraso para conexões CBR e limites de atraso para conexões VBR são derivados e formam a base do controle de admissão proposto.

Uma política de controle de admissão baseada em um esquema de guarda de canais fracionária é proposta em [57]. Seu objetivo é diminuir o número de bloqueios de *handoffs* e de novas chamadas.

Seis algoritmos de controle de admissão são avaliados em [58]. Destes, três são propostos pelos autores: *Weighted Sum Scheme*, *Probability Index Scheme* e *Hybrid Control Scheme*. Eles são comparados a três algoritmos já existentes: *Reservation Scheme*, *Linear Weighting Scheme* e *Distributed Admission Control Scheme*. Como medida de QoS, são adotadas a probabilidade de bloqueio de nova chamada e a probabilidade de término forçado de conexão. As conclusões indicam melhor desempenho do *Hybrid Control Scheme* e um comportamento estável e bom do algoritmo de menor complexidade *Reservation Scheme*.

Uma avaliação de vários métodos de escalonamento de tráfego é mostrada em [56],

incluindo FIFO, SPS (*Static priority Scheduling*) e EDF (*Earliest Deadline First*), com o objetivo de satisfazer requisitos de atraso de algumas classes de QoS do GPRS. Resultados obtidos através de simulações mostram que EDF atende melhor aos requisitos em situações de congestionamento.

Usando um modelo de tráfego do UMTS e um simulador para protocolo IP específico, Lindemann [48] apresenta curvas de desempenho para várias medidas de QoS, ilustrando os benefícios do P-MIB (*Performance Management Information Base*), ferramenta de ajuste dinâmico no escalonamento de pacotes e controle de admissão na interface aérea UMTS.

Como resultado de um trabalho de pesquisa da Siemens, Becher [11] apresenta uma visão de um sistema por trás de uma rede de terceira geração, que compreende uma combinação de diferentes sistemas de acesso em uma plataforma de rede de acesso central baseada em IP. O trabalho visa dar contribuições à padronização do IMS (*IP based Multimedia Subsystem*), sob responsabilidade do 3GPP. Além disso, a arquitetura “Internet on Air”, proposta pela Siemens, engloba os conceitos de redes totalmente IP de quarta geração.

Em [10], há uma descrição do trabalho da Alcatel, que desenvolveu um conjunto de mecanismos de QoS tanto para redes fixas quanto para redes móveis de terceira geração. Para a infra-estrutura, a Alcatel desenvolveu mecanismos Diffserv em nível IP, funções MPLS no nível 2, controle de admissão de chamadas e gerência de QoS para a rede. No UTRAN, a Alcatel desenvolveu capacidades de rádio que contribuem para a QoS fim-a-fim, otimizando o uso de recursos de rádio e terrestres.

Em [14], é feita uma análise de problemas que ocorrem na combinação de Diffserv com Redes Móveis IP, que levou a dois componentes de integração destes dois protocolos: Adaptabilidade da comunicação e Protocolo de Sinalização Diffserv, tendo como objetivo possibilitar a um nodo móvel IP a indicação do tipo de serviço desejado, assim



como funções de contabilização de pacotes transmitidos.

Como resultado de um trabalho de pesquisa da Ericsson, em [49], sistemas WLAN (*Wireless LAN*) são analisados como acesso via rádio complementar às redes 3G. O artigo descreve a interação de QoS entre a rede central UMTS e uma rede de acesso via rádio baseada em HIPERLAN/2.

Como resultado de um trabalho de pesquisa da Nokia, em [43], é feita uma descrição de uma arquitetura emergente de QoS orientada a pacotes para a próxima geração de sistemas celulares. Partindo do princípio que um modelo de rede GPRS foi adotado no UMTS como base de serviços orientados a pacotes, o contexto PDP é tratado como ponto central no provisionamento de QoS para serviços de pacotes em UMTS. Neste artigo, é feita uma descrição dos serviços de suporte do UMTS, além dos módulos de QoS necessários à realização destes serviços, levando a uma arquitetura de referência, contendo estes componentes.

Também como resultado de um trabalho de pesquisa da Nokia, Dixit [21] discute cenários de evolução das redes móveis, descrevendo as classes de QoS do UMTS, além de QoS em interfaces aéreas 3G, redes de acesso via rádio e redes centrais.

Uma arquitetura para provisionamento de QoS em uma rede móvel baseada em IP é proposta como parte do projeto BRAIN em [15]. Desafios, dificuldades e motivações para o desenvolvimento da arquitetura são descritos no artigo, assim como possíveis extensões ao trabalho.

A arquitetura Camaleão é proposta em [39], tendo como objetivo a entrega de serviços fim-a-fim na Internet, sem sofrer interferência do número de domínios ou da tecnologia utilizada.

Em [74], é apresentada a necessidade e possíveis soluções para uma gerência de QoS dinâmica em ambiente VHE (*Virtual Home Environment*), conceito introduzido pelo 3GPP para UMTS, que possibilita a portabilidade de serviços personalizados aos

usuários entre terminais e fronteiras da rede.

MPLS e DiffServ são propostos em [16] como soluções de QoS na rede central GPRS do UMTS.

Em [19], é proposto um ambiente para controle de acesso em sistemas CDMA 3G. Ele é implementado como um controle de dois níveis: em uma escala de tempo de chegadas de chamadas (controle de admissão) e em uma escala de tempo de um *slot* (duração de um *frame*) (controle de fluxo). O controle de admissão garante que as requisições de QoS em termos de taxas de erros e atrasos, podem ser atendidas por todas os tipos de chamada do sistema: voz, vídeo, dados e sessões de busca na WEB. O mecanismo de controle de fluxo garante QoS pelo escalonamento de recursos disponíveis entre usuários de tempo real e não real. Três diferentes modos de implementação de controle de fluxo são apresentadas e comparadas pelos autores.

Em [32], um *framework* é apresentado para estudos de mapeamento de QoS entre vários níveis da pilha de protocolos de transporte de informações. Os autores propõem uma plataforma para avaliação de regras de mapeamento através de medidas de QoS para uma rede ATM e uma aplicação multimídia.

Orvis, em [55], explica características de QoS em GPRS, citando testes possíveis para medidas de QoS.

Uma proposta e análise de uma metodologia para prover garantias estatísticas de QoS estão em [73]. Um modelo analítico de DiffServ com escalonamento estático prioritário é usado no trabalho.

A proposta de uma nova técnica de análise de desempenho de redes móveis é proposta em [51]. Ela é baseada em tecnologias FDMA/TDMA (como GSM/GPRS), nas quais a utilização de duas bandas de frequência separadas levam a uma estrutura celular complexa com sobreposição de micro e macrocélulas. Tempos de duração e interrupção de chamadas são descritos por variáveis randômicas seguindo diversas distribuições. A

técnica de análise é baseada em Markov.

## 3.4 Resumo do Capítulo

Existem diversas soluções de QoS propostas para redes IP, dentre elas DiffServ, IntServ e MPLS. O DiffServ, adotado neste trabalho, faz classificação e marcação de pacotes IP. Esta marcação, EF, AF, ou BE, denominada DSCP, determina o comportamento agregado de pacotes em domínios de uma rede DiffServ.

As especificações de redes GPRS apresentam 4 classes de QoS: precedência, atraso, confiabilidade e vazão. As especificações de redes UMTS também apresentam 4 classes de QoS que serão adotadas como padrão neste trabalho: *conversational*, *streaming*, *interactive* e *background*. No próximo capítulo, será apresentada a proposta de mapeamento da arquitetura DiffMobil entre as classes de serviço GPRS, as classes de serviço UMTS e as classes de serviço DiffServ.

Duas referências da literatura apresentam propostas de QoS fim-a-fim para redes GPRS. Stuckmann [67] propõe a adoção de DiffServ no núcleo IP e de um algoritmo de controle de admissão e escalonamento por prioridades de classes de serviço na interface aérea. Limites  $P$  e  $I$  foram definidos para priorizar as classes *premium* e *interactive* no trabalho de Stuckmann. Através de simulações, ele apresenta resultados experimentais pouco consistentes, pois utiliza apenas dois tipos de aplicação: e-mail e navegação na web. Não apresenta dados de desempenho de rede como um todo e não segue o padrão de classes de serviço proposto por ele no início de seu trabalho. Ele consegue manter o nível de utilização da rede em 75%, bloqueando acesso de chamadas do tipo *background*, quando há congestionamento. Oliver [76] apresenta os algoritmos GQ-MAC e AP-CAC para fazer reserva dedicada de canais e controle de admissão dinâmico e adaptativo. Para o núcleo IP, Oliver propõe a adoção de MPLS e IntServ. Em seus resultados, obtidos através de simulações, Oliver mostra, como principal contribuição

de seus algoritmos, uma redução de bloqueios de *handoffs*. Apesar de Oliver apresentar um trabalho consistente, há a deficiência do controle de congestionamento, através da monitoração do desempenho da rede.

# Capítulo 4

## A Arquitetura DiffMobil

### 4.1 Descrição da Arquitetura

A arquitetura de QoS fim-a-fim DiffMobil herda características funcionais do padrão proposto pelo 3GPP para redes GPRS/UMTS [3], apresentando soluções para as camadas de rádio e do núcleo IP do GPRS. Estas soluções são propostas através da adoção de DiffServ no núcleo IP e da implementação dos seguintes novos módulos: negociação de perfil de QoS, controle de congestionamento (composto por controle de admissão, diagnóstico de rede e controle de fluxo), ilustrados pela Figura 4.1. As Figuras 4.2 e 4.3 representam diagramas UML de caso de uso e de dependência de módulos, respectivamente. Nas Figuras 4.2 e 4.3, a seta ( $\rightarrow$ ) representa uma dependência entre módulos. No diagrama de caso de uso, um assinante da rede requisita QoS através do módulo de negociação. Este, por sua vez, passa esta requisição a um *middleware* hipotético, que não foi implementado, que traduz os parâmetros do usuário em parâmetros de rede, passando-os ao módulo de diagnóstico, no pedido de ativação de contexto PDP. O módulo de diagnóstico, verifica o estado atual de tráfego da rede e permite ou não que o contexto PDP seja ativado para aquele usuário, com a classe de serviço solicitada. A partir deste procedimento, os módulos DiffServ, CAC e Controle de Fluxo, atuam em

seus segmentos baseados nesta classe de serviço negociada, que está no contexto PDP. Como o Controle de Fluxo estende a marcação DiffServ do tráfego *downlink* para a interface aérea, ele é um módulo que depende da implementação do DiffServ no núcleo IP. O DiffServ, por sua vez, depende das classes de serviço negociadas e marcadas no contexto PDP pelo módulo Diagnóstico e pela ativação de contexto PDP do SGSN. O controle de admissão baseado em guarda também depende destas classes de serviço.

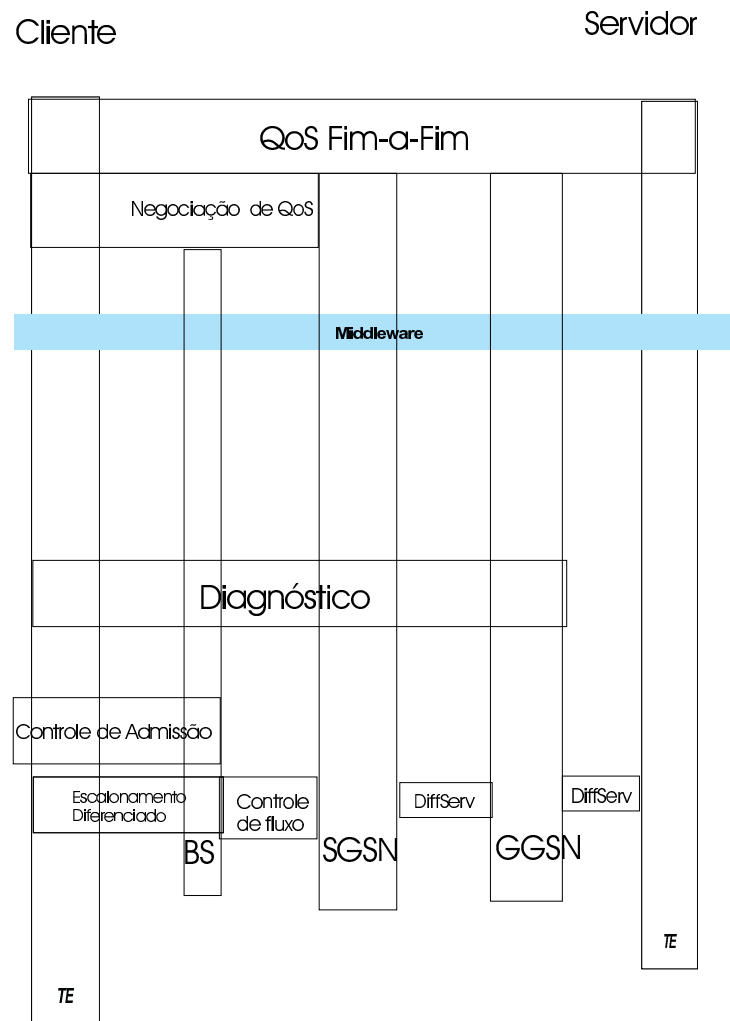


Figura 4.1: Arquitetura DiffMobil

As Figuras 4.1, 4.2 e 4.3 representam respectivamente: a arquitetura DiffMobil, um diagrama de caso de uso UML e um diagrama de dependência de módulos. Nelas destacam-se:

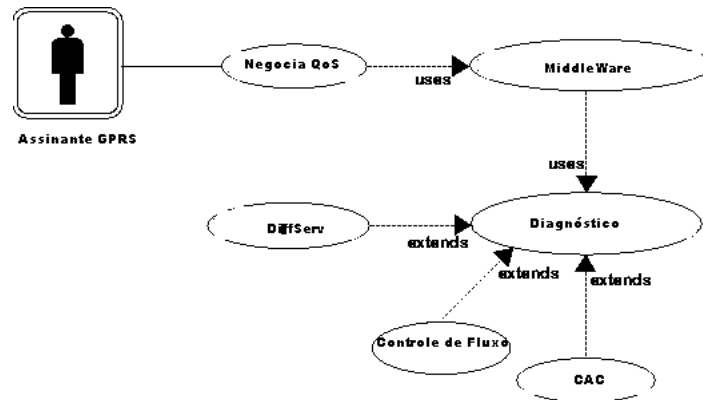


Figura 4.2: Diagrama UML de Caso de Uso DiffMobil

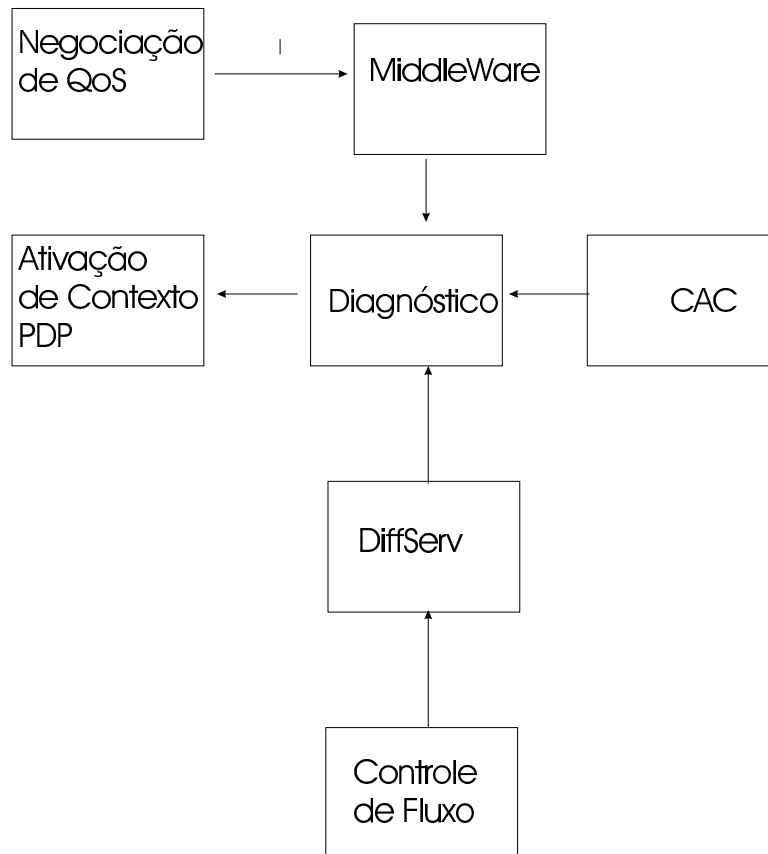


Figura 4.3: Diagrama UML de Dependência entre os Módulos DiffMobil

- Módulo de Negociação de Perfil de QoS: Este módulo cuida da negociação de QoS entre a rede e o usuário. Ele interage com os módulos de controle de admissão (CAC) e de diagnóstico, passando a eles o perfil requisitado pelo usuário.

Dependendo da resposta do diagnóstico, ele retorna ao usuário o perfil de QoS possível de ser atendido.

