Susana Isabel Barreto de Miranda Sargento

Gestão de Recursos em Redes com Suporte de Qualidade de Serviço

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Electrotécnica, realizada sob a orientação científica do Prof. Dr. Rui Valadas, Professor Associado do Departamento de Electrónica e Telecomunicações da Universidade de Aveiro

o júri

Presidente

Prof. Doutor António Manuel Melo de Sousa Pereira

Professor Catedrático da Universidade de Aveiro por delegação do Reitor da Universidade de Aveiro

Prof. Doutor Carlos Alberto de Carvalho Belo

Professor Associado do Instituto Superior Técnico da Universidade Técnica de Lisboa

Prof. Doutor Fernando Pedro Lopes Boavida Fernandes

Professor Associado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Prof. Doutor Rui Jorge Morais Tomaz Valadas

Professor Associado da Universidade de Aveiro (Orientador)

Prof. Doutor Alexandre Júlio Teixeira dos Santos

Professor Associado da Universidade do Minho

Prof. Doutor Amaro Fernandes de Sousa

Professor Auxiliar da Universidade de Aveiro

agradecimentos

Ao Prof. Rui Valadas pela oportunidade que me deu e pelas condições que me proporcionou para a realização deste trabalho. Agradeço todos os conhecimentos que me transmitiu e o acompanhamento neste estudo.

Ao Prof. Amaro de Sousa pela disponibilidade e apoio na primeira fase deste trabalho.

Ao Engº Teixeira de Sousa pela permanente disponibilidade, acompanhamento e transmissão de conhecimentos essenciais na elaboração de uma parte do trabalho apresentado.

Ao Departamento de Electrónica e Telecomunicações e ao Instituto de Telecomunicações, pólo de Aveiro, pelas condições que me permitiram concretizar este trabalho.

À Fundação para a Ciência e Tecnologia pela bolsa de investigação científica do programa PRAXIS que usufruí. Agradeço também o apoio financeiro extra que me concedeu no decorrer do meu estágio nos EUA.

À PTInovação pela possibilidade concedida de participar em projectos de investigação, importantes no desenvolvimento do trabalho.

A todos os meus colegas do Grupo de Sistemas de Banda Larga e do Instituto de Telecomunicações, pelo excelente ambiente de trabalho que me proporcionaram. Um agradecimento especial ao Engº Paulo Salvador pelo apoio e ajuda que me concedeu.

Ao Prof. Edward Knightly e às pessoas que fizeram e que fazem parte do grupo de redes da Universidade de Rice (Houston, EUA), pela experiência que me proporcionaram, pelo constante apoio, e pelos conhecimentos que me transmitiram, sempre com uma grande amizade.

A todos os meus grandes amigos pela paciência que sempre manifestaram quando, nas partes mais difíceis do trabalho, não lhes pude dar o apoio esperado.

Aos meus Pais, irmã e familiares, pela constante motivação, paciência, apoio e carinho.

Ao Victor pela sua ajuda constante a todos os níveis, pelo seu carinho, paciência e amor incondicionais.

resumo

Esta Tese aborda a problemática da gestão de recursos em redes ATM (*Asynchronous Transfer Mode*) e IP (*Internet Protocol*) com suporte de QdS (Qualidade de Serviço). As tecnologias ATM e IP permitem, hoje em dia, integrar diferentes tipos de serviços numa mesma infra-estrutura de rede. No entanto, a diversidade dos serviços a suportar e dos seus requisitos, coloca grandes desafios ao nível da gestão de recursos, a qual se pretende o mais eficiente possível.

A Tese está dividida em duas partes. A primeira parte da Tese foca a gestão de recursos em redes de acesso ATM e IP com suporte de QdS. É dada inicialmente uma perspectiva histórica da evolução das redes de acesso. No que diz respeito às redes de acesso ATM, são propostas estratégias de gestão de recursos com base em VPs (*Virtual Paths*) e são definidas metodologias de dimensionamento que têm em conta requisitos de QdS tanto ao nível da chamada como da célula. As estratégias de gestão de recursos consideradas permitem estabelecer compromissos entre a utilização de recursos e a carga de sinalização. No que diz respeito às redes de acesso IP, é proposta uma nova arquitectura de rede, que constitui uma evolução face às redes de acesso tradicionais ao permitir uma maior partilha de recursos. Esta arquitectura permite diferenciação de QdS e suporte de aplicações multimédia. Em particular é proposta uma solução que inclui o suporte integrado de tecnologias recentemente introduzidas para iniciar e configurar sessões multimédia, gerir as políticas de QdS incluindo as funcionalidades de AAA (*Authentication, Authorization and Accounting*), e reservar recursos.

Na segunda parte da Tese é feita uma abordagem de dois mecanismos para controle de admissão de fluxos escaláveis: mecanismos de *probing* e mecanismos baseados em agregação de reservas individuais. Ambos os mecanismos permitem que a gestão de recursos seja feita sem necessidade de manutenção do estado dos fluxos activos em cada elemento de rede. O mecanismo de *probing* estima o nível de QdS da rede, através da inserção de fluxos de teste, por forma a decidir se um novo fluxo pode ou não ser aceite. É proposto um novo mecanismo de *probing*, denominado de ?-*probing*, que permite minimizar o problema do roubo de recursos que afecta o *probing* simples quando este é utilizado em sistemas com múltiplas classes de serviço. São desenvolvidos modelos analíticos e são efectuados estudos de simulação para analisar o problema do roubo de recursos e os factores que influenciam a estimativa do rácio de perdas efectuada pelos fluxos de *probing* e de ?-*probing*. Verificou-se que o mecanismo de ?-*probing* permite obter simultaneamente uma utilização elevada dos recursos e a diferenciação dos serviços sem roubo de recursos. Os resultados obtidos com este mecanismo foram também validados através do desenvolvimento de um protótipo laboratorial. No mecanismo baseado em agregação de reservas individuais, os elementos do núcleo da rede mantêm apenas o estado de agregados de fluxos (e não de fluxos individuais), sendo a largura de banda dos agregados ajustada de forma dinâmica. São desenvolvidos modelos analíticos e são efectuados estudos de simulação para analisar o compromisso entre a carga de sinalização e a utilização de recursos. Estes estudos mostram que a hierarquização da rede, ou seja, a sua divisão em áreas mais pequenas, e a configuração de agregados entre os *routers* fronteira das áreas, por oposição a agregados extremo-a-extremo entre os *routers* fronteira do domínio, permitem atingir uma utilização de recursos próxima da utilização com sinalização fluxo-a-fluxo, mas com uma carga de sinalização significativamente inferior.

abstract

This Thesis addresses the problem of resource management in ATM (Asynchronous Transfer Mode) and IP (Internet Protocol) networks with QoS (Quality of Service) support. ATM and IP technologies allow, nowadays, the integration of different types of services in the same network infrastructure. However, the variety of services and their requirements, involve big challenges at the resource management level, which needs to be as efficient as possible.

This Thesis is divided in two parts. The first part of the Thesis addresses the resource management of ATM and IP access networks with QoS support. It is presented an historical perspective of the access networks evolution. In terms of ATM access networks, resource management strategies based on VPs (Virtual Paths) are proposed, and dimensioning methodologies are defined, which take into account the QoS requirements at the call and cell level. The considered resource management strategies allow the establishment of trade-offs between the resource utilization and the signaling load. In terms of IP access networks, it is proposed a new network architecture that represents an evolution of the legacy access networks, allowing a larger resource sharing. This architecture allows QoS differentiation and support for multimedia applications. More specifically, it is proposed a solution that includes the integrated support of recently introduced technologies to establish and configure multimedia sessions, manage the QoS policies including AAA (Authentication, Authorization and Accounting) functions, and reserve network resources.

In the second part of the Thesis, two scalable call admission control mechanisms are addressed: probing mechanisms and mechanisms based on the aggregation of individual reservations. Both mechanisms allow resource management without the maintenance of the per-flow state in each network element. The probing mechanism estimates the network QoS level, through the insertion of probe flows, to decide if a new flow can be accepted or not. A new probing mechanism is proposed, denoted by ?-probing, which minimizes the resource stealing problem that exists in the simple probing mechanism when it is applied to a system with multiple service classes. Analytical models are developed and simulation studies are performed to analyze the resource stealing problem and to determine the influencing factors on the estimation of the loss ratio performed by the probing and ?-probing flows. The ?-probing mechanism obtains, simultaneously, a high resource utilization and service differentiation without resource stealing. The results obtained with this mechanism were also validated through the development of a probing test-bed. In the mechanism based on the aggregation of individual reservations, the core network elements only need to maintain the state of the flows' aggregate (instead of individual flows), and the aggregates' bandwidth is dynamically adjusted. Analytical models are defined and simulation studies are performed to analyze the trade-off between signaling load and resource utilization. These studies show that the hierarquization of the network, that is, its partition in smaller areas, and the configuration of aggregates between area border routers, as opposed to end-to-end aggregates between domain border routers, reaches an utilization close to the one of per-flow signaling with a signaling load significantly smaller.

índice

CAPÍTULO 1

INTRODUÇÃO

Nas últimas décadas as redes de telecomunicações têm vindo a sofrer modificações muito profundas. A evolução das tecnologias de transmissão, dos equipamentos terminais, das técnicas de processamento dos sinais de áudio e de vídeo, entre outros, suscita o constante desenvolvimento de novas aplicações e também de inúmeras oportunidades de negócio. Os novos serviços e aplicações são gradualmente mais exigentes, apresentando características e requisitos de transmissão muito diversos. Além disso, com o aparecimento dos serviços multimédia, que integram vários tipos de fluxos de tráfego, cada um com os seus requisitos particulares, torna-se necessário dar um tratamento adequado a cada fluxo. Os diferentes requisitos são quantificados por parâmetros de Qualidade de Serviço (QdS). Por QdS entende-se a capacidade de a rede fornecer recursos a serviços e/ou clientes de forma diferenciada, de acordo com os seus requisitos.

A diversidade de serviços existentes hoje em dia motiva a convergência para uma rede de telecomunicações que suporte qualquer tipo de serviço (rede integradora), e que simultaneamente seja capaz de assegurar o tratamento mais adequado a cada um (rede diferenciadora). A primeira tecnologia que permitiu simultaneamente integrar todos os serviços na mesma infra-estrutura e atribuir-lhes uma QdS diferenciada de acordo com os respectivos requisitos foi o ATM (*Asynchronous Transfer Mode*). A tecnologia ATM,

desenvolvida para dar suporte à RDIS-BL (Rede Digital com Integração de Serviços de Banda Larga), cumpre os requisitos exigidos por esta para implementação de uma rede de telecomunicações multi-serviços, devido aos débitos elevados que lhe estão associados, ao suporte de QdS diferenciada e à capacidade de gestão das ligações. Na tecnologia ATM, a informação digitalizada proveniente dos vários tipos de serviços é transportada em pequenos pacotes denominados de células. O ATM permite suportar serviços desde a emulação de circuitos até ao transporte de dados sem requisitos de QdS. Existem várias vantagens nas redes baseadas em ATM. A primeira vantagem óbvia é o facto de poder integrar numa mesma infra-estrutura de rede uma grande variedade de aplicações com requisitos de QdS diferentes e com tráfego heterogéneo. Além disso, o ATM permite o suporte de comunicação *multicast* e gestão da QdS extremo-a-extremo, apresentando portanto, uma grande versatilidade.

O facto da tecnologia ATM ser orientada à ligação favorece a existência de todas as vantagens descritas anteriormente, mas acarreta consigo uma desvantagem considerável: torna o ATM demasiado complexo devido à necessidade de utilização de procedimentos de sinalização e de gestão para o estabelecimento das ligações. Além desta desvantagem, o pequeno comprimento de cada pacote (53 *bytes* com 5 *bytes* de cabeçalho) faz com que a percentagem da informação não útil (*overhead*) transportada seja elevada. Finalmente, com o crescimento exponencial da *Internet* a que se tem assistido nos últimos anos, espera-se que em pouco tempo o tráfego IP (*Internet Protocol*) seja dominante. Por estas razões, este protocolo tem vindo a ganhar preponderância como tecnologia de transporte nas redes de telecomunicações. Espera-se que com o IP, alguma da complexidade acima referida seja minorada. O IP é também baseado em pacotes, mas de comprimento variável (até um máximo de 65,536 *bytes* no caso do IPv4 e teoricamente ilimitado no caso do IPv6). No entanto, o IP nativamente não suporta qualquer mecanismo de diferenciação de serviços e de suporte de QdS. Por estas razões, ele tem vindo a ser dotado de mecanismos no sentido de permitir o suporte de uma QdS diferenciada. Alguns destes mecanismos limitam a quantidade de tráfego na rede para impedir o congestionamento da mesma, enquanto outros permitem diferenciar entre os diversos tipos de serviços. Estes mecanismos incluem: (i) o controle de admissão, que bloqueia um pedido de utilização da rede, caso não existam recursos suficientes; (ii) o controle do nível de congestão da rede, que limita a taxa máxima à qual são enviados os pacotes, e é responsável pelo descarte de alguns pacotes se a rede se

encontrar em situação de congestionamento e; (iii) o escalonamento dos pacotes, que define a ordem pela qual são servidos em cada nó da rede.

Com o objectivo de dotar a rede IP de suporte e diferenciação de QdS foram definidas pelo *Internet Engineering Task Force* (IETF) [IETF-int] duas arquitecturas: a arquitectura de Integração de Serviços (IntServ) e a de Diferenciação de Serviços (DiffServ). A arquitectura IntServ garante uma QdS diferenciada para cada serviço através da reserva de recursos efectuada individualmente para cada fluxo de tráfego com base em controle de admissão. A arquitectura DiffServ usa uma combinação de mecanismos de diferenciação entre conjuntos de serviços com requisitos de QdS semelhantes (classes de serviço), de atribuição de recursos a cada classe de serviço, e de policiamento (limitação) da informação que é transmitida em cada classe.

A arquitectura IntServ permite um suporte eficiente de QdS diferenciada. No entanto, esta arquitectura apresenta algumas desvantagens relacionadas com a necessidade de efectuar um processamento de admissão e de reserva de recursos, por cada fluxo de tráfego, em todos os elementos da rede (*routers*) que se encontram no percurso do fluxo. Deste modo, os elementos da rede necessitam de manter informação actualizada relativa a todos os fluxos activos. Estas funcionalidades são muito difíceis de implementar numa rede de elevado débito, pois necessitam de elementos com uma elevada capacidade de processamento e armazenamento. A arquitectura DiffServ não apresenta estes problemas, uma vez que trata agregados de fluxos e não fluxos individuais. No entanto, devido a esta mesma característica, os fluxos activos pertencentes a uma mesma classe podem degradar mutuamente as respectivas QdS. Este problema não existe na arquitectura IntServ porque são efectuadas reservas de recursos para cada fluxo de tráfego. Com o objectivo de aproveitar o melhor das duas arquitecturas, têm sido propostas novas arquitecturas e mecanismos que tentam simultaneamente garantir QdS a cada fluxo individual e diminuir os requisitos de processamento e armazenamento de informação dos elementos da rede.

Nesta Tese são abordados diversos aspectos da gestão de recursos em redes ATM e IP com suporte de QdS. A Tese contém duas partes principais. A primeira parte está relacionada com a problemática de gestão de recursos em redes de acesso ATM e IP com suporte de QdS. Ao nível das redes de acesso ATM, são propostas estratégias de gestão de recursos, e metodologias de dimensionamento da rede, tendo em conta requisitos de QdS tanto ao nível da chamada como da célula. Ao nível das redes de acesso IP, é proposta uma

nova arquitectura para redes de acesso com suporte de QdS diferenciada e serviços multimédia. A segunda parte da Tese incide em dois mecanismos de controle de admissão escaláveis (que minimizam o processamento e o estado armazenado nos elementos da rede). O primeiro mecanismo, denominado de *probing*, determina o nível de QdS da rede, através da inserção de fluxos de teste, para decidir se um novo fluxo pode ou não ser aceite. O segundo mecanismo é baseado em agregação de reservas individuais e determina se existem ou não recursos disponíveis na rede para aceitar um novo fluxo, com base na informação de reservas de agregados. Ambos os mecanismos permitem uma gestão de recursos com suporte de QdS diferenciada, minimizando a carga de sinalização nos elementos da rede.

1.1 Estrutura da Tese

Esta Tese está organizada da seguinte forma: os capítulos 2 a 4 são introdutórios e fazem o levantamento do estado-da-arte dos diferentes problemas endereçados na Tese; os capítulos 5 a 9 descrevem o trabalho original efectuado no âmbito da Tese.

O capítulo 2 descreve os aspectos mais importantes da tecnologia ATM, que servem de base aos estudos efectuados em redes de acesso ATM no capítulo 5. Primeiro, são descritos os conceitos básicos do ATM, a célula ATM, o esquema de endereçamento e a construção de redes lógicas. De seguida, são apresentadas as características do ATM que lhe permitem suportar QdS. Em primeiro lugar, são descritas as classes de serviço suportadas pelo ATM e os parâmetros de QdS que as definem. Depois, são referidos os parâmetros de caracterização do tráfego que definem as fontes de tráfego relacionadas com um determinado serviço. Finalmente, são apresentados os mecanismos de controle de tráfego e de congestionamento, ou seja, a gestão dos recursos da rede com base em VPs (*Virtual Paths*), a limitação do número de chamadas activas através de mecanismos de CAC (*Call Admission Control*), e o controle do tráfego enviado pelo utilizador através de mecanismos de UPC (*Usage Parameter Control*). Este capítulo termina com uma pequena descrição das principais vantagens e desvantagens da tecnologia ATM.

O capítulo 3 discute a problemática do suporte de QdS em redes IP. A primeira parte concentra-se na descrição dos diversos mecanismos que permitem dotar o IP de suporte de diferenciação de QdS. Seguidamente, são apresentadas as duas arquitecturas definidas pelo IETF para o efeito, IntServ e DiffServ, os seus modelos de serviço, os seus blocos

constituintes, e as principais vantagens e desvantagens de cada arquitectura. É apresentada também uma outra arquitectura que considera a existência de IntServ nas redes de acesso e DiffServ nas redes do núcleo, sendo os recursos destas últimas geridos centralmente por *Bandwidth Brokers*. Finalmente, são descritas algumas arquitecturas e mecanismos de gestão de recursos e suporte de QdS que tentam estabelecer compromissos entre as arquitecturas IntServ e DiffServ.

O capítulo 4 aborda as redes de acesso e a sua evolução, desde as redes de acesso de circuitos comutados às redes de acesso IP. Nesta perspectiva de evolução é feita uma análise das redes de acesso tradicionais, baseadas em comutação de circuitos, das redes de acesso ADSL (*Asymmetrical Digital Subscriber Line*), das redes de acesso baseadas em fibra com tecnologia ATM, das redes de CATV (*Community Antenna TeleVision*) e redes de acesso IP. Este capítulo termina com uma pequena introdução ao trabalho efectuado sobre redes de acesso no capítulo 6.

O capítulo 5 endereça o dimensionamento de redes de acesso ATM com múltiplas classes de serviço. Primeiro, é apresentada a arquitectura geral da rede de acesso, os seus elementos constituintes e as interfaces físicas. De seguida, é apresentada uma tipificação dos serviços suportados pelas redes de acesso multi-serviços de banda larga. São propostas três estratégias de gestão de recursos que representam diferentes compromissos entre a eficiência na utilização dos recursos e o custo dos elementos de rede. São também apresentadas propostas de metodologias de dimensionamento para cada estratégia definida, que consideram os requisitos de QdS ao nível da chamada e da célula. Finalmente, são feitos estudos comparativos que incluem a avaliação das diferentes estratégias de gestão de recursos, de diferentes métodos de atribuição de largura de banda (ao PBR - *Peak Bit Rate* - e através de larguras de banda efectivas), e de diferentes distribuições espaciais dos utilizadores na rede de acesso.

No capítulo 6 é apresentada uma proposta de uma nova arquitectura para redes de acesso baseadas em IP. Primeiro, são definidos os elementos da rede e as interfaces físicas de ligação entre os mesmos. De seguida, é apresentada uma forma de reutilização do PPP (*Point-to-Point Protocol*) nas redes de acesso com recursos partilhados. Depois, é apresentada a proposta de uma rede de acesso IP que permite suportar QdS diferenciada e aplicações multimédia. Os protocolos e tecnologias escolhidos para o efeito incluem o SIP (*Session Initiation Protocol*) para iniciar e configurar uma sessão, o COPS (*Common Open*

Policy Service) para gerir as políticas de QdS incluindo as funcionalidades de AAA (*Authentication, Authorization and Accounting*), e o RSVP (*resource ReSerVation Protocol*) para reservar os recursos da rede. Como parte integrante da proposta, são determinados os protocolos e tecnologias a suportar em cada elemento de rede.

O capítulo 7 apresenta uma proposta de um mecanismo de controle de admissão baseado na investigação do estado de congestionamento da rede (*probing*), que minimiza o problema do roubo de recursos que afecta sistemas com múltiplas classes de serviço. Primeiro, é feita uma apresentação dos mecanismos de *probing*, sendo discutidos alguns dos problemas destes mecanismos, incluindo o problema do roubo de recursos. Depois, apresenta-se a proposta de um novo mecanismo de *probing* que permite minimizar este problema do roubo de recursos, que se designou por ?-*probing*. É apresentado um primeiro modelo teórico baseado numa distribuição binomial para estudar os factores que influenciam a estimativa do rácio de perdas efectuada pelos fluxos de *probing*. De seguida, é apresentado um segundo modelo baseado em cadeias de *Markov* que permite estudar o problema do roubo de recursos em redes multi-serviço. Logo após, são efectuados estudos numéricos e de simulação para avaliar o desempenho dos mecanismos de *probing* e de ?-*probing*, e determinar as características que os fluxos de *probing* e de ?-*probing* necessitam de possuir para detectar as violações de QdS nas várias classes de serviço definidas.

No capítulo 8 é descrito um estudo experimental dos mecanismos de *probing* que passou pelo desenvolvimento de um protótipo laboratorial. Primeiro, é apresentada a arquitectura do protótipo e são discutidas as opções tomadas na concepção do mesmo. De seguida são descritos os módulos desenvolvidos para permitir a introdução de mecanismos de *probing* numa rede real: um módulo de geração de tráfego e um módulo de *probing* que implementa o processo de *probing* e de ?-*probing*. Finalmente, são descritos os testes efectuados e é feita uma análise comparativa dos resultados experimentais e de simulação.

No capítulo 9 é apresentado um estudo de desempenho de uma rede que suporta agregação de reservas de recursos individuais. Consideram-se dois tipos de sistemas: um domínio plano (uma única área) com agregação extremo-a-extremo, e um domínio hierarquizado (dividido em áreas) com agregação extremo-a-extremo no interior de cada área. Inicialmente, é feita uma abordagem da problemática da sinalização e do estabelecimento de reservas para fluxos individuais e para agregados de fluxos. São desenvolvidos dois modelos para comparar os dois tipos de sistemas. No primeiro modelo,

denominado de modelo de carga por fluxo, a carga oferecida é detalhada ao nível do fluxo. O segundo modelo, denominado de modelo de carga por agregado, considera que o tráfego oferecido tem uma largura de banda agregada média que é variável no tempo. São efectuados estudos numéricos e de simulação para comparar os dois tipos de sistemas ao nível da carga de sinalização, da utilização dos recursos e da probabilidade de bloqueio dos fluxos. Estes estudos são também complementados com estudos de simulação que utilizam medidas de tráfego real.

Finalmente, as conclusões principais do trabalho descrito nesta Tese são apresentadas no capítulo 10. São também apresentadas as propostas de trabalho futuro.

1.2 Enquadramento do trabalho

O trabalho desenvolvido nesta Tese foi parcialmente enquadrado em vários projectos:

? Dimensionamento de redes de acesso ATM – enquadrado no projecto europeu do programa ACTS BBL (*BroadBand Loop*) (AC0038).

? Redes de acesso IP com suporte de QdS – enquadrado nos projectos DIMIP (DIMensionamento de redes de acesso IPng) e GESACESSO (Gestão de redes de acesso IP) do Instituto de Telecomunicações (financiados ambos pela Portugal Telecom Inovação).

? Mecanismos de *probing* – iniciado na Universidade de *Rice* em *Houston*, nos EUA, e continuado no âmbito do projecto TELEMAT (Engenharia de Tráfego para Redes DiffServ/MPLS) do Instituto de Telecomunicações (financiado pela Portugal Telecom Inovação), e do projecto europeu P1112 do EURESCOM NEW DIMENSIONS (NEtWork DIMENSIONing baSed on modeling of internet traffic).

? Agregação em domínios hierárquicos – enquadrado nos projectos TELEMAT e NEW DIMENSIONS.

1.3 Principais contribuições

? Proposta de estratégias de gestão de recursos baseadas em VPs, e das correspondentes metodologias de dimensionamento, para redes de acesso ATM.

- Proposta de uma nova arquitectura para redes de acesso baseadas em IP com suporte de QdS diferenciada e de serviços multimédia. Definição dos mecanismos e tecnologias a suportar em cada elemento de rede.
- Proposta de um novo mecanismo de *probing*, designado por ?-*probing*, que permite atenuar o problema do roubo de recursos que afecta sistemas com múltiplas classes de serviço. Determinação das características dos fluxos de *probing* e de ?-*probing* que permitem simultaneamente obter diferenciação na QdS e minimizar o problema do roubo de recursos. Validação experimental do desempenho destes mecanismos através da implementação de um protótipo laboratorial.
- Proposta de modelos teóricos para analisar o desempenho da agregação de fluxos em domínios planos e domínios hierarquizados. Estudo dos compromissos existentes entre carga de sinalização e utilização de recursos.

CAPÍTULO 2

ATM E QUALIDADE DE SERVIÇO

A tecnologia ATM (*Asynchronous Transfer Mode*) permite a implementação de uma rede multi-serviços com suporte de QdS (Qualidade de Serviço), devido aos débitos elevados que lhe estão associados, ao suporte de QdS diferenciada e à capacidade de gestão das ligações. O ATM é essencialmente um modo de transporte baseado em comutação de pacotes que usa pacotes pequenos com comprimento fixo. A tecnologia ATM permite suportar serviços desde a emulação de circuitos até ao transporte de dados sem requisitos de QdS.

2.1 Conceitos básicos

O ATM é um modo de transferência orientado à ligação que permite transportar todo o tipo de serviços sobre uma mesma rede física. O facto de ser orientado à ligação significa que apenas pode haver transferência de informação entre dois terminais após se ter estabelecido uma ligação virtual entre eles através de procedimentos de sinalização e/ou gestão.

Na tecnologia ATM a informação de cada fluxo de tráfego é inserida em pacotes pequenos de comprimento fixo, 53 octetos, denominados de células. O facto de os pacotes terem comprimento fixo simplifica significativamente os mecanismos de sincronismo de pacote e permite uma comutação rápida dos mesmos. Cada célula é enviada para a rede

após ter sido completamente preenchida, sendo transmitida ao ritmo máximo permitido pela capacidade das ligações. As células são multiplexadas na rede física através de multiplexagem assíncrona no tempo, ou seja, a partilha dos recursos da rede entre as várias chamadas activas é feita no tempo, e não é necessariamente atribuída uma largura de banda fixa de transmissão a cada chamada. Assim, cada fonte de tráfego associada a cada chamada pode transmitir a um ritmo variável desde que a capacidade das ligações atravessadas não seja excedida, e pode usar a largura de banda que não está a ser utilizada por outras fontes de tráfego activas. Além disso, como cada fonte pode usar a largura de banda que não está a ser usada por outras, uma ligação física pode suportar várias fontes de ritmo variável mesmo que a soma dos ritmos de pico de cada fonte seja superior à sua capacidade. Este processo de aproveitamento dos recursos numa rede de pacotes é denominado de multiplexagem estatística. Através da multiplexagem estatística das células ATM é possível diminuir significativamente a largura de banda necessária para suportar um conjunto de fontes de tráfego.

2.2 Modelo de referência

A descrição das funcionalidades do ATM baseia-se no modelo de referência protocolar da RDIS-BL (Rede Digital com Integração de Serviços de Banda Larga) definido pela recomendação I.321 [ITU-T91, ITU-T95]. O modelo de referência da tecnologia ATM é ilustrado na Figura 2-1.

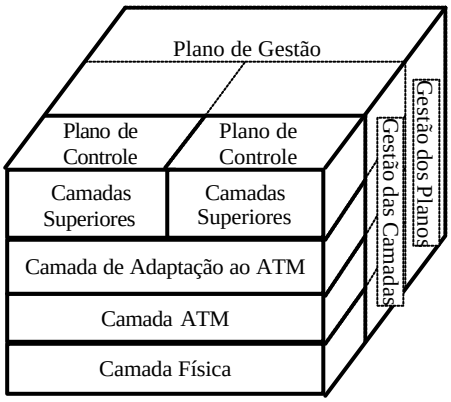


Figura 2-1 : Modelo de referência protocolar da RDIS-BL (norma I.321 ITU-T).

As três camadas inferiores (física, ATM e de adaptação) correspondem às duas camadas inferiores (física e da ligação lógica) do modelo OSI (*Open Systems Interconnection*). No equipamento terminal, a informação do plano do utilizador é mapeada de camadas protocolares superiores para a camada de adaptação do ATM (AAL – *ATM Adaptation Layer*). Esta camada é responsável por assegurar as características dos serviços quando estes são transportados em ATM. É a camada AAL que executa as funções específicas referentes a cada tipo de serviço, que segmenta a informação, e que envia à camada ATM a informação que será colocada no campo de informação de 48 octetos das células ATM, e a informação relevante para formar o cabeçalho de cada célula. Existem vários tipos de AAL [ITU-T93c, ITU-T93d], definidas para diferentes tipos de serviços.

A camada ATM é comum a todos os serviços e é responsável pelo serviço de transporte das células. Esta camada realiza a multiplexagem e desmultiplexagem das células das diversas chamadas activas, gera e extrai os cabeçalhos das células, controla o fluxo de informação, e nos comutadores facilita o processo de encaminhamento através da translação dos valores de identificador de caminho virtual (VPI - *Virtual Path Identifier*) e de identificador de canal virtual (VCI - *Virtual Channel Identifier*) descritos de seguida. Esta última função permite suportar QdS ao nível da camada ATM.

A camada física define as características eléctricas e as interfaces da rede, garantindo independência entre a camada ATM e o tipo de transporte físico.

Este modelo de referência está também estruturado em 3 planos: gestão, controle e utilizador. As camadas descritas anteriormente pertencem ao plano de utilizador, que tem como algumas das suas responsabilidades a transferência dos dados do utilizador e o seu controle ao nível do fluxo, e a recuperação de erros do cabeçalho. O plano do controle é responsável pelo controle das ligações e chamadas através de procedimentos de sinalização. O plano de gestão coordena a gestão de recursos do sistema.

2.3 Célula ATM

Foi já referido anteriormente que uma célula ATM tem um comprimento fixo de 53 octetos, em que os 5 primeiros pertencem ao cabeçalho e os 48 seguintes transportam informação de dados. O cabeçalho apresenta uma estrutura diferente (Figura 2-2) nos dois tipos de interface de uma rede ATM: UNI (*User to Network Interface*), entre o

equipamento terminal e um comutador ATM, e NNI (*Network to Network Interface*), entre dois comutadores ATM.

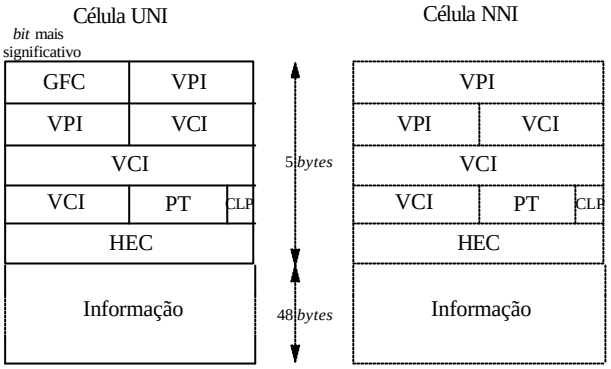


Figura 2-2 : Estrutura de uma célula ATM.

O cabeçalho é constituído pelos campos seguintes:

- GFC (*Generic Flow Control*) – parâmetro independente do meio físico e usado para controlar o fluxo de informação do terminal para a rede.
- VPI/VCI – parâmetros que representam a informação de encaminhamento. Com base nestes parâmetros é possível identificar, numa ligação física, a que ligação lógica pertence uma determinada célula.
- PT (*Payload Type*) – parâmetro que define se a célula transporta dados ou é usada para gestão.
- CLP (*Cell Loss Priority*) – parâmetro que define a prioridade de descarte das células: as células que contêm este *bit* a 1 são descartadas com maior prioridade do que as células que contêm o *bit* a 0. Este parâmetro pode ser alterado pela rede para indicar que o utilizador não está a cumprir o contrato de tráfego negociado no estabelecimento da chamada.
- HEC (*Header Error Control*) – parâmetro que permite a detecção e correcção de erros no cabeçalho.

2.4 Funcionalidades dos VPs e VCs

Numa rede ATM é possível multiplexar numa mesma ligação física, células de diferentes ligações lógicas, uma vez que estas são identificadas através dos identificadores VPI e

VCI. Os identificadores são atribuídos entre dois equipamentos adjacentes no estabelecimento de uma chamada e mantêm-se até à terminação da mesma.

Células pertencentes a chamadas diferentes podem ser agrupadas no mesmo caminho virtual (VP – *Virtual Path*), sendo neste caso identificadas através de um mesmo VPI. Este agrupamento permite reduzir a carga de processamento associada aos mecanismos de controle e gestão de tráfego. Assim, para um determinado conjunto de chamadas, estes mecanismos podem ser executados apenas ao nível do VP para todos os VCs (*Virtual Channels*) que o VP contém. A relação entre VPs e VCs é representada na Figura 2-3.

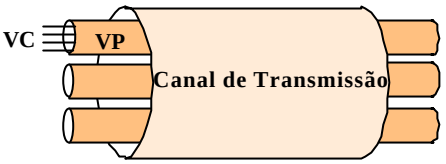
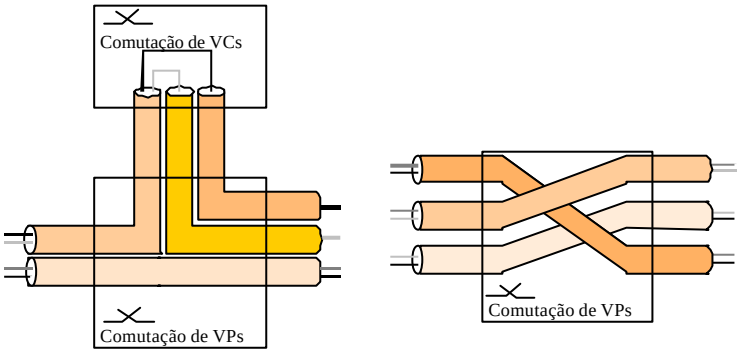


Figura 2-3 : Relação entre VPs e VCs.

As células ATM podem ser comutadas na rede a dois níveis diferentes: ao nível do VP e do VC. Designam-se habitualmente por comutadores ATM os comutadores que têm a capacidade de efectuar a comutação tanto ao nível dos VPs como dos VCs, e por *cross-connects* os que apenas permitem efectuar comutação ao nível dos VPs. A Figura 2-4 ilustra estes diferentes níveis de comutação.



Comutador ATM ***Cross-connect* ATM**

Figura 2-4 : Comutação de células ATM.

Os VPI e VCI definem uma ligação virtual (VCC – *Virtual Channel Connection*). A cada VCC são atribuídos identificadores VPI e VCI em todas as ligações físicas que atravessa.

O ATM permite a criação de várias redes lógicas suportadas por uma mesma rede física através da criação de caminhos virtuais (VPC – *Virtual Path Connection*). Um VPC é caracterizado pela sua origem e destino, pelo percurso que atravessa na rede física e pela largura de banda que lhe é reservada em todas as ligações entre a origem e o destino. A Figura 2-5 ilustra a configuração de duas redes lógicas sobre a mesma rede física. A rede física é constituída por 4 comutadores e 4 ligações, e sobre ela são configurados dois conjuntos de VPCs que dão origem a duas redes lógicas distintas. Cada serviço pode ser transportado numa rede lógica diferente. A topologia de uma rede lógica é independente da topologia da rede física. No entanto, a capacidade reservada para todas as redes lógicas não pode ser superior à capacidade das ligações físicas que as suportam.

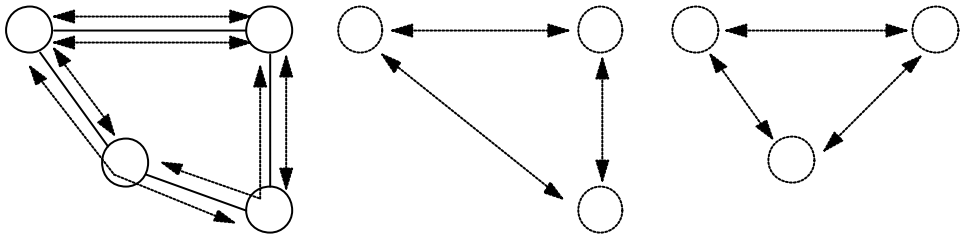


Figura 2-5 : Definição de redes lógicas com VPCs: (a) rede física, (b) rede lógica 1 e (c) rede lógica 2.

Após o estabelecimento de um VPC, vários VCCs podem ser estabelecidos e encaminhados no mesmo VPC. Como um VPC é, do ponto de vista conceptual, um túnel extremo-a-extremo com requisitos de QdS pré-definidos, por cada VCC que é estabelecido é necessário sinalização apenas entre os dois extremos do túnel. Deste modo, o tempo de estabelecimento e a carga de sinalização diminuem. Além disso, é também mais fácil controlar e prevenir situações de congestão da rede.

2.5 Suporte de QdS

Numa rede ATM é possível garantir uma QdS requerida por cada serviço e/ou utilizador. A QdS que é atribuída a uma determinada chamada tem de ser garantida independentemente da quantidade de tráfego correspondente a outras chamadas activas na rede. Por outro lado, deve-se controlar a quantidade de tráfego em cada ligação, de modo a evitar o seu congestionamento. A rede ATM contém funções de controle de tráfego e de congestionamento que estão definidas na camada ATM. Este controle é conseguido através

do uso de mecanismos de controle do número de chamadas activas na rede, denominados de mecanismos de CAC (*Call Admission Control*) e de mecanismos de policiamento de tráfego, denominados de mecanismos de UPC (*Usage Parameter Control*).

De uma forma geral, a QdS é quantificada por parâmetros, em que diferentes combinações de parâmetros definem classes de serviço. As fontes de tráfego são também descritas por parâmetros que caracterizam a forma como as células são transmitidas. O tipo de transferência ATM define o serviço de transferência de informação prestado pela camada ATM e é definido pelos parâmetros de QdS e de tráfego.

Nas secções seguintes vão ser descritos com mais detalhe cada um destes componentes.

2.5.1 Parâmetros de QdS e classes de serviço

Os parâmetros de QdS que podem ser considerados numa rede ATM são o rácio de células com erros (CER - *Cell Error Ratio*), o rácio de células perdidas (CLR - *Cell Loss Ratio*), o ritmo de células mal inseridas (CMR - *Cell Misinsertion Rate*), o rácio de blocos de células gravemente afectados por erros (SECBR – *Severely Errored Cell Block Ratio*), o atraso de transferência (CTD – *Cell Transfer Delay*) e a variação do atraso das células (CDV – *Cell Delay Variation*). Os parâmetros mais relevantes no âmbito do trabalho realizado nesta Tese são o CLR, CTD e CDV.

O CLR é a razão entre o número de células perdidas nas ligações entre o ponto origem e o ponto destino e o número de células transmitidas pelo ponto origem. Uma célula é considerada perdida se, após decorrido um determinado intervalo de tempo desde a emissão da célula num ponto origem, esta não tiver chegado ao seu ponto destino. O CTD é o tempo (médio ou máximo) que decorreu entre a emissão das células num ponto origem e a sua recepção num ponto destino. O CDV pode ser definido de duas formas diferentes consoante o evento em relação ao qual o atraso é medido. O CDV pode ser definido pela variação da diferença entre os tempos de chegada de uma célula e os tempos previstos, considerando que as células são transmitidas ao ritmo definido pelo *Peak Cell Rate* (ritmo de pico) da fonte de tráfego. O CDV pode também ser definido pela variação da diferença entre os tempos de transmissão das células.

Uma classe de serviço é definida por uma combinação de parâmetros de desempenho da rede e os seus limites que deverão ser garantidos pela rede. A classe de serviço é solicitada à rede aquando do estabelecimento da chamada, e os requisitos desta classe têm

de ser mantidos pela rede durante toda a duração da chamada e independentemente do tráfego das outras chamadas activas na rede. Um VPC pode conter VCCs com requisitos diferentes de QdS. No entanto, para não prejudicar as chamadas com requisitos mais estritos de QdS, deverá ser garantida a todos os VCCs a QdS referente aos VCCs com requisitos mais estritos. Claramente esta possibilidade não é a melhor em termos de optimização dos recursos e, em geral, cada VPC deve multiplexar apenas VCCs com requisitos de QdS semelhantes.

Nas normas definidas pelo ITU-T [ITU-T93a, ITU-T93b] e pelo ATM Forum [AForum10, AForum56], as definições das classes de serviço são ligeiramente diferentes.

O ITU-T define 3 classes e considera também uma classe não especificada sem quaisquer garantias de QdS. A classe 1 é a classe mais exigente e especifica limites estritos para o CTD. Esta classe é usada para serviços em que a transferência das células tem de ser efectuada a um ritmo constante e sem atrasos e, deste modo, permite fazer emulação de circuitos dedicados. A classe 2 é mais tolerante e impõe restrições aos parâmetros CLR, CER, CMR e SECBR. A classe 3 impõe restrições ao CLR mas não tem quaisquer limites de atrasos.

O ATM Forum define 4 classes e uma quinta classe sem garantias de QdS. A classe A tem limites de atraso estrito e permite a emulação de circuitos dedicados. A classe B permite ter um ritmo de transmissão variável mas com requisitos em tempo real. As classes C e D não impõem quaisquer restrições de atraso: a classe C permite efectuar uma transferência de dados orientada à ligação enquanto que a classe D permite a transferência de dados não orientada à ligação.

2.5.2 Parâmetros de caracterização do tráfego

As fontes de tráfego são descritas por um conjunto de parâmetros: *Peak Cell Rate* (PCR), *Sustainable Cell Rate* (SCR), *Minimum Cell Rate* (MCR), CDV *Tolerance* (CDVT), *Maximum Burst Size* (MBS) e o tipo de serviço. O PCR define uma taxa máxima de transmissão das células ATM e é calculado como o inverso do intervalo mínimo entre transmissões de células. O SCR define um limite superior da taxa média de transmissão das células numa determinada ligação. O MCR define o ritmo mínimo de transferência necessário pela aplicação. O CDV *Tolerance* indica qual é a tolerância de tempo que é permitida ao intervalo de tempo mínimo entre células. O MBS especifica o número máximo de células consecutivas que podem ser enviadas ao PCR. O tipo de serviço é um

parâmetro que pode ser definido pelo operador e que é usado para descrever, de uma forma implícita, outros quaisquer parâmetros de tráfego necessários.

Estes parâmetros das fontes de tráfego são declarados no estabelecimento da chamada e devem ser cumpridos durante todo o tempo em que esta decorrer. A norma ITU-T I.371 define um algoritmo genérico para determinar se o tráfego está de acordo com o estabelecido no início da chamada. Este algoritmo, denominado de *Generic Cell Rate Algorithm* (GCRA), verifica se o intervalo de tempo entre duas células é inferior ao permitido, e se isso acontecer, classifica a segunda célula como *non-conforming* (não respeitante), ou seja, coloca o *bit* CLP a 1.

2.5.3 Transferência ATM

Os tipos de transferência ATM designam os tipos de ligação oferecidos pela camada ATM. Mais precisamente, designa o conjunto de procedimentos e parâmetros de tráfego e de QdS para um determinado VPC ou VCC. A existência de diversos tipos de transferência permite escolher aquele que é mais adequado a uma aplicação, para rentabilizar os recursos na rede ATM, pois nem todas as aplicações necessitam, por exemplo, de uma largura de banda fixa. Os tipos de transferência definidos pelo ITU-T têm algumas diferenças relativamente aos definidos pelo ATM Forum. Enquanto que o ITU-T usa o termo Capacidades de Transferência ATM (ATC - ATM *Transfer Capabilities*), o ATM Forum usa a designação Categorias de Serviços (*Service Categories*). Neste capítulo são apenas apresentados os tipos de transferência definidos pelo ATM Forum, isto é, as categorias de serviços. O ATM Forum definiu 5 categorias de serviços descritas de seguida:

? Ritmo de transferência constante (CBR – *Constant Bit Rate*) – o CBR é usado em chamadas que necessitam de uma largura de banda constante, sendo possível emular um circuito dedicado. Os parâmetros de tráfego especificados no estabelecimento de cada chamada são o PCR e o CDVT. A QdS nesta categoria é definida pelo CDV, o CTD e CLR máximos.

? Ritmo de transferência variável com requisitos de tempo real (*rt*-VBR – *real time Variable Bit Rate*) – o *rt*-VBR é usado em aplicações com requisitos de tempo real, mas em que o tráfego apresenta um ritmo variável no tempo. No estabelecimento de cada chamada é necessário especificar o PCR e o seu CDVT, o SCR e o MBS. A QdS nesta categoria é definida pelo CDV, o CTD e CLR máximos.

? Ritmo de transferência variável sem requisitos de tempo real (*nrt*-VBR – *non-real time Variable Bit Rate*) – o *nrt*-VBR é, tal como o *rt*-VBR, usado em aplicações em que o perfil de tráfego é variável. No entanto, esta categoria de serviço não é adequada a aplicações com requisitos de tempo real porque não garante atrasos máximos nas transmissões das células. Os parâmetros especificados no estabelecimento da chamada são o PCR, o SCR e o MBS. A QdS nesta categoria é definida pelo CLR.

? Ritmo de transferência disponível (ABR – *Available Bit Rate*) – o ABR é usado por aplicações em que seja possível modificar o ritmo de transferência de dados em função do estado de congestão da rede. Os parâmetros especificados no estabelecimento de uma chamada são o MCR e o PCR. A esta categoria apenas é garantido o MCR, podendo a aplicação usar até ao PCR. Se a rede se tornar congestionada, o ritmo de transmissão pode ser reduzido para o MCR. A QdS nesta categoria é definida pelo CLR, mas este não é tão estrito como nas categorias de serviço descritas anteriormente.

? Ritmo de transferência não especificado (UBR – *Unspecified Bit Rate*) – o UBR é adequado para aplicações sem requisitos de tempo real, tais como a transferência de ficheiros e o correio electrónico. O único parâmetro de tráfego especificado é o PCR, e não é especificado qualquer parâmetro de QdS.

2.5.4 Controle de tráfego e congestionamento

Numa rede ATM coexistem vários tipos de serviços caracterizados por fontes de tráfego diferentes e com requisitos diferentes de QdS. Esta diversidade de serviços na mesma rede permite a existência de multiplexagem estatística e, desta forma, permite aproveitar eficientemente os recursos da rede. No entanto, conseguir aproveitar ao máximo os recursos e garantir ao mesmo tempo uma QdS específica para cada chamada, exige que sejam implementados mecanismos de gestão dos recursos, de controle do tráfego total activo em cada ligação e, caso o tráfego seja excessivo, têm também de ser implementados mecanismos de controle de congestão.

Os mecanismos de controle de tráfego têm como objectivo prevenir que a rede se encontre congestionada com o intuito de garantir que os requisitos de QdS não sejam violados. Os mecanismos de controle de congestão são accionados quando existem

indicações de que a rede está congestionada, ou que está a entrar numa situação de congestão.

Nas secções seguintes são apresentados de uma forma breve estes mecanismos e as suas funções associadas.

2.5.4.1 Gestão dos recursos da rede

A função da gestão dos recursos da rede (NRM – *Network Resource Management*) é configurar os recursos de rede existentes e distribuí-los da forma mais eficaz pelo conjunto dos serviços que são suportados pela rede, de modo a que sejam rentabilizados. Como já foi referido na secção 2.4, numa rede ATM é possível agrupar os serviços com as mesmas características ou requisitos de QdS num mesmo VPC, permitindo simplificar os mecanismos de controle de tráfego, distribuir de uma forma conjunta as mensagens de operação dos mecanismos de controle de tráfego, e executar as funções dos mecanismos de controle de congestão ao conjunto de ligações agregadas.

2.5.4.2 Controle de admissão de chamadas

O CAC tem como função determinar se uma chamada pode ou não ser aceite na rede. A chamada apenas pode ser aceite se existirem recursos disponíveis suficientes em todas as ligações que serão atravessadas pelo tráfego da chamada, para garantir, por um lado, a QdS requerida pela nova chamada, e por outro, a QdS que foi previamente acordada para as chamadas já activas na rede. Este último ponto tenta impedir que o tráfego de uma nova chamada possa degradar a QdS de outras chamadas previamente admitidas na rede.

Normalmente, o CAC é efectuado no processo de estabelecimento de uma chamada. No entanto, o ATM permite que os parâmetros acordados nesta fase possam ser renegociados no decorrer da chamada.

Para determinar se uma chamada pode ou não ser aceite, são implementados algoritmos de CAC nos nós da rede. Estes algoritmos têm normalmente como parâmetros de entrada a capacidade das ligações, os parâmetros de tráfego das chamadas activas e da chamada que se pretende admitir, e os parâmetros de QdS das respectivas chamadas. No capítulo segunite são descritos com maior detalhe alguns destes algoritmos.

2.5.4.3 Controle dos parâmetros do utilizador

O UPC tem a função de comparar os parâmetros do tráfego que é enviado com os parâmetros acordados aquando do estabelecimento da chamada, e tomar medidas no sentido de colocar o tráfego não respeitante dentro do perfil acordado. Desta forma, o UPC protege o tráfego dentro do perfil do tráfego fora do perfil, de modo a impedir que este afecte o desempenho do serviço das chamadas em curso.

O UPC necessita, numa primeira fase, de monitorizar o tráfego correspondente às várias chamadas, para verificar se este está de acordo com os parâmetros que foram estabelecidos no início da chamada. O procedimento usado para verificar a conformidade dos parâmetros pode ser o GCRA ou um outro qualquer procedimento semelhante. Numa segunda fase, quando se verifica que uma célula (ou mais) não respeita os parâmetros de tráfego acordados, esta pode ser transferida sem se executar nenhuma acção, pode ser-lhe aumentada a prioridade de descarte através do campo CLP do cabeçalho da célula, ou pode ser eliminada. As funções UPC estão assim associadas a indicações específicas para a gestão de recursos, para renegociar a largura de banda atribuída, para a eliminação de blocos de células não respeitantes, ou para a reformatação do tráfego (*traffic shaping*) com o objectivo de o colocar dentro do perfil adequado. A reformatação do tráfego altera as características dos fluxos de tráfego de modo a atrasar algumas células que estão fora do perfil, de tal forma que, por exemplo, o valor do PCR não seja violado ou que o valor do CDV seja minimizado. A reformatação do tráfego é normalmente efectuada pela inserção do tráfego numa fila de espera servida a uma taxa máxima de transmissão pré-definida.

2.6 Vantagens e desvantagens do ATM

A descrição da tecnologia ATM apresentada mostra que o ATM apresenta uma elevada funcionalidade. A possibilidade de se poder integrar na mesma infra-estrutura qualquer tipo de serviços e aplicações, beneficiando da multiplexagem estatística entre células, e o consequente aproveitamento eficiente dos recursos com garantias de diferenciação de QdS, representou, na altura de aparecimento do ATM, uma mudança drástica nas redes de telecomunicações.

No entanto, a tecnologia ATM tem os seus pontos menos bons. A evolução e implementação massiva de equipamentos ATM por parte dos operadores de telecomunicações não ocorreu como inicialmente previsto devido a alguns destes pontos

menos bons. Em primeiro lugar a tecnologia ATM é orientada à ligação, o que requer a existência de sinalização relativamente complexa em todos os elementos da rede para estabelecer e libertar os caminhos virtuais. Além disso, o pequeno comprimento de cada célula diminui significativamente a quantidade de informação útil que é transportada, e consequentemente, diminui a utilização e o aproveitamento dos recursos (em quase 10%). Finalmente, como não existem serviços que sejam implementados em ATM, é necessário usar sempre a camada de adaptação AAL que introduz ainda uma redução adicional na utilização dos recursos.

CAPÍTULO 3

QUALIDADE DE SERVIÇO EM REDES IP

A rede *Internet*, tal como é conhecida, é um meio de troca de informação partilhado em que os recursos são repartidos de uma forma equitativa por todos os utilizadores e aplicações. A troca de informação entre os vários elementos da rede é baseada no modelo de datagrama, que é implementado ao nível da camada de rede pelo protocolo IP (*Internet Protocol*). Neste modelo, os pacotes são tratados individualmente pela rede, independentemente do utilizador que os envia e da aplicação à qual pertencem. Este modelo, para além da sua simplicidade, caracteriza-se também por se adaptar automaticamente a mudanças na topologia da rede. No entanto, a partilha de recursos efectuada actualmente na *Internet*, não tem em consideração o tipo de aplicação utilizada, as suas características e requisitos funcionais, nem permite que seja dada uma prioridade mais elevada a um determinado serviço ou utilizador. Este modelo de serviço, que se caracteriza por dar o melhor tratamento possível ao tráfego através de uma partilha equitativa dos recursos por todos os utilizadores, é denominado de modelo de melhor esforço (*Best Effort*). Neste modelo, enquanto os recursos disponíveis forem suficientes para servir todos os utilizadores, não haverá degradação na comunicação, mas logo que surja uma situação de congestão, todo o tráfego dos diferentes utilizadores será afectado.

Este tipo de serviço foi, até há poucos anos, satisfatório, pois as aplicações então utilizadas (*login* remoto, transferência de ficheiros através do FTP - *File Transfer Protocol*, *e-mail*) não eram muito sensíveis a degradação nas ligações, e os possíveis erros e perdas existentes eram corrigidos pelo protocolo da camada de transporte TCP (*Transmission Control Protocol*). O controle de erros no TCP, efectuado através da retransmissão de pacotes, introduz um atraso na transmissão dos mesmos. Para as aplicações que não são sensíveis ao atraso, o TCP é ideal no sentido em que garante um rácio de perdas de pacotes muito pequeno.

Contudo, o rápido crescimento e expansão da *Internet*, conduziu a um interesse crescente no desenvolvimento de novos serviços, e rapidamente foram criadas e identificadas oportunidades de negócio. Estas oportunidades dão origem ao aparecimento e desenvolvimento de novos serviços e aplicações multimédia com requisitos que a rede *Internet* não está preparada para sustentar. Muitos destes serviços e aplicações têm requisitos de tempo real, isto é, uma grande sensibilidade a oscilações na taxa de transmissão e a atrasos elevados e variáveis. Por isso, estas aplicações correm sobre o protocolo de transporte UDP (*User Datagram Protocol*) que é um protocolo muito mais simples que não suporta controle de erros, mas que permite ter atrasos na transmissão inferiores aos do TCP.

Nem todos os serviços e aplicações necessitam de uma elevada qualidade na transmissão. Além disso, aplicações diferentes necessitam de requisitos diferentes. Por exemplo, uma aplicação de transferência de ficheiros é tolerante aos atrasos de transmissão dos pacotes. No entanto, uma aplicação de telefonia pela *Internet* é intolerante a atrasos e principalmente a variações nos atrasos, mas é tolerante à existência de algumas perdas de pacotes desde que essas perdas estejam abaixo de um determinado limiar. Além dos requisitos dos serviços, também os requisitos dos clientes são diferentes. As empresas que realizam processos críticos para o seu funcionamento através da *Internet* estão dispostas a pagar mais por um serviço melhor. Neste sentido, existe uma necessidade crescente de os fornecedores de serviço implementarem mecanismos de diferenciação dos seus serviços, e de fornecerem múltiplos níveis de qualidade adequados às exigências dos seus diversificados clientes. Estas funcionalidades acrescidas dos fornecedores vêm tornar possível a atribuição e diferenciação de Qualidade de Serviço (QdS).

A QdS em redes de telecomunicações pode ser identificada por vários parâmetros, sendo os mais importantes o rácio de perdas, o atraso extremo-a-extremo e a variação do atraso (*jitter*) dos pacotes. Estes parâmetros, que se situam ao nível da transmissão, têm influência directa na percepção por parte do utilizador da QdS que está a ser oferecida a determinada aplicação. Um utilizador descontente com o serviço prestado é um potencial cliente perdido.

Para responder aos novos desafios do mundo das telecomunicações, o actual modelo da *Internet* tem de se adaptar de modo a fornecer a tão desejada QdS. Só desta forma será possível a utilização da *Internet* para os serviços emergentes como são a Voz sobre IP (*Voice over IP* – VoIP) ou a distribuição e visualização em tempo real de conteúdos multimédia, pois estes apresentam requisitos bastante superiores aos que a *Internet* pode hoje em dia fornecer.

A implementação de capacidades de QdS na rede *Internet* tem sido um desafio na evolução da própria rede e tem sido objecto de estudo e investigação nos últimos 10 anos.

A forma mais fácil de garantir que todos os utilizadores e aplicações têm acesso aos recursos de rede de que necessitam, em qualquer altura, com as necessárias características de QdS, é dimensionar a rede de modo a que esta tenha recursos disponíveis para uma situação de pior caso. Por situação de pior caso entende-se um cenário em que todos os utilizadores estão a usar em simultâneo todos os serviços a que têm direito com os requisitos necessários para cada serviço. Embora esta solução seja muito simples de implementar, não é de modo algum eficiente, pois a probabilidade de num dado instante todas as aplicações necessitarem da totalidade de recursos que contrataram é muito pequena e, deste modo, os recursos da rede não são rentabilizados.

Com o intuito de simultaneamente fornecer QdS e rentabilizar ao máximo os recursos da rede, esta tem de conter diferentes mecanismos de gestão de tráfego. Estes mecanismos incluem o controle de admissão de fluxos, o controle do nível de congestão da rede, e a ordenação dos pacotes para serem servidos em cada nó através de algoritmos de escalonamento. O controle de admissão limita o número de fluxos activos na rede para prevenir que esta entre em situação de congestão. Um fluxo é uma sequência de pacotes caracterizada por nó origem, nó destino e modelo de geração de tráfego ao nível do pacote. Os fluxos que estão sujeitos a controle de admissão só serão admitidos se a rede puder fornecer a QdS necessária (sem afectar a QdS dos restantes fluxos previamente admitidos).

O controle do nível de congestão da rede limita a taxa máxima à qual são enviados os pacotes pertencentes a um ou vários fluxos, e é responsável pelo descarte de alguns pacotes se a rede se encontrar em situações de congestionamento. O algoritmo de escalonamento tem como objectivo ordenar a transmissão dos pacotes em cada nó da rede.

As secções seguintes, 3.1 a 3.3, descrevem em detalhe estes mecanismos de suporte de QdS. Na secção 3.4 são descritas algumas arquitecturas para suporte de QdS que vão fazer uso dos mecanismos descritos anteriormente, as arquitecturas IntServ (Integração de Serviços) e DiffServ (Diferenciação de Serviços). São também apresentadas na mesma secção, algumas propostas recentes de arquitecturas e mecanismos desenvolvidos com o objectivo de resolver os problemas das arquitecturas IntServ e DiffServ.

3.1 Controle de admissão de fluxos

A função do controle de admissão é aceitar ou rejeitar pedidos de estabelecimento de fluxos, consoante os recursos disponíveis na rede, os requisitos de largura de banda do fluxo que pede admissão e os seus requisitos de QdS. Um fluxo é aceite apenas se se verificar que a rede o pode suportar com os seus requisitos de QdS sem prejudicar os fluxos já existentes. Existem vários tipos de algoritmos de controle de admissão. De acordo com [Knightly99], os algoritmos podem ser tipificados nas seguintes categorias:

- Algoritmos baseados nos valores de taxa média e de pico - Nestes algoritmos assume-se um sistema sem fila de espera e fontes de tráfego *on-off*. A distribuição da taxa agregada de chegada de pacotes das fontes multiplexadas é dada pela convolução da distribuição da taxa de chegada de cada uma das fontes. Esta distribuição é função do PBR (*Peak Bit Rate*) e do SBR (*Sustainable Bit Rate*) de cada fluxo previamente estabelecido e do fluxo que pede admissão. Existem perdas de pacotes no servidor se a taxa agregada de chegadas exceder a capacidade da ligação. Um fluxo é aceite se o valor da probabilidade de perda de pacotes for inferior a um limiar pré-definido em todas as ligações atravessadas pelo fluxo.
- Algoritmos baseados em larguras de banda efectivas - A largura de banda efectiva é a taxa à qual deve ser servido um fluxo, quando multiplexado com outros na mesma fila de espera, por forma a obter uma determinada QdS. A largura de banda efectiva é obtida através de cálculos probabilísticos a partir do modelo da fonte de tráfego, do rácio de pacotes perdidos especificado e do comprimento da fila de espera de

multiplexagem estatística. O seu valor situa-se entre os valores da taxa de pico e da taxa média do fluxo. As larguras de banda efectivas podem ser aditivas, em que a largura de banda efectiva de um fluxo é a mesma independentemente do número de fluxos multiplexados na fila de espera, ou não-aditivas, em que a largura de banda efectiva de um fluxo diminui com o aumento do número de fluxos multiplexados. A largura de banda efectiva é função da probabilidade de perda de pacotes. Um fluxo é aceite se a soma das larguras de banda efectivas dos fluxos já admitidos com o fluxo que pede admissão for igual ou inferior à capacidade das ligações atravessadas pelo fluxo.

? Algoritmos baseados nas diferentes regiões da curva de perdas – Verificou-se que existem duas zonas com declives diferentes na curva que relaciona o logaritmo da probabilidade de perdas de pacotes e o comprimento da fila de espera. Cada uma destas zonas é determinada por um modelo distinto. Na primeira zona, a probabilidade de perdas pode ser aproximada pelo modelo do histograma [Skelly93]; na segunda zona, a probabilidade de perdas pode ser aproximada por uma curva do tipo exponencial [Shroff98]. A curva de probabilidade de perdas de pacotes é função do número de fontes de tráfego e do seu modelo. Um fluxo é aceite se o valor da probabilidade de perdas determinada em todas as ligações atravessadas pelo fluxo for inferior a um determinado limiar.

? Algoritmos baseados na variância máxima da ocupação das filas de espera – Este método baseia-se na modelação da taxa de ocupação de uma fila de espera para um agregado de fontes de tráfego. Assume-se que a taxa de ocupação segue uma distribuição *Gaussiana*, tendo em vista o teorema do limite central. Para determinar a probabilidade de perdas de pacotes recorre-se à variância máxima da ocupação das filas de espera. Um fluxo é aceite se a probabilidade de perdas de pacotes em todas as ligações atravessadas pelo fluxo for inferior a um determinado limiar.

? Algoritmos baseados na observação do estado de congestão da rede – A observação do estado de congestão da rede permite determinar o nível de QdS que é possível oferecer a um fluxo. Estes algoritmos incluem mecanismos activos e passivos. No primeiro caso, cada nó faz uma medição da taxa agregada dos fluxos activos (*Measurement-Based Admission Control* - MBAC). Um fluxo é aceite numa ligação se a soma da largura de banda total medida com a largura de banda do novo fluxo se

encontra abaixo da capacidade da ligação; um fluxo é aceite na rede se for aceite em todas as ligações do seu percurso. No segundo caso, o emissor envia um fluxo de teste para o receptor, para determinar o atraso, rácio de perdas, ou rácio de marcação de pacotes do fluxo de teste. Os valores de atraso, rácio de perdas ou de marcação determinados correspondem ao atraso ou perdas que o fluxo sofreria caso fosse aceite. Se estes estiverem abaixo de um determinado limiar, o fluxo é aceite. Estes últimos mecanismos são denominados de mecanismos de *probing.* Estes mecanismos serão abordados na secção 3.4.4.4 e serão alvo de estudo nos capítulos 7 e 8.

3.2 Controle do nível de congestão da rede

A função de controle do nível de congestão da rede é limitar a quantidade de tráfego enviado na rede (depois de ter sido admitido), de modo a que o seu congestionamento não exceda um determinado limiar. As formas de efectuar o controle de congestionamento podem ser tipificadas, de acordo com [Keshav00], em tipos diferentes: em ciclo aberto, em ciclo fechado ou híbrido. Num esquema em ciclo aberto, cada fonte de tráfego é formatada, antes de entrar na rede, de acordo com os parâmetros acordados no estabelecimento da sessão. Neste caso, se existir um processo de controle de admissão correcto e se cada fonte de tráfego for formatada, não haverá congestão da rede. Num esquema em ciclo fechado, a taxa de transmissão de uma fonte ou a quantidade de informação enviada podem ser actualizadas dinamicamente de acordo com indicações enviadas pelos nós da rede ou pelo receptor, de uma forma implícita ou explícita. Um esquema híbrido pode conter as duas formas de controlar a congestão.

O esquema em ciclo aberto pode ser implementado por um *leaky bucket*. O *leaky bucket* é um regulador de tráfego que o formata por forma a garantir que a taxa média e o tamanho máximo da rajada não excedem b e p, respectivamente. Deste modo, o *leaky bucket* assegura que o tráfego que entra num dado nó segue um perfil previamente contratado entre o utilizador e o fornecedor do serviço (este assunto vai ser detalhado na secção 3.4.2). O *leaky bucket* é constituído por um reservatório de senhas (*tokens*) e por um servidor de pacotes (Figura 3-1). O reservatório é uma fila de espera com capacidade finita p. A taxa do servidor é definida pela taxa de chegada b das senhas ao reservatório. Cada pacote necessita, para ser servido, de um número de senhas proporcional ao seu comprimento. Se o reservatório tiver senhas suficientes para enviar o pacote, ele é enviado

de imediato e o número de senhas é decrementado. Se não houver senhas suficientes, o pacote é armazenado numa fila de espera até que o número de senhas seja suficiente para enviar o pacote. Deste modo, a quantidade de tráfego de cada fluxo à saída do regulador será limitada em qualquer intervalo de duração *t* por *b*?*t*+*p*.

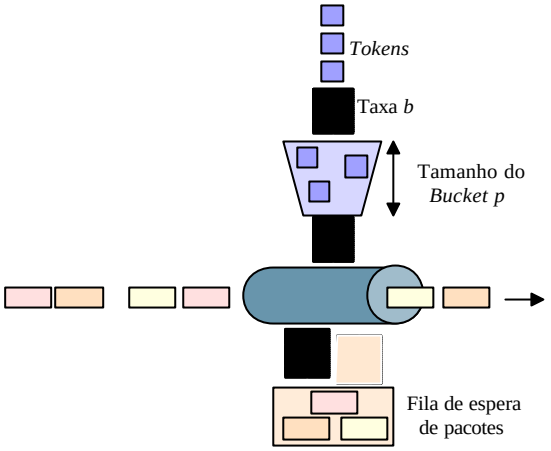


Figura 3-1 : ***Leaky Bucket*** **e seus parâmetros.**

O esquema em ciclo fechado pode ser tipificado em três formas complementares:

? Implícito ou explícito – No primeiro caso, os emissores utilizam medidas de desempenho para inferir dinamicamente o estado de congestão da rede ou a largura de banda disponível; no segundo caso, os elementos da rede usam mensagens de controle para comunicar, a todos os emissores, o seu estado de congestão ou a largura de banda disponível.

? Janela dinâmica ou taxa dinâmica - A janela de controle define o número de pacotes que podem ser enviados pelo emissor, e a taxa define a taxa de envio de pacotes por parte do emissor. O tamanho da janela e a taxa de transmissão são actualizados dinamicamente de acordo com as indicações de congestionamento ou largura de banda disponível recebidas no emissor.

? Nó a nó ou extremo-a-extremo – As indicações de congestionamento ou de largura de banda disponível podem ser enviadas para o emissor por cada nó da rede ou apenas pelo receptor da informação.

Um dos exemplos de controle de congestão em ciclo fechado é o mecanismo de controle de congestão do protocolo TCP. Este controle é implícito, extremo-a-extremo e efectuado com base numa janela dinâmica. Neste mecanismo, a fonte de tráfego ajusta a

sua janela de controle em resposta a sinais implícitos de congestionamento da rede. Mais especificamente, uma fonte aumenta a sua janela até detectar perda de pacotes. Neste ponto, a fonte reduz o tamanho da sua janela. Se as perdas de pacotes deixarem de acontecer, a fonte aumenta novamente a janela de controle, e assim sucessivamente. Numa situação de congestão da rede, se um grande número de fontes activas sofrer perdas, todas elas reduzirão a sua taxa de transmissão, resultando num mau aproveitamento dos recursos. Para prevenir as situações de congestionamento são introduzidos mecanismos de descarte aleatório de pacotes de acordo com a ocupação das filas de espera dos nós, ou seja, antes de a rede entrar em situação de congestão: RED (*Random Early Detection*) [Floyd93] ou RIO (*Red with In and Out*) [Clark98]. Estes mecanismos procedem à monitorização do estado de ocupação das filas de espera (Figura 3-2). O RED mantém dois limiares de ocupação: limiar mínimo e máximo. Quando o comprimento médio da fila de espera é inferior ao limiar mínimo, nenhum pacote é descartado. Quando o comprimento excede o limiar mínimo, os pacotes que chegam à fila de espera são descartados com uma determinada probabilidade que aumenta linearmente com o comprimento médio da fila de espera. Quando o comprimento excede o limiar máximo, todos os pacotes são descartados.

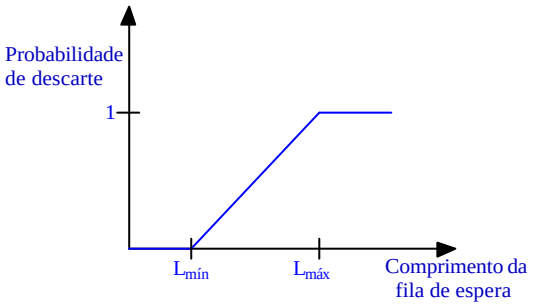


Figura 3-2 : Funcionamento do RED.

O RIO é semelhante ao RED mas distingue pacotes que estão dentro do perfil de tráfego acordado e os pacotes que estão fora do perfil. Cada um destes tipos de pacotes tem associado um limiar mínimo e máximo. Os limiares associados aos pacotes fora do perfil são inferiores aos associados aos pacotes dentro do perfil. Deste modo, os pacotes fora do perfil são descartados primeiro do que os pacotes dentro do perfil. O RED pode também ser utilizado para distinguir entre várias classes de serviço através da associação de limiares diferentes a classes diferentes. Nestes mecanismos, a congestão é detectada de

forma implícita através da perda de pacotes. No entanto, eles podem também ser utilizados para detectar a congestão de uma forma explícita através de marcação de pacotes.

3.3 Algoritmos de escalonamento

O algoritmo de escalonamento decide qual o próximo pacote que será servido na fila de espera. Este algoritmo é um dos mecanismos responsáveis por distribuir a largura de banda da ligação pelos diferentes fluxos (atribuindo a cada fluxo a largura de banda que foi pedida pelo utilizador e aceite pela rede).

Um algoritmo de escalonamento pode ser do tipo *work-conserving* ou *non-work-conserving*. No primeiro caso, o servidor "trabalha" sempre, isto é, havendo pacotes em espera, eles serão sempre transmitidos. No segundo caso, um nó só pode transmitir um pacote quando este se torna *elegível*, isto é, quando o tempo necessário para ele se manter em espera termina. Se no nó apenas se encontrarem pacotes não elegíveis em espera, então o servidor manter-se-á inactivo. Este tipo de algoritmos de escalonamento destinam-se a aplicações que não toleram variações no atraso de transmissão. A desvantagem óbvia destes algoritmos é o desperdício de largura de banda durante os períodos em que apenas existem pacotes não *elegíveis* em espera.

A classificação dos algoritmos de escalonamento pode ser efectuada de acordo com os princípios que definem a ordem de envio dos pacotes [Santiago02]: por ordem de chegada, de uma forma estrita, de uma forma rotativa, por aproximação ao sistema de fluídos, e em tempos pré-definidos.

3.3.1 Ordem de chegada

O algoritmo de escalonamento que serve os pacotes por ordem de chegada é o FIFO (*First In First Out*). Este algoritmo é muito simples de implementar mas não permite diferenciação de QdS, assim como não garante a existência de limites máximos nos atrasos. Os fluxos de tráfego que recebem um serviço melhor são os que geram mais tráfego.

3.3.2 Forma estrita

Neste algoritmo, o escalonador é constituído por várias filas de espera, cada uma com uma prioridade diferente. Os fluxos são classificados em diferentes níveis de prioridade e

associados a uma determinada fila de espera de acordo com a sua prioridade. Este algoritmo serve o tráfego por ordem de prioridade. O tráfego de prioridade mais elevada é servido sempre antes do tráfego com prioridade inferior. Este algoritmo é simples de implementar e permite diferenciação de QdS. No entanto, se não existir nenhum mecanismo de controle de admissão dos fluxos com maior prioridade, uma grande quantidade de pacotes de elevada prioridade pode impedir completamente o serviço de pacotes com menor prioridade (este fenómeno é usualmente denominado de *starvation*). Este mecanismo de prioridade simples deve ser usado apenas para tráfego que exija garantias muito estritas de QdS.

3.3.3 Forma rotativa

Nestes algoritmos, o escalonador é constituído por várias filas de espera com a mesma prioridade e os fluxos são associados a uma determinada fila de espera de acordo com a sua classificação. Estes algoritmos seleccionam o tráfego de uma forma rotativa. No algoritmo mais simples, *Round Robin* (RR), o sistema selecciona um pacote de cada fila de espera de uma forma rotativa. Este algoritmo é também muito simples mas favorece os fluxos que contêm pacotes com maior comprimento, pois o pacote é servido independentemente do seu comprimento. Existem algumas variantes deste sistema que permitem utilizar pacotes de tamanho variável sem prejudicar os mais pequenos, e ainda, atribuir uma largura de banda pesada a cada fila de espera. De entre esses mecanismos destacam-se o *Weighted Round Robin* (WRR) e o *Deficit Round Robin* (DRR) [Shreedhar95].

3.3.4 Aproximação do sistema de fluídos

Este conjunto de algoritmos pressupõe também a existência de escalonadores com várias filas de espera, e distribui a largura de banda pelas diversas filas de espera de uma forma pesada ou de uma forma equitativa (*Fair Queuing* [Demers95]) mas sem a noção de rotatividade. Estes algoritmos tentam emular, num sistema de pacotes, o sistema de fluídos denominado de *Generalized Processor Sharing* (GPS) [Parekh92a, Parekh91b].

3.3.4.1 GPS

No GPS o tráfego é considerado como sendo infinitamente divisível. Esta característica permite considerar que as várias filas de espera são servidas simultaneamente. Cada fila de

espera utiliza, em cada instante, uma percentagem da capacidade da ligação que é proporcional ao peso que lhe é atribuído. Considere-se que ? é o conjunto de filas de espera, e ?' ? ?? é o conjunto das filas de espera não vazias durante o intervalo de tempo (?,*t*). Assume-se que a taxa de serviço é *C*, e que a cada fila de espera *i* é associado um peso ?*i*. Numa disciplina GPS, em qualquer intervalo (?,*t*) a fila de espera *i* recebe uma taxa mínima garantida proporcional ao seu peso dada por $C?_i / ?_{j? \in ?} ?_j$. Além disso, a uma fila *i* que tenha continuamente tráfego para transmitir é garantida uma taxa mínima proporcional ao seu peso, dada por

$$c_i ? \frac{?_i}{?_{j? \in ?} ?_j} C \qquad (3\text{-}1)$$

Esta taxa mínima será designada, ao longo da Tese, por largura de banda mínima garantida.

3.3.4.2 WFQ

O algoritmo WFQ (*Weighted Fair Queuing*) [Parekh92b], também denominado de *Packet Generalized Processor Sharing* (PGPS) tenta emular, numa rede de pacotes, o sistema GPS. No WFQ, um pacote é servido pela ordem em que terminaria serviço no correspondente sistema GPS.

No WFQ é necessário calcular o instante em que o pacote deixaria o servidor num sistema GPS. Os pacotes vão posteriormente ser servidos por ordem destes instantes de partida. No caso geral, sendo F_i^k o instante de partida do pacote *k* pertencente à fila de espera *i*,

$$F_i^k ? \max(F_i^{k?1}, V(t)) ? \frac{l_i^k}{?_i} \qquad (3\text{-}2)$$

em que l_i^k é o comprimento do pacote *k* pertencente à fila de espera *i* e *V(t)* é o chamado *round number* descrito em [Parekh92b]. O WFQ permite obter um limite máximo para o atraso de cada pacote de um fluxo quando utilizado em conjugação com um formatador de tráfego do tipo *leaky bucket*. O atraso máximo dos pacotes de um fluxo *i*, cujo tráfego de entrada no primeiro nó é limitado por um *leaky bucket* com parâmetros (b_i, p_i), tal que a largura de banda mínima garantida c_i é igual ou superior a b_i, é dado por

$$\frac{p_i}{c_i} + \frac{(a-1)l_{max}}{c_i} + \sum_{j=1}^{a} \frac{l_{max}}{C_j} \quad (3\text{-}3)$$

em que a é o número de nós por onde passa o pacote, l_{max} é o comprimento máximo do pacote, e C_j é a taxa de serviço no nó j.

O cálculo da função $V(t)$ é bastante complexo, sendo difícil a sua implementação em *routers* de elevado débito. Existem outros algoritmos de escalonamento que são variantes do WFQ, os quais são menos complexos e utilizados na maior parte das implementações. De entre os algoritmos destacam-se o *Self-Clocking Fair Queuing* (SCFQ) [Golestani94] e o *Virtual Clock* (VC) [Zhang90].

3.3.4.3 SCFQ

O SCFQ propõe uma aproximação simples para calcular o instante de partida no correspondente sistema GPS. Quando um pacote chega a uma fila de espera vazia, o SCFQ usa como $V(t)$ o instante de partida do pacote que está nesse momento em serviço (F_{actual}). O instante de partida é então dado por

$$F_i^k = \max\left(F_i^{k-1}, F_{actual}\right) + \frac{l_i^k}{\phi_i} \quad (3\text{-}4)$$

Este algoritmo é de fácil implementação em redes de elevado débito, mas pode em alguns casos específicos, não ser justo em pequenos intervalos de tempo.

3.3.4.4 VC

O VC tenta emular um sistema TDM (*Time Division Multiplexing*). No VC, a função $V(t)$ é substituída pelo instante de chegada do pacote, se a fila de espera se encontrar vazia. A reserva de uma taxa de serviço para cada fila de espera é efectuada por um valor E_i, acordado previamente entre as fontes de tráfego e o servidor, que representa o valor esperado para os intervalos entre chegadas à fila de espera i. O servidor atende as filas de espera por ordem do *tempo virtual* baseado no parâmetro E_i. O *tempo virtual* do pacote k na fila de espera i é dado pela seguinte expressão:

$$VT_i^k = \max\left(VT_i^{k-1}, Tempo\right) + E_i \quad (3\text{-}5)$$

em que *Tempo* é o instante de chegada do pacote k à fila de espera i.

A implementação deste algoritmo é bastante simples. Além disso, em [Figueira95] prova-se que uma rede com servidores VC pode oferecer as mesmas garantias de atraso a

um fluxo que uma rede com servidores WFQ. No entanto, em determinadas situações, o VC pode não ser uma disciplina justa: o VC tem o problema de punir os fluxos que, durante algum tempo, são servidos a uma taxa superior à taxa previamente acordada (definida pelo parâmetro E_i).