- Controle de congestionamento: implementado na estação base em dois níveis: controle de admissão e diagnóstico. Visa manter sob controle a alocação de recursos e o tráfego da interface aérea, entre a estação base e os diversos clientes de uma célula.
  1. Módulo de Controle de Admissão: responsável por atender ou não às requisições de alocação de canais por novas chamadas ou por *handoffs*. O controle de admissão, no DiffMobil, é implementado através de guarda de canais para a classe *conversational* e para *handoffs*. O número de canais de guarda foi estipulado em 2% para *handoffs* e em 1% para novas chamadas da classe *conversational*, padrão adotado na configuração inicial de [76]. Se o número de canais livres é superior a 3%, todas as chamadas são atendidas. Se apenas 3% dos canais estão livres, somente chamadas da classe *conversational* e *handoffs* são atendidas. Se o número de canais livres é inferior a 2%, somente *handoffs* são atendidos. Como o controle de admissão baseado em guarda não faz policiamento de tráfego, a priorização de algumas classes pode causar danos às demais. Por este motivo, a arquitetura DiffMobil implementa um módulo de diagnóstico de uso da rede, descrito a seguir.
  2. Módulo de Diagnóstico de Rede: Implementado no SGSN, este módulo faz um policiamento de todo o tráfego que entra na rede. Para este policiamento, o módulo mantém um controle de utilização da rede. Ele verifica a taxa de transmissão requisitada pelo cliente na ativação de contexto PDP e decide se a rede ainda comporta mais este tráfego, ativando ou não o contexto PDP. Quando o módulo de diagnóstico detecta *thresholds* de utilização da rede, ele começa a bloquear novas chamadas, seguindo a prioridade das



políticas implementadas, descritas adiante. Medir o desempenho de uma rede em tempo real não é uma tarefa trivial, principalmente quando se leva em consideração a escassez de recursos de redes sem fio. O módulo de diagnóstico deve manter um controle de todo o tráfego em um determinado segmento de uma rede GPRS, usando informações do controle de admissão e do perfil de QoS negociado para fazer estimativas do tráfego gerado em cada conexão e da utilização dos canais de comunicação. A utilização dos canais é uma informação muito importante, uma vez que já foi mostrado em [22, 23, 24] que a rede GPRS possui pontos de gargalo específicos, a partir dos quais medidas de contenção de tráfego devem ser tomadas. O algoritmo 1, descrito em alto nível, mostra como o diagnóstico controla o tráfego de entrada no SGSN.

Os *thresholds* adotados no módulo de diagnóstico foram obtidos através de experimentos exaustivos [23] cujo objetivo era o de congestionar a rede, identificando gargalos. Inicialmente, são propostos *thresholds* fixos, que podem ser adaptativos, dependendo de um padrão de comportamento de tráfego. O algoritmo 1 implementa a seguinte idéia: a admissão na rede GPRS é livre até a marca de 50% de utilização da mesma. Quando a rede atinge 50% de utilização, novas chamadas da classe *background* devem ser negadas. Quando a rede atinge 56%, além das chamadas da classe *background*, as chamadas da classe *interactive* também são negadas. O mesmo ocorre com as novas chamadas da classe *streaming*, que são rejeitadas se a utilização da rede ultrapassa 62%. Entre 62% e 68%, apenas novas chamadas da classe *conversational* e *handoffs* são admitidos. Os *thresholds* utilizados no algoritmo 1 foram extraídos dos resultados apresentados em [24]. Nele, experimentos de análise de desempenho fim-a-fim de uma rede GPRS simulada mostraram

pontos de saturação da rede para tráfego crescente, como pode ser visto nos gráficos das Figuras 4.4 e 4.5. Pelo gráfico, percebe-se que a rede começa a ficar congestionada em torno de 90 kbps (56%), quando a vazão média deixa de acompanhar o crescimento da taxa de geração de tráfego. Em 100 kbps (62,5%), ocorre um leve declínio e, a partir de 110 kbps (68,75%) degradação de desempenho.

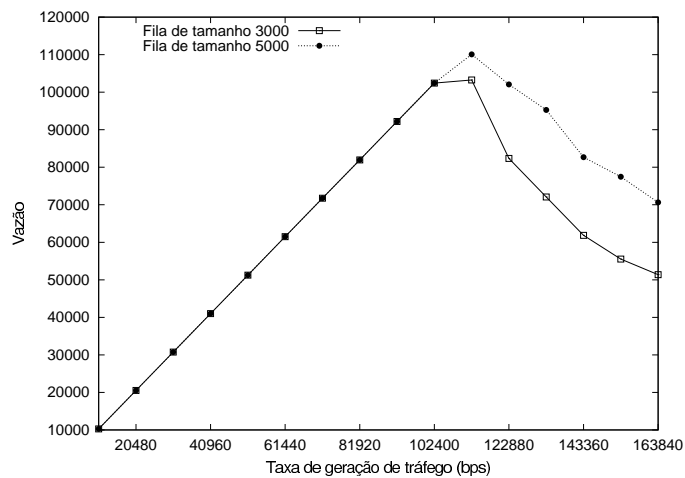


Figura 4.4: Vazão Média  $\times$  Taxa de Geração de Tráfego

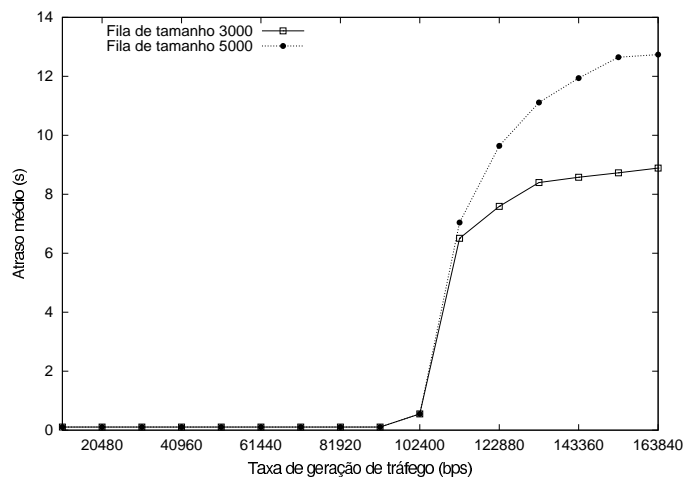


Figura 4.5: Atraso Médio por pacote  $\times$  Taxa de Geração de Tráfego

- **Módulo de Controle de Fluxo:** Este módulo visa marcar os quadros da interface aérea seguindo a marcação de pacotes que o DiffServ faz no núcleo IP. O controle

de fluxo promove o escalonamento diferenciado de quadros, de maneira similar ao DiffServ no núcleo IP. Ele está mais relacionado ao escalonamento de tráfego e visa evitar congestionamento na rede para fluxos de tráfego já admitidos. Por apresentar maior escassez de recursos que o núcleo IP de uma rede GPRS, a interface aérea deve ter um controle de congestionamento rígido, para garantir QoS fim-a-fim. A estação base deve ter este controle de congestionamento implementado de maneira integrada ao controle de congestionamento do núcleo IP, sendo esta uma das propostas da arquitetura DiffMobil. A idéia do controle de fluxo é a de aproveitar a marcação feita no núcleo pelo DiffServ e escalonar os pacotes da estação base seguindo esta marcação. Para isto, o serviço *relay*, da pilha de protocolos da estação base é usado para “traduzir” a marcação DiffServ dos cabeçalhos IP em marcação de quadros TDMA. A marcação DiffServ chega à estação base através do protocolo BSSGP, que transfere pacotes entre o SGSN e a estação base. Na estação base, o *relay* transfere esta marcação para as camadas RLC/MAC [4], para que o escalonamento de quadros siga a mesma prioridade que os pacotes IP. O controle de fluxo faz policiamento de tráfego na interface aérea, como DiffServ faz no núcleo IP. Ele formata o tráfego, direcionando pacotes para *buffers* de tamanho finito. Os algoritmos implementados serão descritos adiante.

Os novos módulos propostos neste trabalho visam a diferenciação de serviço fim-a-fim, de maneira integrada ao DiffServ do núcleo IP. Para isto, as classes de serviço de cada segmento da rede devem ser mapeadas para as classes utilizadas nos padrões 3GPP: *conversational*, *interactive*, *streaming* e *background*. A Tabela 4.1 mostra o mapeamento entre as classes de serviço do GPRS, do 3GPP e do DiffServ, com valores adotados para os perfis de QoS em uma rede GPRS, nas políticas implementadas pela arquitetura DiffMobil. No mapeamento, os valores adotados foram os mesmos das especificações do GSM [40]. A classe *conversational*, mapeada para a classe EF do

DiffServ, requer baixo atraso e baixa variação de atraso. A classe *streaming*, mapeada para a classe AF4 do DiffServ, também requer baixo atraso e baixa variação de atraso, apresentando requisitos de vazão garantida. A classe *interactive*, mapeada para AF3 e AF2, é sensível a perdas, requer garantia de vazão, mas não é intolerante a atraso e variação de atraso. A classe *background*, mapeada para AF1 e BE, não apresenta requisitos de QoS.

Definição de Perfis de QoS					
Classes GPRS	Atributos de QoS	Perfis de QoS Implementados			
		Conv. EF	<i>Stream.</i> AF4	<i>Interactive</i> AF3 e AF2	<i>Background</i> AF1 e BE
Precedência	Prioridade em congestionamento	Alta	Média	Baixa	N/A
Atraso	Latência	Limitado (<0,5s)	Limitado (<5s)	Limitado (<50s)	N/A
	Jitter	Baixo	Baixo	N/A	N/A
Confiabilidade	Probabilidade de perda de pacotes	Tolerável (< 10 <sup>-2</sup> )	Tolerável (< 10 <sup>-4</sup> )	Sensível a perdas (< 10 <sup>-9</sup> )	N/A
Vazão média e máxima	Vazão Nível Rajadas	N/A Alto	Garantida Baixo	Garantido Altíssimo	N/A N/A

Tabela 4.1: Mapeamento de Classes de QoS na Arquitetura DiffMobil

As políticas da arquitetura DiffMobil foram definidas por camada de serviço, conforme explicado a seguir.

## 4.2 Política Implementada pelo DiffMobil no Núcleo IP

Existem diversas soluções possíveis para se garantir QoS, provisionamento e engenharia de tráfego no núcleo IP de uma rede GPRS. O DiffServ, por apresentar boa escalabilidade [14], é a solução proposta para ser um componente do DiffMobil, garantindo

QoS no núcleo IP de uma GPRS. Esta seção apresenta um modelo para a adaptação do DiffServ no núcleo do GPRS. No núcleo IP do GPRS, deve haver diferenciação de pacotes nos roteadores de borda, com marcação de acordo com a classe à qual pertence o pacote. Um módulo de negociação de QoS é executado no momento de ativação de contexto PDP. A marcação de pacotes é feita de acordo com o perfil de QoS negociado na ativação de contexto PDP, da seguinte maneira:

1. Tráfego *Conversational*: Deve receber a marcação EF nos roteadores de borda da rede de núcleo do GPRS. A marcação EF, conforme mencionado no capítulo anterior, recebe maior prioridade de roteamento nos domínios DiffServ. Assim, como exemplos de aplicações conversacionais, podem ser citadas: transmissão de voz e telnet, que receberão prioridade EF, tendo como objetivo terem baixos atrasos na transmissão.
2. Tráfego *Streaming*: Deve receber a marcação AF4, nos roteadores de borda da rede de núcleo do GPRS. A marcação AF4 garante precedência de transmissão e vazão superior às demais aplicações marcadas com AF3, AF2, ou AF1. *Real Audio* é um exemplo de aplicação desta classe.
3. Tráfego *Interactive*: Deve receber a marcação AF3 ou AF2 nos roteadores de borda da rede de núcleo do GPRS. AF3 e AF2 são marcações que garantem um nível de precedência superior ao melhor esforço, mas que não competem por recursos com aplicações das classes *conversational* e *streaming*. Como exemplo de aplicação interativa, pode ser citada navegação Web.
4. Tráfego *Background*: Recebe marcação AF1 ou BE nos roteadores de borda da rede de núcleo do GPRS, correspondendo ao nível de melhor esforço na alocação de recursos. Envio de e-mail é um exemplo de aplicação desta classe.

A adaptação do DiffServ para GPRS faz um tratamento diferenciado dos tráfegos. O esquema de QoS tenta garantir uma utilização mais racional da banda. Ideal para situações onde é desejável prover um nível de qualidade de serviço aceitável, mesmo com cargas altas de tráfego prioritário (e.g., *streaming*) na rede.

Para adaptar a arquitetura DiffServ à rede GPRS um mapeamento entre os principais componentes das duas redes foi feito. Em resumo, os principais componentes do DiffServ são os roteadores de borda e os de núcleo. A rede GPRS é composta por três elementos principais: o SGSN, o GGSN e o roteadores de núcleo.

Os roteadores de borda do DiffServ, responsáveis pela marcação de Pacotes, são mapeados para o SGSN e GGSN. São eles a porta de entrada dos pacotes de dados para as redes GPRS. Os roteadores de núcleo da arquitetura de QoS, encarregados do encaminhamento dos pacotes, respeitando a marcação desses, serão mapeados para os roteadores de núcleo da rede GPRS. A figura 4.6 ilustra essa adaptação. Dessa forma, o conjunto SGSN, GGSN e roteadores de núcleo constituem um domínio DiffServ.

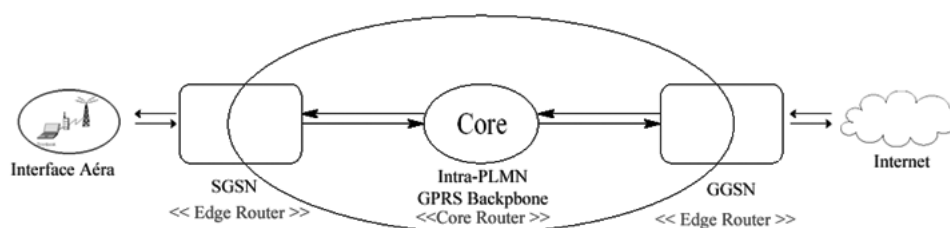


Figura 4.6: Proposta de adaptação DiffServ ao núcleo IP GPRS

Além desses componentes, outras características também devem ser mapeadas. A primeira delas se refere ao nível de diferenciação de tráfego, incluindo as suas classes e outros atributos QoS oferecidos pela operadora GPRS que devem ser mapeados para os grupos PHBs equivalentes. Definidos os grupos PHBs e suas características, os roteadores DiffServ são configurados de modo que eles possam produzir e lidar com esses PHBs.

A seguir, aspectos relacionados ao mapeamento QoS para os diferentes grupos PHBs e a estrutura dos roteadores relacionada ao suporte dessa função são apresentados. Em relação a essa estrutura, os componentes dos roteadores DiffServ e suas funcionalidades são descritas e, para cada um, uma arquitetura e um algoritmo são propostos.

### **Classificação e Mapeamento**

Cada DSCP identifica um comportamento agregado. O DiffServ propõe diferenciação no enfileiramento e encaminhamento de pacotes nos roteadores internos de um domínio, tendo como base o DSCP. Os grupos de comportamento agregado utilizados são aqueles padrões definidos pela IETF, apresentados no Capítulo 3.

Para o DiffServ ser utilizado no núcleo IP GPRS, o mapeamento das classes de QoS de uma rede GPRS para os grupos PHBs é necessário. Podemos utilizar uma adaptação das classes de tráfego definidas pelo *Universal Mobile Telecommunications System* (UMTS) para mapear os grupos PHBs. A Tabela 4.1 mostra esse mapeamento. A classificação dos pacotes será feita nos nós de borda da rede GPRS (SGSN e GGSN).

### **Elementos de Descarte**

RED (*Random Early Detection*) é um dos mecanismos de controle de fila mais utilizados em roteadores IP [54]. RED descarta pacotes aleatoriamente, baseado no tamanho médio da fila. Conforme explicado no capítulo 2, no DiffServ, o DSCP informa ao roteador qual é o PHB a ser aplicado no pacote. O PHB AF define quatro classes de

encaminhamento de pacotes. Os pacotes das diferentes classes são processados em filas físicas diferentes, gerenciadas por um mecanismo de escalonamento. Sob congestionamento, pacotes marcados com prioridade mais alta (menor precedência de descarte) devem ser os últimos a serem descartados e vice-versa. O algoritmo RIO (RED with In and Out) é uma extensão do RED [54]. Ele usa dois conjuntos de parâmetros para diferenciar descartes de pacotes de entrada e de saída. Para decidir quando descartar pacotes de saída, o RIO utiliza o tamanho médio da fila formada por pacotes In e Out. Para pacotes de entrada, ele utiliza o tamanho de uma fila virtual formada apenas por pacotes In. RIO-C (*Coupled*) [54] é uma extensão do RIO para possibilitar o tratamento de  $n > 2$  precedências. A probabilidade de descarte de pacotes com precedência  $1 \leq j < n$  depende do tamanho médio de uma fila virtual contendo somente pacotes com precedência entre 1 e  $j$ . Para pacotes com precedência  $n$  (prioridade mais baixa), a probabilidade de descarte é uma função da ocupação média da fila física (total). Nas simulações descritas no capítulo 5, o DiffServ [8] adotado no núcleo IP utiliza o algoritmo *Rio Coupled* (RIO-C) [61].

## Escalonador

O escalonador tem impacto no tratamento dos pacotes. É a peça chave na diferenciação dos serviços na arquitetura do DiffServ. Diversos esquemas de escalonamento estão presentes na literatura [?, ?], tais como *WFQ*, *Round Robin* e *SCFQ*.

WFQ (*Weighted Fair Queueing*) [41] é um algoritmo do tipo GPS (*Generalized Processor Sharing*) que trata o tráfego pacote a pacote. Por este motivo, ele é considerado de implementação complexa e inadequado para redes de alta velocidade. O SCFQ (*Self-Clocked Fair Queueing*) foi proposto para melhorar a complexidade de implementação do WFQ. Ele calcula *timestamps* para os pacotes de acordo com o nível de serviço do sistema. Andreozzi utiliza o SCFQ como algoritmo de escalonamento em sua simulação do DiffServ no ns-2 [8]. O critério de escolha foi a simplicidade de im-



plementação quando comparado ao WFQ [17]. Portanto, o escalonamento de pacotes no núcleo GPRS, proposto para a arquitetura DiffMobil segue o algoritmo SCFQ.

## 4.3 Política Implementada pelo DiffMobil na Interface Aérea

Duas políticas distintas são implementadas na interface aérea: uma para *downlink* e outra para *uplink*.

### 4.3.1 *Downlink*: Diferenciação de Serviços Fim-a-Fim

Considerando-se que o tráfego de maior volume normalmente é no sentido *downlink*, a arquitetura DiffMobil propõe uma extensão do DiffServ para o acesso aos escassos recursos de rádio. Para isto, além da aplicação de regras de mapeamento, há uma necessidade de se passar a marcação de pacotes feita nos roteadores de borda da rede IP para a estação base. Conforme mencionado anteriormente, com a marcação recebida, a estação base pode priorizar o acesso aos recursos de rádio, adotando diferenciação por classes no escalonamento de quadros TDMA. Desta maneira, a arquitetura DiffMobil permite uma esquema de prioridade de transmissão fim-a-fim.

Para a transmissão de marcação IP para TDMA, o SGSN passa ao protocolo BSSGP a informação necessária para o provimento de QoS a uma unidade móvel específica. A camada *relay* da estação base armazena os pacotes e repassa as informações do BSSGP às camadas RLC/MAC, para que a transmissão ocorra de maneira diferenciada. A primitiva de serviço utilizada é a RL-DL-UNIDATA que inclui os seguintes parâmetros [5]:

- Capacidade da UM: Define sua capacidade de rádio. Pode ou não estar presente na DL-UNIDATA PDU.

- Perfil de QoS: Define a vazão de pico, o tipo de PDU do BSSGP (sinalização ou dados), o tipo de quadro LLC (ACK, SACK, ou não), a classe de precedência e o modo de transmissão.
- *PDU Lifetime*: Define o período de tempo durante o qual uma PDU é considerada válida na estação base. Quando este tempo expira, a PDU é descartada.

A especificação GSM08.18 do 3GPP [5] deixa em aberto para implementação nas operadoras, o algoritmo do escalonador para acesso aos recursos de rádio, que leva em consideração vazão de pico, precedência, ou *PDU Lifetime*. A proposta para a arquitetura DiffMobil, é a de um escalonador baseado em precedência. O motivo desta escolha é a extensão das políticas baseadas em precedência implementadas no DiffServ do núcleo IP.

O 3GPP prevê a criação de PFC's (*Packet Flow Contexts*) para a estação base. Os fluxos de tráfego do SGSN até a estação base podem conter identificadores do tipo PFI (*Packet Flow Identifier*), associados a eles. A especificação TS48018 [7] apresenta quatro identificadores de fluxo: um para melhor esforço, um para sinalização, um para SMS e outro para mensagens para controle de enlaces. A arquitetura DiffMobil propõe um algoritmo que crie contextos de fluxo na estação base, seguindo a QoS negociada entre a unidade móvel e a rede e utilizada na diferenciação de pacotes no núcleo IP. A função PFM (*Packet Flow Management*), é prevista no SGSN, na camada *relay*.

O algoritmo 2 cria um controle de fluxo de tráfego entre o SGSN e a estação base. Ele está localizado na camada *relay* do SGSN e da estação base, correspondendo à função PFM. Inicialmente, ocorre a localização da unidade móvel. Logo após, ocorre a criação do contexto de fluxo. Para o algoritmo em alto nível foram utilizadas primitivas do BSSGP.

O algoritmo 3, descrito em alto nível, representa o recebimento de uma PDU LLC na camada *relay* da estação base. Este algoritmo distribui as PDUs recebidas em quatro

filas distintas, de acordo com as classes adotadas na arquitetura DiffMobil. Cada fila armazena pacotes de uma única classe, de acordo com o contexto de fluxo associado à PDU. As filas *conversational*, *streaming*, *interactive* e *background* recebem dados destas classes em ordem de chegada. O algoritmo 4, descrito em alto nível, promove o escalonamento de quadros, seguindo a precedência das filas:

O algoritmo 4 dá prioridade de escalonamento para as classes *conversational*, *streaming*, *interactive* e *background*, nesta ordem. Pela simplicidade de implementação, a idéia é similar à do Weighted Round Robin [42].

Os modelos de serviço do SGSN e da estação base, com os relacionamentos entre seus protocolos, estão ilustrados na figura 4.7. O SGSN passa os pacotes marcados para o protocolo BSSGP. Estes pacotes são encaminhados à estação base. O BSSGP recebe estes pacotes e repassa os quadros às camadas inferiores da estação base e a marcação é utilizada para controle de fluxo enviado às camadas RLC e MAC. O objetivo é que a camada MAC siga a prioridade associada a cada quadro, no envio de TBFs. O escalonamento de quadros seguirá, então, a mesma marcação feita no DiffServ do núcleo IP.

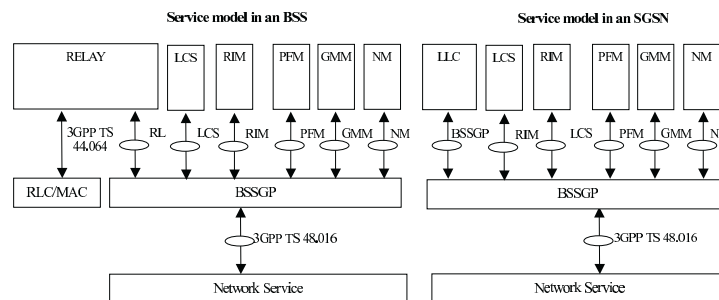


Figura 4.7: Serviços do BSSGP

### 4.3.2 *Uplink*

No sentido *uplink*, o tráfego, na maior parte das vezes, é de requisições de transmissões ou de serviços. A estação base deve ser capaz de distribuir os canais entre os usuários, seguindo uma prioridade de acesso. Para isto, o DiffMobil inclui, em sua camada de serviços de rádio, o algoritmo de controle de admissão, descrito a seguir.

O algoritmo de controle de admissão visa garantir qualidade de serviço de uma maneira adaptativa e dinâmica sobre enlaces sem fio. Ele deve ser implementado em cada estação rádio base do sistema da rede GPRS. O protocolo de controle de admissão visa maximizar o número total de fontes de tráfego admitidas, garantindo seus requisitos de QoS. A idéia do algoritmo apresentado neste trabalho é reservar e alocar canais GPRS sob demanda, minimizando a probabilidade de bloqueio de uma requisição de acesso de alta prioridade, sem deixar canais ociosos.

Para iniciar uma chamada, uma requisição *Call-Init* é enviada no canal S-PRACH, descrito no capítulo 2, através do protocolo *Slotted Aloha*. Quando recebe uma requisição de acesso, o algoritmo verifica se há recursos suficientes para suportar esta chamada sem degradar o serviço já garantido a outros e, se admissível, aloca *slots* para aquela chamada. O tratamento dado pelo controle de admissão a cada tipo de tráfego no DiffMobil, será descrito a seguir.

1. *Handoff*: O algoritmo trabalha com reserva de 2% dos canais para requisições de *handoff*, que têm prioridade sobre novas chamadas, independentemente da classe de tráfego.
2. Tráfego *Conversational*: Depois de admitido, demanda um canal somente quando tem dados a enviar. A requisição é enviada no U-PRACH. Se nenhum recurso estiver disponível, a estação rádio base vai alocar recursos guardados para esta classe. Se todos os recursos reservados estiverem alocados para tráfego conversacional, a estação rádio base rejeita a requisição e a unidade móvel é obrigada

a enviar uma nova requisição de transmissão, após descartar o primeiro pacote da fila. Quando não há dados a serem transmitidos, o recurso alocado é mantido por um período equivalente a três quadros TDMA, após o qual uma mensagem de liberação do canal é enviada.

3. Tráfego Interativo e *Streaming*: A unidade móvel deve enviar à estação rádio base a taxa de transmissão requisitada. A estação rádio base verifica quantos PDTCHs deve alocar para satisfazer a esta requisição e, dependendo do número de canais disponíveis, aceita ou não a transmissão. Cada unidade móvel contém uma fila finita de pacotes de saída. Assim que consegue acesso ao canal, a unidade móvel começa a transmitir estes pacotes da fila.
4. Tráfego *Background*: Fontes de tráfego da classe *background*, depois de serem admitidas, verificam no canal PAGCH quais *slots* estão alocados para elas em *frames uplink*. A estação base aloca para elas os PDTCHs não usados, seguindo o algoritmo *round-robin*.
5. *Handoff*: Possui prioridade máxima na alocação de canais, qualquer que seja sua classe de serviço. Se nenhum recurso estiver disponível, a estação rádio base vai alocar recursos guardados para esta classe. O objetivo é minimizar a taxa de bloqueio de chamadas iniciadas em outras células.

**Algorithm 1** Algoritmo de Controle de Tráfego

---

```

// Captura a QoS requisitada
usedBandwidth = usedbandwidth + qos
// Calcula o impacto da QoS requisitada na utilização da rede
usedPercentage = ((usedBandwidth/networkBandwidth)*100)
// Se a rede estiver com utilização inferior a 50% atende à requisição
if usedPercentage ≤ 50 then
    return(1)
else
    // Se a rede estiver com utilização entre 50% e 56% barra chamadas background
    if usedPercentage ≤ 56 then
        if (classe = Conversational) or (classe = Handoff) or (classe = Streaming) or
        (classe = Interactive) then
            return 1
        else
            usedBandwidth -= qos
            // Se a rede estiver com utilização entre 56% e 62% barra chamadas back-
            ground e interactive
            if usedPercentage ≤ 62 then
                if (classe = Conversational) or (classe = Handoff) or (classe = Streaming)
                then
                    return 1
                else
                    usedBandwidth = usedbandwidth - qos
                    return 0
            else
                // Se a rede estiver com utilização entre 62% e 68% barra chamadas
                background, interactive e streaming
                if usedPercentage ≤ 68 then
                    if (classe = Conversational) or (classe = Handoff) then
                        return 1
                    else
                        // Se a rede estiver com utilização superior a 68% não aceita novas
                        chamadas
                        usedBandwidth = usedbandwidth - qos
                        return 0
                    end if
                end if
            end if
        end if
    end if
end if
end if
end if

```

---

---

**Algorithm 2** Criação de Contexto de Fluxo e Envio de uma PDU *Downlink*

---

```
// requisita localização da unidade móvel
LCS-LOCATE.req(BVCI,NSEI)
// recebe localização da unidade móvel
LCS-LOCATE.res(BVCI,NSEI)
// repasse da marcação DiffServ para PFI
if TOS = EF then
    PFI = 1
else
    if TOS = AF4 then
        PFI = 2
    else
        if TOS = AF3 or TOS = AF2 then
            PFI = 3
        else
            PFI = 4
        end if
    end if
end if
// requisição de criação do contexto de fluxo
PFM-CREATE-BSS-PFC.ind(BVCI,NSEI)
// criação do contexto de fluxo
PFM-CREATE-BSS-PFC.res(BVCI,NSEI)
// envio de dados idownlink pelo BSSGP
BSSGP-DL-UNIDATA.req(BVCI, NSEI, LSP, DL-UNIDATA, PFI)
```

---

---

**Algorithm 3** Algoritmo de Distribuição de PDU's em Filas na Camada *Relay* da Estação Base

---

```
// recebimento de dados downlink pelo relay
RL-DL-UNIDATA.ind(BVCI, NSEI, LSP, DL-UNIDATA, PFI)
// PDUs LLC de identificador de fluxo 1 vão para fila da classe conversational
if PFI = 1 then
    conversational[ic++] = LLC-PDU
else
    // PDUs LLC de identificador de fluxo 2 vão para fila da classe streaming
    if PFI = 2 then
        streaming[is++] = LLC-PDU
    else
        // PDUs LLC de identificador de fluxo 3 vão para fila da classe interactive
        if PFI = 3 or TOS = AF2 then
            interactive[ii++] = LLC-PDU
        // PDUs LLC de identificador de fluxo 4 vão para fila da classe background
        else
            background[ib++] = LLC-PDU
        end if
    end if
end if
```

---



---

**Algorithm 4** Algoritmo de Escalonamento por Prioridade na camada *relay*

---

```
while 1 do
  // Peso de Escalonamento RR
  weigth = 10
  while weigth > 0 do
    // Peso 4 para escalonamento conversational
    if weigth > 6 and ic > 0 then
      RLC-DL-UNIDATA.req(conversational[ic - 1])
    else
      // Peso 3 para escalonamento streaming
      if weigth > 3 and is > 0 then
        RLC-DL-UNIDATA.req(streaming[ic - 1])
      else
        // Peso 2 para escalonamento interactive
        if weigth > 1 and is > 0 then
          RLC-DL-UNIDATA.req(interactive[is - 1])
        else
          // Peso 1 para escalonamento background
          if ib > 0 then
            RLC-DL-UNIDATA.req(background[is - 1])
            // Contabiliza peso para cada pacote escalonado
            weigth = weigth - 1
          end if
        end if
      end if
    end if
  end while
end while
```

---

## Capítulo 5

# Simulações, Resultados e Avaliação da Arquitetura

Neste capítulo serão descritos os modelos de simulação da rede GPRS e da arquitetura DiffMobil, implementados através da ferramenta de simulação ns-2 [62]. Quatro tipos de experimentos foram realizados. No primeiro, foi avaliada a eficácia do Diff-Serv no núcleo IP de uma rede GPRS. Este experimento procurou mostrar o impacto da diferenciação de serviços por prioridade, no núcleo IP da rede GPRS simulada. No segundo, o controle de admissão baseado em guarda foi implementado e testado. Resultados comparam o desempenho de aplicações em uma rede GPRS com melhor esforço e com o controle de admissão baseado em guarda. Neste experimento, a contribuição principal mostrada é a redução do percentual de chamadas bloqueadas. Uma avaliação do desempenho da rede com CAC é mostrada, indicando a necessidade de inclusão de um novo módulo, o de diagnóstico. No terceiro experimento, o módulo de diagnóstico foi acrescentado à rede simulada. O desempenho das aplicações foi medido e comparado ao desempenho obtido com controle de admissão baseado em guarda, mostrando que o diagnóstico consegue controlar o desempenho da rede, complementando o CAC. No quarto experimento, foi acrescentado à rede simulada um mecanismo de controle

de fluxo, que faz escalonamento de quadros TDMA baseado em marcações DiffServ recebidas do núcleo IP. Este módulo completa a arquitetura DiffMobil. Resultados deste experimento mostram uma avaliação de desempenho do GPRS com e sem DiffMobil, além de uma avaliação da eficácia do controle de fluxo, comparando os resultados aos obtidos no terceiro experimento. Com a inclusão do controle de fluxo, a arquitetura DiffMobil consegue garantir taxas de desempenho requisitadas pelos usuários. Para se obter cada ponto dos gráficos de resultados, todas as simulações foram executadas 33 vezes, gerando gráficos com um intervalo de confiança de 95%. Foram utilizados os módulos MAC-GPRS, *core IP* e Diffserv do ns-2 descritos em [36, ?, 8], respectivamente. A seguir, serão mostrados os resultados da inclusão de cada módulo de QoS no GPRS, de acordo com a seqüência dos experimentos realizados.

## 5.1 Modelo de Simulação

A rede GPRS foi simulada com duas PLMNs (Redes móveis terrestres públicas), que compõem o cenário dos experimentos, cada uma com um SGSN e um GGSN conectados através de um *backbone* IP. Cada PLMN contém 4 células. Cada célula foi configurada com 10 canais, na simulação. Este cenário teve como objetivo possibilitar a mobilidade dos assinantes entre células e entre PLMN's. O tráfego principal da rede ocorre entre os nós móveis e um nó correspondente de uma rede fixa. O controle de mobilidade é feito através do protocolo MIPv6 [63]. Os enlaces fixos foram modelados com largura de banda de 5 Mbps e os móveis com a largura dada pelo tipo de codificação na interface aérea, CS-4 [36] que, conforme foi descrito no capítulo 2, alcança taxas teóricas de até 171 kbps. O valor de largura de banda da interface aérea visa a obtenção da máxima largura de banda possível na interface aérea. O valor de 5 Mbps foi escolhido para a largura de banda do núcleo IP não representar um gargalo na rede. A figura 5.1 ilustra o cenário.