3.3.5 Tempos pré-definidos

Este conjunto de algoritmos é do tipo *non-work conserving*. A cada pacote é associado um tempo elegível e um tempo limite. Os pacotes só podem ser servidos depois de o instante de tempo actual exceder o tempo elegível do pacote. Deste modo, garante-se que todos os pacotes de um mesmo fluxo têm aproximadamente o mesmo atraso. O escalonador ordena os pacotes elegíveis para transmitir com base nos seus tempos limite. O tempo limite é o instante de tempo em que o pacote deve ser servido. Os algoritmos de escalonamento apresentados nesta secção são o *Jitter-Virtual Clock* (JVC) [Verma91, Zhang94, Zhang90] e o *Core-Jitter-Virtual Clock* (CJVC) [Stoica99]. O CJVC é utilizado no âmbito de uma arquitectura que será apresentada na secção 3.4.4.1.

3.3.5.1 JVC

No JVC, o tempo elegível do pacote *k* da fila de espera *i* no nó *j* é dado por:

$$TE_{i,j}^{1} = TC_{i,j}^{1} \tag{3-6}$$

$$TE_{i,j}^{k} = \max\left(TC_{i,j}^{k} + TSL_{i,j-1}^{k}, TL_{i,j}^{k-1}\right) \tag{3-7}$$

em que $TC_{i,j}^{k}$ é o tempo de chegada do pacote ao nó *j*, $TSL_{i,j-1}^{k}$ é a diferença entre o tempo limite e o instante em que o pacote foi transmitido no nó anterior, e $TL_{i,j}^{k-1}$ é o tempo limite do pacote anterior. Note-se que o parâmetro $TSL_{i,j-1}^{k}$ não é do conhecimento do nó *j*, e por isso este parâmetro tem de ser marcado no pacote *k*.

O tempo limite do pacote no nó *j* é dado por

$$TL_{i,j}^{k} = TE_{i,j}^{k} + \frac{l_i^k}{c_i} \tag{3-8}$$

Este algoritmo tenta minorar a variação no atraso dos pacotes. Se um pacote for transmitido muito tempo antes do tempo limite num determinado nó, ele vai ter de esperar algum tempo no nó seguinte para compensar esse facto. Em [Georgiadis96] prova-se que

uma rede com JVC e *leaky bucket* fornece as mesmas garantias de atraso que uma rede com VC, por isso o JVC fornece as mesmas garantias que o WFQ.

3.3.5.2 CJVC

No JVC, o tempo elegível e o tempo limite de um pacote dependem do tempo limite do pacote anterior na mesma fila de espera, $TL_{i,j}^{k?1}$. A dependência dos tempos limites de pacotes anteriores na mesma fila de espera requer um cálculo iterativo deste parâmetro em cada nó. Deste modo, os nós ao longo do percurso dos pacotes têm de armazenar a informação referente a pacotes anteriores. A disciplina de escalonamento CVJC foi proposta para eliminar da disciplina JVC, o parâmetro $TL_{i,j}^{k?1}$ no sentido de tornar a disciplina de escalonamento mais fácil de implementar.

Pela observação da equação 3-7 verifica-se que o parâmetro $TL_{i,j}^{k?1}$ é apenas usado numa operação de cálculo de um máximo. Assim, é possível eliminar a necessidade de existência do $TL_{i,j}^{k?1}$ se se garantir que o outro termo do máximo é sempre superior a $TL_{i,j}^{k?1}$. A ideia deste algoritmo é associar uma variável *slack* $?_i^k$ a cada pacote que não varia ao longo do percurso, tal que para qualquer nó *j* ao longo do percurso, a seguinte inequação é sempre verdadeira:

$$TC_{i,j}^k ? TSL_{i,j?1}^k ? ?_i^k ? TL_{i,j}^{k?1} \quad (3\text{-}9)$$

Deste modo, o cálculo do tempo elegível reduz-se à seguinte expressão:

$$TE_{i,j}^k ? TC_{i,j}^k ? TSL_{i,j?1}^k ? ?_i^k \quad (3\text{-}10)$$

Construído deste modo, este algoritmo não necessita do armazenamento do estado do fluxo em cada nó ao longo do percurso. Além disso, calculando o valor de $?_i^k$ através da fórmula derivada em [Stoica99], o atraso máximo garantido por este algoritmo é o mesmo que o do JVC.

3.4 Arquitecturas para a *Internet* com suporte de QdS

Nas secções anteriores foram apresentados os mecanismos fundamentais que permitem adicionar à rede *Internet* capacidades de fornecer e diferenciar a QdS. Nesta secção vão ser descritas algumas arquitecturas para suporte de QdS que vão fazer uso dos mecanismos

descritos anteriormente. Muitas destas arquitecturas implicam mudanças na arquitectura básica da *Internet*. O *Internet Engineering Task Force* (IETF) [IETF-int] definiu duas novas arquitecturas de diferenciação de QdS: a arquitectura IntServ e DiffServ. Estas arquitecturas propõem novos modelos de serviço, para além do modelo de melhor esforço, que possibilitam a existência de garantias de recursos mínimos e a diferenciação de serviços para diferentes aplicações ou utilizadores. A arquitectura IntServ fornece garantias de recursos mínimos através da reserva de recursos para cada fluxo de tráfego relacionado com uma determinada aplicação e utilizador. A arquitectura DiffServ usa uma combinação de policiamento de tráfego na fronteira da rede, de provisão de largura de banda (atribuição de largura de banda na rede superior à do tráfego esperado), e de mecanismos de diferenciação entre classes de serviço.

3.4.1 Integração de serviços

A arquitectura IntServ [Braden94] foi desenvolvida com o objectivo de obter modelos de serviço superiores ao actual modelo de melhor esforço, através da reserva de recursos para cada fluxo de tráfego de acordo com as suas características. Com IntServ é possível manter o modelo de datagrama usado nas redes baseadas em IP e, simultaneamente, suportar aplicações em tempo real através da reserva de recursos e da diferenciação do tráfego de cada utilizador e/ou serviço.

Nestas arquitecturas presume-se que o processo de reserva de recursos é efectuado após se ter definido o percurso do fluxo, em que este é definido por outros protocolos. A reserva de recursos para um fluxo de tráfego consiste em vários passos: (1) a aplicação especifica o fluxo, isto é, indica quais as características de tráfego e os requisitos de QdS do fluxo; (2) o pedido de reserva é enviado para a rede; (3) cada *router* ao receber o pedido decide, através de um algoritmo de controle de admissão, se existem recursos suficientes para aceitar este fluxo com as características pretendidas. Quando a reserva é efectuada e bem sucedida, a informação acerca do fluxo e do seu estado de reserva é colocada numa tabela de reserva de recursos existente nos *routers*. Esta informação vai ser posteriormente utilizada quando o *router* receber pacotes correspondentes a este fluxo de tráfego. Os pacotes são identificados e colocados na respectiva fila de espera, e o escalonador de pacotes atribui recursos para os diferentes fluxos de acordo com a informação de reserva de cada um deles.

Para criar uma reserva de recursos para um fluxo em cada elemento da rede ao longo do percurso definido, é necessário um protocolo de estabelecimento de reservas para instalar o estado de reserva do fluxo. Este protocolo distribui a informação sobre as características do fluxo e os seus requisitos de QdS em cada nó ao longo do percurso, para que possa ser determinado em cada elemento da rede se o novo pedido de reserva pode ou não ser aceite. O protocolo de sinalização utilizado na arquitectura IntServ é o RSVP (*resource ReSerVation Protocol*) [Zhang93]. O RSVP inclui mecanismos que permitem fazer o "refrescamento" da reserva. Deste modo, o RSVP adapta-se automaticamente a alterações nas rotas definidas.

Antes de uma reserva ser aceite, esta tem de passar o teste do controle de admissão em cada nó da rede. As reservas só podem ser aceites se existirem recursos disponíveis em cada ligação. O controle de admissão pode ser baseado nos parâmetros dos fluxos, por exemplo nos parâmetros do *leaky bucket*, ou baseado em medidas da carga de tráfego da rede. A primeira hipótese torna-se difícil de implementar quando o tráfego não tem um modelo bem definido e apresenta grandes variações ao longo do tempo. A segunda hipótese é mais pesada em termos de carga computacional, pois é necessário, repetidamente, monitorizar a quantidade de tráfego em cada ligação.

Para além de manter o processo de reserva de recursos cada nó, ao receber um novo pacote, necessita de, em tempo real, identificar o fluxo a que este pertence, e consequentemente, decidir para que fila de espera o deve encaminhar. No contexto IntServ, a identificação é efectuada através de 5 campos do cabeçalho do pacote IP: endereços IP da origem e destino, identificação do protocolo, e portos de origem e destino. Depois de se identificar a que fluxo pertencem os pacotes, o escalonador decide a ordem pela qual eles são servidos.

A interacção entre estas funções implementadas nos nós de uma rede IntServ é ilustrada na Figura 3-3.

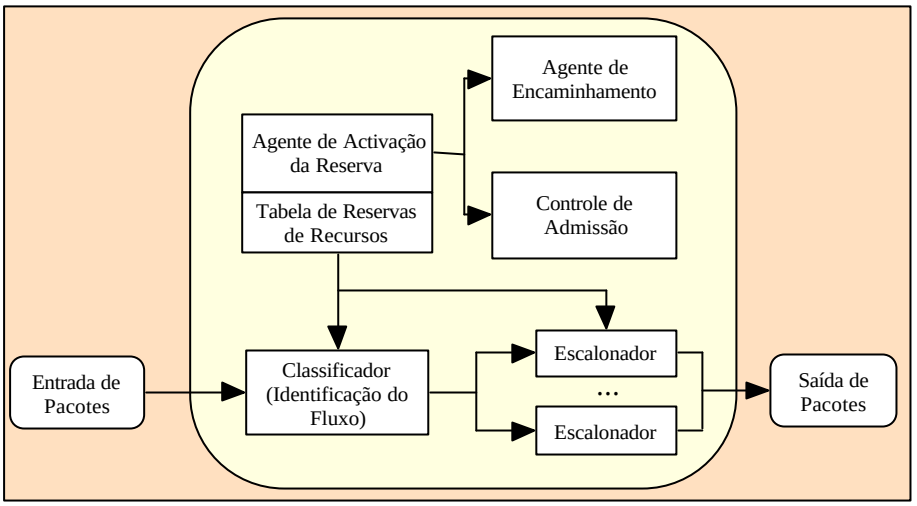


Figura 3-3 : Modelo de referência IntServ.

3.4.1.1 Modelos de serviço

Os modelos de serviço descrevem, por um lado, os serviços que os utilizadores podem utilizar na rede e, por outro lado, os compromissos de garantia de recursos que a rede pode oferecer. Na arquitectura IntServ, além do serviço de melhor esforço, existem dois outros modelos de serviço: serviço garantido e serviço de carga controlada. Antes da apresentação destes serviços, serão primeiro descritos os parâmetros utilizados na especificação dos fluxos, que incluem a caracterização do tráfego e dos requisitos de QdS. Os parâmetros que definem o fluxo de tráfego serão designados de parâmetros de descrição do tráfego; os parâmetros que definem os requisitos de QdS serão designados de parâmetros de especificação do serviço. Estes parâmetros serão detalhados na secção seguinte.

3.4.1.1.1 Especificação dos fluxos

A especificação dos fluxos é como que um contrato de serviço entre o utilizador e a rede, no qual se descreve o fluxo de tráfego que será enviado e os serviços e recursos que a rede se compromete a fornecer. Se o utilizador, por exemplo, envia mais tráfego do que o contratado, a rede não poderá cumprir o nível de QdS esperado. Normalmente, o tráfego é modificado antes de entrar na rede para garantir que está de acordo com o perfil definido. Na arquitectura IntServ, o tráfego é descrito através dos parâmetros (b, p) do *leaky bucket*.

Para uma aplicação pedir um serviço tem de especificar à rede os parâmetros de descrição do tráfego e de especificação do serviço.

Os parâmetros de descrição do tráfego incluem os parâmetros do *leaky bucket*, a taxa de pico, e os comprimentos máximo e mínimo dos pacotes.

Os parâmetros de especificação do serviço dependem do tipo de serviço, da sua sensibilidade a atrasos e a perda de pacotes. Os parâmetros normalmente utilizados são a largura de banda mínima (que será garantida por algoritmos de escalonamento de pacotes, por exemplo, o WFQ), o atraso (que pode ser especificado como atraso médio ou atraso máximo), variações no atraso, e o rácio de perdas de pacotes.

3.4.1.1.2 Serviço garantido

O serviço garantido [Shenker97] é um serviço que fornece garantias estritas de largura de banda e limites bem definidos para os atrasos extremo-a-extremo nas filas de espera. É a classe de serviço que pode oferecer melhores garantias de qualidade de serviço, devendo ser usada em aplicações que requeiram elevadas garantias ao nível de largura de banda e atrasos. O comportamento extremo-a-extremo de um percurso onde são transportados serviços garantidos pode ser equiparado a um circuito virtual com largura de banda garantida.

Os fluxos de tráfego devem estar de acordo com os parâmetros do *leaky bucket* em todos os períodos de tempo em que o fluxo está activo. Se não estiverem de acordo, os seus pacotes são policiados e reformatados antes de entrarem na rede. Os pacotes que não estiverem de acordo com o perfil de tráfego são tratados como pacotes do serviço de melhor esforço, e podem também ser marcados com uma prioridade elevada de descarte em caso de congestão da rede. Numa arquitectura IntServ, o policiamento é feito à entrada de cada domínio.

3.4.1.1.3 Serviço de carga controlada

Para algumas aplicações menos exigentes, um modelo de serviço garantido não é apropriado porque, além de dar garantias estritas não necessárias, diminui em muito a utilização da rede pelo facto de os recursos terem de ser reservados para o pior caso. Para essas aplicações menos exigentes, é preferível um modelo de serviço com menores garantias e de mais baixo custo. O modelo de serviço de carga controlada [Wroclaw97]

tem um comportamento semelhante ao de um serviço de melhor esforço numa rede que não se encontra congestionada, ou em que o nível de congestionamento é muito baixo.

Este modelo de serviço possibilita a multiplexagem estatística entre os diversos fluxos de tráfego. Assim, a determinação dos recursos que se encontram disponíveis para decidir a aceitação ou rejeição de um novo fluxo pode ser feita, por exemplo, em função da taxa máxima agregada medida em intervalos de tempo anteriores, relativa a todos os fluxos pertencentes a este modelo de serviço. Deste modo, como o controle de admissão é realizado com base em medidas efectuadas em intervalos de tempo anteriores, este serviço permite a existência de aumentos pontuais nas perdas e atrasos do tráfego processado de acordo com esse modelo. No entanto, num serviço de carga controlada, a probabilidade de existência destes eventos deve ser baixa. Também neste modelo, o tráfego fora do perfil é tratado como pertencente ao modelo de melhor esforço.

3.4.1.2 RSVP

Ao contrário do modelo de melhor esforço, nos modelos de serviço garantido e de carga controlada definidos pela arquitectura IntServ, antes de uma aplicação começar a transmitir tráfego para a rede, tem de iniciar um pedido de reserva de recursos e obter uma resposta positiva por parte de todos os elementos da rede. O protocolo de reserva de recursos desenvolvido pelo IETF e usado pela arquitectura IntServ é o RSVP [Zhang93, Braden97]. Os utilizadores usam o protocolo para comunicar à rede as características de tráfego e os requisitos de QdS dos serviços, e os nós da rede usam-no para estabelecer o estado de reserva ao longo do percurso dos fluxos.

As reservas efectuadas são unidireccionais, sendo por isso necessário, numa comunicação bidireccional, que os dois extremos da comunicação estabeleçam a reserva nas duas direcções. Os estados de reserva de cada fluxo presentes nos nós da rede ao longo de um percurso são designados de *soft state*, devido ao facto de estes estados terem um tempo de vida associado. Se não forem enviadas mensagens de refrescamento (*refresh*) periodicamente, o estado do fluxo é automaticamente apagado. Deste modo, o RSVP adapta-se facilmente a alterações das rotas definidas.

3.4.1.2.1 Mensagens RSVP

O RSVP tem vários tipos de mensagens: PATH, RESV, PATHErr, RESVErr, PATHTear e RESVTear. A mensagem PATH instala o “estado do percurso” em cada nó do percurso.

Este "estado do percurso" inclui, pelo menos, o endereço IP do nó anterior, que será utilizado para encaminhar a mensagem RESV no sentido oposto. A mensagem PATH contém também a descrição do tráfego que será gerado pela origem. A mensagem RESV é uma mensagem de pedido de recursos que inclui informação sobre o fluxo de tráfego e QdS especificada. A mensagem PATHErr/RESVErr é uma mensagem enviada quando se encontra uma situação de erro no processamento da mensagem. A mensagem PATHTear/RESVTear remove os estados de encaminhamento e de reserva previamente estabelecidos nos nós.

3.4.1.2.2 Modo de operação

O RSVP funciona da seguinte forma (Figura 3-4): (1) o utilizador origem envia uma mensagem PATH ao utilizador destino com as características de tráfego e os requisitos de QdS, e cada nó ao longo do percurso retransmite a mensagem para o nó que se lhe segue em direcção ao destino. A mensagem PATH é marcada com o endereço de cada nó que atravessa, para que este seja armazenado no nó seguinte. (2) Após receber a mensagem PATH, o receptor envia uma mensagem RESV de pedido de reserva de recursos para o fluxo. Cada nó ao longo do percurso pode aceitar ou rejeitar o pedido consoante existam ou não recursos disponíveis. Se o pedido é rejeitado num nó, este envia uma mensagem de erro ao receptor e o processo de sinalização é terminado. Se a reserva é aceite, é reservada largura de banda para o fluxo. Após receber uma mensagem RESV com sucesso, a origem pode iniciar a transmissão de pacotes ao longo do percurso que contém a largura de banda reservada.

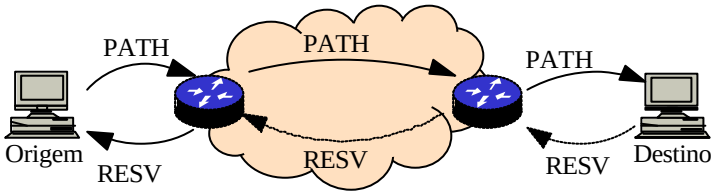


Figura 3-4 : Mensagens RSVP.

3.4.1.3 Análise da arquitectura IntServ

A arquitectura IntServ permite transformar a rede *Internet* actual numa rede com mecanismos de suporte de QdS e de diferenciação do tratamento associado a cada serviço. No entanto, esta arquitectura apresenta um conjunto de limitações:

? O processo de criação de uma reserva para cada fluxo é um processo moroso que apenas faz sentido em sessões cuja duração seja relativamente elevada. As aplicações baseadas na WWW (*World Wide Web*) são dominantes hoje em dia, e a maior parte do tráfego *Web* é resultante de sessões de pequena duração.

? Cada nó da rede tem de manter armazenado o estado de cada fluxo. No caso de redes de núcleo, o número de fluxos pode ser muito elevado e as tabelas de armazenamento tornam-se demasiado grandes para que cada nó as possa suportar.

? Cada nó da rede tem de implementar classificação por fluxo, através da inspecção dos cinco elementos do cabeçalho IP, e da procura desses cinco elementos na tabela de armazenamento. Além disso, o escalonamento de pacotes é também efectuado por fluxo. Mais uma vez, com milhares de fluxos activos em cada nó, pode ser complicado efectuar estas operações a elevados débitos e em tempo real.

Por estas razões, a arquitectura IntServ é de muito difícil implementação em redes de núcleo. Nas redes de acesso os problemas são atenuados porque o número de fluxos activos em cada nó pode ser muito inferior.

No sentido de fornecer suporte de QdS às redes de núcleo, o IETF definiu uma nova arquitectura, a arquitectura DiffServ que limita as funcionalidades de QdS a mecanismos de tratamento diferenciado de classes de serviço em função de acordos de nível de serviço (*Service Level Agreement* - SLA). Deste modo, esta arquitectura é de fácil implementação tanto em redes de acesso como de núcleo.

3.4.2 Diferenciação de serviços

A arquitectura DiffServ [Carlson98, Nichols97] utiliza um conjunto de blocos simples que permitem processar conjuntos de agregados de tráfego com uma determinada QdS, denominados de classes de serviço. Os fluxos são agregados num número limitado de classes, e é dado o mesmo tratamento a tráfego pertencente a uma mesma classe. A atribuição de recursos é efectuada por classe. Numa rede DiffServ, os nós fronteira têm responsabilidades diferentes das dos nós do núcleo (Figura 3-5). Os nós fronteira têm as

tarefas de classificação dos pacotes e de condicionamento do tráfego. Os nós do núcleo necessitam apenas de classificar os pacotes com base nas classes de serviço mapeadas no cabeçalho dos pacotes.

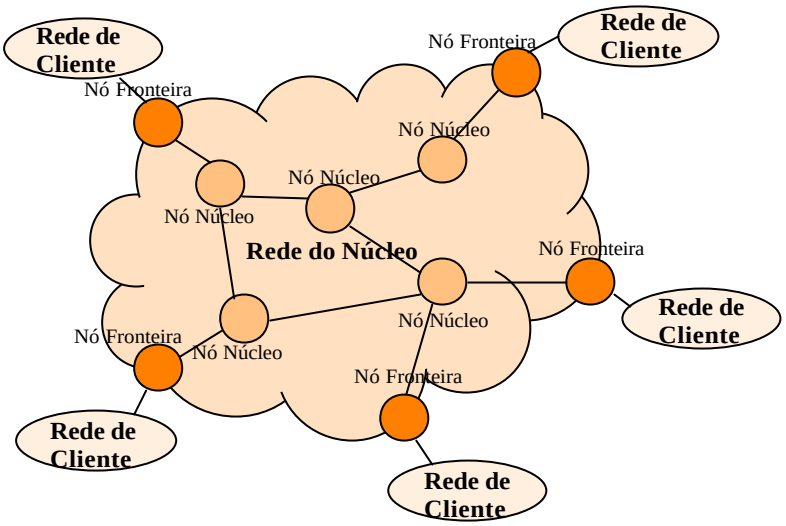


Figura 3-5 : Arquitectura geral de uma rede DiffServ.

A garantia de recursos para uma determinada classe é efectuada através da atribuição de largura de banda na rede suficiente para o tráfego esperado nessa classe e pelo tratamento diferenciado de cada classe. Deste modo, são criados níveis de serviço e garantias de recursos mínimos para cada classe.

Nesta arquitectura os clientes mantêm acordos de nível de serviço (SLAs) com os fornecedores de serviço (*Internet Service Provider* - ISP). Este acordo inclui as características dos serviços utilizados pelos clientes, tais como as características do tráfego e os seus requisitos de QdS, e o custo dos serviços. Desde que um cliente cumpra o acordo, o fornecedor de serviço também cumpre a sua parte do acordo, atribuindo a QdS a que o utilizador tem direito. Se um cliente não cumpre o acordo, este pode ser imposto pelos nós fronteira, através do policiamento de tráfego (reformatação e descarte de pacotes). As classes de serviço são definidas para um domínio e não extremo-a-extremo. Entre domínios existem acordos bilaterais para mapear as definições de classes.

3.4.2.1 Serviços

Um serviço descreve o tratamento que o tráfego de um cliente deverá ter num domínio DiffServ ou extremo-a-extremo. Os serviços são então definidos pelos SLAs entre os clientes e os ISPs. Um elemento importante do SLA é o acordo de condicionamento de tráfego (*Traffic Conditioning Agreement* - TCA). O TCA inclui os perfis de tráfego descritos, por exemplo, através de parâmetros do *leaky bucket*, as métricas de desempenho, tais como a taxa mínima de serviço ou o atraso, acções para tráfego fora do perfil de tráfego acordado, e marcação e reformatação adicional efectuada pelo ISP.

3.4.2.1.1 Comportamento em cada nó (*Per Hop Behavior* – PHB)

Numa arquitectura DiffServ, o tratamento que é dado num nó a um agregado de tráfego pertencente a uma classe de serviço é designado por PHB. Todos os pacotes que pertencem a fluxos com o mesmo PHB são denominados de BA (*Behavior Aggregate*). Cada PHB é representado por 6 *bits* no cabeçalho IP designados por DiffServ *codepoint* (DSCP). O DSCP é colocado no campo de DS (DiffServ) [Nichols98] do cabeçalho IP.

O IETF normalizou dois grupos de PHB para a arquitectura DiffServ: Envio Garantido (*Assured Forwarding* – AF) e Envio Expedito (*Expedited Forwarding* – EF).

3.4.2.1.2 Expedited Forwarding

O grupo de PHB EF [Jacobson99] é caracterizado por assegurar atrasos e perdas muito baixos. Este PHB é necessário para aplicações que necessitam de garantias estritas no rácio máximo de perdas, no atraso e na sua variação. A taxa de serviço do agregado de tráfego que pertence a este PHB tem de ser sempre igual ou superior à taxa configurada em cada nó DiffServ, e esta taxa não deve depender da intensidade do tráfego nos outros PHB. Definido desta forma, o PHB EF emula um circuito virtual garantido porque proporciona um envio de pacotes com baixos atrasos e perdas.

Uma forma de implementar este PHB é utilizando um algoritmo de escalonamento de prioridade estrita em que a fila de espera com prioridade mais elevada é reservada para o tráfego EF [Wang01]. Na entrada desta fila de espera é implementado um *leaky bucket* para impedir que tráfego EF excessivo impeça outro tráfego de ser servido.

3.4.2.1.3 Assured Forwarding

O grupo de PHB AF [Heinanen99] contém quatro classes de serviço, cada uma com três níveis de precedência de descarte de pacotes. Para cada classe são atribuídos uma largura de banda mínima e um espaço mínimo nas filas de espera para armazenar os pacotes da classe respectiva. Ambos estes parâmetros podem ser configurados nos nós da rede. Os três níveis de precedência de descarte de pacotes facilitam a selecção dos pacotes a descartar quando a rede se encontra congestionada e a largura de banda mínima reservada para a classe de serviço não é suficiente. Os pacotes que tiverem maior prioridade de descarte serão os primeiros a ser descartados.

A implementação do grupo de PHB AF pode ser conseguida com uma partição de largura de banda entre as classes e com prioridades de descarte dentro de cada classe. A partição de largura de banda pode ser efectuada, por exemplo, através de WFQ, que associa pesos de acordo com os requisitos mínimos de largura de banda. As prioridades de descarte podem ser implementadas, por exemplo, pelo RED ou pelo RIO.

3.4.2.2 Classificação e condicionamento de tráfego

Os nós fronteira de uma rede DiffServ implementam as funcionalidades de classificação e condicionamento de tráfego. Estes nós são responsáveis pelo mapeamento dos pacotes numa das classes de serviço suportadas pela rede, e por garantir que o tráfego que é injectado na rede está de acordo com o perfil definido pelo SLA. Nos nós do núcleo, a atribuição de recursos é baseada apenas nas classes de serviço. As funções realizadas pelos nós fronteira, classificação e condicionamento de tráfego, estão representadas na Figura 3-6.

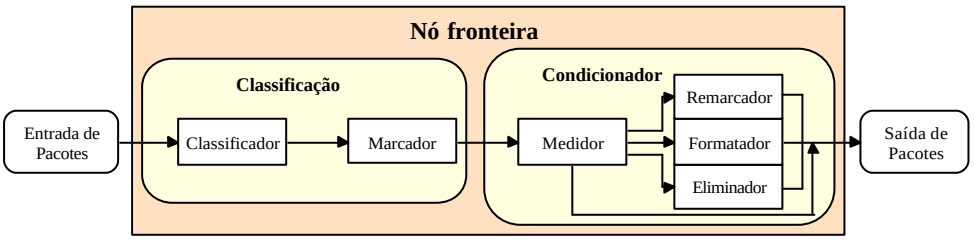


Figura 3-6 : Elementos de um nó fronteira.

No módulo de classificação, o pacote é classificado numa classe de serviço (pelo classificador), sendo depois marcado com o correspondente DSCP (pelo marcador). Existem dois tipos de classificação: com base no BA ou em *Multi-Field* (MF). Nos nós do núcleo a classificação é efectuada apenas com base no BA. Nos nós fronteira podem existir os dois tipos de classificação. A marcação consiste em colocar o campo DS do pacote de acordo com o valor DSCP. Após ter sido marcado, o pacote é encaminhado para a fila de espera correspondente. A marcação pode ocorrer em diversos locais: nas redes dos clientes, se estas suportarem DiffServ, ou apenas nos nós fronteira das redes DiffServ. Se cada domínio usar DSCPs diferentes, é necessário remarcar os pacotes na fronteira entre dois domínios.

O módulo de condicionamento de tráfego tem a função de policiar o tráfego para o obrigar a estar de acordo com o TCA definido entre os clientes e os fornecedores de serviço. Este módulo é constituído por quatro elementos: medidor, remarcador, formatador e eliminador. O medidor mede o fluxo de tráfego em cada classe de serviço para verificar se ele está de acordo com o perfil acordado. Os pacotes que estão dentro do perfil passam directamente para a rede DiffServ, enquanto que os outros são condicionados nos módulos seguintes. A remarcação é efectuada quando um pacote está fora do perfil e, nesse caso, este pode ser marcado com um DSCP diferente, por exemplo, para lhe associar uma prioridade de descarte mais elevada. Os formatadores atrasam os pacotes que estão fora do perfil de tráfego, com o objectivo de tornar o tráfego de acordo com o perfil acordado. Como o tráfego pode ser modificado devido à passagem nos diversos nós, pode ser necessário formatá-lo de novo na fronteira entre domínios DiffServ. O eliminador descarta os pacotes que estão fora de perfil. Este módulo é mais fácil de implementar que um formatador, porque não necessita de nenhuma fila de espera para atrasar o pacotes.

3.4.2.3 Análise da arquitectura DiffServ

A arquitectura DiffServ implementa mecanismos muito simples para suportar QdS. Na arquitectura DiffServ:

- Cada nó da rede armazena apenas o estado de um número reduzido de classes.
- A classificação dos pacotes nos nós do núcleo é efectuada apenas com base no campo DSCP.

? Não existe reserva de recursos para cada fluxo no início de cada sessão. Os recursos estão previamente atribuídos a cada classe, embora esta atribuição possa ser dinâmica. Não existem garantias de recursos mínimos por fluxo.

Para proporcionar um boa gestão de recursos extremo-a-extremo, foi definida uma proposta que integra ambas as arquitecturas, em que os recursos da rede de um domínio DiffServ são geridos por um *Bandwidth Broker* (BB).

3.4.3 Solução integrada IntServ/DiffServ

Numa solução integrada [Bernet98], a arquitectura DiffServ é utilizada nas redes do núcleo, enquanto que a arquitectura IntServ é usada nas redes de cliente e de acesso, ou em ISPs de pequena dimensão, onde existem menos fluxos a percorrer cada nó. Estas duas arquitecturas devem coexistir de modo a oferecer de uma forma integrada um conjunto de níveis de QdS extremo-a-extremo.

Numa solução deste tipo os clientes utilizam o RSVP para efectuar o pedido de reserva de recursos à rede. A rede total inclui domínios IntServ, em que a classificação é baseada em MF e o controle de tráfego é efectuado por fluxo, e domínios DiffServ, em que a classificação é baseada no DSCP e o controle de tráfego é efectuado por classes de serviço. A integração dos dois modelos tem sido objecto de discussão no IETF, o qual apresentou já a estrutura que se apresenta na Figura 3-7 [Bernet98].

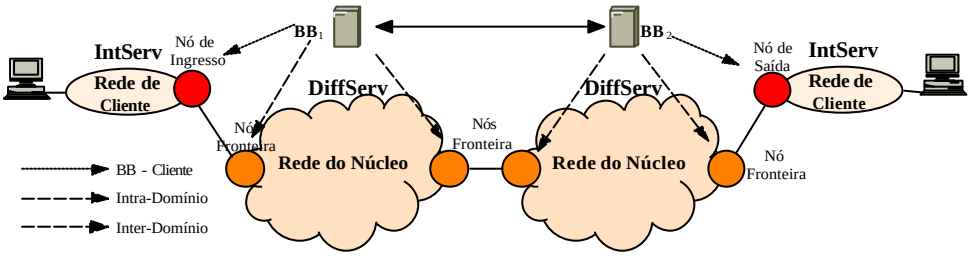


Figura 3-7 : Solução integrada IntServ/DiffServ e gestão de recursos em DiffServ com BB.

Os utilizadores origem e destino estão directamente ligados a uma rede de cliente IntServ. Cada rede IntServ tem um *router* na fronteira com a rede DiffServ. A rede de núcleo é constituída por um conjunto de *routers* que suportam a arquitectura DiffServ. Do ponto de vista das redes IntServ, a nuvem DiffServ é apenas uma ligação virtual.

Os pedidos IntServ têm de ser mapeados nas capacidades da rede DiffServ: (1) é seleccionado o PHB apropriado para os serviços pedidos, (2) é efectuado um policiamento

adequado na fronteira da rede DiffServ, e (3) os parâmetros DiffServ são mapeados nos parâmetros IntServ.

Ambos os clientes emissor e receptor usam o protocolo RSVP para controlar a QdS. Quando o emissor inicia o pedido de estabelecimento de um fluxo, troca mensagens PATH e RESV com o receptor. As mensagens RSVP são ignoradas no interior da rede DiffServ. Quando as mensagens passam pelo *router* de entrada/saída da rede DiffServ, este consulta um BB do domínio DiffServ para decidir se admite ou não o novo fluxo. O BB mantém informação sobre o estado das reservas existentes e baseia-se nelas para admitir ou rejeitar o novo pedido. Pode também decidir a forma como é efectuado o mapeamento entre pedidos IntServ e o modelo DiffServ. Os BB são também utilizados para fazer a negociação de recursos entre domínios DiffServ [Neilson99]. A relação entre dois domínios é efectuada por SLAs entre domínios, isto é, pelo contrato que especifica o perfil do tráfego que atravessa os nós fronteira e a QdS atribuída. Para determinar se um pedido pode ser aceite, o BB_1 (BB_2) comunica com o BB_2 (BB_1), para verificar se ambos os domínios DiffServ têm recursos suficientes para aceitar o pedido. Os BB actuam como agentes de gestão de recursos para os seus domínios DiffServ. Eles são constituídos por um bloco de controle de admissão que toma decisões com base na utilização dos recursos da rede, e um bloco de controle de políticas de QdS, que definem que utilizador tem acesso a que serviço e com que QdS. As reservas efectuadas pelos BB são baseadas em agregados de fluxos e não em fluxos individuais.

Nas secções seguintes são apresentadas novas propostas de arquitecturas e mecanismos para resolver os problemas de ambas as arquitecturas IntServ e DiffServ, que tentam estabelecer diferentes compromissos entre as mesmas.

3.4.4 Novas arquitecturas e mecanismos de controle de admissão, atribuição e gestão de recursos

A arquitectura DiffServ é muito simples de implementar e os seus modelos de serviço têm qualidade bastante superior ao modelo de serviço de melhor esforço. No entanto, se uma classe fica congestionada, todos os fluxos pertencentes a essa classe podem ver a sua QdS degradada. Por outro lado, na arquitectura IntServ, as garantias de QdS são dadas através de reserva de recursos fluxo-a-fluxo, mas a implementação desta arquitectura em redes de grande dimensão e de elevado débito é muito complexa. As redes com integração de IntServ e DiffServ e gestão de recursos em DiffServ com BB resolvem alguns dos

problemas, mas este processo centralizado implica um conhecimento da topologia da rede e do percurso de cada fluxo. Estas tarefas são complicadas de implementar e manter num só elemento, se a rede DiffServ for de grandes dimensões.

Com o objectivo de resolver estes problemas, foram propostas novas arquitecturas e mecanismos que tentam aproveitar a escalabilidade da arquitectura DiffServ e os modelos de serviço eficazes da arquitectura IntServ, das quais se destacam: DPS (*Dynamic Packet State*) [Stoica98, Stoica99], em que a informação do estado do fluxo é inserida no cabeçalho do pacote IP, em vez de ser armazenada nos *routers*; controle de admissão no *router* de saída (*Egress Admission Control*) [Cetinkaya00], baseado em monitorização passiva da rede apenas no *router* de saída; agregação de reservas individuais [Baker01], em que o controle de admissão é efectuado para um grupo de fluxos, reduzindo assim a quantidade de mensagens de sinalização e armazenamento de estado nos *routers*; e investigação do estado de congestão da rede (*probing*) [Bianchi00, Elek00, Gibbens99, Kelly00], em que é inserido um fluxo de teste na rede para averiguar o seu nível de congestão. Uma descrição destes mecanismos é apresentada em [Sargento01a]. As secções seguintes descrevem estes mecanismos com mais detalhe.

3.4.4.1 Dynamic Packet State

Esta nova arquitectura foi designada pelos autores de SCORE (*Scalable Core*) [Stoica98]. É semelhante à arquitectura DiffServ, mas na SCORE todos os *routers* realizam controle de admissão e gestão de recursos por fluxo sem, no entanto, necessitarem de manter armazenado o estado de cada fluxo. A técnica utilizada para implementar a rede SCORE é o DPS, em que o *router* de entrada adiciona ao cabeçalho de cada pacote a informação necessária sobre o fluxo ao qual o pacote pertence. Assim, os *routers* de entrada da rede têm de armazenar o estado de cada fluxo e efectuar controle de admissão por fluxo. Os *routers* do núcleo processam cada pacote com base no estado do fluxo transportado no cabeçalho respectivo e actualizam este estado antes de enviarem o pacote para o próximo nó. O processo de controle de admissão nestes *routers* é efectuado com base apenas na informação transportada no pacote e numa estimativa da largura de banda agregada reservada mantida por cada *router*.

O algoritmo de escalonamento considerado nesta proposta [Stoica99] é o CJVC apresentado na secção 3.3.5. Para poder determinar o tempo elegível de cada pacote sem armazenar a informação necessária ao seu cálculo nos *routers* do núcleo, os *routers*

fronteira armazenam esta informação no cabeçalho dos pacotes antes de estes entrarem na rede SCORE. Os *routers* do núcleo efectuam o escalonamento com base na informação armazenada nos pacotes. Esta informação inclui a largura de banda reservada para o fluxo ao qual o pacote pertence (b_i), a diferença entre o tempo limite do pacote e o instante de tempo de transmissão do pacote no *router* anterior ($TSL^k_{i,j?1}$), e a variável *slack* ($?^k_i$). Estes dois últimos parâmetros são alterados na passagem por cada *router*.

O controle de admissão é efectuado localmente em cada nó fronteira e do núcleo. Enquanto que nos *routers* fronteira o controle de admissão é efectuado com base na informação referente a cada fluxo activo na rede, nos *routers* do núcleo é efectuado com base apenas numa estimativa da largura de banda agregada e na largura de banda do fluxo que pede admissão. Esta arquitectura pressupõe a utilização, dentro de um domínio, de um protocolo de sinalização específico definido para o efeito. O protocolo de sinalização extremo-a-extremo pode ser, por exemplo, o RSVP. Os *routers* de entrada e saída da rede SCORE funcionam como interface destes dois protocolos. As mensagens PATH e RESV do RSVP são processadas apenas pelos *routers* fronteira e passam transparentemente na rede SCORE. A utilização de um protocolo de sinalização diferente do RSVP dentro da rede SCORE está relacionado com o facto de nesta rede não ser necessário instalar o estado das reservas individuais nos *routers*. Após receber a mensagem RESV, o *router* de entrada envia uma mensagem de sinalização especial em direcção ao *router* de saída. No percurso, cada *router* do núcleo, ao receber esta mensagem, decide localmente se pode ou não aceitar o pedido do fluxo, com base numa estimativa do limite superior da largura de banda reservada do agregado de fluxos, e com base na largura de banda do fluxo que pede admissão. Se a soma das larguras de banda for superior à capacidade da ligação, o fluxo é rejeitado em cada *router* do núcleo; caso contrário será aceite. Quando a mensagem de sinalização chega ao *router* de saída, é enviada no sentido contrário para o *router* de entrada (passando transparentemente no núcleo) para este tomar a decisão final.

Em [Stoica99] propõe-se que estes parâmetros, necessários ao funcionamento do algoritmo de escalonamento e ao mecanismo de controle de admissão, sejam introduzidos no campo de ToS e em 13 *bits* do campo *ip_off* usado para fragmentação dos pacotes. O único parâmetro que é mantido nos *routers* do núcleo é a largura de banda do agregado de fluxos. Desta forma, é possível ter uma rede semelhante à rede IntServ sem manutenção de estado por fluxo no seu interior. Esta arquitectura apresenta, no entanto, diversos

problemas e obstáculos: (i) é necessário que todos os *routers* no percurso do fluxo implementem o mesmo algoritmo de escalonamento; (ii) os *routers* do núcleo têm ainda uma considerável carga de processamento, embora menor do que na arquitectura IntServ; e (iii) impossibilita a utilização de protocolos que incluem compressão de cabeçalhos e segurança.

3.4.4.2 Controle de admissão no router de saída

Neste mecanismo [Cetinkaya00], as decisões de controle de admissão são efectuadas apenas no nó de saída, sem manutenção do estado por fluxo em nenhum dos nós, nem coordenação do estado entre os nós de núcleo e o de saída. Neste sentido, as decisões de admissão são efectuadas com base nas medidas do tráfego que atravessa o *router* de saída. A técnica utilizada consiste em estimar a taxa de serviço mínima disponível no percurso extremo-a-extremo, com base nos atrasos sofridos pelos pacotes, e verificar se esta é ou não suficiente para aceitar o novo fluxo (com os seus requisitos de QdS). As medidas podem incorporar os efeitos do *cross traffic*, sem o medir ou controlar explicitamente. O *cross traffic* é todo o tráfego que no seu percurso interage com o tráfego que se está a medir, mas que tem um *router* de saída diferente do tráfego em questão.

O início do processo de controle de admissão é efectuado, tal com no DPS, com o envio de uma mensagem de sinalização de reserva de recursos, por exemplo através do RSVP, especificando as características do fluxo que pede admissão. A mensagem passa transparentemente nos *routers* do núcleo e é processada apenas pelo *router* de saída do domínio. O *router* de saída toma uma decisão de aceitação ou rejeição do fluxo e notifica o emissor e o receptor da sua decisão. A decisão é baseada nas medidas efectuadas pelo *router* de saída, em intervalos de tempo anteriores, na taxa máxima de chegada de pacotes (com base no número de *bits* recebidos nos intervalos de tempo de medição) e na taxa de serviço mínimo disponível (com base nos atrasos sofridos pelos pacotes). Este mecanismo é independente do algoritmo de escalonamento implementado nos *routers*.

Neste processo de controle de admissão é utilizada a teoria dos *envelopes* [Qiu99]. Consideram-se dois tipos de *envelopes*: de chegada e de serviço. Os *envelopes* de chegada correspondem à taxa máxima de chegadas de pacotes. Os *envelopes* de serviço correspondem à taxa mínima de serviço disponível. Os *envelopes* são calculados para cada par *router* de entrada/*router* de saída. Além deste cálculo, é também determinada a variância das medidas efectuadas para determinar o grau de confiança destas. As medidas

realizadas no *router* de saída baseiam-se no intervalo de tempo que os pacotes demoram a percorrer o percurso entre um *router* de entrada da rede e um *router* de saída. Para que o instante de partida de um pacote no *router* de entrada seja conhecido no *router* de saída, o *router* de entrada tem de marcar o instante de partida no cabeçalho do pacote, e os dois elementos têm de estar sincronizados.

Após calcular o *envelope* de chegada e o *envelope* do serviço mínimo disponível, o nó de saída executa um algoritmo de controle de admissão para aceitar ou rejeitar o novo fluxo. O fluxo é admitido no sistema se o serviço mínimo disponível for suficiente para garantir o atraso máximo admissível e para garantir que os requisitos de QdS dos fluxos já admitidos no sistema não são violados. A decisão de admissão baseia-se numa equação que é função do envelope de chegada e da sua variância, do envelope de serviço e da sua variância, da taxa de pico do fluxo que pede admissão, do atraso máximo admissível e do grau de confiança que se quer ter na decisão. Os detalhes da equação são apresentados em [Cetinkaya00].

O único *router* que necessita de participar no processo de controle de admissão é o *router* de saída. No entanto, apenas uma previsão do nível de congestão da rede é efectuada. As condições da rede podem modificar-se e a QdS dos fluxos admitidos pode ser degradada. O parâmetro do grau de confiança relaciona-se com este facto. Com um elevado grau de confiança, um fluxo é aceite apenas se houver garantias de que a probabilidade da rede sofrer congestão é muito baixa.

Note-se que apenas é possível prever a taxa de serviço disponível para um determinado fluxo, se existirem fluxos activos que atravessam o mesmo par *router* de entrada/*router* de saída. Se isso não acontecer, o novo fluxo, se for admitido, pode roubar recursos que estão previamente atribuídos a fluxos pertencentes a outros pares origem/destino.

3.4.4.3 Agregação de reservas individuais

O controle de admissão baseado em agregação de reservas individuais é um mecanismo que permite ter, para cada fluxo, as mesmas garantias estritas obtidas com a arquitectura IntServ, sem no entanto exigir uma sinalização fluxo-a-fluxo. Nesta técnica, numa região de agregação de tráfego, mantém-se uma reserva para cada agregado de fluxos; a largura de banda desta reserva varia ao longo do tempo muito mais lentamente do que o intervalo entre pedidos de reservas dos fluxos individuais. Os *routers* de entrada da região de

agregação mantêm o estado de reserva fluxo-a-fluxo, e verificam, para cada fluxo que pede admissão, se existe ou não largura de banda suficiente no agregado. Devido à existência prévia de uma reserva, os fluxos que tentam admissão necessitam de sinalizar apenas estes nós para verificar se na reserva do agregado existem ou não recursos disponíveis para os aceitar. Quando a largura de banda reservada para o agregado não é suficiente para comportar o novo fluxo, ou é muito superior à largura de banda utilizada, esta pode ser actualizada com uma inserção ou libertação de um volume de largura de banda no agregado. Este volume de largura de banda, que será designado ao longo desta Tese por *bulk*, é normalmente bastante superior à largura de banda de cada fluxo. Os *routers* do núcleo apenas são sinalizados para efectuar estas actualizações.

Em [Baker01] é proposta uma extensão do RSVP para permitir efectuar reservas de recursos por agregados. Nesta extensão, quando não é necessário actualizar a largura de banda do agregado, as mensagens de sinalização RSVP são "escondidas" dos *routers* do núcleo da rede para estes não participarem no processo de reserva de recursos fluxo-a-fluxo. Na entrada da região de agregação, o tipo de protocolo presente no cabeçalho das mensagens RSVP é alterado de RSVP para RSVP-E2E-IGNORE e restituído na saída da região de agregação. Desta forma, os *routers* no interior da região de agregação ignoram a mensagem e apenas a enviam para o *router* seguinte. Quando é necessário actualizar a largura de banda de um ou mais agregados, para poder admitir um novo fluxo, o valor do tipo de protocolo IP não é alterado, a mensagem RSVP não será ignorada pelos *routers* no percurso de cada agregado, e todos eles verificam se existe largura de banda disponível nas ligações para inserir um novo *bulk* no agregado. Se a resposta for positiva em todos os *routers* no percurso de cada agregado que não tem largura de banda disponível, o fluxo é aceite e um novo *bulk* é reservado para esse agregado. Quando a largura de banda disponível num agregado é superior ao tamanho de um *bulk*, porque alguns fluxos saíram do sistema, é enviada uma mensagem RSVP para todos os nós da rede para libertar um ou mais *bulks* do agregado e actualizar o valor da reserva.

Em [Pan00] é apresentada uma forma de efectuar agregação de reservas de recursos apenas nas ligações entre domínios. Para o efeito, é definido um protocolo de reserva de recursos entre domínios, o *Border Gateway Reservation Protocol* (BGRP). Para o seu funcionamento, o BGRP recolhe informação do protocolo de encaminhamento entre domínios, o *Border Gateway Protocol* (BGP) [Rekhter95]. O BGP define uma topologia

em árvore em que a raiz se encontra no domínio destino (*sink tree*). As reservas de agregados são construídas tendo como base esta topologia. O BGRP opera apenas entre *routers* fronteira e estabelece as reservas apenas entre estes *routers*. Cada domínio pode usar outro qualquer protocolo para reserva de recursos dentro do domínio. Desta forma, os *routers* no interior de cada domínio mantêm apenas a informação da *sink tree*, e o número total de reservas em cada *router* é proporcional ao número de domínios da *Internet*.

O funcionamento do BGRP é muito semelhante ao do RSVP. No BGRP são definidas as mensagens PROBE e GRAFT que têm as mesmas funções que as mensagens PATH e RESV do RSVP, respectivamente. A diferença essencial é o facto de as reservas serem estabelecidas apenas entre *routers* fronteira, e para agregados de fluxos. Estas mensagens são activadas apenas quando é necessário actualizar a reserva do agregado, com uma largura de banda significativamente superior à largura de banda dos fluxos. A mensagem PROBE é enviada pelo *router* fronteira do domínio de origem em direcção ao *router* fronteira do domínio destino para coleccionar informação de encaminhamento ao longo do percurso da reserva. A mensagem GRAFT é enviada no sentido contrário para estabelecer a reserva no percurso definido pelos *routers* fronteira. Ao processar a mensagem GRAFT, cada *router* fronteira realiza a interface com os protocolos de estabelecimento de reservas suportados pelo domínio para estabelecer as reservas dentro do domínio.

[Schelén98] propõe a utilização de agentes por domínio para realizar o controle de admissão e efectuar a reserva de recursos para agregados de fluxos. Os agentes têm o mapa da topologia do seu domínio através de obtenção de informação do protocolo de encaminhamento (por exemplo, o OSPF - *Open Shortest Path First*). A informação sobre as larguras de banda das ligações no domínio é obtida através de pedidos enviados aos *routers* (por exemplo, por SNMP - *Simple Network Management Protocol*). Quando um agente comunica com um agente vizinho para efectuar uma reserva, o endereço origem incluído no pedido de reserva determina o ponto da rede (a interface do *router*) em que o tráfego atravessa a fronteira entre os domínios. Para agregar o tráfego no seu domínio, o agente vizinho necessita apenas de saber o ponto de entrada e de saída no seu domínio. O tráfego com o mesmo ponto de entrada e de saída no domínio é agregado na mesma reserva. Assim, o tráfego de diferentes origens pode ser agregado numa única reserva num domínio, desde que este não seja o domínio de origem.

Além das propostas de controle de admissão e reserva de recursos por agregados de fluxos existentes na literatura, existem também alguns estudos analíticos relativos a caracterização de agregados de fluxos. Um agregado de fluxos é caracterizado analiticamente de uma forma precisa em [Schmitt99], tendo em conta que os parâmetros que caracterizam um fluxo individual são os parâmetros de um *leaky bucket*. Em [Schmitt99] são derivadas fórmulas para calcular a largura de banda do agregado e o atraso sofrido pelos pacotes de um agregado na região de agregação.

A agregação implica um compromisso. Com maior agregação, isto é, com *bulks* de largura de banda de maior dimensão, mais fluxos deixarão se ser admitidos e a utilização decresce; com menor agregação, a utilização cresce mas o número de mensagens de sinalização aumenta. Se o tráfego oferecido for relativamente constante, os nós do núcleo raramente têm de ser sinalizados; caso contrário, a carga de sinalização será semelhante à de IntServ. Em [Fu01] foi realizado um estudo de desempenho para analisar o impacto das características do tráfego na eficácia da agregação.

Em domínios com dimensão relativamente elevada, a hierarquização de domínios, ou seja, a sua divisão em áreas mais pequenas e a configuração de agregados por área em oposição a agregados extremo-a-extremo num domínio, permite aumentar a utilização dos seus recursos. No capítulo 9 é efectuado um estudo de desempenho da agregação ao nível da carga de sinalização e da utilização de recursos em domínios hierarquizados.

3.4.4.4 Probing (investigação do nível de congestionamento da rede)

Com o objectivo de evitar o uso de protocolos de sinalização e diminuir a carga de processamento dos *routers* fronteira e do núcleo, foi proposto um mecanismo de controle de admissão de fluxos [Bianchi00, Elek00, Gibbens99, Kelly00] baseado na investigação do estado de congestão da rede, através da inserção de fluxos de teste (*probing*).

Neste mecanismo, o controle de admissão é efectuado pelos equipamentos terminais origem e destino ou pelos *routers* de entrada e de saída de um domínio, através da medição do estado de congestão da rede no percurso do fluxo. Os parâmetros de QdS normalmente utilizados são o atraso ou a variação do atraso dos pacotes, o rácio de pacotes perdidos, ou o rácio de pacotes marcados (os pacotes são marcados se as filas de espera que atravessam tiverem uma ocupação superior a um determinado limiar). O mecanismo de *probing* permite avaliar qual o impacto no nível de QdS da rede de admitir um novo fluxo. Se o desempenho da rede com a inserção do novo fluxo ainda proporcionar a QdS requerida por

cada fluxo previamente admitido e pelo fluxo que pede admissão (por exemplo, se o rácio de perdas de pacotes ou o atraso forem inferiores ao máximo admissível), ele é admitido.

O mecanismo de *probing* compreende os seguintes passos. Quando um fluxo pede admissão na rede, o emissor envia, durante um determinado intervalo de tempo, denominado de tempo de *probing*, uma sequência de pacotes com as mesmas características do fluxo que pede admissão. No final do tempo de *probing*, o receptor envia ao emissor um pacote com informação sobre os pacotes recebidos (por exemplo, o atraso ou o número de pacotes), e o emissor, após determinar o nível de QdS da rede, decide se admite ou não o novo fluxo.

As propostas de implementação deste processo de controle de admissão apresentam algumas diferenças. Em [Elek00] e [Bianchi00] os fluxos de *probing* são enviados num nível de prioridade inferior ao dos fluxos de dados. O parâmetro de QdS utilizado para realizar uma decisão de controle de admissão é o rácio de perdas de pacotes. No fim do tempo de *probing*, o receptor envia ao emissor uma mensagem que contém informação sobre o número de pacotes recebidos. Esta mensagem é enviada com uma elevada prioridade para haver garantias de que é transmitida sem perdas. O emissor decide se aceita ou não o fluxo com base nessa informação.

Em [Benameur01] é proposta a inserção de pacotes de teste na rede e medida a sua taxa de perdas para efectuar o controle de admissão de tráfego elástico que é enviado através de TCP. Os autores argumentam que mesmo para tráfego TCP é necessário limitar o número de sessões activas para evitar situações de congestão, e que o processo de controle de admissão tem de ser efectuado num intervalo de tempo muito pequeno dado o reduzido tamanho das sessões. Os pacotes de teste são enviados para a rede, e basta um deles ser perdido para indicar que existe alguma congestão na rede e a sessão ser rejeitada.

As propostas [Gibbens99, Kelly00] são baseadas na atribuição de um preço às marcas de congestão ECN (*Explicit Congestion Notification*) [Floyd99]. Todos os pacotes (por exemplo, pertencentes a tráfego de melhor esforço, a tráfego com requisitos em tempo real, e a tráfego de teste) são tratados de uma forma idêntica, e são marcados se a ocupação das filas de espera que atravessam exceder um determinado limiar. Os utilizadores podem enviar o tráfego desejado, mas pagam um preço extra (além do preço que pagam pelo serviço) pelos pacotes que são marcados. Quando são enviados os pacotes de teste para a rede, o utilizador é informado da percentagem de pacotes marcados. Se esta percentagem

for elevada, o utilizador apenas inicia a transmissão de informação se estiver disposto a pagar por todos os seus pacotes que forem marcados durante a transmissão de informação.

Em [Breslau00b] é efectuado um estudo mais aprofundado destas técnicas aplicadas a cenários diferentes. Os estudos de simulação realizados incluem, além dos mecanismos de *probing* simples descritos anteriormente nesta secção (com perdas e marcação de pacotes), mecanismos de *probing* que permitem reduzir a perturbação que os próprios fluxos de *probing* introduzem na rede. Quando muitos fluxos realizam o processo de *probing* em simultâneo, o nível de QdS pode ser reduzido a um ponto que impeça qualquer fluxo de ser aceite. Esta situação em que todos os fluxos se debatem para entrar e nenhum consegue é designada de *thrashing*. Os mecanismos apresentados em [Breslau00b] para minimizar este problema são o *Slow Start Probing* e o *Early Reject*. No mecanismo de *Slow Start Probing*, o tempo de *probing* é dividido em intervalos, e em cada intervalo é introduzido um fluxo de *probing* com uma largura de banda diferente (e crescente). No primeiro intervalo, a largura de banda do fluxo de *probing* é muito inferior à largura de banda do fluxo que pede admissão. No último intervalo, a largura de banda do fluxo de *probing* é igual à do fluxo que pede admissão. Se no final de cada intervalo o nível de QdS determinado é superior ao limiar, o processo de *probing* continua (no máximo até ao final do tempo de *probing*) e a largura de banda do fluxo de *probing* é aumentada no próximo intervalo; caso contrário, o fluxo é automaticamente rejeitado e o processo de *probing* é terminado. No mecanismo de *Early Reject*, a largura de banda do fluxo de *probing* é sempre igual à do fluxo que pede admissão mas, em intervalos mais pequenos que o tempo de *probing*, verifica-se se o nível de QdS observado se encontra abaixo do limiar, e se assim for, o fluxo é rejeitado e o processo de *probing* é imediatamente terminado. [Breslau00b] conclui que, num sistema com uma classe de serviço, o mecanismo de *probing* realiza um controle de admissão correcto e possibilita a existência de um modelo de serviço de carga controlada como o definido em IntServ. No entanto, [Breslau00b] identificou um problema de roubo de recursos a fluxos já admitidos, que ocorre devido ao facto de não haver uma reserva explícita de largura de banda para cada fluxo.

Mais recentemente, em [Kelly01] é proposto um protocolo de controle de admissão, semelhante ao apresentado em [Breslau00b], baseado na inserção de marcas ECN em fluxos de teste. Geralmente, o número de pacotes perdidos é muito inferior ao número de marcas. Por isso, em [Kelly01] afirma-se que é possível reduzir significativamente os

tempos de *probing* utilizando os limiares de marcação de pacotes, pois como estes são superiores, o número de pacotes de teste que é necessário enviar para ter uma determinada confiança estatística nas estimativas obtidas diminui. Um problema inerente a um controle de admissão baseado em marcas de congestão é o facto de ser muito difícil relacionar o limiar de marcação de pacotes com o rácio de perdas atingido, podendo resultar num mau aproveitamento de recursos.

Uma questão fulcral a colocar é a da aplicabilidade deste tipo de mecanismos. O mecanismo de *probing* tenta prever o nível de QdS da rede, e através desta previsão, efectua uma decisão de admissão dos fluxos. Note-se que esta previsão é baseada em medidas efectuadas num pequeno intervalo de tempo, cujo resultado é função do congestionamento da rede nesse intervalo. Além disso, como o intervalo de tempo é pequeno, não é possível estimar adequadamente rácios de perdas nulos ou muito pequenos. Nos estudos de simulação efectuados em [Breslau00b] e nos estudos que serão apresentados nos capítulos 7 e 8, verifica-se que os mecanismos de *probing* não são adequados para serviços que necessitam de garantias muito estritas de QdS, e que, com estes mecanismos é possível garantir (não estritamente) um rácio de perdas da ordem de 1% ou superior. Assim, estes mecanismos não são aplicáveis aos serviços que se enquadram no modelo de serviço garantido da arquitectura IntServ, e no modelo EF da arquitectura DiffServ. Os mecanismos de *probing* são aplicáveis aos serviços denominados de *soft real time*, ou seja, serviços com requisitos em tempo real, mas que toleram algumas perdas ou atrasos na transmissão dos seus pacotes. Estes serviços têm uma QdS superior aos serviços que pertencem ao modelo de serviço de melhor esforço, mas não exigem garantias estritas. Em [Breslau00b] verificou-se que o desempenho dos mecanismos de *probing* é semelhante ao desempenho obtido com mecanismos de controle de admissão baseados em estimativas da taxa agregada de fluxos activos (MBAC). O controle de admissão baseado em MBAC é aplicável a serviços pertencentes ao modelo de serviço de carga controlada. Os mecanismos de *probing* podem também ser aplicados a classes AF da arquitectura DiffServ utilizadas para serviços *soft real time*.

Para verificar mais concretamente qual a possibilidade de aplicação dos mecanismos de *probing* aos serviços existentes hoje em dia, apresenta-se uma tabela de requisitos de serviços definidos pelo ITU-T [ITU-T]. Observa-se que o rácio de perdas tem de ser aproximadamente nulo para serviços de Telemetria e de comércio pela *Internet*. Os

mecanismos de *probing* não podem ser aplicados a estes serviços. No entanto, para os restantes serviços, os rácios de perdas são da ordem de 1% ou 3%, valores que são possíveis de garantir com o mecanismo de *probing* [Breslau00b]. Pela observação da tabela, verifica-se que os mecanismos de *probing* podem ser aplicados a uma grande parte das aplicações existentes actualmente.

Tipo de serviço	Serviço	Taxa dos dados (Kb/seg)	Atraso	Variação no atraso (mseg)	Rácio de perdas (%)
Conversacional em tempo real	Voz conversacional	4-25	<150 mseg	<1	<3
	Videofone	32-384	<150 mseg		<1
	Telemetria (controle)	<28.8	<250 mseg		~0
	Jogos	<1	<250 mseg		<3
Interactivo	Mensagens de voz	4-13	<1 seg	<1	<3
	Procura info na *Internet*		4 seg/página		
	Comércio na *Internet*		4 seg		~0
Streaming	Fluxos de áudio	32-384	<10 seg	<1	<1
	Vídeo	32-384	<10 seg		<1
	Telemetria (monitorização)	<28.8	<10 seg		~0

Tabela 3-1 : Qualidade de Serviço [ITU-T].

No capítulo 7 serão apresentados os mecanismos de *probing*, o seu problema de roubo de recursos, e a proposta de um novo mecanismo de *probing*, denominado de ?-*probing* [Sargento01c], que elimina o roubo de recursos num sistema com múltiplas classes de serviço. No capítulo 8 é descrita a implementação de um protótipo laboratorial com uma arquitectura de base DiffServ, que inclui ambos os mecanismos de *probing* e de ?-*probing*.

CAPÍTULO 4

REDES DE ACESSO

As redes de telecomunicações podem ser classificadas em redes de cliente (por exemplo, a rede de uma mesma empresa), redes de acesso e redes de transporte. A rede de acesso é portanto, a interface entre o cliente e a rede de transporte.

As primeiras redes de acesso não eram mais do que as centrais de comutação local e os pares de cobre que as ligavam aos telefones instalados nos clientes. A evolução tecnológica levou ao desenvolvimento das redes de acesso e criou oportunidades para o aparecimento de novos serviços. Actualmente, uma rede de acesso deve ser capaz de simultaneamente integrar qualquer tipo de serviços na mesma infra-estrutura de acesso, e atribuir uma QdS (Qualidade de Serviço) diferenciada consoante os requisitos de cada serviço e/ou cliente.

Este capítulo apresenta uma perspectiva evolutiva das redes de acesso, ao nível da infra-estrutura física de suporte e ao nível das tecnologias de transporte dos serviços. A descrição evolutiva inicia-se com as redes de acesso de circuitos e culmina nas redes de acesso de recursos partilhados com o IP (*Internet Protocol*) como tecnologia de transporte.

Este capítulo é organizado da seguinte forma. Na secção 4.1 é apresentado o modelo de negócios, ou seja, a interligação entre as diferentes entidades lógicas participantes no

processo de pedido, aceitação, e fornecimento de um serviço e de recursos para que o mesmo seja acedido. No modelo de negócios é também apresentado o papel de cada entidade. Na secção 4.2 é apresentada uma perspectiva histórica de evolução das redes de acesso fixas e com fios. É também descrita a forma como estas redes estão ligadas às redes PSTN (*Public Switched Telephone Network*) e RDIS (Rede Digital com Integração de Serviços). Um exemplo de redes de acesso fixas e sem fios são as redes de acesso baseadas no sistema LMDS (*Local Multipoint Distribution Service*) [Int-LMDS]. As redes de acesso sem fios não se encontram no âmbito desta Tese, e como tal, não serão comtempladas nesta análise. Finalmente, na secção 4.3 apresenta-se uma perspectiva do futuro das redes de acesso.

4.1 Modelo de negócios

Fornecer um serviço a um cliente requer uma entidade para providenciar o serviço, o fornecedor de serviço (SP – *Service Provider*), e uma outra entidade para providenciar a infra-estrutura de ligação onde o serviço é transportado para o cliente, o fornecedor da rede de acesso (NAP - *Network Access Provider*). O NAP pode também assumir as funções de um SP. No entanto, neste trabalho considera-se que as funções de cada entidade são separadas.

O SP fornece os serviços de acordo com um contrato que estabelece com o cliente, o acordo de nível de serviço (SLA – *Service Level Agreement*). Um SLA pode incluir o custo de um serviço, as suas características e requisitos de QdS. Podem existir diferentes SLAs entre o mesmo SP e diferentes clientes. Deste modo, os SPs têm de suportar uma variada gama de serviços de acordo com os requisitos necessários para cada cliente.

O NAP fornece os recursos de rede e os mecanismos necessários para garantir a QdS contratada para cada cliente, de modo a que o contrato com o SP seja cumprido. Numa escala temporal bastante superior à da duração das chamadas, o NAP pode redimensionar a sua rede e incrementar os recursos de acordo com um possível aumento da sua utilização.

Na Figura 4-1 é apresentado o modelo de negócios de interligação das várias entidades definido no *TeleManagement* (TM) *Forum* [TMForum97]. Os NAPs usam os recursos disponíveis nas redes de acesso para fornecer os serviços de suporte (ou serviços de rede). Os SPs fornecem os níveis diferentes de serviços, avançados e básicos,

construídos através dos serviços de suporte. Finalmente, os clientes usam aplicações de acordo com os serviços disponíveis.

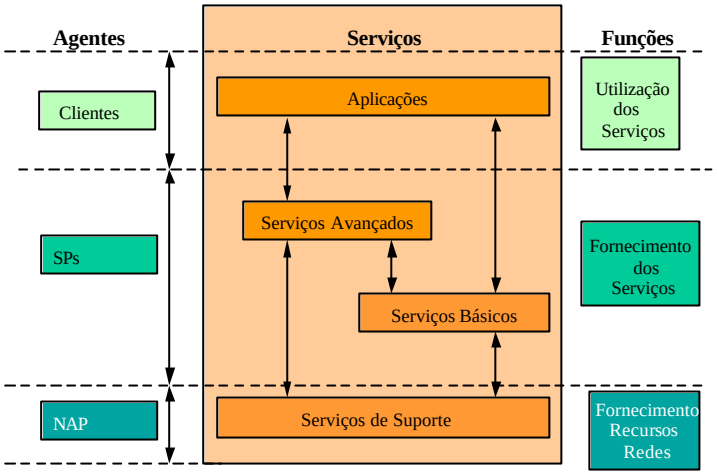


Figura 4-1 : Modelo de negócios.

4.2 Evolução das redes de acesso

A evolução das redes de acesso está intimamente associada com a evolução das infra-estruturas de suporte físico e das tecnologias de transporte: desde circuitos dedicados baseados em pares de cobre, até uma infra-estrutura partilhada em que é utilizado o IP como tecnologia de transporte. Esta perspectiva é apresentada nas secções seguintes.

4.2.1 Redes de acesso tradicionais

As redes de acesso tradicionais, suportadas em pares de cobre, ainda são as redes mais comuns actualmente (Figura 4-2). Os fios de cobre ligam o equipamento terminal, o cliente, à central de comutação local (LE - *Local Exchange*), analógica ou digital. Tradicionalmente estas redes eram utilizadas apenas para o transporte de voz (serviço telefónico). A topologia da rede de acesso é uma estrela, pois existe um par de fios de cobre em cada ligação entre um cliente e a central de comutação. Um telefone analógico pode ser ligado directamente ao par de fios da rede de acesso, enquanto que um telefone digital necessita de um equipamento de interface, terminador de rede (NT – *Network Terminator*). O NT realiza, além de outras funções, a conversão de sinais de um par de fios

(linha de assinante) para dois pares de fios, característica das comunicações digitais, e a implementação de um código de linha adequado ao transporte da informação digital entre o cliente e a estação RDIS associada. Com o crescimento das novas tecnologias, as redes de acesso tradicionais passaram a ser utilizadas também para o transporte de dados. Numa primeira fase, como os dados são digitais e a rede de acesso suporta apenas transmissão analógica, os terminais ligam-se à rede através de um *modem* para fazer a conversão. Numa rede de acesso com transmissão digital, como acontece na RDIS, os terminais de dados podem ser ligados à rede de acesso através do interface RDIS.

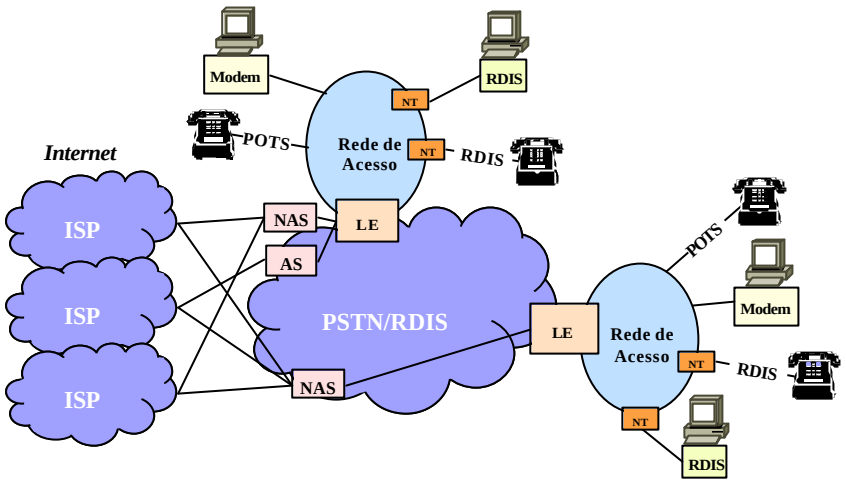


Figura 4-2 : Redes de acesso com comutação de circuitos.

Num acesso à *Internet* através de *dial-up*, é estabelecido um circuito entre um terminal e um servidor de acesso (AS – *Access Server*) através da rede PSTN/RDIS. O AS implementa o ponto de presença (PoP – *Point of Presence*) na rede de acesso do fornecedor de serviço *Internet* (ISP – *Internet* SP) e, deste modo, é o elemento de interface entre a rede de acesso e a rede IP do ISP. Ele fornece o acesso a um determinado ISP, autentica e autoriza o utilizador. Como estratégia comercial, o AS pode ser substituído por um servidor de acesso de rede (NAS – *Network* AS) que pode representar vários ISPs e fornecer o acesso a qualquer um deles simultaneamente. O NAS permite que o terminal escolha o ISP fornecedor de serviço.

Num acesso *dial-up* é atribuída uma largura de banda fixa e reservada ao utilizador durante toda a sessão de acesso à *Internet*. Normalmente, a duração típica de uma sessão deste tipo é muito superior à duração de uma chamada telefónica, da ordem das dezenas de

minutos. Ter um circuito permanentemente atribuído durante todo este tempo para uma única sessão é pouco eficiente do ponto de vista de utilização dos recursos da rede. Além disso, o número de utilizadores residenciais e empresariais a usar os serviços de *Internet* tem vindo a crescer de uma forma exponencial. Deste modo, o serviço de acesso à *Internet* exige muito em termos de recursos da rede PSTN. Assim, para que o desperdício de recursos seja o menor possível, convém que o AS (ou o NAS) esteja colocado o mais próximo possível da central de comutação.

4.2.2 Redes de acesso ADSL

No cenário anterior, a ligação do cliente ao ISP é suportada por circuitos de banda estreita analógicos ou digitais. Este tipo de circuitos não é compatível com os requisitos dos serviços de banda larga emergentes nas redes de telecomunicações, tais como, por exemplo, os serviços de distribuição de vídeo. Em primeiro lugar, os requisitos de largura de banda de alguns serviços são superiores aos disponíveis em cada circuito. Em segundo lugar, a maior parte dos serviços de banda larga são assimétricos, com maior largura de banda no sentido descendente. Deste modo, numa ligação simétrica os recursos reservados no sentido ascendente poderão ser desperdiçados.

Tendo em conta a assimetria do tráfego e aproveitando a infra-estrutura de cobre existente, apareceram novas soluções tecnológicas designadas por *x*DSL (*Digital Subscriber Line*). A tecnologia *x*DSL (ADSL - *Asymmetrical* DSL, VDSL - *Very high data rate* DSL e HDSL – *High data rate* DSL) inclui uma grande variedade de métodos de transmissão, para acesso digital ponto-a-ponto, sobre pares de fios de cobre usados inicialmente para transmitir voz. A tecnologia ADSL define uma taxa de transmissão no sentido ascendente de 16 a 640 Kb/seg, e uma taxa de transmissão no sentido descendente de 1.5 a 6 Mb/seg, e distâncias máximas da ordem de 5.5 Km às taxas mais baixas e de 3 Km às mais elevadas. A tecnologia VDSL define uma taxa de transmissão no sentido ascendente de 1.5 a 13 Mb/seg, uma taxa de transmissão no sentido descendente de 12.96 a 51.85 Mb/seg, e distâncias máximas da ordem de 1.4 Km às taxas mais baixas e de 0.3 Km às mais elevadas. As taxas de transmissão em VDSL são mais elevadas, mas as distâncias máximas são inferiores. A tecnologia HDSL define uma taxa de transmissão de 2 Mb/seg até distâncias de 4.8 Km.

Esta secção apresenta redes de acesso baseadas em ADSL (Figura 4-3). A topologia da rede continua a ser em estrela, com ligações em cobre ponto-a-ponto do DSLAM (DSL

Access Multiplexer) para cada terminal. O DSLAM é um *multiplexer* que agrega vários acessos ADSL. Esta rede de acesso permite a coexistência de serviços de banda estreita e de banda larga no mesmo par de cobre através de multiplexagem na frequência. No DSLAM, o tráfego dos diferentes serviços é separado através de *splitters*: o tráfego IP é encaminhado para o servidor de acesso de banda larga (BAS – *Broadband* AS) e o tráfego de voz é encaminhado para a central de comutação PSTN ou RDIS. O BAS é um AS de banda larga que suporta serviços com requisitos elevados de largura de banda.

Numa ligação estão envolvidos dois *modems* ADSL: o ATU-C que realiza a interface entre a terminação ADSL e o DSLAM, e o ATU-R que realiza a interface entre a terminação ADSL e o cliente.

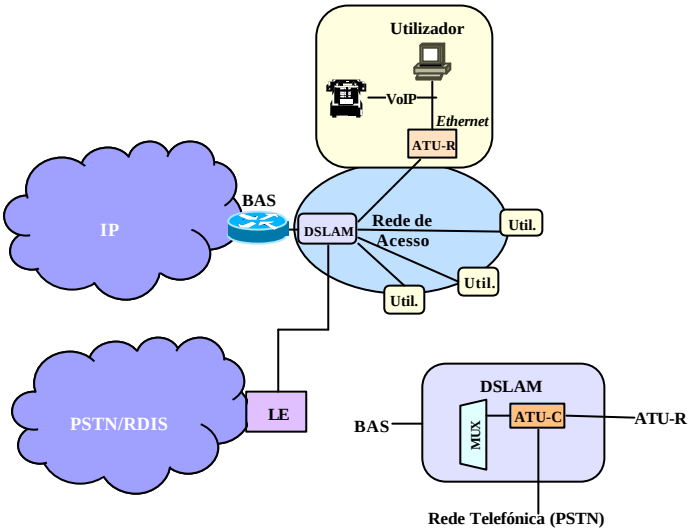


Figura 4-3 : Redes de acesso ADSL.

A tecnologia ADSL (e todas as variantes *x*DSL) apresenta várias vantagens para os operadores de rede: reutiliza a infra-estrutura existente de pares de cobre, e fornece uma largura de banda dedicada para cada cliente. No entanto, esta tecnologia tem o problema de as distâncias envolvidas imporem uma limitação na largura de banda disponível. Uma forma de ter os DSLAMs o mais próximos possível das instalações dos clientes é instalar anéis SDH (*Synchronous Digital Hierarchy*) na rede de acesso. Esta solução (Figura 4-4) permite aumentar as distâncias na rede de acesso. Assim, as ligações entre os utilizadores e os *Add Drops* dos anéis SDH são suportadas em *x*DSL sobre pares de cobre. Um *Add Drop* é um ponto de entrada no anel SDH que faz a interface entre o anel SDH e qualquer outra

tecnologia. Uma unidade de inter-funcionamento (IWU – *InterWorking Unit*) encontra-se entre o DSLAM e o *Add Drop* para "recolher" as ligações de banda estreita. Tipicamente, estas ligações *x*DSL em cobre são mais curtas dada a distribuição de pontos de acesso via *Add Drop*, sendo possível por isso utilizar *modems x*DSL com débitos mais elevados.

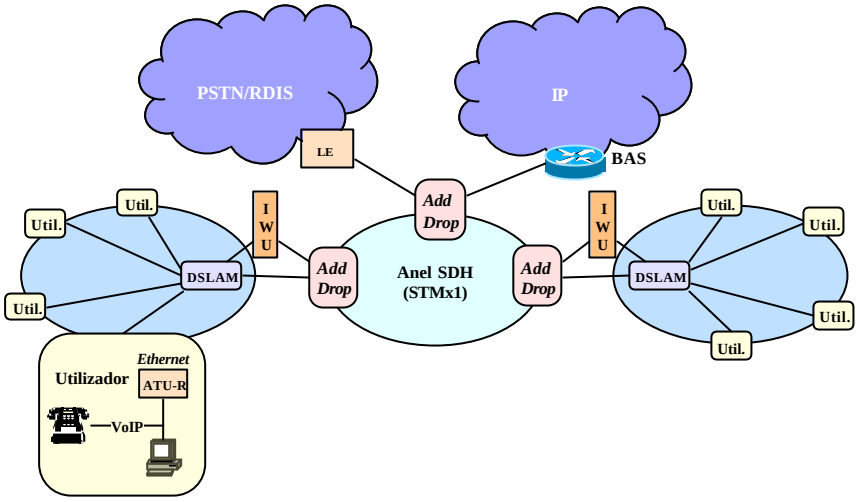


Figura 4-4 : Redes de acesso ADSL com anéis SDH.

Na rede de acesso baseada em ADSL não é possível tirar partido da multiplexagem estatística entre tráfego gerado pelos diferentes terminais, e tráfego com destino a diferentes terminais, e obter economia da largura de banda utilizada, uma vez que as ligações físicas entre os terminais e o DSLAM são ponto-a-ponto (dedicadas). Quando são usados anéis SDH, as ligações em estrela do utilizador ao DSLAM são muito mais pequenas, pois o DSLAM encontra-se próximo do *Add Drop* do anel SDH. Neste cenário é possível ter multiplexagem estatística entre as ligações no anel SDH e assim rentabilizar os recursos da rede de acesso.

4.2.3 Redes de acesso FITL com tecnologia de transporte ATM

Na rede de acesso FITL (*Fiber Into The Loop*) com tecnologia de transporte ATM podem ser usados anéis SDH primários e secundários em redes de grandes dimensões. Dependente do tipo de situação (distância dos anéis até ao cliente e quantidade de tráfego que tem de ser suportado), os anéis SDH secundários podem ser substituídos por PONs (*Passive Optical Network*) ou ligações *x*DSL directas aos clientes. Em qualquer dos casos, o troço

final até ao cliente será normalmente feito em pares de cobre por ser a solução mais económica. Uma PON é uma rede de distribuição partilhada baseada em *splitters* de fibra óptica.

Nestas redes a tecnologia de transporte é, nas implementações disponíveis, o ATM (Figura 4-5). As comunicações individuais são identificadas na rede de acesso através dos valores do VPI (*Virtual Path Identifier*) e VCI (*Virtual Channel Identifier*). A rede de acesso em questão é uma rede multi-serviços que transporta serviços de banda estreita e banda larga. Como a tecnologia de transporte é o ATM, a rede de acesso terá que ser ligada a uma estação ATM. Os serviços de banda estreita terão que ser "passados" para a rede PSTN e os pacotes IP das comunicações de banda larga transportados sobre ATM serão encaminhados para a rede IP.