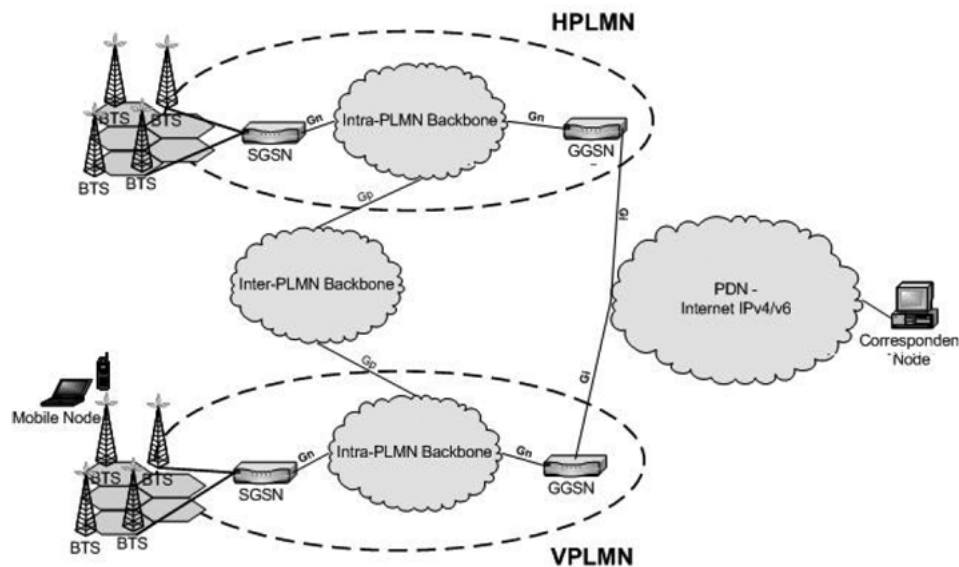


Figura 5.1: Cenário GPRS Modelado

## 5.2 DiffServ no Núcleo IP

O primeiro experimento utiliza parte do cenário simulado e procura refletir o impacto de aplicações menos prioritárias sobre as demais numa eventual sobrecarga da rede. Inicialmente, foram simuladas quatro aplicações, uma de cada classe de QoS DiffServ, gerando tráfego *downlink* entre o nó correspondente e quatro nós móveis de uma única célula. O impacto do crescimento do tráfego do tipo *background* foi avaliado sobre as aplicações das classes *conversational* e *streaming*. O comportamento da rede, neste primeiro experimento, foi observado com e sem aplicação do DiffServ no núcleo IP.

Os resultados deste experimento dão uma visão do potencial da adaptação do DiffServ no núcleo IP da rede GPRS, para diferenciação de aplicações. Nos experimentos executados com DiffServ no núcleo IP do cenário modelado, foram inseridas as quatro classes de tráfego adotadas como padrão na arquitetura DiffMobil, e apenas a menos prioritária (*background*) sofreu variação na sua taxa de geração.

Para se avaliar o comportamento do DiffServ no núcleo IP GPRS, foi simulado um

tráfego crescente do tipo *Background*, na rede GPRS, variando sua taxa de 2 kbps a 20 kbps, sem a aplicação de políticas de qualidade de serviço. As taxas das aplicações, neste experimento, foram de 20 kbps para *conversational*, 8 kbps para *streaming* e 7 kbps para *interactive*. Estas taxas procuraram mostrar como um tráfego de mais baixa prioridade (*background*) pode prejudicar um de mais alta (*conversational* e *streaming*), quando não há um mecanismo de diferenciação. Um tráfego CBR adicional de 4.9 Mbps foi inserido somente no núcleo IP (entre o SGSN e o GGSN), de maneira que a diferenciação de serviço no núcleo IP agisse em uma situação de congestionamento (98% de utilização). O resultado pode ser conferido no gráfico da figura 5.2.

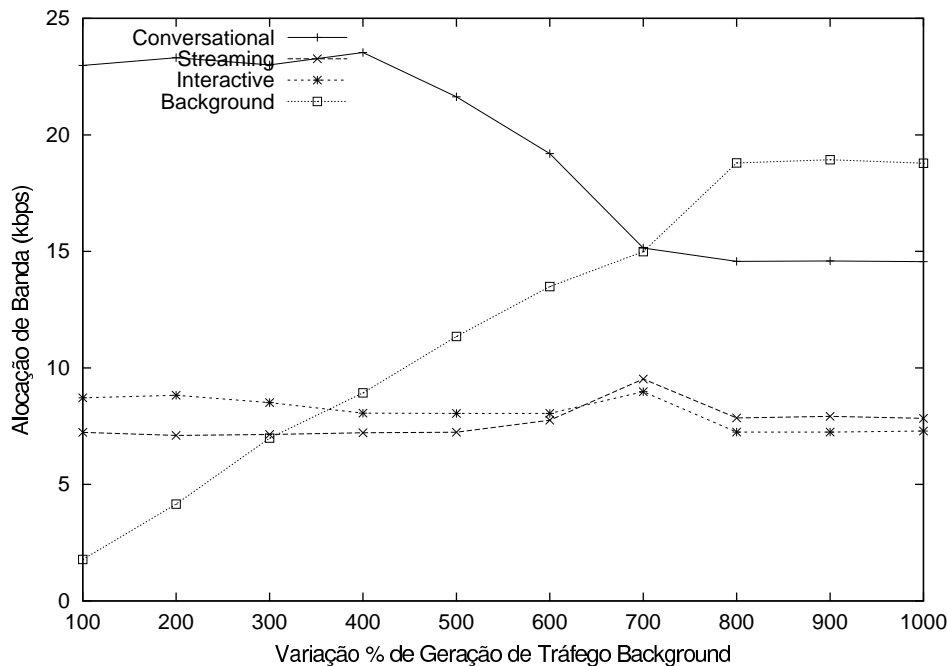


Figura 5.2: Alocação de Banda x Variação % de Carga Background sem DiffServ

Pelo gráfico da figura 5.2, observa-se que o tráfego *background*, menos prioritário, influencia os tráfegos *conversational*, *streaming* e *interactive*, que possuem maior prioridade. Esta influência é indicada pela queda de vazão das classes mais prioritárias. Esta queda de vazão é causada pela falta de priorização de tráfego e causa perda de desempenho de uma maneira geral, acarretando aumento de atraso de pacotes e variação

de atraso, uma vez que há aumento de tempo de espera em fila e descarte de pacotes em casos de filas cheias.

Com o DiffServ, a alocação de banda para tráfegos mais prioritários deve ser independente da carga dos demais. Esse resultado pode ser observado através da figura 5.3. Mesmo com um crescimento da carga do tráfego menos prioritário (CBR, do tipo *background*), a alocação de banda para os tráfegos mais prioritários não sofre influência. O gráfico da figura mostra um crescimento na alocação do tráfego *background* até 300 segundos. A partir daí sua alocação permanece constante. Este resultado é proporcionado pela adoção de uma política do QoS, que limita a banda dos tráfegos a fim de manter uma qualidade de serviço para aplicações que requisitam mais QoS, principalmente no que se refere a largura de banda.

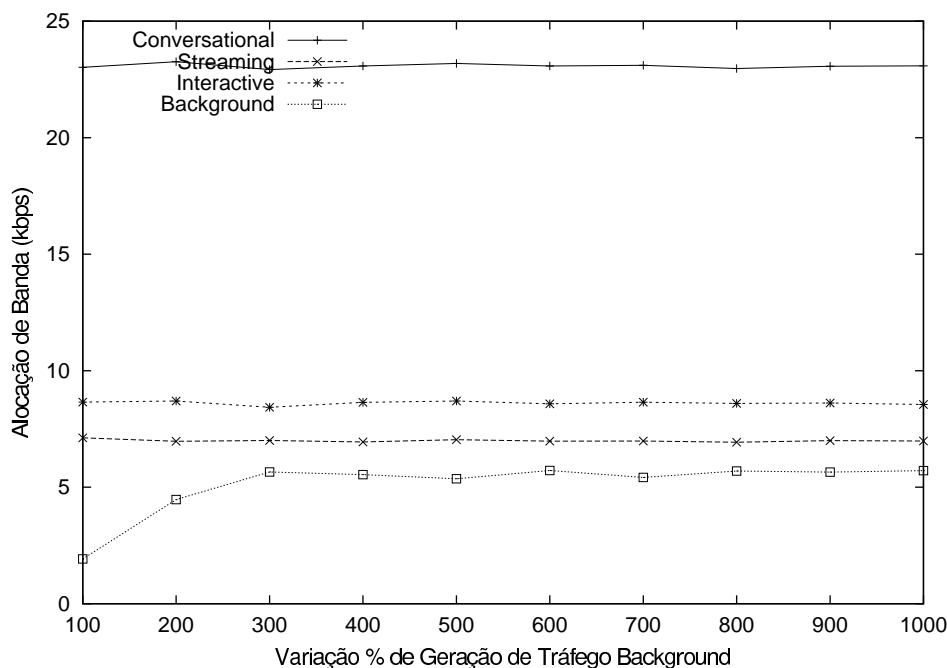


Figura 5.3: Alocação de Banda x Variação % de Carga Background sem DiffServ

Estes experimentos foram realizados com o intuito de avaliar a eficácia da adaptação do DiffServ no núcleo IP, tendo como roteadores de borda SGNS's e GGSN's. Como, após sua adoção, o tráfego *background* não degradou o desempenho e não prejudicou a

alocação de banda das classes mais prioritárias, comprovou-se, no cenário simulado, a eficácia do DiffServ, que é um dos módulos da arquitetura DiffMobil. Os experimentos com DiffServ utilizaram o protocolo MIPv6 que, como mostra Rosalen [63], otimiza o roteamento de redes GPRS, proporcionando redução de até 50% em atraso médio dos pacotes.

### 5.3 Controle de Admissão Baseado em Guarda e Diagnóstico da Rede

Inicialmente, somente o módulo de controle de admissão baseado em guarda de canais (CAC) foi implementado. Como foi mencionado anteriormente, o objetivo desse módulo é maximizar a alocação de recursos da rede, minimizando o bloqueio de acesso a canal por novas chamadas da classe *conversational* e de *handoffs*, que requerem rápida conexão com a rede. A classe *conversational*, por ser considerada prioritária, e a classe de *handoffs*, devido à necessidade de se manter uma conexão iniciada em outra célula anteriormente. Como o CAC não visa garantir desempenho das aplicações sobre a rede GPRS, a arquitetura DiffMobil implementa um módulo adicional ao controle de admissão: o de diagnóstico da rede. Este módulo, conforme explicado no capítulo anterior, mantém um diagnóstico do tráfego da rede a todo instante, bloqueando requisições de ativação de contexto PDP em casos de congestionamento, seguindo a prioridade de classes da arquitetura DiffMobil. Os experimentos relacionados aos módulos CAC e diagnóstico são descritos nas seções a seguir.

## 5.4 CAC: Controle de Admissão Baseado em Guarda

O controle de admissão foi implementado na arquitetura DiffMobil tendo como objetivo diminuir a taxa de bloqueios de acesso de *handoffs* e de novas chamadas da classe conversacional. Para isto, foi implementada uma guarda de canais para as novas chamadas da classe *conversational* e para *handoffs*. Oliver [76], em seu cenário inicial, reservou, em um conjunto de 100 canais, 3 para *handoffs* e não fez reserva para a classe prioritária *conversational*, ou seja, 3% dos canais foram reservados para *handoffs*. O controle de admissão da arquitetura DiffMobil, baseado nestes valores, faz reserva de 2% dos canais para *handoffs* e de 1% para novas chamadas da classe *conversational*. A opção por estes valores levou em conta a necessidade de reserva para novas chamadas prioritárias (*conversational*), uma vez que taxa de bloqueios de acesso de novas chamadas obtida por Oliver foi elevada, entre 4% e 17%.

A implementação do CAC aloca o máximo possível de *slots* vagos, preservando apenas os canais reservados. Esta abordagem não leva em consideração a degradação de desempenho causada por um excesso de transmissões simultâneas. Em consequência disso, há uma tendência de congestionamento, com todos os *slots* de todos os canais ocupados, principalmente com tráfegos prioritários, muitas vezes com taxas de transmissão elevadas, como no caso de aplicações *Real Audio*, da classe *streaming* provenientes de *handoffs*. Nos experimentos, foi simulada uma taxa crescente de chegada de unidades móveis: a cada quatro novas chamadas (uma de cada classe), chega um *handoff* a uma célula. Com esta taxa de chegada crescente, há um aumento do tráfego e uma tendência de esgotamento dos recursos disponíveis. Tráfego de aplicações do tipo Telnet, *Real Audio*, exponencial e CBR, correspondentes às classes *conversational*, *streaming*, *interactive* e *background*, respectivamente, foram simulados. As taxas variaram de 4 kbps a



20 kbps. Para uma célula, foram simulados dez canais de comunicação. Com esta taxa de chegada crescente, há um aumento do tráfego e uma tendência de esgotamento dos recursos disponíveis. Tráfego de aplicações do tipo Telnet, *Real Audio*, exponencial e CBR, correspondentes às classes *conversational*, *streaming*, *interactive* e *background*, respectivamente, foram simulados. As taxas variaram de 4 kbps a 20kbps. Para uma célula, foram simulados dez canais de comunicação. As Figuras 5.4 e 5.5 apresentam, como exemplo, uma amostra da distribuição de eventos da simulação por classe de aplicação, em 500 segundos de congestionamento (100% de utilização), sem e com CAC, respectivamente, em um canal. Em 500 segundos, ocorrem em um canal eventos que geram tráfego das classes *conversational*, *streaming*, *interactive*, *background* e dois tipos de *handoff*: *background* e *streaming*. Os eventos foram gerados em ordem aleatória, de maneira que se obtivesse uma carga heterogênea em um canal, com disputa entre classes de prioridades diferentes. No ambiente de simulação sem CAC, conseguem acesso imediato ao canal: três aplicações *background*, duas aplicações *interactive*, uma *conversational* e um *handoff background*, ocorrendo bloqueio de uma aplicação *conversational*, uma *background* e um *handoff streaming*. No ambiente de simulação com CAC, como há reserva para *conversational* e *handoff*, conseguem acesso imediato ao canal: duas aplicações *background*, uma aplicação *interactive*, duas *conversational*, um *handoff background* e um *handoff streaming*, ocorrendo bloqueio de duas aplicações *background* e uma *interactive*.

Conforme ilustra a Tabela 5.1, os resultados de uma simulação de 500 segundos, mostram que, para a simulação de melhor esforço, 15% das tentativas de acesso de novas chamadas e 10% de *handoffs* foram bloqueados, em média. Para a simulação da guarda de canais do DiffMobil, 1% da classe *conversational* e 4% de *handoffs* foram bloqueados, em média, contra 20% de bloqueios de novas chamadas sem prioridade. Oliver [76], obteve bloqueio de novas chamadas entre 4% e 17% e bloqueio de *handoffs*

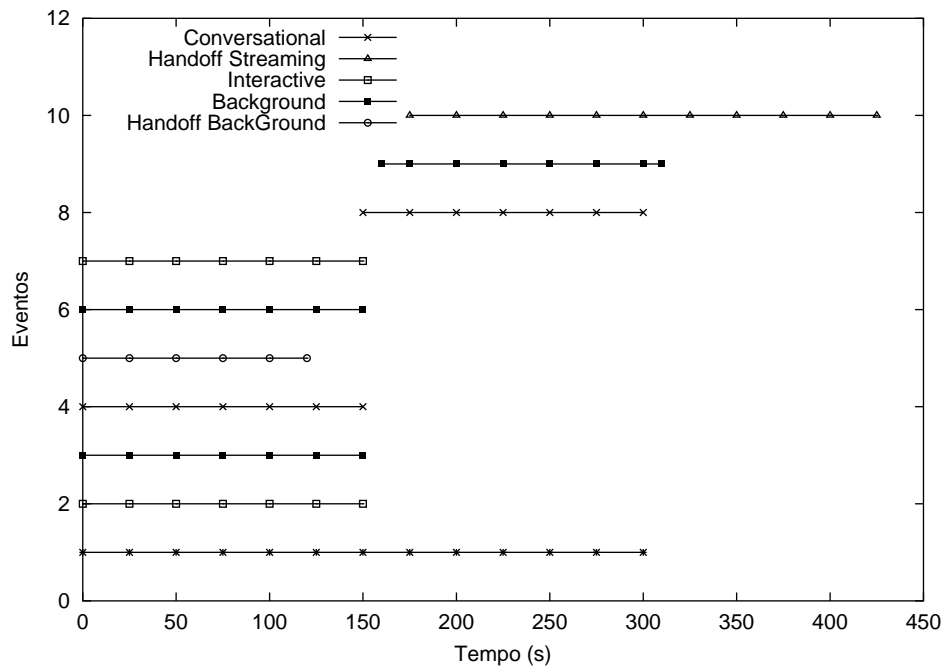


Figura 5.4: Eventos da Simulação sem CAC

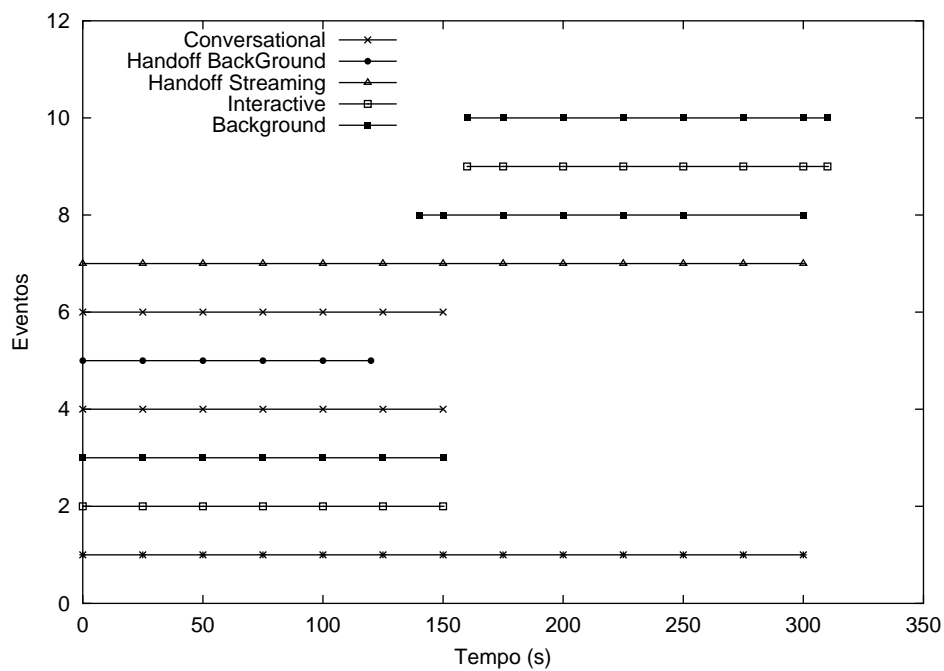


Figura 5.5: Eventos da Simulação com CAC

entre 1% e 4%. As comparações são feitas considerando-se um congestionamento da rede, com utilização de 100% dos canais livres. A vantagem da proposta da arquitetura

DiffMobil é a reserva para a classe *conversational*, com distinção entre as demais novas chamadas, o que o AP-CAC [76] não faz. Esta distinção deve-se às características das aplicações da classe *conversational* que, muitas vezes, têm necessidade de acesso imediato e em tempo real aos recursos da rede.