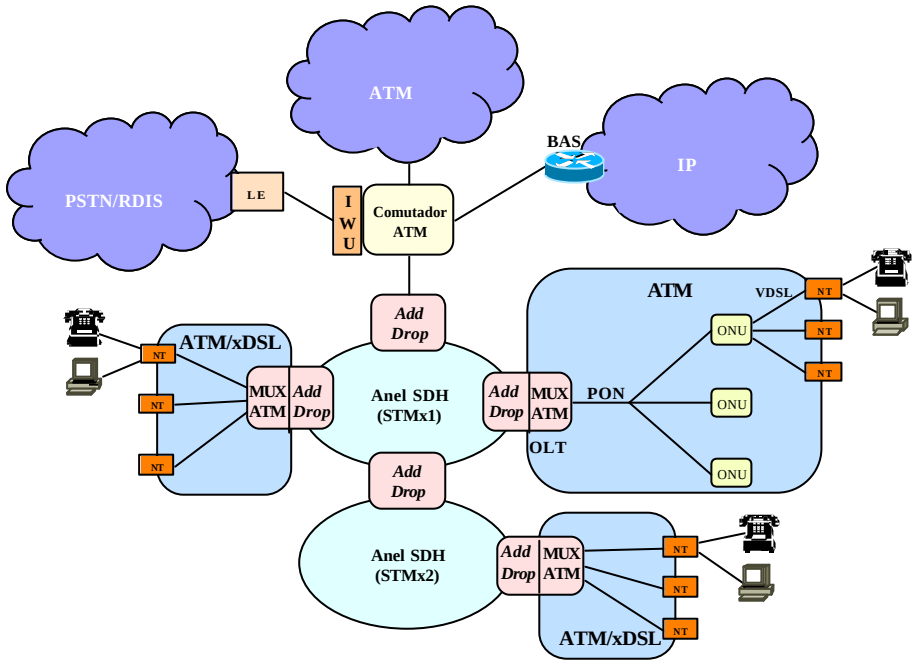


Figura 4-5 : Redes de acesso FITL com ATM como tecnologia de transporte.

A rede PSTN é ligada ao comutador ATM através de uma unidade de inter-funcionamento (IWU) que realiza a interface entre a rede ATM e a rede PSTN de banda estreita. A rede IP é ligada ao comutador ATM através do BAS que faz a interligação às várias redes de acesso ligadas através do anel SDH.

Além das redes de acesso ilustradas na Figura 4-5, é também possível ter redes de acesso mais pequenas apenas baseadas em cobre ou PONs ligadas directamente ao comutador ATM. As PONs têm sido objecto de estudo de diversos operadores de telecomunicações por serem infra-estruturas de acesso de banda larga bastante acessíveis, tendo em conta os débitos elevados que proporcionam. Estas redes vão ser detalhadas na secção seguinte.

4.2.3.1 PON

Um conjunto de operadores de todo o mundo estabeleceu um grupo, o FSAN (*Full Services Access Network*) [Quayle98], com o objectivo de normalizar um conjunto de redes de acesso (de banda larga) com as características pretendidas para o transporte de serviços de banda larga e compatível com os serviços de banda estreita já existentes. A topologia de acesso escolhida pelo FSAN e normalizada pelo ITU-T (*International Telecommunication Union – Telecommunication Standardization Sector*) nas recomendações G. 983.1, G.983.2 e G.983.3 [ITU-T98, ITU-T99 e ITU-T01] é uma PON. Em relação à tecnologia de transporte, esta pode ser qualquer tecnologia compatível com o WDM (*Wavelength Division Multiplexing*).

Um sistema PON de banda larga (Figura 4-6) consiste numa OLT (*Optical Line Termination*), várias ONUs (*Optical Network Unit*) ou ONTs (*Optical Network Termination*) e uma rede de distribuição baseada em *splitters* ópticos (ODN – *Optical Distribution Network*) [Maeda01]. A OLT é um equipamento terminal que faz a interface da rede de acesso com a rede de transporte e que concentra ligações de vários ONUs/ONTs através da ODN. A ONU é um equipamento que termina a parte óptica da rede de acesso e faz a interface com o par de cobre do lado do utilizador. O NT é um terminador de rede que funciona como equipamento de interface entre a rede de acesso (em cobre) e a rede de cliente. O ONT realiza numa só unidade as funções da ONU e do NT. Utilizam-se ONTs sempre que a (grande) dimensão da rede de cliente ou a exigência do serviço prestado o justifiquem. A rede de distribuição da ONU/ONT até ao utilizador numa rede de acesso FTTCab (*Fiber To The Cabinet*) / C (*Curb*) / B (*Building*) é em cobre, na qual se podem utilizar *modems* VDSL dada a proximidade das instalações dos clientes. Numa rede de acesso FTTH (*Fiber To The Home*) a rede de distribuição é constituída por fibra até ao cliente.

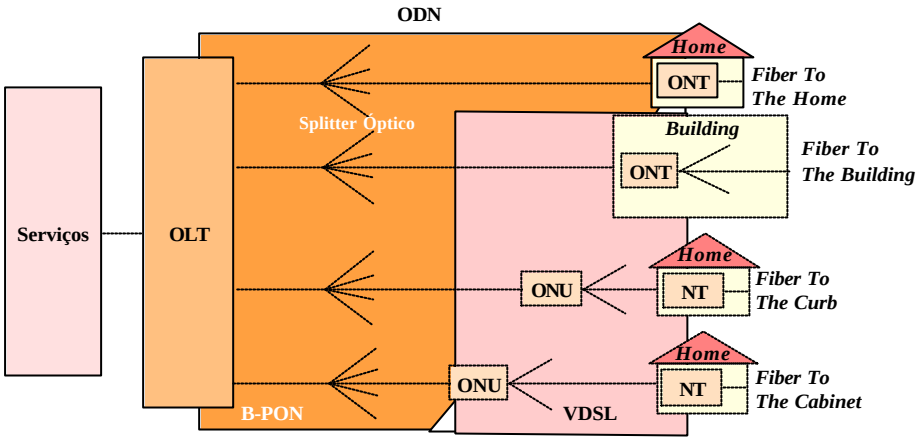


Figura 4-6 : PON na rede de acesso [Maeda01].

A tecnologia de transporte escolhida aquando da formação do FSAN foi o ATM. Actualmente, o ATM é apenas uma das tecnologias possíveis. Deste modo, o nome B-PON (*Broadband* PON) abrange as diversas tecnologias de transporte possíveis numa rede de acesso baseada em PONs.

4.2.4 Redes de distribuição CATV

Um outro tipo de redes de acesso que apresenta uma grande importância hoje em dia são as redes de distribuição de TV por cabo (CATV - *Community Antenna TeleVision*) que foram criadas inicialmente para distribuir canais de TV através de cabo coaxial. A arquitectura típica destas redes é uma arquitectura HFC (*Hybrid Fiber-Coax*) que combina a utilização de fibra óptica desde o *head end* (ponto de ligação aos SPs – *Service Providers*) até ao passeio ou edifício (*Fiber To The Curb* ou *Fiber To The Building*), e cabo coaxial desde estes pontos até ao cliente (Figura 4-7). A introdução de fibra óptica aumenta a largura de banda disponível e reduz o número de amplificadores que são necessários no percurso do *head end* até ao cliente devido à superior qualidade do sinal.

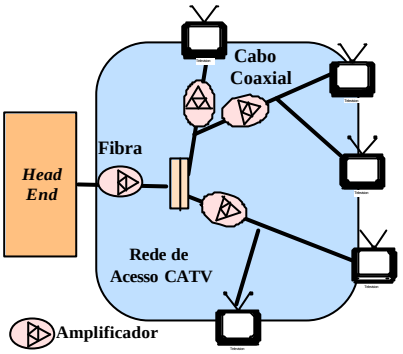


Figura 4-7 : Arquitectura da rede HFC.

Na sua origem, as redes CATV apenas permitiam transportar serviços distributivos. No entanto, as elevadas larguras de banda disponíveis nestas redes justificaram o seu aproveitamento para transportar também serviços interactivos de banda larga.

Como as redes CATV foram construídas tendo em conta o suporte de serviços distributivos, foi necessário modificá-las para o suporte de dados e de serviços interactivos. Mais especificamente, foi necessário implementar uma técnica de acesso múltiplo para a partilha do meio físico de transmissão, comum aos vários utilizadores. Outra modificação muito importante para o transporte de serviços interactivos foi a implementação do canal de retorno para cada cliente (comunicação no sentido ascendente).

As redes de distribuição de TV por cabo já instaladas e modificadas de modo a terem conectividade no sentido ascendente são, hoje em dia, capazes de suportar o tráfego de banda larga e os novos serviços multimédia e, além disso, podem ainda substituir as linhas telefónicas, através da utilização de soluções de Voz sobre IP (VoIP). Estas redes permitem oferecer uma maior largura de banda com custos aceitáveis e o serviço analógico distributivo pode ser mantido em simultâneo com a evolução para o suporte dos novos serviços de dados, vídeo e interactivos.

Existem diversos sistemas proprietários normalizados para o suporte de serviços interactivos e de dados em redes CATV. Os dois mais importantes são o DVB/RCC (*Digital Video Broadcasting / Return Control Channel*) ou ETS (*European Telecommunications Standard*) 300 800 [Ricken98] e o DOCSIS (*Data Over Cable System Interface Specification*) [Int-DOCSIS]. Estes dois sistemas foram desenvolvidos separadamente: o DVB/RCC é um sistema europeu, enquanto que o DOCSIS é americano.

Estes sistemas apresentam diferenças a nível protocolar e de arquitectura. Na Europa é também utilizado o sistema DOCSIS com algumas modificações, o EuroDOCSIS.

Nesta secção é apresentado com mais detalhe o sistema DOCSIS. A norma DOCSIS especifica as interfaces para serviços de comunicação de dados bidireccionais em redes CATV e define um serviço de transporte transparente para tráfego IP entre o CMTS (*Cable Modem Termination System*) e o CM (*Cable Modem*) representados na Figura 4-8.

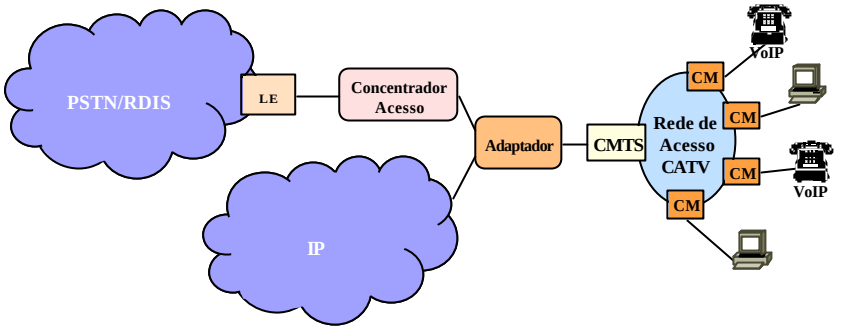


Figura 4-8 : Arquitectura de referência *Data over Cable* em DOCSIS.

Nas fronteiras da rede de acesso encontram-se o CMTS no lado do fornecedor de serviço e o CM no lado de cada utilizador. Para poder ligar o CMTS à rede IP é necessário ter um elemento de interface que realiza a adaptação entre as duas redes. A interface entre a rede PSTN/RDIS/RDIS-BL e a rede CATV é realizada por um concentrador de acesso.

A trama MAC (*Media Access Control*) é a unidade básica de transferência de dados, entre o CMTS e o CM, e suporta o transporte de diferentes tipos de protocolos, por exemplo, pacotes *Ethernet*. Além das funções transportadoras dos pacotes, a trama MAC transporta também mensagens de gestão para a operação do sistema e mensagens para sincronização dos *modems*. Por exemplo, os parâmetros de transmissão são configurados pelo CM ou CMTS através de envio de mensagens MAC. A trama MAC suporta, por exemplo, o transporte de pacotes *Ethernet*, sendo a informação de tipo de protocolo incluída no cabeçalho. As tramas MAC são transportadas no sentido descendente no campo de informação dos pacotes MPEG (*Moving Picture Experts Group*) e são misturadas com os pacotes MPEG que transportam vídeo (em TDM – *Time Division Multiplexing*). Cada CM recebe as tramas MAC e filtra a informação que é destinada aos seus utilizadores. Deste modo, os pacotes IP são transportados transparentemente na rede de CATV.

No sentido ascendente a informação é transportada em *minislots* com um formato específico definido pela norma. A informação é transportada para o CMTS em TDMA (*Time Division Multiple Access*) ou FDMA (*Frequency Division Multiple Access*). A unidade de transporte de informação neste sentido é também a trama MAC. Tal como no caso do sentido descendente, a trama *Ethernet* é inserida na trama MAC. O PDU da trama pode ser fragmentado e inserido em vários *minislots*, que são partilhados por vários utilizadores. Cada *minislot* tem um comprimento máximo de 255 *bytes*, e cada utilizador pode usar até um máximo de 255 *minislots*. Assim, o comprimento máximo da rajada enviada por um utilizador é o correspondente à informação que pode ser transportada em 255 *minislots*. O número de *minislots* atribuídos a cada utilizador é negociado no estabelecimento da sessão. No entanto, o atraso dos pacotes de cada utilizador depende da largura de banda disponível no sentido ascendente, da quantidade de informação enviada por cada utilizador e do número de utilizadores. Quanto maior for o número de utilizadores, maior será o atraso. Existem duas possibilidades para impedir que este atraso exceda um determinado limiar: limitando o número de utilizadores na rede de acesso, ou limitando o número de sessões activas através de mecanismos de controle de admissão.

A norma DOCSIS suporta classes de serviço definidas por parâmetros específicos: taxa máxima, prioridade de canal, taxa mínima garantida, comprimento máximo de rajada, etc. A QdS é suportada através de um diálogo MAC entre cada CMTS e CM. Os mecanismos de QdS suportados são ao nível do controle de admissão, considerando que o mecanismo de transporte não é limitativo. No sentido descendente associa-se uma largura de banda bem definida do CMTS ao CM. No sentido ascendente, é atribuída uma largura de banda a cada CM. No mecanismo de acesso ao meio pode ser dada uma prioridade a um CM em relação a outros, dependendo da sua QdS.

Na norma EuroDOCSIS é possível ter uma taxa de transmissão ascendente em cada canal de 0,32 a 10 Mb/seg. Na norma DVB/RCC esta taxa é de 0,256 a 6,176 Mb/seg.

4.3 O futuro das redes de acesso

Um dos trabalhos efectuados nesta Tese, descrito no capítulo 5, centra-se em redes de acesso baseadas na tecnologia ATM [Sargento99a e Sargento99b]. Este trabalho foi desenvolvido no âmbito de um projecto Europeu do programa ACTS (*Advanced Communications Technology and Services*) 038, o *BroadBand Loop* (BBL) [Andersen97],

que definiu e desenvolveu uma rede de acesso baseada em ATM com integração de serviços e suporte de QdS.

O objectivo primordial no desenvolvimento do ATM era ter aplicações suportadas em ATM nativo, com o intuito de ter o ATM como tecnologia de transporte extremo-a-extremo. No entanto esse objectivo nunca foi atingido, sendo a razão principal a não existência de aplicações de utilizador que usassem ATM nativo. Com o IP passou-se precisamente o oposto. O crescente número de aplicações que utilizam o IP, levou à necessidade de integrar este protocolo nas redes baseadas em ATM através da utilização de soluções de IP sobre ATM. Uma arquitectura deste tipo foi já desenvolvida e testada em diversos cenários diferentes, por exemplo, no projecto Europeu do programa ACTS, o *Broadband Trial Integration* (BTI) [Azcorra98]. No entanto, uma solução de IP sobre ATM não é muito efectiva do ponto de vista da redução na utilização dos recursos da rede, devido ao *overhead* introduzido pelas células ATM. Haverá ainda uma redução adicional na utilização dependente do AAL que seja usado (normalmente o AAL5) para o transporte do tráfego IP. Além disso, a criação, gestão e terminação de circuitos virtuais em redes de grandes dimensões requer o protocolo de sinalização VB5.2 [ETS-V5.2], que é extremamente complexo e envolve muitas funcionalidades por parte de todos os elementos da rede de acesso e de transporte.

No futuro, uma solução muito mais atractiva será ter uma rede de acesso com IP como tecnologia de transporte. Uma solução com IP directamente sobre fibra é a solução mais desejada, mas não é ainda possível actualmente. Nesta solução, a tecnologia de transporte na PON seria o IP. Uma outra solução de rede de acesso IP implementável actualmente é uma solução baseada em anéis SDH com ligações de cobre dos anéis até aos clientes. No capítulo 6, este tipo de redes de acesso com tecnologia IP vai ser descrito e estudado com maior detalhe. Mais concretamente, é proposta uma arquitectura para redes de acesso IP, com suporte de QdS diferenciada, integrando na mesma infra-estrutura de rede um conjunto variado de serviços e aplicações, incluindo os serviços multimédia [Sargento01b].

CAPÍTULO 5

DIMENSIONAMENTO DE REDES DE ACESSO ATM

No capítulo 4 apresentou-se uma breve descrição das redes de acesso baseadas na tecnologia ATM (*Asynchronous Transfer Mode*). O trabalho apresentado neste capítulo concentra-se nas metodologias de dimensionamento de redes de acesso ATM. O dimensionamento de uma rede ATM é determinado pela topologia da rede, pelas estratégias de gestão de recursos e pelas características dos serviços suportados [Valadas98]. As estratégias de gestão de recursos implementadas na rede dependem de factores relacionados com o custo da mesma. A sinalização na rede aumenta o custo do sistema mas permite implementar estratégias de gestão de recursos mais eficientes. O custo da rede e as estratégias de gestão de recursos implementadas dependem dos elementos em que são implementados mecanismos de CAC (*Call Admission Control*) e de UPC (*Usage Parameter Control*).

No caso concreto da rede de acesso em estudo, considera-se que os recursos são geridos através do uso de VPs (*Virtual Paths*). Estes VPs podem ter como origem e destino, respectivamente a origem e o destino das chamadas, sendo neste caso designados por VPCs (*Virtual Path Connections*) extremo-a-extremo, ou podem existir diversos VPCs no percurso de uma chamada. Estas soluções diferentes impõem requisitos diferentes na

sinalização. Em VPCs extremo-a-extremo, a sinalização necessária é mínima, enquanto que com vários VPCs entre a origem e o destino é possível ter um aproveitamento mais eficiente dos recursos, mas a carga de sinalização na rede aumenta.

Neste estudo de dimensionamento considera-se que o encaminhamento é fixo. O dimensionamento da rede é efectuado a dois níveis: dimensionamento ao nível da chamada e ao nível da célula. O dimensionamento ao nível da chamada determina o número de VCCs (*Virtual Channel Connections*) necessários em cada VPC para que seja atingido um determinado grau de serviço (GdS – *Grade of Service*). O GdS corresponde, neste caso, à probabilidade de bloqueio dos VCCs. O dimensionamento ao nível da célula determina a largura de banda que é necessário atribuir a cada VCC para obter uma determinada QdS (Qualidade de Serviço) ao nível da célula. A QdS ao nível da célula define, no caso específico deste trabalho, o rácio de perdas de células admissível, a CLR (*Cell Loss Ratio*). Nos estudos de dimensionamento considera-se a sobreposição de várias redes lógicas numa rede física. Numa primeira fase determina-se a capacidade de cada VPC, ou seja, dimensionam-se as redes lógicas. Seguidamente determina-se a largura de banda necessária em cada ligação da rede física para suportar as diversas redes lógicas existentes.

Este capítulo apresenta metodologias de dimensionamento de redes de acesso ATM, considerando diferentes estratégias de gestão de recursos. A secção 5.1 apresenta a arquitectura da rede de acesso baseada em ATM e a secção 5.2 faz uma tipificação dos serviços a suportar. Na secção 5.3 são propostas as estratégias de gestão de recursos que vão ser estudadas no dimensionamento da rede de acesso. As secções 5.4 e 5.5 apresentam, respectivamente, os métodos de dimensionamento ao nível da chamada e da célula. Com o intuito de ilustrar os métodos de dimensionamento, são considerados diversos casos de estudo na secção 5.7 que incluem cenários residenciais, empresariais e mistos. As conclusões são apresentadas na secção 5.8.

5.1 Arquitectura da rede de acesso

A arquitectura da rede global, rede de transporte e rede de acesso, que vai servir de base ao dimensionamento, é representada na Figura 5-1. Esta rede de acesso é um caso particular das redes de acesso apresentadas no capítulo 4 na secção sobre redes de acesso FITL (*Fiber Into The Loop*), em que é usada uma PON (*Passive Optical Network*) entre a rede de transporte e a rede de cliente [Quayle97] e [Andersen97].

Considera-se que a rede de transporte é baseada num anel SDH (*Synchronous Digital Hierarchy*) com um ou mais comutadores ATM ligados aos fornecedores de serviço (SPs). No caso geral, o sistema pode incluir várias redes de acesso e o anel SDH fornece conectividade entre as redes de acesso e os comutadores ATM. A rede de acesso é ligada aos SPs através de contentores VC-4 SDH.

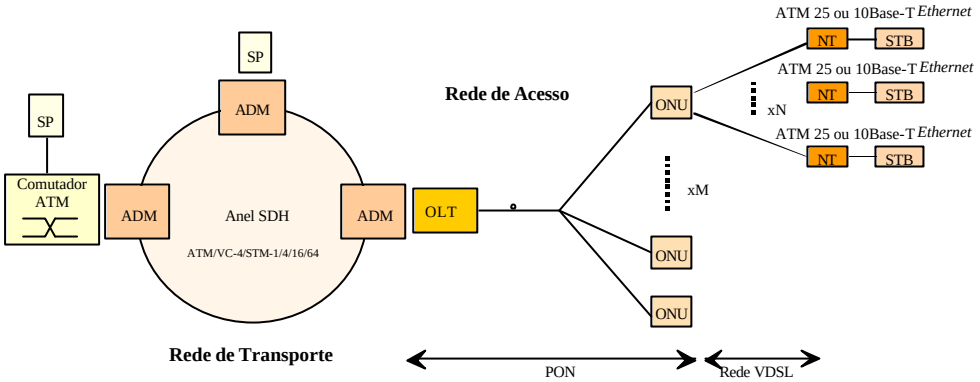


Figura 5-1 : Arquitectura da rede.

Cada rede de acesso consiste numa rede de distribuição de fibra óptica (PON) com uma OLT (*Optical Line Termination*) e com ONUs (*Optical Network Units*) nas terminações dos clientes. Cada ONU é capaz de servir vários NTs (*Network Terminations*). A tecnologia de transmissão na rede de distribuição entre a ONU e os NTs numa rede de acesso FTTCab/C/B (*Fiber To The Cabinet/Curb/Building)* é VDSL (*Very high data rate Digital Subscriber Line*). Numa rede de acesso FTTH (FTT *Home*), a ligação entre a OLT e o cliente é implementada através de fibra [Huish96]. A interface física entre o NT e o cliente pode ser ATM a 25 Mb/seg ou *Ethernet* a 10 ou 100 Mb/seg.

No sentido descendente da PON, ou seja, da OLT para as ONUs, a informação é transportada em banda base. Esta informação é difundida para todas as ONUs através de *splitters* ópticos nas fibras da PON. As ONUs recebem o fluxo de tráfego e filtram as células que contenham os valores de VPI (*Virtual Path Identifier*) e VCI (*Virtual Channel Identifier*) correspondentes a chamadas de utilizadores a elas ligados. Deste modo, a largura de banda atribuída a cada ONU é configurável ao nível do ATM. No sentido ascendente, das ONUs para a OLT, a transmissão da informação é efectuada através de técnicas de acesso múltiplo SCMA (*Sub-Carrier Multiple Access*) [Andersen97] ou TDMA (*Time Division Multiple Access*) [PONB97].

Com o aumento do número de utilizadores e a introdução de novos serviços pode haver necessidade de aumentar a largura de banda da PON. Uma das grandes vantagens destas redes de acesso é a possibilidade de serem inseridos módulos adicionais nas ligações ascendentes e descendentes entre cada ONU e a OLT sempre que a quantidade de tráfego que atravessa essas ligações o justifique. No caso concreto da rede de acesso desenvolvida no âmbito do projecto *BroadbandLoop* (BBL) [Andersen97] do programa Europeu ACTS (AC0038), a transmissão da informação colocada nos módulos adicionais é baseada em SCM (*Sub-Carrier Multiplexing*), e deste modo, a OLT coloca a informação em portadoras, cada uma com uma frequência diferente correspondente a cada ONU (Figura 5-2). No sentido ascendente, a transmissão de informação nos módulos adicionais, assim como nos módulos base, é também baseada em SCM. Cada ONU tem direito a um módulo base, que ocupa uma banda de frequência não partilhada, para transmitir a informação dos seus utilizadores e, se necessário, tem direito a um módulo adicional.

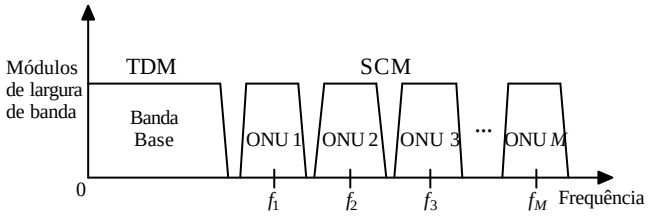


Figura 5-2 : Transmissão de informação em banda base e em SCM no sentido descendente.

Um dos objectivos do dimensionamento é (consoante a quantidade de utilizadores de cada serviço e os seus requisitos) determinar se é necessário introduzir módulos adicionais na rede.

Na rede de acesso considera-se que existem dois tipos diferentes de ONUs que diferem quanto à largura de banda das ligações VDSL. As ONUs simétricas estão ligadas a linhas VDSL com a mesma largura de banda nos dois sentidos. As ONUs assimétricas estão ligadas a linhas VDSL com maior largura de banda no sentido descendente que no sentido ascendente. Estas soluções (assimétrica/simétrica) foram definidas principalmente por questões de custo.

5.2 Caracterização dos serviços

Os serviços são tipificados nesta Tese em comutados *versus* permanentes, conversacionais *versus* distributivos, e simétricos *versus* assimétricos. Nos serviços comutados os VCCs são estabelecidos a pedido pelo plano de controle do ATM, enquanto que nos serviços permanentes os VCCs entre os extremos estão sempre disponíveis e configurados ao nível do plano de gestão. Nos serviços distributivos os VCCs são pedidos apenas por um utilizador ligado à PON para um destino exterior. Nos serviços conversacionais os VCCs podem também ser pedidos entre utilizadores da mesma PON. Nos serviços simétricos os VCCs necessitam dos mesmos recursos nos dois sentidos da chamada, e nos serviços assimétricos são necessários mais recursos no sentido descendente do que no sentido ascendente. Geralmente, os serviços residenciais são todos assimétricos e os serviços empresariais são maioritariamente simétricos. Consequentemente, uma ONU simétrica é adequada para utilizadores empresariais enquanto que uma ONU assimétrica é mais apropriada para utilizadores residenciais.

No contexto dos casos de estudo que serão apresentados na secção 5.7 deste capítulo, foi seleccionado um conjunto de serviços residenciais e empresariais [Sargento98]. Os serviços residenciais são o serviço de acesso rápido à *Internet*, VoD MPEG-2 (*Video-on-Demand* com MPEG-2) e de distribuição de vídeo MPEG-2 (DVB - *Digital Video Broadcast* com MPEG-2). Os serviços de acesso à *Internet* e VoD são comutados, distributivos e assimétricos. O serviço de distribuição de vídeo é permanente, distributivo e assimétrico. Os serviços empresariais são o acesso rápido à *Internet*, videoconferência e o serviço de interligação de LANs (*Local Area Networks*). O serviço de videoconferência é comutado, conversacional, e simétrico, e o serviço de interligação de LANs é permanente, conversacional, e simétrico. [Menendez97] considerou também os seguintes serviços para as redes de acesso: videotelefonia, ensino à distância, telemedicina, POTS/RDIS e ligações virtuais comutadas (SVCs – *Switched Virtual Connections*) ATM.

5.3 Estratégias de gestão de recursos

Existem duas questões importantes relacionadas com a gestão de recursos: a forma como os VPCs são configurados e a forma como os VCCs (permanentes ou comutados) são multiplexados. Em relação à configuração dos VPCs, uma consideração que tem um

impacto directo no custo do sistema é se são ou não implementados mecanismos de controle de tráfego e congestionamento (CAC e UPC) nos elementos da rede. Nesta Tese são propostas três estratégias diferentes de arquitectura dos VPCs e de multiplexagem dos VCCs:

? Estratégia I - VPCs entre cada ONU e um Comutador ATM no Exterior (CAE). Esta estratégia (Figura 5-3 (a)) pode ser usada quando a OLT e as ONUs não implementam funções de CAC. Cada terminal tem de possuir um VCI diferente para permitir que as suas células sejam identificadas no CAE. Assume-se que existem funções de UPC nas ONUs, funções estas que são configuráveis ao nível do plano de gestão, para impedir que tráfego fora do perfil contratado influencie a QdS do restante tráfego. Nesta estratégia, o tráfego dos VCCs pertencentes a terminais de uma mesma ONU são multiplexados no mesmo VPC. Como os VPCs são extremo-a-extremo entre cada ONU e o CAE, a OLT é um *cross-connect* ATM, isto é, não comuta VCCs. Considerando um caso concreto de uma PON com 16 ONUs, em que os utilizadores de cada ONU acedem a 3 serviços diferentes, e em que são configurados VPCs por serviço, o número de VPCs entre cada ONU e a OLT é de 3, e o número de VPCs na ligação entre a OLT e o CAE é de 3? 16 = 48.

? Estratégia II - VPCs entre cada ONU e a OLT (Figura 5-3 (b)) e um VPC entre a OLT e o CAE para suportar as comunicações com o exterior. Esta estratégia é usada quando são implementadas funções de CAC na OLT. Cada terminal tem de possuir um VCI diferente para permitir que as suas células sejam identificadas na OLT. Assim como na estratégia anterior, assume-se que a ONU implementa funções de UPC. Em relação à multiplexagem de VCCs, como existe comutação na OLT, existe multiplexagem no mesmo VPC entre a OLT e a ONU dos VCCs com origem ou destino numa mesma ONU, e no mesmo VPC entre o CAE e a OLT dos VCCs com origem ou destino numa mesma PON. Esta maior multiplexagem permite aumentar a partilha de recursos da rede. Considerando o mesmo caso concreto da estratégia anterior, o número de VPCs entre cada ONU e a OLT é de 3, e o número de VPCs na ligação entre a OLT e o CAE é também de 3, pois o tráfego pertencente ao mesmo serviço de ONUs diferentes é multiplexado no mesmo VPC.

? Estratégia III - VPCs entre cada par de ONUs (Figura 5-3 (c)) e VPCs entre cada ONU e o CAE para suportar as comunicações para o exterior. Esta estratégia apenas

é possível quando são implementadas funções de CAC (e de UPC) apenas nas ONUs. Nesta estratégia, assim como na estratégia I, a OLT actua apenas como um *cross-connect*. Assim, existe apenas multiplexagem entre VCCs de terminais diferentes em cada ONU, mas não existe multiplexagem de VCCs com origem ou destino em ONUs diferentes na OLT. Considerando o mesmo caso concreto das estratégias anteriores, o número de VPCs entre cada ONU e a OLT é de 16?3 = 48, devido à existência de VPCs extremo-a-extremo, e o número de VPCs na ligação entre a OLT e o CAE é também de 16?3 = 48.

Outra alternativa é ter funções de CAC implementadas nas ONUs e na OLT. Nesta situação, existem VPCs entre cada ONU e a OLT e um VPC entre a OLT e o comutador ATM no exterior, tal como na estratégia II de gestão de recursos, mas com a diferença de que o tráfego entre utilizadores da mesma ONU é comutado na própria ONU. Os procedimentos de dimensionamento que serão utilizados na estratégia II podem ser facilmente adaptados a este caso, e não são por isso alvo deste estudo.

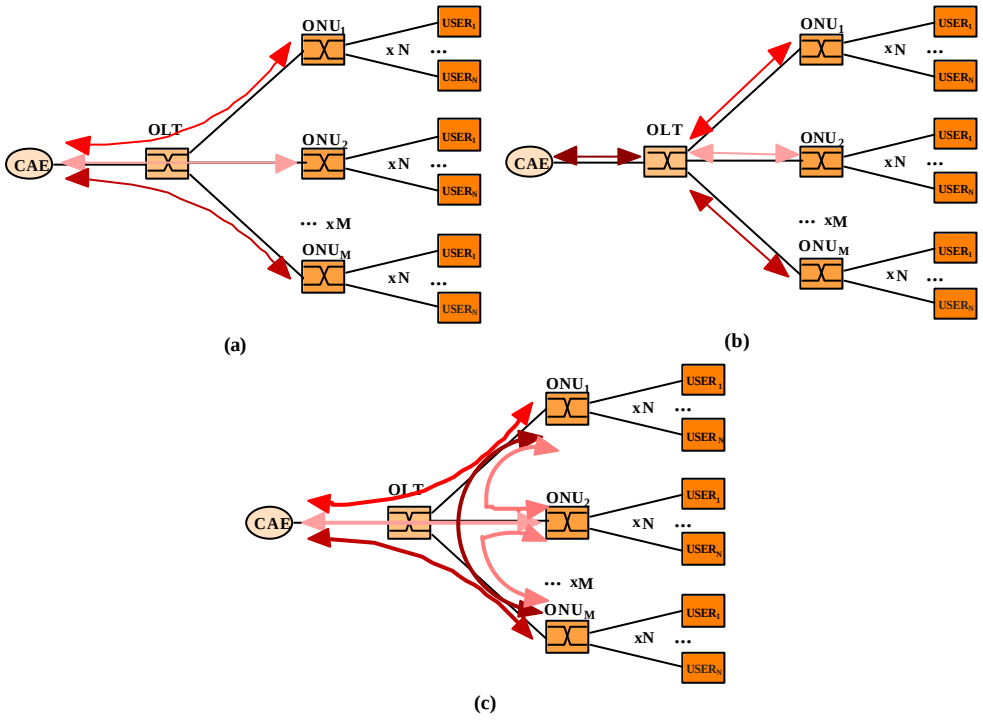


Figura 5-3 : VPCs entre os elementos da rede de acesso e o comutador ATM: (a) VPCs entre cada ONU e um comutador ATM no exterior, (b) VPCs entre cada ONU e a OLT, e (c) VPCs entre cada par de ONUs.

Em relação à multiplexagem de VCCs referida, os elementos da rede de acesso (ONU, OLT) podem ser configurados com uma fila de espera por classe de serviço. Esta solução permite estabelecer uma rede lógica com recursos reservados por classe de serviço (com qualquer uma das estratégias de gestão de VPCs apresentadas anteriormente). Embora esta situação de ter uma fila de espera por classe de serviço facilite a gestão de recursos, existem outras alternativas que podem ser mais vantajosas. Assim, vão ser considerados os diferentes casos de ter: (i) uma fila de espera por classe de serviço (ii) uma fila de espera para um determinado conjunto de classes de serviço, ou (iii) uma fila de espera para todas as classes de serviço.

5.4 Dimensionamento ao nível da chamada

O dimensionamento ao nível da chamada calcula o número de VCCs necessários em cada VPC para fornecer um conjunto de serviços com um determinado GdS e, através desta informação, determina o número de VCCs necessários em cada ligação física. Neste estudo de dimensionamento considera-se como parâmetro de GdS a probabilidade de bloqueio de chamadas.

O sistema pode ser modelado ao nível da chamada por um processo de nascimento e morte multi-dimensional, em que o estado do sistema corresponde ao número de chamadas activas em cada sessão [Ross95]. Uma sessão de tráfego é caracterizada pelo par origem/destino, pelo percurso, pela largura de banda de cada chamada e pela intensidade de tráfego. Os pares origem/destino podem ser, no caso concreto da rede de acesso, duas ONUs, no caso de o tráfego ser interno à PON, ou uma ONU e o CAE, no caso de o tráfego ser externo à PON. Nos casos de estudo que serão considerados na secção 5.7 assumem-se duas condições de simetria: (i) a percentagem de tráfego que sai da PON é igual à percentagem de tráfego que entra na PON; e (ii) a percentagem de tráfego que sai de uma ONU é igual à percentagem de tráfego que entra na mesma ONU.

Como será descrito nas secções seguintes, em alguns casos, o modelo permite o cálculo das probabilidades de bloqueio exactas (por exemplo, através da fórmula de *Erlang B*). Noutros casos, a complexidade associada ao cálculo da probabilidade de bloqueio através de processos de nascimento e morte multi-dimensionais implica que seja necessário recorrer a aproximações de carga reduzida, tanto para redes com um serviço como para redes multi-serviço. Note-se que as probabilidades de bloqueio poderiam ser sempre

calculadas através da determinação dos estados possíveis do processo de nascimento e morte multi-dimensional e do cálculo das respectivas probabilidades limite. Este método de cálculo não foi considerado por questões de eficiência computacional.

Neste estudo de dimensionamento vão ser consideradas duas situações extremas relacionadas com a forma de combinação dos serviços nas filas de espera de saída dos nós em que é efectuada multiplexagem estatística: (i) uma fila de espera por serviço, em que existe apenas multiplexagem estatística entre as células pertencentes ao mesmo serviço, situação que é denominada de segregação de serviços, e (ii) uma fila de espera para todos os serviços, em que existe multiplexagem estatística entre as células de todos os serviços, situação que é denominada de agregação de serviços. A primeira e segunda situações foram endereçadas respectivamente em [Sargento99a] e [Sargento99b]. Na primeira situação os algoritmos de CAC são mais simples, mas a largura de banda necessária em cada ligação poderá ser ligeiramente superior, porque existe menos tráfego a ser multiplexado na fila de espera. Nesta situação existe um VPC por serviço, enquanto que na segunda situação, o mesmo VPC transporta todos os serviços. Uma solução intermédia seria ter várias filas de espera e um número de serviços superior ao número de filas de espera. Neste caso teriam de ser consideradas regras para agregar serviços em filas de espera [Mocci94]. Estas regras incluem, por exemplo, a separação de tráfego que contenha fluxos com larguras de banda ou durações separadas de mais do que uma ordem de grandeza, e a prevenção de situações em que exista um grande volume de tráfego que tenha uma QdS muito superior à necessária, apenas porque alguns fluxos nesse volume de tráfego requerem essa QdS.

Na prática, os utilizadores são arbitrariamente distribuídos pelas ONUs da PON. O impacto desta distribuição em relação ao dimensionamento vai ser estudado através da consideração de dois casos extremos: (i) PON com o número máximo possível de ONUs e com os utilizadores igualmente distribuídos pelas ONUs, que se designa por *maxONUs* (neste caso, a largura de banda nas ligações entre as ONUs e a OLT é mínima, mas o número de ONUs necessárias é máximo), e (ii) PON com o número mínimo de ONUs possível, e com os utilizadores concentrados nestas ONUs, que se designa por *minONUs* (neste caso, a largura de banda necessária nas ligações entre as ONUs e a OLT é máxima, mas o número necessário de ONUs é mínimo). Nesta situação existem dois tipos de ONUs:

as ONUs que têm a si ligados o número máximo de utilizadores possível (tipo 1), e as ONUs que apenas têm a si ligados alguns utilizadores (tipo 2).

5.4.1 Modelo do tráfego oferecido

Considere-se que M é o número de ONUs da PON, M_{max} é o número máximo de ONUs suportadas pela PON e N_{max} é o número máximo de utilizadores suportados por ONU. Assim, dentro de cada PON pode existir um máximo de $M_{max}N_{max}$ utilizadores. Supondo que os utilizadores podem aceder a S serviços, em que $? : = \{1,...,S\}$ é o conjunto dos serviços, considera-se que apenas uma percentagem $?^s$ dos utilizadores subscrevem um determinado serviço s ? ?. Esta percentagem $?^s$ é denominada de factor de penetração do serviço. O número efectivo de utilizadores na PON a subscrever um determinado serviço s ? ? é $??^s M_{max}N_{max}?$, em que $?x?$ é o menor inteiro superior a x.

Assume-se que todos os utilizadores exteriores à PON estão ligados ao CAE. Considera-se que as chamadas do serviço s chegam ao sistema de acordo com um processo de *Poisson* com taxa $?^s$ e têm um tempo de permanência exponencialmente distribuído de média $1/?^s$. Desta forma, a intensidade do tráfego oferecido é dada por $?^s = ?^s/?^s$.

Considera-se que $? : = \{1,...,M\}$ é o conjunto das ONUs. Numa situação *minONUs* existem dois tipos de ONUs: do tipo 1 e do tipo 2. Assim, definem-se os sub-conjuntos $?_1$ e $?_2$ ($?_1$? ? e $?_2$? ?), que são, respectivamente, os conjuntos das ONUs do tipo 1 e do tipo 2. Considera-se também que ? representa a OLT e ? representa o exterior.

Neste capítulo adoptou-se uma notação menos compacta do que seria possível, com o objectivo de tornar mais clara a formulação do problema. Assim, considera-se que as ligações, físicas ou lógicas (VPCs), são representadas por um par origem i e destino j, em que i e j podem pertencer a ?, $?_1$ ou $?_2$ ou ser iguais a ? ou ?. As sessões de tráfego são também representadas por um par origem l e destino m, em que l e m podem pertencer a ?, $?_1$ ou $?_2$ ou ser iguais a ?. As ligações do interior da PON (entre as ONUs e a OLT no sentido ascendente e entre a OLT e as ONUs no sentido descendente) vão ser designadas por ligações interiores, e as ligações do exterior da PON (entre a OLT e o CAE no sentido ascendente e entre o CAE e a OLT no sentido descendente) vão ser designadas por ligações exteriores. As ligações exteriores são constituídas pelos contentores VC-4 do anel SDH.

Para poder proceder ao dimensionamento da rede assume-se que: (i) cada utilizador no interior da PON gera um tráfego por serviço s de intensidade λ^s, (ii) um factor α do tráfego oferecido por cada utilizador tem como destino os utilizadores internos da PON (e um factor 1-α tem como destino o exterior), (iii) o tráfego interno oferecido por cada utilizador é repartido uniformemente por todos os destinos possíveis, e (iv) o tráfego oferecido pelo CAE à PON é igual ao tráfego oferecido pela PON ao CAE.

Na situação denominada de *maxONUs*, o número de ONUs é $M = M_{max}$ e o número de utilizadores por ONU é $N^s = \lfloor N_{max}^s \rfloor$ para um determinado serviço. Cada utilizador tem um total de N^sM-1 destinos internos, dos quais N^s-1 são da própria ONU e $(M\text{-}1)N^s$ são destinos de outras ONUs. Resulta então

$$\lambda_{ij}^s = \alpha \lambda^s N^s \left\lfloor \frac{N^s - 1}{N^s M - 1} \right\rfloor, \quad i, j = 1, \ldots, M, \quad i = j \tag{5-1}$$

$$\lambda_{ij}^s = \alpha \lambda^s N^s \left\lfloor \frac{N^s}{N^s M - 1} \right\rfloor, \quad i, j = 1, \ldots, M, \quad i \neq j \tag{5-2}$$

$$\lambda_{ie}^s = \alpha \lambda^s N^s \left\lfloor 1 - \frac{1}{N^s M - 1} \right\rfloor, \quad i = 1, \ldots, M, \quad e = 1, \ldots \tag{5-3}$$

em que λ_{ij}^s, $i = j$, é o tráfego gerado por uma ONU que tem como destino utilizadores da mesma ONU, λ_{ij}^s, $i \neq j$, é o tráfego gerado por uma ONU que tem como destino outra ONU na mesma PON, e λ_{ie}^s é o tráfego gerado por uma ONU que tem como destino o exterior.

O tráfego total do serviço s gerado por cada ONU i que tem como destino o interior da PON é $\sum_{j=1}^{M} \lambda_{ij}^s = \alpha \lambda^s N^s$ e o tráfego que tem como destino utilizadores exteriores à PON é $\lambda_{ie}^s = \lambda^s N^s (1\text{-}\alpha)$. Considerando, como já foi referido atrás, que a quantidade de tráfego oferecido à PON pelo CAE é igual à quantidade de tráfego oferecido pela PON ao CAE, o tráfego total do serviço s oferecido à PON é $\lambda^s N^s M \alpha + 2\lambda^s N^s M (1\text{-}\alpha)$.

Na situação denominada de *minONUs*, o tráfego é distribuído pelo número mínimo de ONUs possível, isto é, algumas ONUs contêm o número máximo possível de utilizadores e uma ONU contém o número restante de utilizadores. Assim, existem dois tipos diferentes de ONUs: as ONUs que têm a si ligados o número máximo de utilizadores possível (tipo 1), e as ONUs que apenas têm a si ligados alguns utilizadores (tipo 2). Note-

se que nesta situação existe apenas uma ONU do tipo 2 em cada PON para o serviço *s*. As ONUs do tipo 1 têm $N_1^s = N_{max}$ utilizadores; as ONUs do tipo 2 têm $N_2^s < N_{max}$ utilizadores. O número de ONUs da PON é neste caso inferior ao número máximo admissível

$$M^s = \left\lfloor \frac{M_{max} N_{max}}{N_{max}} \right\rfloor \tag{5-4}$$

Como cada utilizador de uma ONU do tipo 1 tem um total de $(M^s\text{-}1)\ N_1^s + N_2^s\text{-}1$ destinos, dos quais $N_1^s\text{-}1$ são pertencentes à mesma ONU, N_1^s são pertencentes a uma outra ONU do tipo 1, e N_2^s são pertencentes à ONU do tipo 2, resulta que no caso das ONUs do tipo 1

$$\lambda_{i_1 j_1}^s = \lambda^s N_1^s \left(\frac{N_1^s - 1}{(M^s - 1) N_1^s + N_2^s - 1} \right), \quad i_1, j_1 \in \aleph_1, \ i_1 \neq j_1 \tag{5-5}$$

$$\lambda_{i_1 j_1}^s = \lambda^s N_1^s \left(\frac{N_1^s}{(M^s - 1) N_1^s + N_2^s - 1} \right), \quad i_1, j_1 \in \aleph_1, \ i_1 \neq j_1 \tag{5-6}$$

$$\lambda_{i_1 j_2}^s = \lambda^s N_1^s \left(\frac{N_2^s}{(M^s - 1) N_1^s + N_2^s - 1} \right), \quad i_1 \in \aleph_1, \ j_2 \in \aleph_2, \ i_1 \neq j_2 \tag{5-7}$$

$$\lambda_{i_1 e}^s = \lambda^s N_1^s \left(1 - \frac{1}{M^s} \right), \quad i_1 \in \aleph_1, \ e \notin \aleph \tag{5-8}$$

O tráfego total do serviço *s* gerado por cada ONU do tipo 1 que tem como destinos os utilizadores da mesma PON é $\sum_{j_1 \neq 1}^{M^s - 1} \lambda_{i_1 j_1}^s = \lambda_{i_1 j_1}^s + \lambda_{i_1 j_2}^s + \lambda^s N_1^s$, e o tráfego que tem como destino utilizadores exteriores à PON é $\lambda_{i_1 e}^s = \lambda^s N_1^s \left(1 - \frac{1}{M^s} \right)$.

Considerando a ONU do tipo 2, $N_2^s\text{-}1$ são utilizadores destino pertencentes à mesma ONU e N_1^s são utilizadores destino pertencentes a uma ONU do tipo 1. Assim

$$\lambda_{i_2 j_2}^s = \lambda^s N_2^s \left(\frac{N_2^s - 1}{(M^s - 1) N_1^s + N_2^s - 1} \right), \quad i_2, j_2 \in \aleph_2, \ i_2 \neq j_2 \tag{5-9}$$

$$\lambda_{i_2 j_1}^s = \lambda^s N_2^s \left(\frac{N_1^s}{(M^s - 1) N_1^s + N_2^s - 1} \right), \quad i_2 \in \aleph_2, \ j_1 \in \aleph_1, \ i_2 \neq j_1 \tag{5-10}$$

$$\lambda_{i_2 e}^s = \lambda^s N_2^s \left(1 - \frac{1}{M^s} \right), \quad i_2 \in \aleph_2, \ e \notin \aleph \tag{5-11}$$

O tráfego total do serviço *s* que tem como destino os utilizadores da mesma PON e que é gerado por ONUs do tipo 2 é $\sum_{j_1=1}^{M^s-1} \rho^s_{i_2 j_1} \rho \rho^s_{i_2 j_2} \rho \rho^s N_2^s \rho$, e o tráfego que tem como destino utilizadores exteriores à PON é $\rho^s_{i_2 e} \rho \rho^s N_2^s \rho 1 \rho \rho$.

O tráfego do serviço *s* gerado pelas ONUs de tipo 1 que tem como destino os utilizadores exteriores à PON é $\rho^s N_1^s$ (1-ρ) e o gerado pela ONU do tipo 2 é $\rho^s N_2^s$ (1-ρ). Assim, o tráfego total do serviço *s* gerado pela PON que tem como destino os utilizadores exteriores à PON é $\rho^s N_1^s$ (M^s-1)(1-ρ) + $\rho^s N_2^s$ (1-ρ). Esta quantidade de tráfego é idêntica à da situação *maxONUs*. Considerando os mesmos pressupostos da situação *maxONUs,* o tráfego total do serviço *s* oferecido à PON é $\rho^s N_1^s$ (M^s-1)ρ^s + $\rho^s N_2^s \rho \rho$+ 2$\rho^s N_1^s$ (M^s-1)(1-ρ) + 2$\rho^s N_2^s$ (1-ρ).

5.4.2 Estratégia I - VPs entre cada ONU e um comutador ATM no exterior

Nesta estratégia de gestão de recursos os VPCs são construídos extremo-a-extremo, entre a ONU e o comutador ATM no exterior. Deste modo, existem apenas dois tipos de VPCs na rede de acesso: VPCs no sentido ascendente e no sentido descendente.

5.4.2.1 Segregação de serviços

Nesta secção assume-se que existe um VPC por serviço. Assim, não existe multiplexagem entre tráfego de serviços diferentes.

5.4.2.1.1 Serviços distributivos

A Figura 5-4 ilustra as diferentes sessões de tráfego de cada serviço num cenário com serviços distributivos. Todos os VPCs são extremo-a-extremo, isto é, cada sessão de tráfego percorre apenas um VPC entre a origem e o destino e, por isso, o dimensionamento ao nível da chamada pode ser baseado na fórmula de *Erlang B*. As probabilidades de bloqueio das sessões de tráfego de cada serviço *s* numa situação *maxONUs* são então dadas por

$$B^s_{ie} \rho ER\left(\rho^s_{ie}, v^s_{ie}\right), i \rho \forall e \rho \forall \quad (5\text{-}12)$$

$$B^s_{ej} \rho ER\left(\rho^s_{ej}, v^s_{ej}\right), j \rho \forall e \rho \forall \quad (5\text{-}13)$$

em que B_{ie}^{s} e B_{ej}^{s} são as probabilidades de bloqueio das sessões de tráfego pertencentes ao serviço *s*, $?_{ie}^{s}$ e $?_{ej}^{s}$ representam a intensidade de tráfego do serviço *s* definidos na secção anterior, e v_{ie}^{s} e v_{ej}^{s} representam o número de VCCs atravessados pelas sessões do serviço *s*. *ER*[.,.] representa a fórmula de *Erlang B*, isto é,

$$ER[?, v] = \frac{\frac{?^{v}}{v!}}{\sum_{k=0}^{v} \frac{?^{k}}{k!}} \tag{5-14}$$

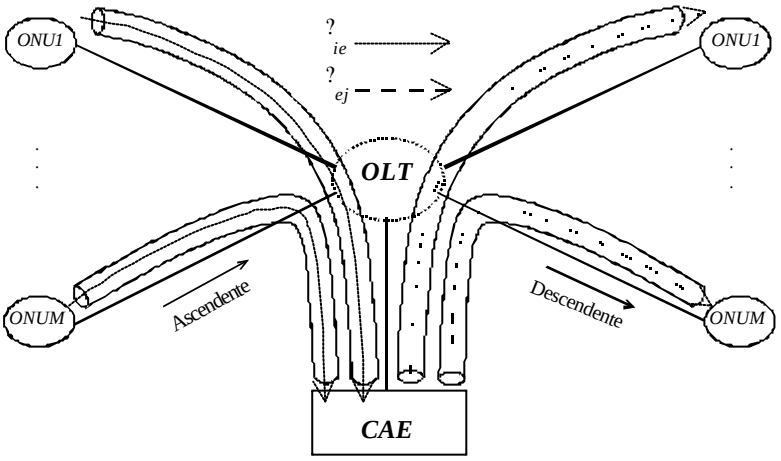


Figura 5-4 : Sessões de tráfego – estratégia I, serviços distributivos.

O número de VCCs nas ligações interiores é dado por

$$V_{it} = \sum_{s} v_{ie}^{s}, \; i = 1, 2, \ldots, t = 1, 2, \ldots, e = 1, 2, \ldots \tag{5-15}$$

$$V_{tj} = \sum_{s} v_{ej}^{s}, \; j = 1, 2, \ldots, t = 1, 2, \ldots, e = 1, 2, \ldots \tag{5-16}$$

O número de VCCs nas ligações exteriores é dado por

$$V_{te} = \sum_{s} \sum_{i=1}^{M} v_{ie}^{s}, \; i = 1, 2, \ldots, t = 1, 2, \ldots, e = 1, 2, \ldots \tag{5-17}$$

$$V_{et} = \sum_{s} \sum_{j=1}^{M} v_{ej}^{s}, \; j = 1, 2, \ldots, t = 1, 2, \ldots, e = 1, 2, \ldots \tag{5-18}$$

Numa situação *minONUs*, o tráfego que é gerado por cada ONU varia consoante a ONU é do tipo 1 ou 2. Assim, existem dois tipos de VPCs em cada sentido para cada serviço *s*: um correspondente às ONUs do tipo 1, com número de VCCs $v^s_{i_1e}$ ou $v^s_{ej_1}$, e em que a intensidade de tráfego que atravessa cada VPC é $?^s_{i_1e}$ ou $?^s_{ej_1}$; e outro correspondente à ONU do tipo 2, com número de VCCs $v^s_{i_2e}$ ou $v^s_{ej_2}$, e em que a intensidade de tráfego que atravessa cada VPC é $?^s_{i_2e}$ ou $?^s_{ej_2}$. As probabilidades de bloqueio das sessões de tráfego do serviço *s* são, neste caso, obtidas pelas seguintes expressões:

$$B^s_{i_1e} \; ? \; ER\left(?^s_{i_1e}, v^s_{i_1e}\right), \; i_1 \; ? \; \mathbb{Z}_1, \; e \; ? \; \mathbb{Z} \tag{5-19}$$

$$B^s_{ej_1} \; ? \; ER\left(?^s_{ej_1}, v^s_{ej_1}\right), \; j_1 \; ? \; \mathbb{Z}_1, \; e \; ? \; \mathbb{Z} \tag{5-20}$$

$$B^s_{i_2e} \; ? \; ER\left(?^s_{i_2e}, v^s_{i_2e}\right), \; i_2 \; ? \; \mathbb{Z}_2, \; e \; ? \; \mathbb{Z} \tag{5-21}$$

$$B^s_{ej_2} \; ? \; ER\left(?^s_{ej_2}, v^s_{ej_2}\right), \; j_2 \; ? \; \mathbb{Z}_2, \; e \; ? \; \mathbb{Z} \tag{5-22}$$

O número de VCCs necessários nas ligações interiores é dado por

$$V_{i_1t} \; ? \; \sum_{s? \mathbb{Z}} v^s_{i_1e}, \; i_1 \; ? \; \mathbb{Z}_1, \; t \; ? \; \mathbb{Z}, \; e \; ? \; \mathbb{Z} \tag{5-23}$$

$$V_{i_2t} \; ? \; \sum_{s? \mathbb{Z}} v^s_{i_2e}, \; i_2 \; ? \; \mathbb{Z}_2, \; t \; ? \; \mathbb{Z}, \; e \; ? \; \mathbb{Z} \tag{5-24}$$

$$V_{tj_1} \; ? \; \sum_{s? \mathbb{Z}} v^s_{ej_1}, \; j_1 \; ? \; \mathbb{Z}_1, \; t \; ? \; \mathbb{Z}, \; e \; ? \; \mathbb{Z} \tag{5-25}$$

$$V_{tj_2} \; ? \; \sum_{s? \mathbb{Z}} v^s_{ej_2}, \; j_2 \; ? \; \mathbb{Z}_2, \; t \; ? \; \mathbb{Z}, \; e \; ? \; \mathbb{Z} \tag{5-26}$$

O número de VCCs necessários nas ligações exteriores é dado por

$$V_{te} \; ? \; \sum_{s? \mathbb{Z}} \sum_{i_1?1}^{M^s?1} v^s_{i_1e} \; ? \; \sum_{s? \mathbb{Z}} v^s_{i_2e}, \; i_1 \; ? \; \mathbb{Z}_1, \; i_2 \; ? \; \mathbb{Z}_2, \; t \; ? \; \mathbb{Z}, \; e \; ? \; \mathbb{Z} \tag{5-27}$$

$$V_{et} \; ? \; \sum_{s? \mathbb{Z}} \sum_{j_1?1}^{M^s?1} v^s_{ej_1} \; ? \; \sum_{s? \mathbb{Z}} v^s_{ej_2}, \; j_1 \; ? \; \mathbb{Z}_1, \; j_2 \; ? \; \mathbb{Z}_2, \; t \; ? \; \mathbb{Z}, \; e \; ? \; \mathbb{Z} \tag{5-28}$$

5.4.2.1.2 Serviços conversacionais

No cenário com serviços conversacionais, as sessões de tráfego que têm como origem e destino utilizadores da mesma PON são comutadas no comutador ATM, sendo cada uma transportada em dois VPCs sucessivos, um no sentido ascendente e outro no sentido descendente, como é ilustrado na Figura 5-5. O tráfego para o exterior ou do exterior apenas é transportado pelo VPC no sentido ascendente ou pelo VPC no sentido descendente, respectivamente. Como se observa na Figura 5-5, o tráfego do mesmo serviço gerado por cada ONU pode pertencer a 3 tipos de sessões de tráfego diferentes: sessões com origem e destino na mesma ONU, com origem e destino em ONUs diferentes pertencentes à mesma PON, e com origem e destino em PONs diferentes; a mesma situação acontece em relação ao tráfego que tem como destino uma determinada ONU. A fórmula de *Erlang B* não pode ser usada porque as sessões de tráfego atravessam mais do que um VPC. Sendo assim, é necessário recorrer a uma aproximação de carga reduzida [Ross95] para efectuar o dimensionamento ao nível da chamada.

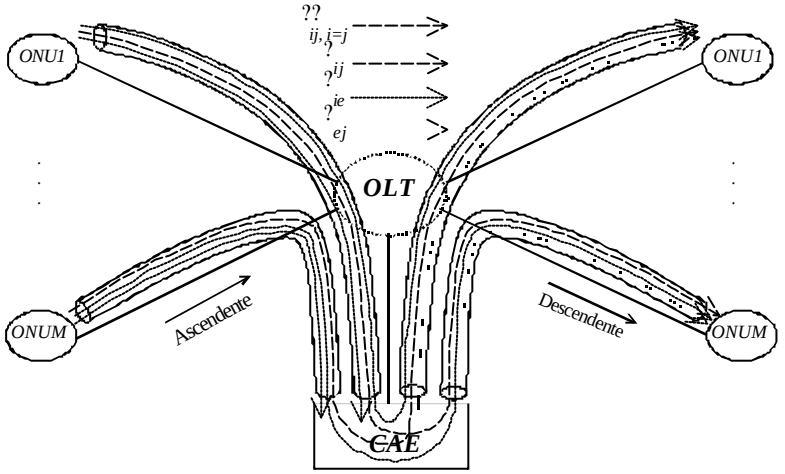


Figura 5-5 : Sessões de tráfego – estratégia I, serviços conversacionais.

As equações gerais de cálculo da probabilidade de bloqueio pela aproximação de carga reduzida são apresentadas no anexo A. Como as sessões de tráfego que partilham um VPC pertencem ao mesmo serviço *s*, é considerada a aproximação de carga reduzida num sistema de serviço único (secção A.4).

Considerando a situação *maxONUs*, as probabilidades de bloqueio aproximadas dos diferentes tipos de sessões de tráfego são dadas por

$$B_{ij}^{s} = 1 - (1 - L_{ie}^{s})(1 - L_{ej}^{s}),\ i,\ j = \overline{1, M},\ e = \overline{1, E} \tag{5-29}$$

$$B_{ie}^{s} = L_{ie}^{s},\ i = \overline{1, M},\ e = \overline{1, E} \tag{5-30}$$

$$B_{ej}^{s} = L_{ej}^{s},\ j = \overline{1, M},\ e = \overline{1, E} \tag{5-31}$$

em que L_{ie}^{s} e L_{ej}^{s} representam as probabilidades de bloqueio dos VPCs que transportam as sessões de tráfego do serviço *s*, respectivamente, nas ligações nos sentidos ascendente e descendente. Estas probabilidades são calculadas através do seguinte conjunto de equações de ponto fixo:

$$L_{ie}^{s} = ER\left(\sum_{j=1}^{M} \left[\gamma_{ij}^{s}\left(1 - L_{ej}^{s}\right)\right], \gamma_{ie}^{s}, v_{ie}^{s}\right),\ i,\ j = \overline{1, M},\ e = \overline{1, E} \tag{5-32}$$

$$L_{ej}^{s} = ER\left(\sum_{i=1}^{M} \left[\gamma_{ij}^{s}\left(1 - L_{ie}^{s}\right)\right], \gamma_{ej}^{s}, v_{ej}^{s}\right),\ i,\ j = \overline{1, M},\ e = \overline{1, E} \tag{5-33}$$

O número de VCCs necessários nas ligações interiores e exteriores é dado por

$$V_{it} = \sum_{s=\overline{1,S}} v_{ie}^{s},\ i = \overline{1, M},\ t = \overline{1, T},\ e = \overline{1, E} \tag{5-34}$$

$$V_{tj} = \sum_{s=\overline{1,S}} v_{ej}^{s},\ j = \overline{1, M},\ t = \overline{1, T},\ e = \overline{1, E} \tag{5-35}$$

$$V_{te} = \sum_{s=\overline{1,S}} \sum_{i=1}^{M} v_{ie}^{s},\ i = \overline{1, M},\ t = \overline{1, T},\ e = \overline{1, E} \tag{5-36}$$

$$V_{et} = \sum_{s=\overline{1,S}} \sum_{j=1}^{M} v_{ej}^{s},\ j = \overline{1, M},\ t = \overline{1, T},\ e = \overline{1, E} \tag{5-37}$$

Numa situação *minONUs* existem, tal como no caso dos serviços distributivos, dois tipos de VPCs, correspondentes às ONUs do tipo 1 e do tipo 2. As probabilidades aproximadas de bloqueio podem, neste caso, ser calculadas através das seguintes expressões:

$$B_{i_1 j_1}^{s} = 1 - (1 - L_{i_1 e}^{s})(1 - L_{e j_1}^{s}),\ i_1,\ j_1 = \overline{1, M_1},\ e = \overline{1, E} \tag{5-38}$$

$$B_{i_1 j_2}^{s} = 1 - (1 - L_{i_1 e}^{s})(1 - L_{e j_2}^{s}),\ i_1 = \overline{1, M_1},\ j_2 = \overline{1, M_2},\ e = \overline{1, E} \tag{5-39}$$

$$B_{i_2 j_1}^{s} = 1 - (1 - L_{i_2 e}^{s})(1 - L_{e j_1}^{s}),\ i_2 = \overline{1, M_2},\ j_1 = \overline{1, M_1},\ e = \overline{1, E} \tag{5-40}$$

$$B^s_{i_2 j_2} = 1 - \left(1 - L^s_{i_2 e}\right)\left(1 - L^s_{e j_2}\right), \; i_2, j_2 \neq \varepsilon_2, e \neq \varepsilon \tag{5-41}$$

$$B^s_{i_1 e} = L^s_{i_1 e}, \; i_1 \neq \varepsilon_1, e \neq \varepsilon \tag{5-42}$$

$$B^s_{i_2 e} = L^s_{i_2 e}, \; i_2 \neq \varepsilon_2, e \neq \varepsilon \tag{5-43}$$

$$B^s_{e j_1} = L^s_{e j_1}, \; j_1 \neq \varepsilon_1, e \neq \varepsilon \tag{5-44}$$

$$B^s_{e j_2} = L^s_{e j_2}, \; j_2 \neq \varepsilon_2, e \neq \varepsilon \tag{5-45}$$

em que as probabilidades de bloqueio dos VPCs são dadas por

$$L^s_{i_1 e} = ER\left(\sum_{j_1=1}^{M^s} \left(1 - \left(1 - L^s_{e j_1}\right)\left(1 - B^s_{i_1 j_1}\right)\right), v^s_{i_1 e}\right), \quad i_1, j_1 \neq \varepsilon_1, j_2 \neq \varepsilon_2, e \neq \varepsilon \tag{5-46}$$

$$L^s_{i_2 e} = ER\left(\sum_{j_1=1}^{M^s} \left(1 - \left(1 - L^s_{e j_1}\right)\left(1 - B^s_{i_2 j_1}\right)\right), v^s_{i_2 e}\right), \quad i_2, j_2 \neq \varepsilon_2, j_1 \neq \varepsilon_1, e \neq \varepsilon \tag{5-47}$$

$$L^s_{e j_1} = ER\left(\sum_{i_1=1}^{M^s} \left(1 - \left(1 - L^s_{i_1 e}\right)\left(1 - B^s_{i_2 j_1}\right)\right), v^s_{e j_1}\right), \quad i_1, j_1 \neq \varepsilon_1, i_2 \neq \varepsilon_2, e \neq \varepsilon \tag{5-48}$$

$$L^s_{e j_2} = ER\left(\sum_{i_1=1}^{M^s} \left(1 - \left(1 - L^s_{i_1 e}\right)\left(1 - B^s_{i_2 j_2}\right)\right), v^s_{e j_2}\right), \quad i_1 \neq \varepsilon_1, i_2, j_2 \neq \varepsilon_2, e \neq \varepsilon \tag{5-49}$$

O número de VCCs necessários em cada ligação é obtido pelas mesmas expressões que as apresentadas nos serviços distributivos (5-23 a 5-28).

5.4.2.2 Agregação de serviços

Na secção anterior considerou-se que existia um VPC por serviço. No entanto, o caso geral é ter vários serviços agrupados no mesmo VPC. O agrupamento tem de obedecer a um determinado conjunto de regras já descritas em 5.4. Nesta secção vai ser estudada uma situação extrema em que todos os serviços são agregados, e o seu tráfego é multiplexado no mesmo VPC.

5.4.2.2.1 Serviços distributivos

Nesta situação não é possível utilizar a fórmula de *Erlang B* apresentada anteriormente, pois esta apenas se aplica a um serviço. Quando existe mais do que um serviço multiplexado no mesmo VPC é necessário recorrer a outras metodologias de cálculo das probabilidades de bloqueio [Nilsson99], [Kaufman81], [Labourdett92] e [Siebenhaar95]. Nesta secção recorre-se ao método directo de *Knapsack* [Kaufman81]. Este método encontra-se descrito no anexo A (secção A.2). As probabilidades de bloqueio são obtidas pelas seguintes expressões:

$$B_{ie}^{s} = \sum_{c=c_{it}-b^{s}+1}^{c_{it}} q(\rho_{ie}, b^{s}, c) \quad i = 1, t = 1, e = 1 \tag{5-50}$$

$$B_{ej}^{s} = \sum_{c=c_{tj}-b^{s}+1}^{c_{tj}} q(\rho_{ej}^{s}, b^{s}, c) \quad j = 1, t = 1, e = 1 \tag{5-51}$$

em que c_{it} e c_{tj} são as capacidades dos VPCs interiores, b^s é a largura de banda do serviço *s*, e a função $q(\rho, b, c)$ está definida no anexo A (secção A.2). As capacidades das ligações interiores e exteriores são dadas por

$$C_{it} = c_{it}, \quad i = 1, t = 1 \tag{5-52}$$

$$C_{tj} = c_{tj}, \quad j = 1, t = 1 \tag{5-53}$$

$$C_{te} = \sum_{i=1}^{M} c_{it}, \quad i = 1, t = 1, e = 1 \tag{5-54}$$

$$C_{et} = \sum_{j=1}^{M} c_{tj}, \quad j = 1, t = 1, e = 1 \tag{5-55}$$

Numa situação *minONUs*, o método de cálculo das probabilidades de bloqueio é semelhante.

5.4.2.2.2 Serviços conversacionais

Na situação em que existe um VPC apenas entre a ONU e o comutador ATM para transportar todos os serviços, a aproximação de carga reduzida usada na secção 5.4.2.1 não pode ser aplicada. Assim, é necessário usar generalizações da aproximação de carga reduzida para sistemas multi-serviço, como por exemplo, as aproximações de *Knapsack* [Chung93] e de *Erlang* [Kelly86]. Neste capítulo recorre-se à aproximação de carga

reduzida de *Erlang* para redes multi-serviços. O método de cálculo das probabilidades de bloqueio e as fórmulas gerais associadas são apresentadas no anexo A (secção A.5). As probabilidades aproximadas de bloqueio de uma determinada sessão de tráfego num determinado VPC são obtidas pelas seguintes expressões:

$$L^s_{ie,ij} = Q^s_{ij}(\rho^s_{ij}+1, L^s_{ej,ij}, v_{it}, b^s),\ i, j = \overline{1,M},\ t = \overline{1,T},\ e = \overline{1,E} \tag{5-56}$$

$$L^s_{ie,ie} = Q^s_{ie}(\rho^s_{ie}, v_{it}, b^s),\ i = \overline{1,M},\ t = \overline{1,T},\ e = \overline{1,E} \tag{5-57}$$

$$L^s_{ej,ij} = Q^s_{ij}(\rho^s_{ij}+1, L^s_{ie,ij}, v_{tj}, b^s),\ i, j = \overline{1,M},\ t = \overline{1,T},\ e = \overline{1,E} \tag{5-58}$$

$$L^s_{ej,ej} = Q^s_{ej}(\rho^s_{ej}, v_{tj}, b^s),\ j = \overline{1,M},\ t = \overline{1,T},\ e = \overline{1,E} \tag{5-59}$$

em que $L^s_{ij,ml}$ é a probabilidade de bloqueio da sessão de tráfego pertencente ao serviço *s*, com origem em m e destino em l, no VPC ascendente com origem em i e destino em j. A função $Q(\rho, v, b)$ está definida no anexo A (secção A.5).

As probabilidades aproximadas de bloqueio dos diferentes tipos de sessões de tráfego do serviço *s* são obtidas pelas seguintes equações:

$$B^s_{ij} = 1 - (1 - L^s_{ie,ij})(1 - L^s_{ej,ij}),\ i, j = \overline{1,M},\ e = \overline{1,E} \tag{5-60}$$

$$B^s_{ie} = L^s_{ie,ie},\ i = \overline{1,M},\ e = \overline{1,E} \tag{5-61}$$

$$B^s_{ej} = L^s_{ej,ej},\ j = \overline{1,M},\ e = \overline{1,E} \tag{5-62}$$

Nesta situação, como existe apenas um VPC em cada ligação no interior da PON, o número de VCCs nos VPCs das ligações interiores é dado por

$$V_{it} = v_{it},\ i = \overline{1,M},\ t = \overline{1,T} \tag{5-63}$$

$$V_{tj} = v_{tj},\ j = \overline{1,M},\ t = \overline{1,T} \tag{5-64}$$

O número de VCCs nas ligações exteriores é dado por

$$V_{te} = \sum_{i=1}^{M} v_{it},\ i = \overline{1,M},\ t = \overline{1,T},\ e = \overline{1,E} \tag{5-65}$$

$$V_{et} = \sum_{j=1}^{M} v_{tj},\ j = \overline{1,M},\ t = \overline{1,T},\ e = \overline{1,E} \tag{5-66}$$

Numa situação *minONUs* o método de cálculo é semelhante.

5.4.3 Estratégia II - VPs entre cada ONU e a OLT

Nesta estratégia a OLT possui funcionalidades de CAC e de UPC. Deste modo, a OLT pode terminar e iniciar VPCs. Assim, podem ser considerados quatro tipos de VPCs: interiores no sentido ascendente; interiores no sentido descendente; exteriores no sentido ascendente; e exteriores no sentido descendente. Todas as sessões de tráfego são comutadas na OLT e atravessam dois VPCs no seu trajecto entre a ONU e o comutador ATM, ou entre ONUs. A Figura 5-6 ilustra os VPCs e os tipos de sessões de tráfego existentes nesta estratégia no caso de serviços conversacionais. As sessões de tráfego correspondentes a serviços distributivos podem ser consideradas como casos particulares das sessões apresentadas na figura. Nestes serviços apenas existem os tráfegos $?^{s}_{ie}$ e $?^{s}_{ej}$ que também atravessam dois VPCs distintos.

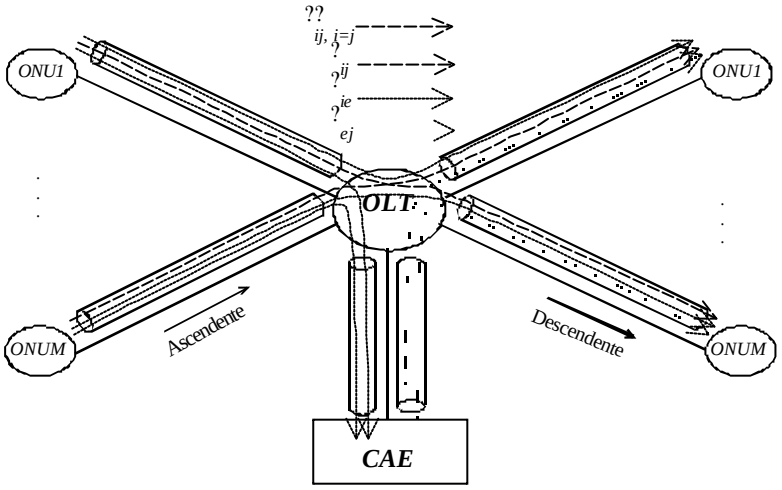


Figura 5-6 : Sessões de tráfego – estratégia II, serviços conversacionais.

5.4.3.1 Segregação de serviços

5.4.3.1.1 Serviços distributivos

No caso dos serviços distributivos, a topologia lógica da rede é reduzida a uma árvore, tanto no sentido ascendente como descendente. Nesta situação particular, o dimensionamento ao nível da chamada pode ser baseado num algoritmo de convolução que fornece o valor exacto para a probabilidade de bloqueio de cada sessão de tráfego do serviço *s*, baseado no *Knapsack* estocástico generalizado [Tsang90]. Este algoritmo é

apresentado no anexo A (secção A.3). As probabilidades de bloqueio são dadas pelas seguintes expressões:

$$B^s_{ie} = G(v^s_{te}, v^s_{it}, \ldots, v^s_{ie}, \ldots), \quad i = 1, \ldots, t = 1, \ldots, e = 1, \ldots \tag{5-67}$$

$$B^s_{ej} = G(v^s_{et}, v^s_{tj}, \ldots, v^s_{ej}, \ldots), \quad j = 1, \ldots, t = 1, \ldots, e = 1, \ldots \tag{5-68}$$

em que $G(V, v, \ldots)$ é uma função que se encontra definida no anexo A (secção A.3). Note-se que v^s_{te} e v^s_{et} representam os VCCs necessários em cada VPC do "tronco" da árvore, respectivamente, nos sentidos ascendente e descendente, e v^s_{it} e v^s_{tj} representam os VCCs dos VPCs dos "ramos" da árvore. Considerando esta estratégia, o número de VCCs no conjunto de VPCs é dado por

$$V_{it} = \sum_{s=1}^{\ldots} v^s_{it}, \quad i = 1, \ldots, t = 1, \ldots \tag{5-69}$$

$$V_{tj} = \sum_{s=1}^{\ldots} v^s_{tj}, \quad j = 1, \ldots, t = 1, \ldots \tag{5-70}$$

$$V_{te} = \sum_{s=1}^{\ldots} v^s_{te}, \quad t = 1, \ldots, e = 1, \ldots \tag{5-71}$$

$$V_{et} = \sum_{s=1}^{\ldots} v^s_{et}, \quad t = 1, \ldots, e = 1, \ldots \tag{5-72}$$

Numa situação *minONUs* a topologia da rede também se reduz a uma árvore. Pode ser aplicado o mesmo algoritmo convolucional nesta situação; o "ramo" ligado à ONU do tipo 2 terá um tráfego e uma capacidade inferiores aos dos "ramos" ligados às ONUs do tipo 1.

As probabilidades de bloqueio aproximadas dos diferentes tipos de sessões de tráfego, nesta rede com topologia em árvore, podem também ser determinadas por uma aproximação de carga reduzida. Através de um estudo comparativo efectuado, verificou-se que as probabilidades de bloqueio obtidas pelos dois métodos são semelhantes, mas que os tempos computacionais são muito diferentes. Através da aplicação dos dois métodos a dois cenários diferentes, obtiveram-se tempos computacionais do método convolucional aproximadamente 300 vezes superiores aos da aproximação de carga reduzida. Dado que os resultados obtidos pelos dois métodos são semelhantes, optou-se por utilizar a aproximação de carga reduzida nos casos de estudo apresentados na secção 5.7.

5.4.3.1.2 Serviços conversacionais

No caso de serviços conversacionais, a topologia já não é em árvore, porque neste caso existe tráfego entre utilizadores da mesma PON e, assim sendo, este tráfego atravessa os "ramos" da suposta árvore, mas não atravessa o "tronco" da mesma. Assim, nesta situação as probabilidades de bloqueio aproximadas são determinadas por aproximações de carga reduzida. Estas probabilidades são obtidas pelas seguintes expressões:

$$B_{ij}^{s} = 1 - \left(1 - L_{it}^{s}\right)\left(1 - L_{tj}^{s}\right),\ i,\ j \neq t \tag{5-73}$$

$$B_{ie}^{s} = 1 - \left(1 - L_{it}^{s}\right)\left(1 - L_{te}^{s}\right),\ i \neq t,\ e \neq t \tag{5-74}$$

$$B_{ej}^{s} = 1 - \left(1 - L_{tj}^{s}\right)\left(1 - L_{et}^{s}\right),\ j \neq t,\ e \neq t \tag{5-75}$$

em que as probabilidades de bloqueio dos VPCs são dadas pelas seguintes equações de ponto fixo:

$$L_{it}^{s} = ER\left(\sum_{j=1}^{M} a_{ij}^{s}\left(1 - L_{tj}^{s}\right) + a_{ie}^{s}\left(1 - L_{te}^{s}\right), v_{it}^{s}\right),\ i,\ j \neq t,\ e \neq t \tag{5-76}$$

$$L_{tj}^{s} = ER\left(\sum_{i=1}^{M} a_{ij}^{s}\left(1 - L_{it}^{s}\right) + a_{ej}^{s}\left(1 - L_{et}^{s}\right), v_{tj}^{s}\right),\ i,\ j \neq t,\ e \neq t \tag{5-77}$$

$$L_{te}^{s} = ER\left(\sum_{i=1}^{M} a_{ie}^{s}\left(1 - L_{it}^{s}\right), v_{te}^{s}\right),\ i \neq t,\ e \neq t \tag{5-78}$$

$$L_{et}^{s} = ER\left(\sum_{j=1}^{M} a_{ej}^{s}\left(1 - L_{tj}^{s}\right), v_{et}^{s}\right),\ j \neq t,\ e \neq t \tag{5-79}$$

O número de VCCs necessários nas ligações interiores e exteriores é dado pelas mesmas expressões apresentadas para os serviços distributivos.

Numa situação *minONUs*, o método de cálculo do número de VCCs é semelhante.

5.4.3.2 Agregação de serviços

Nesta situação, assim como na secção 5.4.2.2, recorre-se à aproximação de carga reduzida de *Erlang* para redes multi-serviço. Como exemplo, apresentam-se as expressões aproximadas das probabilidades de bloqueio, considerando serviços conversacionais. Um cenário apenas com serviços distributivos é um caso particular deste cenário. As probabilidades aproximadas de bloqueio são obtidas pelas seguintes expressões:

$$L^s_{it,ij} = Q^s_{ij}(\rho^s_{ij}, 1) \otimes L^s_{tj,ij}(v_{it}, b^s), \; i, \; j \in \mathcal{E}, t \in \mathcal{E} \tag{5-80}$$

$$L^s_{it,ie} = Q^s_{ie}(\rho^s_{ie}, 1) \otimes L^s_{te,ie}(v_{it}, b^s), \; i \in \mathcal{E}, \; t \in \mathcal{E}, e \in \mathcal{E} \tag{5-81}$$

$$L^s_{tj,ij} = Q^s_{ij}(\rho^s_{ij}, 1) \otimes L^s_{it,ij}(v_{tj}, b^s), \; i, \; j \in \mathcal{E}, t \in \mathcal{E} \tag{5-82}$$

$$L^s_{tj,ej} = Q^s_{ej}(\rho^s_{ej}, 1) \otimes L^s_{et,ej}(v_{tj}, b^s), \; j \in \mathcal{E}, t \in \mathcal{E}, e \in \mathcal{E} \tag{5-83}$$

$$L^s_{te,ie} = Q^s_{ie}(\rho^s_{ie}, 1) \otimes L^s_{it,ie}(v_{te}, b^s), i \in \mathcal{E}, \; t \in \mathcal{E}, e \in \mathcal{E} \tag{5-84}$$

$$L^s_{et,ej} = Q^s_{ej}(\rho^s_{ej}, 1) \otimes L^s_{tj,ej}(v_{et}, b^s), \; j \in \mathcal{E}, \; t \in \mathcal{E}, \; e \in \mathcal{E} \tag{5-85}$$

As probabilidades de bloqueio dos diferentes tipos de sessões de tráfego do serviço *s* são obtidas pelas seguintes equações:

$$B^s_{ij} = 1 - (1 - L^s_{it,ij}) \cdot (1 - L^s_{tj,ij}), \; i, \; j \in \mathcal{E}, t \in \mathcal{E} \tag{5-86}$$

$$B^s_{ie} = 1 - (1 - L^s_{it,ie}) \cdot (1 - L^s_{te,ie}), i \in \mathcal{E}, \; t \in \mathcal{E}, e \in \mathcal{E} \tag{5-87}$$

$$B^s_{ej} = 1 - (1 - L^s_{tj,ej}) \cdot (1 - L^s_{et,ej}), \; j \in \mathcal{E}, t \in \mathcal{E}, e \in \mathcal{E} \tag{5-88}$$

O número de VCCs nas ligações interiores e exteriores é dado por

$$V_{it} = v_{it}, i \in \mathcal{E}, t \in \mathcal{E} \tag{5-89}$$

$$V_{tj} = v_{tj}, \; j \in \mathcal{E}, t \in \mathcal{E} \tag{5-90}$$

$$V_{te} = v_{te}, e \in \mathcal{E}, t \in \mathcal{E} \tag{5-91}$$

$$V_{et} = v_{et}, e \in \mathcal{E}, t \in \mathcal{E} \tag{5-92}$$

A situação *minONUs* utiliza métodos de dimensionamento semelhantes.

5.4.4 Estratégia III - VPs entre cada par de ONUs

Nesta estratégia considera-se que são implementadas funções de CAC (e de UPC) nas ONUs. Na estratégia I a ONU actuava apenas como agregador de tráfego no sentido ascendente e distribuidor no sentido descendente. Nesta estratégia a ONU possui funcionalidades de CAC que lhe permitem encaminhar o tráfego directamente para o seu destino. A OLT actua apenas como um *cross-connect*. Note-se que, nesta estratégia, a OLT possui as mesmas funções que na estratégia I. Devido à existência de funcionalidades de CAC na ONU, o tráfego entre utilizadores da mesma PON, mas de diferentes ONUs, é

comutado na OLT. Este tráfego atravessa apenas as ligações entre a ONU e a OLT nos sentidos ascendente e descendente. Assim, todos os VPCs são extremo-a-extremo, mas o tráfego entre utilizadores da mesma ONU é comutado na própria ONU. Nesta estratégia existem dois tipos de VPCs: VPCs interiores entre cada ONU, e VPCs exteriores entre cada ONU e o comutador ATM. Estes VPCs estão ilustrados na Figura 5-7. As sessões de tráfego entre utilizadores da mesma ONU não são apresentadas, pois estas são comutadas na ONU e não atravessam os VPCs apresentados.

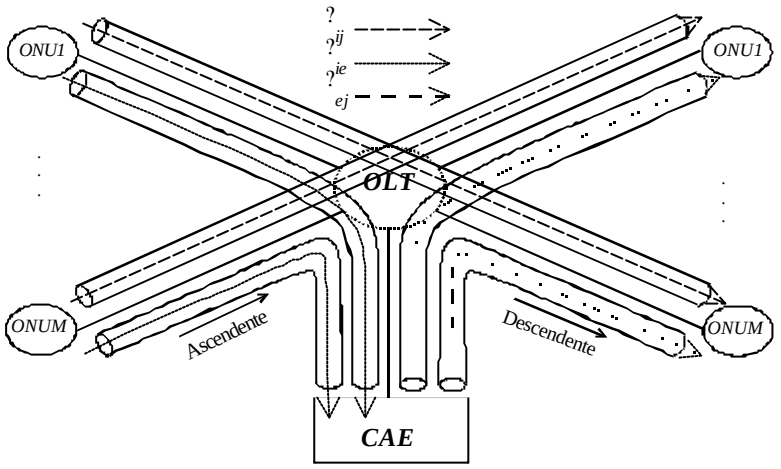


Figura 5-7 : Sessões de tráfego – estratégia III, serviços conversacionais.

5.4.4.1 Segregação de serviços

Considerando que apenas existem serviços distributivos, as sessões de tráfego (incluindo a intensidade de tráfego) e os VPCs existentes são exactamente os mesmos que os da estratégia I. Assim, os métodos de dimensionamento e os resultados obtidos são iguais aos apresentados na estratégia I.

Considerando que os serviços são conversacionais, existem VPCs extremo-a-extremo entre ONUs, e VPCs entre cada ONU e o exterior. Como cada sessão de tráfego atravessa apenas um VPC entre a sua origem e o destino, é possível usar a fórmula de *Erlang B* para determinar as probabilidades de bloqueio das sessões de tráfego. Estas probabilidades são obtidas pelas seguintes expressões:

$$B_{ie}^{s} = ER\left(\rho_{ie,}^{s} v_{ie}^{s}\right),\ i = 1, \ldots, e = 1, \ldots \tag{5-93}$$

$$B_{ej}^{s} = ER\left(\rho_{ej}^{s}, v_{ej}^{s}\right), \quad j = e \tag{5-94}$$

$$B_{ij}^{s} = ER\left(\rho_{ij}^{s}, v_{ij}^{s}\right), \quad i, j = e, \quad j \neq i \tag{5-95}$$

Nesta estratégia, o número de VCCs nas ligações interiores e exteriores é dado por:

$$V_{it} = \sum_{s=1}^{S} \left(\sum_{\substack{j=1 \\ j \neq i}}^{M} v_{ij}^{s} + v_{ie}^{s} \right), \quad i, j = t, e \tag{5-96}$$

$$V_{tj} = \sum_{s=1}^{S} \left(\sum_{\substack{i=1 \\ i \neq j}}^{M} v_{ij}^{s} + v_{ej}^{s} \right), \quad i, j = t, e \tag{5-97}$$

$$V_{te} = \sum_{s=1}^{S} \sum_{i=1}^{M} v_{ie}^{s}, \quad i = t, e \tag{5-98}$$

$$V_{et} = \sum_{s=1}^{S} \sum_{j=1}^{M} v_{ej}^{s}, \quad j = t, e \tag{5-99}$$

Numa situação *minONUs*, as metodologias de dimensionamento são semelhantes.

5.4.4.2 Agregação de serviços

Considerando apenas os serviços distributivos, os VPCs existentes e as sessões de tráfego que os atravessam são exactamente os mesmos que na estratégia I. Assim, a capacidade dos VPCs é igual nas estratégias I e III.

Num sistema com serviços conversacionais, os VPCs são diferentes dos existentes na estratégia I, pois na estratégia III existem VPCs extremo-a-extremo entre ONUs. O dimensionamento pode ser efectuado através de um dos métodos de cálculo da probabilidade de bloqueio num sistema com múltiplos serviços. Como exemplo, são apresentadas de seguida as fórmulas das probabilidades de bloqueio utilizando o método de *Knapsack*:

$$B_{ij}^{s} = \sum_{c=c_{ij} \cdot b^{s}+1}^{C_{ij}} q\left(\rho_{ij}^{s}, b^{s}, c\right), \quad i, j = e, \quad j \neq i \tag{5-100}$$

$$B_{ie}^{s} = \sum_{c=c_{ie} \cdot b^{s}+1}^{C_{ie}} q\left(\rho_{ie}^{s}, b^{s}, c\right), \quad i = e \tag{5-101}$$

$$B_{ej}^{s} = \sum_{c=c_{ej}-b^{s}+1}^{c_{ej}} q(c, b^{s}, c_{ej}^{s}) \quad j = 1, \ldots, M \qquad (5\text{-}102)$$

em que c_{ij} é a capacidade dos VPCs extremo-a-extremo entre ONUs, e c_{ie} e c_{ej} são as capacidades dos VPCs entre cada ONU e o exterior, respectivamente, nos sentidos ascendente e descendente. A função $q(?, b, c)$ encontra-se definida no anexo A, (secção A.2).

As capacidades das ligações interiores e exteriores são dadas por

$$C_{it} = \sum_{\substack{j=1 \\ i \neq j}}^{M} c_{ij} = c_{ie}, \; i, \; j = 1, \ldots, t = 1, \ldots, e = 1 \qquad (5\text{-}103)$$

$$C_{tj} = \sum_{\substack{i=1 \\ i \neq j}}^{M} c_{ij} = c_{ej}, \; i, \; j = 1, \ldots, t = 1, \ldots, e = 1 \qquad (5\text{-}104)$$

$$C_{te} = \sum_{i=1}^{M} c_{ie}, \; i = 1, \ldots, t = 1, \ldots, e = 1 \qquad (5\text{-}105)$$

$$C_{et} = \sum_{j=1}^{M} c_{ej}, \; j = 1, \ldots, t = 1, \ldots, e = 1 \qquad (5\text{-}106)$$

Numa situação *minONUs*, o método de cálculo das capacidades é semelhante.

5.5 Dimensionamento ao nível da célula

Os métodos de dimensionamento ao nível da chamada permitem calcular o número de VCCs necessários em cada VPC. Para determinar a largura de banda necessária em cada ligação é necessário calcular a largura de banda que tem de ser atribuída a cada VCC. Este é o objectivo do dimensionamento ao nível da célula.

No dimensionamento ao nível da célula são considerados dois métodos de atribuição de largura de banda: atribuição ao PBR (*Peak Bit Rate*) e atribuição de uma largura de banda efectiva a cada fonte de tráfego. A atribuição ao PBR assume que a fonte de tráfego se encontra sempre a transmitir à taxa máxima. A atribuição de largura de banda efectiva explora os ganhos potenciais da multiplexagem estatística entre as fontes de tráfego que se encontram no mesmo VPC. A largura de banda efectiva de uma fonte representa o débito (equivalente) que lhe deve ser atribuído, por forma a garantir que a QdS obtida quando esta

fonte é multiplexada com outras fontes numa fila de espera seja superior a um limiar pré-definido.

Normalmente, a largura de banda efectiva depende do número de fontes multiplexadas, do modelo das fontes de tráfego, da capacidade da fila de espera e dos parâmetros de QdS ao nível da célula. Os métodos de cálculo da largura de banda efectiva podem ser aditivos ou não-aditivos. Os métodos aditivos conduzem a uma largura de banda efectiva que é independente do número de fontes multiplexadas. Neste caso, a largura de banda total de uma ligação é a soma das larguras de banda efectivas de todas as fontes que atravessam a ligação. Nos métodos não-aditivos, a largura de banda efectiva de cada fonte de tráfego diminui com o aumento do número de fontes multiplexadas. Este último conjunto de métodos permite, normalmente, obter melhores aproximações, mas à custa de uma eficiência computacional inferior.

O modelo de cálculo de largura de banda efectiva considerado neste capítulo assume que o parâmetro de QdS ao nível da célula é o rácio de perda de células, isto é, o CLR. Este modelo, denominado de modelo de fluídos aditivo [Kesidis93], considera o problema geral da existência de uma largura de banda efectiva para fontes estacionárias e ergódicas. A sua determinação é efectuada a partir de uma aproximação baseada em grandes desvios. Este modelo assume que as fontes de tráfego de cada serviço são representadas ao nível da célula por um modelo de fluídos de *Markov*, definidas pela taxa máxima de transmissão (PBR), pelo tempo médio de permanência nos estados *on* e *off*, e pelo requisito de QdS (CLR). Em [Elwalid93] é apresentada uma extensão simples do modelo para um sistema com múltiplos serviços. Este modelo, além de ser aditivo e de ser aplicado a um sistema com múltiplos serviços, pode também ser estendido para redes com múltiplas filas de espera em cascata [Walrand98] (isto é, a largura de banda efectiva de uma fonte após a multiplexagem da fonte numa cascata de filas de espera nunca é superior à largura de banda efectiva da fonte na fila de espera de entrada). Este resultado vai ser necessário na estratégia II, pois nesta estratégia existem vários elementos de multiplexagem entre a origem e o destino das chamadas.

O número de elementos que efectuam multiplexagem estatística das fontes de tráfego no interior da rede de acesso depende da estratégia de gestão de recursos considerada. Em todas as estratégias considera-se que, no sentido ascendente, existe na ONU multiplexagem das fontes de tráfego (do mesmo serviço ou de todos os serviços) dos utilizadores ligados a

essa ONU. Na estratégia I de gestão de recursos, a OLT é apenas um *cross-connect* e não possui, por isso, capacidade de comutação de VCs e, consequentemente, capacidade de multiplexagem do tráfego de diferentes VPCs. Na estratégia II, a OLT é um elemento de comutação de VPs e de VCs, e por isso, é também um elemento de multiplexagem na rede de acesso. Na estratégia III, a OLT é, tal como na estratégia I, um *cross-connect*, e por isso, não tem a capacidade de multiplexagem de tráfego de diferentes VPCs. Na ONU, além da multiplexagem de tráfego de diferentes utilizadores, existe ainda, na estratégia III, a capacidade de comutação do tráfego entre utilizadores da mesma ONU.

5.6 Dimensionamento das ligações

O dimensionamento da rede de acesso consiste em determinar a largura de banda necessária em cada ligação física exterior e interior para garantir a cada serviço a QdS requerida tanto ao nível da chamada como da célula. Após a determinação da largura de banda das ligações, calcula-se o número de contentores VC-4 necessários no anel SDH para suportar a PON, e (se necessários) o número de módulos adicionais no interior da PON em ambos os sentidos ascendente e descendente.

Nas ligações exteriores, cada contentor VC-4 do anel SDH pode suportar um número inteiro de VCCs. Tendo em conta este facto, o cálculo do número de contentores VC-4 baseia-se na determinação do número de VCCs suportados em cada VC-4.

Na PON o dimensionamento necessita de ter em conta outros factores. No sentido descendente existe uma largura de banda base disponível na PON que pode ser partilhada por todas as ONUs. Esta largura de banda é multiplexada ao nível da camada ATM. Quando esta largura de banda não é suficiente, podem ser inseridos módulos adicionais em cada ONU. Como a transmissão da informação colocada nos módulos adicionais é baseada em SCM, não existe multiplexagem estatística entre os fluxos de tráfego inseridos no módulo de banda base e em cada um dos módulos adicionais. Os fluxos de tráfego que atravessam o mesmo VPC são multiplexados no interior dos módulos base e multiplexados no interior dos módulos adicionais. No sentido ascendente, os módulos (módulo base e adicional) de uma ONU não podem ser partilhados por tráfego de outras ONUs.

Tendo em conta todos estes factores, apresenta-se um pequeno algoritmo para efectuar o dimensionamento. No sentido descendente determina-se a largura de banda total necessária no interior da PON, através do cálculo do número de VCCs e da sua largura de

banda. Se a largura de banda total for inferior à capacidade do módulo de banda base, toda a informação é colocada neste módulo e não existe necessidade de inserir módulos adicionais. Caso contrário, é necessário introduzir alguns módulos adicionais na PON. Para rentabilizar cada módulo adicional, e dado que o módulo que pertence a uma ONU não pode ser partilhado por outras, determina-se a ONU com maiores necessidades de largura de banda e insere-se um módulo adicional nesta ONU. A largura de banda que é inserida no módulo adicional é subtraída à largura de banda do módulo base. Este processo repete-se até que a largura de banda necessária no módulo de banda base seja igual ou inferior à capacidade do módulo de banda base. Note-se, uma vez mais, que o número de VCCs introduzidos em cada módulo adicional tem de ser um número inteiro. Se o módulo adicional não tiver capacidade para comportar todo o tráfego com destino a uma determinada ONU, o tráfego restante é transportado em banda base. No sentido ascendente, o dimensionamento é mais simples. Como cada módulo base pertence a uma determinada ONU, quando este módulo não é suficiente para comportar esse tráfego, é inserido um módulo adicional.

O dimensionamento das ligações depende dos modelos de cálculo da largura de banda efectiva utilizados. O uso de modelos não-aditivos, que dependem do número de fontes multiplexadas, complicam o dimensionamento ao nível da célula. Neste caso específico, a largura de banda de cada ligação depende do número de módulos (contentores VC-4, módulos base e os módulos adicionais) utilizados. Por um lado, a largura de banda efectiva é alterada quando se altera o número de VCCs em cada módulo. Por outro lado, o número de VCCs que podem ser suportados por cada módulo também depende da largura de banda efectiva dos serviços. Deste modo, é necessário um procedimento iterativo, que converge ao fim de um número pequeno de iterações, para determinar o número de módulos. Se for usado um método de atribuição de largura de banda ao PBR ou considerando uma largura de banda efectiva aditiva, esta é independente do número de VCCs suportados em cada módulo e, por isso, o cálculo do número de módulos torna-se mais rápido.

5.7 Casos de estudo

Nesta secção são apresentados alguns casos de estudo baseados na rede desenvolvida como parte do projecto BBL. O estudo de dimensionamento aplica-se também ao caso da rede

FSAN (*Full Services Access Networks*) [Quayle97], dada a semelhança das duas arquitecturas. Assim, nos casos de estudo considera-se o caso específico de uma rede de acesso BBL com 16 ONUs e um máximo de 32 utilizadores por ONU. O anel SDH é composto por vários contentores VC-4 cada um com 155.52 Mb/seg. A largura de banda base da PON no sentido descendente é de 155.52 Mb/seg com a possibilidade de instalação de módulos adicionais de 51.84 Mb/seg em cada ONU. No sentido ascendente, a largura de banda no módulo base disponível em cada ONU é de 9.72 Mb/seg, com a possibilidade de colocar módulos adicionais de 51.84 Mb/seg.

Nos casos de estudo consideram-se três cenários diferentes: (i) um cenário com utilizadores do tipo residencial, (ii) um cenário com utilizadores do tipo empresarial (iii) e um cenário misto residencial e empresarial.

5.7.1 Parametrização dos serviços

Nesta secção são descritos os serviços seleccionados e os parâmetros dos modelos ao nível da chamada e da célula.

O serviço de acesso à *Internet* é considerado residencial e empresarial. É um serviço comutado, distributivo e assimétrico. Assume-se, nestes casos de estudo, que em cada acesso à *Internet* a fonte de tráfego é *on-off* com um PBR de 1 Mb/seg e uma percentagem de permanência no estado *on* de 50%. O parâmetro de QdS é um CLR de 10^{-6}. Considera-se então que o serviço é configurado como um VBR com largura de banda de pico de 1 Mb/seg e largura de banda média de 0.5 Mb/seg. Assume-se também que cada utilizador acede à *Internet* em média 4 vezes por dia, e que cada sessão tem uma duração média de meia hora. No sentido ascendente, considera-se que a quantidade de informação enviada é muito pequena e que, por isso, a taxa de envio de informação por utilizador é desprezável.

O serviço de *Video on Demand* (VoD) é residencial, comutado, distributivo e assimétrico. As características de largura de banda dependem essencialmente da norma de compressão: com MPEG1 a taxa é de 1.5 Mb/seg, e com MPEG2 a taxa varia entre 3 e 6 Mb/seg. Considera-se que é utilizado o MPEG2 para codificar a informação e que, por isso, cada chamada tem uma taxa de pico de 6 Mb/seg e uma taxa média de 4 Mb/seg [Gall91]. A fonte de tráfego considerada é um *on-off* com um PBR de 6 Mb/seg e uma percentagem de permanência no estado *on* de 66.(6)%. O parâmetro de QdS é um CLR de 10^{-9} para garantir uma boa qualidade de imagem. O serviço é configurado como um VBR com largura de banda de pico de 6 Mb/seg e largura de banda média de 4 Mb/seg.

Considera-se que o número médio de chamadas por utilizador é de 5 vídeos por semana e que a duração média de cada filme é de 2 horas. No sentido ascendente, considera-se também que a largura de banda necessária é desprezável.

O serviço de distribuição de vídeo é um serviço ponto-multiponto. Este serviço é residencial, permanente, distributivo e assimétrico. Como é um serviço distributivo, a informação é transmitida apenas no sentido descendente. A codificação considerada é MPEG2, tal como no serviço de VoD. Como o serviço de distribuição de vídeo é um serviço permanente, assume-se que a largura de banda necessária é sempre de 4 Mb/seg. O serviço é configurado como um CBR de 4 Mb/seg.

O serviço de videoconferência é um serviço empresarial, comutado, conversacional e simétrico. Existem actualmente soluções comerciais para implementar o serviço com base em canais RDIS de 64 Kb/seg. Os serviços de videoconferência baseados em 2 acessos básicos (4 canais RDIS) permitem obter uma qualidade razoável para serviços com cenas estáticas, enquanto que para cenas dinâmicas são necessários 3 acessos básicos (6 canais RDIS) [Int-Vidconf]. Sendo assim, considera-se que a fonte de tráfego é um *on-off* com um PBR de 384 Kb/seg e uma percentagem de permanência no estado *on* de 66.(6)%. A QdS ao nível da célula é caracterizada por um CLR de 10^{-9}. O serviço é configurado como um VBR com largura de banda de pico de 384 Kb/seg e largura de banda média de 256 Kb/seg. Considera-se que o número de chamadas por utilizador é em média de 1 chamada em cada duas horas e que a duração média de cada sessão de videoconferência é de 1 hora.

O serviço de interligação de LANs é um serviço de ligação permanente entre duas redes locais. Este serviço é empresarial e, como se considera a troca de informação entre as duas redes locais (serviço conversacional), considera-se que o tráfego é simétrico nos sentidos ascendente e descendente. Assume-se que estão permanentemente atribuídos 4 Mb/seg para este serviço nos dois sentidos. Nos casos de estudo apresentados considera-se que, nos cenários mistos e empresariais, 4 dos utilizadores de cada ONU empresarial têm largura de banda reservada para este serviço. Considera-se também que 25% das ligações são entre utilizadores (LANs) de PONs distintas, e que 75% das ligações são entre utilizadores (LANs) da mesma PON.

A Tabela 5-1 e a Tabela 5-2 apresentam uma listagem dos parâmetros ao nível da chamada e da célula considerados para os serviços que foram seleccionados para serem suportados pela rede de acesso.

	Serviços				
Parâmetros	**VoD**	**Acesso à *Internet***	**Distribuição de Vídeo**	**Interligação de LANs**	**Videoconferência**
??? **(seg)**	7200	1800	--	--	3600
?????? **(seg^{-1})**	0.03	0.17	--	--	0.5
GdS (%)	0.1	0.1	--	--	0.1

Tabela 5-1: Parâmetros dos serviços ao nível da chamada.

	Serviços				
Parâmetros	**VoD**	**Acesso à *Internet***	**Distribuição de Vídeo**	**Interligação de LANs**	**Videoconferência**
PBR (Mb/seg)	6	1	4	4	0.384
Permanência *on* (%)	66.(6)	50	100	100	66.(6)
QdS	10^{-9}	10^{-6}	--	--	10^{-9}

Tabela 5-2 : Parâmetros dos serviços ao nível da célula.

A Tabela 5-3 apresenta a largura de banda efectiva das fontes de tráfego dos serviços quando é atribuída a largura de banda baseada no PBR, e quando é considerada multiplexagem estatística entre fontes. LBEFI (Largura de Banda Efectiva I) corresponde aos parâmetros de tráfego ao nível da célula apresentados na Tabela 5-2, e LBEFII corresponde ao caso em que a percentagem de tempo de permanência no estado *on* é 3 vezes inferior à existente no caso LBEFI, mantendo a largura de banda de pico. Este caso é apresentado para analisar o efeito do *burstiness* das fontes de tráfego no dimensionamento. Os valores apresentados na Tabela 5-3 consideram que a fila de espera de multiplexagem das fontes de tráfego tem um comprimento de 100 células. Pode-se observar na tabela que a largura de banda efectiva diminui com a diminuição da percentagem do tempo de permanência no estado *on.*

Serviços	**PBR (Mb/seg)**	**LBEFI (Mb/seg)**	**LBEFII (Mb/seg)**
VoD	6	5.9322	5.7314
Acesso à *Internet*	1	0.6559	0.4518
Videoconferência	0.384	0.3353	0.3193

Tabela 5-3 : Atribuição de largura de banda ao PBR e considerando larguras de banda efectivas.

5.7.2 Resultados do dimensionamento

Nesta secção são apresentados os resultados do dimensionamento da rede de acesso considerando cenários com características diferentes.

Na Figura 5-8 é ilustrada a diferença entre os casos *maxONUs* e *minONUs*, considerando a estratégia I de gestão de recursos num cenário residencial com um requisito

de GdS de 0.1 %. As colunas 1 e 2 correspondem ao número de contentores VC-4 do anel SDH com *maxONUs* (coluna 1) e *minONUs* (coluna 2); as colunas 3 e 4 correspondem ao número de módulos adicionais necessários no sentido descendente com *maxONUs* (coluna 3) e *minONUs* (coluna 4). Para uma penetração de serviço de 100% são obtidos os mesmos resultados em ambos os casos *maxONUs* e *minONUS*. Nesta situação a PON está a operar com o número máximo possível de utilizadores e por isso, a configuração da rede é igual em ambos os casos. Com uma pequena penetração de serviço, o caso *minONUs* necessita de um número inferior de contentores VC-4 e módulos adicionais. Por exemplo, para uma penetração de serviço de 25%, são necessários 13 módulos adicionais e 4 contentores VC-4 no caso *maxONUs*, mas apenas 2 módulos adicionais e 2 contentores VC-4 no caso *minONUs*. Note-se que, com esta percentagem de penetração de serviço, a rede de acesso contém 16 ONUs cada uma com 8 utilizadores no caso *maxONUs*, enquanto que no caso *minONUs* a rede contém 4 ONUs com o número máximo de utilizadores (cada uma com 32 utilizadores). O caso *maxONUs* necessita de um maior número de módulos adicionais porque o tráfego oferecido que excede a largura de banda base é distribuído por mais ONUs. Assim, a largura de banda necessária entre cada ONU e a OLT é menor no caso *maxONUs* em comparação com o caso *minONUs*, e pode ser muito inferior à largura de banda do módulo adicional e, por isso, a largura de banda que fica disponível em cada módulo adicional pode ser elevada. O caso *maxONUs* necessita de um maior número de contentores VC-4 porque as mesmas sessões de tráfego são distribuídas por um maior número de VPCs, conduzindo a uma menor partilha de recursos ao nível da chamada.

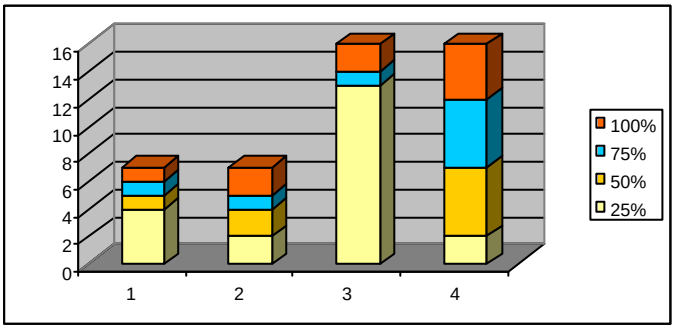


Figura 5-8 : Contentores VC-4 (coluna 1 e 2) e módulos adicionais residenciais no sentido descendente (coluna 3 e 4) em *maxONUs* (1 e 3) *vs minONUs* (2 e 4) (PBR).

Na Figura 5-9 é estudado o impacto no dimensionamento dos ganhos de multiplexagem estatística, considerando a estratégia I de gestão de recursos num cenário residencial, com um requisito de GdS de 0.1%. Considera-se o caso *maxONUs* e uma penetração de serviço de 75%. As colunas 1 a 3 correspondem ao número de contentores VC-4 necessários para uma atribuição de largura de banda ao PBR (coluna 1), LBEFI (coluna 2) e LBEFII (coluna 3); as colunas 4 a 6 correspondem ao número de módulos adicionais para uma atribuição de largura de banda ao PBR (coluna 4), LBEFI (coluna 5) e LBEFII (coluna 6). Em cada coluna 1 a 3 é representada (a escuro) a percentagem de ocupação do último contentor VC-4 que é colocado na rede de acesso (o único que tem alguma largura de banda disponível); em cada coluna 4 a 6 é representada (a escuro) a percentagem da largura de banda base no sentido descendente utilizada após se terem introduzido os módulos adicionais necessários.

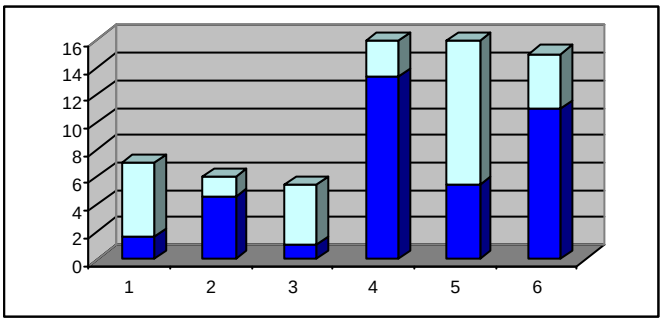


Figura 5-9 : Contentores VC-4 (colunas 1 a 3) e módulos adicionais residenciais no sentido descendente (colunas 4 a 6) com PBR (1 e 4) *vs* LBEFI (2 e 5) *vs* LBEFII (3 e 6) (*maxONUs*).

Os resultados ilustram a vantagem de os recursos serem atribuídos aos VPCs assumindo a multiplexagem estatística entre sessões que atravessam o mesmo VPC. O número de contentores VC-4 é de 7 com atribuição de largura de banda ao PBR, e 6 com LBEFI e LBEFII. Note-se que embora ambos LBEFI e LBEFII necessitem do mesmo número de contentores VC-4, a ocupação do último contentor VC-4 é muito inferior no caso de LBEFII. As mesmas conclusões podem ser retiradas quanto aos módulos adicionais.

Na Figura 5-10 são comparadas as três estratégias de gestão de recursos num cenário residencial com uma penetração de serviço de 100%. As colunas 1 a 3 correspondem ao número de contentores VC-4 nas estratégias I (coluna 1), II (coluna 2) e III (coluna 3). As

colunas 4 a 6 correspondem ao número de módulos adicionais nas estratégias I (coluna 4), II (coluna 5) e III (coluna 6). Em cada coluna representa-se o número necessário de contentores VC-4 e módulos adicionais para os casos de atribuição de largura de banda ao PBR, LBEFI e LBEFII. O número de contentores VC-4 e de módulos adicionais é o mesmo nas estratégias I e III porque, dado que os serviços residenciais são do tipo distributivo, não existem VPCs internos na estratégia III, e portanto esta torna-se igual à estratégia I. Na estratégia II o número de contentores VC-4 é muito inferior porque todas as sessões de tráfego de cada serviço partilham o mesmo VPC entre a OLT e o CAE. O número de módulos adicionais na estratégia II é aproximadamente igual ao número necessário nas estratégias I e III. Note-se que em ambas as estratégias I e II os VPCs internos transportam as mesmas sessões de tráfego. A única diferença entre estas duas estratégias num cenário residencial é o facto de na estratégia II o tráfego entre PONs diferentes atravessar 2 VPCs distintos, enquanto que na estratégia I atravessam apenas um VPC. Deste modo, o número de VCCs necessários nos VPCs internos na estratégia II é sempre igual ou superior ao número necessário na estratégia I (mas, em geral, não existem diferenças significativas). Observa-se também na figura que nas estratégias I e III o número de módulos adicionais necessários apenas aumentam de uma unidade quando se passa de uma atribuição LBEFII para PBR. Este facto deve-se a este cenário ser dominado pelo serviço de VoD (os outros são os serviços de distribuição de vídeo e de acesso à *Internet*), o qual, como pode ser observado na Tabela 5-3, não beneficia significativamente da multiplexagem estatística.

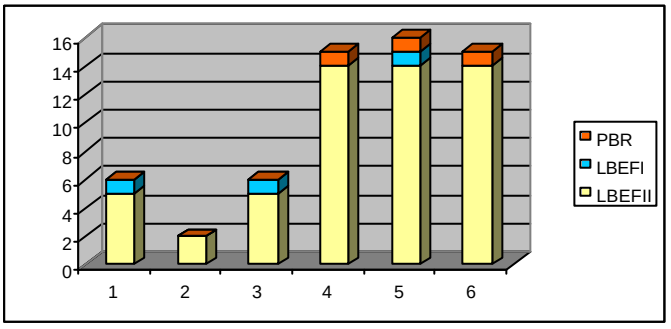


Figura 5-10 : Contentores VC-4 (colunas 1 a 3) e módulos adicionais residenciais no sentido descendente (colunas 4 a 6) com as estratégias I (1 e 4), II (2 e 5) e III (3 e 6) (*maxONUs*).

Na Figura 5-11 é ilustrada a diferença entre as situações de segregação de serviços e agregação de serviços, considerando a estratégia I de gestão de recursos num cenário residencial. As colunas 1 e 2 correspondem ao número de contentores VC-4 do anel SDH com segregação de serviços (coluna 1) e com agregação de serviços (coluna 2); as colunas 3 e 4 correspondem ao número de módulos adicionais necessários no sentido descendente com segregação de serviços (coluna 3) e com agregação de serviços (coluna 4). Em cada coluna representa-se o número necessário de contentores VC-4 e módulos adicionais para os casos de atribuição de largura de banda ao PBR, LBEFI e LBEFII. Na situação com agregação de serviços, o número de contentores VC-4 e de módulos adicionais é sempre ligeiramente inferior porque existe multiplexagem estatística entre todas as sessões de tráfego de todos os serviços. Por exemplo, para uma atribuição de largura de banda LBEFII, são necessários 15 módulos adicionais na situação com segregação de serviços e apenas 14 na situação com agregação de serviços. Note-se que, neste caso, como todos os serviços atravessam a mesma fila de espera, a QdS (o limiar de CLR) de todos os serviços é igual e corresponde à QdS mais estrita, equivalente a um CLR de 10^{-9} para todos os serviços, incluindo o de acesso à *Internet*. Assim, a largura de banda efectiva dos serviços que necessitam de uma QdS inferior aumenta, e por isso, as diferenças no número de módulos adicionais e contentores necessários são relativamente pequenas, no máximo de 1 unidade.

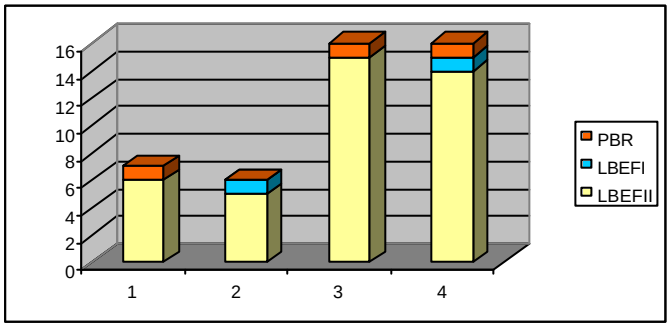


Figura 5-11 : Contentores VC-4 (colunas 1 e 2) e módulos adicionais residenciais no sentido descendente (colunas 1 e 2) com segregação (1 e 3) e agregação (2 e 4) de serviços (*maxONUs* e estratégia I).

Na Figura 5-12 são comparadas as três estratégias de gestão de recursos num cenário empresarial com 4 ligações de interligação de LANs por ONU, uma penetração de serviço

de 100%, e com 25% do tráfego de todos os serviços entre utilizadores de PONs diferentes. As colunas 1 a 3 correspondem ao número de contentores VC-4 nas estratégias I (coluna 1), II (coluna 2) e III (coluna 3). As colunas 4 a 6 correspondem ao número de módulos adicionais nas estratégias I (coluna 4), II (coluna 5) e III (coluna 6). Em cada coluna representa-se o número necessário de contentores VC-4 e módulos adicionais para os casos de atribuição de largura de banda ao PBR, LBEFI e LBEFII. O número de contentores VC-4 é mais uma vez bastante inferior na estratégia II. Existem duas grandes diferenças deste cenário em relação ao cenário residencial no que respeita ao número de módulos adicionais. Primeiro, o número de módulos adicionais necessários na estratégia III é superior ao necessário nas outras duas. Este facto acontece porque na estratégia III não existe partilha de recursos, uma vez que todos os VPCs internos transportam apenas uma sessão de tráfego. Segundo, os ganhos de multiplexagem estatística são superiores neste cenário para todas as estratégias, isto é, existe uma diferença de 3 a 4 módulos entre uma atribuição de largura de banda ao PBR e LBEFII. Neste cenário, o serviço que ocupa a maior parte dos recursos é o serviço de acesso à *Internet* que, como se observa na Tabela 5-3, beneficia significativamente da multiplexagem estatística.

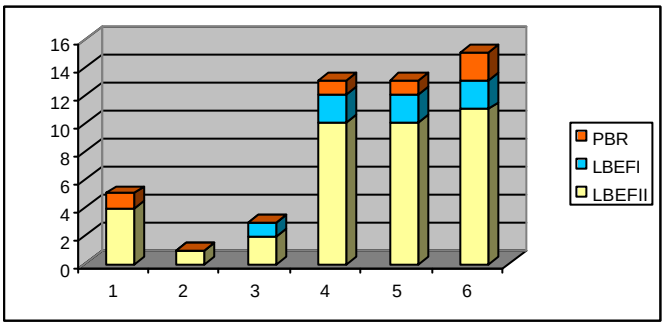


Figura 5-12 : Contentores VC-4 (colunas 1, 2 e 3) e módulos adicionais empresariais no sentido descendente (colunas 4, 5 e 6) com as estratégias I (1 e 4), II (2 e 5) e III (3 e 6) (*maxONUs*).

A Figura 5-13 apresenta o número de contentores VC-4 (coluna 1) e de módulos adicionais (colunas 2 a 4) necessários num cenário misto com 12 ONUs residenciais e 4 empresariais. São apresentados os módulos adicionais necessários nas ONUs residenciais no sentido descendente (coluna 4) e empresariais nos sentidos ascendente (coluna 2) e descendente (coluna 3). Considera-se que em cada ONU empresarial existem 4 utilizadores que usam o serviço de interligação de LANs. A percentagem de tráfego entre PONs

distintas é de 25%, a atribuição de largura de banda é efectuada ao PBR, e é considerado o caso *maxONUs*. Neste cenário existe a necessidade de incluir um determinado número de módulos adicionais no sentido ascendente nas ONUs empresariais porque dois dos serviços empresariais são simétricos. Para uma penetração de serviço de 25% e 50% não existe a necessidade de introduzir módulos adicionais no sentido descendente nas ONUs empresariais porque os serviços residenciais ocupam uma maior quantidade de largura de banda (a colocação de largura de banda em módulos adicionais tem início nas ONUs com maior requisitos de largura de banda). Apenas como exemplo, para uma penetração de serviço de 25%, as ONUs residenciais necessitam de 48 Mb/seg e as empresariais necessitam de 23 Mb/seg. Para todos os valores de penetração de serviço apresentados, todas as ONUs residenciais necessitam de módulos adicionais no sentido descendente, enquanto que apenas existe a necessidade de ter módulos adicionais em todas as ONUs empresariais para uma penetração de serviço de 100%.

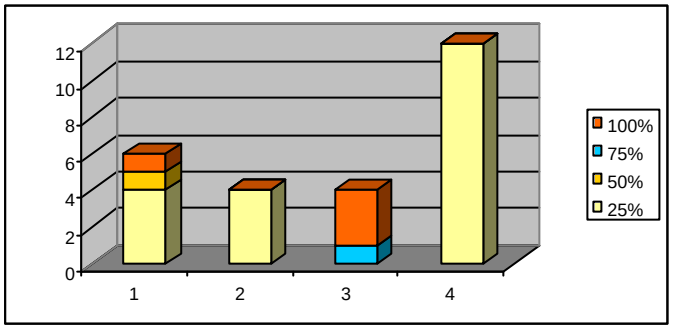


Figura 5-13 : Contentores VC-4 (coluna 1) e módulos adicionais residenciais no sentido descendente (coluna 4) e empresariais nos sentidos ascendente (coluna 2) e descendente (coluna 3) com a estratégia I (PBR e *maxONUs*).

Na Figura 5-14 é apresentada a diferença entre as situações com segregação de serviços (colunas 1, 3 e 5) e com agregação de serviços (colunas 2, 4 e 6), considerando a estratégia I de gestão de recursos num cenário misto. A figura apresenta o número de contentores VC-4 (colunas 1 e 2), o número de módulos adicionais no sentido descendente nas ONUs empresariais (colunas 3 e 4), e o número de módulos adicionais no sentido descendente nas ONUs residenciais (colunas 5 e 6). Mais uma vez se verifica que numa situação com agregação de serviços o número de contentores VC-4 e de módulos adicionais é ligeiramente inferior (no máximo existe uma diferença de 1 módulo) devido

aos superiores ganhos de multiplexagem estatística. Note-se que, nos módulos residenciais, não existe diferença entre as duas situações, pois as ONUs residenciais necessitam sempre do número máximo de módulos. Em relação aos módulos empresariais e número de contentores VC-4 existe uma pequena redução na largura de banda necessária em agregação de serviços.

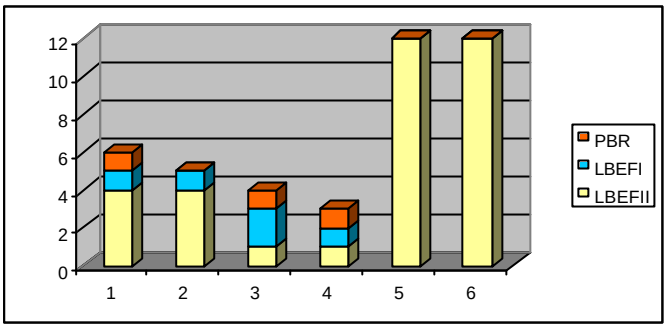


Figura 5-14 : Contentores VC-4 (colunas 1 e 2) e módulos adicionais residenciais (colunas 3 e 4) e empresariais (colunas 5 e 6) no sentido descendente com segregação (1, 3 e 5) e agregação (2, 4 e 6) de serviços (*maxONUs* e a estratégia I).

5.8 Conclusões

Este capítulo propôs estratégias de gestão de recursos com base em VPs e definiu metodologias de dimensionamento para redes de acesso ATM. O dimensionamento é realizado em dois passos: dimensionamento ao nível da chamada e ao nível da célula. No dimensionamento ao nível da chamada é determinado o número de VCCs necessários em cada VPC para garantir um GdS pré-definido. No dimensionamento ao nível da célula é determinada a largura de banda necessária em cada VCC para garantir o transporte dos serviços com uma determinada QdS.

Os métodos de dimensionamento propostos são função da estratégia de gestão de recursos implementada, das características dos serviços suportados (permanentes ou comutados, distributivos ou conversacionais e simétricos ou assimétricos), da distribuição espacial dos utilizadores pelas ONUs, da agregação ou segregação de serviços nos VPCs, e da forma de atribuição de largura de banda aos VPCs (PBR ou largura de banda efectiva).

A estratégia de gestão de recursos que permitiu ter uma maior economia de largura de banda foi a estratégia II. Considerando as ligações do interior da PON, as necessidades

de largura de banda são semelhantes em todas as estratégias, sendo as da estratégia III ligeiramente superiores. Considerando as ligações do exterior da PON, verificou-se que a largura de banda necessária na estratégia II apresenta uma redução da ordem de 67% em cenários residenciais e de 67% a 80% em cenários empresariais, comparativamente às outras estratégias. No entanto, esta estratégia requer funcionalidades de CAC e UPC na OLT, ou seja, um aumento no seu custo. Em última análise, a decisão da estratégia a implementar depende do compromisso que for possível obter entre custo de sinalização e de largura de banda.

Os ganhos obtidos considerando multiplexagem estatística são significativos, principalmente no caso dos serviços empresariais que consideram fontes de tráfego com maior *burstiness*. Nos estudos efectuados verificou-se que em algumas situações foi possível diminuir o número de módulos adicionais em 75% e diminuir o número de módulos VC-4 necessários no anel SDH em 33%.

Nos casos apresentados, o número de módulos VC-4 e de módulos adicionais foi sempre ligeiramente inferior quando se considerou a agregação de todos os serviços no mesmo VPC. Esta vantagem da agregação de serviços está relacionada com o facto de o CLR ser semelhante para os diferentes serviços (entre 10^{-6} e 10^{-9}). No entanto, se as diferenças entre os CLRs forem elevadas, é necessário atribuir o CLR mais estrito a todos os serviços, se estes forem agregados no mesmo VPC. Nesta situação, a largura de banda necessária em cada ligação aumenta e pode ser superior à necessária considerando segregação de serviços.

Os métodos de dimensionamento constituem uma ferramenta que pode ser usada por um operador de rede de acesso para: (i) projectar a sua rede de acordo com o tráfego esperado ao longo do tempo; (ii) tomar decisões acerca das funcionalidades requeridas nos elementos da sua rede de acordo com os custos específicos dos elementos e de qualquer módulo adicional de largura de banda e; (iii) ter (em qualquer altura) uma estimativa da QdS e probabilidade de bloqueio de um serviço acedido por um utilizador específico.

CAPÍTULO 6

REDES DE ACESSO IP COM SUPORTE DE QDS

No capítulo 4 apresentou-se uma perspectiva histórica de evolução das redes de acesso. Devido ao número crescente de aplicações que utilizam o IP (*Internet Protocol*), a sua introdução como tecnologia de transporte nas redes de acesso representa uma tendência natural. O problema inerente às redes IP tradicionais é o facto de o modelo de serviço fornecido por estas redes ser o modelo de melhor esforço, que não suporta QdS (Qualidade de Serviço). No entanto, existem diversas arquitecturas e mecanismos, normalizados ou em fase de investigação (alguns dos quais descritos no capítulo 3), que permitem dotar uma rede IP de suporte de QdS.

O trabalho realizado neste capítulo centra-se na proposta de uma arquitectura para redes de acesso de nova geração baseadas em IP, com suporte de QdS diferenciada, integrando na mesma infra-estrutura de rede um conjunto variado de serviços e aplicações, incluindo serviços multimédia. A definição de uma solução tecnológica para uma rede de acesso, qualquer que ela seja, está fortemente condicionada por questões de custos, devido ao número de elementos de rede que têm de ser substituídos ou instalados de raiz. Como já foi referido no capítulo 4, as tecnologias de rede de acesso tradicionais permitem uma partilha dos recursos da rede muito limitada. Por exemplo, nas redes de circuitos comutados (com acesso *dial-up*) os recursos atribuídos aos utilizadores estão

completamente isolados uns dos outros; nas redes ADSL (*Asymmetrical Digital Subscriber Line*) os serviços de voz e de banda larga são isolados ao nível da camada física, através de multiplexagem na frequência. Esta característica deve-se essencialmente ao custo da sinalização que seria necessário introduzir na rede para a tornar numa rede partilhada. A partilha dos recursos na rede de acesso permite diminuir o custo da largura de banda das ligações, mas requer equipamentos com mais funcionalidades, ou seja, equipamentos mais complexos e com um maior custo. Nas redes de acesso futuras é necessário estabelecer um bom compromisso entre o custo e a funcionalidade dos equipamentos. A arquitectura das redes de acesso IP que é proposta neste capítulo foi projectada tendo em conta o factor custo, ao mesmo tempo que permite fornecer um suporte integrado de serviços multimédia de banda larga. Além da definição da arquitectura das redes de acesso, são ainda apresentadas as tecnologias e protocolos necessários ao suporte de reserva de recursos, sinalização das sessões, e gestão das políticas de QdS que incluem as funcionalidades de AAA (*Authentication, Authorization and Accounting*).

Uma possibilidade de evolução das redes de acesso com circuitos comutados para redes com recursos partilhados pode passar pela reutilização do PPP (*Point-to-Point Protocol*) [Simpson94] nas redes de acesso. A título ilustrativo, será descrita muito sumariamente neste capítulo, uma forma de reutilizar o PPP nas redes de acesso IP. Uma solução com PPP tem o problema de, pelo facto de este não estar preparado para ser utilizado directamente numa rede com recursos partilhados, ser necessário o recurso a protocolos de *tunneling*, por exemplo o L2TP (*Layer 2 Tunneling Protocol*) [Townsley99] e o PPPoE (PPP *over Ethernet*) [Mamakos99]. A introdução destes protocolos aumenta a complexidade da rede de acesso e o *overhead* introduzido em cada pacote de informação.

A arquitectura da rede de acesso proposta permite suportar diferenciação de QdS e serviços multimédia através de um conjunto de tecnologias e protocolos introduzidos recentemente. Estes protocolos incluem o SIP (*Session Initiation Protocol*) [Handley99] para iniciar e configurar uma sessão, o COPS (*Common Open Policy Service*) [Boyle00] e o DIAMETER [Calhoun01] para gerir as políticas de QdS, incluindo as funções de AAA, e o RSVP (*resource ReSerVation Protocol*) para reservar recursos. A rede de acesso proposta nesta solução pode ser configurada de acordo com uma arquitectura IntServ (Integração de Serviços) ou DiffServ (Diferenciação de Serviços), dependendo da escolha do operador da rede. Em qualquer dos casos, assume-se que o protocolo de reserva de

recursos tem como base o RSVP. Numa arquitectura IntServ é usado o protocolo RSVP base, e em DiffServ é usada uma extensão do protocolo RSVP [Baker01] que permite efectuar reservas de agregados de fluxos. À data da escrita desta Tese encontra-se já implementado um demonstrador desta rede de acesso proposta [Salgado02]. Foram realizadas várias experiências para validar o funcionamento do demonstrador e o seu desempenho. Como exemplo das experiências efectuadas, foi utilizado o demonstrador da rede de acesso para transportar em simultâneo serviços de áudio (música) e de dados numa rede congestionada. Os resultados obtidos com o demonstrador validam o funcionamento da rede, e o seu suporte de integração dos serviços com uma QdS diferenciada.

Este capítulo é organizado da seguinte forma. Na secção 6.1 são descritos os elementos constituintes da rede de acesso e as interfaces físicas entre cada elemento. Na secção 6.2 apresentam-se os esquemas de endereçamento dos elementos da rede e os protocolos de encaminhamento necessários. A secção 6.3 apresenta, a título ilustrativo, uma forma de reutilizar o PPP na rede de acesso. A arquitectura de rede de acesso proposta é apresentada na secção 6.4, incluindo a descrição das funcionalidades que cada elemento da rede de acesso tem de suportar e das tecnologias e protocolos utilizados para suportar essas funcionalidades. Nesta secção é também apresentado um exemplo do fluxo de mensagens de uma sessão multimédia de acesso à *Internet*. Finalmente, na secção 6.5 são apresentadas as conclusões mais importantes.

6.1 Elementos e interfaces físicas de uma rede de acesso IP com recursos partilhados

A Figura 6-1 apresenta os elementos da rede de acesso IP com recursos partilhados. Considera-se que um utilizador pode ser um terminal ou uma rede local (LAN - *Local Area Network*). O NT (*Network Terminator*) faz a adaptação entre a interface de utilizador e a rede de acesso. O tráfego de diferentes NTs é agregado em *multiplexers* (MUX) de nível 3, designados por MUX-IP, sendo possíveis vários andares de multiplexagem. A opção pela utilização de MUX-IP na rede de acesso (em vez de, por exemplo, *routers* IP) justifica-se por questões de custo e pelo facto de não ser necessária redundância de percurso. Considera-se que na rede de acesso é apenas necessário garantir redundância no nível físico. Assim, a topologia lógica da rede de acesso reduz-se a uma árvore, o que simplifica as funções de encaminhamento e de endereçamento. Considera-se que a rede de acesso é

baseada em tecnologia de nível 3, por exemplo, em IP. Nesta rede todos os elementos implementam funções de nível 3. O BAS (*Broadband Access Server*) realiza a interface entre a rede de acesso e o ISP (*Internet Service Provider*). As funções do BAS incluem a autenticação e a autorização do utilizador, e a contabilização dos serviços por ele acedidos (actuando como um *proxy* de AAA do ISP), a atribuição de endereços IP aos utilizadores (actuando como um servidor de DHCP – *Dynamic Host Configuration Protocol*), e a selecção de ISPs por parte do utilizador. Note-se que na figura estão delimitadas as entidades descritas no capítulo 4 (utilizador, operador da rede de acesso e ISP) e os elementos que a elas pertencem. Todos os elementos entre o NT e o BAS inclusivé são geridos pelo operador da rede de acesso.

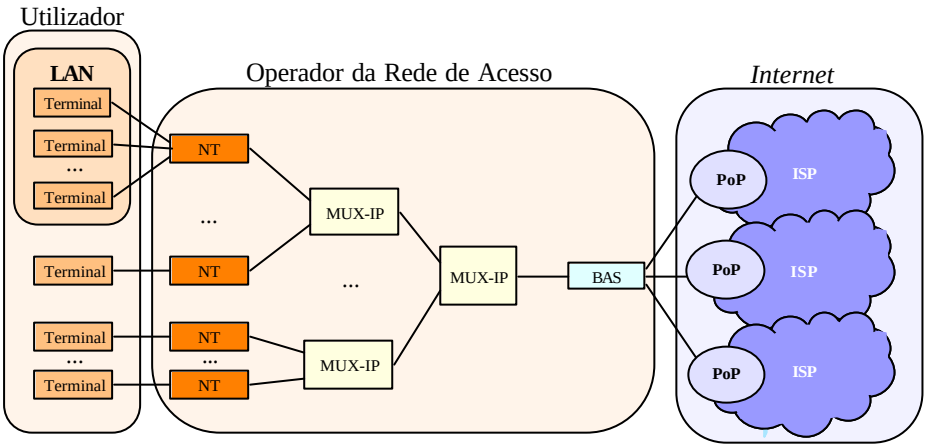


Figura 6-1 : Rede de acesso baseada em IP.

A interface física entre os NTs e os MUX-IP pode ser *x*DSL ou fibra (FTTH - *Fiber To The Home*). A primeira opção permitiria reaproveitar a infra-estrutura de cobre existente. Ao nível da interface física na rede de acesso entre os MUX-IP e entre estes e o BAS podem ser consideradas duas topologias diferentes: topologia em anel e em estrela. Na primeira topologia, ilustrada na Figura 6-2, a interface física é constituída por um anel SDH (*Synchronous Digital Hierarchy*). O BAS e os *Add Drops* encontram-se ligados através de tributários (STMs - *Synchronous Transmission Modules*) do anel SDH. O SDH fornece protecção ao nível da camada física. Prevê-se que as tecnologias de rede venham a evoluir no sentido de permitir o transporte de IP directamente sobre fibra.

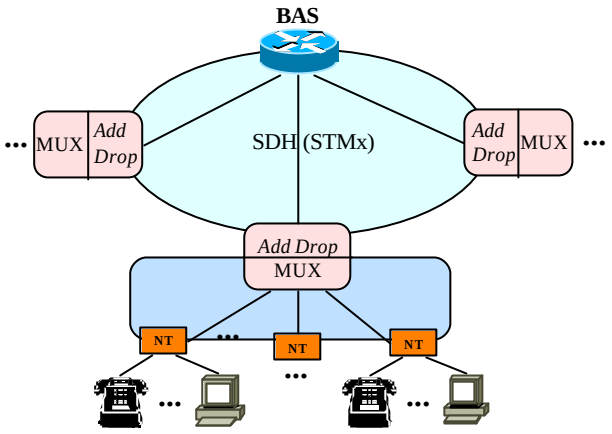


Figura 6-2 : Topologia em anel.

Numa topologia em estrela, existem ligações de fibra ponto-a-ponto entre o BAS e cada MUX-IP. Neste caso, a protecção pode ser conseguida por duplicação de fibras e de equipamentos. Obviamente, esta solução é economicamente menos viável.

A rede de acesso proposta pode conter vários andares de MUX-IP. A escolha do número de andares depende da quantidade de tráfego e da dispersão geográfica dos utilizadores. No que se segue, considera-se uma topologia com 2 andares de MUX-IP.

6.2 Endereçamento e encaminhamento

Dado que a topologia da rede de acesso é uma árvore, o tráfego pode ser enviado no sentido ascendente sem qualquer informação de encaminhamento, pois o percurso é único. O BAS e os MUX-IP necessitam apenas de manter tabelas de encaminhamento para enviar o tráfego no sentido descendente. Assim, a informação de encaminhamento tem apenas de ser anunciada no sentido ascendente. Não é necessário envolver métricas no processo de encaminhamento e os elementos de rede (MUX-IP ou BAS) não necessitam de incluir algoritmos para determinação dos percursos de menor custo, porque não existe mais do que um percurso possível entre o BAS e cada utilizador. As tabelas de encaminhamento são muito simples, porque para um determinado elemento na rede de acesso, é necessária apenas uma entrada na tabela para cada rede ou utilizador que é atingível no sentido descendente a partir desse elemento.

As funcionalidades de encaminhamento podem ser implementadas por uma versão simplificada de um protocolo baseado em vectores de distância, tal como o RIP (*Routing Internet Protocol*) [Malkin98]. Neste caso, o protocolo pode ser simplificado porque não existe o problema da contagem-para-infinito, que complica os protocolos baseados em vectores de distância, e também pelo facto de não serem necessárias métricas de encaminhamento. O endereçamento dentro da rede de acesso também pode ser simplificado, pois não é necessário associar um endereço IP a cada interface. É necessário apenas um endereço IP por cada elemento de rede para suportar o funcionamento do protocolo de encaminhamento. É também conveniente que o endereçamento dos elementos da rede de acesso seja privado, porque além de não existir necessidade que estes endereços sejam públicos e conhecidos em redes exteriores, também não é desejável por questões de segurança.

Nos utilizadores o esquema de endereçamento é diferente. Estes podem ter endereços privados ou públicos. Se os endereços são privados, o NT utiliza o protocolo NAT (*Network Address Translation*) [Srisuresh01] para efectuar a tradução de endereços. Existem duas formas de atribuir endereços IP (IPv4) aos terminais quando estes se ligam à rede de acesso: de uma forma manual pelo administrador da rede ou por DHCP [Droms97]. Nesta segunda forma, o servidor de DHCP atribui um endereço IP a cada utilizador de entre o conjunto de endereços IP disponíveis no ISP (associados aos utilizadores de uma determinada rede de acesso).

6.3 Reutilização do PPP na rede de acesso

O acesso do tipo *dial-up* é actualmente realizado pelo PPP e pelo conjunto de protocolos associados que permitem efectuar a autenticação do utilizador e a atribuição de endereços IP. O processo inicia-se com a autenticação do utilizador. O BAS pode autenticar o utilizador, funcionando como um *proxy* de AAA do ISP. O protocolo utilizado para efectuar a autenticação pode ser o RADIUS (*Remote Access Dial In User Service*) [Rigney00]. Se o utilizador não for aceite, a ligação PPP é terminada. Se for aceite, o ISP atribui um endereço IP ao utilizador através de DHCP. Considerando que o BAS pode fornecer o acesso a vários ISPs, no processo de autenticação o utilizador envia ao BAS o seu *login* e *password* para ser autenticado num determinado ISP. O *login* do utilizador contém também o nome do domínio (ISP) onde o utilizador se pretende ligar. Este

processo permite que um terminal seleccione um determinado ISP para lhe fornecer o serviço.

A partir deste momento, o utilizador tem um endereço IP atribuído e encontra-se ligado ao ISP através da sessão PPP. Numa rede de acesso com comutação de circuitos isto seria o suficiente para ser possível iniciar a transmissão de pacotes IP. No entanto, numa rede de acesso IP com recursos partilhados a situação é diferente. O PPP foi desenvolvido para funcionar sobre circuitos comutados ponto-a-ponto. Desta forma, ele não pode ser usado directamente sobre uma rede com recursos partilhados. Neste caso, será necessário realizar o *tunneling* das tramas PPP para possibilitar o seu transporte numa rede partilhada. O protocolo de criação e gestão de túneis normalmente utilizado em redes de nível 3 é o L2TP, e em redes de nível 2 é o PPPoE. Estes protocolos de *tunneling* permitem transportar sessões PPP sobre a *Ethernet* (PPPoE) ou sobre o IP (L2TP). Eles definem a forma como os túneis PPPoE ou L2TP são estabelecidos, geridos e terminados, e a forma como as tramas PPP são encapsuladas em tramas *Ethernet* ou em pacotes IP.

O PPPoE inicia-se com uma fase de descoberta: antes de ser iniciado o processo de configuração e estabelecimento da sessão PPP, o utilizador necessita de descobrir o endereço *Ethernet* do BAS para poder encaminhar as tramas PPP na rede de acesso. Nesta fase de descoberta o utilizador envia uma trama *Ethernet* para o endereço de *broadcast*, indicando, além de outras informações, o nome do serviço que pretende receber. O nome do serviço pode ser, por exemplo, o nome de um ISP ou uma classe de serviço. O formato da trama *Ethernet* enviada e o cabeçalho da trama PPP encapsulada na trama *Ethernet* são ilustrados na Figura 6-3 e na Figura 6-4, respectivamente. O campo tipo da trama *Ethernet* indica o tipo de protocolo transportado na trama, como por exemplo, o IP ou o PPPoE. Como na fase de descoberta ainda não foi iniciada a sessão PPPoE, o valor do identificador da sessão (ID Sessão na Figura 6-4) toma o valor 0. O BAS recebe a trama *Ethernet*, verifica se pode responder afirmativamente de acordo com o serviço pedido, e caso seja possível, envia uma resposta ao utilizador. Na possibilidade de existirem vários BAS, cada um mantendo contratos com ISPs diferentes, o utilizador pode receber mais do que uma resposta. Se isso acontecer, o utilizador escolhe um deles e envia uma mensagem de pedido de início de uma sessão PPPoE ao BAS escolhido. Neste ponto, o BAS prepara-se para iniciar a sessão PPPoE. Para o efeito, gera um identificador único da sessão e envia essa

informação ao utilizador. Após este processo, a sessão PPP e as suas fases de configuração da ligação, autenticação e autorização do utilizador podem começar.

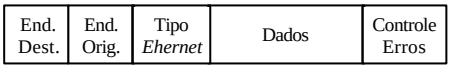


Figura 6-3 : Trama *Ethernet.*

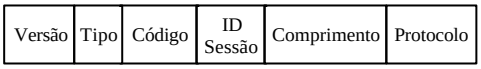


Figura 6-4 : Cabeçalho PPP.

Na Figura 6-5 ilustra-se o encapsulamento de IP em PPPoE numa trama *Ethernet*.

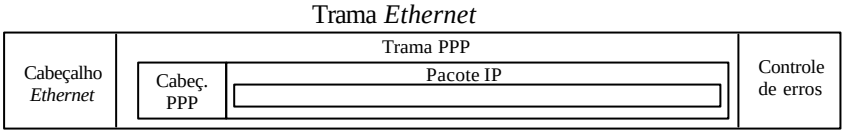


Figura 6-5 : Trama *Ethernet* numa ligação PPPoE.

A criação de túneis L2TP necessita do conhecimento prévio dos endereços IP dos elementos em cada extremo do túnel. Estes endereços IP são incluídos no cabeçalho de um datagrama IP que é utilizado para transportar as tramas L2TP, como ilustrado na Figura 6-7. O cabeçalho de uma trama L2TP é ilustrado na Figura 6-6.

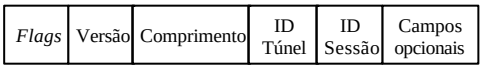


Figura 6-6 : Cabeçalho L2TP.

A identificação de um túnel com uma determinada origem e destino é efectuada pelo ID do túnel. O identificador da sessão é negociado aquando do estabelecimento da sessão. Assim, tal como no PPPoE, no início da configuração da sessão o identificador toma o valor 0. O transporte dos datagramas IP em L2TP é ilustrado na Figura 6-7. Por comparação com a Figura 6-5, verifica-se que o *overhead* introduzido é maior em L2TP do que em PPPoE.

Como todos os elementos da rede de acesso são de nível 3, e os elementos entre o BAS e o ISP podem também, alguns deles, ser de nível 3, os túneis que serão suportados

na rede de acesso e entre o BAS e o ISP serão túneis L2TP. Na rede local pode ser necessária a utilização de túneis PPPoE.

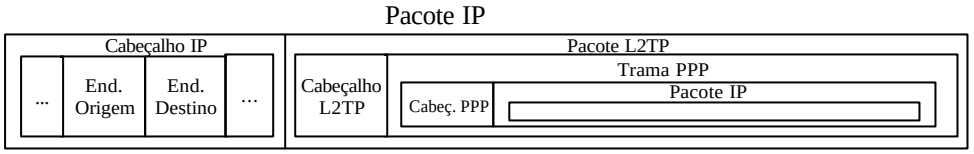


Figura 6-7 : Encapsulamento protocolar para suportar o PPP numa rede IP.

Nesta rede de acesso vão ser considerados dois casos diferentes: sessões PPP iniciadas nos terminais dos utilizadores e sessões PPP iniciadas no NT.

6.3.1 Sessões PPP iniciadas no terminal

Neste caso, o utilizador é a entidade que inicia o estabelecimento de uma sessão PPP. No processo de estabelecimento da sessão, o ISP atribui um endereço IP público ao utilizador.

As sessões PPP podem ser transportadas dos terminais até ao BAS através do estabelecimento de um túnel PPPoE entre o utilizador e o NT, e um túnel L2TP do NT até ao BAS. Considerando que o utilizador é constituído apenas por um terminal, o túnel L2TP pode começar no utilizador. Uma solução com túneis L2TP do NT até ao ISP, sem terminação no BAS, apenas permite a multiplexagem, no mesmo túnel, de sessões entre o mesmo ISP e utilizadores ligados ao mesmo NT. Para permitir no BAS a multiplexagem das sessões PPP (provenientes de diferentes NTs) de um mesmo ISP podem ser utilizados dois túneis L2TP, um entre o NT e o BAS e outro entre o BAS e o ISP.

Na Figura 6-8 apresentam-se, como exemplo, as camadas protocolares em cada elemento de rede quando as sessões PPP são iniciadas no terminal. Observam-se na figura as diversas camadas e encapsulamentos necessários para transportar a informação. Quando existem vários terminais na mesma LAN, é estabelecido um túnel PPPoE entre os utilizadores da LAN e o NT; o NT e o BAS mantêm um túnel L2TP partilhado por todos os utilizadores ligados ao mesmo NT. Nos MUX-IP, os pacotes passam transparentemente como se fossem pacotes IP normais. Como se considera que a rede de acesso é um domínio privado, os endereços IP utilizados para transportar os túneis L2TP na rede de acesso são privados. O BAS termina o túnel L2TP da rede de acesso, e mantém um outro túnel L2TP entre ele e o ISP, enviando por esse túnel as tramas PPP de todos os utilizadores dessa rede

de acesso que sejam clientes desse ISP. Note-se que, como as soluções *x*DSL actualmente disponíveis utilizam o ATM, é necessário transportar os pacotes IP em ATM através de AAL5. Uma solução futura muito mais atractiva é a possibilidade de transportar os pacotes IP directamente em *x*DSL.

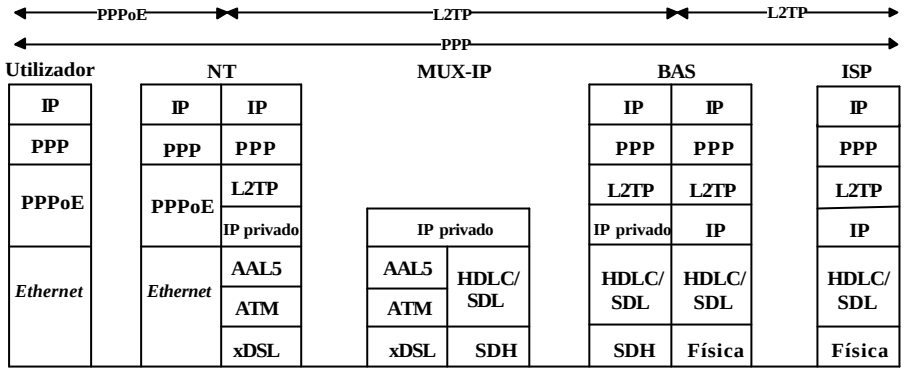


Figura 6-8 : Camadas protocolares dos acessos PPPoE e L2TP.

6.3.2 Sessões PPP iniciadas no NT

Neste caso, como as sessões PPP são iniciadas pelos NTs, os utilizadores têm endereços IP privados previamente atribuídos. No NT é utilizado o NAT para proceder à tradução entre endereços privados e endereços públicos, quando necessário. O estabelecimento de sessões PPP visa, neste caso, a obtenção de endereços IP públicos. Sempre que um utilizador necessita de aceder a um determinado ISP, o NT, se não possuir nenhum endereço IP disponível, estabelece uma nova sessão PPP com o ISP, sendo-lhe atribuído um novo endereço IP público. Note-se também que o NT pode ter configurados um conjunto de endereços IP públicos.

Uma outra característica importante deste cenário é o facto de não existir autenticação por utilizador, mas sim pelo NT que inicia a sessão. O NT inicia sessões PPP em nome de um conjunto de utilizadores.

Tal como no caso anterior podem ser estabelecidos dois túneis L2TP para transportar sessões PPP, um entre o NT e o BAS e outro entre o BAS e o ISP (Figura 6-9).

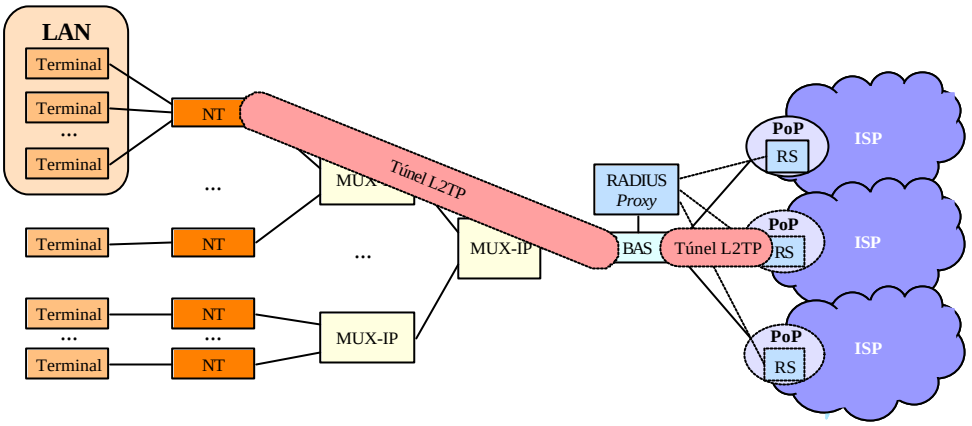


Figura 6-9 : Túneis L2TP.

6.3.3 Análise das soluções de reutilização do PPP na rede de acesso

As soluções com PPP anteriormente descritas colocam alguns problemas de implementação. O suporte do PPP requer que se usem protocolos de criação e gestão de túneis para poder encaminhar as tramas PPP numa rede partilhada. Os protocolos de *tunneling* introduzem vários problemas. Primeiro, os terminais, os NTs e o BAS necessitam de suportar um *software* específico para estabelecer, gerir e terminar os túneis. Este requisito introduz complexidade nos elementos da rede de acesso. Segundo, os sucessivos encapsulamentos da informação introduzem um *overhead* muito elevado, diminuindo a utilização dos recursos da rede. Ao nível deste último problema, a solução com PPPoE é mais eficaz. Além disso, existem já implementações disponíveis de PPPoE para soluções ADSL. Este tipo de soluções ainda se tornariam mais complexas com a introdução de suporte de QdS. Face a estas limitações, na secção seguinte vai ser descrita e discutida uma proposta de redes de acesso que resolve muitos dos problemas associados a esta solução com PPP. A rede de acesso proposta é baseada num conjunto de protocolos introduzidos recentemente, que permitem também o suporte de QdS diferenciada e a integração de serviços multimédia.

6.4 Redes de acesso IP com suporte de QdS

Uma rede de acesso para transporte de múltiplos serviços (incluindo os serviços multimédia) com suporte de QdS diferenciada requer um número de funcionalidades acrescidas: reserva de recursos, sinalização da sessão sincronizada com a reserva de recursos, gestão das políticas de QdS e funcionalidades de AAA numa base utilizador/serviço/QdS. Nas secções seguintes serão descritos os protocolos que deverão ser suportados pelos diversos elementos da rede de acesso para garantir todas estas funcionalidades.

6.4.1 Reservas de recursos (RSVP e extensões do RSVP)

O RSVP, descrito no capítulo 3, é o protocolo de reserva de recursos utilizado normalmente na arquitectura IntServ. Nesta arquitectura as reservas de recursos são estabelecidas e mantidas fluxo-a-fluxo. Como já foi referido no capítulo 3, a carga de sinalização dos elementos da rede é muito elevada, pois o número de mensagens processadas é proporcional ao número de sessões activas na rede. Uma carga de sinalização elevada pode degradar o desempenho de cada elemento da rede.

Recentemente foi proposta uma extensão do RSVP para implementar a agregação de reservas individuais e suportar uma arquitectura DiffServ. Nesta extensão, descrita no capítulo 3 e estudada no capítulo 9, é reservada uma quantidade de largura de banda entre dois elementos da rede para um agregado de fluxos. O controle de admissão é efectuado apenas para um agregado de fluxos, e por isso, os elementos do núcleo da rede necessitam apenas de manter o estado de reserva do agregado e processar as mensagens de sinalização quando a largura de banda do agregado necessita de uma actualização. No entanto, os elementos de entrada de um agregado continuam a necessitar de manter o estado de cada fluxo e efectuar reservas fluxo-a-fluxo. No caso da rede de acesso em questão, considerando que podem ser configurados agregados entre os NTs e o BAS, apenas estes elementos necessitam de manter o estado de reserva fluxo-a-fluxo. Os MUX-IP mantêm apenas o estado da reserva dos agregados. Ainda assim, estes elementos necessitam de suportar o RSVP para efectuar as reservas e actualizações das reservas dos agregados.

Dado que a maior parte das aplicações multimédia estão a ser desenvolvidas assumindo que o protocolo de reserva de recursos é o RSVP, considera-se que o suporte do RSVP na rede de acesso é praticamente inevitável. Além disso, como o RSVP foi

actualizado para efectuar a reserva de recursos para fluxos individuais e agregados de fluxos (possibilitando o suporte de ambas as arquitecturas IntServ e DiffServ), o seu suporte na rede de acesso permite estabelecer compromissos entre custo e desempenho da rede: MUX-IP capazes de processar um número maior de fluxos terão um custo maior, mas permitirão atingir uma maior utilização dos recursos.

6.4.2 Escalonamento de pacotes

A rede de acesso tem de suportar vários tipos de tráfego simultaneamente: o tráfego de sinalização, o tráfego reservado (utilizado para fornecer ligações semi-permanentes), o tráfego de sessões comutadas (*on-demand*), e o tráfego de melhor esforço. As disciplinas de escalonamento de pacotes nos elementos da rede de acesso necessitam de dois níveis hierárquicos. No primeiro nível, o tratamento que é atribuído a cada tipo de tráfego é diferenciado com base em prioridade estrita com três prioridades: uma para tráfego de sinalização, outra para tráfego reservado e tráfego de sessões comutadas, e outra para tráfego de melhor esforço. O tráfego de sinalização, sendo o que necessita de garantias mais estritas, terá a prioridade mais elevada. No entanto, no sentido de proteger a rede de utilizadores maliciosos, é necessário associar um limitador de taxa a esta prioridade nos pontos de entrada da rede de acesso, ou seja, nos NTs para tráfego no sentido ascendente e no BAS para tráfego no sentido descendente. Ao tráfego de melhor esforço será atribuída a prioridade mais baixa, pois este não tem requisitos de largura de banda, atrasos ou perdas. Finalmente, é atribuída a prioridade do meio ao tráfego reservado e de sessões comutadas. O tráfego nesta prioridade tem ainda de ser diferenciado utilizando o algoritmo WFQ ou uma das suas variantes (apresentados no capítulo 3), que permitem atribuir uma largura de banda mínima para cada fluxo ou agregado de fluxos. Deste modo, o escalonamento de pacotes deste tráfego na rede de acesso será efectuado através de um mecanismo de prioridade estrita num primeiro nível e de WFQ num segundo nível.

6.4.3 SIP

Nesta arquitectura de redes de acesso, o protocolo de sinalização de sessões escolhido é o SIP. O SIP foi definido pelo IETF para estabelecer, modificar e terminar sessões multimédia através da *Internet*. Este protocolo foi desenvolvido no seu início para Voz sobre IP (VoIP) e tem vindo a sofrer alterações e actualizações para ser aplicável a outros serviços e aplicações multimédia. Os elementos principais da arquitectura do SIP são

utilizadores (*User Agents* - UA), servidores de registo (*registrars*), servidores *proxy* e servidores de redireccionamento (*redirect*). Os UA são os extremos das sessões; os servidores de registo são responsáveis por realizar o processo de registo dos UAs; os servidores *proxy* são *routers* da camada de aplicação que efectuam a expedição das mensagens SIP; e os servidores de redireccionamento retornam localizações alternativas dos UAs ou servidores. Algumas das funções oferecidas pelo protocolo SIP são de extrema importância nas redes de acesso. Por exemplo, o SIP suporta o registo dos UAs, suporta portabilidade (isto é, permite estabelecer sessões a partir de um mesmo terminal em várias localizações diferentes), e pode ser sincronizado com os processos de reserva de recursos e de AAA.

No SIP os utilizadores são identificados por URLs (*Uniform Resource Locators*), e são semelhantes a endereços de *e-mail* (como exemplo, sus@domínio-A.org). Em cada sessão, o URL é traduzido para um endereço IP através do uso de um servidor *proxy* do SIP e de uma procura DNS (*Domain Name System*) [Postel94]. Note-se que, ao contrário dos endereços IP, o URL de um utilizador é sempre o mesmo.