CAC	Bloqueios		
	<i>Handoffs</i>	<i>Conversational</i>	Novas Chamadas
Melhor Esforço	10%	15%	15%
<i>Oliver</i>	entre 1% e 4%	entre 4% e 17%	entre 4% e 17%
DiffMobil	4%	1%	20%

Tabela 5.1: Comparação entre a Arquitetura DiffMobil, Melhor Esforço e o AP-CAC de Oliver

Experimentos compararam o desempenho das aplicações em um ambiente com CAC e em outro com melhor esforço. O comportamento da simulação revelou o que se esperava: mais bloqueios de novas chamadas *interactive*, *streaming* e *background* e menos bloqueios de *handoffs* e novas chamadas da classe *conversational*. Os resultados indicaram que houve a redução de bloqueio de novas chamadas da classe conversacional e de *handoffs*, com uma utilização superior da rede por estas classes prioritárias, mas houve também uma degradação no desempenho, com aumento de atraso médio e variação de atraso, para a maioria das aplicações. Isto se explica pelo fato do CAC minimizar o bloqueio de classes prioritárias, mas não controlar o nível de utilização da rede. As próximas subseções apresentam comparações de resultados de análise de desempenho do ambiente simulado sem e com CAC. De uma maneira geral, conforme será mostrado, com a guarda implementada pelo CAC, a utilização da rede passa a ser maior por aplicações de classes prioritárias, o que garante vazão para estas aplicações. Porém o CAC não oferece garantias de atraso e variação do atraso, como pode ser visto a seguir.

### 5.4.1 Atraso Médio no tempo, para diferentes tipos de tráfego

Foram realizados experimentos para medir o atraso médio dos pacotes sem e com controle baseado em guarda, conforme ilustram os gráficos das figuras 5.6 e 5.7. Como o CAC não controla o nível de congestionamento da rede, o fato de aplicações com taxas de transmissão mais elevadas terem acesso garantido aos canais faz com que ocorra uma degradação de desempenho. Como pode ser visto nos gráficos das figuras 5.6 e 5.7, o atraso médio por pacote das classes *conversational* e *streaming* pioram com a introdução do CAC, passando de 0,4 para 1 segundo e de 0,5 para 1,2 segundos. Este resultado indica a necessidade de se introduzir na arquitetura um módulo que controle o desempenho da rede, que é o módulo de diagnóstico, que será descrito na seção 5.5.

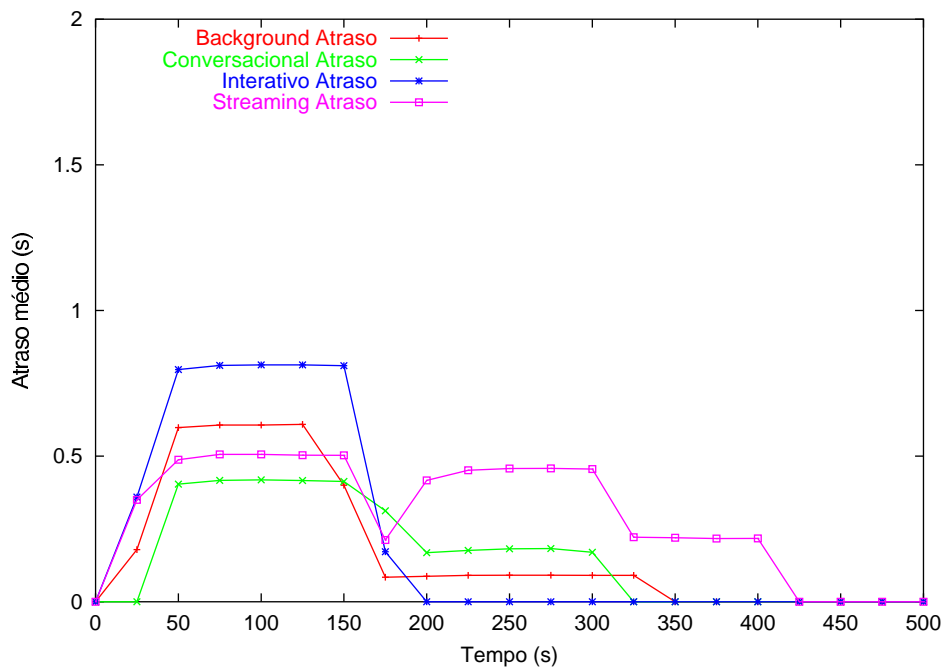


Figura 5.6: Atraso Médio sem CAC

### 5.4.2 Variação do Atraso Médio no Tempo

Foram realizados experimentos para medir a variação de atraso dos pacotes com admissão baseada em guarda, conforme ilustram os gráficos das figuras 5.8 e 5.9. Uma

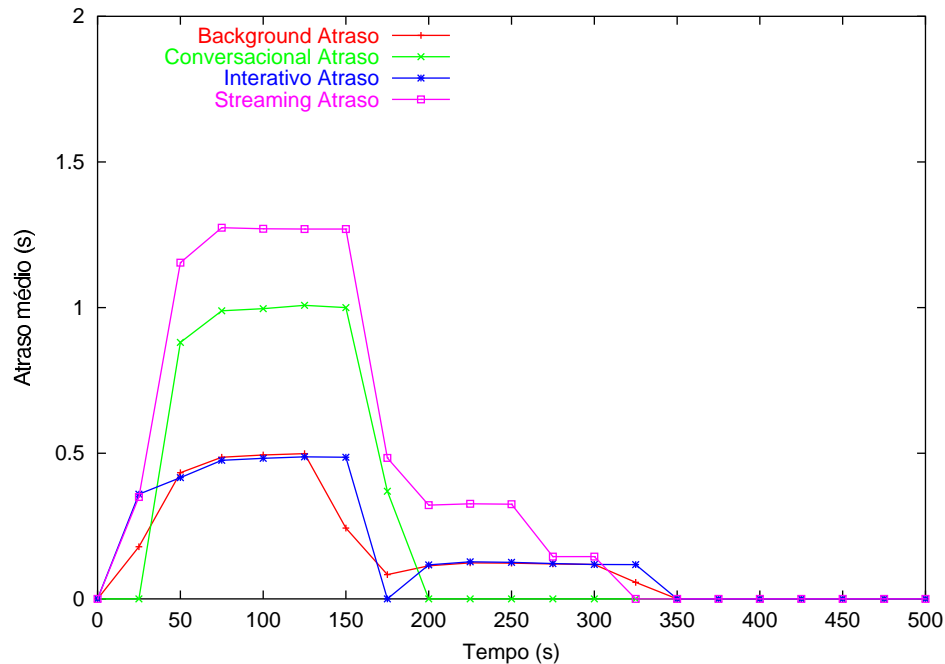


Figura 5.7: Atraso Médio com CAC

degradação do desempenho das classes prioritárias pode ser notada após a introdução do CAC. Como principalmente a classe *streaming* é sensível à variação de atraso, este resultado indica, mais uma vez, a necessidade de se implementar o módulo de diagnóstico. As classes *interactive* e *background*, apesar de não serem sensíveis à variação de atraso, foram incluídas nos gráficos para efeitos comparativos com as demais classes.

### 5.4.3 Vazão Média no tempo, para diferentes tipos de tráfego

Foram realizados experimentos para medir a vazão obtida pelas classes de aplicações com admissão baseada em guarda, conforme ilustram os gráficos das figuras 5.10 e 5.11. Como pode ser observado no gráfico, a vazão de pico para aplicações das classes *conversational* e *streaming* aumenta de 4 kbps para 8 kbps e de 9 kbps para 17 kbps, respectivamente. Em compensação, houve uma queda na vazão de pico para aplicações das classes menos prioritárias *interactive* (de 9 kbps para 4 kbps) e *background* (de 4 kbps para 2 kbps). Este fato se explica pela falta de disponibilidade de canais para

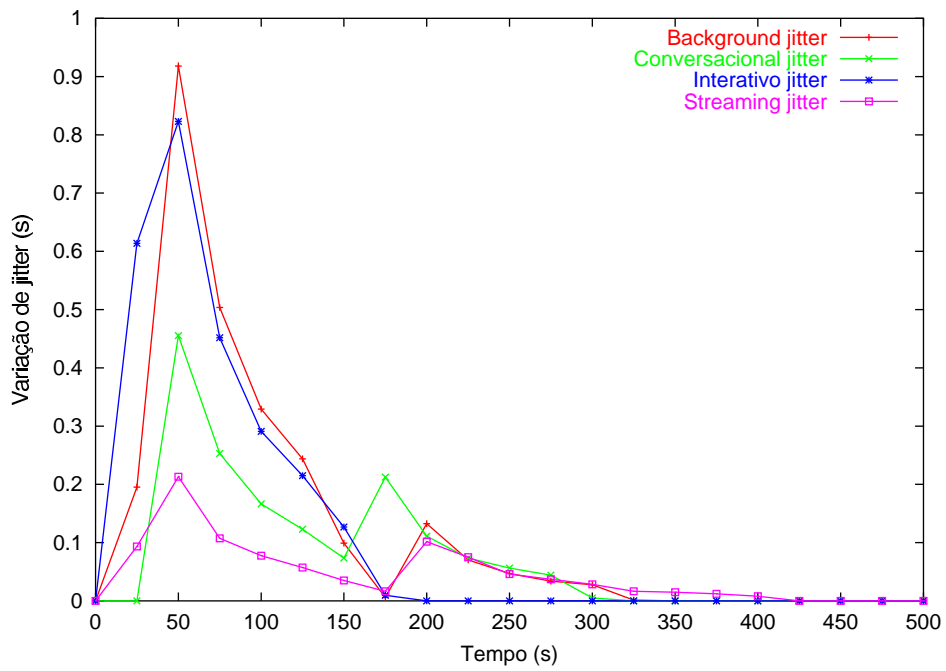


Figura 5.8: Variação do Atraso Médio sem CAC

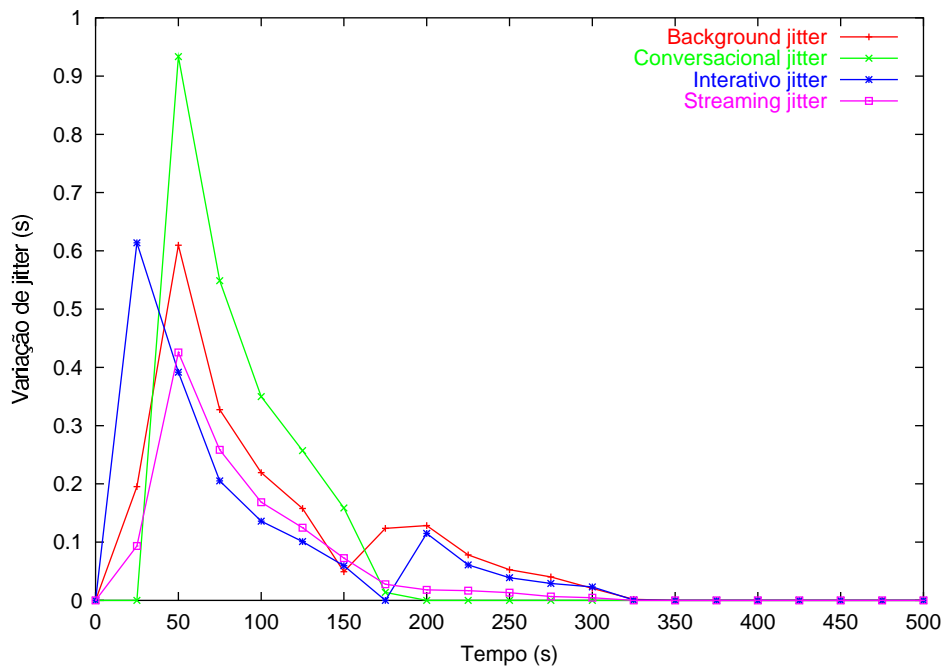


Figura 5.9: Variação do Atraso Médio com CAC

estas aplicações.

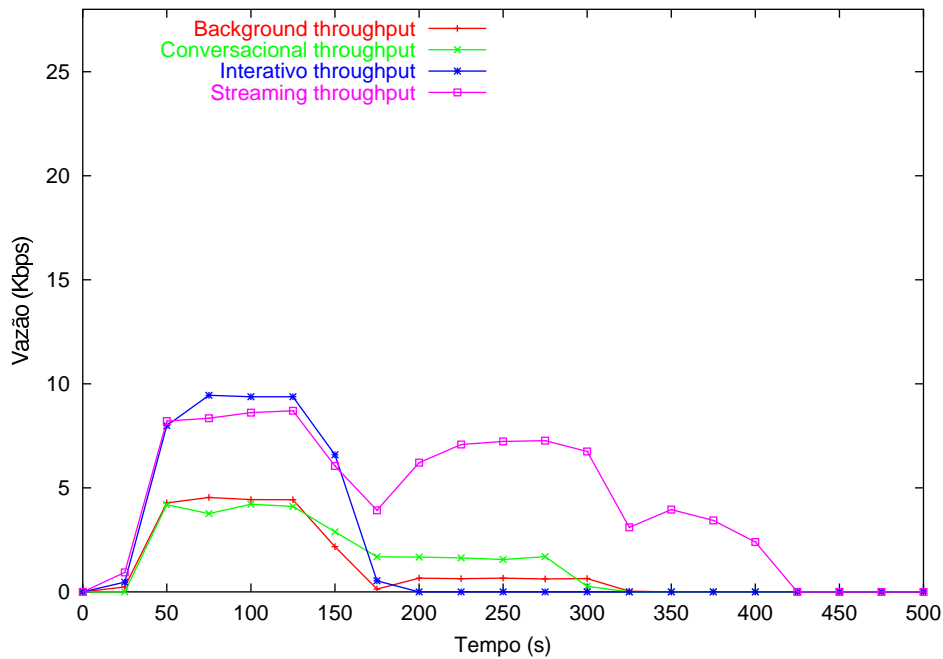


Figura 5.10: Vazão sem CAC

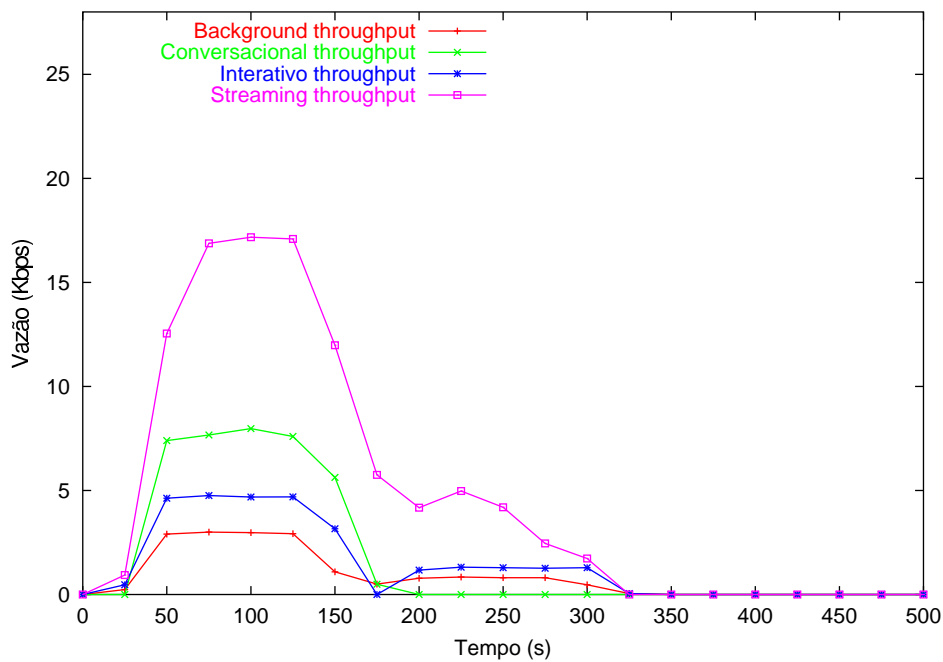


Figura 5.11: Vazão com CAC

#### 5.4.4 Utilização da Rede, para diferentes tipos de tráfego

Os gráficos das figuras 5.12 e 5.13 mostram a utilização dos canais sem e com CAC.

Por eles, percebe-se um domínio de tráfego *interactive* e *background*, no início da simu-

lação, quando não há reserva de canais para as classes prioritárias. Com CAC, ocorre uma elevação da utilização da rede pelas classes prioritárias *conversational* e *streaming* (esta última devido à ocorrência de *handoffs* desta classe). Estas classes passam suas utilizações de pico da rede de 5% para 10% e de 12% para 20%, respectivamente.