O SIP tem dois tipos de mensagens: as mensagens de pedido e as mensagens de resposta ao pedido. As mensagens de pedido pedem uma acção específica por parte do UA destino ou do servidor *proxy*. As mensagens INVITE, ACK, BYE e REGISTER são alguns exemplos de mensagens de pedido. As mensagens de resposta às mensagens de pedido são as mensagens com números 1*xx*, 200, e de 3*xx* até 6*xx*. A mensagem INVITE, enviada pelo UA origem da sessão, é a mensagem que inicia o processo de estabelecimento de uma sessão. A mensagem ACK, enviada pelo UA origem da sessão no final do processo de estabelecimento da sessão, é uma mensagem de confirmação da aceitação da sessão. A mensagem BYE, enviada por um dos UA, é a mensagem que inicia o processo de terminação da sessão. A mensagem REGISTER, enviada por cada UA, é a mensagem que inicia o processo de registo do UA. As mensagens 1*xx* são de informação; a mensagem 200 é uma mensagem de aceitação do pedido; as mensagens 3*xx* são de redireccionamento da chamada; as 4*xx* indicam existência de erro no cliente; as 5*xx* indicam erros no servidor; e as 6*xx* indicam erros globais.

Considere-se como exemplo a troca de mensagens de sinalização no processo de iniciação de uma sessão entre utilizadores de redes de acesso diferentes (A e B), ilustrada na Figura 6-10. Este processo envolve dois servidores *proxy*, um que coordena o processo

de sinalização no domínio da origem da sessão, e outro no domínio do destino. O processo de sinalização é iniciado pela mensagem INVITE (mensagem 1). Esta mensagem contém uma descrição do tipo de sessão pedida. Uma sessão multimédia pode conter vários fluxos de tráfego. Por exemplo, uma aplicação de videoconferência pode conter fluxos de áudio, vídeo e dados, tendo cada fluxo de tráfego requisitos diferentes de QdS. O SIP utiliza o protocolo SDP (*Session Description Protocol*) [Handley98] para descrever as capacidades de envio e de recepção de cada terminal (ou equipamento terminal), e as características das sessões e dos seus fluxos de tráfego associados. A mensagem INVITE inclui também o URL SIP do destino, que vai ser usado para encaminhar o pedido. Se o endereço IP do UA destino é conhecido, a mensagem INVITE pode ser enviada directamente para o destino; caso contrário, é enviada para o servidor *proxy* do domínio origem. Após receber esta mensagem, o *proxy* usa o DNS para determinar o endereço IP do servidor *proxy* do domínio destino (mensagens 2 e 3). A mensagem INVITE é então enviada para este *proxy*, que determina o endereço IP do UA destino através de uma procura na base de dados do servidor de localização associado (mensagens 5 e 6). O UA destino responde com uma mensagem 180 RINGING (mensagem 8) indicando que a mensagem INVITE foi recebida e que está a tomar as devidas acções para decidir se pode aceitar o pedido. Quando o UA destino decide aceitar a sessão, envia uma mensagem 200 OK (mensagem 11) indicando também que suporta o tipo de sessão proposta pelo UA origem. O passo final do estabelecimento da sessão é o envio da mensagem ACK (mensagem 14) pelo UA origem, que é uma confirmação de que também pode aceitar a sessão. Note-se que esta mensagem já não atravessa os servidores *proxy*, pois o UA origem tem conhecimento do endereço destino. Nesta fase, a sessão é iniciada (mensagem 15) através de outro protocolo, por exemplo o RTP (*Real-Time Transport Protocol*). A sessão é terminada quando um dos UA (origem ou destino) envia uma mensagem BYE, e após esta ser confirmada por uma mensagem 200 OK.

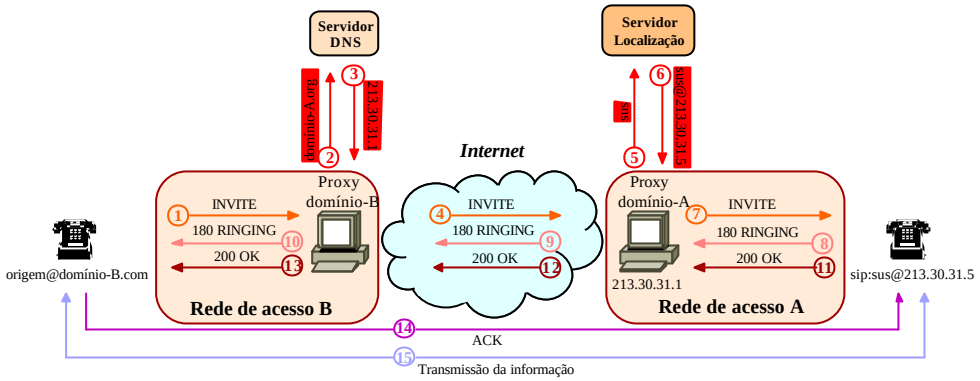


Figura 6-10 : Processo de iniciação da sessão com SIP.

O processo de registo de um utilizador é ilustrado na Figura 6-11. Quando é atribuído um endereço IP a um utilizador, este envia uma mensagem de pedido REGISTER para o servidor *proxy* ou de redireccionamento, para o informar dos endereços IP que está a utilizar (mensagem 1). O servidor *proxy* envia a mensagem para o servidor de registo, que pode estar ou não localizado no mesmo elemento do *proxy*. O servidor de registo envia o endereço IP e o URL para o servidor de localização que armazenará esta informação numa base de dados (mensagem 2). Mais tarde, esta informação pode ser utilizada pelo servidor *proxy* para localizar o utilizador.

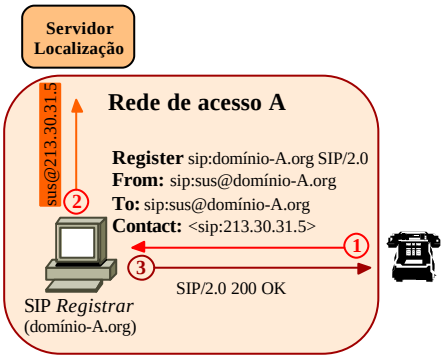


Figura 6-11 : Registo de um utilizador com SIP.

Na rede de acesso em definição, o BAS deverá incluir um servidor *proxy* SIP. É também necessário um servidor de registo, mas este pode servir mais do que um BAS. O servidor *proxy* terá acesso, através do servidor de localização, a uma base de dados que contém os endereços IP dos utilizadores no seu domínio.

6.4.4 Endereçamento SIP e selecção de um ISP

Uma função essencial da rede de acesso que pode ser facilmente realizada utilizando o SIP é a selecção de um ISP. Para o efeito, o utilizador especifica no URL o domínio que identifica o ISP a seleccionar. Por exemplo, se existirem dois ISPs disponíveis, isp-X.com e isp-Y.com, e o utilizador deseja aceder a um serviço através de isp-X, ele utilizará o URL utilizador@isp-X.com e o pedido será encaminhado pelo domínio do isp-X.

Recentemente, foi proposta uma extensão ao DHCP que permite que um único servidor de DHCP anuncie um ou mais servidores *proxy* do SIP [Schulzrin02a]. Para o efeito, foi introduzida no DHCP a opção "SIP *server*" que transporta os endereços IP ou os nomes do domínio DNS, que podem ser mapeados em um ou mais servidores *proxy*. Este é um dos métodos utilizados para que seja possível ao UA obter os endereços de um servidor *proxy*. Assim, quando é atribuído um endereço IP ao UA, este é também informado dos vários ISPs que estão disponíveis para servir os utilizadores da rede de acesso onde se encontra ligado.

6.4.5 Suporte de QdS e segurança com SIP

O SIP inclui mecanismos, denominados de pré-condições, para coordenar a sinalização da sessão com o estabelecimento de reservas de recursos extremo-a-extremo, ou com o estabelecimento de túneis de segurança. Uma pré-condição é uma condição que terá de ser obrigatoriamente verificada por um ou mais utilizadores. A continuação do processo de sinalização pode estar (ou não) dependente da satisfação desta condição. O suporte de QdS e de segurança são funcionalidades adicionais do protocolo SIP base.

Existem duas opções para efectuar a interacção entre a sinalização da sessão e a reserva de recursos, resultando em sessões com QdS garantida e sessões com QdS opcional [Marshall01, Sinnreich01]. Nas sessões com QdS garantida, o emissor só receberá uma mensagem 180 RINGING, indicando que a sessão pode ter início, quando o processo de reserva de recursos tiver terminado com sucesso. Este tipo de serviço é equivalente ao serviço de telefonia em redes de circuitos comutados. O tempo de estabelecimento da sessão é mais elevado do que quando não é exigida esta sincronização, devido à dependência do estabelecimento da reserva de recursos. Em sessões com QdS opcional, o estabelecimento da sessão e o processo de reserva de recursos são independentes e podem prosseguir de uma forma concorrente. No caso de o estabelecimento das reservas falhar, o

emissor é notificado, e tem a opção de continuar a sessão com o serviço de melhor esforço. O tempo de estabelecimento da sessão é inferior e o modelo de melhor esforço pode ainda ser aceitável para os utilizadores.

As pré-condições são atributos que podem ser incluídos nas mensagens SIP. O atributo de QdS indica se a reserva de recursos extremo-a-extremo é opcional ou obrigatória, e em que sentido (sentido do envio, recepção ou ambos). O atributo pode também conter um parâmetro de confirmação que, quando presente na mensagem, solicita a confirmação do outro extremo da comunicação de que as pré-condições foram satisfeitas, através do envio de uma mensagem denominada de COMET. Se o atributo de QdS é obrigatório, o UA que recebe a mensagem SIP não deve continuar com o processo de sinalização enquanto não for efectuada a reserva de recursos no sentido indicado na mensagem. Caso contrário, se a pré-condição é opcional, o processo de sinalização continua mas o participante da comunicação deve tentar reservar os recursos no sentido indicado. Como uma sessão pode conter vários fluxos de tráfego, existem pré-condições diferentes para cada fluxo.

```
…
m=audio 49170 RTP/AVP 0
a=qos: mandatory recv confirm
m=video 51372 RTP/AVP 31
a=secure: mandatory sendrecv
m=application 32416 udp wb
a=orient: portrait
a=qos: optional sendrecv
a=secure: optional sendrecv
…
```

Figura 6-12 : Atributos de QdS e segurança numa mensagem INVITE.

Considere-se a Figura 6-12 que apresenta uma parte do corpo da mensagem SDP que pode ser incluída na mensagem INVITE. Os nomes associados ao tráfego (áudio, vídeo, etc.) e os endereços de transporte são identificados pela letra *m*, e os seus atributos (QdS ou segurança) pela letra *a*. No exemplo, existem três fluxos de tráfego diferentes, áudio, vídeo e dados, cada um com os seus requisitos de QdS e de segurança. O fluxo de áudio solicita uma QdS obrigatória no sentido da recepção e uma confirmação do UA destino após a reserva ter sido efectuada com sucesso; o fluxo de vídeo solicita uma segurança obrigatória nos dois sentidos; e na aplicação UDP (*User Datagram Protocol*), as pré-condições de QdS e de segurança são opcionais. Neste caso, a sinalização da sessão pode prosseguir

logo após a reserva de recursos para o fluxo de áudio e a configuração dos túneis de segurança do fluxo de vídeo terem sido efectuados com sucesso.

6.4.6 COPS e servidores de AAA

Uma rede com suporte de múltiplos serviços e de QdS diferenciada coloca requisitos adicionais nas funcionalidades de AAA. Estas funções têm agora de ser realizadas a diferentes níveis: utilizador, serviço e QdS. Por exemplo, um serviço pode ser autorizado apenas a alguns utilizadores, entre pontos autorizados, com determinados requisitos de QdS (que dependem do utilizador a solicitar o serviço), apenas em determinados intervalos de tempo [QoSForum98]. O utilizador tem também de ser taxado de acordo com as características do serviço que recebe.

A tendência actual é incluir as funcionalidades de AAA no âmbito mais geral da arquitectura de gestão de políticas de QdS. Existem dois elementos arquitecturais principais no âmbito do controle de políticas: o PEP (*Policy Enforcement Point*) e o PDP (*Policy Decision Point*) [Yavatkar00] (Figura 6-13). O PDP é o elemento que toma decisões com base em políticas que estão descritas num repositório de políticas, nos servidores de AAA e em outras entidades. O PEP é o elemento que implementa na rede as decisões tomadas pelo PDP. O PEP é um componente do nó da rede e o PDP é uma entidade remota que pode residir num servidor de políticas. Normalmente existe um PDP num domínio de rede e vários PEPs (Figura 6-14). O repositório de políticas é uma base de dados remota (um serviço de directórios ou um sistema de ficheiros de rede). Para trocar informação com o repositório de políticas e para armazenar ou retirar informação relacionada com as políticas de QdS, o PDP usa o protocolo LDAP (*Lightweight Directory Access Protocol*) [Wahl97].

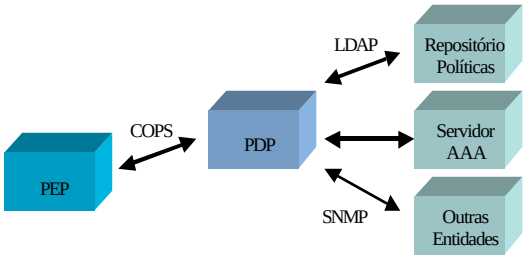


Figura 6-13 : Modelo de políticas com COPS.

O PDP pode recorrer a diversos mecanismos e protocolos para tomar as suas decisões (autenticação e autorização de um utilizador, contabilização do serviço acedido pelo utilizador de acordo com os seus requisitos, armazenamento de informação relacionada com as políticas de QdS, etc.). O COPS [Boyle00] é o protocolo normalmente utilizado para efectuar a troca de informação entre o PEP e o PDP. A interacção entre o PEP e o PDP é ilustrada na Figura 6-15. Existem três tipos de mensagens COPS: mensagens de pedido, mensagens de decisão, e mensagens de relatório. A mensagem de pedido COPS é enviada do PEP ao PDP. Esta mensagem pode conter um ou mais elementos de política. Um exemplo de um elemento de política é a informação de autenticação. O PDP responde com uma decisão que é colocada numa mensagem de decisão COPS, que especifica a acção que o PEP deve tomar. As mensagens de relatório podem ser utilizadas para efectuar a contabilização do utilizador.

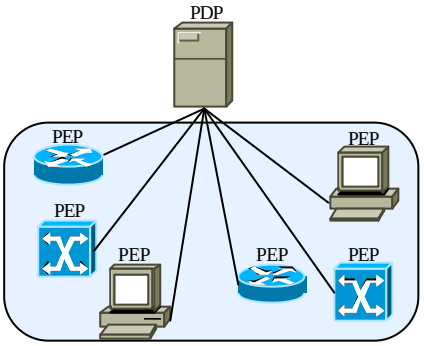


Figura 6-14 : Vários PEPs a partilhar um PDP.

O PEP notifica o seu PDP de todos os eventos que requerem uma decisão baseada em políticas de QdS. Deste modo, o PDP é um elemento de agregação que pode ser utilizado para monitorizar a actividade da rede e para armazenar a informação relacionada com a contabilização dos serviços acedidos pelos utilizadores. Em relação a este último ponto, o PEP pode enviar mensagens COPS de relatório ao PDP para este efectuar a contabilização dos serviços.

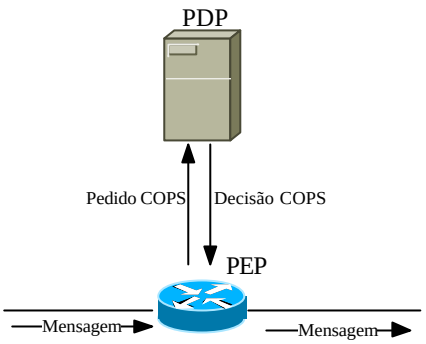


Figura 6-15 : Interacção entre o PEP e o PDP.

Uma funcionalidade importante do COPS é a capacidade de fazer o *download* das configurações de QdS nos elementos da rede. O PDP necessita apenas de saber que um elemento usa um determinado conjunto de regras, e faz o *download* destas regras para o elemento. As configurações de políticas de QdS incluem os mecanismos para classificação dos pacotes, definição das taxas dos formatadores de tráfego, definição das classes de serviço (no caso de uma rede DiffServ), as acções a tomar no caso de o tráfego não estar de acordo com um determinado perfil, os mecanismos de escalonamento e de preferência de descarte dos pacotes (para aplicar aos pacotes de acordo com a sua classificação).

Os servidores de AAA acedidos pelo PDP são normalmente o RADIUS, e mais recentemente, o DIAMETER [Calhoun01]. O protocolo RADIUS tem sido usado para fornecer serviços de AAA a sessões PPP *dial-up* e acessos a servidores de terminais. O DIAMETER é um protocolo que substituirá o RADIUS, pois apresenta uma maior escalabilidade e suporta mobilidade. A informação distribuída pelo protocolo DIAMETER encontra-se, assim como no RADIUS, na forma de pares atributo/valor (AVP – *Attribute Value Pair*). O conjunto dos AVPs presentes nas mensagens DIAMETER são utilizados para transportar informação específica de autenticação dos utilizadores, autorização do serviço e informação de utilização dos recursos.

O protocolo DIAMETER básico concentra-se apenas nas capacidades de negociação e no envio das suas mensagens. Existem três tipos de mensagens DIAMETER: as mensagens de pedido, de resposta ao pedido e de indicação. Além do protocolo básico, existem extensões específicas do DIAMETER para cada aplicação.

Na Figura 6-16 é ilustrada a extensão do DIAMETER para mensagens SIP. O servidor *proxy* contém um cliente DIAMETER. Quando o servidor *proxy* recebe uma

mensagem SIP que necessita de uma resposta do servidor de AAA, contacta o servidor DIAMETER através do cliente DIAMETER.

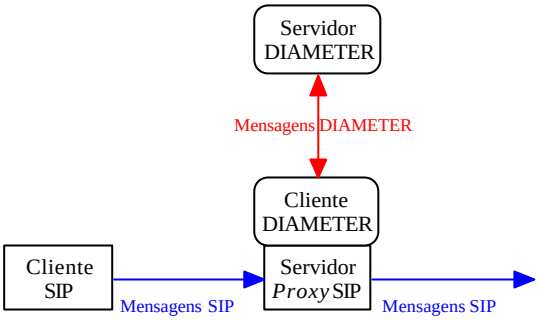


Figura 6-16 : Extensão DIAMETER para SIP.

Usando o COPS para realizar a gestão das políticas de QdS, o PDP tem acesso a um servidor de AAA para implementar as funções de autenticação. Este cenário, em relação ao anterior, adiciona a gestão das políticas de QdS. Tanto nos processos de registo como de início de uma chamada, o servidor *proxy* envia uma mensagem de pedido COPS ao PDP, e este contacta um servidor DIAMETER através de mensagens DIAMETER de pedido e resposta (Figura 6-17).

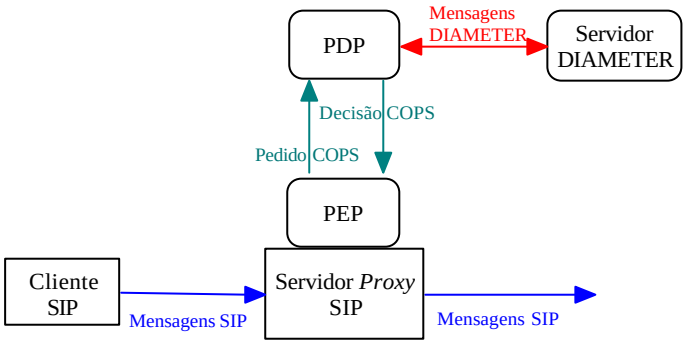


Figura 6-17 : COPS e DIAMETER.

Em [Basilier01] são descritos os requisitos a que devem obedecer os futuros protocolos de AAA (por exemplo, uma nova extensão do DIAMETER) para serem compatíveis com o SIP e ao mesmo tempo realizar também algumas funções que neste momento estão a cargo do COPS. Estes requisitos incluem: (1) a distribuição da informação de políticas do utilizador ou da rede para os servidores *proxy* SIP; (2) a

implementação de qualquer política para autorizar a sessão (por exemplo, período horário e requisitos de QdS); (3) a distribuição dos perfis do utilizador, na forma de políticas, para os servidores SIP, com base nos serviços que acede e suas características, e com base no nível de segurança requerido para cada serviço; (4) a atribuição de um servidor SIP a um utilizador no instante de registo, por exemplo, com base em políticas de QdS e distribuição da carga. Quando todos estes requisitos forem implementados numa extensão do DIAMETER, este poderá ser usado directamente com o SIP, realizando as funções que na rede de acesso definida estão atribuídas ao COPS. Como não existe ainda uma implementação do DIAMETER que inclua todos estes requisitos, considera-se que na rede de acesso proposta, o COPS realiza todas as funções relacionadas com as políticas de QdS e recorre aos servidores de DIAMETER para efectuar a autenticação do utilizador.

Na rede de acesso proposta todos os elementos, exceptuando os UA, incluem um PEP. Os PEPs nos NTs e MUX-IP necessitam apenas de suportar o *download* das configurações de QdS. O PEP no BAS necessita de suportar a activação das mensagens de COPS em eventos específicos relacionados com o processo de sinalização SIP. Note-se que na rede de acesso, todos os fluxos de pacotes atravessam o BAS, mesmo os fluxos pertencentes a sessões entre utilizadores da mesma rede de acesso. Deste modo, o único elemento que necessita de trocar informação de contabilização com o PDP é o BAS.

6.4.7 Sumário das principais características da rede de acesso

Nas secções anteriores foram descritos os protocolos e as tecnologias que vão ser utilizados na rede de acesso. Nesta secção é apresentada a solução geral com base nos protocolos e tecnologias anteriores. Na Figura 6-18 são apresentados os módulos de *software* que cada elemento da rede deverá implementar. Para simplificar a figura, o repositório de políticas não é representado.

Como já foi referido anteriormente, os terminais e todos os elementos da rede de acesso devem suportar o RSVP. Se a rede de acesso suportar a arquitectura DiffServ, o NT e o BAS têm de possuir a funcionalidade de configurar agregados de reservas individuais e aceitar ou rejeitar um novo fluxo com base na largura de banda reservada para os agregados; os MUX-IP necessitam apenas de manter o estado de cada agregado e efectuar a sua actualização. O BAS contém também um servidor *proxy* para controlar todos os pedidos de estabelecimento de sessões, e efectuar a localização de um utilizador no seu domínio. O BAS contém ainda um PEP para suportar a autenticação e a autorização dos

utilizadores, e trocar informação de contabilização com o PDP. Os NTs e os MUX-IP incluem também PEPs, mas apenas para suportar o *download* das configurações de QdS. O PDP pode estar incluído num elemento da rede de acesso ou fazer parte de um servidor de políticas que gere várias redes de acesso.

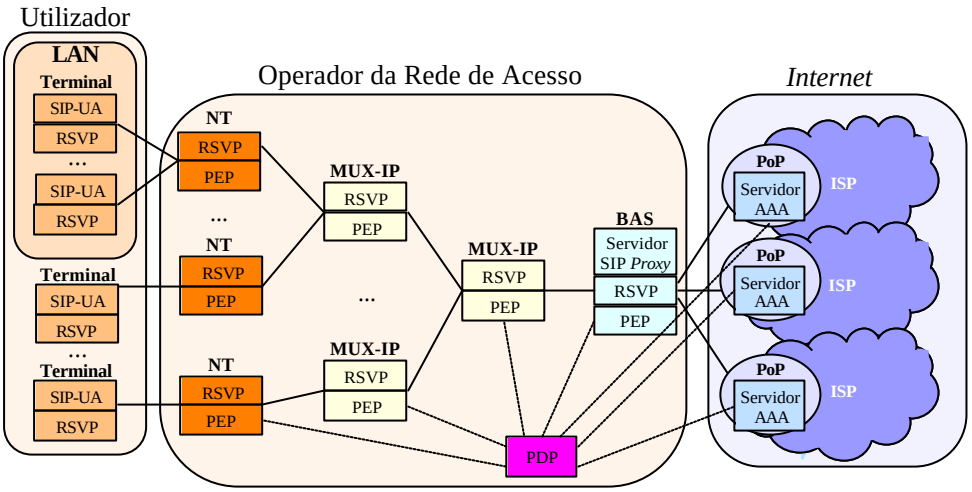


Figura 6-18 : Elementos da rede de acesso e protocolos suportados.

Todos estes protocolos são necessários na rede de acesso quando os utilizadores acedem a serviços a pedido (*on-demand*) e com requisitos de QdS. Quando se trata, por exemplo, de um serviço de acesso à *Internet* sem requisitos de QdS, é utilizado o modelo de serviço de melhor esforço, e por isso, não se recorre ao RSVP. No entanto, o utilizador necessita de ser autenticado antes de aceder ao serviço, e deste modo, é necessário suportar o COPS para fornecer estas funções. Além disso, o SIP é necessário para activar o COPS e configurar a sessão.

No caso de serviços semi-permanentes existe uma reserva prévia de largura de banda para estes serviços entre o utilizador e o ISP, e por isso, o RSVP não é necessário. As políticas de QdS relacionadas com esta reserva de largura de banda são configuradas nos elementos da rede através de *download* usando o COPS. Ambos os formatadores do NT e do BAS, respectivamente no sentido ascendente e descendente, têm de ser configurados com uma taxa igual à largura de banda reservada.

6.4.8 Exemplo de fluxo de mensagens na rede de acesso

A Figura 6-19 ilustra um exemplo de um fluxo de mensagens (desde o início até à terminação de uma sessão) numa rede de acesso no caso de uma sessão bidireccional com QdS garantida. As mensagens entre o ISP e o UA destino não são representadas. Considera-se que a reserva de recursos é efectuada nos dois sentidos.

Quando um utilizador deseja iniciar uma sessão, envia uma mensagem INVITE em direcção ao destino. Esta mensagem inclui um atributo de QdS que indica que a reserva de recursos é obrigatória nos dois sentidos. O servidor *proxy* captura a mensagem INVITE e envia uma mensagem de pedido COPS ao PDP para efectuar a autenticação do utilizador com base nas políticas de QdS. A mensagem de pedido inclui informação que é lida da mensagem INVITE, como por exemplo, os endereços SIP da origem e do destino, o identificador da sessão SIP, informação sobre as características dos fluxos de tráfego e os atributos de QdS. Após receber uma resposta positiva através da mensagem de decisão COPS, a mensagem INVITE é enviada para o destino através da rede do ISP seleccionado pelo utilizador (isp-X). Como a mensagem INVITE inclui pré-condições, o utilizador destino tem de responder com uma mensagem 183 SESSION PROGRESS utilizando o mecanismo de fiabilidade no envio desta mensagem [Rosenberg01]. A mensagem 183 inclui o endereço IP do utilizador destino e contém os atributos de QdS e de segurança para cada fluxo de tráfego que tem associado uma pré-condição. Se o receptor é capaz de suportar as pré-condições (por exemplo, se é capaz de reservar recursos para a sessão e de configurar os túneis de segurança), copia apenas os atributos de QdS e de segurança da mensagem INVITE para a mensagem 183; caso contrário, pode propor novos atributos com requisitos diferentes (por exemplo, pode querer mudar o requisito de QdS de obrigatório para opcional) e, neste caso, começa uma fase de negociação com o envio de uma nova mensagem INVITE por parte do utilizador origem. O mecanismo de fiabilidade utilizado no envio da mensagem 183 é necessário porque o envio desta mensagem é crucial para o estabelecimento da sessão. De acordo com o mecanismo de fiabilidade, o utilizador origem tem de responder à mensagem 183 com uma mensagem PRACK e o utilizador destino acusa a recepção da mensagem com o envio da mensagem 200 OK. O utilizador destino também solicita ao utilizador origem, na mensagem 183, uma confirmação, que terá de ser enviada quando as pré-condições estiverem estabelecidas.

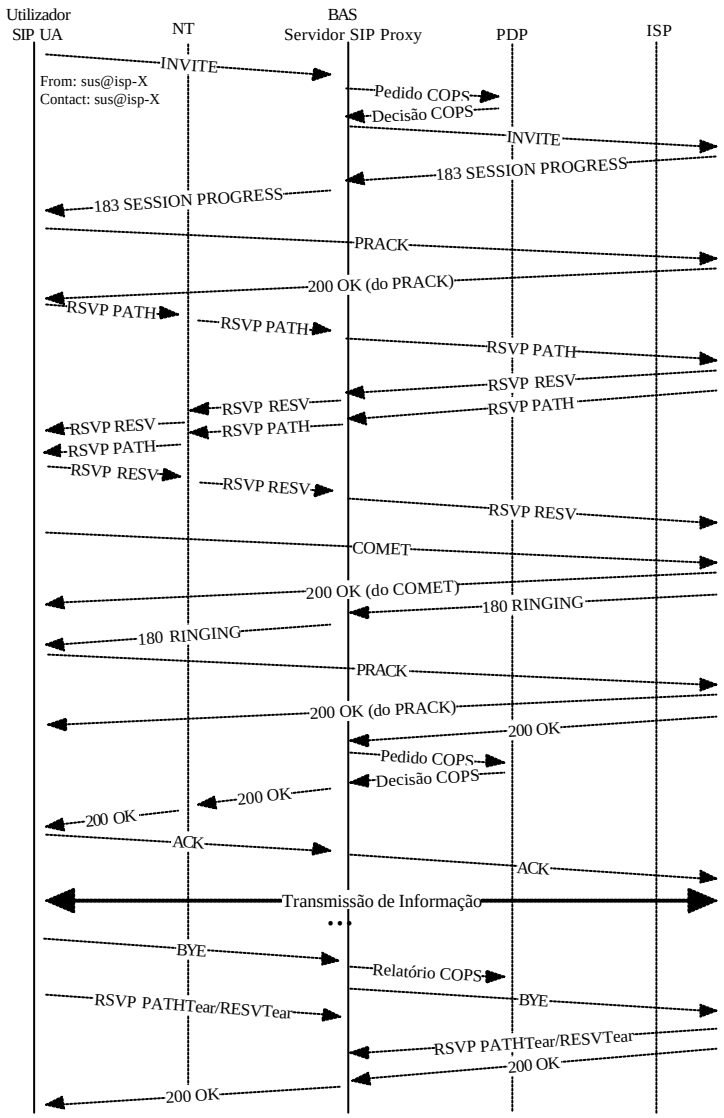


Figura 6-19 : Diagrama temporal de uma sessão multimédia de acesso à *Internet* com SIP, COPS e RSVP.

Nesta fase, inicia-se o processo RSVP através do envio de mensagens PATH e RESV entre o emissor e o receptor [Schulzrinne99]. As mensagens RSVP nos dois sentidos podem ser sobrepostas. Como foi referido anteriormente, o RSVP pode ser utilizado em ambas as arquitecturas IntServ e DiffServ. Nas arquitecturas DiffServ (que é o caso deste exemplo), os MUX-IP processam apenas as mensagens de reserva dos agregados; o processamento fluxo-a-fluxo é apenas efectuado nos NTs e no BAS. Numa

arquitectura IntServ, as mensagens RSVP não passariam transparentemente nos MUX-IP, e necessitariam de ser processadas fluxo-a-fluxo nestes. Após ter sido efectuado o processo de reserva de recursos com sucesso no sentido do emissor para o receptor, o emissor envia uma mensagem COMET para o receptor. Esta mensagem é uma extensão do SIP que inclui o corpo SDP da mensagem, indicando o estado de cada uma das pré-condições como sendo de sucesso ou insucesso ("success" ou "failure"). O receptor ao receber a mensagem COMET indicando o sucesso das pré-condições, e ao ter conhecimento do sucesso da reserva de recursos no sentido do receptor para o emissor, determina que todas as pré-condições foram estabelecidas com sucesso nos dois sentidos, e envia uma mensagem 180 RINGING para o emissor indicando que a sessão pode começar. Neste caso, a mensagem 180 também necessita do mecanismo de fiabilidade no seu envio (PRACK e 200 OK).

Finalmente, o receptor envia uma mensagem 200 OK (em resposta ao INVITE). Esta mensagem contém informação das características e requisitos da sessão que foi aceite. As características da sessão acordada, na mensagem 200 OK, podem ser diferentes das características da sessão incluídas na mensagem INVITE no início do processo de estabelecimento da sessão. Como já foi referido, se o UA destino não consegue suportar algumas características e requisitos da sessão propostos pelo UA origem, pode ser realizado um processo de negociação. Para poder actualizar a informação resultante deste processo, o servidor *proxy* no BAS captura a mensagem 200 OK e troca informação com o PDP para actualizar algumas modificações que tenham sido efectuadas na informação da sessão, e iniciar as funcionalidades de contabilização. Após o envio da mensagem ACK, a informação da sessão pode ser transmitida entre os dois UAs.

Para terminar a sessão, um dos utilizadores envia uma mensagem BYE e o outro responde com 200 OK. As reservas de recursos têm também de ser removidas em ambos os sentidos. O servidor *proxy* no BAS captura a mensagem BYE e envia uma mensagem de relatório COPS para o PDP para terminar a contabilização da sessão.

Num cenário em que a QdS seja opcional, o destinatário da sessão, em vez de enviar ao utilizador origem uma mensagem 183 SESSION PROGRESS, envia uma mensagem 180 RINGING e a sessão pode iniciar-se sem que o processo de reserva de recursos esteja finalizado. As mensagens PATH e RESV são enviadas sem sincronismo com o estabelecimento da sessão. Neste cenário, as mensagens 183 SESSION PROGRESS, COMET e PRACK apresentadas na Figura 6-19 não são utilizadas.

O diagrama temporal de uma sessão de acesso à *Internet* sem requisitos de QdS é semelhante ao apresentado na Figura 6-19 sem as mensagens RSVP, 183 SESSION PROGRESS, COMET e PRACK.

6.5 Conclusões e trabalho futuro

Neste capítulo foi proposta uma arquitectura para redes de acesso IP com suporte de QdS diferenciada e integração de serviços multimédia. Os elementos principais da rede proposta são os NTs, os *multiplexers* IP e o BAS. Os NTs fazem a interface entre a rede de acesso e a rede local de utilizadores, e o BAS faz a interface entre a rede de acesso e os ISPs. Considerou-se que a rede fornece redundância apenas ao nível físico, por exemplo, através do SDH, o que permitiu simplificar um número de funções, tal como o endereçamento e o encaminhamento, e permitiu reduzir o custo da rede de acesso. O endereçamento no interior da rede de acesso pode ser privado e o encaminhamento é muito simplificado porque existe apenas um percurso possível entre o BAS e o utilizador final.

Considerou-se inicialmente a possibilidade de reutilizar o PPP neste tipo de redes de acesso. Verificou-se que este tipo de solução, mesmo sem suportar QdS, introduzia uma grande complexidade adicional devido à necessidade de utilização dos protocolos de criação e gestão dos túneis. Em face destas limitações foi proposta uma nova arquitectura para suporte de QdS em redes de acesso, que incorpora algumas tecnologias e protocolos recentemente introduzidos, os quais permitem suportar a integração dos serviços multimédia de banda larga com um controle completo dos recursos da rede e da sua gestão, das regalias dos utilizadores, e dos serviços prestados e da sua qualidade. Estas tecnologias e protocolos incluem: SIP para sinalizar o início da sessão, COPS e DIAMETER para gestão das políticas de QdS e das funcionalidades de AAA, e RSVP para reserva de recursos. A introdução recente de uma extensão do RSVP para suportar a agregação de fluxos individuais permite estabelecer compromissos entre custo e desempenho da rede de acesso. Na arquitectura proposta os elementos de rede têm de possuir as seguintes funções: todos os elementos necessitam de suportar o RSVP e incluir um elemento que implemente as políticas de QdS (designado por PEP); o BAS contém um servidor *proxy* de SIP e tem acesso a um elemento que toma decisões de políticas de QdS (designado por PDP), o qual contacta os servidores de AAA dos ISPs e um repositório de políticas; todos os elementos têm um escalonador para diferenciar entre os vários tipos de fluxos de tráfego. Para

verificar o funcionamento destas tecnologias, foi ilustrado o fluxo de mensagens na rede de acesso do estabelecimento de uma sessão multimédia.

Neste capítulo apresentou-se uma solução completa para redes de acesso IPv4. Prevê-se que a introdução do IPv6 nas redes de acesso facilite alguns dos mecanismos descritos anteriormente. O estudo de uma solução completa para redes de acesso IPv6 é um tópico em aberto para investigação futura.

CAPÍTULO 7

MECANISMOS DE *PROBING*

Por forma a combinar a escalabilidade da arquitectura DiffServ (Diferenciação de Serviços) e as garantias superiores de QdS (Qualidade de Serviço) da arquitectura IntServ (Integração de Serviços), uma nova abordagem tem vindo a ser proposta em diversos trabalhos de investigação [Bianchi00, Cetinkaya00, Elek00, Gibbens99, Kelly00, Kelly01]. Nesta abordagem, o controle de admissão é efectuado pelos extremos da comunicação, sejam eles o próprio equipamento terminal do utilizador ou os *routers* de entrada e de saída das redes de cliente. Estas propostas têm como denominador comum o recurso à investigação do estado de congestão da rede através da inserção de fluxos de teste. Estes mecanismos de controle de admissão são designados de *probing.*

Estes mecanismos podem ser descritos através de um exemplo, usando a rede ilustrada na Figura 7-1. Para estabelecer um fluxo entre os utilizadores *H* e *H'*, o utilizador *H* transmite primeiro uma sequência de pacotes para a rede, denominados de *probes*, durante um certo intervalo de tempo, designado por tempo de *probing.* A sequência de pacotes enviados denomina-se de fluxo de *probing*. Após este intervalo de tempo, o receptor envia para o emissor uma mensagem (relatório) contendo informação estatística dos pacotes recebidos. Esta informação permite ao emissor estimar o nível de QdS que o fluxo sofreria caso fosse admitido na rede. Se o nível da QdS estimada se encontrar acima

de um limiar pré-definido, o fluxo é admitido. Caso contrário, o fluxo é rejeitado. Note-se que o objectivo do fluxo de *probing* é determinar se um fluxo de dados pode ser aceite. Para que este objectivo seja atingido, as características de tráfego dos fluxos de *probing* necessitam de ser semelhantes às dos fluxos de dados que pedem admissão. Uma implementação com base nestes mecanismos é escalável porque todas as funcionalidades de monitorização de QdS são implementadas nos extremos, removendo a necessidade de qualquer sinalização e armazenamento de estado por fluxo, e evitando aumentar a complexidade do núcleo da rede.

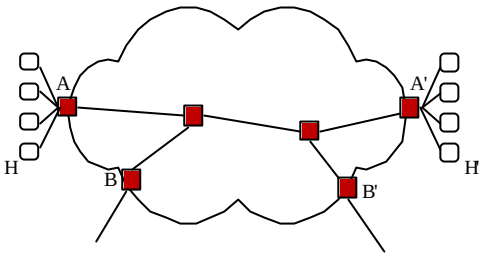


Figura 7-1 : Ilustração dos mecanismos de *probing*.

Alguns dos parâmetros de QdS que podem ser considerados nos mecanismos de *probing* são o rácio de perdas de pacotes, o atraso sofrido pelos pacotes, a variação no atraso (*jitter*), e o rácio de marcação de pacotes. Por exemplo, se o factor de decisão é apenas o rácio de perdas de pacotes, o relatório que o receptor envia ao emissor contém informação sobre o número de pacotes recebidos. Com base nesta informação e na informação que ele próprio tem acerca do número de pacotes enviados, o emissor determina o rácio de perdas. Se por outro lado, se pretende admitir um fluxo, dependendo essa admissão do atraso dos pacotes, uma forma possível de implementação consiste em o emissor inserir em cada pacote o instante de envio. Esta informação permite que o receptor (com o relógio devidamente sincronizado com o emissor) determine o atraso sofrido por cada pacote. Após o tempo de *probing*, o receptor pode então enviar informação relacionada com o atraso sofrido pelos pacotes ou com a variação do atraso. O rácio de marcação de pacotes indica o nível de ocupação das filas de espera dos elementos entre o emissor e o receptor. Em cada elemento, um pacote de *probing* é marcado se, ao ser inserido na fila de espera, esta apresenta uma ocupação superior a uma determinada percentagem. O receptor envia ao emissor o relatório indicando o número de pacotes

marcados recebidos, e este conjuntamente com a informação de pacotes transmitidos, determina o rácio de marcação de pacotes. Na secção 7.5.2.2 é comparado o desempenho dos mecanismos de *probing* quando são considerados o rácio de perdas e o rácio de marcação de pacotes como parâmetros de QdS relevantes para admissão de um determinado fluxo. Exceptuando a secção 7.5.2.2, ao longo deste capítulo e do capítulo 8, a discussão do desempenho dos mecanismos de *probing* e de ?-*probing* será efectuada apenas considerando as perdas de pacotes entre o emissor e receptor.

Uma questão que se pode colocar é se o tráfego de *probing* deve estar no mesmo nível de prioridade estrita dos fluxos de dados ou num nível inferior. Ambas as possibilidades apresentam vantagens e desvantagens. Se o tráfego de *probing* tiver o mesmo nível de prioridade do tráfego de dados, a estimativa obtida da QdS fornecida pela rede é mais precisa, pois o tráfego de *probing* será sujeito às mesmas condicionantes que o tráfego de dados. No entanto, a inserção do tráfego de *probing* na rede pode alterar e degradar a QdS que está a ser fornecida ao tráfego de dados. Assim, utilizando este método, o tráfego de *probing* deve ser limitado em relação ao tráfego de dados para atenuar a sobrecarga que este introduz. Se o tráfego de *probing* for inserido num nível de prioridade inferior ao nível do tráfego de dados, o problema da sobrecarga introduzida pelo tráfego de *probing* é resolvido, mas o nível de QdS estimada será inferior ao nível de QdS real do tráfego de dados. O trabalho desenvolvido neste capítulo considera que o tráfego de *probing* e o tráfego de dados coexistem no mesmo nível de prioridade.

O facto dos fluxos de *probing* e de dados usarem o mesmo nível de prioridade estrita pode conduzir o sistema a uma situação de *thrashing*, como já foi discutido na secção 3.4.4.4. Nos estudos de simulação efectuados neste capítulo considera-se a utilização de um mecanismo que visa atenuar este problema. Este mecanismo interrompe o fluxo de *probing* quando detecta que o rácio de pacotes perdidos excedeu o limiar de perdas. Para isso são inseridos números de sequência nos pacotes de *probing* e o receptor quando verifica, com base nos pacotes de *probing* recebidos, que já não é possível atingir o limiar de perdas pré-definido solicita ao emissor que termine o fluxo de *probing*. Como exemplo, considere-se que o tempo de *probing* é de 2 segundos, o limiar de perdas é de 0.5% e que neste tempo são enviados 1000 *probes*. Se ao fim de 0.5 segundos o número de pacotes perdidos é de 51 (atingiu e ultrapassou os 0.5%), o receptor verifica que o rácio de perdas será superior ao limiar, e envia informação ao emissor para terminar de imediato o

processo de *probing* e rejeitar o pedido de acesso. Note-se que, em redes que não garantam que os pacotes são recebidos pela ordem em que foram enviados, este mecanismo é conservativo. No entanto, nas experiências realizadas neste capítulo garante-se que as chegadas são na mesma ordem dos envios.

Um problema adicional dos mecanismos de *probing* é o roubo de recursos em sistemas com múltiplas classes de serviço, que será descrito na secção 7.1. O objectivo primordial deste capítulo é estudar este problema e propor um novo mecanismo de *probing*, denominado de ?-*probing*, que permite minorar os seus efeitos. Os estudos que serão efectuados neste capítulo mostram que, numa rede com diferenciação de QdS, o mecanismo de ?-*probing* permite atingir uma elevada utilização e atenuar o problema do roubo de recursos.

Este capítulo está organizado da seguinte forma. Na secção 7.1 é discutido o problema do roubo de recursos. Na secção 7.2 é proposto um novo mecanismo de *probing*, o ?-*probing* para minorar os seus efeitos. Na secção 7.3 é apresentado um modelo de fluídos que permite estudar a influência do rácio de perdas e do número de pacotes de *probing* e de ?-*probing* na estimativa efectuada pelos fluxos de *probing* e de ?-*probing*. De seguida, na secção 7.4, é apresentado um modelo baseado em cadeias de *Markov*, que possibilita determinar analiticamente a utilização e a probabilidade de roubo de recursos, e analisar comparativamente os mecanismos de *probing* (e de ?-*probing*) aplicados a redes com diferentes algoritmos de escalonamento. Na secção 7.5 são apresentados estudos numéricos e de simulação para averiguar o desempenho dos mecanismos de *probing* e de ?-*probing* num conjunto extenso de cenários diferentes. Na secção 7.5.1 são efectuados estudos numéricos e de simulação ao nível do fluxo, e na secção 7.5.2 são realizados estudos de simulação ao nível do pacote através da utilização do simulador de redes *ns*-2 (*network simulator*) [NS-2End]. Para finalizar, na secção 7.6, são apresentadas algumas conclusões.

7.1 Mecanismos de *probing* e o roubo de recursos

No processo de *probing*, o emissor aceita ou rejeita um novo fluxo com base apenas nas estimativas obtidas pelo fluxo de *probing*. À primeira vista, se o nível de QdS estimado pelo fluxo de *probing* for superior ao limiar, o fluxo poderá ser aceite sem degradar a QdS

da rede. No entanto, existem situações específicas, dependentes do algoritmo de escalonamento utilizado nos elementos da rede, em que o fluxo infere que pode ser admitido, mas a sua admissão introduz uma degradação na QdS de fluxos que tenham sido previamente admitidos. Quando tal situação acontece, diz-se que há roubo de recursos, pois o fluxo ao ser admitido vai "roubar", sem se aperceber, recursos previamente atribuídos a outros fluxos.

Considerem-se em primeiro lugar algoritmos de escalonamento que, quando utilizados em redes que processam o controle de admissão através de mecanismos de *probing*, apresentam problemas de roubo de recursos. Este problema pode ser descrito através de exemplos simples. Considere-se o cenário de dois fluxos a partilhar uma ligação com capacidade C. Os fluxos atravessam um *router* com algoritmo de escalonamento FQ - *Fair Queuing* (o mesmo exemplo é aplicável ao algoritmo RR - *Round Robin*). Supõe-se que o primeiro fluxo requer uma largura de banda de $0.75C$ e é admitido no sistema inicialmente sem carga. Posteriormente, um novo fluxo que requer uma largura de banda de $0.5C$ pede admissão no sistema. O correspondente fluxo de *probing* terá um tratamento que lhe garantirá 50% de recursos (FQ com dois fluxos a competirem) e o novo fluxo será admitido. No entanto, embora $0.5C$ seja a taxa atribuída a cada fluxo pelo algoritmo FQ, o objectivo de preservar a QdS dos fluxos já admitidos não é atingido. Assim, o primeiro fluxo vai ter uma redução na taxa de serviço em relação à taxa com que foi inicialmente admitido, devido à introdução do segundo fluxo. Este exemplo simples ilustra um aspecto importante: ao ser efectuado o *probing* de um fluxo, apenas é estimado o nível de QdS desse fluxo, e não é possível ter informação do impacto da aceitação desse fluxo nos restantes fluxos previamente admitidos na rede.

Considere-se um escalonador baseado em classes, em que no interior de cada classe é implementado o algoritmo de escalonamento FIFO (*First In First Out*) e entre as várias classes é implementado o algoritmo de escalonamento WFQ - *Weighted* FQ (o mesmo exemplo é aplicável ao WRR - *Weighted* RR e DRR - *Deficit* RR). Considera-se que a cada classe de serviço é associado um peso de 50% (Figura 7-2). Assume-se que o tráfego oferecido a cada classe é inicialmente de $0.8C$ na classe 1 e de $0.2C$ na classe 2. Devido à natureza *work-conserving* do escalonador, a classe 2 pode "emprestar" os seus recursos à classe 1, permitindo aceitar todo o tráfego da classe 1 sem que ocorram perdas. Se neste momento vários fluxos tentam admissão na classe 2, representando no total $0.3C$, eles são

admitidos e servidos sem perdas no escalonador, pois o tráfego total oferecido à classe 2 é de 0.5*C* (igual ao peso da classe). Esta admissão de fluxos na classe 2 reduz a taxa de serviço da classe 1 de 0.8*C* para 0.5*C*, o que tem como consequência a perda de 30% dos pacotes da classe 1, pertencentes a fluxos previamente admitidos no sistema. Também neste caso, o objectivo de preservar a QdS dos fluxos já admitidos não é atingido.

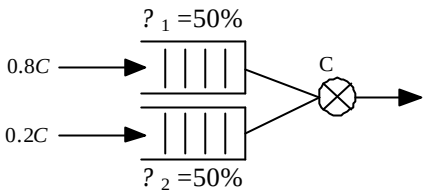


Figura 7-2 : Ilustração do roubo de recursos num escalonador WFQ baseado em classes.

Finalmente, considere-se o exemplo anterior com 2 classes e o algoritmo de escalonamento de prioridade estrita, ou seja, a classe 1 tem uma prioridade superior à da classe 2 (o mesmo exemplo é aplicável quando o escalonamento é efectuado com base em fluxos). Sempre que um fluxo tenta admissão na classe 1, os pacotes do fluxo de *probing* serão servidos antes dos pacotes dos fluxos previamente admitidos na classe 2. Se este fluxo é aceite pode roubar recursos previamente atribuídos aos fluxos da classe 2.

Considerem-se agora algoritmos de escalonamento que não apresentam roubo de recursos, por exemplo, o FIFO, o WFQ, o WRR e o DRR baseados em fluxos e suas variantes. No FIFO os pacotes são servidos por ordem de chegada e não existe diferenciação de QdS. Um fluxo de *probing* é aceite se a capacidade de cada ligação for suficiente para admitir o novo fluxo. Deste modo, este algoritmo não apresenta problemas de roubo de recursos, mas não pode ser utilizado em redes com suporte de diferenciação de serviços. O WFQ baseado em fluxos (assim como o WRR e o DRR) garante uma largura de banda mínima a cada fluxo. Assim, se os pacotes de um fluxo forem enviados para a rede a uma taxa igual ou superior à sua largura de banda mínima, a sua taxa mínima de serviço corresponderá à largura de banda mínima garantida para o fluxo, enquanto que no algoritmo FQ a taxa de serviço de cada fluxo dependerá do número de fluxos activos na rede. Estes algoritmos não apresentam roubo de recursos e suportam diferenciação de QdS. No entanto, esta diferenciação é efectuada fluxo-a-fluxo.

Os exemplos apresentados anteriormente permitem concluir que os algoritmos de escalonamento que apresentam roubo de recursos, quando utilizados em redes que

efectuam controle de admissão através de mecanismos de *probing*, podem ser tipificados da seguinte forma: algoritmos *work-conserving* com garantias relativas por fluxo (FQ, RR e prioridade estrita), e com garantias absolutas por classes (WFQ, WRR e DRR baseados em classes e suas variantes) ou relativas por classes (prioridade estrita). As garantias absolutas estão associadas ao conceito de largura de banda mínima garantida a cada classe ou fluxo, tal como definida na secção 3.3.4.1. As garantias relativas não asseguram nenhuma largura de banda mínima, mas sim um tratamento diferenciado entre fluxos ou entre classes.

Note-se que o roubo de recursos entre classes é possível devido à existência de partilha de recursos entre elas (escalonador *work-conserving*). Se a largura de banda de cada classe for limitada (escalonador *non-work-conserving*), o problema do roubo de recursos não existe, mas a utilização dos recursos da rede depende da proporcionalidade entre o tráfego oferecido e a largura de banda reservada para cada classe. Na secção seguinte é proposto um novo mecanismo de *probing*, o ?-*probing*, que permite a partilha de recursos entre classes sem roubo de recursos, e com diferenciação de QdS.

7.2 Mecanismo de ?-*probing*

Em [Sargento01a] foi proposto um mecanismo de *probing*, o ?-*probing*, que permite resolver o problema do roubo de recursos que afecta sistemas com múltiplas classes de serviço. O método usado consiste em observar o impacto da admissão do novo fluxo nas restantes classes através da inserção de fluxos de *probing* de pequena largura de banda nessas classes. Estes fluxos que são enviados nas outras classes são denominados de fluxos de ?-*probing*. Os pacotes que são enviados nos fluxos de ?-*probing* denominam-se de ?-*probes*. O fluxo é admitido se o nível de QdS estimado pelo fluxo de *probing* e pelos fluxos de ?-*probing* for superior ao limiar admissível. O objectivo dos fluxos de ?-*probing* é determinar se, quando é inserido um fluxo de *probing* numa classe, a QdS nas outras classes se mantém acima de um determinado limiar. Deste modo, a largura de banda dos fluxos de ?-*probing* não necessita de ser tão elevada como a dos fluxos de *probing*. No exemplo do escalonador WFQ baseado em classes apresentado na secção anterior (ao longo deste capítulo, este algoritmo será denominado de CBQ - *Class-Based Queuing*), o problema do roubo de recursos é resolvido porque, em simultâneo com a inserção do fluxo

de *probing* na segunda classe, é também inserido um fluxo de ?-*probing* na primeira. Este fluxo de ?-*probing* vai detectar as perdas da primeira classe motivadas pelo fluxo de *probing* na segunda classe. O nível de QdS estimado pelo fluxo de *probing* é superior ao limiar na segunda classe, mas o nível de QdS estimado pelo fluxo de ?-*probing* é inferior ao limiar na primeira classe, e por isso, o fluxo não será admitido.

Para efectuar uma primeira análise do mecanismo de ?-*probing*, vai ser utilizado um modelo simplificado de fluídos e sem fila de espera, como o utilizado em [Breslau00b]. Assume-se também que o processo de *probing* é perfeito no sentido em que os fluxos de *probing* conseguem inferir a QdS da rede de uma forma exacta.

Considere-se um *router* que implementa um algoritmo de escalonamento WFQ (Figura 7-3) com *K* classes. A classe *k* tem um peso associado de $?_k$ e requisito de QdS de rácio máximo de perdas $?_k$, designado por limiar de perdas. De acordo com a definição de WFQ, a largura de banda utilizada por cada classe *k* é dada por

$$c'_k ? \frac{?_k}{? \; _{i? \not\in ?} ?_i} C \qquad (7\text{-}1)$$

em que ?' é o conjunto das filas de espera não vazias. A largura de banda do tráfego oferecido à classe *k* é designada por r_k. Se $r_1 ? ... ? r_K ? C$, então vão existir perdas no sistema, sendo o rácio de perdas da classe *k* dado por

$$?_k ? ?r_k ? c'_k ?? / r_k \qquad (7\text{-}2)$$

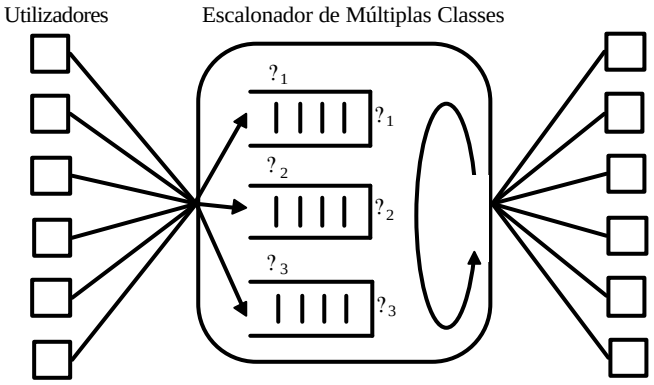


Figura 7-3 : Rede com escalonador de múltiplas classes.

Quando um fluxo com requisito de largura de banda b_k pede admissão e envia um fluxo de *probing* apenas na sua classe k, o novo fluxo é admitido se o rácio de perdas medido é inferior ao limiar da classe, isto é, se $?_k ? ?_k$. O fluxo será admitido desde que

$$r_k ? b_k ? C\frac{?_k}{? \sum_{i? \neq ?} ?_i} \frac{1}{1 ? ?_k ?} \tag{7-3}$$

Esta condição indica que, independentemente do nível de congestão das outras classes, o fluxo é admitido numa classe k desde que o rácio de perdas nessa classe seja inferior ao limiar. Deste modo, pode ocorrer roubo de recursos entre classes porque o fluxo de *probing* não consegue "observar" se os requisitos de QdS estão a ser satisfeitos nas outras classes. Este mecanismo será denominado de *probing* simples.

No mecanismo de ?-*probing* o novo fluxo só é admitido se a condição $?_k ? ?_k$ é satisfeita para todos os k = 1, …, K. Considerando que é transmitida uma largura de banda $b_{?,j}$ em cada fluxo de ?-*probing*, $j? k$, em simultâneo com a transmissão do fluxo de *probing* na classe k, o fluxo é admitido se forem satisfeitas, simultaneamente, as condições 7-3 e 7-4:

$$r_j ? b_{?,j} ? C\frac{?_j}{? \sum_{i? \neq ?} ?_i} \frac{1}{1 ? ?_j ?}, j ? k \tag{7-4}$$

O rácio de perdas admissível em cada fluxo de ?-*probing* pode ser o limiar de rácio de perdas da classe onde é introduzido, ou pode ser um limiar único, idêntico para todas as classes e aprovado globalmente.

O mecanismo de ?-*probing* é aplicável tanto a algoritmos de escalonamento WFQ, WRR e DRR baseados em classes (e as suas variantes), como a algoritmos de escalonamento de prioridade estrita. Neste último, só é necessário injectar fluxos de ?-*probing* nos níveis de prioridade inferior ao nível de prioridade da classe à qual pertence o fluxo que está a pedir admissão, pois apenas estes podem ver os seus recursos roubados. Deste modo, a quantidade de tráfego adicional que é introduzido no sistema é menor num algoritmo de escalonamento de prioridade estrita em relação a um algoritmo CBQ.

7.3 Modelo teórico baseado numa distribuição binomial

Nesta secção é efectuado um estudo teórico para determinar os factores que influenciam a estimativa do rácio de perdas por parte dos fluxos de *probing*.

O objectivo dos fluxos de *probing* é estimar as perdas do sistema, e a partir dessa estimativa, determinar se um novo fluxo pode ou não ser aceite. Considere-se um sistema ideal em que os fluxos de *probing* estimam as perdas sem perturbar o sistema (isto é, apenas observam as perdas) e em que as perdas dos pacotes são independentes. Neste caso, a probabilidade de *i* pacotes perdidos em *n* enviados é descrita por uma distribuição binomial dada por

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i} \tag{7-5}$$

onde p é a probabilidade de perda de pacotes.

O rácio de perdas de pacotes estimado por um fluxo de *probing* com *n* pacotes enviados é de *i*/*n*. A probabilidade de o rácio estimado ser inferior a um determinado limiar de perdas p, ou seja, a probabilidade de aceitação de um fluxo no sistema, é dada por

$$P_a = \sum_{i/n \le p} P(X = i) \tag{7-6}$$

Neste sistema ideal, se $p_{real} \le p$, as perdas do sistema são inferiores ao limiar de perdas e, por isso, todos os fluxos devem ser aceites; caso contrário, todos os fluxos devem ser rejeitados.

Na Figura 7-4 são apresentadas as curvas da probabilidade de aceitação em função de p_{real}/p, para 2 valores diferentes de *n* (64 e 128), e para 3 valores diferentes de p (0.5%, 5% e 15%). Note-se que o comportamento ideal do sistema corresponde a uma probabilidade de aceitação de 1 para $p_{real}/p < 1$, e de 0 para $p_{real}/p > 1$. Os resultados mostram que o comportamento tende para o ideal quando aumenta o número de *probes* e quando aumenta o limiar de perdas. Note-se que o menor rácio de perdas não nulo que uma sequência de *n* pacotes consegue estimar é 1/*n*. Isto ilustra que um pequeno número de pacotes enviados não consegue detectar perdas muito pequenas.

Considere-se agora o caso de um limiar de perdas nulo, $p = 0$. Neste caso, a probabilidade de aceitação de um fluxo vem dada por $(1-p)^n$, ou seja, mesmo com um limiar de perdas nulo ($p = 0$) e perdas reais não nulas ($p_{real} > 0$), existe alguma probabilidade

de um fluxo ser aceite. Uma conclusão importante a extrair deste estudo é que é difícil garantir um rácio de perdas pequeno ou nulo, através de mecanismos de *probing.*

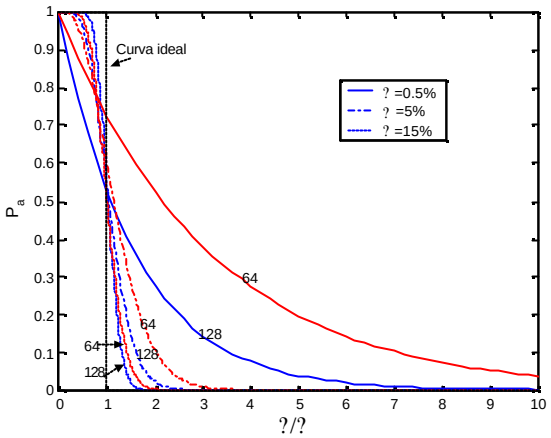


Figura 7-4 : Probabilidade de aceitação de um fluxo. *?*

7.4 Modelo teórico baseado em cadeias de *Markov*

Nesta secção é desenvolvido um modelo analítico baseado em cadeias de *Markov* em tempo contínuo para estudar o problema do roubo de recursos em redes com múltiplas classes. Os algoritmos de escalonamento que vão ser considerados são o FIFO, o FQ baseado em fluxos, os limitadores de taxas e o CBQ.

No modelo, cada estado identifica o número actual de fluxos admitidos em cada classe, tal que com K classes, a cadeia de *Markov* tem K dimensões. O sistema considerado tem apenas uma ligação com C unidades de capacidade e cada fluxo da classe k ocupa b_k unidades. Note-se que no caso geral, a largura de banda dos fluxos pertencentes à mesma classe pode ser diferente; no entanto, esta suposição é necessária para a formulação matemática através do modelo de *Markov*. Assume-se que as chegadas de fluxos novos ao sistema são modeladas por um processo de *Poisson* com um intervalo médio entre chegadas de $1/?_k$, e o tempo de vida de cada fluxo é exponencialmente distribuído com média $1/?_k$. Com o objectivo de simplificar o modelo, considera-se que os fluxos de *probing* estimam a QdS de forma exacta e instantânea, tendo portanto uma duração nula. Todas as classes têm um limiar de perdas nulo.

O número de fluxos da classe k no sistema é designado por n_k e a quantidade de recursos ocupados no sistema por todos os fluxos pertencentes a todas as classes é dada por (**b.n**), em que $\mathbf{b} = (b_1,...,b_K)$, $\mathbf{n} = (n_1,...,n_K)$, e $(\mathbf{b}.\mathbf{n}) = \sum_{k=1}^{K} b_k n_k$.

Os estados admissíveis na cadeia de *Markov* são dependentes do algoritmo de escalonamento usado. O espaço de estados que corresponde ao algoritmo de escalonamento genérico *AE* será designado por Ω_{AE}. A utilização média do sistema é dada por

$$U = \frac{1}{C} \sum_{\mathbf{n} \in \Omega_{AE}} \min((\mathbf{b}.\mathbf{n}), C) \pi_{\mathbf{n}} \tag{7-7}$$

em que $\pi_{\mathbf{n}}$ é a probabilidade limite do estado $\mathbf{n}$ que pode ser determinada a partir das equações de balanço [Ross97].

Como o espaço de estados do FIFO contém todos os estados possíveis que não apresentam roubo de recursos, este vai ser considerado como cenário de referência. Todos os estados que apresentam roubo de recursos (noutros escalonadores) são os que não são admissíveis no FIFO.

A probabilidade de roubo de recursos é calculada como a percentagem de largura de banda assegurada para determinados fluxos que é roubada por outros fluxos, ou seja,

$$p_{st}^{AE} = 1 - \frac{\sum_{\mathbf{n} \in \Omega_{FIFO}} (\mathbf{b}.\mathbf{n}) \pi_{\mathbf{n}}}{\sum_{\mathbf{n} \in \Omega_{AE}} \min((\mathbf{b}.\mathbf{n}), C) \pi_{\mathbf{n}}} \tag{7-8}$$

De seguida, vão ser apresentados os modelos de *Markov*, considerando os diferentes algoritmos de escalonamento e mecanismos de *probing*. O cálculo de $\pi_{\mathbf{n}}$ requer a determinação do espaço de estados e também do espaço de transições, o que será efectuado nas secções seguintes. Na discussão que se segue, é considerado um exemplo simples com duas classes de serviço, em que a largura de banda total da ligação é C = 3 unidades e a largura de banda dos fluxos é $b_1 = 1$ e $b_2 = 2$.

7.4.1 FIFO

Nesta secção considera-se um sistema com algoritmo de escalonamento FIFO. No exemplo com duas classes descrito anteriormente, um fluxo é admitido se $b_1 n_1 + b_2 n_2 \le C$. A Figura 7-5 ilustra o correspondente diagrama de transição de estados.

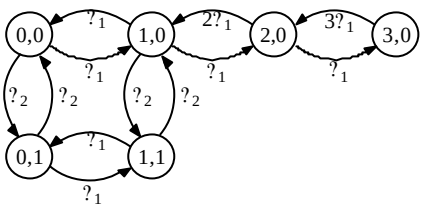


Figura 7-5 : Diagrama de transição de estados de FIFO.

No caso geral, o espaço de estados é representado por $\mathcal{S}_{FIFO} = \{(n) : n \in \mathbb{N}_0 \wedge n \leq C\}$ em que $\mathbb{N}_0$ é o conjunto dos inteiros não negativos. A probabilidade de roubo de recursos é nula, pois os fluxos de *probing* só são admitidos se existir largura de banda suficiente na ligação.

No caso de CBQ com $\mathbb{N}_0$*-probing*, todas as classes são investigadas para garantir que não existe roubo de recursos entre elas. Considerando os fluxos de *probing* e de $\mathbb{N}_0$*-probing* ideais, os estados admissíveis em CBQ com $\mathbb{N}_0$*-probing* são os mesmos que os admissíveis em FIFO com *probing* simples. A utilização é também a mesma. Num sistema real com duração de fluxos de *probing* e de $\mathbb{N}_0$*-probing* não nulas, a utilização é ligeiramente inferior. A diferença essencial entre FIFO com *probing* simples e CBQ com $\mathbb{N}_0$*-probing* é a diferenciação entre classes de serviço que este último permite.

7.4.2 *Fair queuing*

Na secção 7.1 referiu-se que, num sistema com algoritmo de escalonamento FQ baseado em fluxos, os fluxos com largura de banda superior podem ver os seus recursos roubados por fluxos de largura de banda inferior. A Figura 7-6 ilustra o diagrama de transição de estados de um sistema com FQ. O espaço de estados de FQ inclui todos os estados tais que o número de fluxos multiplicado pela menor largura de banda dos fluxos activos no sistema é menor ou igual à capacidade da ligação. Por exemplo, a transição do estado (1,1) para o estado (2,1) é possível porque o FQ garante 1 unidade de largura de banda a cada fluxo, e esta unidade é suficiente para fluxos pertencentes à classe 1. Por outro lado, a transição do estado (2,0) para o (2,1) não é possível porque os fluxos da classe 2 requerem uma largura de banda de 2 unidades. Deste modo, o roubo de recursos ocorrerá sempre que o sistema estiver a transmitir à taxa máxima e forem oferecidos fluxos com largura de banda inferior a fluxos activos no sistema. No diagrama de transição de estados, os estados em que se verifica roubo de recursos são representados por uma cruz.

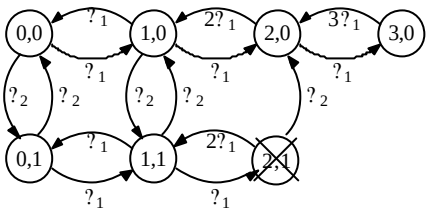


Figura 7-6 : Diagrama de transição de estados de FQ.

Admitindo que o sistema está no estado $\mathbf{n} = (n_1, ..., n_k, ..., n_K)$, um novo fluxo pertencente à classe k pode ser admitido se se verificar a seguinte condição: $b_1 n_1 + b_2 n_2 + ... + b_k (n_k + 1) + ... + b_K n_K \leq C$ com $b_1 \leq b_2 \leq ... \leq b_k \leq ... \leq b_K$. Quando se observa esta condição, duas situações podem acontecer: se o tráfego oferecido satisfaz a condição $b_1 n_1 + b_2 n_2 + ... + b_k (n_k + 1) + ... + b_K n_K \leq C$, todos os fluxos vão ser admitidos sem roubo de recursos; caso contrário ocorre roubo de recursos no sistema. O algoritmo de escalonamento garante uma largura de banda (relativa) a cada fluxo individualmente e, por isso, os fluxos com uma maior largura de banda do que os fluxos que estão a pedir admissão, vão ter os seus recursos roubados.

No caso específico do exemplo apresentado com duas classes de serviço, o espaço de estados, incluindo os estados de roubo de recursos, é dado por

$$\Omega_{FQ} = \{\mathbf{n} = (n_1, n_2) \in \mathbb{N}_0^2 : b_1 n_1 + b_1 n_2 \leq C \} \cup \{ b_2 n_2 \leq C \} \cup \{ n_1 = 0 \} \cup \{ b_2 n_2 \leq C \} \cup \{ n_1 = 0 \} \quad (7\text{-}9)$$

Para melhor compreensão da expressão do espaço de estados, considere-se um exemplo. O estado (2,1) no diagrama de transição de estados (com $n_1 = 2 > 0$) é um estado possível, embora seja um estado em que existe roubo de recursos, porque $b_1 n_1 + b_1 n_2 \leq 3 \leq C$ e $b_2 n_2 \leq 2 \leq C$. Com $n_1 > 0$, o estado com o número máximo de fluxos é o estado (0,1) porque $b_2 n_2 \leq 2 \leq C$. Esta condição já não será verificada no estado (0,2) e, por isso, este não faz parte do diagrama de transição de estados.

Para generalizar o espaço de estados a K classes de fluxos, considere-se k^* como sendo o menor k tal que $n_{k^*} > 0$. Então o espaço de estados é dado por

$$\Omega_{FQ} = \left\{ \mathbf{n} = (n_1, ..., n_K) \in \mathbb{N}_0^K : \begin{array}{l} b_{k^*} n_{k^*} + b_{k^*} n_{k^*+1} + ... + b_{k^*} n_K \leq C \} \cup \\ \{ b_{k^*+1} n_{k^*+1} + ... + b_{k^*+1} n_K \leq C \} \cup ... \cup \{ b_K n_K \leq C \} \end{array} \right. \quad (7\text{-}10)$$

Para determinar o espaço de transições, considera-se que $\mathbf{n}_k^+$ é um estado que pertence a Ω_{FQ} atingido a partir de $\mathbf{n}$ através do aumento em uma unidade do número de fluxos da sessão k, isto é, $\mathbf{n}_k^+ = (n_1, ..., n_k+1, ..., n_K)$; considera-se que o par $(\mathbf{n},\mathbf{n}_k^+)$ representa uma transição ascendente, do estado $\mathbf{n}$ para o estado $\mathbf{n}_k^+$, e que o par $(\mathbf{n}_k^+,\mathbf{n})$ representa uma transição no sentido descendente. No sentido descendente qualquer transição é possível, enquanto que no sentido ascendente uma transição para o estado $\mathbf{n}_k^+$ é possível apenas se $b_1 n_1 + ... + b_k n_k + 1 \le C$. Assim, o espaço de transições é dado por

$$\begin{aligned}\Lambda_{FQ} = \{&(\mathbf{n},\mathbf{n}_k^+) : \mathbf{n},\mathbf{n}_k^+ \in \Omega_{FQ} \wedge b_1 n_1 + ... + b_k n_k + 1 \le C, k = 1,...,K\} \cup \\ &\{(\mathbf{n}_k^+,\mathbf{n}) : \mathbf{n},\mathbf{n}_k^+ \in \Omega_{FQ}\}\end{aligned} \tag{7-11}$$

7.4.3 Limitadores de taxas

Os limitadores de taxas (*rate limiters*) limitam cada classe de fluxos a usar apenas a largura de banda que lhe é atribuída: os fluxos da classe k podem usar um máximo de C_k unidades de largura de banda. No exemplo corrente, vão ser considerados limitadores de $C_1 = 1$ unidades e $C_2 = 2$ unidades. A Figura 7-7 ilustra o correspondente diagrama de transição de estados. Dado o isolamento completo a que as classes estão sujeitas, o espaço de estados vem simplesmente dado por

$$\Omega_{RL} = \{\mathbf{n} \in \mathbb{N}^K : b_k n_k \le C_k, k = 1,2,...,K\} \tag{7-12}$$

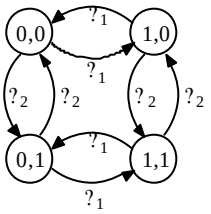


Figura 7-7 : Diagrama de transição de estados de um limitador de taxas.

Com a eliminação de vários estados de elevada utilização, esta é reduzida em comparação com a utilização dos sistemas *work-conserving* apresentados anteriormente. No entanto, esta redução depende da relação entre o tráfego oferecido e os limites definidos para cada classe. Se os limitadores estiverem na proporção do tráfego oferecido, isto é, se

$$C_k \; ? \; C \frac{b_k ?_k / ?_k}{\sum_{i?1}^{K} b_i ?_i / ?_i}, k=1, ..., K \qquad (7\text{-}13)$$

onde $b_k ?_k / ?_k$ é o tráfego oferecido à sessão k e K é o número de classes, então a redução na utilização é mínima.

7.4.4 CBQ

No caso de CBQ com *probing* simples, pode ocorrer roubo de recursos entre classes, quando os fluxos tentam admissão na classe k com um tráfego oferecido menor que $?_k C$, e alguma da largura de banda garantida a esta classe está a ser usada por outras classes. Deste modo, um fluxo da classe k é admitido com largura de banda b_k se, incluindo o fluxo de *probing*, uma das seguintes condições ocorrer: $b_k n_k \; ? \; C$ quando (**b.n**) $? \; C$, ou $b_k n_k \; ? \; C ?_k$ quando (**b.n**) $> C$. O primeiro conjunto de desigualdades define os estados existentes no algoritmo de base FIFO, enquanto que o segundo conjunto define o conjunto de estados em que existe roubo de recursos. No diagrama de transição de estados da Figura 7-8, os pesos associados a cada classe são $?_1 = 2/3$ e $?_2 = 1/3$. Para uma melhor compreensão do diagrama de transições, considere-se como exemplo a transição do estado (2,0) para o estado (2,1). Neste caso, o segundo conjunto de desigualdades é satisfeito porque $b_1 n_1 \; ? \; 2 \; ? \; C ?_1$ e $b_1 n_1 \; ? \; b_2 n_2 \; ? \; 4 \; ? \; C$. Deste modo, a transição para um estado onde existirá roubo de recursos é possível. Do mesmo modo, a transição do estado (1,1) para o (2,1) não é permitida porque no estado (1,1) cada classe utiliza os seus recursos na totalidade. Por isso, não existem recursos disponíveis para serem emprestados à classe 1. Nenhum dos conjuntos de desigualdades definidos anteriormente é satisfeito para esta última transição.

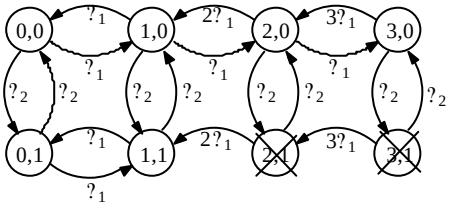


Figura 7-8 : Diagrama de transição de estados de CBQ com *probing* simples.

O espaço de estados neste sistema, $?_{CBQ}$, pode ser decomposto em duas partes. A primeira parte é igual ao espaço de estados do algoritmo FIFO, ou seja, inclui o conjunto

$\{\mathbf{n} \in \mathbb{N}^K : \mathbf{b}.\mathbf{n} \leq C\}$. A segunda parte inclui todos os estados para os quais a largura de banda utilizada em cada classe é igual ou inferior à largura de banda mínima garantida, ou seja, $b_k n_k \leq C/K$. O espaço de estados vem então dado por

$$\Omega_{CBQ} = \{\mathbf{n} \in \mathbb{N}^K : \mathbf{b}.\mathbf{n} \leq C \wedge b_k n_k \leq C/K, k = 1,\ldots,K\} \tag{7-14}$$

Em relação ao espaço de transições, qualquer transição no sentido descendente é possível (assim como em FQ), enquanto que no sentido ascendente uma transição para o estado $\mathbf{n}_k^+$ é possível apenas se $\{\mathbf{b}.\mathbf{n}_k^+ \leq C \wedge b_k n_k + 1 \leq C/K$. O espaço de transições é então dado por

$$\begin{aligned} \Lambda_{CBQ} = \{&(\mathbf{n}, \mathbf{n}_k^+) : \mathbf{n}, \mathbf{n}_k^+ \in \Omega_{CBQ} \wedge \{\mathbf{b}.\mathbf{n}_k^+ \leq C \wedge b_k n_k + 1 \leq C/K, k = 1,\ldots, K\} \cup \\ &\cup \{(\mathbf{n}_k^+, \mathbf{n}) : \mathbf{n}, \mathbf{n}_k^+ \in \Omega_{CBQ}\} \end{aligned} \tag{7-15}$$

Comparando os sistemas que usam FQ e CBQ com *probing* simples no caso do exemplo anterior, o sistema CBQ com *probing* simples tem um maior número de estados de roubo de recursos. Por exemplo, o estado que provoca roubo de recursos (3,1) é possível no sistema CBQ mas não no mesmo sistema FQ. Este último sistema bloqueia o quarto fluxo pois o sistema e o próprio fluxo vão ter perdas consideráveis. Enquanto isso, o sistema CBQ admite o novo fluxo porque a largura de banda de 2 unidades pedida para o fluxo ainda está disponível na classe 1, mesmo que esta admissão roube recursos previamente garantidos a fluxos na classe 1.

7.5 Resultados numéricos e de simulação

Nesta secção é apresentado um conjunto de estudos numéricos e de simulação aplicados a diversos cenários com características diferentes, com o objectivo de avaliar o desempenho dos mecanismos de *probing* simples e de ε-*probing* descritos em 7.2, e validar as conclusões derivadas através dos modelos analíticos apresentados nas secções 7.3 e 7.4.

Os estudos de simulação são divididos em duas partes: simulações ao nível do fluxo (na secção 7.5.1) e simulações ao nível do pacote (secção 7.5.2). As primeiras simulações têm como objectivo comparar os mecanismos de *probing* e os diferentes tipos de algoritmos de escalonamento analisados na secção 7.4. As segundas simulações foram efectuadas no simulador *ns*-2, e têm como objectivo avaliar as características necessárias dos fluxos de *probing* para detectar com precisão as perdas nas diferentes classes de

serviço, e comparar os mecanismos de *probing* simples e de ?-*probing* em diferentes cenários.

7.5.1 Estudos numéricos e de simulação ao nível do fluxo

Este primeiro conjunto de experiências tem o objectivo de estudar o fenómeno do roubo de recursos e a utilização da rede num sistema com mecanismos de *probing* e com os algoritmos de escalonamento considerados na secção 7.4. As experiências efectuadas recorrem ao modelo analítico descrito na secção 7.4 e a simulação de eventos discretos ao nível do fluxo.

O cenário base para realizar o primeiro conjunto de experiências é ilustrado na Figura 7-9. Considera-se que a capacidade da ligação é de C = 4.096 Mb/seg. Em algumas experiências, o *router* contém limitadores de taxa que descartam todos os pacotes que excedem uma taxa pré-definida para a classe. São considerados os algoritmos de escalonamento analisados na secção 7.4, em que FIFO é usado como cenário base para fazer estudos comparativos entre os diversos mecanismos. No simulador de fluxos, os novos fluxos chegam ao sistema com tempos entre chegadas independentes de acordo com um processo de *Poisson*, e têm durações exponencialmente distribuídas. Uma vez que a simulação não é ao nível do pacote, as decisões de controle de admissão efectuadas pelo simulador de fluxos baseiam-se apenas em larguras de banda. Os processos de decisão são em tudo idênticos aos apresentados na secção 7.4.

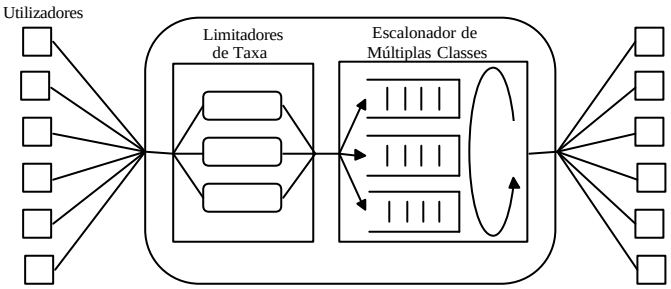


Figura 7-9 : Cenário de simulação.

Este cenário contém três classes de serviço com requisitos de largura de banda b_1 = 256 Kb/seg, b_2 = 512 Kb/seg e b_3 = 1024 Kb/seg, respectivamente.

Na Figura 7-10 (a e b) é comparada a utilização atingida pelos mecanismos de *probing* em diferentes escalonadores, considerando o *probing* ideal (fluxos de *probing*

instantâneos) na Figura 7-10 (a) e fluxos de *probing* diferentes de zero na Figura 7-10 (b). No primeiro caso, os resultados apresentados são obtidos pelas equações derivadas através do modelo de *Markov* da secção 7.4. Os resultados apresentados na segunda figura são obtidos através do simulador de fluxos. Neste último caso, para admitir cada fluxo é enviado um fluxo de *probing* à taxa de admissão desejada e, quando é usado ?-*probing*, são enviados também fluxos de ?-*probing* com $b_? = 64$ Kb/seg. A duração dos fluxos de *probing* é de 2 segundos. Na Figura 7-10 (c) compara-se o roubo de recursos existente em FQ e CBQ com *probing* simples (para *probing* ideal e *probing* de 2 segundos). São consideradas três variantes do sistema ilustrado na Figura 7-9. A curva de FQ (denominada no gráfico de "FQ") representa o cenário em que o escalonador atribui a largura de banda de uma forma justa aos fluxos, isto é, o *N*-ésimo fluxo de *probing* será admitido no sistema se a sua largura de banda necessária for inferior a C/N. Por outro lado, as curvas denominadas de "Limitadores de Taxa 1", "Limitadores de Taxa 2" e "CBQ" representam escalonadores baseados em classes. No primeiro caso, a largura de banda de cada classe é limitada proporcionalmente à largura de banda dos seus fluxos, pois $?_k/?_k$ é igual para todas as classes. Na curva "Limitadores de Taxa 2", os limitadores das classes 2 e 3 limitam a largura de banda destas classes a metade da largura de banda reservada em "Limitadores de Taxa 1", e a largura de banda restante vai ser utilizada pela classe 1. No último caso, as classes não têm limitador e o escalonador efectua CBQ com *probing* simples, tendo cada classe um peso proporcional à largura de banda dos seus fluxos. São também apresentadas as curvas de CBQ com ?-*probing*. Com *probing* ideal, a utilização é a mesma que em FIFO, enquanto que para fluxos de *probing* diferentes de zero a utilização é menor em CBQ.

Pela observação das figuras verifica-se que, comparando os resultados de utilização com *probing* ideal e com *probing* de duração e largura de banda não nulas, verifica-se que a utilização com *probing* ideal é obviamente superior, porque a ligação não está ocupada com fluxos de *probing*, estando totalmente disponível para os fluxos de dados. Como a utilização é maior, a probabilidade de o tráfego admitido no sistema exceder a capacidade da ligação aumenta, e o roubo de recursos em FQ e CBQ é também superior. As diferenças entre a utilização com *probing* ideal e real são da ordem dos 2% a 3%. Estas diferenças correspondem ao *overhead* introduzido pelos fluxos de *probing*. A redução na utilização

não é significativa, o que permite concluir que com os mecanismos de *probing* é possível obter utilizações dos recursos elevadas.

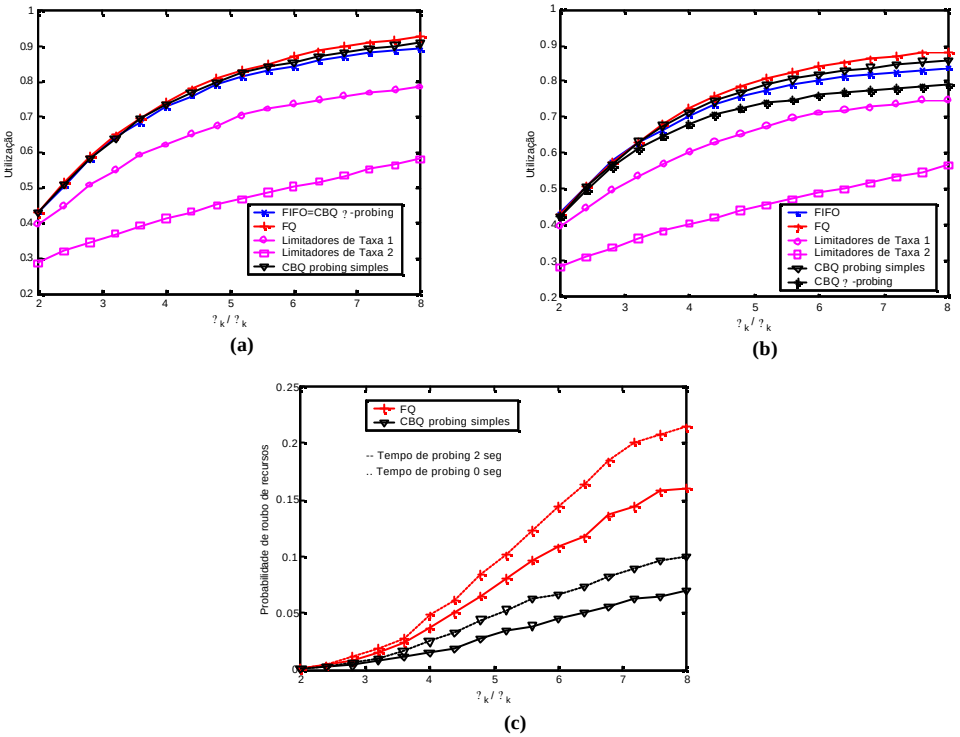


Figura 7-10 : Comparação entre mecanismos de *probing* e entre escalonadores: (a) utilização com *probing* ideal (modelo de *Markov*), (b) utilização com *probing* real e (c) probabilidade de roubo de recursos.

Da Figura 7-10 (a) verifica-se que os algoritmos de FQ, CBQ e FIFO atingem maiores utilizações do que os limitadores de taxa devido à natureza *non-work-conserving* destes. Os limitadores de taxa impedem que os fluxos sejam admitidos numa classe quando esta já excedeu a largura de banda para si reservada, mesmo que exista largura de banda disponível noutras classes. Além disso, quando o tráfego oferecido não é proporcional à largura de banda atribuída a cada classe, a utilização pode decrescer bastante, como é o caso da curva "Limitadores de Taxa 2" em que os limitadores restringem a utilização do sistema a menos de 60%. No entanto, na Figura 7-10 (c), observa-se que a elevada utilização dos algoritmos de FQ e CBQ é atingida com o custo adicional do roubo de recursos, que é de 16% e 7% (para um tráfego oferecido de 8), respectivamente.

Pela observação da Figura 7-10 (c), verifica-se que FQ tem um maior roubo de recursos que CBQ. Com o primeiro mecanismo, os fluxos com largura de banda inferior podem sempre roubar recursos a fluxos com largura de banda superior previamente admitidos no sistema. Este facto resulta numa maior utilização dos recursos, porque mais fluxos vão ser admitidos, mas também numa maior probabilidade de roubo de recursos. Com CBQ, o roubo de recursos só ocorre quando, estando uma ligação saturada, uma classe excede a sua largura de banda garantida e um fluxo ou mais de outra classe requer admissão. Este evento ocorre com menos frequência, porque o tráfego médio oferecido a cada classe é constante ao longo da experiência.

Finalmente, observe-se que comparado com CBQ com *probing* simples, o CBQ com ?-*probing* diminui ligeiramente a utilização do sistema. Existem dois factores que contribuem para esta diminuição. Primeiro, os fluxos de ?-*probing*, embora tenham uma largura de banda muito pequena, introduzem um tráfego oferecido adicional no sistema. Segundo, sendo bloqueados fluxos que iriam resultar em roubo de recursos, existem menos fluxos no sistema em média com ?-*probing*. No entanto, esta pequena redução na utilização tem a vantagem de eliminar o roubo de recursos.

Enquanto que a redução na utilização dos limitadores de taxa pode ser elevada dependendo da proporcionalidade ou não entre o tráfego oferecido e a largura de banda atribuída às classes, a eficiência do mecanismo de ?-*probing* não depende da sintonização correcta dos limitadores. Deste modo, o mecanismo de ?-*probing* num sistema CBQ é capaz de simultaneamente eliminar o roubo de recursos, permitir a existência de múltiplas classes de serviço e diferenciação entre elas, e permitir uma partilha de recursos entre classes.

Estas experiências ilustram também que o único par mecanismo de *probing*/escalonador que possibilita uma elevada utilização, diferenciação de serviços, e um modelo de serviço sem roubo de recursos, é o CBQ com ?-*probing*.

7.5.2 Estudos de simulação ao nível do pacote

Esta secção tem os seguintes objectivos: (i) avaliar o desempenho dos mecanismos de *probing*, (ii) verificar que o roubo de recursos num sistema CBQ com *probing* simples pode ser muito elevado se o tráfego médio oferecido varia ao longo do tempo, (iii) e determinar as características que os fluxos de *probing* e de ?-*probing* têm de possuir para

detectar as violações de QdS nas várias classes de serviço. A investigação de outros assuntos importantes relacionados com o desempenho dos mecanismos de *probing*, como a sua interacção com fluxos TCP (*Transmission Control Protocol*) e a sua aplicabilidade em redes de grandes dimensões, podem ser encontrados em [Breslau00b].

Todos os estudos de simulação ao nível de pacote recorrem ao simulador de redes *ns*-2 [NS-2End]. O *ns* é um simulador de eventos discretos direccionado à investigação na área de redes de telecomunicações. O *ns*, comparativamente a alguns simuladores de redes existentes, tem a vantagem de fornecer um suporte substancial para implementação dos protocolos TCP, comunicação *multicast*, e protocolos de encaminhamento em redes com e sem fios. Para realizar estes conjuntos de experiências foi necessário efectuar algumas extensões ao *ns*. Estas extensões incluem a implementação do mecanismo de controle de admissão (de acordo com as perdas observadas numa ou em todas as classes de serviço), o envio dos fluxos de *probing* e de ?-*probing* do emissor para o receptor na classe respectiva, assim como da informação das estatísticas do receptor para o emissor, a atribuição de um número de sequência aos pacotes, e a definição do espaço das filas de espera em *bytes* em vez de pacotes.

O tráfego de cada sessão k é modelado ao nível do fluxo por um processo de chegadas de *Poisson* com uma taxa média de $?_k$ fluxos por segundo, e o tempo de vida de cada fluxo é considerado exponencialmente distribuído com média $1/?_k$ segundos. A relação $?_k = ?_k/?_k$ é designada por intensidade de tráfego. Define-se tráfego oferecido a uma sessão k como sendo $b_k?_k/?_k$. Ao nível do pacote, as fontes de tráfego consideradas são CBR (*Constant Bit Rate*), que transmitem sempre à mesma taxa, e fontes *on-off*, que transmitem enquanto no estado *on* e estão inactivas quando se encontram no estado *off*. Os tempos de permanência nos estados *on* e *off* podem ter uma distribuição exponencial ou de Pareto. Quando uma fonte de tráfego se encontra activa, os pacotes são gerados com intervalos de tempo entre chegadas fixos. As fontes de tráfego (e respectivos parâmetros) que serão utilizadas nas experiências seguintes são apresentadas na Tabela 7-1.

Fonte de tráfego	Taxa de pico (Kb/seg)	Tempo *on* (mseg)	Tempo *off* (mseg)	Taxa média (Kb/seg)	? ?
CBR_16	16	_	_	16	_
CBR_32	32	_	_	32	_
CBR_64	64	_	_	64	_
CBR_128	128	_	_	128	_
CBR_256	256	_	_	256	_
CBR_512	512	_	_	512	_
CBR_640	640	_	_	640	_
CBR_1024	1024	_	_	1024	_
EXP_128_50	128	500	500	64	_
EXP_256_50	256	500	500	128	_
EXP_512_12.5	512	125	875	64	_
EXP_1024_12.5	1024	125	875	128	_
POO_128_50	128	500	500	64	1.2
POO_256_50	256	500	500	128	1.2

Tabela 7-1 : Características das fontes de tráfego.

Na designação das fontes de tráfego, as primeiras três letras referem o tipo de fonte (EXP para exponencial e POO para Pareto), o número seguinte refere a taxa de pico (em Kb/seg), e o último número refere o *burstiness* da fonte, isto é, a percentagem de tempo em que se encontra a transmitir. A função densidade de probabilidade de uma distribuição de Pareto é dada por

$$f(x) = \frac{\alpha \beta^{\alpha}}{x^{\alpha+1}}, \quad x \geq \beta \tag{7-16}$$

onde α é o parâmetro de forma e β o parâmetro de localização.

Quando nada for dito em contrário, a capacidade das ligações é de 10 Mb/seg, o comprimento dos pacotes é de 125 *bytes*, as filas de espera têm espaço suficiente para um máximo de 200 pacotes (25,000 *bytes*), e o atraso de propagação em cada ligação é de 20 mseg.

Os estudos efectuados nesta secção estão divididos em dois conjuntos. No primeiro é considerado um sistema com uma única classe e sem controle de admissão. O objectivo deste estudo é averiguar as características necessárias dos fluxos de *probing* para estimar com uma determinada precisão as perdas reais do sistema. O segundo conjunto de experiências considera sistemas com controle de admissão com uma classe, em que é

aplicado o mecanismo de *probing* simples, e com múltiplas classes, em que são aplicados os mecanismos de *probing* simples e de ?-*probing*.

7.5.2.1 Análise das características dos fluxos de probing

Este conjunto de experiências baseia-se num cenário simples com uma rede constituída apenas por uma ligação. O nó de comutação suporta uma classe de serviço com algoritmo de escalonamento FIFO. O objectivo destas experiências é estudar as características necessárias dos fluxos de *probing* para estimar com uma determinada precisão as perdas reais na rede. Para atingir este objectivo, é injectado no sistema tráfego oferecido constante (o número de fluxos é mantido constante ao longo da experiência). Além disso, apenas é injectado um fluxo de *probing* de cada vez no sistema, sendo o intervalo entre o fim de um fluxo e o início do seguinte fixado em 1 segundo. Pretendeu-se desta forma estudar o processo de *probing* num ambiente em que a interferência do tráfego de *probing* é mínima. O critério de paragem das simulações é o número de fluxos de *probing* já processados (1,000 no caso concreto). Em cada simulação, o tráfego oferecido é ajustado por forma a garantir um rácio de perdas de pacotes pré-definido.

Para realizar este conjunto de experiências foi necessário modificar o simulador *ns*, por forma a impedir que uma decisão de controle de admissão positiva (com base no fluxo de *probing*) provocasse a entrada de um novo fluxo no sistema. O que se fez neste caso foi garantir que estes fluxos teriam uma duração nula.

Nas experiências descritas de seguida é definido um critério de precisão para as estimativas do rácio de perdas obtidas pelos fluxos de *probing* (as perdas estimadas): considera-se que uma estimativa é suficientemente precisa quando se observam duas condições em simultâneo: (i) o comprimento do intervalo de confiança a 95% das estimativas é inferior a 40% do valor médio das perdas reais na rede e; (ii) o intervalo de confiança contém o valor médio das perdas reais na rede. O número de réplicas consideradas na determinação do intervalo de confiança é de 500.

Neste estudo considera-se a variação de diversos parâmetros para determinar qual a influência destes no desempenho dos fluxos de *probing*. São variadas as características dos fluxos de *probing*, dos fluxos de dados e do próprio sistema. Ao nível dos fluxos de *probing* o estudo incide sobre a largura de banda, o comprimento dos pacotes e o tempo de *probing*. Em relação aos fluxos de dados o estudo incide sobre o modelo da fonte de tráfego. O parâmetro de QdS do sistema que se vai ter em consideração é o rácio de perdas.

Como já foi referido, o tráfego oferecido é controlado de modo a ter um determinado rácio de perdas reais na rede. Os cálculos teóricos para determinar, com base no número de fluxos activos no sistema, o rácio de perdas reais podem ser efectuados com base em vários modelos. Uma das possibilidades é assumir um modelo de fluídos. Como exemplo, considerando 158 fluxos CBR_64 activos no sistema, o tráfego oferecido a uma ligação com 10 Mb/seg de largura banda é de 10.112 Mb/seg. De acordo com um modelo de fluídos, o rácio médio de perdas reais na rede seria de 1.12%.

O tempo de aquecimento de cada experiência é de 1,000 segundos.

7.5.2.1.1 Largura de banda dos fluxos de dados

Neste conjunto de experiências é determinado, para fluxos de dados com diferentes larguras de banda, o tempo de *probing* necessário para os fluxos de *probing* (também com diferentes larguras de banda) estimarem o rácio de perdas do sistema com a precisão definida acima. O tempo de *probing* necessário é o tempo mínimo que permite garantir a observação do critério de precisão. Este estudo vai ser efectuado para os fluxos CBR seguintes: CBR_16, CBR_64, CBR_128, CBR_256 e CBR_640. O rácio médio de perdas reais na rede é de 1%.

Na Figura 7-11 (a) estão representadas as curvas de rácio de perdas estimadas pelos fluxos de *probing* (e o respectivo intervalo de confiança a 95%) e de perdas reais na rede, para fluxos de dados CBR_64 e fluxos de *probing* de 64 Kb/seg (64 pacotes/seg). A Figura 7-11 (b) apresenta, para cada tipo de fluxos de dados e diferentes larguras de banda dos fluxos de *probing*, o tempo de *probing* necessário para que os fluxos de *probing* estimem com a precisão definida anteriormente, o rácio de perdas reais na rede. Na Figura 7-11 (c) encontra-se representado, para algumas fontes de tráfego, o correspondente número de *probes*.

Na Figura 7-11 (a) o rácio de perdas reais aumenta ligeiramente com o aumento do tempo de *probing*, devido ao tráfego adicional injectado pelos fluxos de *probing*. Em relação à curva do rácio de perdas estimado pelos fluxos de *probing*, verifica-se que esta segue a curva de perdas reais na rede apenas para tempos de *probing* superiores a 7 segundos. O comprimento do intervalo de confiança é inicialmente muito elevado e decresce com o tempo de *probing*, reflectindo uma melhoria na precisão da estimativa das perdas reais. Estes resultados estão de acordo com as conclusões da secção 7.3. Aplicando o critério de precisão definido anteriormente verifica-se que este é satisfeito para tempos

de *probing* iguais ou superiores a 3 segundos. Sendo assim, o tempo mínimo de *probing* necessário é, neste caso, de 3 segundos.

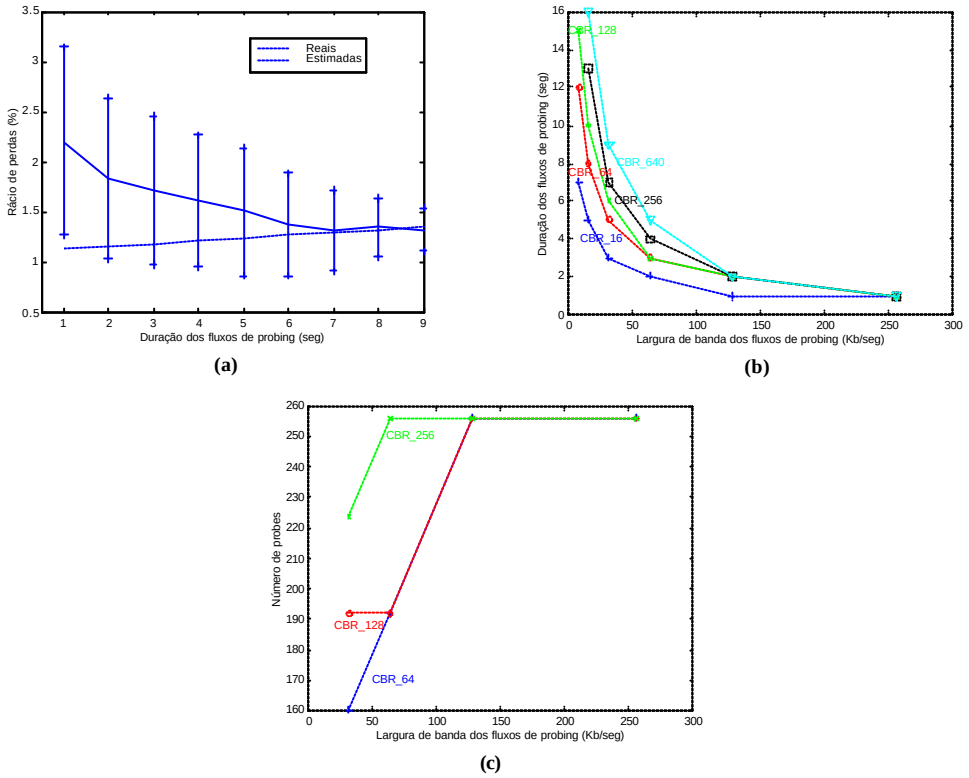


Figura 7-11 : Requisitos dos fluxos de *probing* em função da largura de banda dos fluxos de dados: (a) rácio de perdas *vs* tempo de *probing* ($b_?$= 64 Kb/seg, CBR_64), (b) tempo mínimo de *probing vs* largura de banda dos fluxos de *probing* e (c) número de *probes vs* largura de banda dos fluxos de *probing*.

Na Figura 7-11 (b), observa-se que, aumentando a largura de banda dos fluxos de *probing*, o tempo de *probing* necessário para estimar o rácio de perdas com a precisão definida diminui. Este facto deve-se ao aumento do número de *probes* enviados no mesmo intervalo de tempo. A Figura 7-11 (b) ilustra também o impacto da largura de banda dos fluxos de dados na precisão das medidas obtidas pelos fluxos de *probing*. Verifica-se que, para uma largura de banda dos fluxos de *probing* pequena, fluxos de dados com maior largura de banda necessitam de maior tempo de *probing*. Isto significa que a precisão das medidas efectuadas pelos fluxos de *probing* depende não apenas da largura de banda destes, mas também da largura de banda dos fluxos de dados. Observa-se também que existe uma tendência de aproximação das curvas com o aumento da largura de banda dos

fluxos de *probing*. Por exemplo, considerando fluxos de *probing* de 128 Kb/seg, 2 segundos são suficientes para quase todos os fluxos de dados (menos para o CBR_16), e considerando fluxos de *probing* de 256 Kb/seg, os valores de tempos de *probing* necessários convergem para 1 segundo. Ou seja, quando o número de *probes* enviados é muito elevado, a dependência entre larguras de banda de *probing* e dados referida anteriormente é atenuada.

A Figura 7-11 (c) ilustra o facto de o número de *probes* necessários para estimar o rácio de perdas na rede aumentar com o aumento da largura de banda dos fluxos de *probing*. No entanto, a partir de um determinado valor de largura de banda dos fluxos de *probing*, o número de *probes* mantém-se constante e independente da largura de banda dos dados. Note-se que para fluxos de *probing* com largura de banda mais elevada (superior a 256 Kb/seg) o número de *probes* necessários manter-se-ia, pois os valores de tempo de *probing* necessário diminuiriam linearmente na Figura 7-11 (b). Estes valores não se encontram nos gráficos da Figura 7-11 (b) e Figura 7-11 (c) devido à granularidade do tempo de *probing*, o qual não toma valores abaixo de 1 segundo na simulação realizada.

7.5.2.1.2 Comprimento dos *probes*

Este conjunto de experiências é semelhante ao anterior, mas considera que o parâmetro a variar é o comprimento dos *probes*. A Figura 7-12 (a, b, c e d) apresenta o tempo de *probing* necessário para estimar o rácio de perdas reais na rede em função da largura de banda dos fluxos de *probing*, para diferentes comprimentos dos *probes*, considerando diferentes fluxos de dados (CBR_64 e CBR_128), e diferentes comprimentos dos pacotes de dados (125, 225 e 75 *bytes*). A Figura 7-12 (e) apresenta as curvas de rácio de perdas reais na rede e o rácio estimado pelos fluxos de *probing*, quando os *probes* têm um comprimento de 50 *bytes* e os pacotes de dados são de 125 *bytes*. O rácio médio de perdas reais é de 1%.

Observa-se nas figuras que o tempo de *probing* necessário é superior para comprimentos dos *probes* maiores do que o comprimento dos pacotes de dados. Este resultado deve-se a dois factores. Primeiro, um aumento no comprimento dos *probes* diminui o número de *probes* enviados no mesmo intervalo de tempo, o que contribui para uma diminuição na precisão da estimativa do rácio de perdas. Segundo, numa situação limite em que a rede está congestionada e a fila de espera se encontra cheia, a probabilidade de perda de pacotes com comprimento superior é maior. A título ilustrativo

considere-se o caso em que os pacotes de dados e os *probes* têm comprimento de 125 e 225 *bytes*, respectivamente. Sempre que o espaço disponível na fila de espera se encontra entre 125 e 224 *bytes*, os *probes* são descartados, enquanto que os pacotes de dados são servidos.

Considerando agora a situação oposta em que os *probes* são um pouco mais pequenos que os pacotes de dados, verifica-se que o tempo de *probing* necessário diminui. Nesta situação, a diminuição do comprimento dos *probes* aumenta o número de *probes* enviados no mesmo intervalo de tempo, e consequentemente, aumenta a precisão das estimativas do rácio de perdas reais na rede. No entanto, considerando apenas os casos em que os *probes* têm um comprimento inferior aos pacotes de dados, verifica-se nas situações estudadas, que o tempo de *probing* aumenta com a diminuição do comprimento dos *probes*. Isto pode explicar-se pela tendência dos *probes* mais pequenos detectarem perdas inferiores às dos dados. Na Figura 7-12 (e) observa-se numa situação extrema que, para fluxos de dados CBR_64, largura de banda dos fluxos de *probing* de 64 Kb/seg e 256 Kb/seg, e comprimento dos pacotes de dados e dos *probes* de 125 e 50 *bytes*, respectivamente, o rácio de perdas reais na rede não se encontra dentro dos limites definidos pelo intervalo de confiança do rácio de perdas estimadas. Para fluxos de *probing* de 256 Kb/seg, existe uma maior aproximação entre a curva de perdas reais e a curva de perdas estimadas. Foram também testadas situações com perdas reais da ordem dos 6% que obtiveram os mesmos resultados. Para todos os tempos de *probing* testados (até 15 segundos) não foi possível encontrar uma situação em que as perdas fossem detectadas.

Nas experiências em que se varia o comprimento dos pacotes de dados, verifica-se que os resultados obtidos são semelhantes no que diz respeito à necessidade de um maior tempo de *probing* quando os *probes* têm comprimento maior do que o dos dados e vice-versa. Pela observação da curva da Figura 7-12 (b) correspondente a *probes* de 225 *bytes*, verifica-se que é necessário um tempo de *probing* inferior ao da mesma curva da Figura 7-12 (a). Nestas situações, o número de *probes* enviados no tempo de *probing* é o mesmo. No entanto, no caso da Figura 7-12 (b), o comprimento dos pacotes de dados e dos *probes* é idêntico, pelo que a probabilidade de os *probes* e os pacotes de dados serem descartados é igual, enquanto que no caso da Figura 7-12 (a), a probabilidade de os *probes* serem descartados é maior, e por isso, é necessário um maior tempo de *probing* para obter a

mesma estimativa de perdas. Na Figura 7-12 (c) observa-se a mesma situação no que respeita aos *probes* de 75 *bytes*.

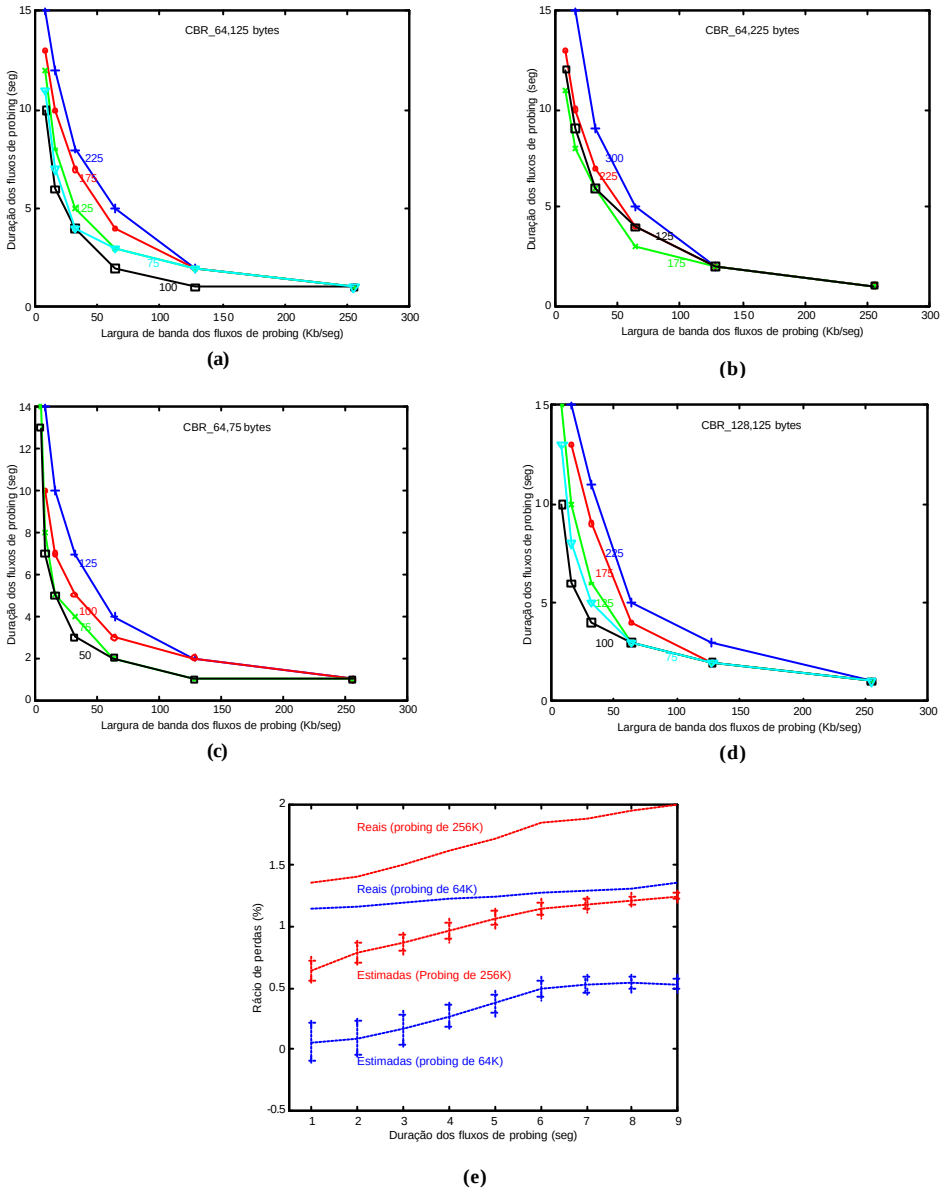


Figura 7-12 : Requisitos dos fluxos de *probing* em função do comprimentos dos *probes*: (a, b, c e d) tempo mínimo de *probing vs* largura de banda dos fluxos de *probing* e (e) rácio de perdas *vs* tempo de *probing* ($b_?$?= 64 Kb/seg e $b_?$?= 256 Kb/seg, CBR_64).

A experiência apresentada na Figura 7-12 (d) é semelhante à da Figura 7-12 (a) em que apenas é alterada a largura de banda dos fluxos de dados para 128 Kb/seg (CBR_128). Verifica-se que existe um ligeiro aumento no tempo de *probing* necessário devido ao aumento da largura de banda dos dados, mas a relação entre as diversas curvas de diferentes comprimentos de *probes* é a mesma que a apresentada na Figura 7-12 (a).

7.5.2.1.3 Rácio de perdas reais

Neste conjunto de experiências vai ser avaliado o impacto de aumentar o rácio de perdas reais na rede, de 1% para 2.5% e 5%. O objectivo é investigar se, com perdas superiores e com a mesma largura de banda dos fluxos de *probing*, o tempo necessário para estimar o rácio de perdas na rede é inferior. O comprimento dos pacotes é de 125 *bytes* em todos os fluxos e em todas as experiências.

A Figura 7-13 (a e b) apresenta os valores de tempo de *probing* necessários para os fluxos de *probing* estimarem com a precisão definida o rácio de perdas reais na rede, usando como fluxos de dados CBR_64 e CBR_256. A Figura 7-13 (c) apresenta as curvas de rácio de perdas reais e rácio de perdas estimado pelos fluxos de *probing* (incluindo o intervalo de confiança a 95%), para os três valores especificados de rácio de perdas, considerando fluxos de dados CBR_64 e fluxos de *probing* de 64 Kb/seg.

Observa-se nas figuras que, para a mesma largura de banda dos fluxos de *probing*, o tempo de *probing* necessário diminui com o rácio de perdas reais na rede. Mais uma vez, estes resultados estão de acordo com as conclusões da secção 7.3. Na Figura 7-13 (c) observa-se com maior detalhe que, com o aumento do rácio de perdas reais na rede, o tempo de *probing* necessário para que as perdas estimadas consigam seguir as perdas reais diminui.

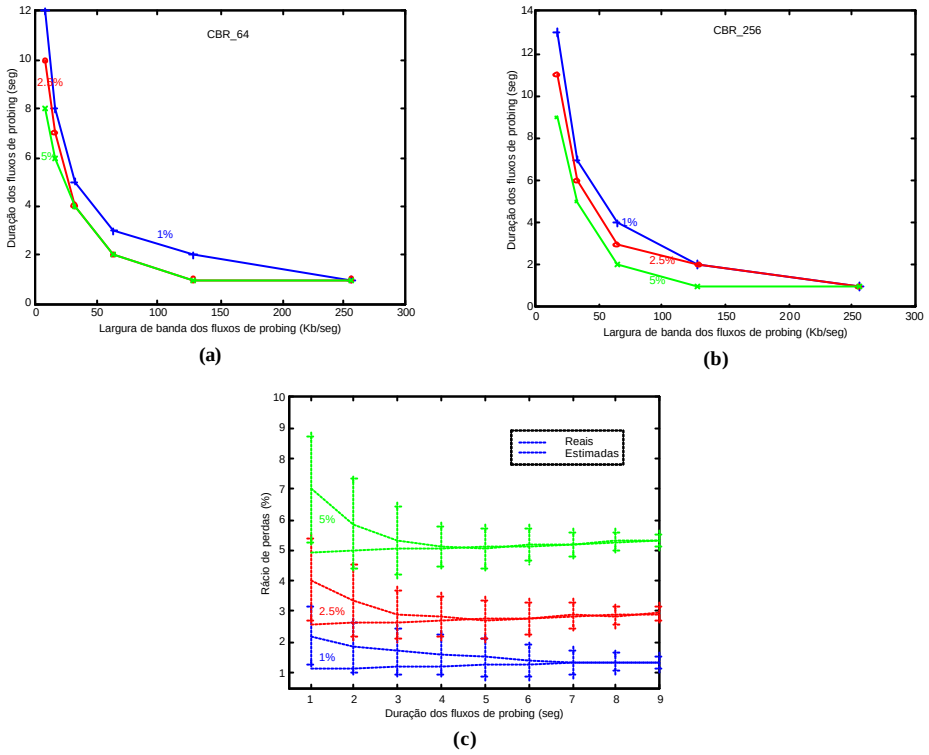


Figura 7-13 : Requisitos dos fluxos de*probing* em função das perdas reais: (a e b) tempo mínimo de *probing vs* largura de banda dos fluxos de *probing* e (c) rácio de perdas *vs* tempo de *probing* ($b_?$?= 64 Kb/seg, CBR_64).

7.5.2.1.4 Tipos de fontes de tráfego

Por último, é estudada a situação em que as fontes de tráfego dos fluxos de dados são *on-off* exponenciais e de Pareto. Na Figura 7-14 (a e b) são apresentadas as estimativas de rácio de perdas para fontes de tráfego EXP_128_50 e POO_128_50. Note-se que estas fontes têm a mesma taxa média que a fonte CBR_64. Também os fluxos de *probing* são iguais aos utilizados na experiência da Figura 7-11 (a). Desta forma estas curvas podem ser comparadas directamente com as da Figura 7-11 (a)**.**

Pela observação das figuras anteriores e por comparação com a Figura 7-11 (a), verifica-se que é necessário um tempo de *probing* mais elevado quando as fontes de tráfego são *on-off* para que os fluxos de *probing* detectem as perdas reais na rede com a precisão definida. Devido à variabilidade do tráfego, a variância das estimativas obtidas pelos fluxos de *probing* é maior quando são utilizadas fontes *on-off*, sendo necessário um

maior tempo de *probing* para obter a precisão definida. A distribuição de Pareto apresenta caudas mais longas, dependentes do parâmetro ?, provocando uma maior variância dos tempos *on* e *off*. Isto explica que o intervalo de confiança com fontes *on-off* de Pareto seja superior ao das fontes *on-off* exponenciais.

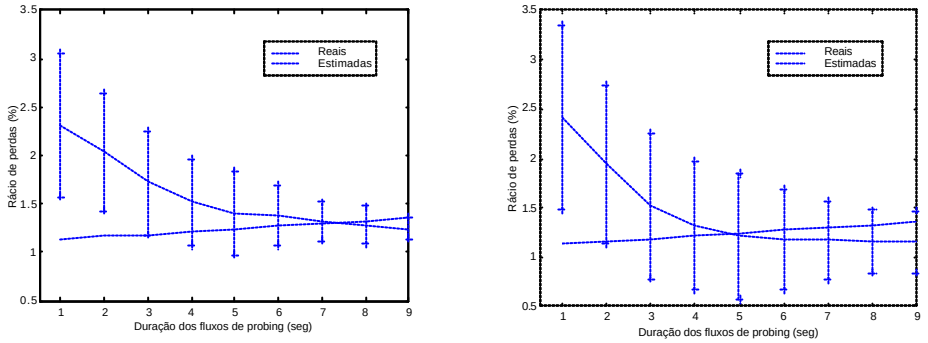


Figura 7-14 : Rácio de perdas em função do tempo de *probing*: (a) fontes *on-off* exponenciais (*b*? ?= 64 Kb/seg, EXP_128_50) e (b) fontes de Pareto (*b*? ?= 64 Kb/seg, POO_128_50).

7.5.2.1.5 Principais conclusões

? A precisão das estimativas efectuadas por um fluxo de *probing* melhora com a largura de banda destes e com o tempo de *probing*; o aumento do nível de congestionamento da rede também melhora a precisão das estimativas.

? Os *probes* devem ter um comprimento semelhante ao dos pacotes de dados; pacotes demasiado pequenos podem conduzir a erros elevados, sendo as perdas estimadas inferiores às perdas reais.

? A precisão das estimativas efectuadas por um fluxo de *probing* diminui com fontes de tráfego *on-off*, e diminui com o aumento da variância entre os tempos *on* e *off*.

7.5.2.2 Sistema com controle de admissão

Neste conjunto de experiências são realizados estudos de simulação com controle de admissão numa rede congestionada, em sistemas com uma ou duas classes de serviço. Nas experiências com uma classe é utilizado o algoritmo de escalonamento FIFO. Nas experiências com duas classes, o algoritmo de escalonamento que vai ser usado é o CBQ.

O conjunto de experiências efectuado no caso de sistemas com duas classes é dividido em dois sub-conjuntos: no primeiro, o tráfego oferecido é constante ao longo do

tempo; no segundo, o tráfego oferecido é variável no tempo. A motivação para considerar este segundo conjunto de experiências é aumentar intencionalmente a probabilidade de roubo de recursos, de forma a poder averiguar, num cenário de pior caso, o desempenho dos mecanismos de *probing* simples e de ?-*probing*.

As simulações apresentadas nesta secção e nas seguintes consideram o mecanismo que termina o processo de *probing* assim que o receptor verificar que o número de pacotes perdidos atingiu um valor para o qual a percentagem de perdas total é superior ao limiar.

Nesta secção considera-se que as perdas sofridas pelos fluxos de dados são as perdas reais e que as perdas dos fluxos de *probing* e de ?-*probing* são as perdas estimadas.

Quando não é dito nada em contrário, considera-se um tempo máximo de *probing* de 2 segundos. Os valores dos gráficos apresentados nesta secção correspondem a uma média de 5 valores obtidos em diferentes corridas da mesma experiência. O tempo de cada experiência é de 10,000 segundos, sendo o tempo de aquecimento de 1,000 segundos.

7.5.2.2.1 Experiências com uma classe

O objectivo deste conjunto de experiências é estudar o mecanismo de *probing* como um mecanismo de controle de admissão num ambiente sem problemas de roubo de recursos. Para valores diferentes de limiar de perdas, é estudada a influência, no rácio de perdas reais e estimadas, no bloqueio e na utilização, de diferentes variantes do *probing* (através de perdas ou marcas, com fluxos de *probing* no mesmo nível de prioridade dos fluxos de dados ou separados em dois níveis diferentes), do tempo de *probing* e dos modelos das fontes de tráfego dos fluxos de dados. A utilização é baseada apenas em pacotes de dados, sendo o tráfego de *probing* considerado tráfego adicional que diminui a utilização do sistema. Todas as curvas apresentadas nesta secção têm seis ou cinco pontos em que cada um corresponde a um limiar de perdas ou marcação diferentes. Da esquerda para a direita dos gráficos, os valores de limiar de perdas são de 0%, 1%, 2%, 3%, 4% e 5%. Quando são também consideradas as marcações de pacotes, os valores de limiar de marcação são de 0%, 5%, 10%, 15% e 20%.

Quando nada for dito em contrário, a média do intervalo entre chegadas dos fluxos é de $1/? = 3.5$ segundos e a duração média de cada fluxo é de $1/? = 300$ segundos. Considerando fontes de tráfego com uma taxa média de 128 Kb/seg (será este o caso nesta

secção), o tráfego médio oferecido é superior à largura de banda da ligação, e deste modo, a rede encontra-se congestionada.

Variantes do *probing* – Nesta primeira experiência é efectuado um estudo de simulação para averiguar a eficácia das várias formas de efectuar *probing*. Nos estudos comparativos que vão ser efectuados, é usado o algoritmo MBAC (*Measurement-Based Admission Control*) de soma média (*Measured Sum*) [Breslau00a]. Neste algoritmo, a largura de banda do conjunto dos fluxos activos é derivada através de um estimador de janela temporal. A janela temporal tem um comprimento *T* que é dividido em diversos períodos de amostragem. No final do tempo *T*, utiliza-se a média mais elevada da largura de banda agregada medida nos diversos períodos de amostragem como estimativa da largura de banda ocupada na próxima janela temporal. A decisão de controle de admissão é efectuada com base nesta estimativa e na largura de banda dos fluxos que pedem admissão.

Em [Breslau00b] são apresentadas duas formas de efectuar *probing*, as quais foram descritas na introdução deste capítulo: no mesmo nível de prioridade que o tráfego de dados ou num nível inferior. Outra variante é o parâmetro de QdS utilizado para tomar decisões de controle de admissão: pacotes perdidos ou marcados. Nesta secção foi replicado um dos cenários de simulação apresentado em [Breslau00b] para comparar o desempenho do mecanismo de *probing* considerando estas variantes. Assim, considera-se que a fonte de tráfego utilizada é a EXP_256_50 e o tempo de *probing* é de 5 segundos. Considera-se também que um pacote é marcado quando é inserido numa fila de espera com uma ocupação igual ou superior a 90%.

A Figura 7-15 apresenta o rácio de perdas reais em função da utilização da ligação. Os resultados das simulações são semelhantes aos apresentados em [Breslau00b]. As várias curvas encontram-se na zona da curva MBAC (com excepção da curva que corresponde a marcação de pacotes com fluxos de *probing* num nível de prioridade inferior) e muito próximas entre si. Para um dado valor de utilização, as perdas reais obtidas com os mecanismos de *probing* são muito semelhantes às perdas obtidas com o algoritmo MBAC. Este resultado mostra que os mecanismos de *probing* são apropriados para efectuar o controle de admissão no modelo de serviço de carga controlada da arquitectura IntServ. Verifica-se também que os valores de rácio de perdas e utilização em cada curva de mecanismo de *probing* são muito diferentes. Num extremo apresenta-se o *probing* efectuado num nível de prioridade inferior ao dos fluxos de dados em que o parâmetro de

QdS é o rácio de marcação de pacotes. Este mecanismo permite obter um rácio de perdas de 0.001% mas sacrifica bastante a utilização (74%). No outro extremo apresenta-se o *probing* efectuado no mesmo nível de prioridade que o dos fluxos de dados em que o parâmetro de QdS é o rácio de perdas de pacotes. Este mecanismo apresenta rácios de perdas superiores a 0.3% com utilizações elevadas de 88%.

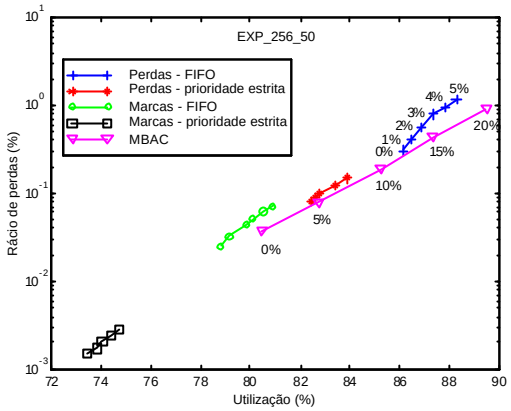


Figura 7-15 : Rácio de perdas em função da utilização para diferentes valores de limiar (algoritmos MBAC e mecanismos de *probing*.

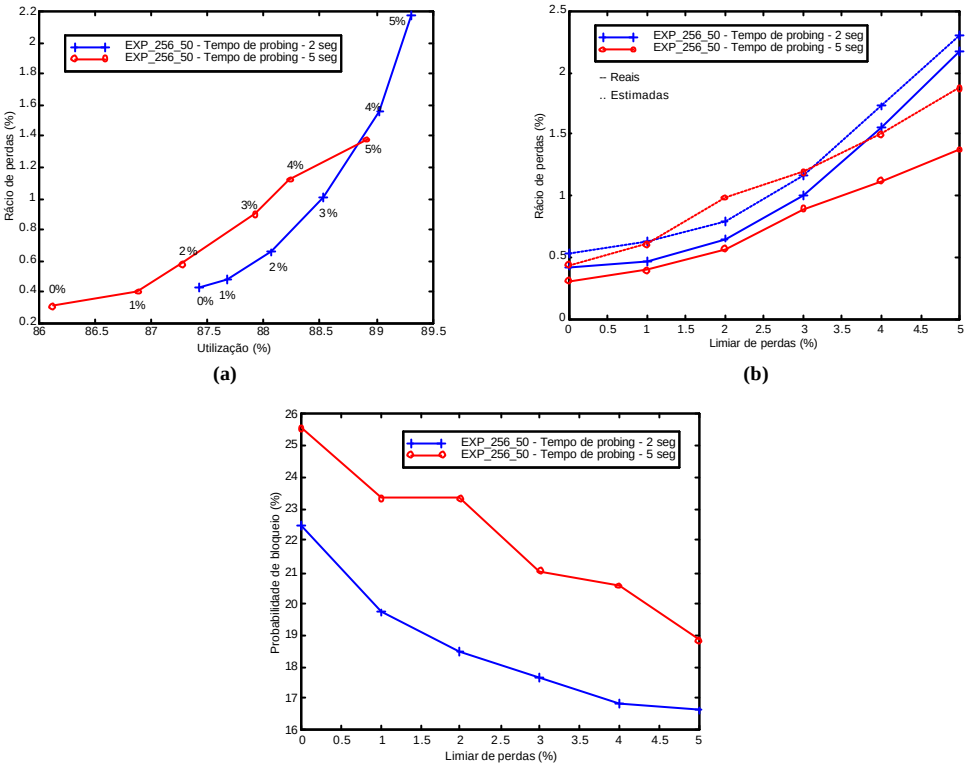
Embora o *probing* que considera como parâmetro de QdS o rácio de marcação de pacotes permita obter um rácio de perdas muito pequeno, estas perdas são muito inferiores ao limiar de decisão. Este resultado significa que é muito difícil para estes mecanismos relacionar o limiar de decisão com o nível de perdas desejado. Nestes dois casos, o limiar de decisão é um limite superior muito impreciso do nível de perdas desejado.

Nos outros dois casos, e principalmente quando o tráfego de *probing* tem o mesmo nível de prioridade que o de dados, o rácio de perdas é mais elevado, mas exceptuando o valor de perdas para limiar de decisão nulo, todos os valores são inferiores ao limiar. Isto significa que, a menos do limiar nulo, este mecanismo efectua um controle de admissão correcto ao mesmo tempo que permite utilizar os recursos de uma forma eficiente. Note-se que nenhum dos mecanismos, incluindo o MBAC, atingiu perdas nulas para um limiar nulo.

Tempo de *probing* - Nesta experiência é estudada a influência do tempo de *probing* no desempenho do mecanismo de *probing.* São considerados dois valores de tempo de *probing*: 2 e 5 segundos. Os fluxos de dados são EXP_256_50. A Figura 7-16 apresenta (a)

o rácio de perdas reais e o respectivo valor de utilização da rede para valores diferentes de limiar, (b) o rácio de perdas reais e estimadas em função do limiar, e (c) as respectivas probabilidades de bloqueio. Observa-se nas figuras que um aumento no limiar de perdas induz um aumento no rácio de perdas reais e na utilização dos recursos do sistema. Um aumento no limiar permite que mais fluxos entrem no sistema, aumentando tanto a utilização como as perdas reais. Verifica-se também que, com excepção do limiar nulo, as perdas reais nunca excedem o limiar. O maior rácio de perdas obtido é de 2.2% para um limiar de 5%, uma utilização de 89.3%, e tempo de *probing* de 2 segundos. Estes valores permitem concluir que o sistema está a realizar um controle de admissão correcto, isto é, não permite que o rácio de perdas reais se sobreponha ao limiar, ao mesmo tempo que a utilização do sistema se mantém em níveis elevados. Note-se também que o rácio de perdas estimadas é sempre superior ao das perdas reais (Figura 7-16 (b)). Isto deve-se à própria sobrecarga introduzida pelos fluxos de *probing*, os quais fazem aumentar o rácio de perdas durante o tempo de *probing*. Quando as perdas estimadas são superiores ao limiar, os fluxos são bloqueados e as perdas reais mantêm-se inferiores ao limiar.

Verifica-se também que com o aumento do tempo de *probing*, o rácio de perdas reais e a utilização dos recursos diminuem para o mesmo limiar de perdas. Um tempo de *probing* mais longo conduz assim, a uma maior diferença entre o limiar de perdas e as perdas reais observadas. Seria de esperar que um maior tempo de *probing* permitisse aproximá-las, em consonância com os resultados das secções 7.3 e 7.5.2.1. Um tempo de *probing* mais longo conduz também a uma maior diferença entre perdas reais e estimadas, e a uma maior probabilidade de bloqueio (Figura 7-16 (b e c)). Este comportamento pode ser explicado da seguinte forma. Conforme já referido, o processo de *probing* perturba o próprio sistema que pretende analisar. Cada novo fluxo de *probing* introduz um novo conjunto de pacotes no sistema, fazendo aumentar o tráfego. O tráfego adicional devido aos fluxos de *probing* aumenta com o tempo de *probing*. Não havendo controle de admissão o rácio de pacotes de dados perdidos aumentaria. No entanto, o mecanismo de controle de admissão reage, bloqueando um maior número de fluxos, e diminuindo o rácio de perdas reais. Este efeito sobrepõe-se ao que conduziria a um maior rácio de perdas reais (mais próximo do limiar). A mesma situação foi observada em [Breslau00b].



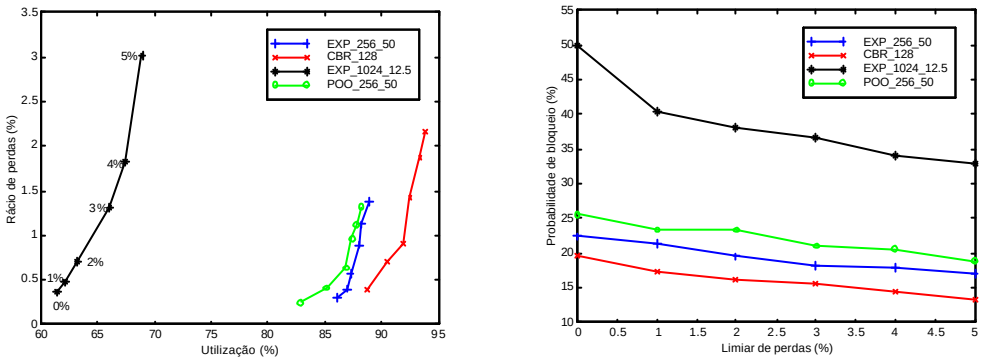
(a) **(b)**

(c)

Figura 7-16 : Desempenho do mecanismo de *probing* em função do tempo de *probing*: (a) rácio de perdas reais *vs* utilização, (b) rácio de perdas reais e estimadas *vs* limiar e (c) probabilidade de bloqueio *vs* limiar.

Tipos de fontes de tráfego - Na experiência seguinte é estudado o impacto do modelo das fontes de tráfego (Figura 7-17 (a e b)). As fontes de tráfego utilizadas nesta experiência são CBR_128, EXP_256_50, POO_256_50 e EXP_1024_12.5, todas com a mesma taxa média. Observa-se, na Figura 7-17 (a), que a maior utilização é atingida com as fontes CBR_128, é menor nas fontes *on-off* de Pareto do que nas fontes *on-off* exponenciais, e diminui com o aumento do *burstiness* da fonte, ou seja, com o aumento da diferença entre a taxa média e a taxa de pico. Este comportamento pode ser explicado da seguinte forma. Embora a taxa média das fontes seja a mesma, a largura de banda efectiva das fontes *on-off* é superior à largura de banda média e, consequentemente, é superior à largura de banda das fontes CBR. Este facto introduz uma diminuição no número de fluxos admitidos quando estes são do tipo *on-off* (Figura 7-17 (b)), e consequentemente, uma diminuição na

utilização. Como a distribuição de Pareto apresenta caudas mais longas (com $?$=1.2) que a distribuição exponencial, a variância dos tempos *on* e *off* é maior. Deste modo, como existe alguma probabilidade de os tempos *on* numa distribuição de Pareto serem mais longos que na distribuição exponencial (e o mesmo se aplica aos tempos *off*), menos fluxos POO_256_50 são admitidos (superior probabilidade de bloqueio na figura), e por isso a utilização do sistema é ligeiramente inferior à de fontes *on-off* exponenciais EXP_256_50. Em relação ao rácio de perdas reais observado, verifica-se que este também depende da taxa de pico das fontes de tráfego. Existem assim, dois efeitos opostos. Por um lado, uma menor utilização da rede indica que existem menos fluxos activos no sistema e por isso, o rácio de perdas reais é menor. É este efeito que predomina na diferença entre as perdas reais das fontes CBR_128, EXP_256_50 e POO_256_50. Por outro lado, uma maior taxa de pico aumenta a sobrecarga no sistema introduzida pelos fluxos de *probing*. Note-se que a fonte EXP_1024_12.5 apresenta uma taxa de pico aproximadamente 10 vezes inferior à capacidade da ligação (10 Mb/seg). Assim, cada fluxo que tenta admissão pode introduzir perdas consideráveis. É este efeito que predomina na diferença entre as perdas reais das fontes EXP_256_50/POO_256_50 e EXP_1024_12.5.



(a) **(b)**

Figura 7-17 : Desempenho do mecanismo de *probing* em função dos tipos de fontes de tráfego: (a) rácio de perdas reais *vs* utilização e (b) probabilidade de bloqueio *vs* limiar.

Largura de banda dos fluxos de dados - Este último conjunto de experiências num sistema com uma classe tem o objectivo de verificar o impacto da variação da largura de banda dos fluxos de dados. Os fluxos de dados são CBR_64, CBR_128 e CBR_512. Para ter o mesmo tráfego oferecido nas três experiências, $1/?$ é de 1.75, 3.5 e 14 segundos,

respectivamente para os fluxos CBR_64, CBR_128 e CBR_512, enquanto 1/? é mantido constante e igual a 300 segundos. A Figura 7-18 apresenta o rácio de perdas reais em função da utilização para os diversos limiares de perdas. Para fluxos de dados com menor largura de banda é possível atingir utilizações mais elevadas e perdas menores. Para o mesmo limiar de perdas, os fluxos mais pequenos atingem maiores utilizações porque, havendo pouca largura de banda disponível no sistema, a probabilidade de um fluxo com menor largura de banda ser admitido é maior do que a de um fluxo com maior largura de banda. Ainda para o mesmo limiar de perdas, o rácio de perdas reais aumenta com a largura de banda dos fluxos de dados (exceptuando o limiar de perdas nulo). Este comportamento deve-se ao facto de, quanto maior for a largura de banda dos fluxos de dados, maior é a dos fluxos de *probing*, e assim, maior é a sobrecarga no sistema gerada por um fluxo de *probing*. Note-se também que, a menos do limiar nulo, as perdas reais são sempre inferiores ao limiar.

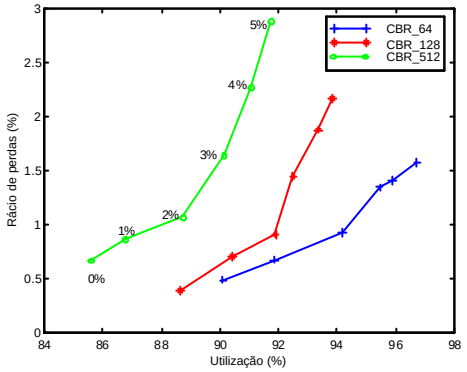


Figura 7-18 : Rácio de perdas reais em função da utilização (fluxos de dados com larguras de banda diferentes).

Principais conclusões - As experiências realizadas nesta secção permitem confirmar algumas conclusões de [Breslau00b]:

? O mecanismo de *probing* não permite garantir perdas nulas. O rácio de perdas reais tende a ser sempre inferior ao limiar de perdas, exceptuando os casos de limiares muito baixos ou nulos.

? A utilização de uma rede em que o controle de admissão é efectuado por mecanismos de *probing* diminui com o *burstiness* das fontes de tráfego e com o aumento na largura de banda dos fluxos de dados.

7.5.2.2.2 Experiências com duas classes e tráfego oferecido constante

Neste conjunto de experiências, cada cenário contém sempre duas classes de serviço com requisitos de QdS diferentes ou com características dos fluxos de dados diferentes. O tráfego oferecido a cada classe é constante, isto é, a taxa de chegada de fluxos e o tempo médio de permanência no sistema não variam ao longo da experiência. O objectivo é estender os estudos anteriores a cenários com várias classes em que o roubo de recursos possa existir, e analisar o comportamento dos mecanismos de *probing* e de ?-*probing*. Este conjunto de experiências baseia-se num cenário simples com uma rede constituída apenas por uma ligação. O peso atribuído a cada classe no escalonador CBQ é de 50%.

O tráfego oferecido a cada classe é sempre proporcional aos pesos, com excepção da última experiência. Em todas as experiências que consideram uma largura de banda da ligação de 10 Mb/seg, com excepção da última, 1/ ?₁= 3.2 segundos e 1/ ?₁= 300 segundos na classe 1, e 1/ ?₂= 2.2 segundos e 1/ ?₂= 100 segundos na classe 2. As fontes de tráfego que serão utilizadas na classe 1 e na classe 2 têm, respectivamente, uma largura de banda média de 64 e 128 Kb/seg. Exceptuando estas diferenças, todos os outros parâmetros do sistema são os mesmos que os usados nas experiências com uma classe. O limiar de perdas para os fluxos de *probing* e de ?-*probing* é de 0%, sempre que nada for dito em contrário.

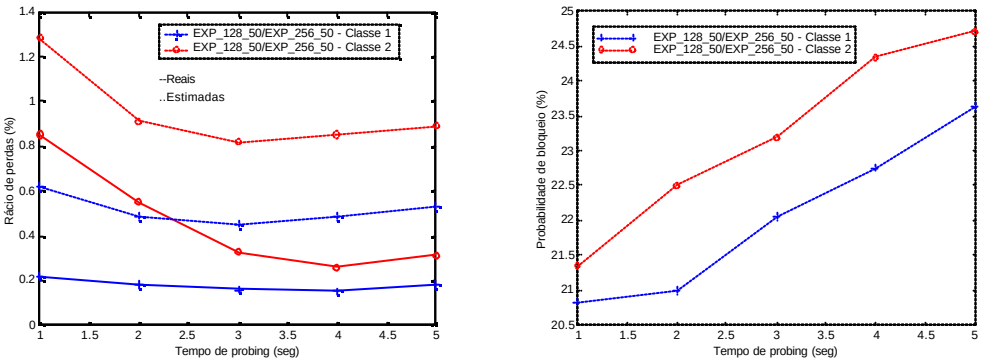
O mecanismo de *probing* simples é representado nas figuras pelos valores de largura de banda nula dos fluxos de ?-*probing.*

Tempo de *probing* - Na primeira experiência, o parâmetro do sistema que é variável é o tempo de *probing,* entre 1 e 5 segundos. As fontes de tráfego utilizadas na classe 1 e 2 são, respectivamente, EXP_128_50 e EXP_256_50. A largura de banda dos fluxos de ?-*probing* é de 32 Kb/seg. A Figura 7-19 (a) apresenta as curvas de rácio de perdas reais e estimadas em função do tempo de *probing*. A Figura 7-19 (b) apresenta as curvas de probabilidade de bloqueio das duas classes. Verifica-se que o rácio de perdas reais é sempre superior ao limiar, pois este é de 0%. O rácio de perdas reais nas duas classes é sempre inferior ao das estimadas.

Para tempos de *probing* pequenos, as perdas estimadas são elevadas, devido ao facto de o tempo de *probing* não ser suficiente para estimar com a precisão necessária as perdas reais na rede. Este efeito foi também observado na secção 7.5.2.1. Com o aumento do tempo de *probing* até 3 segundos, a precisão das estimativas aumenta e as perdas estimadas

diminuem. No entanto, para tempos de *probing* entre 3 e 5 segundos, as perdas estimadas aumentam ligeiramente. Isto explica-se pelo facto de, como já foi referido nas experiências com uma classe, o processo de *probing* perturbar o próprio sistema que pretende analisar e, assim sendo, o aumento do tempo de *probing* introduz um aumento nas perdas dos fluxos de *probing* e de ?-*probing*. Este efeito foi também observado nas experiências com uma classe de serviço e nas experiências apresentadas em [Breslau00b]. As perdas reais diminuem para tempos de *probing* até 4 segundos, devido ao facto de o mecanismo de controle de admissão bloquear um maior número de fluxos à medida que o tempo de *probing* aumenta (Figura 7-19 (b)). Para tempos de *probing* maiores do que 4 segundos, existe um ligeiro aumento no rácio de perdas reais. Este comportamento deve-se ao facto de o tráfego de *probing* se tornar significativo, contribuindo ele próprio para a degradação das perdas reais.

O rácio de perdas reais na classe 2 é sempre superior ao da classe 1 porque, como a largura de banda dos seus fluxos é maior, mais *probes* são gerados no mesmo tempo de *probing*, introduzindo uma maior sobrecarga no sistema.



(a) **(b)**

Figura 7-19 : Desempenho do mecanismo de ?-*probing* com a variação do tempo de *probing*: (a) rácio de perdas reais e estimadas *vs* tempo de *probing*, (b) probabilidade de bloqueio *vs* tempo de *probing*.

Diferenciação de serviços - Na experiência seguinte o tempo de *probing* é fixo, de 2 segundos, e é usado um limiar de perdas diferente para cada classe: 2% e 0%. Na Figura 7-20 é apresentado, para fluxos de dados *on-off* exponenciais (EXP_128_50 na classe 1 e EXP_256_50 na classe 2), o rácio de perdas reais em cada uma das classes, com limiares diferentes, em função da largura de banda dos fluxos de ?-*probing*.

O rácio de perdas reais é superior ao limiar quando é considerado o mecanismo de *probing* simples, mas decresce rapidamente com o aumento da largura de banda dos fluxos de ?-*probing*. A existência de um rácio de perdas superior ao limiar (com *probing* simples) deve-se ao roubo de recursos, induzido pela partilha de recursos entre classes. Com fluxos de ?-*probing* iguais ou superiores a 2 Kb/seg o rácio de perdas reais na classe 1 passa a ser inferior ao limiar. Na classe 2, tal como observado nas experiências com uma classe de serviço, o limiar de perdas nulo nunca é completamente atingido. No entanto, o aumento na largura de banda dos fluxos de ?-*probing* permite reduzir o rácio de perdas reais na classe 1 para menos de 0.3%. Este estudo mostra que o mecanismo de ?-*probing* reduz o roubo de recursos entre classes, e consequentemente, diminui o rácio de perdas reais nas duas classes.

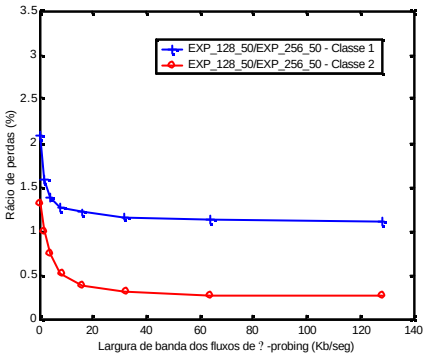


Figura 7-20 : Rácio de perdas em função da largura de banda dos fluxos de ?-*probing* (diferenciação entre classes de serviço).

Pela observação da figura é também possível verificar que existe diferenciação entre as classes de serviço, isto é, o rácio de perdas reais observado nas duas classes é distinto de acordo com os limiares de perdas pré-definidos. O mecanismo de ?-*probing* é, deste modo, capaz de proporcionar a diferenciação de QdS e permitir a partilha de recursos entre classes, minimizando o roubo de recursos.

Tipos de fontes de tráfego - Na Figura 7-21 é ilustrada a eficiência de ambos os mecanismos de *probing* e de ?-*probing* em sistemas com fluxos de dados modelados por fontes de tráfego diferentes: CBR, EXP e POO. As fontes utilizadas em cada classe têm a mesma largura de banda média de 64 Kb/seg na classe 1 e de 128 Kb/seg na classe 2. O rácio de perdas reais apresentado na figura corresponde a uma média das duas classes.

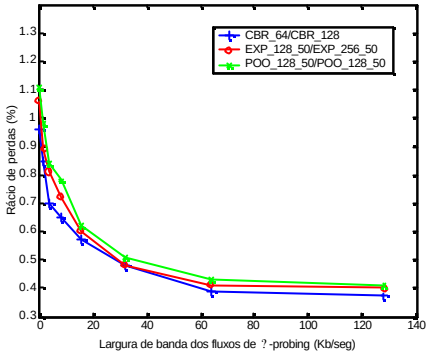


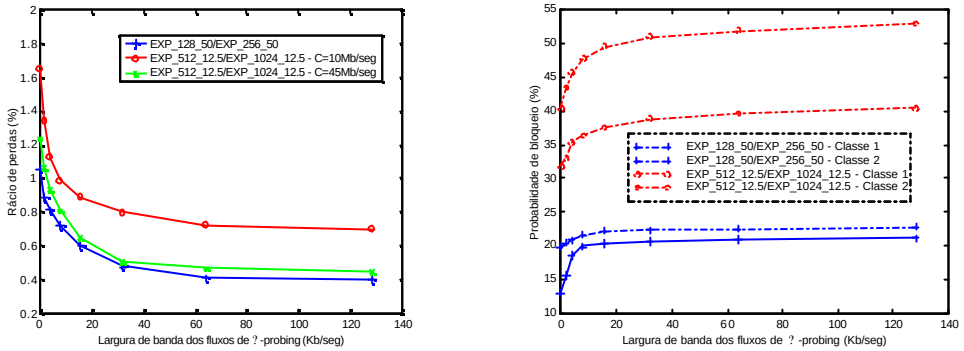
Figura 7-21 : Rácio de perdas em função da largura de banda dos fluxos de ?-*probing* (tipos de fontes de tráfego).

Na figura verifica-se que as curvas têm o mesmo comportamento, existindo apenas ligeiras diferenças entre elas. O rácio de perdas reais das fontes *on-off* é ligeiramente superior. Recorde-se que, nas experiências com uma classe de serviço apresentadas nesta mesma secção, o rácio de perdas das fontes *on-off* era ligeiramente inferior. Estas ligeiras diferenças podem ser atribuídas ao roubo de recursos entre classes que existe nestas experiências com duas classes.

***Burstiness* das fontes de tráfego** - Na Figura 7-22 é analisado o efeito de aumentar o *burstiness* das fontes *on-off* exponenciais. As fontes utilizadas são EXP_128_50 e EXP_256_50, respectivamente na classe 1 e na classe 2, com *burstiness* de 50%, e fontes EXP_512_12.5 e EXP_1024_12.5 (classe 1 e classe 2) com *burstiness* de 12.5%. Vão ser considerados dois casos diferentes para o segundo conjunto de fontes de tráfego. No primeiro caso, a capacidade da ligação é de 10 Mb/seg e os valores de $?_1$ e $?_2$ são iguais aos utilizados no primeiro conjunto de fontes; no segundo caso, a capacidade da ligação aumenta para 45 Mb/seg e os valores de $?_1$ e $?_2$ aumentam 4.5 vezes em relação aos valores anteriores. O rácio de perdas reais apresentado na figura corresponde a uma média das duas classes.

Observa-se que o rácio de perdas reais num sistema com fontes de 512 e 1024 Kb/seg e capacidade da ligação de 10 Mb/seg é aproximadamente o dobro do rácio de perdas do sistema com fontes de 128 e 256 Kb/seg. Existem duas razões que explicam este comportamento. A primeira está relacionada com o facto de a taxa de pico das fontes com *burstiness* de 12.5% ser 4 vezes maior do que a das outras fontes, o que introduz maiores

perdas nas filas de espera. A segunda razão centra-se no facto de os fluxos de *probing*, enviados numa situação em que o *burstiness* das fontes de tráfego é de 12.5%, terem maior largura de banda, introduzindo uma maior sobrecarga no sistema. Em relação às curvas de bloqueio, observa-se que com o aumento da taxa de pico das fontes, para a mesma capacidade da ligação de 10 Mb/seg, a probabilidade de bloqueio passa de valores da ordem dos 20% para valores da ordem dos 40%. Estes valores estão de acordo com os observados na Figura 7-17 (b) nas experiências com apenas uma classe. A probabilidade de bloqueio da classe 2 é superior à da classe 1 devido à maior largura de banda dos fluxos da classe 2 (maiores fluxos introduzem perdas maiores que se reflectem numa maior probabilidade de bloqueio).



(a) **(b)**

Figura 7-22 : Desempenho dos mecanismos de *probing* e de ?-*probing* em função do *burstiness* das fontes de tráfego: (a) rácio de perdas *vs* largura de banda dos fluxos de ?-*probing* e (b) probabilidade de bloqueio *vs* largura de banda dos fluxos de ?-*probing*.

O aumento da capacidade da ligação para 45 Mb/seg reduz o rácio de perdas reais observado, aproximando-o do rácio de perdas existente no sistema com fontes EXP_128_50 e EXP_256_50. O aumento da capacidade da ligação tem dois efeitos. Primeiro, a multiplexagem de um maior número de fontes (que ocorre para uma capacidade da ligação maior) atenua as perdas nas filas de espera provocadas pelo elevado *burstiness* das fontes. Segundo, como a diferença entre a largura de banda dos fluxos de *probing* e a da ligação aumenta, as perdas introduzidas pelos fluxos de *probing* diminuem.

Um outro aspecto importante a salientar das curvas de probabilidade de bloqueio é o facto de o aumento do bloqueio com o aumento da largura de banda dos fluxos de ?-*probing* ser significativo apenas para fluxos de ?-*probing* inferiores a 32 Kb/seg. Para este

caso, fluxos de 32 Kb/seg são suficientes para detectar a possibilidade de existir roubo de recursos entre classes.

Tráfego oferecido não proporcional aos pesos do CBQ - Na última experiência deste conjunto, os pesos do escalonador CBQ não são proporcionais ao tráfego oferecido. O objectivo desta experiência é aumentar o roubo de recursos existente entre classes. Relativamente às experiências anteriores, o tráfego oferecido à classe 1 é aumentado ($1/?_1$ diminui de 3.2 para 2.8 segundos) e o da classe 2 é diminuído ($1/?_2$ aumenta de 2.2 para 2.6 segundos). As fontes de tráfego usadas são CBR_64 e CBR_128, respectivamente, na classe 1 e 2. Assim sendo, neste caso o tráfego oferecido à classe 1 é maior que o oferecido à classe 2 e, como os pesos das classes são iguais, existirá alguma tendência para que a classe 1 use parte da largura de banda garantida à classe 2.

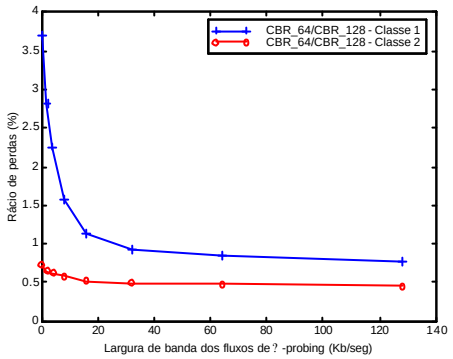


Figura 7-23 : Rácio de perdas em função da largura de banda dos fluxos de ?-*probing* (desajuste do tráfego oferecido com os pesos do CBQ).

Na Figura 7-23 são apresentadas as curvas de rácio de perdas reais em ambas as classes em função da largura de banda dos fluxos de ?-*probing*. Com o mecanismo de *probing* simples, o rácio de perdas na classe 1 atinge quase 4%. Neste caso, sempre que o tráfego da classe 2 é inferior à largura de banda garantida, a classe 1 tenta usar a sua largura de banda disponível com sucesso. Os fluxos da classe 1 vão ver os seus recursos roubados porque nesta situação, os novos pedidos na classe 2 vão ser aceites devido à largura de banda usada por eles estar abaixo da largura de banda garantida. O custo desta aceitação é o roubo de recursos na classe 1.

Com o aumento da largura de banda dos fluxos de ?-*probing*, estes analisam o estado de congestão na outra classe (na classe 1) e impedem que os recursos que estão a ser utilizados por essa classe sejam atribuídos à classe 2. Observa-se que para valores de fluxos de ?-*probing* de 32 Kb/seg, o rácio de perdas das duas classes torna-se muito próximo.

Principais conclusões

? Em sistemas com várias classes de serviço, os mecanismos de ?-*probing* são capazes de garantir diferenciação de QdS.

? A sobrecarga introduzida pelos próprios fluxos de *probing* no sistema aumenta quando diminui a diferença entre a capacidade da ligação e a taxa de pico das fontes de tráfego.

? Em sistemas com várias classes e tráfego oferecido constante, a partilha de recursos entre classes introduz algum roubo de recursos com o mecanismo de *probing* simples. O mecanismo de ?-*probing* consegue atenuar este problema.

7.5.2.2.3 Experiências com duas classes e tráfego oferecido variável no tempo

Neste conjunto de experiências é efectuada uma investigação de como os mecanismos de *probing* simples e de ?-*probing* reagem a perturbações no sistema. Uma perturbação é definida no contexto deste estudo como uma variação temporal do tráfego oferecido que visa potenciar o roubo de recursos. Concretamente, o tráfego oferecido à classe 2 alterna entre um valor reduzido (em que a classe 1 utiliza recursos da classe 2) e um valor elevado (em que a classe 2 tenta recuperar os recursos, roubando-os à classe 1).

A topologia da rede utilizada para efectuar estes estudos encontra-se na Figura 7-24. Consideram-se duas classes de serviço. A classe 1 tem tráfego oferecido $r_{1,AD}$ constante ao longo da experiência, com $?_1$ = 1/2.0 e $?_1$ = 1/300, ocupando 80% da capacidade da ligação. A classe 2 é constituída por duas sessões de tráfego com percursos diferentes: $r_{2,AD}$ constante ao longo da experiência com $?_{2A}$ = 1/6.8 e $?_{2A}$ = 1/100, ocupando 20% da capacidade da ligação; e $r_{2,BD}$ que é activado e desactivado periodicamente com $?_{2B}$ = 1/0.8 e $?_{2B}$ = 1/20. Nesta sessão de tráfego os fluxos são activados aos 1,000 segundos, os tempos de *on* são de 100 segundos e os tempos de *off* são de 200 segundos. O tráfego total oferecido à classe 2 na ligação CD em função do tempo de simulação encontra-se representado na Figura 7-25.

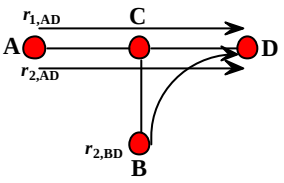


Figura 7-24 : Topologia da rede num cenário com perturbações.

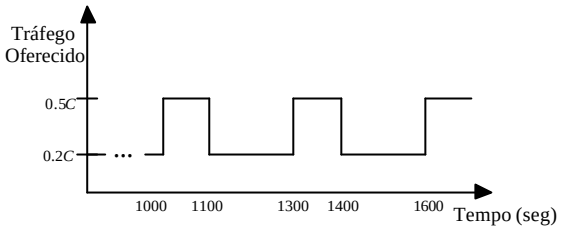


Figura 7-25 : Tráfego oferecido à classe 2.

Exceptuando estas diferenças nos tráfegos oferecidos a cada classe, todos os outros parâmetros do sistema são os mesmos que os usados nas experiências anteriores. Os valores apresentados nos gráficos seguintes são uma média dos valores obtidos a partir do instante de tempo de 1,000 segundos e num conjunto de 5 experiências.

Para analisar os resultados das experiências, além do rácio de perdas reais, são utilizadas duas métricas de desempenho diferentes: a percentagem de decisões erradas e a percentagem de largura de banda roubada. A primeira é a percentagem de fluxos que são aceites quando a largura de banda de todos os fluxos é superior à largura de banda da ligação. A última é a percentagem de largura de banda que é roubada pela admissão de novos fluxos, quando esta admissão é uma decisão errada. O cálculo destas métricas é feito da seguinte forma: sempre que existe uma decisão de admissão positiva, calcula-se a largura de banda ocupada por todos os fluxos admitidos, com base no número de fluxos e na sua largura de banda. Se o resultado é superior à capacidade da ligação (incluindo a tolerância permitida pelo limiar de perdas), considera-se que a decisão é errada. A percentagem de largura de banda roubada é uma média temporal determinada pelo rácio entre a largura de banda que é roubada em cada decisão errada e a largura de banda dos fluxos admitidos. A métrica de largura de banda roubada inclui toda a informação necessária sobre roubo de recursos. A métrica de decisões erradas vai ser utilizada apenas para mostrar a quantidade de decisões erradas que podem ser tomadas com o mecanismo

de *probing* simples. Nestes estudos de simulação são utilizadas fontes de tráfego CBR por forma a que a largura de banda dos fluxos esteja sempre bem caracterizada, independentemente de efeitos de multiplexagem estatística. Quando nada for dito em contrário, o limiar de perdas é nulo para todos os fluxos de *probing* e de ?-*probing*.

Sistema com e sem perturbações – Na Figura 7-26 (a, b, c) comparam-se situações sem perturbações com uma situação em que se considera perturbação do sistema. Para o efeito, é considerada a mesma experiência com fluxos de dados CBR apresentada na Figura 7-21, denominada na Figura 7-26 de "Mesmo tráfego oferecido", e a mesma experiência apresentada na Figura 7-23, denominada na Figura 7-26 de "Tráfego oferecido diferente". A designação de "Mesmo tráfego oferecido" significa que as duas classes apresentam o mesmo tráfego médio oferecido. A designação de "Tráfego oferecido diferente" significa que a primeira classe tem tráfego médio oferecido superior ao da segunda classe, exactamente com os mesmos parâmetros da experiência apresentada na Figura 7-23. A designação de "Perturbações" refere a situação em que é introduzida periodicamente uma perturbação no sistema.