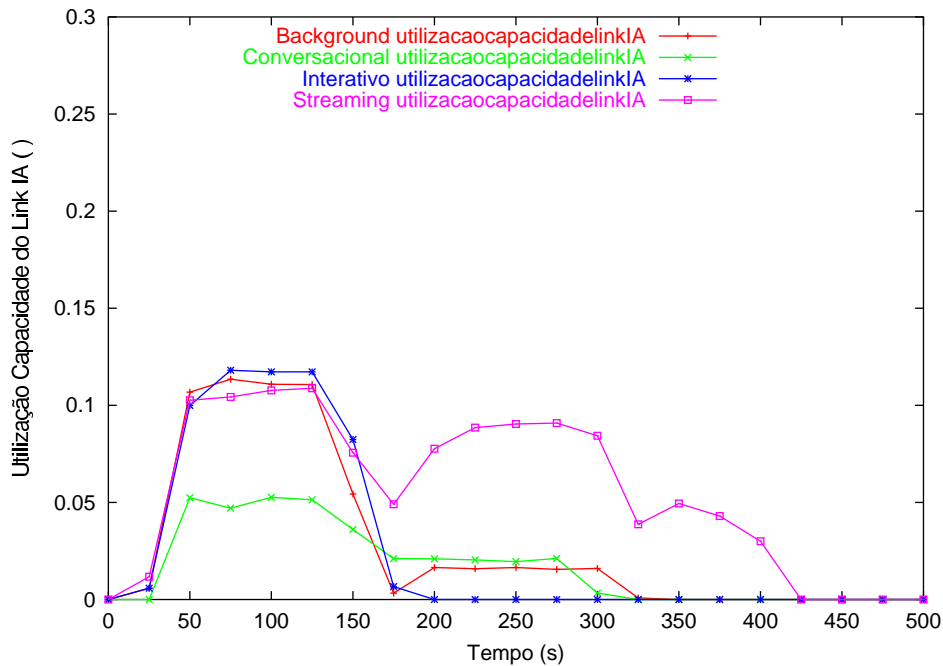


Figura 5.12: Utilização de Canal sem CAC

## 5.5 Módulo de Diagnóstico

Conforme mostrado anteriormente, o algoritmo CAC não avalia o desempenho da rede ao alocar recursos de rádio para novas chamadas ou *handoffs*. Em consequência disto, a rede pode sofrer degradação de desempenho quando o objetivo do controle de admissão é apenas minimizar o número de bloqueios, como o proposto por Oliver [76]. A arquitetura DiffMobil trata este problema de congestionamento, associando um módulo de diagnóstico ao CAC, que bloqueia chamadas de classes inferiores mesmo que o número de canais livres seja superior ao de guarda, quando a utilização da rede ultrapassa um



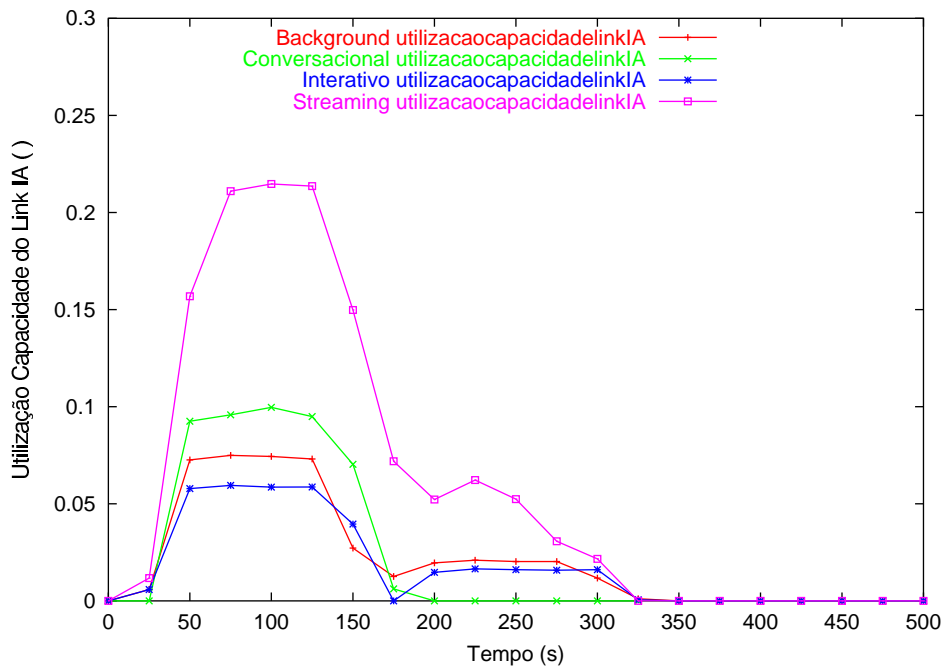


Figura 5.13: Utilização de Canal com CAC

certo *threshold*. Com isto, assegura-se a manutenção do desempenho negociado entre a aplicação e a rede, no momento da ativação do contexto PDP. Os experimentos realizados após a introdução do CAC foram repetidos para verificação da eficiência do algoritmo de diagnóstico de rede. Para estes experimentos, o mesmo cenário foi simulado, com tráfego crescente de várias unidades móveis, de maneira a se esgotar os canais disponíveis e haver competição entre elas. Os resultados mostram o desempenho do diagnóstico, nas próximas subseções.

### 5.5.1 Atraso Médio no tempo, para diferentes tipos de tráfego

Foram realizados experimentos para medir o atraso médio dos pacotes com controle baseado em guarda e diagnóstico, conforme ilustra o gráfico da figura 5.14. Conforme explicado no capítulo anterior, o diagnóstico tem *thresholds* de utilização que detectam pontos de congestionamento. A partir destes pontos, ele bloqueia chamadas, seguindo uma ordem de prioridade. Estas chamadas são admitidas em momentos posteriores,

o que proporciona uma melhor distribuição de tráfego menos prioritário no tempo (as classes menos prioritárias só ganham acesso quando a rede está descongestionada).

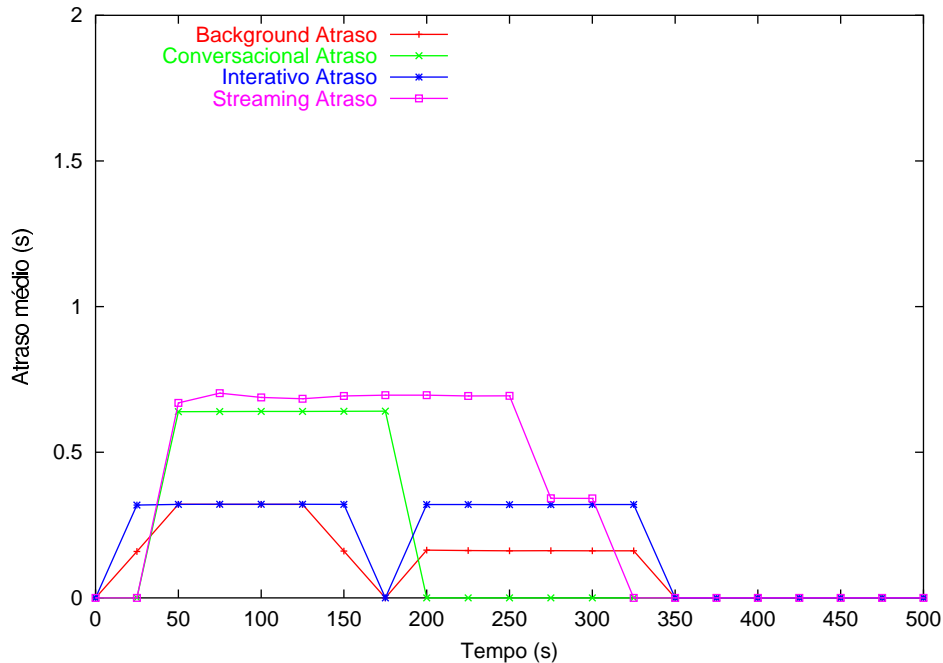


Figura 5.14: Atraso Médio com Diagnóstico

### 5.5.2 Variação do Atraso Médio no Tempo

Foram realizados experimentos para medir a variação de atraso dos pacotes com admissão baseada em guarda e diagnóstico, conforme ilustra o gráfico da figura 5.15. Pequenas variações são notadas após a introdução do diagnóstico. Como as taxas de transmissão continuam as mesmas, as variações de atraso médio quase não se alteram. Há uma leve piora na variação de atraso para aplicações das classes *interactive* e *background*, menos prioritárias, o que indica uma eficácia do módulo de diagnóstico.

### 5.5.3 Vazão Média no tempo, para diferentes tipos de tráfego

Foram realizados experimentos para medir a vazão obtida pelas classes de aplicações com admissão baseada em guarda e diagnóstico. Como pode ser observado no gráfico

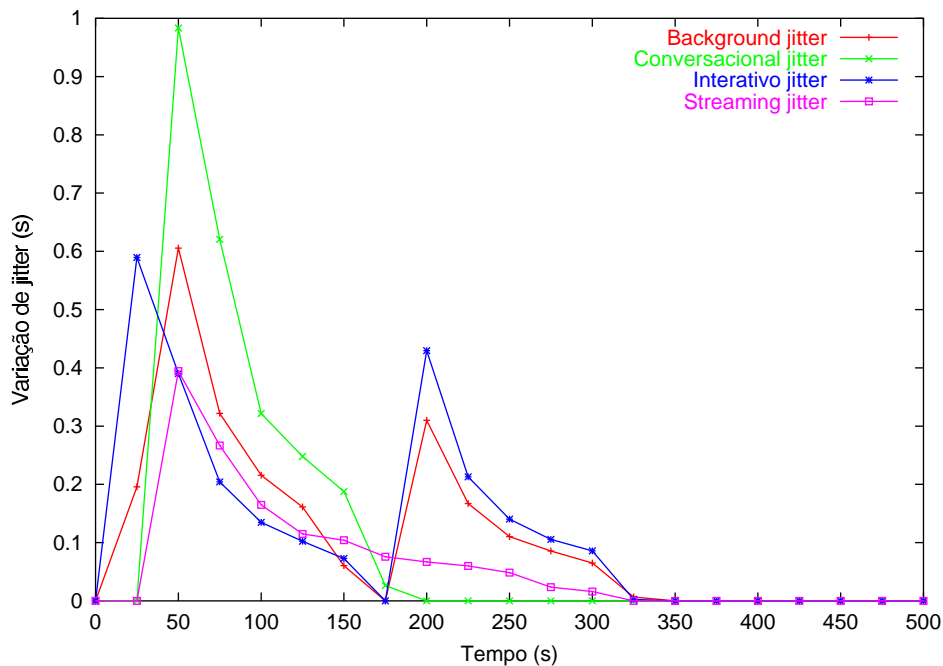


Figura 5.15: Variação do Atraso Médio com Diagnóstico

da figura 5.16, a vazão média para aplicações das classes *conversational* e *streaming* se manteve em torno de 8 kbps e 20 kbps, muito próximas às taxas nas quais foram geradas. Em compensação, houve uma queda na vazão média para aplicações das classes *interactive* e *background*, que são aplicações menos prioritárias. No caso de aplicações *interactive*, há a necessidade de uma renegociação de taxa, em casos de congestionamentos, uma vez que esta é uma aplicação sensível a vazão.

#### 5.5.4 Utilização da Rede, para diferentes tipos de tráfego

Foram realizados experimentos para medir a utilização da rede com admissão baseada em guarda e diagnóstico, conforme ilustra o gráfico da figura 5.17. Estes experimentos visam mostrar a priorização de tráfego pelo módulo de diagnóstico. Se compararmos os resultados com os apresentados no gráfico da figura 5.13, há uma maior distribuição de tráfego no tempo, aumentando o nível de utilização da rede pelas classes prioritárias, uma vez que a admissão das classes menos prioritárias é postergada em momentos de

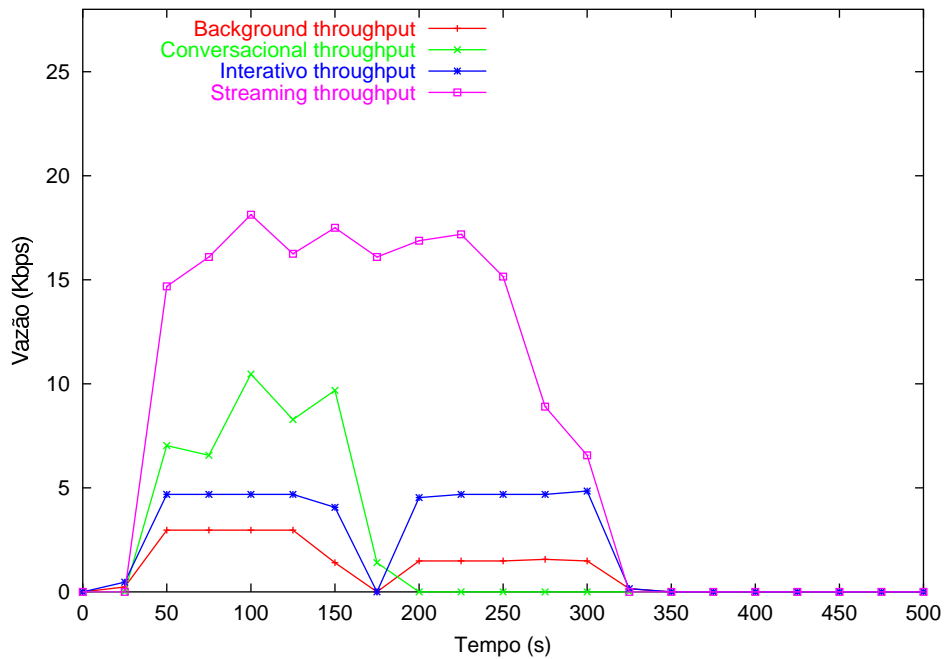


Figura 5.16: Vazão com Diagnóstico

congestionamento.

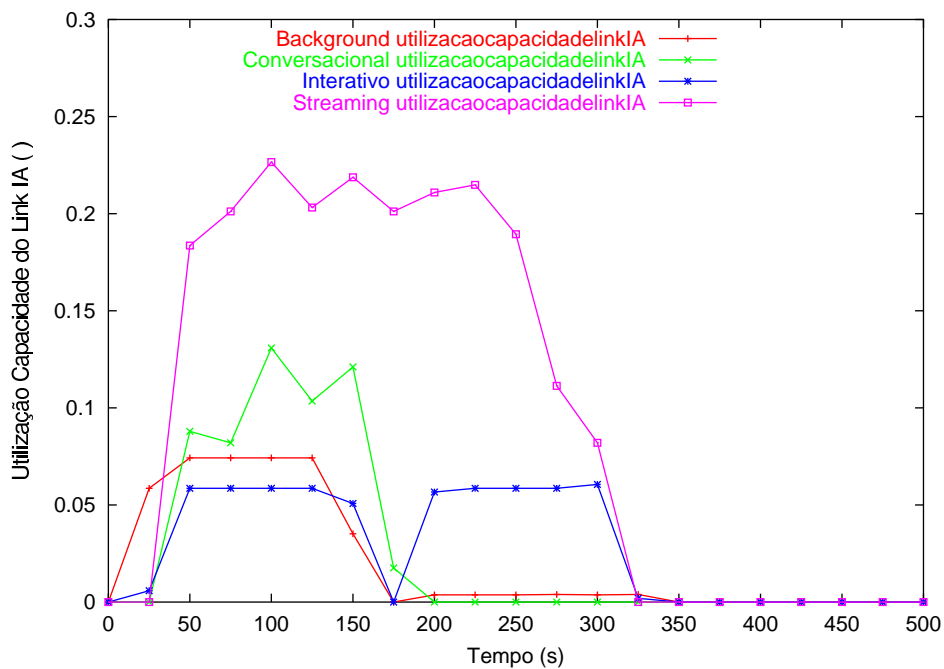


Figura 5.17: Utilização de Canal com Diagnóstico

## 5.6 Módulo de Controle de Fluxo

Foram realizados experimentos sem o controle de fluxo, que promove o escalonamento de quadros TDMA seguindo prioridades passadas pelo DiffServ do núcleo IP. Os resultados são mostrados nas subseções a seguir.

### 5.6.1 Atraso Médio no tempo, para diferentes tipos de tráfego

Foram realizados experimentos para medir o atraso médio dos pacotes com controle de fluxo, conforme ilustra o gráfico da figura 5.18, que será comparado com o gráfico da figura 5.14. O controle de fluxo reduz o atraso médio da classe *conversational* de 0,7 segundos para 0,25 segundos, representando ganhos médios de 60%, na classe mais sensível a atraso. O controle de fluxo prioriza o escalonamento de quadros *conversational* e *streaming*, em detrimento das classes *interactive* e *background*. Com isto, o atraso médio das menos prioritárias aumenta, à medida que o das mais prioritárias diminui, conforme o esperado.