Observa-se na Figura 7-26 (a) que a percentagem de decisões erradas é elevada quando apenas se considera o mecanismo de *probing* simples. Com o mecanismo de ?-*probing*, e com o aumento da largura de banda dos fluxos de ?-*probing*, a percentagem de decisões erradas diminui rapidamente. Com *probing* simples, a percentagem de decisões erradas atinge 44% no cenário em que o tráfego oferecido a cada classe não é proporcional ao peso da classe, e perto de 50% no cenário com perturbações. Neste caso, no sistema com perturbações, a largura de banda roubada (Figura 7-26 (b)) é aproximadamente de 5%. Verifica-se que para fluxos de ?-*probing* entre 16 e 32 Kb/seg existe uma tendência para as curvas estabilizarem. Para fluxos de ?-*probing* de 32Kb/seg, a percentagem de decisões erradas é inferior a 10% e a percentagem de largura de banda roubada é inferior a 0.2%. Este conjunto de experiências mostra que, mesmo numa situação de pior caso em que são inseridas propositadamente perturbações destinadas a potenciar o roubo de recursos, o mecanismo de ?-*probing* previne a admissão de mais tráfego do que aquele que pode ser servido pelo sistema.

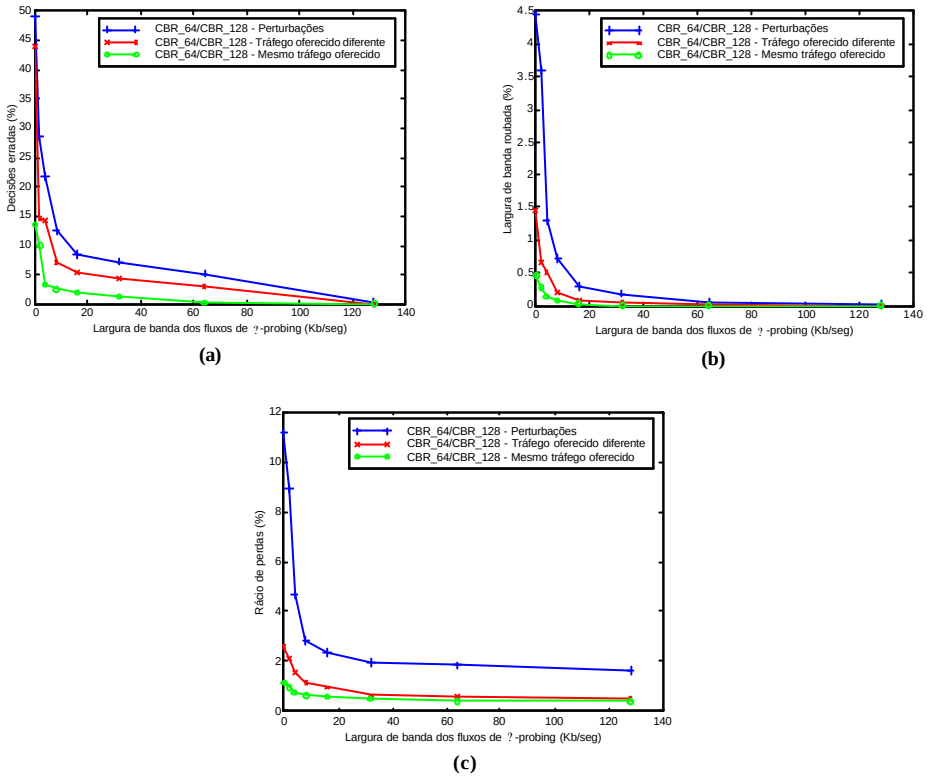


Figura 7-26 : Desempenho dos mecanismos de *probing* e de ?-*probing* face a perturbações introduzidas no sistema: (a) percentagem de decisões erradas, (b) percentagem de largura de banda roubada e (c) rácio de perdas reais, *vs* largura de banda dos fluxos de ?-*probing*.

Em relação ao rácio de perdas reais, observa-se na Figura 7-26 (c) que numa situação com perturbações e *probing* simples, este é muito elevado, superior a 10%. Com o aumento da largura de banda dos fluxos de ?-*probing*, estes impedem que a classe 2 roube recursos que estão a ser usados por fluxos previamente admitidos na classe 1, e deste modo, o rácio de perdas diminui bastante. No entanto, observe-se que esta taxa é sempre superior a 1.6%. Um dos problemas do mecanismo de *probing* é o facto de os próprios fluxos de *probing* introduzirem perdas no sistema. Neste caso, as perdas são superiores às existentes numa situação sem perturbações, porque a taxa de chegadas de fluxos na classe 2 é muito elevada. O intervalo entre chegadas de fluxos na sessão da classe 2 entre *B* e *D* é de 0.8 segundos e o tempo de vida de cada fluxo é de apenas 20 segundos, ou seja, apenas

10 vezes superior ao tempo de *probing*. Nesta situação, existem muitos fluxos de *probing* a pedir admissão simultaneamente.

Largura de banda dos fluxos de dados - Para avaliar o desempenho dos mecanismos de *probing* em sistemas em que a largura de banda dos fluxos é diferente, são usadas as fontes de tráfego de CBR_32 a CBR_1024. Os resultados são apresentados na Figura 7-27 (a, b e c). As percentagens de decisões erradas e de largura de banda roubada apresentadas consideram um limiar de perdas de 0%. O rácio de perdas reais apresentado considera limiares de perdas de 0% e de 5%.

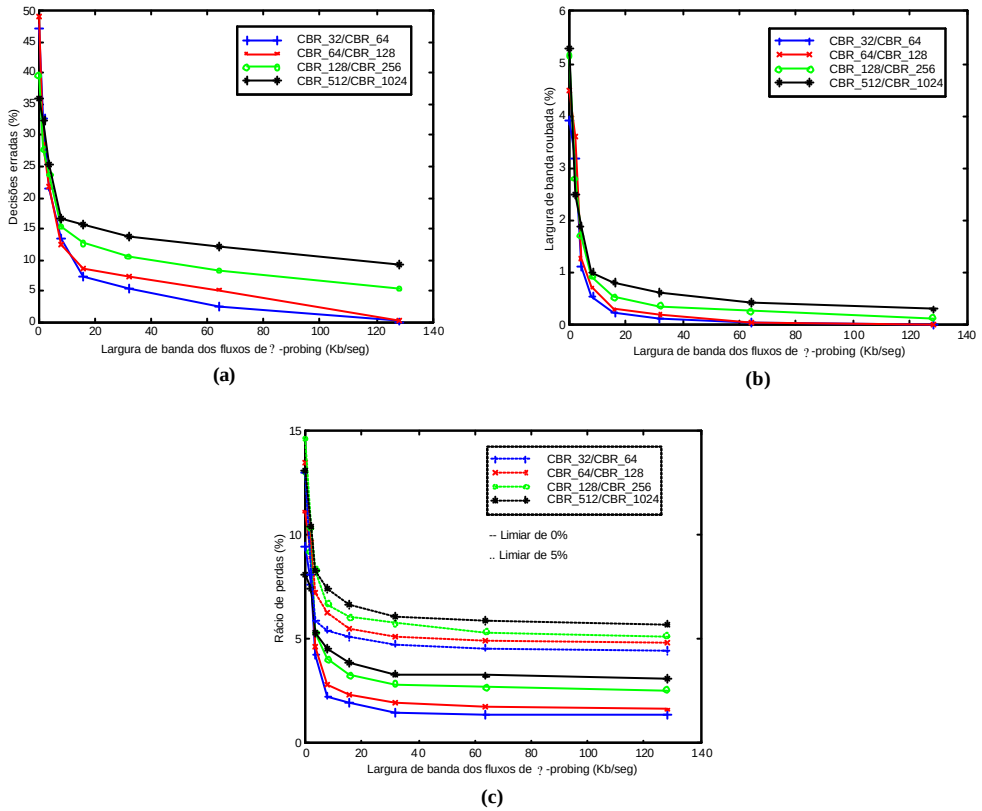


Figura 7-27 : Desempenho dos mecanismos de *probing* e de ?-*probing* em função da largura de banda dos fluxos de dados: (a) percentagem de decisões erradas, (b) percentagem de largura de banda roubada e (c) rácio de perdas reais, *vs* largura de banda dos fluxos de ?-*probing* (limiar de 0% e 5%).

Pela visualização dos gráficos verifica-se que, embora a percentagem de decisões erradas diminua com o aumento da largura de banda dos fluxos de ?-*probing*, esta aumenta

com o aumento da largura de banda dos fluxos de dados, sendo de 9% para fluxos de dados CBR_512/CBR_1024 e para fluxos de ?-*probing* de 128 Kb/seg. O mesmo comportamento está presente nas curvas de percentagem de largura de banda roubada. Um aumento na largura de banda dos fluxos de dados introduz um aumento na largura de banda das perturbações, que por sua vez aumentam a largura de banda roubada. Para fluxos de dados com largura de banda elevada e para os valores dos fluxos de ?-*probing* apresentados, verifica-se que existem sempre decisões erradas e largura de banda roubada.

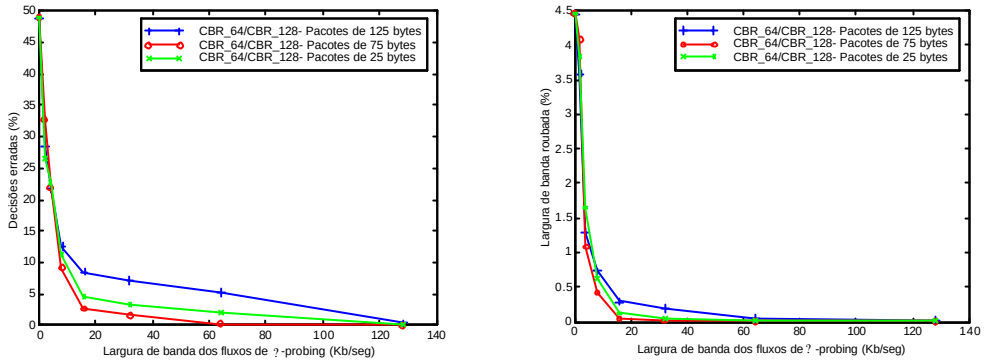
Em relação ao rácio de perdas reais, verifica-se que este é sempre superior ao limiar, quando este é de 0%. Para limiares de 5%, o rácio de perdas reais é inferior ao limiar apenas quando a largura de banda dos fluxos de dados é inferior a 128 Kb/seg. Com o aumento da largura de banda dos fluxos de dados, as perturbações introduzidas pelos fluxos de *probing* tornam-se muito elevadas. Uma vez mais se verifica que o mecanismo de ?-*probing* não permite obter garantias estritas de QdS, mesmo para valores relativamente elevados de limiares de perdas.

Assim como nas experiências da Figura 7-26, também nestas experiências as curvas tendem a estabilizar para valores de largura de banda dos fluxos de ?-*probing* entre 16 e 32 Kb/seg. As perdas na rede nestes casos são muito elevadas, e por isso, estes valores de largura de banda são suficientes para detectar a existência de perdas nas classes de serviço, e assim reduzir o roubo de recursos.

Comprimentos dos ?-*probes* - Nas experiências anteriores, o comprimento dos pacotes de dados, dos *probes* e dos ?-*probes* foi sempre de 125 *bytes*. Se a largura de banda de um fluxo é constante, a diminuição do comprimento dos pacotes do fluxo aumenta o número de pacotes que são gerados e enviados num mesmo período de tempo. Com mais ?-*probes* gerados no tempo de *probing*, é possível ter melhor confiança estatística no rácio de perdas de pacotes estimado. No entanto, é necessário saber se este facto tem por consequência diminuir a largura de banda roubada. Na Figura 7-28 (a e b) são apresentadas as mesmas métricas de desempenho (decisões erradas e largura de banda roubada) para diferentes comprimentos de ?-*probes*, mantendo todos os outros pacotes em 125 *bytes*.

Pela análise das figuras é possível confirmar que, com o aumento do número de pacotes gerados, o desempenho do sistema aumenta: ambas as percentagens de decisões erradas e de largura de banda roubada diminuem. No entanto observa-se que o desempenho

para ?-*probes* de 25 *bytes* é inferior ao desempenho para ?-*probes* de 75 *bytes*. Como já foi referido na secção 7.5.2.1, embora o número de pacotes gerados seja muito superior com pacotes de 25 *bytes*, no limite de uma fila de espera congestionada, a probabilidade de descarte de pacotes de 125 *bytes* é superior à de pacotes de 25 *bytes*. Neste caso, o rácio de perdas dos ?-*probes* será inferior ao rácio de perdas dos *probes* e dos pacotes de dados. Como as perdas detectadas na classe em que poderá existir roubo de recursos serão menores, a largura de banda roubada será maior para ?-*probes* muito pequenos.



(a) **(b)**

Figura 7-28 Desempenho dos mecanismos de *probing* e de ?-*probing* em função do comprimento dos ?-*probes*: (a) percentagem de decisões erradas e (b) percentagem de largura de banda roubada, *vs* largura de banda dos fluxos de ?-*probing*.

Na secção 7.5.2.1 verificou-se nas experiências em que as perdas reais são da ordem de 1%, que um comprimento de *probes* de 50 *bytes* impede que estas sejam detectadas com a precisão aí definida. Nestas experiências com perturbações, as perdas das ligações são de tal forma elevadas quando a perturbação é inserida, que a inserção de ?-*probes* de 25 *bytes* ainda permitem obter um melhor desempenho na redução do roubo de recursos que os ?-*probes* de 125 *bytes*. No entanto, como normalmente os valores das perdas na rede não são tão elevados, o comprimento médio dos *probes* e ?-*probes* deve ser semelhante ao comprimento médio dos pacotes de dados, para diminuir a discrepância entre as medidas realizadas. Este facto pode constituir uma limitação na aplicabilidade destes mecanismos.

Principais conclusões gerais

- Com tráfego oferecido variante no tempo, o roubo de recursos pode ser elevado. Nesta situação o mecanismo de ?-*probing* permite atenuar de forma significativa este problema.
- O comprimento dos ?-*probes* deve ser semelhante ao dos pacotes de dados. Quando as perdas introduzidas numa classe são muito elevadas, como foi o caso nesta secção, os pacotes pequenos conseguem detectá-las. No entanto, quando a rede se encontra pouco congestionada, apenas os pacotes com o mesmo comprimento que os de dados detectam as mesmas perdas que os pacotes de dados.

7.6 Conclusões e trabalho futuro

Este capítulo abordou o problema do roubo de recursos inerente aos mecanismos de *probing*, quando estes são utilizados em sistemas com múltiplas classes de serviço. Com o objectivo de atenuar este problema foi proposto um novo mecanismo de *probing*, denominado de ?-*probing*, que introduz simultaneamente fluxos de *probing* na classe a que pertence o fluxo que pede admissão, e fluxos de ?-*probing* de largura de banda menor nas restantes classes. Deste modo, o mecanismo de ?-*probing* impede que uma classe utilize recursos que não estão disponíveis, através da investigação do estado de congestão de todas as classes. Os estudos efectuados neste capítulo mostram que, numa rede com diferenciação de QdS, o mecanismo de ?-*probing* permite atingir uma elevada utilização e atenuar o problema do roubo de recursos.

Verificou-se nos estudos efectuados que a largura de banda dos fluxos de ?-*probing* pode ser inferior à largura de banda dos fluxos de dados. A largura de banda necessária dos fluxos de ?-*probing* é dependente da dos fluxos de dados. Em situações em que o tráfego oferecido é moderado (em que o bloqueio é da ordem dos 20%), verificou-se que, para reduzir significativamente o roubo de recursos existente em sistemas com múltiplas classes, a largura de banda dos fluxos de ?-*probing* deve encontrar-se entre 25% e 50% da largura de banda dos fluxos de dados. Para estes casos, o número de *probes* e de ?-*probes* enviados não deve ser inferior a valores da ordem de 256 e 64, respectivamente, para permitir que seja enviado um número de *probes* suficiente para estimar as perdas das classes e de ?-*probes* suficientes para detectar perdas nas classes. Por outro lado, o número

de *probes* e de ?-*probes* não deve ser superior a valores da ordem de 512 e 128, respectivamente, para limitar o tráfego de *probing* e de ?-*probing* que é introduzido na rede. Para valores de largura de banda dos fluxos de *probing* elevados, a precisão das estimativas depende principalmente do número de *probes* enviados e não da largura de banda dos fluxos de dados. Em relação ao comprimento dos pacotes, verificou-se que tanto os *probes* como os ?-*probes* devem ter o mesmo comprimento que os pacotes de dados, para não existirem discrepâncias nas medições efectuadas. Se se considerarem situações em que a largura de banda do tráfego oferecido é muito superior à capacidade da ligação (probabilidade de bloqueio muito elevada), e em que o intervalo entre chegadas de fluxos é muito inferior ao tempo de *probing*, as perdas reais tornam-se muito elevadas. Nestes casos verificou-se que a introdução de um número de ?-*probes* entre 32 e 64 permitiu reduzir significativamente o roubo de recursos entre classes, para uma gama de valores da largura de banda dos fluxos de dados entre 32 Kb/seg e 1024 Kb/seg.

Os problemas inerentes aos mecanismos de *probing* são o elevado tempo de estabelecimento das sessões, dependente do tempo de *probing*, e a sobrecarga introduzida na rede pelos próprios fluxos de *probing*. A última desvantagem não é característica de todos os algoritmos de escalonamento. O uso de uma prioridade inferior para os fluxos de *probing* impede que estes interfiram no tráfego já existente. No entanto, com estes mecanismos é mais difícil relacionar o limiar de perdas admissível com as que realmente se fazem sentir na rede. Deste modo, os recursos podem não ser utilizados da melhor forma.

Como conclusão geral, o mecanismo de ?-*probing* é um mecanismo extremamente simples de controle de admissão, que permite fazer uma boa utilização dos recursos da rede, fornecer diferenciação de serviços e garantias relativas de QdS a cada classe, e atenuar significativamente o problema de roubo de recursos em sistemas com múltiplas classes.

Em relação à aplicabilidade dos mecanismos de *probing*, estes possibilitam a implementação de serviços de carga controlada definidos na arquitectura IntServ. Existe um elevado número de serviços e aplicações que toleram o nível de perdas que se consegue discriminar com os mecanismos de *probing*.

O trabalho realizado neste capítulo pode ser complementado com uma análise teórica completa deste problema. Esta análise deveria conter a derivação de expressões para a utilização, para o roubo de recursos e para o rácio de perdas em função das características dos fluxos de dados, dos fluxos de *probing* e de ?-*probing*, e das características do sistema a analisar. Esta análise constitui um tópico em aberto para investigação futura.

CAPÍTULO 8

ESTUDO EXPERIMENTAL DOS MECANISMOS DE *PROBING*

O processo de controle de admissão baseado em mecanismos de *probing* é simples de implementar, permite atingir uma elevada utilização dos recursos garantindo simultaneamente o suporte de QdS (Qualidade de Serviço) por fluxo. No capítulo 7 foi proposto um novo mecanismo de *probing*, o ?-*probing*, com o intuito de resolver o problema do roubo de recursos que afecta os sistemas com múltiplas classes, quando estas partilham os recursos entre si. Neste capítulo foi também descrito um conjunto extenso de estudos baseados em simulações de eventos discretos. Com o objectivo de validar os resultados obtidos no capítulo 7, foi implementado um sistema experimental com uma arquitectura de base DiffServ (Diferenciação de Serviços) que suporta várias classes de serviço, e que inclui os mecanismos de controle de admissão *probing* e ?-*probing* [Além01, Sargento02a].

Este capítulo descreve os diversos passos desta implementação e apresenta os resultados obtidos nas diversas experiências realizadas. A secção 8.1 apresenta a arquitectura geral do sistema experimental e descreve algumas opções tomadas na sua definição. Nas secções 8.2 e 8.3 são descritos os módulos de *software* desenvolvidos. Na secção 8.4 é apresentado o cenário experimental utilizado nas experiências, e na secção 8.5 são apresentados e discutidos os resultados experimentais. Para finalizar, na secção 8.6 são

apresentadas as conclusões dos estudos efectuados e as tarefas propostas para trabalho futuro.

8.1 Arquitectura do sistema experimental

Nesta secção é descrita a arquitectura do sistema experimental que foi projectado com o objectivo de investigar o desempenho do mecanismo de ?-*probing* proposto. Este sistema tenta replicar uma rede DiffServ operacional.

O objectivo principal dos estudos experimentais é observar o comportamento dos mecanismos de *probing* e de ?-*probing* numa rede congestionada, isto é, com tráfego oferecido superior ao que a rede pode transportar. Idealmente, o tráfego gerado deveria partir de equipamentos terminais distintos. No entanto, devido às limitações inerentes ao facto de se tratar de um cenário laboratorial, restrito em termos de disponibilidade de equipamento, optou-se por uma solução alternativa. Assim, foi desenvolvido um módulo de *software* que gera todo o tráfego para cada classe de serviço ao nível do fluxo e do pacote. Este módulo foi designado de módulo de geração de tráfego. Devido a questões de desempenho do equipamento utilizado, no cenário experimental foi utilizado um terminal distinto para gerar o tráfego de cada classe de serviço.

Foi também desenvolvido um módulo para realizar todas as funções de *probing* e de ?-*probing* de todos os utilizadores, que se designou de módulo de *probing*. Na arquitectura do sistema experimental considera-se que o módulo de *probing* é instalado num PC dedicado, denominado de servidor de *probing*, que se encontra ligado a uma rede local delimitada por dois *routers*: um *router* de acesso e um *router* fronteira (Figura 8-1). O módulo de *probing* opera em modo promíscuo, isto é, processa todos os pacotes injectados na rede local. O módulo aceita pedidos de estabelecimento de fluxos provenientes do gerador de tráfego e implementa os mecanismos de *probing* e de ?-*probing*; este módulo é também responsável por marcar os pacotes de dados enviados pelo gerador de tráfego de acordo com a QdS pedida. O *router* fronteira da rede DiffServ necessita de implementar classificação e escalonamento de pacotes, funções estas que se podem encontrar actualmente em qualquer *router* IP (*Internet Protocol*), mesmo nos de menor custo. O *router* de acesso é usado apenas para isolar o tráfego entre a rede de acesso e a rede DiffServ. Deste modo, o conjunto formado pelos dois *routers* e pelo servidor de *probing*

emula um *router* fronteira DiffServ que inclui controle de admissão de fluxos baseado em mecanismos de *probing* e de ?-*probing*.

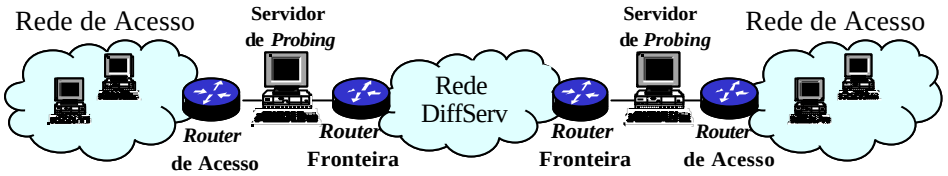


Figura 8-1 : Arquitectura do sistema experimental.

A interacção entre os vários elementos de rede é realizada através de protocolos da camada de aplicação desenvolvidos para o efeito. Os protocolos de transporte utilizados foram o TCP (*Transmission Control Protocol*), para a troca de informação de controle entre o gerador de tráfego e o módulo de *probing*, e o UDP (*User Datagram Protocol*) para troca de informação de controle entre os módulos de *probing*.

O fluxo de mensagens no sistema é ilustrado na Figura 8-2. O gerador de tráfego pede a admissão de um novo fluxo, estabelecendo uma ligação TCP com o módulo de *probing*, e enviando uma mensagem *REQUEST*. O objectivo desta mensagem é despoletar o processo de *probing* numa determinada classe de serviço. Esta mensagem inclui os endereços IP da origem e do destino, os portos UDP/TCP da origem e do destino, o tipo de protocolo e a classe de serviço à qual pertence o novo fluxo. Esta informação é necessária para que o módulo de *probing* possa identificar completamente o fluxo. Após receber a mensagem *REQUEST*, o módulo de *probing* de entrada da rede DiffServ inicia o processo de *probing*. Para o efeito, envia uma mensagem *PROBE START*, seguida da sequência de *probes*, e finaliza o processo de envio através da mensagem *PROBE STOP*. Todas estas mensagens são endereçadas ao utilizador destino e transportadas sobre UDP. Como já foi mencionado no capítulo 7, existem dois tipos de *probes*: os *probes* que são enviados apenas na classe à qual pertence o fluxo, e os ?-*probes* enviados nas restantes classes. O módulo de *probing* de saída da rede DiffServ escuta promiscuamente os pacotes de controle e os *probes*, e conta o número de *probes* recebidos em cada classe entre as mensagens *PROBE START* e *PROBE STOP*. Quando escuta a mensagem *PROBE STOP*, envia a mensagem *STATISTICS* para o módulo de *probing* de entrada, com as estatísticas dos pacotes recebidos. Se a mensagem *STATISTICS* não é recebida dentro de um *timeout* pré-definido, o fluxo é rejeitado e a ligação TCP com o gerador de tráfego é fechada. Se é

recebida dentro do *timeout*, o módulo de *probing* efectua uma decisão de controle de admissão com base na contagem dos *probes* e ?-*probes* transportados na mensagem *STATISTICS*, na contagem dos pacotes transmitidos e no limiar de perdas da classe respectiva. Se o fluxo é aceite, o módulo de *probing* envia uma mensagem *AUTHORIZE* ao gerador de tráfego, e fecha a ligação TCP que mantinha com este; se não é aceite, é enviada uma mensagem *REJECT* e a ligação TCP é também fechada. Se o fluxo é aceite, o gerador de tráfego inicia o envio de pacotes de dados que, no caso deste sistema experimental, são transportados sobre UDP. Para sinalizar o fim da transmissão, o módulo gerador de tráfego abre uma nova ligação TCP com o módulo de *probing* de entrada e envia uma mensagem *END SESSION*.

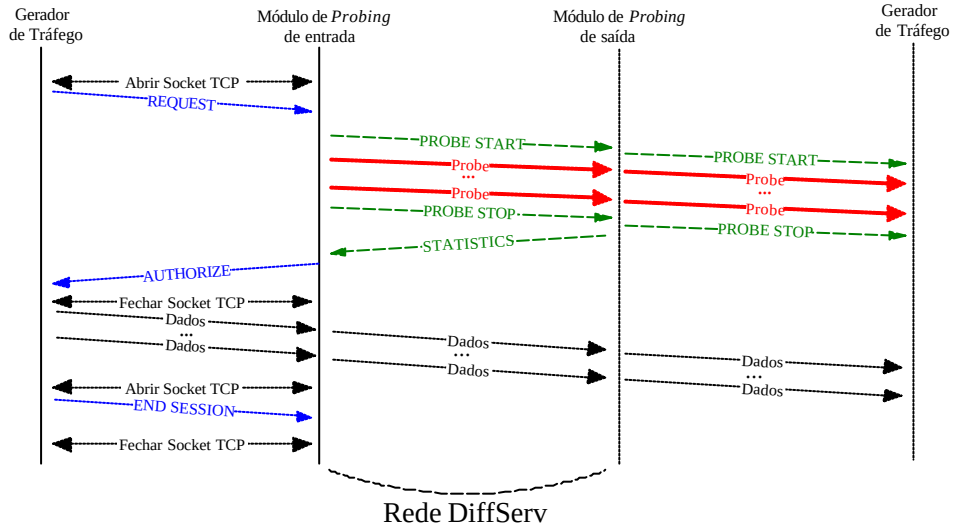


Figura 8-2 : Fluxos de mensagens entre os módulos gerador de tráfego e de *probing*.

As mensagens *REQUEST* e *END SESSION* têm ambas o mesmo formato e são identificadas por uma *flag* que é enviada juntamente com os restantes campos que foram descritos anteriormente. As mensagens *AUTHORIZE* e *REJECT* têm também ambas o mesmo formato, e são também identificadas por uma *flag* que toma o valor "1" na mensagem *AUTHORIZE* e "0" na mensagem *REJECT*. O cliente, após ter efectuado o pedido, fica à espera da decisão de aceitação ou rejeição do fluxo, e ao receber a mensagem com a *flag* a "1" ou "0", sabe que pode iniciar ou não a sessão. As mensagens de controle do processo de *probing*, *PROBE START*, *PROBE STOP* e *STATISTICS*,

incluem três campos: o primeiro campo identifica o tipo de mensagem; o segundo indica a classe de serviço; e o último campo é usado para transportar, na mensagem *STATISTICS*, os contadores de *probes* e de ?-*probes* em cada classe. Note-se que a informação trocada ao nível da camada de aplicação não é suficiente para definir completamente um fluxo. Os endereços IP e os portos TCP/UDP são também necessários. No entanto, dado que esta informação se encontra presente nos cabeçalhos IP e TCP/UDP, e como a camada de aplicação tem acesso a estes cabeçalhos, ela não é incluída no campo de informação, reduzindo assim o comprimento dos pacotes de controle.

A diferenciação entre classes de serviço é feita através dos *bits* ToS (*Type of Service*) contidos no *byte* ToS do pacote IP. É assumido na implementação que as mensagens de controle injectadas na rede DiffServ têm um nível de prioridade estrita superior a todas as outras classes que suportam os mecanismos de *probing*. Os *bits* de precedência do *byte* ToS são usados para diferenciar entre pacotes de controle, *probes*, ?-*probes* e pacotes de dados. Mais concretamente, é associado o valor 110 aos pacotes de controle, 010 aos *probes*, 100 aos ?-*probes* e 000 aos pacotes de dados. Nesta arquitectura, a manipulação do byte ToS é realizada pelo módulo de *probing*.

Ambos o gerador de tráfego e o módulo de *probing* foram desenvolvidos para funcionar sobre *Microsoft Windows* 2000. O *software* foi desenvolvido usando o *Microsoft Visual C++*, *Windows Sockets* 2.0, e recorre a técnicas de programação de *multiple thread*. Na implementação dos módulos cada fluxo é processado por um *thread* e, dentro de cada um destes *threads*, as tarefas são executadas de uma forma concorrente dando origem a novos *threads*. O uso de *Windows Sockets* 2.0 e *Microsoft SDK* permite que o módulo de *probing* opere em modo promíscuo, e permite também a manipulação dos campos do cabeçalho dos pacotes em camadas inferiores (camada de rede e de ligação lógica). É de notar que o mesmo tipo de funcionalidades estão disponíveis para desenvolvimentos e implementações em *UNIX*.

Nas próximas secções vão ser descritos com maior detalhe os módulos gerador de tráfego e de *probing*.

8.2 Módulo gerador de tráfego

Este módulo gera o tráfego de cada dasse de serviço ao nível do fluxo e do pacote. Para especificar o tráfego é necessário caracterizar os instantes de chegada e as durações dos fluxos; para além disso, é também necessário caracterizar a forma como são gerados os pacotes de um fluxo quando esse fluxo é aceite na rede, isto é, os instantes de chegada e os comprimentos dos pacotes. O módulo permite a geração de fluxos de acordo com um processo de *Poisson* e com durações exponencialmente distribuídas. A implementação do gerador de tráfego possibilita a escolha entre diversos modelos para o processo de chegadas de pacotes e para o comprimento dos pacotes. Estes podem ser CBR (*Constant Bit Rate*) ou *on-off* com durações *on* e *off* exponenciais ou de Pareto. As fontes de tráfego CBR são apenas caracterizadas pela taxa de chegada de pacotes. As fontes *on-off* requerem que, além da taxa de transmissão no tempo *on*, sejam também especificados os tempos médios de permanência nos estados *on* e *off*. A distribuição de Pareto requer ainda um parâmetro adicional denominado de parâmetro de forma (*shape*). O comprimento dos pacotes pode também ser fixo, exponencial ou de Pareto. A Figura 8-3 apresenta a janela de configuração do gerador de tráfego. Além dos parâmetros de tráfego descritos anteriormente, são também configurados nesta janela o endereço IP e o endereço do porto do módulo de *probing*, e o endereço IP destino do fluxo.

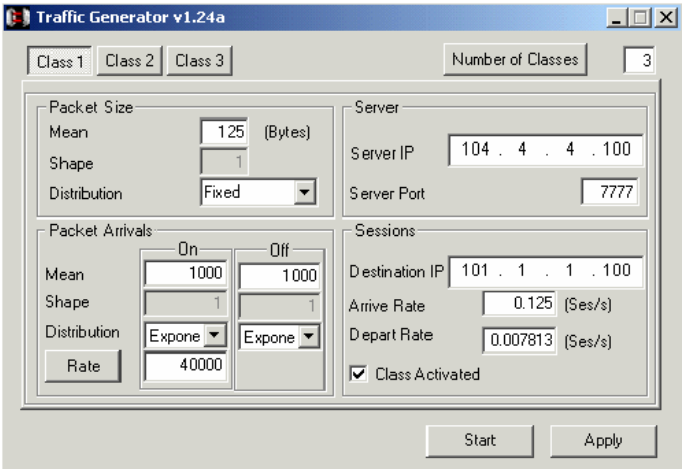


Figura 8-3 : Janela de configuração do gerador de tráfego.

O gerador de tráfego recorre a dois tipos de *sockets*: um *socket* TCP para a troca de informação de controle com o módulo de *probing*, e um *socket* UDP para transmissão de

dados. Existe também um *thread* por classe de serviço que agenda o instante de chegada do próximo fluxo e determina a sua duração. Cada vez que chega um novo fluxo, é activado um *thread* que fica responsável pela geração dos pacotes desse fluxo e pelas mensagens de controle trocadas entre o gerador de tráfego e o módulo de *probing*.

8.3 Módulo de *probing*

O módulo de *probing* é responsável por implementar o processo de *probing* e por efectuar a marcação dos pacotes. Como foi referido anteriormente, o módulo de *probing* escuta de forma promíscua todos os pacotes que foram injectados na sua rede local. Este módulo é implementado usando *raw sockets*, pois estes permitem a manipulação dos campos do cabeçalho do pacote IP. No lado da entrada da rede DiffServ, o módulo de *probing* captura os pacotes de dados e injecta-os de novo na rede local após lhes ter modificado o campo de ToS (de acordo com a sua classe de serviço, a qual foi indicada ao módulo de *probing* através da mensagem *REQUEST*) e de controle de erros do cabeçalho IP. Como o *Microsoft Windows* 2000 não suporta nativamente a manipulação do *byte* ToS, foi desenvolvido um pequeno programa para o efeito. Para além dos *raw sockets*, o módulo de *probing* recorre também a *sockets* TCP para troca de informação com os outros módulos de *probing*, e *sockets* UDP para transmissão dos pacotes de dados, de *probes* e de ?-*probes*, e transmissão das mensagens de controle do processo de *probing*. Neste módulo existe um *thread* permanentemente à escuta de pedidos de estabelecimento de novos fluxos num porto específico. Quando o módulo de *probing* recebe um pedido do gerador de tráfego, o *thread* de escuta dá origem a um novo *thread* que fica responsável por esse fluxo. Para aumentar o desempenho do sistema, são utilizados *sockets* UDP assíncronos para evitar que se esteja permanentemente a verificar o estado do *socket*. Os *sockets* TCP usados na implementação são do tipo bloqueante. Neste caso, o sistema suspende a execução de outras tarefas sempre que o *socket* TCP está a executar uma operação.

A janela de configuração principal do módulo de *probing* é apresentada na Figura 8-4. Esta janela permite que sejam configurados vários parâmetros genéricos: o porto do servidor para comunicação com o gerador de tráfego, os endereços dos *gateways* no sentido da rede de acesso ou da rede DiffServ, o tempo de *probing* e o tempo máximo (*timeout*) que o módulo espera pela mensagem *STATISTICS* após ter enviado a mensagem *PROBE STOP*, e a capacidade da ligação série. A capacidade da ligação é apenas

necessária para calcular alguns parâmetros de desempenho que são facultados na janela de estatísticas (Figura 8-6).

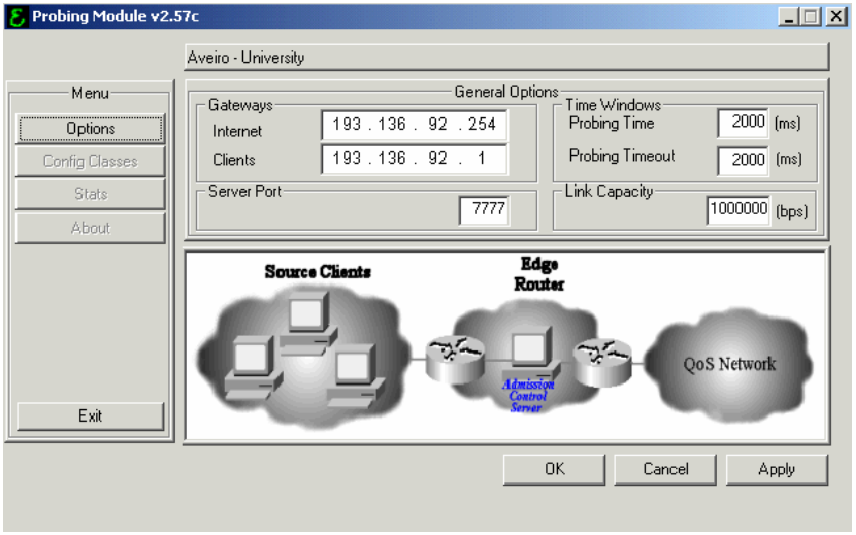


Figura 8-4 : Janela de configuração do módulo de *probing.*

A Figura 8-5 apresenta a janela de configuração do tráfego de *probing* (*Config Classes*), onde são configurados os parâmetros e os modelos de tráfego usados na geração dos fluxos de *probing* e de ?-*probing*. Existe uma janela de configuração para cada classe de serviço. A janela inclui também o limiar de perdas de ambos os fluxos, em cada classe, e uma opção para desactivar os fluxos de ?-*probing*. Note-se que, em princípio, as características das fontes de tráfego dos fluxos de dados e de *probing* (não considerando os de ?-*probing*) devem ser as mesmas. Embora os parâmetros dos fluxos de *probing* pudessem ser enviados directamente pelo gerador de tráfego, evitando a necessidade de os configurar no módulo de *probing*, a sua configuração neste módulo permitiu uma maior flexibilidade nos testes efectuados. Note-se que no exemplo apresentado na Figura 8-5, a largura de banda dos fluxos de *probing* é igual à dos fluxos de dados (apresentada na Figura 8-3) e que a largura de banda dos fluxos de ?-*probing* é menor que a dos fluxos de dados.

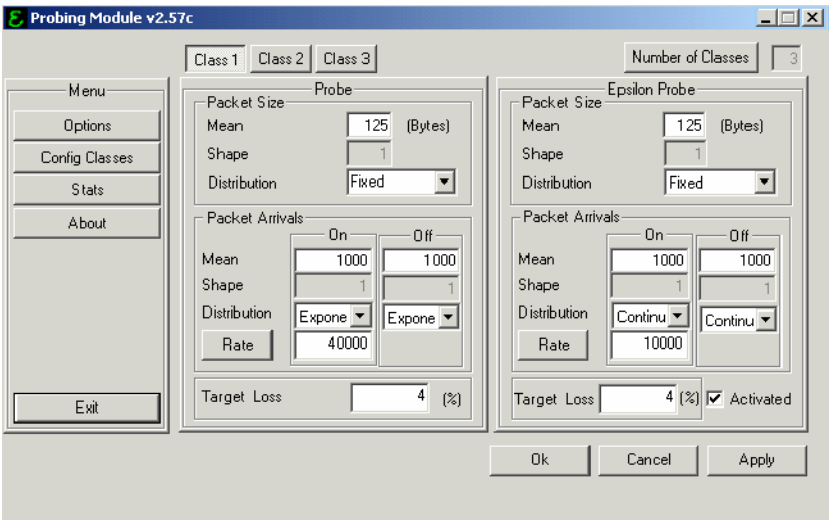


Figura 8-5 : Janela de configuração do tráfego de *probing* no módulo de *probing*.

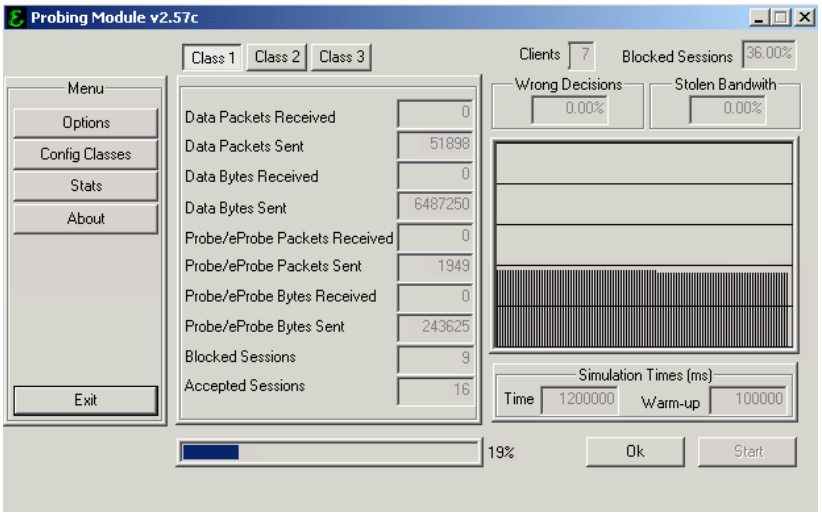


Figura 8-6 : Janela de estatísticas do módulo de *probing*.

A Figura 8-6 apresenta a janela de estatísticas do módulo de *probing* (*Stats*). Existe uma janela de estatísticas para cada classe de serviço. Cada uma inclui as estatísticas de números de pacotes de dados, de *probes* e ?-*probes* enviados e recebidos, o número de fluxos aceites e rejeitados, o número total de fluxos activos no sistema, e as percentagens de decisões erradas e de largura de banda roubada quando são inseridas perturbações no sistema. Nesta janela pode observar-se também a curva de evolução da probabilidade de

bloqueio ao longo do tempo. Todos os parâmetros apresentados são actualizados em tempo real. É ainda possível configurar nesta janela a duração total da experiência e o tempo de aquecimento para armazenamento das estatísticas (tempo necessário para estabilização do sistema).

8.4 Cenário experimental

As experiências a realizar consideram a existência de uma ou de duas classes de serviço. Todos os conjuntos de experiências recorrem ao cenário experimental ilustrado na Figura 8-7. Cada utilizador A e B gera o tráfego total de uma das classes de serviço. Nas experiências com apenas uma classe, só os utilizadores A são usados. A nível de equipamento, os utilizadores A e B são PCs *Pentium* a 120 MHz com 64 *Mbytes* de RAM. Devido a questões de desempenho, são usados dois servidores de *probing* no lado de entrada da rede DiffServ. O servidor de *probing* A é um *Pentium* a 350 MHz com 128 *Mbytes* de RAM, e o servidor B é um *Pentium* a 733 MHz com 256 *Mbytes* de RAM. O servidor no lado de saída da rede DiffServ é um *Pentium* a 933 MHz com 256 *Mbytes* de RAM. O sistema operativo dos utilizadores origem e destino A e B é o *Windows* NT 4.0, e o dos servidores de *probing* é o *Windows* 2000 *Professional*.

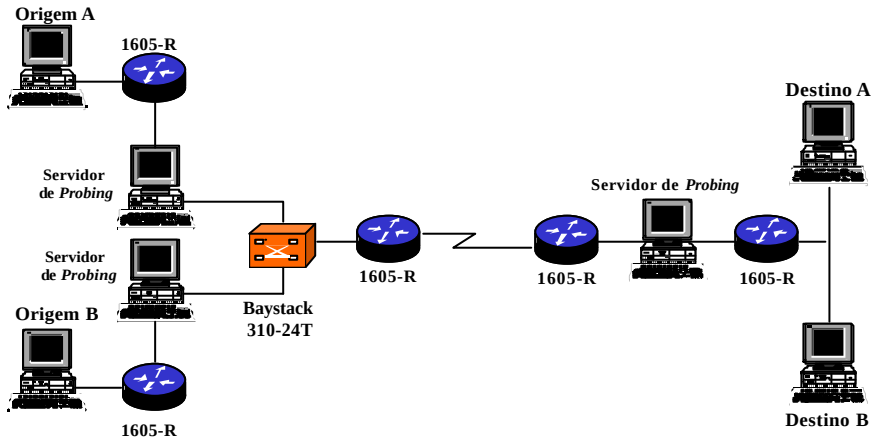


Figura 8-7 : Cenário experimental.

Todos os *routers* utilizados nas experiências são *Cisco* 1605 R que correm a versão de sistema operativo IOS 12.0(7)T. Os *routers* de entrada e de saída da rede DiffServ são

ligados por uma ligação série por forma a permitir um controle flexível da capacidade da ligação. Em experiências que consideram apenas uma classe de serviço, estes *routers* são configurados com a disciplina de escalonamento de prioridade estrita (denominada de *Priority Queuing* nos *routers* da *Cisco*) com dois níveis de prioridade. A prioridade mais elevada é atribuída às mensagens de controle. As experiências que consideram duas classes recorrem ao *Custom Queuing* da *Cisco*. Este mecanismo mantém um máximo de 16 filas de espera que são divididas em dois grupos, em que o primeiro usa prioridade estrita e o segundo usa o algoritmo *Deficit Round Robin* apresentado no capítulo 3 (secção 3.3.3); este segundo grupo tem uma prioridade inferior. No caso específico das experiências realizadas, uma fila de espera é configurada com mecanismo de prioridade estrita (para o tráfego de controle) e duas filas de espera são configuradas com *Deficit Round Robin* (para o tráfego de dados e de *probing* de cada classe). A gestão das filas de espera é realizada de uma forma dinâmica, ou seja, o comprimento das filas de espera depende em cada instante da sua ocupação. A classificação dos pacotes nestes *routers* é baseada na análise dos *bits* de precedência e nos *bits* de ToS e recorre às listas de acesso da *Cisco*. É usado um comutador *Ethernet* (*Baystack* 310-24T) para multiplexar o tráfego vindo dos utilizadores A e B do lado de entrada da rede DiffServ.

8.5 Resultados experimentais e discussão

Nesta secção são apresentados e discutidos três conjuntos de experiências semelhantes aos apresentados no capítulo 7. O primeiro conjunto considera apenas uma classe de serviço. O segundo considera duas classes de serviço em que o tráfego médio oferecido a cada classe é constante. O último conjunto considera também duas classes de serviço mas com tráfego médio oferecido numa das classes variável ao longo do tempo.

Os parâmetros utilizados nestas experiências diferem dos utilizados nos estudos de simulação descritos no capítulo 7. Uma das diferenças centra-se na capacidade da ligação. Nos estudos de simulação esta capacidade era de 10 Mb/seg (e de 45 Mb/seg numa das experiências realizadas). Nas experiências laboratoriais, a capacidade máxima da ligação é de 1 Mb/seg, devido a restrições impostas pelo *hardware* disponível. A diminuição da capacidade da ligação impõe uma diminuição na largura de banda média das fontes de tráfego utilizadas e no número médio de fluxos activos no sistema, em relação aos valores utilizados nos estudos de simulação do capítulo 7.

As fontes de tráfego (e respectivos parâmetros) utilizadas nas experiências são CBR (*Constant Bit Rate*) e *on-off* exponenciais e estão descritas na Tabela 8-1.

Fonte de Tráfego	Taxa de Pico	Tempo *On*	Tempo *Off*	Taxa Média
CBR_40	40 Kb/seg	_	_	40 Kb/seg
CBR_64	64 Kb/seg	_	_	64 Kb/seg
CBR_100	100 Kb/seg	_	_	100 Kb/seg
EXP_128_50	128 Kb/seg	500 mseg	500 mseg	64 Kb/seg

Tabela 8-1 : Características das fontes de tráfego.

Assim como nos estudos de simulação, os fluxos são gerados de acordo com um processo de *Poisson* com uma taxa média de chegadas ? e têm uma duração média no sistema de 1/?. A intensidade de tráfego é ? ?= ?/?. O comprimento dos pacotes é de 125 *bytes*, sempre que nada for dito em contrário. Nos *routers* utilizados nas experiências, o comprimento máximo do conjunto das filas de espera que contêm os pacotes de dados, *probes* e ?-*probes*, é de 24,000 *bytes*.

A duração de cada experiência é de 1,200 segundos. Considera-se que o tempo de aquecimento é o dobro da maior das durações médias dos fluxos. Cada experiência descrita de seguida é repetida 3 vezes e os resultados apresentados correspondem à média dos valores obtidos.

Por forma a controlar os resultados experimentais, este cenário experimental foi simulado através do simulador de redes *ns*-2 (*network simulator*), utilizando parâmetros semelhantes aos das experiências laboratoriais. A forma como são geridas as filas de espera é diferente nos estudos de simulação e nos *routers* utilizados. Nos estudos de simulação é atribuído um comprimento fixo para o conjunto das filas de espera. Nos *routers* utilizados, o comprimento do conjunto das filas de espera é dinâmico e dependente da sua ocupação.

No sistema experimental implementado, o receptor tem apenas informação sobre os pacotes recebidos, e envia a mensagem *STATISTICS* para o emissor apenas quando o tempo de *probing* termina. Para fazer corresponder esta implementação aos estudos de simulação, o simulador é modificado para terminar o envio de *probes* apenas no final do tempo de *probing* e para só então ser tomada a decisão de controle de admissão.

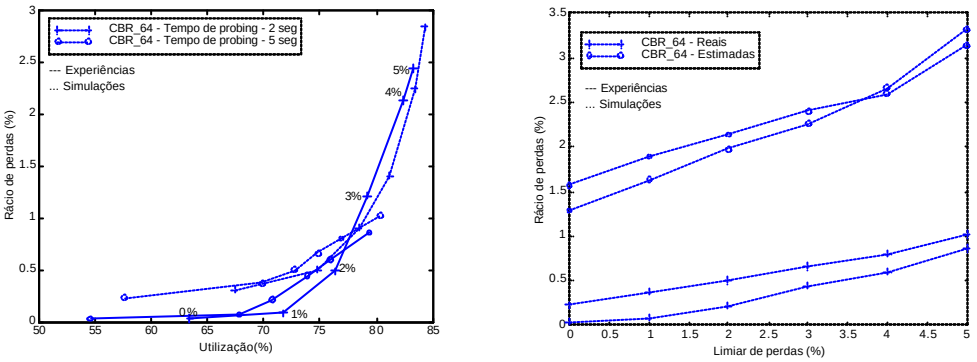
8.5.1 Experiências com uma classe

Nestas experiências é averiguado, para valores diferentes de limiar de perdas, o desempenho do mecanismo de *probing* com a variação do tempo de *probing*, da largura de banda dos fluxos de dados e do modelo das fontes de tráfego utilizado. As figuras apresentadas de seguida mostram a variação do rácio de perdas em função da utilização para valores diferentes do limiar de perdas. A utilização é baseada apenas em pacotes de dados, sendo o tráfego de *probing* considerado tráfego adicional (que diminui a utilização do sistema). Todas as curvas apresentadas nesta secção têm seis pontos em que cada um corresponde a um limiar de perdas diferente. Da esquerda para a direita dos gráficos, os valores de limiar de perdas são de 0%, 1%, 2%, 3%, 4% e 5%. A capacidade da ligação é de 800 Kb/seg. Este valor de capacidade é escolhido tendo em conta as limitações do cenário experimental implementado.

Os fluxos de dados utilizados nestas experiências são CBR_64, CBR_100 e EXP_128_50. Para fluxos CBR_64 e EXP_128_50, ? ?= ??; para fluxos CBR_100, ? ?= ???

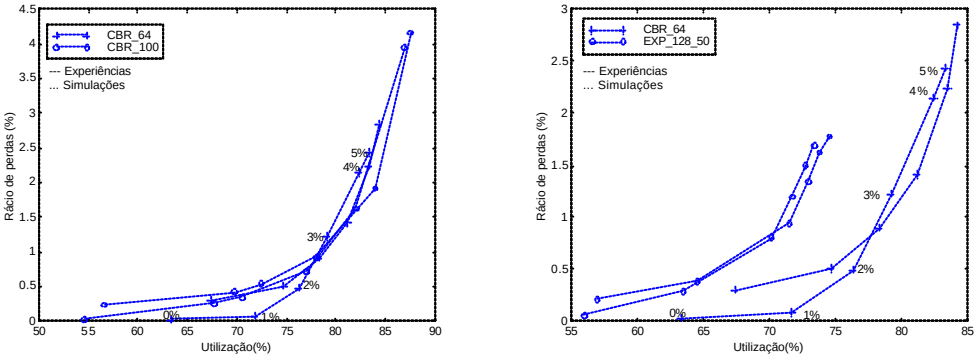
De seguida são apresentados os resultados das experiências realizadas com uma classe de serviço e os correspondentes resultados obtidos por simulação do mesmo sistema. A Figura 8-8 (a) ilustra o rácio de perdas reais e o respectivo valor de utilização do sistema para valores diferentes do limiar de perdas, e para 2 e 5 segundos de tempo de *probing*, e a Figura 8-8 (b) apresenta o rácio de perdas reais e estimadas em função do limiar de perdas, para um tempo de *probing* de 5 segundos. ? Figura 8-9 (a e b) apresenta as curvas de perdas reais em função da utilização para diferentes fluxos de dados: (a) CBR_64 e CBR_100, e (b) CBR_64 e EXP_128_50. Os resultados observados estão de acordo com os obtidos nas experiências com uma classe de serviço apresentadas no capítulo 7.

As figuras apresentam os resultados das experiências laboratoriais e das correspondentes simulações. Pela análise comparativa das curvas verifica-se que os resultados são semelhantes, mas que o rácio de perdas reais e a utilização são ligeiramente menores nas experiências laboratoriais em relação às simulações. Verificou-se nas experiências laboratoriais que ocorreram sempre algumas perdas de pacotes de controle. Nas situações em que os pacotes de controle não chegam ao emissor dentro do *timeout*, os fluxos são rejeitados. Este facto contribui para diminuir o número de fluxos activos no sistema, e consequentemente, diminuir a utilização e o rácio de perdas reais.



(a) **(b)**

Figura 8-8 : Desempenho do mecanismo de *probing* em função do tempo de *probing*: (a) rácio de perdas reais *vs* utilização, (b) rácio de perdas reais e estimadas em função do limiar.



(a) **(b)**

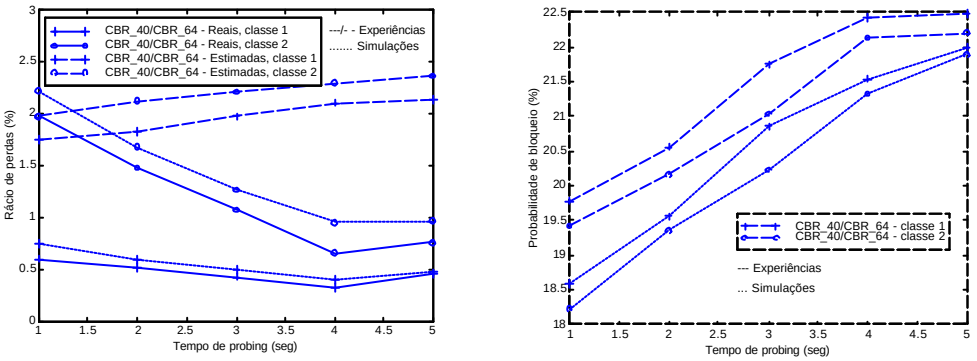
Figura 8-9 : Rácio de perdas em função da utilização: (a) fontes CBR e (b) fontes CBR e exponencial.

8.5.2 Experiências com duas classes e tráfego oferecido constante

Este conjunto de experiências aborda duas classes de serviço e um tráfego oferecido constante, isto é, a taxa de chegada e a duração média dos fluxos no sistema não variam ao longo da experiência. Em todas as experiências o limiar de perdas para os fluxos de *probing* e de ?-*probing* é de 5%. Os fluxos da classe 1 são CBR_40 e os da classe 2 são CBR_64. Os pesos associados às classes 1 e 2 são, respectivamente, de 20% e 80%. Estes pesos são diferentes dos usados nos estudos de simulação apresentados no capítulo 7 devido a restrições impostas pelo *hardware* disponível.

De seguida são apresentados os resultados obtidos em duas experiências semelhantes às apresentadas no capítulo 7. Na primeira, avalia-se o desempenho do mecanismo de ?-*probing* com a variação do tempo de *probing*, considerando tráfego oferecido proporcional aos pesos do escalonador CBQ; na segunda experiência são analisados os mecanismos de *probing* e de ?-*probing* numa situação em que existe desadaptação entre o tráfego oferecido e os pesos do CBQ. Na primeira experiência, $?_1 = 4$ e $?_2 = 11$, e a largura de banda dos fluxos de ?-*probing* em cada classe é de 20 Kb/seg. Na segunda experiência, o tráfego oferecido (definido no capítulo 7 por $b_k ?_k$) à classe 1 é aumentado de 20% para 50% da capacidade da ligação (através do aumento de $?_1$ de 4 para 11). O tráfego da classe 2 não é alterado.

A Figura 8-10 (a) apresenta as curvas de rácio de perdas reais e estimadas em cada classe em função do tempo de *probing*, considerando tráfego médio oferecido proporcional aos pesos do CBQ. A Figura 8-10 (b) apresenta as curvas correspondentes de probabilidade de bloqueio. Na Figura 8-11 são apresentadas as curvas de perdas reais e estimadas em função da largura de banda dos fluxos de ?-*probing*, considerando um desajuste entre o tráfego médio oferecido e os pesos do CBQ. Com o objectivo de simplificar a figura, os resultados de simulação apresentados consideram apenas as perdas reais. Os resultados estão, uma vez mais, de acordo com os apresentados no capítulo 7.



(a) **(b)**

Figura 8-10 : Desempenho do mecanismo de ?-*probing* em função do tempo de *probing*: (a) rácio de perdas *vs* tempo de *probing*, (b) probabilidade de bloqueio vs tempo de *probing*.

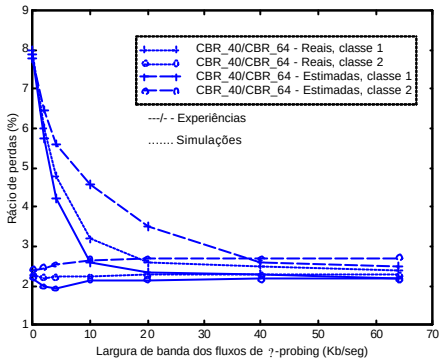


Figura 8-11 : Rácio de perdas em função da largura de banda dos fluxos de ?-*probing*, com tráfego oferecido desajustado com os pesos do CBQ.

As curvas dos resultados experimentais apresentam o mesmo comportamento que as curvas de simulação. Uma vez mais, os valores de rácio de perdas são superiores nas simulações (e os valores de bloqueio inferiores). A explicação para este comportamento é a mesma que a referida nas experiências com uma classe de serviço.

8.5.3 Experiências com duas classes e tráfego oferecido variável no tempo

Neste conjunto de experiências é considerado um tráfego oferecido variável no tempo numa das classes. Em particular, a intensidade do tráfego da classe 1 vai ser aumentada durante a experiência após o tempo de aquecimento, isto é, no instante em que começam a ser armazenados os dados necessários para o cálculo das estatísticas de interesse. Antes da perturbação, o tráfego oferecido à classe 1 representa 20% da capacidade da ligação ($?_1$ = 4). A classe 2 tem tráfego oferecido constante ao longo da experiência, ocupando 80% da capacidade da ligação ($?_2$ = 11). A cada classe vai ser associado no escalonador CBQ um peso de 50%. Por isso, antes da perturbação, a classe 1 tem tráfego oferecido inferior à largura de banda garantida e o oposto acontece na classe 2. Os fluxos de dados em ambas as classes são CBR_64. Nas experiências realizadas, são utilizados dois geradores para o tráfego oferecido à classe 1. Ambos os geradores geram tráfego oferecido constante, mas o segundo só é activado após o tempo de aquecimento. O tempo de *probing* considerado é de 2 segundos e o limiar de perdas é de 5%.

Para analisar os resultados das experiências são utilizadas as mesmas métricas de desempenho que foram consideradas nos estudos de simulação do capítulo 7 (secção 7.5.2.2): a percentagem de decisões erradas e de largura de banda roubada.

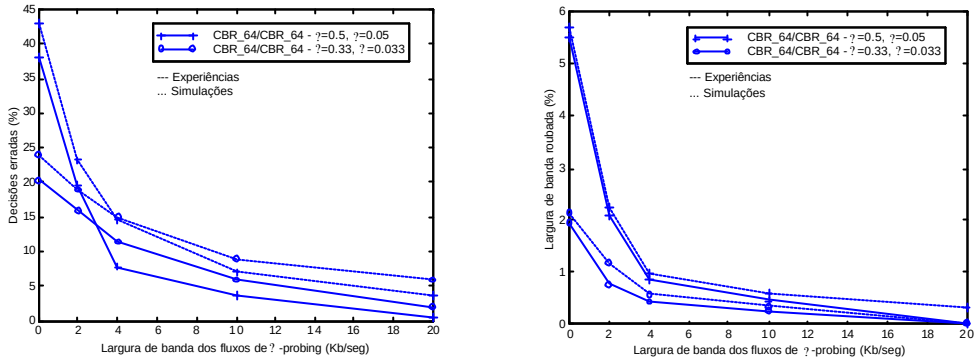
Dado que se pretende analisar o comportamento transitório do sistema, isto é, quando a perturbação é inserida, a duração da experiência é de apenas 200 segundos, contados após o instante em que termina o aquecimento, para que os valores das métricas de desempenho reflictam a situação de roubo de recursos despoletada pelo transitório. Note-se que se considera apenas uma perturbação em cada experiência, isto é, apenas existe uma variação do tráfego oferecido à classe 1. Esta situação é diferente da considerada nos estudos de simulação efectuados no capítulo 7, em que se considerou que o tráfego oferecido à classe 1 apresentava variações periódicas, entre 20% e 50% da capacidade da ligação.

Nesta secção vão ser consideradas duas experiências diferentes. Na primeira é analisado o desempenho do mecanismo de ?-*probing* para diferentes perturbações, e a segunda avalia o impacto da variação do comprimento dos ?-*probes* na reacção à perturbação inserida. Os dois casos de perturbações considerados na primeira experiência apresentam o mesmo ? ? 10, mas a taxa de chegadas do segundo gerador é de ? = 0.5 seg^{-1} no primeiro caso e de ? = 0.33 seg^{-1} no segundo caso. Também foram efectuadas experiências com uma taxa de chegadas de fluxos de ? = 1 seg^{-1}, mas foi atingida a situação de *thrashing*, porque neste caso muitos fluxos pedem admissão ao mesmo tempo, e muito poucos acabam por ser aceites. Na segunda experiência, a taxa de chegada das perturbações é de ? = 0.5 seg^{-1} com ??= 10. Os comprimentos dos ?-*probes* considerados são 125, 200 e 75 *bytes*.

A Figura 8-12 (a e b) e a Figura 8-13 (a e b) apresentam a percentagem de decisões erradas e a percentagem de largura de banda roubada, respectivamente na primeira e segunda experiências.

A Figura 8-12 (a e b) mostra que no mecanismo de *probing* simples (largura de banda nula dos fluxos de ?-*probing*) a percentagem de decisões erradas e de largura de banda roubada são muito elevadas (as decisões erradas são de 38% com a primeira perturbação e de 20% com a segunda; a largura de banda roubada é superior a 5% na primeira perturbação e de quase 2% na segunda). Estes resultados estão também de acordo com os observados no capítulo 7. Verifica-se que ambas as métricas decrescem rapidamente com a largura de banda dos fluxos de ?-*probing*: com 2 Kb/seg apenas, a largura de banda roubada decresce quase para metade, e com 10 Kb/seg (menos de 1/6 da largura de banda dos fluxos admitidos) o roubo de recursos é insignificante. A comparação

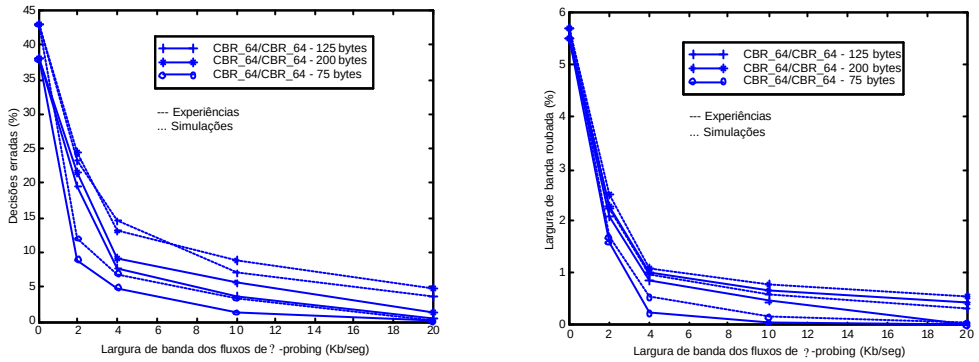
das duas curvas mostra também que uma maior taxa de chegadas na perturbação conduz a um maior roubo de recursos. Os resultados desta experiência em que o roubo dos recursos é aumentado de uma forma intencional, mostram claramente que o mecanismo de ?-*probing* é capaz de atenuar este problema.



(a) **(b)**

Figura 8-12 : Desempenho do mecanismo de ?-*probing* face a perturbações no sistema em função da taxa de chegada dos fluxos de perturbação: (a) percentagem de decisões erradas e (b) percentagem de largura de banda roubada.

Os resultados observados na Figura 8-13 (a e b) estão também de acordo com os apresentados no capítulo 7. Em particular, verifica-se que a redução das percentagens de decisões erradas e de largura de banda roubada com o aumento da largura de banda dos fluxos de ?-*probing* é mais pronunciada quando o comprimento dos ?-*probes* é de 75 *bytes*, ou seja, quando o comprimento dos ?-*probes* é ligeiramente inferior ao dos dados. Para fluxos de ?-*probing* de 4 Kb/seg (1/16 dos fluxos de dados), a largura de banda roubada é já insignificante. É de notar que, como foi referido no capítulo 7, as perdas medidas por pacotes mais pequenos vão ser menores que as medidas pelos dados. Este efeito não é ainda notório para 75 *bytes* porque, com perturbações elevadas, as perdas também são muito elevadas.



(a) **(b)**

Figura 8-13 : Desempenho do mecanismo de ?-*probing* face a perturbações no sistema em função do comprimento dos ?-*probes*: (a) percentagem de decisões erradas e (b) percentagem de largura de banda roubada.

As curvas dos resultados experimentais apresentam o mesmo comportamento que as curvas de simulação, sendo os valores destas últimas superiores. Os factores que influenciam estas diferenças de resultados são os mesmos que os apresentados nas secções anteriores.

8.6 Conclusões e trabalho futuro

Neste capítulo foi apresentada a implementação de uma rede DiffServ com suporte de QdS e controle de admissão de fluxos através de mecanismos de *probing* e de ?-*probing*. Para o efeito foram desenvolvidos dois módulos: módulo de geração de tráfego e módulo de *probing*. Para efectuar a coordenação entre os dois módulos foi definido um protocolo de comunicação simples.

Sobre este protótipo laboratorial foram efectuadas algumas experiências semelhantes às realizadas por simulação no capítulo 7. Os resultados das experiências confirmam que os mecanismos de *probing* são capazes de efectuar correctamente o controle de admissão ao mesmo tempo que permitem atingir utilizações elevadas. Além disso, mostram também que em ambientes com múltiplas classes, o mecanismo de ?-*probing* consegue atenuar o problema do roubo de recursos entre classes. Este mecanismo é então uma solução eficiente para suportar QdS por fluxo sem roubo de recursos e sem necessidade de manter qualquer estado por fluxo nos *routers* da rede. Em particular, nas experiências realizadas,

verificou-se que fluxos de ?-*probing* com larguras de banda inferiores a 1/6 da largura de banda dos fluxos de dados são capazes de atenuar significativamente o problema do roubo de recursos. Comparando os resultados de simulação com os resultados experimentais verifica-se que, nestes últimos, tanto a utilização como o rácio de perdas, como o roubo de recursos diminuem ligeiramente.

No sistema desenvolvido apenas foi implementado o mecanismo de *probing* baseado em perdas, em que o tráfego de *probing* tem o mesmo nível de prioridade do tráfego de dados. A implementação de outros mecanismos, por exemplo, com *probing* num nível inferior, e considerando marcação de pacotes ou atraso de pacotes em vez de perdas, é um tópico em aberto para trabalho futuro.

CAPÍTULO 9

AGREGAÇÃO EM DOMÍNIOS HIERÁRQUICOS

Na arquitectura IntServ (Integração de Serviços) os recursos são reservados para fluxos individuais. Esta reserva de recursos pode ser efectuada através, por exemplo, do RSVP (*resource ReSerVation Protocol*). Nesta arquitectura, reservar recursos fluxo-a-fluxo implica a troca de mensagens de sinalização entre os vários elementos da rede (terminais e *routers*) que se encontram no percurso do fluxo, para determinar se existem recursos disponíveis nestes elementos; além disso, é necessário manter o estado de cada fluxo nos elementos da rede. Estes dois factores contribuem para a falta de escalabilidade normalmente atribuída à arquitectura IntServ.

A reserva de recursos para agregados de fluxos tem sido proposta em diversos trabalhos, [Baker01, Pan00, Schelèn98, Schmitt99, Fu01], como forma de reduzir significativamente a carga de sinalização e a quantidade de informação do estado das reservas armazenada nos *routers*, mas mantendo o suporte da QdS (Qualidade de Serviço) requerida por cada fluxo.

Para implementar a agregação de reservas, foi definida recentemente uma extensão ao RSVP [Baker01]. Esta extensão, descrita no capítulo 3, permite reservar (*à priori*) uma quantidade de largura de banda para um agregado de fluxos entre dois elementos da rede, o *router* de entrada e o *router* de saída do agregado. Esta reserva pode ser actualizada

dinamicamente em quantidades de largura de banda, que designaremos por *bulks*, que são normalmente bastante superiores à largura de banda dos fluxos. Sempre que um fluxo pede admissão, o *router* de entrada do agregado verifica se existe largura de banda disponível no agregado para aceitar o fluxo. Se existirem recursos disponíveis, o fluxo é automaticamente aceite pelo *router* de entrada, sem que os *routers* interiores ao agregado (situados entre o *router* de entrada e de saída do agregado) sejam sinalizados. Nesta situação, o *router* de entrada altera o campo de protocolo da mensagem RSVP para RSVP-E2E-IGNORE, para que esta mensagem seja escondida dos *routers* interiores. A mensagem RSVP passa então transparentemente nestes *routers* sem provocar qualquer processamento. Caso não existam recursos disponíveis no agregado, o *router* de entrada sinaliza os *routers* interiores (ou seja, não esconde as mensagens RSVP destes *routers*) para verificar se estes podem acomodar uma largura de banda adicional igual ao *bulk*. Se esta tentativa é bem sucedida, o fluxo é admitido e a largura de banda do agregado é actualizada; caso contrário, o fluxo é rejeitado. Assim, com agregação, as mensagens de sinalização são processadas nos *routers* interiores apenas quando a largura de banda do agregado necessita de uma actualização. Os *routers* de entrada do agregado processam mensagens de sinalização fluxo-a-fluxo, pois necessitam de verificar a largura de banda disponível no agregado sempre que um fluxo pede admissão.

A eficiência da agregação depende da forma como estão adaptados o tráfego admitido e a reserva do agregado que o suporta. Se o tamanho do *bulk* é demasiado grande, a carga de sinalização será mínima, mas os recursos da rede podem ser sub-utilizados. Caso contrário, com um *bulk* demasiado pequeno, a largura de banda reservada para o agregado de fluxos aproxima-se da largura de banda do tráfego admitido, mas a carga de sinalização torna-se elevada e semelhante à existente em reservas de recursos fluxo-a-fluxo.

Considerando um domínio de rede com agregação, a carga de sinalização no conjunto dos elementos de rede do domínio aumenta com o aumento do número de elementos que necessitam de processar mensagens de sinalização fluxo-a-fluxo. No caso em que os agregados são configurados entre os *routers* fronteira do domínio, denominados de agregados extremo-a-extremo, apenas os *routers* de entrada do domínio processam mensagens de sinalização fluxo-a-fluxo. No entanto, num domínio de rede de dimensão elevada, a necessidade de estabelecer um número elevado de agregados extremo-a-extremo

pode reduzir significativamente a utilização dos recursos da rede. A divisão do domínio em áreas e a configuração de agregados entre os *routers* fronteira das áreas permite atenuar o problema da utilização, mas provoca um aumento na carga de sinalização do domínio. Esta hierarquização do domínio permite aumentar a utilização dos recursos reservados, porque os agregados numa área podem ser partilhados por um maior número de fluxos. No entanto, sempre que uma mensagem de pedido de admissão de um novo fluxo chega a um *router* fronteira de um domínio, existe a necessidade de verificar se existem recursos suficientes em todos os agregados no percurso do fluxo entre os *routers* de entrada e de saída desse domínio. Note-se que o caso extremo de hierarquização de domínios corresponde ao caso em que são configurados agregados entre cada par de *routers*. Este caso é equivalente ao de sinalização fluxo-a-fluxo.

A facilidade de dividir um domínio em áreas está também incluída em diversos protocolos de encaminhamento, como por exemplo o OSPF (*Open Shortest Path First*) [Katz01] e o ISIS (*Intermediate System to Intermediate System Intra Domain Routing Exchange Protocol*) [Smit01], e ainda no MPLS (*Multi-Protocol Label Switching*) [Kompella01]. A implementação da hierarquização de domínios nestes protocolos é também motivada por questões de escalabilidade, que estão neste caso relacionadas com o tamanho das tabelas de encaminhamento.

O objectivo do trabalho descrito neste capítulo é analisar os compromissos entre carga de sinalização e utilização de recursos num domínio de rede que pode ser dividido em áreas. Para atingir este objectivo são desenvolvidos dois modelos analíticos que se complementam [Sargento02b, Sargento02c]. No primeiro modelo, denominado de modelo de carga por fluxo, a carga oferecida é detalhada ao nível do fluxo. O modelo assume que os fluxos chegam de acordo com um processo de *Poisson* e que têm durações exponencialmente distribuídas. Deste modo, o número de fluxos activos num domínio pode ser descrito através de um processo de nascimento e morte multi-dimensional. Note-se que, embora o processo de *Poisson* possa não ser apropriado para o tráfego ao nível do pacote, ele é largamente usado para tráfego ao nível do fluxo, dado que os fluxos são normalmente gerados por um número elevado de utilizadores independentes. O segundo modelo analítico, denominado de modelo de carga por agregado, considera que o tráfego oferecido tem uma largura de banda agregada média que é variável no tempo. Mais especificamente, assume-se neste capítulo que a variação é sinusoidal. Neste modelo o

tráfego não é detalhado ao nível do fluxo. Os dois modelos analíticos distinguem-se quanto a dois aspectos principais: (i) a carga oferecida é invariante (modelo de carga por fluxo) ou variante no tempo (modelo de carga por agregado); e (ii) as actualizações dos agregados são efectuadas em instantes arbitrários (modelo de carga por fluxo), ou no início de intervalos de duração fixa (modelo de carga por agregado). No primeiro modelo, as actualizações dos agregados são efectuadas aquando da chegada de um fluxo, em múltiplos do *bulk*, sempre que necessário e se existir largura de banda suficiente para efectuar as actualizações. No segundo modelo, como o tráfego oferecido não é detalhado ao nível do fluxo, são pré-definidos instantes de actualização dos agregados; toma-se ainda como aproximação que a actualização é feita exactamente de acordo com as necessidades de largura de banda no intervalo de tempo seguinte.

O primeiro modelo permite efectuar um estudo detalhado da carga de sinalização, enquanto que o segundo modelo permite analisar os compromissos entre a taxa de variação do tráfego oferecido e a taxa de actualizações do agregado. Como complemento aos dois modelos analíticos são efectuados estudos de simulação que consideram agregados de fluxos reais. Este estudo é efectuado através de simulação de eventos discretos.

Este capítulo está organizado da seguinte forma. A secção 9.1 apresenta o modelo do sistema e algumas definições necessárias ao desenvolvimento dos modelos analíticos. As secções 9.2 e 9.3 apresentam os dois modelos analíticos: o modelo de carga por fluxo e o modelo de carga por agregado. Na secção 9.4 são apresentados os resultados numéricos obtidos com os modelos e os resultados de simulação de eventos discretos que consideram agregados reais. A secção 9.5 conclui o capítulo e introduz algumas áreas de trabalho futuro.

9.1 Modelo do sistema

Considera-se um domínio de rede dividido em áreas como se ilustra na Figura 9-1. Os *routers* na fronteira do domínio são denominados de DBRs (*Domain Border Routers*), e os *routers* na fronteira de cada uma das áreas são denominados de ABRs (*Area Border Routers*). Os DBRs de uma rede dividida em áreas incluem também as funções dos ABRs.

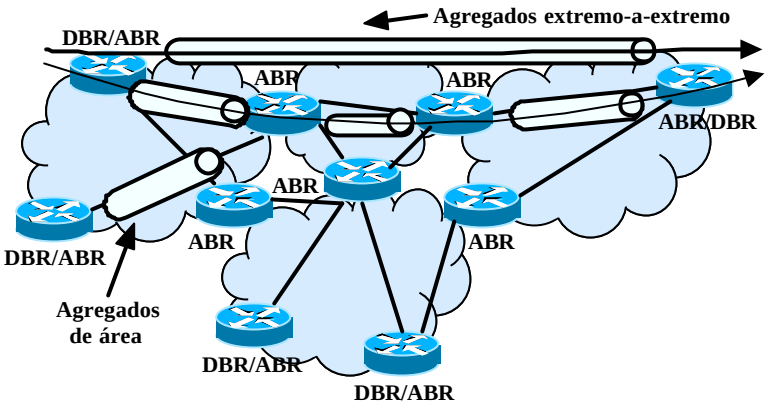


Figura 9-1: Modelo do sistema.

As sessões de fluxos de pacotes são oferecidas entre DBRs e podem ser agregadas em "tubos" de largura de banda reservada, denominados de agregados. São considerados dois casos diferentes relativamente à origem e destino dos agregados: (i) agregados extremo-a-extremo, em que a largura de banda é reservada extremo-a-extremo entre DBRs, ou (ii) agregados de área, em que a largura de banda é reservada extremo-a-extremo entre ABRs. No primeiro caso, os fluxos atravessam um único agregado extremo-a-extremo entre a entrada e a saída do domínio, enquanto que no segundo caso, os fluxos atravessam uma concatenação de agregados de área. Neste caso, as sessões de tráfego que partilham os mesmos ABRs de entrada e de saída das áreas podem ser agrupadas num mesmo agregado de área.

Os *routers* interiores de um agregado são os *routers* que se situam entre o *router* de entrada e o de saída do agregado. Assume-se que a largura de banda de um agregado pode ser ajustada ao longo do tempo através de sinalização apropriada nos *routers* que se encontram no percurso do agregado. O *router* de entrada de cada agregado processa mensagens de sinalização fluxo-a-fluxo, enquanto que os *routers* interiores processam mensagens de sinalização apenas quando é necessário ajustar a largura de banda do agregado.

No caso geral, os agregados atravessam uma ou mais áreas num domínio. Em cada área eles são configurados entre um par de ABRs (o ABR de entrada e de saída da área). Considere-se que ? ? = {1, 2, ..., *J*} é o conjunto dos pares de ABRs no domínio. Assume-se que cada par de ABRs é caracterizado por uma capacidade C_j, que corresponde à menor capacidade de entre todas as ligações do percurso do agregado. A largura de banda do

agregado pode variar ao longo do tempo, mas é limitada pela capacidade C_j dos pares de ABRs que atravessa.