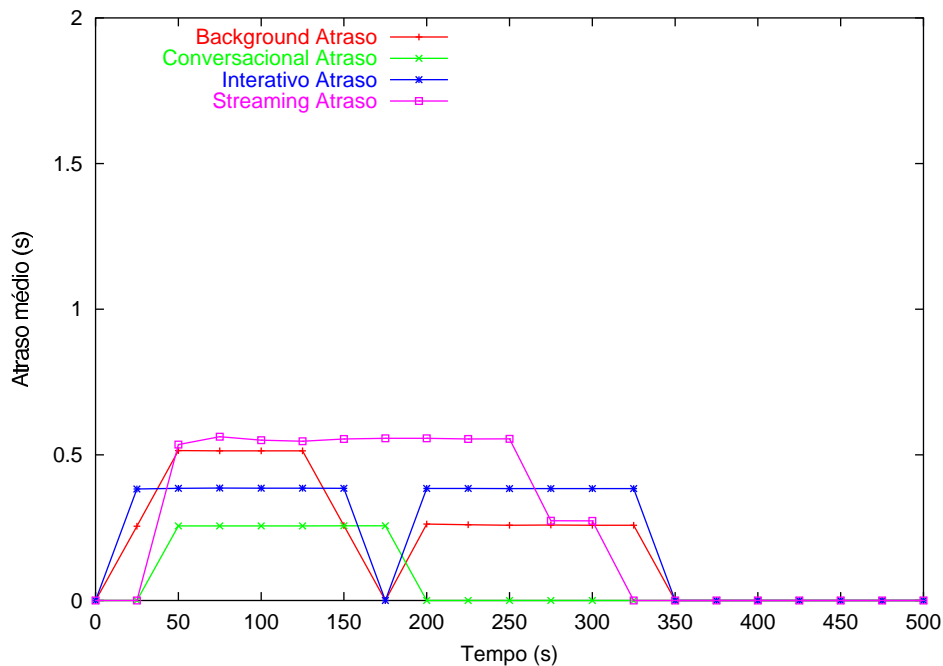


Figura 5.18: Atraso Médio com Controle de Fluxo

### 5.6.2 Variação do Atraso Médio no Tempo

Foram realizados experimentos para medir a variação de atraso dos pacotes com controle de fluxo, conforme ilustra o gráfico da figura 5.19. Percebe-se, com o controle de fluxo, uma redução significativa na variação de atraso das classes mais prioritárias. Este resultado é mais importante para a classe *streaming*, sensível a variação de atraso. Com o controle de fluxo, houve menor variação de atraso para *conversational* (de 1 para 0,4 segundo a maior variação) e *streaming* (de 0,4 para 0,3 a maior variação). Ressalta-se que, com o controle de fluxo, obteve-se uma redução de 20% da variação de atraso para a classe *streaming*, que é a mais sensível a atraso.

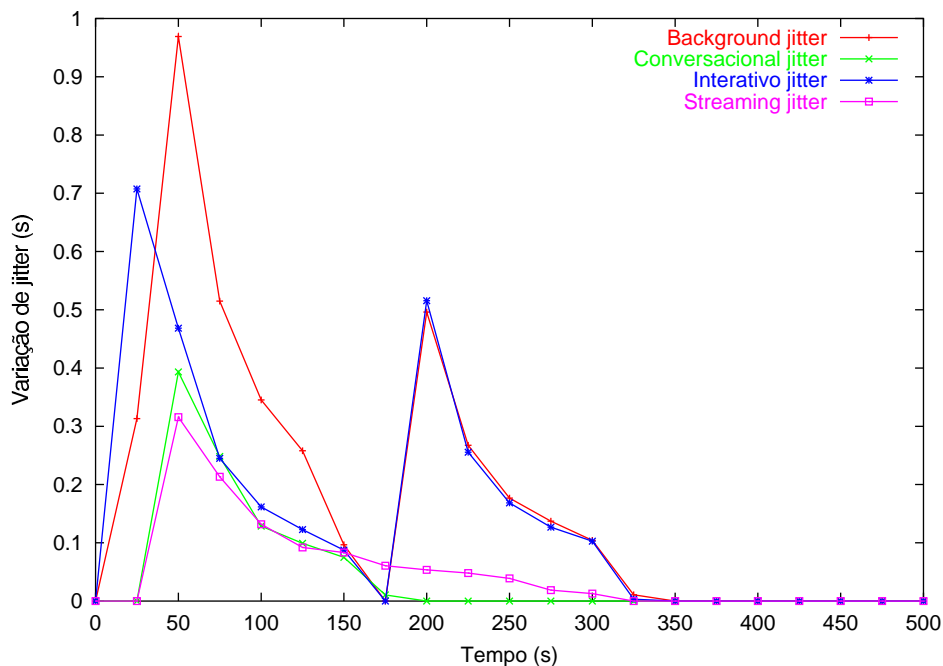


Figura 5.19: Variação do Atraso Médio com Controle de Fluxo

### 5.6.3 Vazão Média no tempo, para diferentes tipos de tráfego

Foram realizados experimentos para medir a vazão obtida pelas classes de aplicações com controle de fluxo, conforme ilustra o gráfico da figura 5.20. Com o controle de fluxo, a vazão média das classes *conversational* e *streaming* se mantém na vazão de

pico, 8 kbps e 20 kbps, respectivamente. Este resultado implica em redução da vazão média para as classes *interactive* e *background*. Para a classe *interactive*, este é um resultado ruim, pois esta é uma classe sensível a atraso. Para resolver este problema, uma renegociação de vazão pode ser feita, ou alteração dos pesos de escalonamento, em caso de necessidade de manutenção da vazão. Como os pesos utilizados foram especificados empiricamente, há a necessidade de um ajuste dos mesmos, de acordo com as necessidades de cada operadora, em determinados momentos. Para isto, um *middleware* deverá ser utilizado.

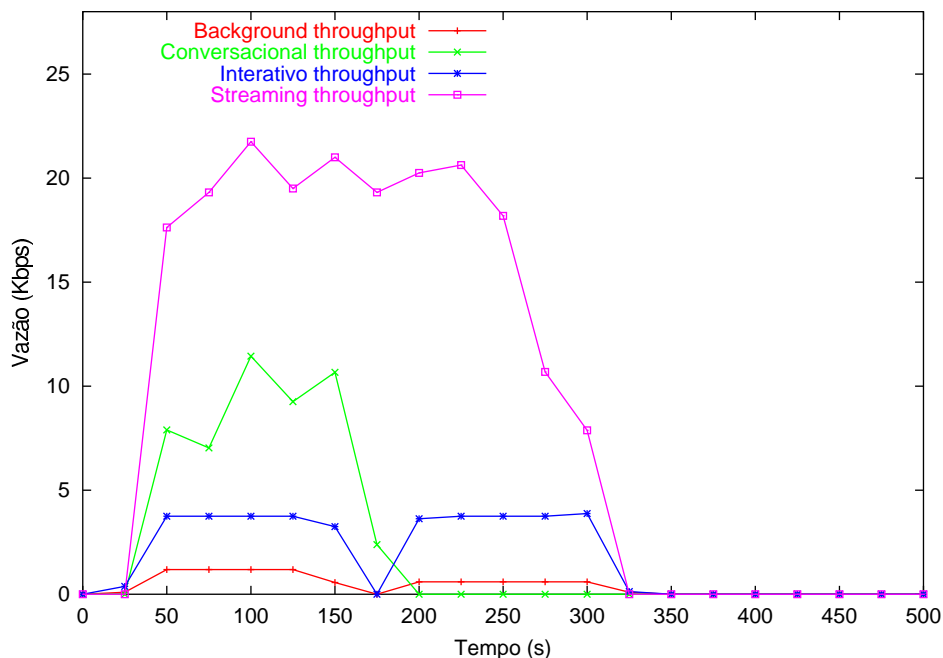


Figura 5.20: Vazão com Controle de Fluxo

#### 5.6.4 Utilização do Canal, para diferentes tipos de tráfego

Foram realizados experimentos para medir a utilização do canal com controle de fluxo, conforme ilustra o gráfico da figura 5.21. Conforme esperado, as classes *conversational* e *streaming* passam a utilizar muito mais os recursos da rede após a aplicação do algoritmo de controle de fluxo. Mais uma vez, aqui se aplica o comentário sobre os

pesos, que devem ser ajustados de acordo com as necessidades.

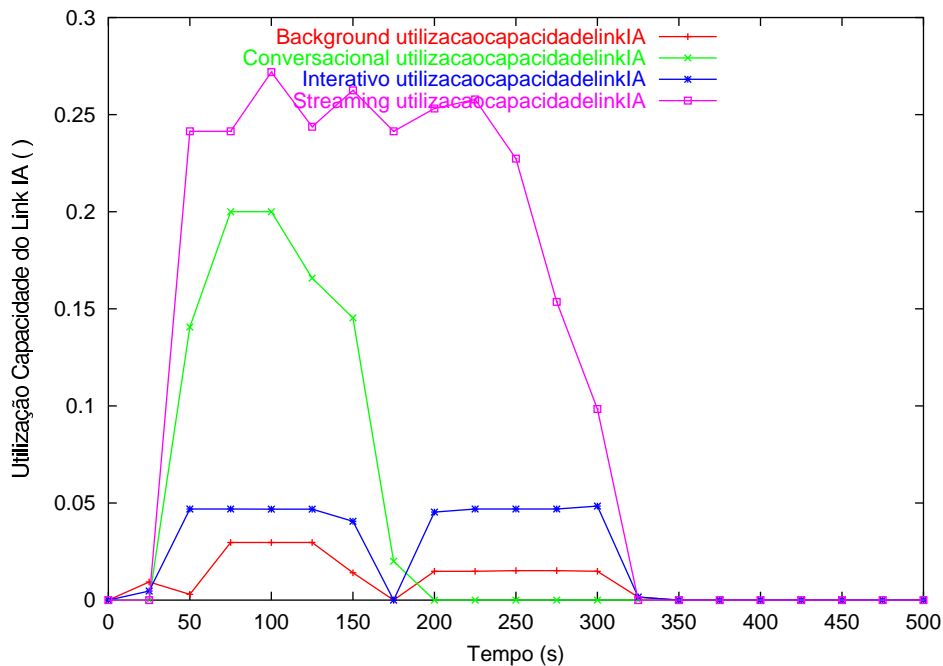


Figura 5.21: Utilização de Canal com Controle de Fluxo

## 5.7 QoS na Interface Aérea e no Núcleo IP

Foram realizados experimentos para a arquitetura DiffMobil completa, com CAC (controle de admissão baseado em guarda), diagnóstico e controle de fluxo na interface aérea e DiffServ no núcleo IP. Resultados mostram que o DiffServ no núcleo IP não interfere no desempenho das aplicações, pois o nível de utilização da rede fixa é baixo, quando comparado ao nível de utilização do enlace de rádio. O DiffServ no núcleo IP é interessante para priorização de tráfego em uma rede que não é dedicada ao GPRS e pode, eventualmente ter congestionamento. Considerando-se que a escassez de recurso se encontra na interface aérea, o desempenho fim-a-fim não se altera com a adoção do Diffserv no núcleo IP.



## 5.8 Conclusões

A arquitetura DiffMobil apresenta soluções integradas de QoS para a interface aérea e para o núcleo IP. Na interface aérea, o controle de admissão baseado em guarda proporciona redução de taxas de bloqueio de classes prioritárias *handoff* e *conversational* mas degrada o desempenho das aplicações de uma maneira geral. O módulo de diagnóstico resolve este problema, bloqueando as ativações de contexto PDP das classes menos prioritárias quando detecta *thresholds* de congestionamento. Apesar de melhorar o desempenho, o diagnóstico ainda não é suficiente para garantir desempenho requisitado pelas aplicações. Para isto, é introduzido o módulo de controle de fluxo, que faz escalonamento de pacotes baseado na marcação DiffServ que recebe do SGSN. Portanto, para se garantir QoS fim-a-fim, é necessário que o núcleo IP implemente diferenciação de serviços, conforme proposto na arquitetura.

# Capítulo 6

## Conclusões

Redes GPRS vêm sendo amplamente utilizadas em todo o mundo. Apesar disso, questões relacionadas à garantia de qualidade de serviço fim-a-fim ainda estão em aberto.

Soluções de QoS visam garantir a satisfação de usuários de uma rede. Esta satisfação pode ser medida pelo acesso garantido ou pelo desempenho das aplicações. Algumas soluções são propostas para implantar QoS em redes GPRS, mas nenhuma delas encontra um equilíbrio entre garantia de acesso e garantia de desempenho.

A principal contribuição deste trabalho foi o desenvolvimento de uma arquitetura de QoS fim-a-fim em um sistema GPRS, denominada DiffMobil. Esta arquitetura oferece soluções que garantem acesso a aplicações prioritárias, mantendo o nível de serviço acordado entre a operadora e o assinante. A arquitetura DiffMobil é composta por módulos de controle de admissão, diagnóstico e controle de fluxo, que também representam contribuições individuais importantes do trabalho. Além dos novos módulos, a arquitetura adota o DiffServ como solução de QoS no núcleo IP de uma rede GPRS, priorizando o roteamento de pacotes de alguns tipos aplicações.

O controle de admissão proposto trabalha com guarda de canais para *handoffs* e tráfego conversacional de novas chamadas. O módulo de diagnóstico da rede implemen-

ta um controle em tempo real do uso de recursos da rede, indicando o percentual de largura de banda disponível em cada segmento da rede. O controle de admissão baseado em guarda e diagnóstico busca um equilíbrio entre reserva de recursos para classes prioritárias e níveis de desempenho e qualidade de serviço acordados com assinantes de uma rede GPRS.

O controle de fluxo proposto promove a diferenciação de quadros TDMA, de maneira a estender a diferenciação de serviços do núcleo IP na interface aérea. Nele, o protocolo BSSGP transfere pacotes marcados do SGSN para a estação base. Na estação base, a camada *relay* classifica quadros em filas individuais e promove escalonamento segundo a prioridade de classe: *conversational*, *streaming*, *interactive*, *background*. Desta maneira, a camada RLC/MAC transmite no enlace de rádio os pacotes seguindo esta ordem prioridade.

Foram desenvolvidos modelos de simulação de uma rede GPRS completa e da arquitetura DiffMobil através da ferramenta ns-2 (*Network Simulator*). Resultados comparam o desempenho de aplicações de classes diferentes, com e sem DiffMobil, em uma rede GPRS. Telnet, Real Audio, navegação na WEB e transferência de e-mails foram as aplicações simuladas para as classes *conversational*, *streaming*, *interactive* e *background*, respectivamente. Estas classes foram adotadas como padrão da arquitetura DiffMobil, seguindo um mapeamento feito para as classes do GPRS. Em uma primeira etapa, foi incluído o DiffServ no núcleo IP da rede GPRS. Resultados mostraram, para este experimento, a possibilidade de se alocar recursos da rede fixa de maneira diferenciada com garantias de vazão média e de pico para as classes prioritárias. Na segunda etapa, foi desenvolvido um mecanismo de garantia de QoS na interface aérea. Este mecanismo na arquitetura DiffMobil é composto pelo controle de admissão, pelo diagnóstico e pelo controle de fluxo. O controle de admissão foi implementado com reserva de canais para *handoffs* e para novas chamadas da classe *conversational*. Dois trabalhos da literatu-

---

ra [67] e [76] foram utilizados como referência. O controle baseado em guarda, reduziu bloqueios de *handoffs* em 60%, quando comparado ao melhor esforço. Também reduziu bloqueios de novas chamadas da classe *conversational* para 1%. Apesar de não haver distinção de classe anteriormente, o melhor esforço bloqueava 15% de novas chamadas. Se considerarmos a possibilidade de 15% de chamadas *conversational* serem bloqueadas, os ganhos obtidos são de 93% para esta classe. Para se garantir requisitos de desempenho, foi implementado o módulo de diagnóstico, que reduziu atraso médio e variação de atraso das aplicações, o que o módulo de guarda não garante. Sobre os resultados obtidos com guarda e diagnóstico, foram realizados experimentos com o controle de fluxo. O controle de fluxo mantém atraso médio e variação de atraso nos patamares desejados pelos assinantes, o que o controle de admissão não conseguia. A utilização da rede e a vazão das classes prioritárias aumentam em até 60% com toda a arquitetura DiffMobil implementada. O DiffServ, além de controlar o tráfego por prioridades no núcleo IP, repassa marcações para o protocolo BSSGP. Este repasse possibilita a marcação de quadros TDMA por prioridades, o que é fundamental ao cumprimento das garantias de QoS negociadas. A Tabela 6.1 mostra valores obtidos pela arquitetura DiffMobil, por classe de aplicação. Comparando estes valores com os da tabela 4.1, conclui-se que a arquitetura DiffMobil garante QoS fim-a-fim, necessitando de uma negociação de vazão para a classe *Interactive*, em casos de congestionamento.

Conforme foi apresentado, a arquitetura DiffMobil supre deficiências das soluções propostas por Oliver [76] e Stuckmann [67], que propõem escalonamento com reserva e controle de admissão baseado em guarda, respectivamente. As duas soluções visam garantir acesso às aplicações prioritárias, mas não apresentam soluções de compromisso com níveis de serviço acordados. A tabela 6.2 mostra uma comparação entre a arquitetura DiffMobil e as soluções propostas por Oliver [76] e Stuckmann [67].

Pela tabela 6.2, conclui-se que a arquitetura DiffMobil é, entre as soluções encon-

Definição de Perfis de QoS					
Classes GPRS	Atributos de QoS	Perfis de QoS Implementados			
		Conv. EF	Stream. AF4	Interactive AF3 e AF2	Background AF1 e BE
Precedência	Prioridade em congestionamento	Alta	Média	Baixa	N/A
Atraso	Latência	Limitado (<0,5s)	Limitado (<5s)	Limitado (<50s)	N/A
	Jitter	Baixo	Baixo	N/A	N/A
Vazão média e máxima	Vazão Nível Rajadas	N/A Alto	Garantida Baixo	Negociado Alto	N/A N/A

Tabela 6.1: Mapeamento de Classes de QoS na Arquitetura DiffMobil

Garantias de QoS	Soluções Propostas		
	<i>DiffMobil</i>	<i>Oliver</i>	<i>Stuckmann</i>
QoS no núcleo IP	DiffServ	MPLS	DiffServ
QoS na Interface Aérea	CAC, Diagnóstico e Controle de Fluxo	CAC Adaptativo	Escalonamento com Reserva
Bloqueios de <i>handoffs</i>	4%	entre 1% e 7%	N/A
Bloqueios de Novas Chamadas	11%	entre 4% e 17%	N/A
Melhora desempenho por prioridade	sim	não	sim
Solução integrada	sim	não	não

Tabela 6.2: Comparação entre a Arquitetura DiffMobil e as soluções de Oliver Stuckman

tradas, a única que controla o nível de bloqueios de classes prioritárias e o desempenho negociado com o assinante, fim-a-fim. Controlando os níveis de serviço, a arquitetura DiffMobil mantém um compromisso entre garantir acesso aos recursos da rede e assegurar bom desempenho às classes prioritárias, com diferenciação de serviço fim-a-fim.

Possíveis extensões ao trabalho incluem experimentos com pesos diferenciados no controle de fluxo, a adaptação do módulo de diagnóstico de maneira que as decisões sejam tomadas de modo dinâmico, levando em consideração um histórico e um modelo de tráfego. Também está prevista a confecção de um *middleware* para passagem de

parâmetros de QoS do usuário para a rede e a elaboração de uma arquitetura similar à DiffMobil, que faça diferenciação de serviços fim-fim em redes móveis de terceira geração UMTS. Estudos sobre redes locais sem fio, interconectadas por *backbones* IP e assuntos relacionados à qualidade de serviço destes ambientes também são assunto de interesse para trabalhos futuros.

# Referências Bibliográficas

- [1] 3GPP. End-to-End QoS Concept and Architecture. TS 23.207 version 5.3.0.
- [2] 3GPP. End-to-End QoS Signalling Flows. TS 29.208 version 1.0.0.
- [3] 3GPP. UMTS QoS Concepts. TS 23.107 version 3.2.0.
- [4] 3GPP. GSM 03.64, 1999.
- [5] 3GPP. GSM 08.18. version 8.1.0, 1999.
- [6] 3GPP. RLC/MAC Protocol. version 8.10.0, 1999.
- [7] 3GPP. BSSGP Protocol. TS 48.018 version 5.10.0, 2004.
- [8] Sergio Andreozzi. Sergio andreozzi ns diffserv patch, <http://www.cnaf.infn.it/sergio/diffserv/ns/download.htm>. 2001.
- [9] Mirza Shahrukh Baig, Vishwanath Sinha, and Peter Stuckmann. Differentiated Services for Mobile Core Networks. In *Symposium on Wireless Personal Multimedia Communications*, pages 675–680, Aalborg, Denmark, September 2001.
- [10] Baudet, S. and Besset-Bathias, C. and Frêne, P. and Giroux, N. Qos implementation in umts networks. In *Alcatel Telecommunications Review*, 1, pages 41–49. 2001.
- [11] Dr. R. Becher, Dr. G. Klas, U. Weiss, and M. Koehler. IMS (IP based Multimedia Subsystem)-milestone to 4G All IP networks. In *IST Mobile Summit*, Barcelona, Spain, September 2001.
- [12] S. Blake, D. Black, W. Weiss, Davies E., and Wang Z. An architecture for differentiated services, <http://www.ietf.org/rfc/rfc2475.txt>. 1998.
- [13] S. Blake, W. Nichols, F. Baker, and Black D. Definition of the differentiated services field (ds field) in the ipv4 and ipv6 headers, <http://www.ietf.org/rfc/rfc2474.txt>. 1998.
- [14] Torsten Braun, Claude Castelluccia, Gunther Stattenberger, and Imad Aad. An Analysis of the Diffserv Approach in Mobile Environments. <http://www.inrialpes.fr/planete/people/MobiQoS/paper2.ps>, unpublished, April 1999.

- [15] Louise Burness, Eleanor Hepworth, Alberto López, and Jukka Manner. Architecture for Providing QoS in an IP-based Mobile Network. In *IST Mobile Summit*, Barcelona, Spain, September 2001.
- [16] Hemant Chaskar and Rajeev Koodli. MPLS and Diffeserv for UMTS QoS in GPRS Core Network Architecture. In *Internet Society's INET Conference*, Stockholm, Sweden, June 2001.
- [17] Byung-Hwan Choi and Hong-Shik Park. Rate proportional scfq algorithm for high-speed packet switched networks. *ETRI*, 22(3):1–9, 2000.
- [18] Sunghyun Choi. *Doctoral Thesis: QoS Guarantees in Wireless/Mobile Networks*. University of Michigan, 1999.
- [19] Cristina Comaniciu, Narayan B. Mandayam, David Famolari, and Prathima Agrawal. QoS Guarantees for Third Generation (3G) CDMA Systems via Admission and Flow Control. In *IEEE Vehicular Technology Conference (VTC)*, Boston, September 2000.
- [20] Kelvin Lopes Dias and Djamel Fauzi Sadok. Internet Móvel: Tecnologias, Aplicações e QoS. In *Minicurso - Simpósio Brasileiro de Redes de Computadores*, Florianópolis, S.C., Brazil, May 2001.
- [21] Sudhir Dixit, Yile Guo, and Zoe Antoniou. Resource management and Quality of Service in Third-Generation Wireless Networks. In *IEEE Communications Magazine*, pages 125–133, February 2001.
- [22] Fátima de L. P. Duarte, Antonio A. L. Loureiro, Leonardo B. Oliveira, and Cláudio M. de S. Vicente. GPRS Systems Performance Analysis. In *IEEE International Telecommunications Symposium*, Natal, RN, Brazil, September 2002.
- [23] Fátima de L. P. Duarte, Kalina Ramos Porto, Antonio A. L. Loureiro, Leonardo B. Oliveira, and Cláudio M. de S. Vicente. GPRS: Simulação do Core IP e Avaliação de Desempenho Fim-a-Fim. In *Workshop de Comunicação sem Fio*, São Paulo, SP, Brazil, October 2002.
- [24] Fátima de L. P. Duarte-Figueiredo, Vinicius Rosalen da Silva, Thiago H. de Paula Figueiredo, and Antonio A. L. Loureiro. Desempenho do GPRS: Definindo *Thresholds* para Políticas de QoS. In *Workshop de Comunicação sem Fio*, São Lourenço, MG, Brazil, October 2003.
- [25] Brahim Ghribi and Luigi Logrippo. Understanding GPRS: the GSM packet radio service. In *Computer Networks*, volume 34, pages 763–779. Elsevier Science B.V., 2000.
- [26] Network Working Group, S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. RFC2475 - An Architecture for Differentiated Services. <http://www.rfc-editor.org/rfc/rfc2475.txt>, December 1998.



- [27] Network Working Group, R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin. RFC 2205 - Resource reservation Protocol (RSVP) - Version 1 Functional Specification. <http://www.rfc-editor.org/rfc/rfc2205.txt>, September 1997.
- [28] Network Working Group and S. Bradner. RFC1242 - Benchmarking Terminology for Network Interconnection Devices. <http://www.rfc-editor.org/rfc/rfc1242.txt>, July 1991.
- [29] William C. Hardy and QSDG Quality of Service Development Group. "QoS" *Measurement and Evaluation of Telecommunications Quality of Service*. John Wiley and Sons LTD, 2001.
- [30] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski. Assured forwarding phb group, <http://www.ietf.org/rfc/rfc2597.txt>. 1999.
- [31] Daehyoung Hong and S. S. Rappaport. Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Nonprioritized Handoff Procedures. In *IEEE Transactions on Vehicular Technology*, volume 35, pages 77–92, August 1986.
- [32] J. F. Huard and A. A. Lazar. On QoS Mapping in Multimedia Networks. In *21th IEEE Annual International Computer Software and Application Conference - COMPSAC'97*, pages 13–15, Washington, D.C, August 1997.
- [33] IETF. MPLS - Multiprotocol Label Switching Architecture. IETF draft-ietf-mpls-arch06.txt, August 1999.
- [34] Telia Finland J. Heinanen and R. Guerin. A single rate three color marker, <http://www.ietf.org/rfc/rfc2697.txt>. 1999.
- [35] V. Jacobson, K. Nichols, and K. Poduri. An expedited forwarding phb, <http://www.ietf.org/rfc/rfc2598.txt>. 1999.
- [36] Ms. Richa Jain. GPRS Simulations using ns-Network Simulator. Dual Degree Dissertation of the India Institute of Technology, June 2001.
- [37] Sevanto Jarkko, Mika Liljeberg, and Kimmo Raatikainen. Introducing Quality-of-Service and Traffic Classes into Wireless Mobile Networks. In *1st ACM International Workshop on Wireless Mobile Multimedia*, pages 21–29, Dallas, Texas, USA, 1998.
- [38] Carlos A. Kamienski and Djamel Sadok. Qualidade de Serviço na Internet. In *Minicurso 2 - Décimo Oitavo Simpósio Brasileiro de Redes de Computadores - SBRC2001*, Belo Horizonte, Minas Gerais, Brazil, May 2000.
- [39] Carlos A. Kamienski and Djamel Sadok. Chameleon: an Architecture for Advanced End-to-End Services. In *Second Latin America Network Operation and Management Symposium*, pages 245–256, Belo Horizonte, Brazil, August 2001.

- [40] Georgios Karagiannis. QoS in GPRS - Ericsson Open Report. <http://ing.ctit.utwente.nl/WU4>, 2000.
- [41] Manolis Katevenis, Evangelos Markatos, and Mavroidis Ioannis. Weighted round-robin scheduler using per-class urgency counters.
- [42] Manolis Katevenis, Dimitrios Serpanos, and Evangelos Markatos. Multi-queue management and scheduling for improved qos in communication networks. In *EMMSEC'97(European Multimedia Microprocessor Systems and Electronic Commerce Conference)*, pages 906–913, Florence, Italy, November 1997.
- [43] Rajeev Koodli and Mikko Puskari. Supporting Packet-Data QoS in Next-Generation Cellular Networks. In *IEEE Communications Magazine*, pages 180–188, February 2001.
- [44] Jussi Laukkanen. Umts quality of service concept and architecture. <http://www.cs.helsinki.fi/u/kraatika/Courses/awcs00s/awcs00s-titles.html>.
- [45] Lee, Lin-Nan and Sun, Feng-Wen and Karimullah, Khalid and Eroz, Mustafa and Hammons, Roger. Third generation wireless communication technologies. In Rajamani Ganesh, Kaveh Pahlavan, and Zoran Zvonar, editors, *Wireless Multimedia Network Technologies*, chapter 2, pages 24–44. Kluwer Academic Publishers, 2000.
- [46] Lundy Lewis. *Service Level Management for Enterprise Networks*. Artech House Publishers, 2000.
- [47] Frank Yong Li and Norvald Stol. A Priority-oriented Call Admission Control Paradigm with QoS Re-negotiation for Multimedia Services in UMTS. In *IEEE Vehicular Technology Conference (VTC'01 spring)*, pages 2021–2025, Rhodes, Greece, May 2001.
- [48] Christoph Lindemann, Marco Lohmann, and Axel Thummler. Adaptive Performance Management for UMTS Networks. In *Computer Networks Journal*, to appear.
- [49] Thorsten Lohmar and Ralf Keller. QoS Interworking between WLAN and UMTS Core Network. In *First Workshop on IP Quality of Service for Wireless and Mobile Networks*, Aachen, Germany, April 1999.
- [50] Indu Mahadevan and Krishna M. Siviligam. An Architecture for QoS Guarantees and Routing in Wireless/Mobile Networks. In *First ACM/IEEE International Workshop on Wireless and Mobile Multimedia (WoWMoM'98)*, pages 11–20, Dallas, October 1998.
- [51] Marco Ajmone Marsan, Gabriele Ginella, Roberta Maglione, and Michela Meo. Performance Analysis of Hierarchical Cellular Networks With Generally Distributed Call Holding Times and Dwell Times. In *IEEE Transactions on Wireless Communications*, volume 3, January 2004.

- [52] Geraldo Robson Mateus and Antonio Alfredo F. Loureiro. *Introdução à Computação Móvel*. Décima Primeira Escola de Computação, 1998.
- [53] Ojanperä, Tero. Perspectives on 3g system development. In Rajamani Ganesh, Kaveh Pahlavan, and Zoran Zvonar, editors, *Wireless Multimedia Network Technologies*, chapter 3, pages 45–55. Kluwer Academic Publishers, 2000.
- [54] Julio Orozco and David Ros. An adaptative rio (a-rio) queue management algorithm. *International Workshop on Quality of Future Internet Services (QoFIS)*, 2003.
- [55] Joel Orvis. GPRS QoS. GN Nettet Network Operator Division on <http://www.watmag.com>.
- [56] Qixiang Pang, Amir Bigbloo, Victor C. M. Leung, and Chris Scholefield. Service Scheduling for General Packet Radio Service. In *IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, USA., 1999.
- [57] Ioannis C. Panoutsopoulos and Stavros Kotsopoulos. Handover and New Call Admission Policy Optimization for G3G Systems. In *Wireless Networks*, volume 8, pages 381–389. Kluwer Academic Publishers, 2002.
- [58] John M. Peha and Arak Sutivong. Admission Control Algorithms for Cellular Systems. In *Wireless Networks*, volume 7, pages 117–125. Kluwer Academic Publishers, 2001.
- [59] Pereira, Jorge. An unified, open architecture to deliver mobile multimedia. In Rajamani Ganesh, Kaveh Pahlavan, and Zoran Zvonar, editors, *Wireless Multimedia Network Technologies*, chapter 4, pages 57–74. Kluwer Academic Publishers, 2000.
- [60] The VINT Project. The network simulator - ns-2, <http://www.isi.edu/nsnam/ns>. 2001.
- [61] The VINT Project. The ns manual (formerly ns notes and documentation), <http://www.isi.edu/nsnam/ns/doc/>. 2003.
- [62] The VINT Projetc. The ns Manual. <http://www.isi.edu/nsnam/ns/ns-documentation.html>, September 2001.
- [63] Vinicius Rosalen. Um Estudo Quantitativo da Integração entre o IPv6 Móvel e as Redes Celulares. Dissertação de Mestrado da Universidade Federal de Minas Gerais, Outubro 2003.
- [64] Peter Rysavy. Network Design Manual: Wide-Area Wireless Computing. <http://www.networkcomputing.com/netdesign/wireless1.html>, September 1997.
- [65] Samukic, Antun. Umts/imt-2000 standardisation. In Rajamani Ganesh, Kaveh Pahlavan, and Zoran Zvonar, editors, *Wireless Multimedia Network Technologies*, chapter 5, pages 75–93. Kluwer Academic Publishers, 2000.

- [66] Vladimir de Lima Santos and Antonio Alfredo F. Loureiro. Qualidade de Serviço em Computação Móvel. In *Minicurso 2 - I Workshop de Comunicação Sem Fio*, Belo Horizonte, Minas Gerais, Brazil, July 1999.
- [67] Peter Stuckmann. Quality of Service Management in GPRS-Based Radio Access Networks. In *Telecommunication Systems*, volume 3–4, chapter 19, pages 515–546. Kluwer Academic Publishers, 2002.
- [68] Peter Stuckmann and Frank Muller. Quality of Service Management in GPRS Networks. In *IEEE International Conference on Networking (ICN'01)*, Colmar, France, July 2001.
- [69] Cisco Systems. DiffServ - The Scalable End-to-End QoS Model. White Paper on <http://www.cisco.com>.
- [70] Megistro Systems. QoS Protection. <http://www.megistro.com/solutions/wirelessReview.html>.
- [71] Agilent Technologies. Understanding General Packet Radio Service (GPRS). White Paper on <http://www.techonline.com/community/>, February 2003.
- [72] Usha Communications Technology. GPRS - General Packet Radio Service. White Paper on <http://www.iec.com>.
- [73] Shengquan Wang, Dong Xuan, Riccardo Bettati, and Wei Zhao. A Study Providing Statistical QoS in a Differentiated Services Network. In *Second IEEE International Symposium on Network Computing and Applications - NCA'03*, Cambridge, USA, April 2003.
- [74] A. Yew, C. Bohoris, A. Liotta, and G. Pavlov. Towards 'Quality of Service aware' services for the Virtual Home Environment. In *Globecom2001*, April 2001.
- [75] C.H. Yoon and C.K. Un. Performance of Personal Portable Radio Telephone Systems With and Without Guard Channels. In *IEEE Journal on Selected Areas in Communications*, volume 11, pages 911–917, August 1993.
- [76] Oliver T. W. Yu. End-to-End Adaptive QoS Provisioning over GPRS Wireless Mobile Network. In *Mobile Networks and Applications*, volume 8, pages 255–267. Kluwer Academic Publishers, 2003.
- [77] H. Zhang. Service disciplines for guaranteed performance service in packet-switching networks. 1996.
- [78] Dongmei Zhao, Xuemin Shen, and Jon W. Mark. QoS Performance Bounds and Efficient Connection Admission Control for Heterogeneous Services in Wireless Networks. In *Wireless Networks*, volume 8, pages 85–95. Kluwer Academic Publishers, 2002.