Considere-se que ? = {1, 2, ..., *H*} é o conjunto de todos os agregados. Os agregados são definidos (i) pela sua largura de banda $r_h(t)$, (ii) pelos seus ABRs de origem e destino, (ii) pelo percurso ?$_h$? ? descrito através dos pares de ABRs que os agregados atravessam e (iv) pelo número de *routers* a_h atravessados pelos agregados.

Como foi referido acima, o tráfego é oferecido a um domínio entre DBRs de entrada e de saída. Considere-se que ? = {1, 2, ..., *K*} é o conjunto das sessões de tráfego. São considerados dois modelos de tráfego oferecido. No primeiro modelo, modelo de carga por fluxo, uma sessão *k* é caracterizada por (i) par de DBRs de entrada e saída, (ii) percurso ?$_k$? ? descrito em termos dos agregados que a sessão atravessa, (iii) largura de banda de cada fluxo de pacotes, b_k, e (iv) intensidade de tráfego ?$_k$ = ?$_k$/?$_k$. Neste modelo assume-se que os fluxos de pacotes chegam ao sistema de acordo com um processo de *Poisson* com uma taxa de chegadas média de ?$_k$ e que têm durações exponencialmente distribuídas com média 1/?$_k$. Considera-se também que a largura de banda do agregado pode ser ajustada em quantidades fixas de largura de banda, referenciadas por q_h. O segundo modelo, modelo de carga por agregado, não detalha o tráfego oferecido ao nível do fluxo. Uma sessão *k* é caracterizada simplesmente por (i) par de DBRs de entrada e de saída, (ii) percurso ?$_k$? ? descrito em termos dos pares de ABRs atravessados pela sessão e (iii) largura de banda agregada da sessão *k*, $r_k(t)$. Neste caso a sessão representa um agregado de fluxos. Considera-se que a largura de banda de um agregado é ajustada no início de intervalos de tempo fixos, correspondendo exactamente à largura de banda necessária para esse intervalo (isto é, ao valor máximo de $r_k(t)$ nesse intervalo).

Definem-se, adicionalmente, os seguintes conjuntos: o conjunto de sessões que atravessam o agregado *h* é representado por ?$_h$? ?, e o conjunto dos agregados que atravessam o par de ABRs *j* é representado por ?$_j$ = {1, 2, ..., H_j} ? ?.

Note-se que podem ser utilizados agregados diferentes para sessões com classes de serviço diferentes. Mais especificamente, podem ser consideradas as classes de serviço definidas na arquitectura DiffServ. Nos modelos seguintes considera-se que um conjunto de sessões é agrupado num mesmo agregado se pertencer à mesma classe de serviço. Por uma questão de facilidade de notação, a classe de serviço não é explicitamente representada.

9.2 Modelo de carga por fluxo

Nesta secção é desenvolvido um modelo analítico baseado em cadeias de *Markov* em tempo contínuo (mais especificamente, em processos de nascimento e morte multi-dimensionais) que permitem caracterizar o estado do sistema, assumindo que os fluxos chegam de acordo com um processo de *Poisson* e têm durações exponencialmente distribuídas. O estado do sistema é caracterizado por um vector $\mathbf{n} = (n_1, n_2, ..., n_K)$, em que n_k representa o número de fluxos no sistema pertencentes à sessão k. Os novos fluxos que pedem admissão no domínio podem ser aceites se existir largura de banda disponível em cada um dos agregados que atravessam; estes fluxos podem também ser aceites se a largura de banda de todos os agregados que não respeitam a condição anterior puder ser ajustada para poder comportar os novos fluxos. Considera-se que a largura de banda de um agregado h pode ser ajustada em múltiplos de uma quantidade fixa que se designa por *bulk* e se representa por q_h. Considera-se também que o *bulk* é idêntico em todos os agregados.

Define-se como métrica para avaliar a carga de sinalização e a quantidade de informação do estado das reservas num domínio, a taxa total de mensagens de sinalização processadas nos *routers*. Estas mensagens de sinalização correspondem a tentativas de actualização do estado de reserva num *router*. Em particular, uma mensagem de sinalização pode ser usada para tentar iniciar (ou finalizar) uma reserva para um fluxo ou agregado, ou pode ser usada para tentar aumentar (ou diminuir) a largura de banda de um agregado já instalado. Através desta métrica é possível capturar, não só o número de reservas efectuadas nos *routers*, mas também a frequência com que as reservas são actualizadas. Esta frequência tem um impacto directo no custo dos *routers*. É usual considerar-se como métrica de sinalização o número médio de fluxos no domínio [Pan00]. Esta métrica captura apenas a quantidade de informação de estado das reservas, o que é claramente insuficiente, especialmente em domínios com agregação.

Esta métrica de sinalização será utilizada em conjunto com outras métricas para avaliar as vantagens e desvantagens de um sistema com agregação extremo-a-extremo face a um sistema com agregação por área. As outras métricas utilizadas são:

- Utilização das reservas - a percentagem de largura de banda reservada que é utilizada pelo tráfego admitido em todos os pares de ABRs do domínio.
- Probabilidade de bloqueio - probabilidade de um novo fluxo tentar admissão no domínio e a sua admissão ser rejeitada.

Num domínio sem agregação (isto é, onde as reservas são fluxo-a-fluxo), assim que um fluxo tenta admissão no sistema, todos os *routers* que pertencem ao percurso do fluxo necessitam de processar uma mensagem de sinalização. Num domínio com agregação, as mensagens de sinalização são sempre processadas pelo *router* de entrada do agregado, e os *routers* interiores que são atravessados pelo agregado apenas processam mensagens de sinalização se existir uma tentativa para actualizar a largura de banda do agregado. Note-se que, no caso de uma sessão atravessar vários agregados, como é o caso de um domínio com áreas, a chegada de um fluxo pode provocar tentativas de actualização da reserva de largura de banda em mais do que um agregado.

Considere-se o exemplo simples ilustrado na Figura 9-2, que corresponde a um domínio dividido em 3 áreas. Neste exemplo existem 2 sessões de tráfego, uma oferecida entre os DBRs A e D e a outra oferecida entre os DBRs B e D. A figura apresenta também os diagramas de transição de estados das cadeias de *Markov* para os casos de agregados extremo-a-extremo e agregados de áreas. No primeiro caso existem dois agregados extremo-a-extremo; no segundo caso existem três agregados de área. Considera-se neste exemplo simples que a largura de banda dos fluxos pertencentes às duas sessões é de $b_1 = b_2 = 1$ unidade, o tamanho do *bulk* é $q_h = 2$ unidades em todos os agregados, e a capacidade de todos os pares de ABRs é $C_j = 4$ unidades.

Na representação das cadeias de *Markov*, os estados admissíveis são agrupados de acordo com a largura de banda dos agregados correspondentes; cada grupo está delimitado por um polígono e a largura de banda respectiva está indicada num dos cantos do polígono. No caso de agregados de área, os polígonos apresentados correspondem às áreas CD (linha a cheio) e BC (linha a tracejado). No caso de agregados extremo-a-extremo, os polígonos correspondem aos dois agregados existentes em todo o domínio. Por exemplo, no caso de agregados extremo-a-extremo, no estado (2,1) ambos os agregados têm 2 unidades de largura de banda reservada; o primeiro agregado, que tem como origem e destino os *routers* A e D, é utilizado a 100%, e o segundo, que tem como origem e destino os *routers* B e D, é utilizado a 50%. Os estados (3,1) e (1,3) são estados permitidos apenas no caso de agregados de área, pois neste caso existe apenas um agregado na área CD partilhado por ambas as sessões, podendo a sua largura de banda aumentar até ao limite de 4 unidades; este exemplo ilustra o facto de este tipo de agregação permitir utilizações mais elevadas.

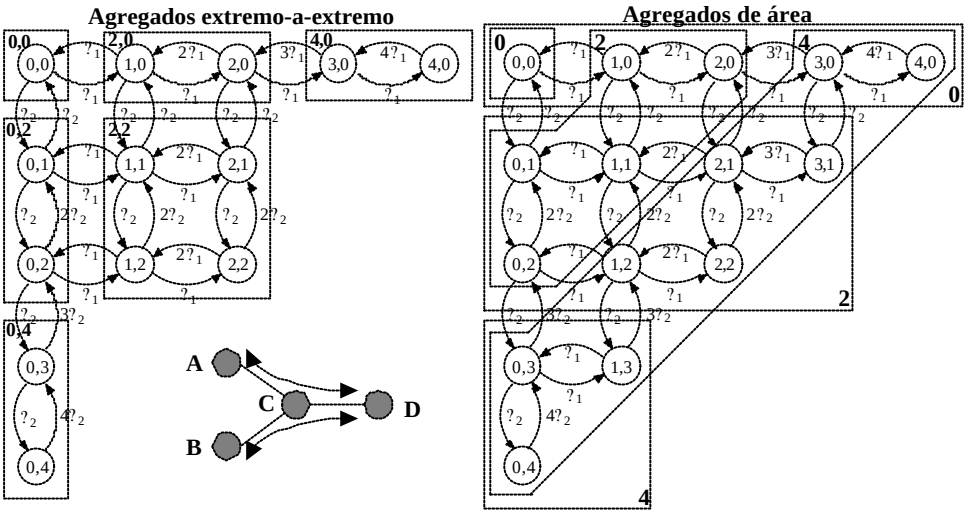


Figura 9-2: Domínio com 3 áreas e diagrama de transição de estados, respectivamente, com agregados extremo-a-extremo e de área.

As mensagens de sinalização processadas na tentativa de actualização da largura de banda de um agregado são induzidas por transições entre estados que pertencem a polígonos diferentes. Note-se que também existem mensagens de sinalização quando existem tentativas de transição de estados fronteira com elevada largura de banda reservada para estados que não são admissíveis. Nestas situações, as mensagens de sinalização são processadas por todos os *routers* que pertencem ao percurso do agregado atravessado pelo fluxo. Em agregados de área, a chegada de um fluxo pode induzir mensagens de sinalização em apenas algumas das áreas atravessadas pelo fluxo. Considere-se o exemplo do estado (1,1). As transições para ambos os estados (2,1) ou (1,2), induzem mensagens de sinalização apenas no agregado da área CD. No entanto, a transição do estado (0,2) para o estado (0,3) induz mensagens de sinalização simultaneamente nas áreas CD e AC (ou BC).

No caso geral, o espaço de estados da cadeia de *Markov* é definido por:

$$\Omega = \left\{ \mathbf{n} \in \mathbb{N}^K : \sum_{k \in \Omega_j} \sum_{h \in \Omega_k} n_k b_k \le C_j,\ j = 1,\ldots,J \right\} \tag{9-1}$$

em que $\mathbb{N}^K$ é o conjunto dos vectores com K elementos inteiros não negativos e $\lceil x \rceil_h$ é o menor múltiplo de q_h superior a x. O somatório interior na definição do espaço de estados corresponde à largura de banda reservada para cada agregado, que tem de ser sempre um

múltiplo de q_h. O somatório exterior corresponde à largura de banda total, reservada para todos os agregados, no par de ABRs j ? ? .

Uma vez definido o espaço de estados, a probabilidade limite do estado **n**, representada por ?**n**, é dada por [Ross97]:

$$? _{\mathbf{n}} ? \frac{\displaystyle ? _{k?1}^{K} \frac{? _k^{n_k}}{n_k!}}{\displaystyle ? _{\mathbf{n}? \preccurlyeq ? } \left(? _{k?1}^{K} \frac{? _k^{n_k} ?}{n_k! ?} \right)}, ? _k ? \frac{? _k}{? _k} \tag{9-2}$$

Para modelar a carga de sinalização, considera-se que $\mathbf{n}_k^?$ é um estado (possivelmente não pertencente a ?) atingido a partir de **n** ? ? através do aumento em uma unidade do número de fluxos da sessão k, isto é, $\mathbf{n}_k^?$ = (n_1, ..., n_k+1, ..., n_K); considera-se que o par $?\mathbf{n},\mathbf{n}_k^??$ representa uma transição ascendente (possivelmente não permitida no espaço de estados) do estado **n** para o estado $\mathbf{n}_k^?$. Existem dois tipos de transições ascendentes: permitidas e proibidas. O conjunto de transições permitidas, ? , e o conjunto de transições proibidas, ? , são definidos por:

$$\preccurlyeq ? ?\mathbf{n},\mathbf{n}_k^? ?:\mathbf{n},\mathbf{n}_k^? ? \preccurlyeq? \tag{9-3}$$

$$\preccurlyeq ? ?\mathbf{n},\mathbf{n}_k^? ?:\mathbf{n} ? \preccurlyeq,\mathbf{n}_k^? ? \preccurlyeq? \tag{9-4}$$

Para descrever a actualização da largura de banda do agregado h atravessado pela sessão k, introduz-se a função indicadora $I_h^{?\mathbf{n},\mathbf{n}_k^??}$. Esta função toma o valor 1 sempre que existe uma transição $?\mathbf{n},\mathbf{n}_k^??$ permitida ou proibida, induzida por uma chegada de um fluxo pertencente à sessão k, e não existe largura de banda suficiente para acomodar este fluxo (sendo portanto necessário actualizar a largura de banda do agregado h). O valor da função é zero sempre que existe uma transição em que não é necessário actualizar a largura de banda do agregado h. Esta função define-se da seguinte forma:

$$I_h^{?\mathbf{n},\mathbf{n}_k^??} ? \begin{cases} ?1, & ? _{k? \preccurlyeq_h} n_k b_k ? b_k ? ? ? ? _{k? \preccurlyeq_h} n_k b_k ? ? ?_h \\ ?0, & \textit{caso contrário} \end{cases} \tag{9-5}$$

Se $I_h^{(\mathbf{n},\mathbf{n}_k')} = 1$ para $(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}$, isto é, para uma transição permitida, existirá uma actualização do agregado com sucesso; caso contrário, se a transição é proibida, a tentativa de actualização do agregado falha. Em ambos os casos, as mensagens de sinalização são processadas em todos os *routers* ao longo do percurso do fluxo. O número de *routers* de um agregado h que tenta actualizar a sua reserva numa transição $(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}$ é $a_h I_h^{(\mathbf{n},\mathbf{n}_k')}+1$, em que o primeiro termo corresponde ao número de *routers* interiores, situados entre o *router* de entrada e de saída do agregado, e o segundo termo corresponde ao *router* de entrada do agregado. Qualquer transição no sentido inverso, $(\mathbf{n}_k',\mathbf{n}) \notin \mathcal{T}$, provoca uma actualização da reserva de recursos num mesmo número de *routers*.

Define-se λ_A^B com sendo a taxa de mensagens de sinalização, em que A indica se apenas são considerados os *routers* interiores (I) ou todos os *routers* do domínio (D), e B indica se são consideradas todas as tentativas de reservas (T) ou apenas as tentativas de reservas com sucesso (S). Usando estas definições:

$$\lambda_I^S = \sum_{(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}} \lambda_k \pi_{\mathbf{n}} \left[n_k + 1 + \pi_{k+1} \pi_{\mathbf{n}_k'} \sum_{h \preceq k} a_h I_h^{(\mathbf{n},\mathbf{n}_k')} \right] \tag{9-6}$$

$$\lambda_I^T = \lambda_I^S = \sum_{(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}} \lambda_k \pi_{\mathbf{n}} \sum_{h \preceq k} a_h I_h^{(\mathbf{n},\mathbf{n}_k')} \tag{9-7}$$

$$\lambda_D^S = \sum_{(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}} \lambda_k \pi_{\mathbf{n}} \left[n_k + 1 + \pi_{k+1} \pi_{\mathbf{n}_k'} \sum_{h \preceq k} a_h I_h^{(\mathbf{n},\mathbf{n}_k')} + 1 \right] \tag{9-8}$$

$$\lambda_D^T = \lambda_D^S = \sum_{(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}} \lambda_k \pi_{\mathbf{n}} \left[\sum_{h \preceq k} a_h I_h^{(\mathbf{n},\mathbf{n}_k')} + 1 \right] \tag{9-9}$$

Note-se que, sempre que $I_h^{(\mathbf{n},\mathbf{n}_k')} = 0$, existem mensagens de sinalização apenas no *router* de entrada do agregado; quando $I_h^{(\mathbf{n},\mathbf{n}_k')} = 1$, todos os *routers* no percurso do agregado necessitam de processar as mensagens de sinalização. Os estudos relacionados com a carga de sinalização têm como referência a taxa de mensagens de sinalização em reservas efectuadas fluxo-a-fluxo (IntServ), que é designada de $\lambda_{A,ref}^B$. Neste caso, sempre que um novo fluxo pede admissão, todos os *routers* no seu percurso processam mensagens de sinalização. Assim:

$$\lambda_{I,ref}^S = \sum_{(\mathbf{n},\mathbf{n}_k') \in \mathcal{T}} \lambda_k \pi_{\mathbf{n}} \left[n_k + 1 + \pi_{k+1} \pi_{\mathbf{n}_k'} \sum_{h \preceq k} a_h \right] \tag{9-10}$$

$$\pi^{T}_{I,ref} = \pi^{S}_{I,ref} = \sum_{\mathbf{n},\mathbf{n}_k \in \Omega} \pi_{\mathbf{n}} \sum_{h \in \mathcal{L}_k} a_h \tag{9-11}$$

$$\pi^{S}_{D,ref} = \sum_{\mathbf{n},\mathbf{n}_k \in \Omega} \pi_{\mathbf{n}} \, n_k \left(1 - \prod_{k \in \mathcal{L}_k} \left(1 - \frac{\mathbf{n}_k}{C_k}\right)^{a_h}\right) \tag{9-12}$$

$$\pi^{T}_{D,ref} = \pi^{S}_{D,ref} = \sum_{\mathbf{n},\mathbf{n}_k \in \Omega} \pi_{\mathbf{n}} \sum_{h \in \mathcal{L}_k} a_h \left(1 - \frac{\mathbf{n}_k}{C_k}\right) \tag{9-13}$$

Note-se que $\sum_{h \in \mathcal{L}_k} a_h \left(1 - \frac{\mathbf{n}_k}{C_k}\right)$ representa o número total de *routers* no percurso do fluxo *k*. Note-se também que a sinalização em reservas fluxo-a-fluxo é um caso particular do caso de sinalização com agregação, em que existe um agregado extremo-a-extremo por sessão e um tamanho do *bulk* igual à largura de banda do fluxo de cada sessão. Finalmente, define-se o ganho de sinalização como sendo a razão entre a taxa de sinalização em reservas fluxo-a-fluxo e a taxa de sinalização com agregação, ou seja, $G^{B}_{A} = \pi^{B}_{A,ref} / \pi^{B}_{A}$.

A utilização das reservas representa a fracção das reservas dos agregados que é utilizada pelo tráfego admitido. Esta métrica de desempenho é definida por:

$$U = \frac{1}{J} \sum_{\mathbf{n} \in \Omega} \sum_{j \in \mathcal{L}} \left(\frac{\sum_{h \in \mathcal{L}_j} \sum_{k \in \mathcal{L}_h} n_k b_k}{\sum_{h \in \mathcal{L}_j} \sum_{k \in \mathcal{L}_h} n_k b_k} \right) \pi_{\mathbf{n}} \tag{9-14}$$

O numerador representa a largura de banda do tráfego admitido no par de ABRs *j*, e o denominador representa a largura de banda que é reservada para o mesmo tráfego nesse par de ABRs *j*. A utilização dos recursos reservados no domínio é uma média pesada das utilizações dos recursos em todos os pares de ABRs.

A probabilidade de bloqueio de uma sessão *k* corresponde à soma das probabilidades de todos os estados que bloqueiam chegadas de fluxos dessa sessão. No caso geral de um processo de nascimento e morte multi-dimensional, a probabilidade de bloqueio de uma sessão *k* é dada por:

$$B_k = 1 - \frac{\sum_{\mathbf{n} \in \Omega_k} \sum_{k \in \mathcal{L}} \frac{\rho_k^{n_k}}{n_k!}}{\sum_{\mathbf{n} \in \Omega} \sum_{k \in \mathcal{L}} \frac{\rho_k^{n_k}}{n_k!}} = 1 - \frac{\sum_{\mathbf{n} \in \Omega_k} \pi_{\mathbf{n}}}{\sum_{\mathbf{n} \in \Omega} \pi_{\mathbf{n}}} \tag{9-15}$$

em que $?_k$ representa o espaço de estados que não são fronteira, isto é, que podem admitir pelo menos mais um fluxo da sessão *k*. $?_k$ é definido por:

$$\mathbf{?}_k = \left\{ \mathbf{n} \in \mathbf{?}^K : \sum_{h \in \mathbf{?}_j} \sum_{n \in \mathbf{?}_h} n_l b_l \le b_k \right\} \quad \forall C_j, j = 1,\ldots,J \tag{9-16}$$

9.3 Modelo de carga por agregado

Esta secção apresenta um modelo que considera uma carga oferecida variável no tempo. Este modelo é uma extensão para múltiplas áreas do modelo descrito em [Fu01]. Em particular, considera-se que o tráfego agregado de uma sessão *k* é caracterizado por uma sinusóide com fase variável:

$$r_k(t) = f_k + e_k \cos\left(\frac{2\pi}{T} t + \varphi_k\right) \tag{9-17}$$

em que f_k é a largura de banda média do agregado, e_k é a amplitude da sinusóide, T é o período da sinusóide, e $?_k$ é a fase aleatória uniformemente distribuída no intervalo [0,2?]. A adopção deste modelo é motivada pelo comportamento de um elevado número de medições de tráfego agregado que exibem uma tendência periódica a longo prazo. Veja-se, por exemplo, a medição de tráfego apresentada na Figura 9-3. Estes dados foram medidos e disponibilizados pelo projecto *QBone* do *Internet2 QoS Working Group* [Int-Qbone]. O período *T* corresponde a 24 horas.

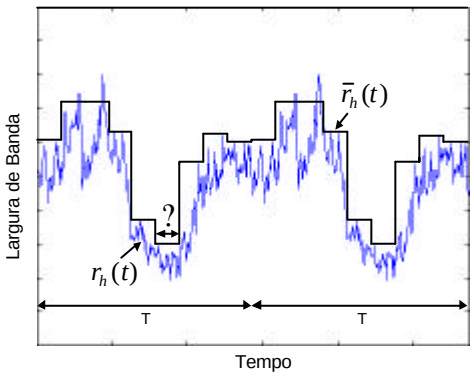


Figura 9-3: Medição de tráfego [Int-Qbone].

Em [Fu01] foram efectuados estudos, com base em agregados reais, de ajustamento de diferentes modelos de tráfego (sinusoidal, triangular, em onda quadrada). Os resultados obtidos mostram que, desde que o ruído do tráfego agregado seja moderado, o modelo de carga oferecida sinusoidal consegue prever adequadamente o comportamento do sistema.

Neste modelo assume-se, como já foi referido anteriormente, que (i) a reserva do agregado *h* é actualizada no início de intervalos de tempo fixos de duração *?* e que (ii) a quantidade de largura de banda que se tenta reservar no início de um intervalo de tempo corresponde à largura de banda necessária para suportar o tráfego nesse intervalo. Concretamente, a largura de banda a reservar no intervalo de tempo x_h = 1, 2, ..., *T*/*?*, no agregado *h*, corresponde à largura de banda máxima oferecida neste intervalo, isto é,

$$r_{h,x_h} \; ? \; \max_{?x_h?1?\,?\,t?\,x_h?} \; ? \; r_k \; ?t? \qquad (9\text{-}18)$$

Note-se que, no início do intervalo de tempo x_h, a largura de banda do agregado pode ou não ser actualizada para r_{h,x_h}, dependendo da largura de banda disponível nos pares de ABRs atravessados pelo agregado *h*. Este facto é ilustrado na Figura 9-3 onde $r_h(t)$ e $\bar{r}_h$ *?t?* representam a carga oferecida e a largura de banda efectivamente reservada ao agregado *h*, respectivamente.

Para simplificar o desenvolvimento do modelo e das suas métricas de desempenho, assume-se que *T*/*?* é um número inteiro. As métricas de desempenho utilizadas neste modelo são a probabilidade de sobrecarga e a utilização das reservas.

A probabilidade de sobrecarga de uma sessão *k*, P_k, é definida como sendo a fracção da carga oferecida pela sessão *k* que não pode ser admitida. Para determinar esta métrica considera-se uma aproximação de carga reduzida [Ross95], em que o tráfego oferecido a um par de ABRs é reduzido de acordo com a sobrecarga sofrida pelas sessões que atravessam esse par de ABRs, nos outros pares de ABRs atravessados pelas mesmas sessões. Partindo do pressuposto que a sobrecarga é independente de ligação para ligação, a carga reduzida oferecida pela sessão de tráfego *k* ao par de ABRs *j* é aproximada por

$$r_k \; ?t? \; ? \; ?1 \; ? \; T_l \; ? \qquad (9\text{-}19)$$
$$l? \not< k \; ?? j?$$

em que T_l é a probabilidade de sobrecarga no par de ABRs *l*. A carga reduzida oferecida ao par de ABRs *j* por todas as sessões de tráfego que o atravessam é

$$\hat{r}_j(t) = \sum_{h \preceq_j} \sum_{k \preceq_h} r_k(t) \sum_{l \preceq_k j} 1 / T_l \tag{9-20}$$

Uma vez que a carga oferecida ao par de ABRs *j* é reduzida, a largura de banda que se pretende reservar no intervalo de tempo x_h vem também reduzida, sendo dada por

$$\hat{r}_{j,x_h} = \sum_{h \preceq_j} \hat{r}_{h,x_h} = \sum_{h \preceq_j} \max_{x_h - 1 \le t \le x_h} \left(\sum_{k \preceq_h} r_k(t) \sum_{l \preceq_k j} 1 / T_l \right) \tag{9-21}$$

Considera-se que existe sobrecarga no intervalo de tempo x_h quando a largura de banda a reservar para o conjunto dos agregados do par de ABRs *j* é superior à capacidade do mesmo, isto é, quando

$$\sum_{h \preceq_j} \hat{r}_{h,x_h} > C_j \tag{9-22}$$

A probabilidade de sobrecarga no par de ABRs *j* é dada pela razão entre a largura de banda de sobrecarga (largura de banda que não é possível reservar para o conjunto dos agregados do par de ABRs *j*) e a largura de banda que seria necessário reservar para o conjunto de agregados no par de ABRs *j*. Esta probabilidade é aproximada pela seguinte expressão:

$$T_j = \frac{\left(\frac{1}{T}\right)^{H_j} \sum_{x_1=1}^{T} \cdots \sum_{x_{H_j}=1}^{T} \left(\sum_{h \preceq_j} \hat{r}_{h,x_h} - C_j \right)^{+}}{\frac{1}{T} \sum_{h \preceq_j} \sum_{x_h=1}^{T} \hat{r}_{h,x_h}}, \quad j = 1,2,\ldots, J \tag{9-23}$$

O numerador corresponde ao valor médio de sobrecarga no par de ABRs *j*. Os somatórios no numerador realizam todas as combinações possíveis que correspondem a desfasamentos entre o tráfego oferecido aos diferentes agregados; $(1/T)^{H_j}$ é a probabilidade de cada combinação; o denominador corresponde à média da largura de banda a reservar no conjunto dos agregados no par de ABRs *j*.

Este conjunto de *J* equações não lineares com *J* incógnitas pode ser resolvido pelo método de repetições sucessivas [Ross95]. O conjunto de equações (9-23) é uma extensão da equação de probabilidade de sobrecarga definida em [Fu01].

A probabilidade de sobrecarga de uma sessão *k* depende das probabilidades de sobrecarga em todos os pares de ABRs atravessados pela sessão *k*. Esta probabilidade é aproximada pela seguinte expressão:

$$P_k = 1 - \sum_{j \preceq_k} (1 - T_j) \tag{9-24}$$

A utilização das reservas é definida como a razão entre a carga média reduzida admitida e a largura de banda média reservada. A utilização é aproximada pela seguinte expressão:

$$U = \frac{1}{J} \sum_{j \preceq} \frac{(1 - T_j) E\left[\sum_{h \preceq_j} \sum_{k \preceq_h} r_k \left(t - \sum_{l \preceq_k \wedge l \neq j} (1 - T_l) \right) \right]}{\sum_{x_1=1}^{H_j} \cdots \sum_{x_{H_j}=1}^{T/r} \min\left(r_{j,x_h}, C_j \right)} \tag{9-25}$$

em que $E(x)$ representa o valor esperado de x. O numerador corresponde ao valor esperado da carga oferecida a cada par de ABRs j. A utilização das reservas tem em consideração apenas a carga média oferecida que é admitida. O denominador da equação corresponde à média da quantidade de largura de banda reduzida reservada em todos os agregados presentes no par de ABRs j. A utilização das reservas no domínio é definida como a média das utilizações em todos os pares de ABRs.

Nas secções seguintes vão ser apresentados resultados numéricos obtidos com os dois modelos desenvolvidos e resultados obtidos através de simulações de eventos discretos que consideram tráfego agregado medido num *router* de acesso da Universidade de *Auckland* [Int-NLANR].

9.4 Resultados numéricos e simulações

Nesta secção são apresentados exemplos numéricos e simulações que têm como objectivo analisar os compromissos entre a carga de sinalização e a utilização dos recursos reservados. Numa primeira fase considera-se a topologia do tipo *Dumbbell* ilustrada na Figura 9-4, com duas áreas periféricas em cada lado do domínio e uma área central. Posteriormente consideram-se dois domínios diferentes: um domínio baseado também na topologia *Dumbbell*, mas de maior dimensão, que representa uma rede de núcleo, e um domínio com topologia em árvore que representa uma rede de acesso.

Nas experiências que utilizam a rede ilustrada na Figura 9-4 existem 4 sessões no domínio, com percursos ACDE, ACDF, BCDF e BCDE, que serão designadas por r_{xy}, em que x representa a origem e y representa o destino. O número de *routers* dentro de cada

área é 4. Assim, cada sessão atravessa 15 *routers* (sem incluir o *router* de saída do domínio). Com a excepção das simulações que consideram tráfego real, a largura de banda de cada par de ABRs é de 32 Mb/seg em todas as áreas.

Nestas experiências são comparados dois tipos de domínios de rede: (i) domínios apenas com agregados extremo-a-extremo e (ii) domínios apenas com agregados de áreas. Nas figuras apresentadas de seguida, a agregação extremo-a-extremo é designada por "Extremo-a-extremo" e a agregação por área é designada por "Área". Em agregação por área consideram-se domínios com uma classe de serviço, designados por "1 classe", e com duas classes de serviço, designados por "2 classes". No caso de duas classes de serviço, considera-se que as sessões r_{AE} e r_{AF} pertencem a uma classe de serviço e que as sessões r_{BE} e r_{BF} pertencem a outra classe de serviço. Assim, existirá apenas um agregado em cada área periférica da esquerda e dois agregados nas restantes áreas. Consideram-se dois casos no que respeita ao tamanho do *bulk* ou a *?/T*: o tamanho do *bulk* (ou *?/T*) é o mesmo em todos os agregados do domínio, denominado de *bulk* fixo (ou *?/T* fixo), e o tamanho do *bulk* (ou *?/T*) é proporcional à carga oferecida aos agregados, designado por "*bulk* prop." (ou "*?/T* prop."). Neste último caso, o eixo dos *xx* representa o tamanho do *bulk* (ou *?/T*) nas áreas periféricas. A motivação de considerar *bulk* e *?/T* proporcionais ao tráfego oferecido está relacionada com o facto de os agregados de área serem atravessados por uma maior quantidade de tráfego.

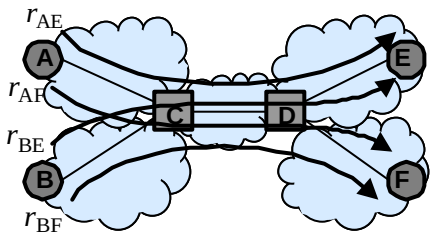


Figura 9-4: Topologia *Dumbbell* com 2 áreas periféricas.

9.4.1 Modelo de carga por fluxo

Nesta secção assume-se que as sessões são caracterizadas por b_k = 1 Mb/seg e $1/?_k$ = 1 segundo. Quando nada é referido em contrário, considera-se $?_k = 8$ seg^{-1}. Neste caso, a largura de banda média oferecida em cada sessão é de 8 Mb/seg.

A Figura 9-5 ilustra os ganhos de sinalização obtidos nos sistemas com agregação extremo-a-extremo e com agregação por área. Os ganhos de sinalização apresentados consideram (i) apenas as tentativas de reservas que resultaram em sucesso (curvas a cheio, designadas por "Sucesso") e (ii) todas as tentativas de reservas (curvas a tracejado, designadas por "Todas"). Os resultados mostram que, quando a largura de banda do *bulk* é igual à largura de banda dos fluxos, a taxa de sinalização iguala a taxa de sinalização fluxo-a-fluxo (obtém-se um ganho unitário). Os ganhos obtidos em relação à sinalização fluxo-a-fluxo aumentam com o tamanho do *bulk*. Os ganhos obtidos com um *bulk* de 8 Mb/seg e considerando apenas as tentativas de reservas com sucesso, são de aproximadamente 15 em agregação extremo-a-extremo e 4 em agregação por área. O ganho máximo possível no domínio é de 15 e de 5, respectivamente, no limite quando o *bulk* é tão elevado que não existe sinalização nos *routers* interiores. Note-se que o número de *routers* que processam mensagens de sinalização fluxo-a-fluxo (sempre que um fluxo chega ou termina), é 1 em agregação extremo-a-extremo (apenas o DBR de entrada), 3 em agregação por área (os ABRs de entrada de cada área) e 15 sem agregação, o que explica os ganhos limite referidos. Os ganhos em agregação extremo-a-extremo quase atingem o limite quando o tamanho do *bulk* é de 8 Mb/seg. Dado que na área central existem 4 agregados, cada um com uma largura de banda média oferecida de 8 Mb/seg, e que a largura de banda de cada par de ABRs é de 32 Mb/seg, a tendência do sistema é ter a largura de banda de todos os agregados ajustada ao tamanho do *bulk* em todos os instantes. Esta tendência faz com que praticamente não exista actualização da largura de banda dos agregados e, deste modo, o número de mensagens de sinalização nos *routers* interiores seja muito pequeno. Os ganhos com agregação por área são muito próximos nas três situações consideradas, sendo ligeiramente superiores no caso de 1 classe de serviço com *bulks* proporcionais. Nesta situação, a utilização de um *bulk* superior na área central (neste caso o *bulk* na área central é o dobro do *bulk* nas áreas periféricas) provoca uma ligeira diminuição na carga de sinalização nesta área.

Considere-se agora a diferença entre os ganhos de sinalização obtidos com (i) as tentativas de reservas com sucesso e (ii) todas as tentativas de reservas. Para um *bulk* de 8 Mb/seg e considerando agregação extremo-a-extremo, o ganho de sinalização decresce de 15 para 8.5. Esta diminuição reflecte o número significativo de pedidos de reservas que não podem ser estabelecidos, isto é, uma probabilidade de bloqueio relativamente elevada.

Com agregação por área, os ganhos de sinalização são aproximadamente iguais, o que indica que quase todas as tentativas de reserva têm sucesso.

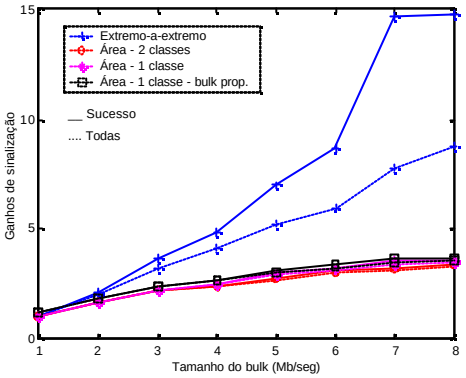


Figura 9-5: Ganhos de sinalização de todos os *routers* do domínio (modelo de carga por fluxo).

A Figura 9-6 considera os ganhos de sinalização atingidos nos *routers* interiores, isto é, sem incluir os DBRs no caso de agregação extremo-a-extremo, e sem incluir os DBRs e ABRs, no caso de agregação por área. Os ganhos correspondem a uma média dos ganhos em todos os *routers* interiores e consideram (i) todas as tentativas de reservas e (ii) apenas as tentativas de reservas com sucesso. Esta métrica é relevante, pois reflecte os ganhos em termos de custo dos *routers* interiores. Note-se que os ganhos de sinalização no domínio (englobando todos os *routers*) podem ser polarizados pelo número de *routers* que necessitam de efectuar sinalização fluxo-a-fluxo, e que o custo dos *routers* não é, em princípio, uma função linear da carga de sinalização.

Considerando apenas as tentativas de reservas com sucesso e *bulk* fixo, observa-se que os ganhos de sinalização são semelhantes em agregação por área com uma e duas classes de serviço (recorde-se que no caso de uma classe de serviço existe apenas um agregado na área central, e no caso de duas classes de serviço existem dois). Dois efeitos opostos explicam estas semelhanças. Primeiro, no caso de uma classe de serviço os agregados são sempre partilhados por mais do que uma sessão; o tráfego total dentro de cada agregado torna-se mais "suave", contribuindo para reduzir a taxa de sinalização. Por outro lado, a maior partilha de recursos em agregação por área com uma classe de serviço aumenta o número de fluxos admitidos, contribuindo para o aumento da taxa de sinalização. Estes dois efeitos opostos quase se anulam neste caso específico, e por isso, os ganhos totais obtidos são semelhantes. No caso de *bulks* proporcionais ao tráfego

oferecido, verifica-se que os ganhos de sinalização aumentam em relação aos ganhos existentes com *bulks* fixos. Isto deve-se ao facto de, com o aumento do tamanho do *bulk* da área central, existir uma diminuição do número de mensagens de sinalização nesta área. Os ganhos de sinalização em agregação extremo-a-extremo são maiores que os de agregação por área para tamanhos do *bulk* maiores do que aproximadamente 4 Mb/seg. Este facto é explicado novamente pela tendência do sistema, em ter a largura de banda de todos os agregados ajustada ao tamanho do *bulk* em todos os instantes, para valores elevados do tamanho do *bulk*.

Considerando a diferença entre os ganhos obtidos com todas as tentativas de reservas e apenas com as tentativas de reservas com sucesso, verifica-se que esta é elevada num sistema com agregação extremo-a-extremo, devido à elevada probabilidade de bloqueio. Com agregação por área as diferenças são muito menores.

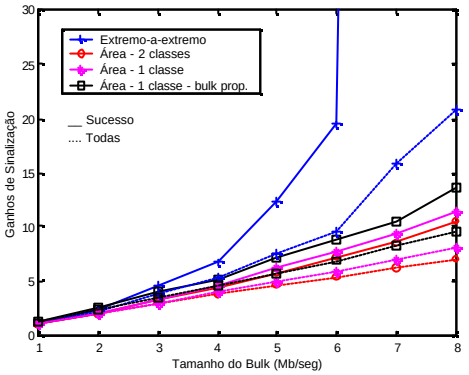


Figura 9-6: Ganhos de sinalização dos *routers* interiores (modelo de carga por fluxo).

A Figura 9-7 ilustra a utilização das reservas. Esta utilização é superior no caso de agregação por área, em comparação com agregação extremo-a-extremo, pois na primeira existe maior partilha de recursos. Em agregação por área com uma classe de serviço as quatro sessões partilham o mesmo agregado na área central. Em agregação extremo-a-extremo existe apenas partilha de recursos entre agregados. Considerando agregação por área e duas classes de serviço, o número de agregados em cada área é maior do que o existente em agregação por área com uma classe de serviço. O número de sessões partilhadas no mesmo agregado é inferior, e por isso, a utilização da reserva de recursos é inferior à obtida quando apenas se considera uma classe de serviço. Considerem-se agora as situações de agregação por área com uma classe de serviço, com *bulks* fixos ou *bulks*

proporcionais. A utilização atingida na situação em que os *bulks* são proporcionais é ligeiramente inferior, mas muito próxima da utilização obtida na situação em que se consideram *bulks* fixos. Esta proximidade reflecte o bom compromisso entre sinalização e utilização que pode ser atingido neste caso.

Apenas como exemplo, para ter uma utilização sempre superior a 84%, é necessário que o tamanho do *bulk* seja inferior a 5 Mb/seg em agregação extremo-a-extremo, e a 6 e 8 Mb/seg em agregação por área com duas ou uma classe de serviço, respectivamente. Pela Figura 9-6 observa-se que, com estes valores do tamanho dos *bulks*, o ganho de sinalização dentro das áreas com agregação extremo-a-extremo é de 7, e com agregação por área (e com *bulks* proporcionais) é de 13. Deste modo, com agregação por área, os *routers* do interior das áreas podem ter requisitos de desempenho mais baixos (por exemplo, em termos de capacidade de processamento), e consequentemente, o seu custo pode ser inferior ao necessário em agregação extremo-a-extremo, sem que a utilização de recursos seja penalizada.

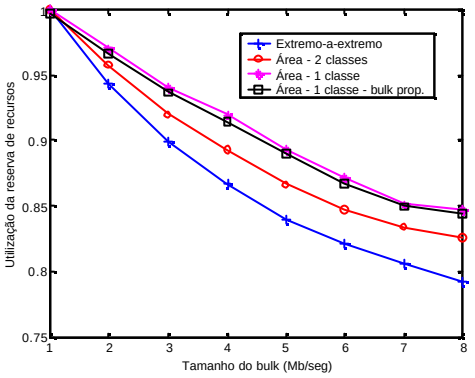


Figura 9-7: Utilização da reserva de recursos (modelo de carga por fluxo).

A probabilidade de bloqueio, ilustrada na Figura 9-8, é apresentada em função da taxa de chegada dos fluxos, para um tamanho do *bulk* de 4 Mb/seg. A razão de ser desta opção prende-se com o facto de a probabilidade de bloqueio ser dependente da relação entre a capacidade das ligações e a largura de banda do *bulk*. Os resultados obtidos comprovam o aumento do desempenho do sistema quando os agregados são configurados por área. Considerando a taxa média de chegada de fluxos de 8 seg^{-1} (igual à das curvas anteriores), verifica-se que a probabilidade de bloqueio em agregação extremo-a-extremo é mais do dobro da probabilidade de bloqueio em agregação por área. Estes resultados

confirmam os resultados anteriores relativos à diferença entre ganhos de sinalização obtidos considerando todas as tentativas de reservas e apenas as tentativas com sucesso.

Um sistema com agregação por área e duas classes de serviço é, assim como na utilização dos recursos, um caso intermédio. Uma situação com *bulks* proporcionais ao tráfego oferecido apresenta uma probabilidade de bloqueio que é apenas ligeiramente superior à existente com *bulks* fixos, o que reflecte o bom compromisso existente entre sinalização e bloqueio.

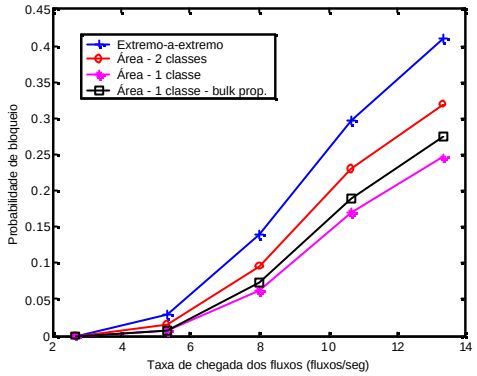
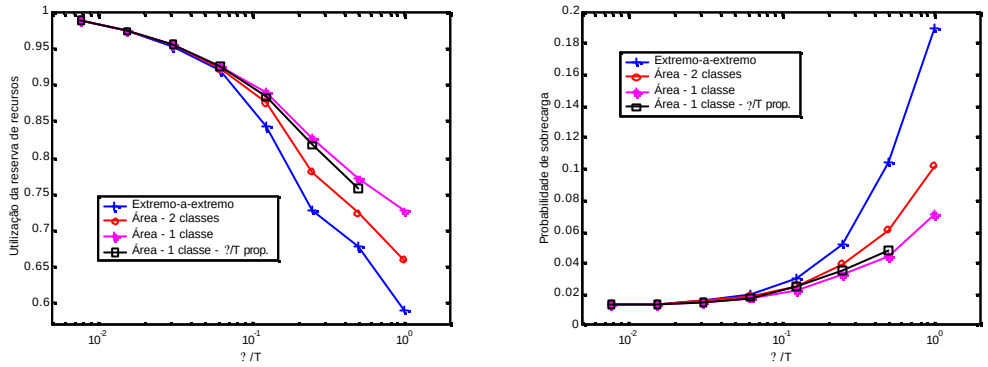


Figura 9-8: Probabilidade de bloqueio (modelo de carga por fluxo).

9.4.2 Modelo de carga por agregado

Com o objectivo de estudar um cenário com alguma sobrecarga, e tendo em conta que a largura de banda de todos os pares de ABRs é de 32 Mb/seg, considera-se neste estudo que a largura de banda média da onda sinusoidal é f_k = 5.3 Mb/seg e a amplitude da onda é também e_k = 5.3 Mb/seg.

A Figura 9-9 apresenta a utilização dos recursos reservados no domínio e a probabilidade de sobrecarga em função de $?/T$. As curvas que correspondem ao caso de agregação por área com uma classe de serviço e $?/T$ proporcional apresentam valores apenas para $?/T$? 1/2. Note-se que os valores de $?/T$ da área central são o dobro dos das áreas periféricas, e o eixo dos *xx* representa $?/T$ nas áreas periféricas.



(a) **(b)**

Figura 9-9: (a) Utilização da reserva de recursos e (b) probabilidade de sobrecarga (modelo de carga por agregado).

Com o aumento de *?/T*, a frequência de actualização das reservas dos agregados diminui, reduzindo a utilização das reservas e aumentando a probabilidade de sobrecarga. Para *?/T* pequeno, a utilização das reservas e a probabilidade de sobrecarga obtidas com os vários tipos de agregação são aproximadamente semelhantes. Note-se que, com agregação extremo-a-extremo, existem quatro agregados na área central (um agregado por sessão), e com agregação por área e uma classe de serviço, existe apenas um agregado na mesma área, sendo os seus recursos partilhados pelas quatro sessões que o atravessam. Com agregação extremo-a-extremo, as quatro sessões partilham os recursos apenas por via dos seus agregados, ou seja, a adaptação entre o tráfego oferecido e a largura de banda reservada ocorre apenas de *?/T* em *?/T*. Em agregação por área, as quatro sessões partilham sempre os recursos independentemente do valor de *?/T*. Para *?/T* pequeno, os agregados são ajustados frequentemente (em relação à taxa de variação do tráfego oferecido), e por isso, existe uma partilha de recursos significativa nos dois tipos de agregação. No entanto, a taxa de sinalização é também elevada. Com o aumento de *?/T* e a diminuição da taxa de sinalização, as reservas são efectuadas para intervalos de tempo maiores, diminuindo a utilização e aumentando a probabilidade de sobrecarga. Esta diminuição da taxa de sinalização afecta com maior intensidade um sistema com agregação extremo-a-extremo porque, como foi referido acima, em agregação por área, as quatro sessões continuam a partilhar os recursos. Como exemplo, para atingir uma utilização superior a 77% e uma probabilidade de sobrecarga inferior a 4%, *?/T* ? 1/2 em agregação por área com uma

classe de serviço, e ?/*T* ? 1/6 em agregação extremo-a-extremo. O caso que considera agregação por área com uma classe de serviço e ?/*T* proporcional apresenta uma utilização ligeiramente inferior (e uma probabilidade de sobrecarga ligeiramente superior), em comparação com o caso de ?/*T* fixo. Este facto deve-se à diminuição da frequência de actualização das reservas na área central. No entanto, as diferenças entre as curvas que correspondem a ?/*T* fixo e proporcional são muito pequenas, o que indica que a consideração de ?/*T* proporcional tem um impacto muito reduzido na utilização e na probabilidade de sobrecarga. O aumento do número de classes de serviço aumenta o número de agregados necessários em cada área. Sendo assim, o caso de duas classes de serviço é um caso intermédio entre agregação extremo-a-extremo e agregação por área com uma classe de serviço.

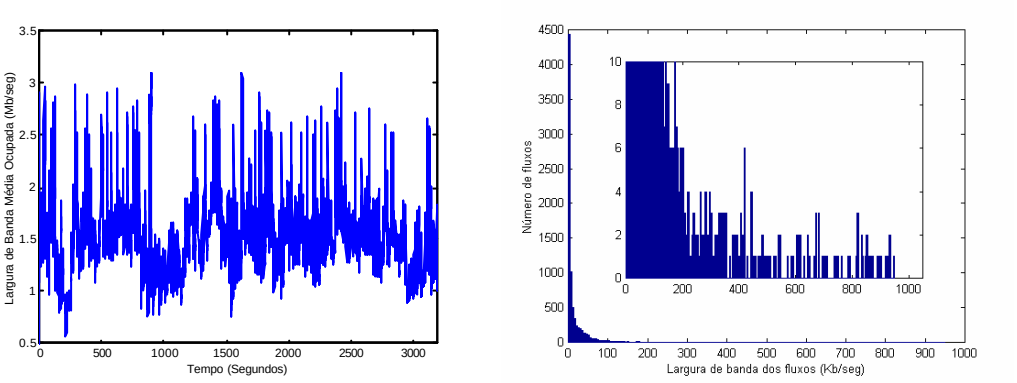
Embora este modelo não detalhe a carga oferecida ao nível do fluxo, é ainda possível derivar uma aproximação (grosseira) dos ganhos de sinalização notando, como em [Fu01], que uma utilização das reservas unitária é atingida quando é efectuada sinalização fluxo-a-fluxo. De acordo com os resultados da Figura 9-9 (a), uma utilização unitária é aproximadamente atingida quando ?/*T* = 1/256. Assim, os ganhos de sinalização serão de 32 quando ??? = 1/8, e de 128 quando ??? = 1/2.

No caso geral, os resultados obtidos com este modelo confirmam os obtidos com o modelo de carga por fluxo. Embora este modelo considere que a carga oferecida é variante no tempo, ele não permite determinar a carga de sinalização exacta. Na próxima secção apresenta-se um estudo de simulação com base em agregados reais, que permite determinar os ganhos de sinalização obtidos quando se consideram cargas oferecidas variantes no tempo.

9.4.3 Resultados com agregados de tráfego real

Nesta secção considera-se uma captura de tráfego medido num *router* de acesso à *Internet* da Universidade de *Auckland* [Int-NLANR]. Este tráfego vai ser utilizado para representar a largura de banda agregada de cada sessão. O desempenho do sistema vai ser determinado através de simulação de eventos discretos. Este tráfego, ilustrado na Figura 9-10 (a), é caracterizado por uma elevada variância e ruído. A informação disponível inclui o instante de chegada, a duração e o número de *bytes* enviados de cada fluxo. O número total de fluxos é de 64087 e a sua largura de banda média é de 19.6 Kb/seg. A largura de banda

média do agregado é de 1.43 Mb/seg e a respectiva variância é de 0.144 Mb/seg. A Figura 9-10 (b) apresenta o histograma da largura de banda dos fluxos. Observa-se que a maior concentração de fluxos se encontra na gama de 0 a 50 Kb/seg. Existem também alguns fluxos com larguras de banda elevadas da ordem das centenas de Kb/seg. Neste agregado cerca de 80% dos fluxos têm largura de banda inferior à média e apenas 20% apresentam largura de banda superior. Os resultados de simulação correspondem a médias efectuadas num total de 20 corridas de simulação; em cada corrida a fase de cada agregado foi definida aleatoriamente.



(a) **(b)**

Figura 9-10: Tráfego medido num *router* de acesso à *Internet* da Universidade de *Auckland*: (a) largura de banda média ocupada e (b) histograma das larguras de banda dos fluxos.

Este estudo permite considerar em simultâneo dois aspectos que foram considerados separadamente nos dois modelos analíticos: tráfego variante no tempo e tráfego detalhado ao nível do fluxo.

Esta secção considera o domínio apresentado na Figura 9-4 da secção 9.4.3.1, e considera também nas secções 9.4.3.2 e 9.4.3.3 dois domínios de maior dimensão. Estes dois domínios representam uma rede de acesso (com topologia em árvore) e uma rede de núcleo (com topologia base *Dumbbell*).

9.4.3.1 Topologia Dumbbell

Neste estudo, a largura de banda dos pares de ABRs é reduzida para 10 Mb/seg, relativamente ao caso das secções anteriores, por forma a garantir a existência de sobrecarga com este agregado real.

A Figura 9-11 apresenta os ganhos de sinalização num domínio com agregação extremo-a-extremo e por área. Os resultados confirmam os obtidos com o modelo de carga por fluxo. Os ganhos de sinalização, considerando apenas as reservas com sucesso, são, uma vez mais, aproximadamente 3 vezes inferiores em agregação por área, comparando com agregação extremo-a-extremo, devido ao número de *routers* que processam mensagens de sinalização fluxo-a-fluxo. Observe-se que, para *bulks* elevados, estes ganhos de sinalização quase atingem o limite que, conforme mencionado na secção 9.4.1, é de 15 em agregação extremo-a-extremo e 5 em agregação por área, reflectindo uma reduzida carga de sinalização nos *routers* interiores dos agregados.

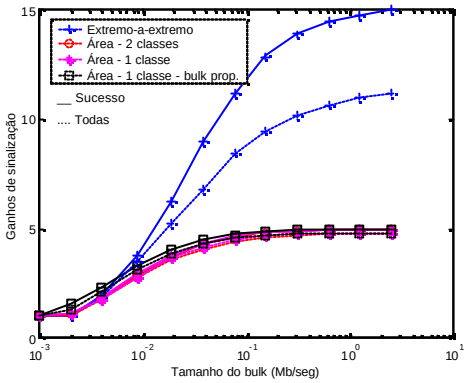


Figura 9-11: Ganhos de sinalização de todos os *routers* do domínio (agregados de tráfego real).

A Figura 9-12 apresenta os ganhos de sinalização dos *routers* interiores. Os ganhos de sinalização aumentam bastante com o tamanho do *bulk*, em todos os tipos de agregação, atingindo valores bastante mais elevados, exceptuando o caso de agregação extremo-a-extremo, que os apresentados na secção 9.4.1. Com um *bulk* de 1.25 Mb/seg e considerando todas as tentativas de reservas, os ganhos de sinalização atingem 500 em agregação extremo-a-extremo, 950 em agregação por área com duas classes de serviço, 1380 em agregação por área com uma classe de serviço e *bulks* fixos, e 1570 em agregação por área com uma classe de serviço e *bulks* proporcionais. Estes ganhos devem-se essencialmente à elevada razão entre o tamanho do *bulk* e a largura de banda dos fluxos.

Observa-se na figura que os ganhos de sinalização são sempre maiores em agregação por área do que em agregação extremo-a-extremo. Este facto é explicado pelo carácter variante do tráfego oferecido que intensifica a partilha de recursos de sessões no mesmo agregado em agregação por área. O efeito produzido por esta maior partilha de recursos

sobrepõe-se ao efeito oposto produzido pelo aumento do número de fluxos admitidos, e como resultado, os ganhos de sinalização em agregação por área são maiores. Note-se que este cenário não inclui a situação limite que foi considerada no modelo de carga por fluxo, em que a largura de banda de todos os agregados é ajustada ao tamanho do *bulk* em todos os instantes, pois o tamanho máximo do *bulk* considerado nas experiências ainda permite uma granularidade suficiente no processo de ajuste da largura de banda dos agregados. Este cenário é, de facto, mais realista. A diferença entre os ganhos obtidos considerando todas as tentativas de reserva e apenas as tentativas de reserva com sucesso é inferior em agregação por área devido ao menor bloqueio de fluxos existente neste tipo de agregação.

Os ganhos de sinalização, obtidos quando a razão entre o tamanho do *bulk* e a largura de banda média dos fluxos é de 8 (isto é, com um tamanho do *bulk* de 156.8 Kb/seg), são de 124 em agregação por área com uma classe de serviço e *bulks* fixos, e de 160 em agregação por área com *bulks* proporcionais. Nas experiências com o modelo de carga por fluxo, para a mesma razão entre larguras de banda (isto é, com um tamanho do *bulk* de 8 Mb/seg), os ganhos eram de 12 e 14, respectivamente. Esta diferença é explicada pela assimetria na distribuição da largura de banda dos fluxos no agregado real, de acordo com a qual 80% dos fluxos têm uma largura de banda abaixo da média (note-se que no modelo de carga por fluxo a largura de banda dos fluxos é fixa). Assim, neste caso, as tentativas de actualização da largura de banda do agregado são induzidas por um número muito menor de fluxos (são induzidas principalmente pelos fluxos com maior largura de banda, que são em menor número), introduzindo um decrescimento na taxa de sinalização.

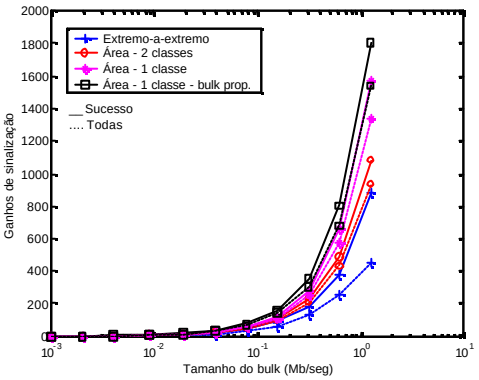
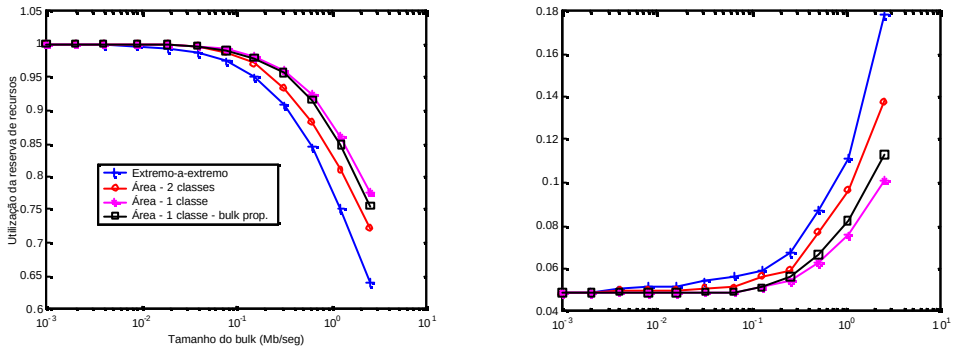


Figura 9-12: Ganhos de sinalização dos *routers* interiores (agregados de tráfego real).

A utilização da reserva de recursos e a probabilidade de bloqueio são apresentadas na Figura 9-13 (a e b). Os resultados mostram que um sistema com agregação por área permite obter uma utilização maior dos recursos e uma menor probabilidade de bloqueio em comparação com agregação extremo-a-extremo. Como exemplo, para atingir uma utilização superior a 90%, o *bulk* deve ser inferior a 800 Kb/seg em agregação por área com uma classe de serviço (com *bulks* fixos ou proporcionais), inferior a 400 Kb/seg em agregação por área com duas classes de serviço, e inferior a 300 Kb/seg em agregação extremo-a-extremo. Da mesma forma, para obter uma probabilidade de bloqueio sempre inferior a 7%, o *bulk* pode atingir 1.25 Mb/seg em agregação por área com *bulks* fixos, 800 Kb/seg em agregação por área com *bulks* proporcionais, 400 Kb/seg em agregação por área com duas classes de serviço, e apenas 250 Kb/seg em agregação extremo-a-extremo. Para estes *bulks*, e observando a Figura 9-12, o ganho de sinalização dos *routers* interiores é de 150 com agregação extremo-a-extremo e de 1500 com agregação por área, uma classe de serviço e *bulks* proporcionais. Estes ganhos mostram claramente que a agregação por área reduz significativamente os requisitos de processamento (e o custo) dos *routers* interiores.



(a) **(b)**

Figura 9-13: (a) Utilização da reserva dos recursos e (b) probabilidade de bloqueio (agregados de tráfego real).

Os resultados destes estudos mostram que agregação por área, em comparação com agregação extremo-a-extremo, permite obter maiores ganhos de sinalização, maior utilização dos recursos e menor probabilidade de bloqueio (ou de sobrecarga), mesmo para redes relativamente pequenas. Mostram também que a utilização de *bulks* proporcionais aumenta os ganhos de sinalização com um impacto muito pequeno na utilização. Nas

próximas duas secções, será estudado o desempenho da agregação em domínios de rede de maior dimensão.

9.4.3.2 Rede de acesso

Nesta secção considera-se uma rede de acesso com uma topologia em árvore, ilustrada na Figura 9-14. O domínio é constituído por 8 áreas, e cada área, à excepção da área da direita, contém dois pares de ABRs. Existem 8 sessões a percorrer o domínio, cada uma com origem num DBR da esquerda diferente e com destino no mesmo DBR da direita. Assume-se que todas as sessões pertencem à mesma classe de serviço. O número de *routers* dentro de cada área é de 4. Assim, cada sessão atravessa 25 *routers* (sem incluir o *router* de saída do domínio). Como todas as sessões têm pares origem/destino diferentes e o domínio suporta apenas uma classe de serviço, existe um agregado por sessão em agregação extremo-a-extremo e um agregado em cada par de ABRs em agregação por área. A largura de banda de cada par de ABRs está representada na figura em Mb/seg. No sistema com agregação por área consideram-se os casos de *bulks* fixos e de *bulks* proporcionais ao tráfego oferecido aos agregados. Neste último caso, o *bulk* da área central é o dobro do *bulk* das 6 áreas da esquerda e o *bulk* da área da direita é o quádruplo do *bulk* das 6 áreas da esquerda.

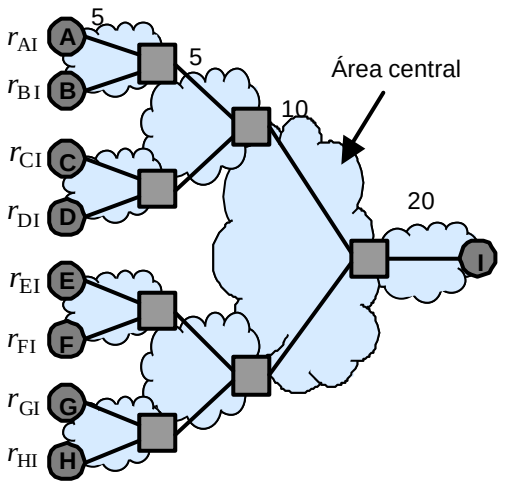


Figura 9-14: Topologia em árvore – rede de acesso.

Os ganhos de sinalização dos *routers* interiores e de todos os *routers* do domínio são apresentados na Figura 9-15. Os ganhos dos *routers* interiores são bastante elevados, aproximadamente de 2000 em ambas as situações de agregação por área. Existe também um aumento nos ganhos de sinalização quando se consideram *bulks* proporcionais. Note-se que, considerando todas as tentativas de reservas e *bulks* de 1.25 Mb/seg, os ganhos obtidos em agregação por área são aproximadamente 4 vezes superiores aos obtidos com agregação extremo-a-extremo. Os ganhos no domínio (considerando apenas as tentativas de reservas com sucesso) tendem para 25 em agregação extremo-a-extremo, pois apenas 1 *router* (o DBR de entrada) processa mensagens de sinalização fluxo-a-fluxo, e tendem para 5 em agregação por área, pois 5 *routers* (o DBR de entrada e os ABRs no percurso das sessões) processam mensagens de sinalização fluxo-a-fluxo.

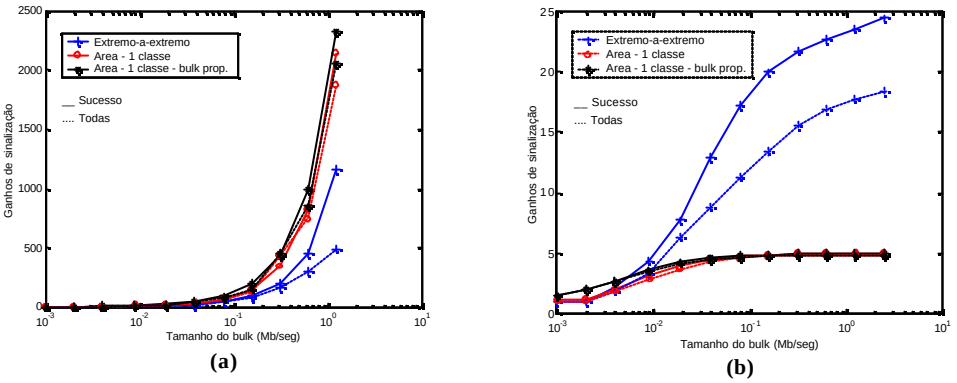
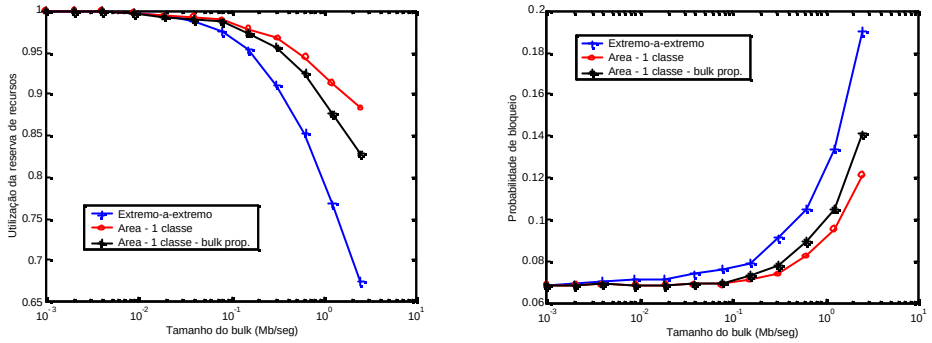


Figura 9-15: Ganhos de sinalização dos (a) *routers* interiores e de (b) todos os *routers* do domínio (rede de acesso).

A Figura 9-16 apresenta a utilização da reserva de recursos e a probabilidade de bloqueio. Considerando um sistema com agregação por área e *bulks* fixos, a utilização das reservas é sempre superior a 88%, e a probabilidade de bloqueio é sempre inferior a 12%. Note-se que, na área da direita, 8 sessões partilham um mesmo agregado em agregação por área, o que explica estes valores elevados de utilização. Um sistema com agregação por área e *bulks* proporcionais apresenta uma utilização ligeiramente inferior (e uma probabilidade de bloqueio ligeiramente superior), comparativamente ao caso de *bulks* fixos, mas bastante superior à utilização obtida com agregação extremo-a-extremo. Como exemplo, para garantir uma utilização sempre superior a 92%, o *bulk* pode atingir 1.25 Mb/seg em agregação por área com *bulks* fixos, e apenas 312 Kb/seg em agregação

extremo-a-extremo. Da mesma forma, para obter uma probabilidade de bloqueio sempre inferior a 10%, o *bulk* pode atingir 1.25 Mb/seg em agregação por área com *bulks* fixos, e apenas 350 Kb/seg em agregação extremo-a-extremo.



(a) **(b)**

Figura 9-16: (a) Utilização da reserva dos recursos e (b) probabilidade de bloqueio (rede de acesso).

9.4.3.3 Rede do núcleo

A rede do núcleo é baseada na topologia *Dumbbell* ilustrada na Figura 9-17. O domínio contém 5 áreas, com 2 pares de ABRs em cada área periférica e um par de ABRs na área central. O número de *routers* dentro de cada área é também de 4. Existem 16 sessões a atravessar o domínio: cada DBR da esquerda é origem de uma sessão com destino em cada DBR da direita. Assim como no caso da rede de acesso, existe um agregado por sessão em agregação extremo-a-extremo e um agregado em cada par de ABRs em agregação por área. Note-se que todas as 16 sessões atravessam a área central. A largura de banda de cada par de ABRs está representada na figura em Mb/seg. Em agregação por área consideram-se também os casos de *bulks* fixos e proporcionais. Neste caso, o *bulk* na área central é o quádruplo do *bulk* nas áreas periféricas.

Nesta topologia os ganhos de sinalização apresentados na Figura 9-18 são, uma vez mais, bastante elevados nos *routers* interiores, sendo significativamente maiores quando se consideram *bulk* proporcionais.

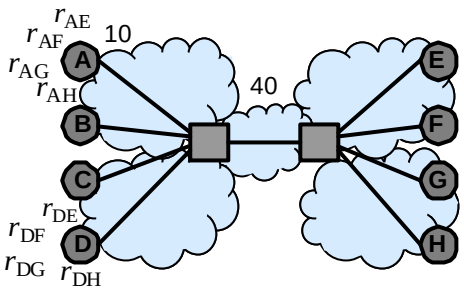


Figura 9-17: Topologia *Dumbbell* – rede do núcleo.

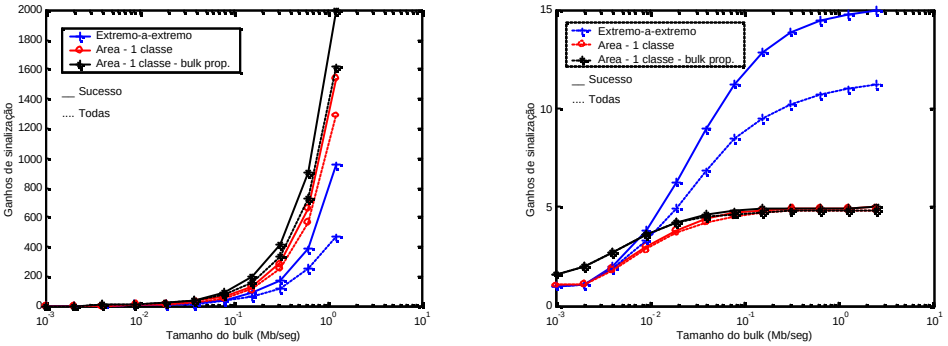
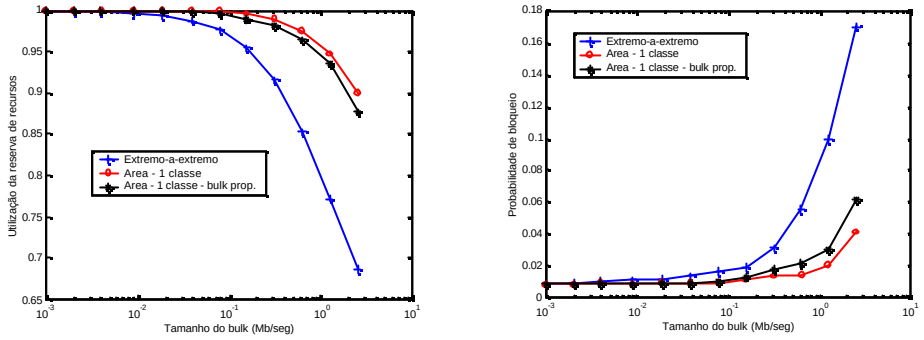


Figura 9-18: Ganhos de sinalização dos (a) *routers* interiores e de (b) todos os *routers* do domínio (rede do núcleo).

A utilização das reservas e a probabilidade de bloqueio são apresentadas na Figura 9-19. As diferenças entre as curvas correspondentes a agregação extremo-a-extremo e agregação por área são muito elevadas. A utilização obtida com agregação por área e *bulks* fixos é sempre superior a 90%; a probabilidade de bloqueio é sempre inferior a 4%. Além disso, a utilização de *bulks* proporcionais não degradam a utilização e a probabilidade de bloqueio. A partilha de recursos elevada em agregação por área permite atingir uma utilização das reservas sempre superior a 95% com *bulks* de 1.25 Mb/seg em agregação por área, enquanto que em agregação extremo-a-extremo o *bulk* pode atingir apenas os 150 Kb/seg.

Os resultados obtidos em redes de acesso e de núcleo mostram que os ganhos de desempenho globais de agregação por área em relação a agregação extremo-a-extremo aumentam com o tamanho da rede e podem atingir valores muito elevados.



(a) **(b)**

Figura 9-19: (a) Utilização da reserva dos recursos e (b) probabilidade de bloqueio (rede do núcleo).

9.5 Conclusões e trabalho futuro

Este capítulo abordou a utilização de agregação em domínios que podem ser estruturados em áreas. Consideraram-se duas formas de configuração dos agregados: agregados configurados entre os *routers* fronteira do domínio, denominados de agregados extremo-a-extremo, e agregados configurados entre os *routers* fronteira das áreas, denominados de agregados de área. Foram desenvolvidos dois modelos analíticos diferentes para estudar os compromissos entre carga de sinalização e utilização das reservas de recursos. O primeiro modelo, designado por modelo de carga por fluxo, é baseado em processos de nascimento e morte multi-dimensionais. Neste modelo, a carga oferecida é detalhada ao nível do fluxo, permitindo determinar a carga de sinalização com precisão. O segundo modelo analítico, designado por modelo de carga por agregado, considera tráfego oferecido variante no tempo, permitindo analisar os compromissos entre a taxa de variação do tráfego oferecido e a taxa de actualizações do agregado.

Os resultados obtidos com o modelo de carga por fluxo mostram que a carga de sinalização diminui significativamente quando é considerada a agregação de reservas individuais. Os ganhos de sinalização dos *routers* situados entre o *router* de entrada e de saída dos agregados (*routers* interiores), em cada um dos tipos de agregação, dependem da partilha de recursos existente entre sessões, do número de fluxos admitidos no sistema, e da proximidade entre o tamanho dos *bulks* e a largura de banda dos agregados. Para tamanhos do *bulk* elevados, os ganhos de sinalização dos *routers* interiores num sistema

com agregação extremo-a-extremo são maiores, devido ao tamanho do *bulk* ser de tal forma elevado, que não permite a existência de actualizações dos agregados. Os ganhos de sinalização do domínio são menores em agregação por área, pois todos os *routers* de entrada do agregado processam mensagens de sinalização fluxo-a-fluxo. A utilização das reservas é maior em agregação por área, pois esta permite partilha de recursos entre sessões num mesmo agregado. Verificou-se também que uma solução com *bulks* proporcionais ao tráfego oferecido permite obter ganhos de sinalização muito elevados com um impacto quase nulo na utilização de reservas, relativamente a uma situação com agregados extremo-a-extremo.

O modelo de carga por agregado tem a vantagem de permitir estudar o efeito de tráfego oferecido variante no tempo. Para intervalos de tempo entre actualizações de reservas pequenos, verificou-se que o desempenho do sistema é semelhante com qualquer tipo de agregação porque existe partilha de recursos entre agregados, em agregação extremo-a-extremo, e entre sessões no mesmo agregado, em agregação por área. Este caso é semelhante a utilizar um pequeno *bulk* no modelo de carga por fluxo. Para intervalos entre reservas maiores, a partilha de recursos em agregação extremo-a-extremo é muito limitada, e as diferenças entre os dois tipos de agregação são muito maiores. Confirmam-se mais uma vez as vantagens de agregação por área em relação a agregação extremo-a-extremo.

Para complementar os resultados obtidos com os modelos analíticos, realizaram-se conjuntos de simulações de eventos discretos usando agregados reais como tráfego oferecido. Os resultados obtidos reforçam as vantagens de agregação por área. Quando se consideraram domínios de maior dimensão, verificou-se que a utilização das reservas obtida com agregação por área, para *bulks* com largura de banda 100 vezes superior à largura de banda média dos fluxos, torna-se próxima da utilização obtida num sistema com reservas de recursos fluxo-a-fluxo (90% da utilização obtida num sistema sem agregação). Além disso, a carga de sinalização dos *routers* interiores é reduzida de aproximadamente 2000 vezes. Isto significa que, com agregação por área, o custo dos *routers* interiores pode ser significativamente menor. No entanto, note-se também que o número de *routers* que necessitam de processar sinalização fluxo-a-fluxo aumenta com o aumento do número de áreas. A escolha da configuração a adoptar depende, portanto, das características particulares de cada caso.

Face às condições particulares de uma rede, torna-se necessário determinar onde se deve agregar e desagregar o tráfego, isto é, determinar a topologia de rede lógica de agregados que optimiza a utilização dos recursos da rede e a capacidade de processamento (e custo) de cada elemento de rede. Este estudo de optimização é um tópico em aberto para investigação futura. Outro tópico em aberto é a determinação do *bulk* óptimo que minimize a carga de sinalização e maximize a utilização da reserva de recursos.

Formatted: Bullets and Numbering

CAPÍTULO 10

CONCLUSÕES

Neste capítulo são apresentadas as principais conclusões do trabalho efectuado e descrito nesta Tese. São também apresentados os tópicos em aberto, e que se consideram importantes como propostas de trabalho futuro.

10.1 Principais conclusões

Esta Tese abordou a problemática da gestão de recursos em redes ATM (*Asynchronous Transfer Mode*) e IP (*Internet Protocol*) com suporte de QdS (Qualidade de Serviço). A primeira parte concentrou-se nas redes de acesso. Mais concretamente, foram propostas estratégias de gestão de recursos com base em VPs (*Virtual Path*) em redes de acesso ATM, e foram propostas metodologias de dimensionamento das redes de acesso ATM. Foi também proposta uma rede de acesso IP, com suporte de serviços multimédia e QdS diferenciada. A segunda parte incidiu em dois mecanismos de controle de admissão escaláveis (que minimizam o processamento e o estado armazenado nos elementos da rede). Em particular, foi proposto um novo mecanismo de *probing* da rede, designado por ?-*probing*, que atenua o roubo de recursos existente em sistemas com múltiplas classes de serviço. Finalmente, em sistemas que efectuam a agregação de reservas individuais, foi

estudada a hierarquização de domínios de grandes dimensões, para permitir aumentar a utilização dos recursos da rede sem penalizar a carga de sinalização.

Nos capítulos 2 a 4 foi realizado o levantamento do estado-da-arte dos diferentes problemas endereçados na Tese.

No capítulo 5 foi apresentado um estudo de dimensionamento de redes de acesso ATM, constituídas por numa rede de distribuição de fibra óptica (PON) com uma OLT (*Optical Line Termination*) que realiza a interface com a rede de transporte, e com ONUs (*Optical Network Units*) nas terminações dos clientes. A estratégia de gestão de recursos escolhida tem um grande impacto na largura de banda necessária em cada VPC (*Virtual Path Connection*) e em cada ligação. Das estratégias apresentadas, aquela que permite a maior partilha de recursos da rede é a que considera que a OLT (*Optical Line Termination*) possui funcionalidades de CAC (*Call Admission Control*) e de UPC (*Usage Parameter Control*). Deste modo, a OLT pode terminar e iniciar VPCs. Esta estratégia economiza em largura de banda, principalmente no número de contentores VC-4 (*Virtual Containers*) necessários no anel SDH (*Synchronous Digital Hierarchy*), mas aumenta a complexidade da OLT e, consequentemente, o seu custo. Nas outras estratégias apresentadas, a carga de sinalização na OLT é menor, mas a partilha de recursos existente é mais limitada. Existe assim um compromisso entre economia de largura de banda e carga de sinalização nos elementos da rede de acesso. O operador da rede deverá ponderar os prós e os contras de cada estratégia e decidir qual a mais vantajosa no seu caso concreto.

O ATM permite a existência de multiplexagem estatística entre as células do tráfego agregado em cada VPC. Se o dimensionamento da rede tiver em conta esta multiplexagem, é possível reduzir significativamente a largura de banda necessária em cada ligação, e consequentemente, o número de módulos adicionais necessários na rede de acesso e de módulos VC-4 do anel SDH. O dimensionamento considerou o caso geral em que o número de classes de serviço em cada VPC pode ser ou não unitário. No caso de se considerar mais do que uma classe de serviço num VPC, a partilha de recursos é maior, pois o VPC contém uma maior quantidade de tráfego. No entanto, caso os parâmetros de QdS dos serviços agrupados no mesmo VPC sejam muito distintos, é necessário atribuir a todos os serviços a QdS mais estrita do conjunto. Nesta situação, a largura de banda necessária em cada ligação aumenta e pode ser superior à necessária considerando segregação de serviços em diferentes VPCs.

O capítulo 6 centrou-se na evolução das redes de acesso actuais, e propôs uma futura rede de acesso partilhada, utilizando o protocolo IP como única tecnologia de rede. A rede de acesso IP proposta foi definida ao nível dos equipamentos constituintes, interfaces físicas e protocolos e tecnologias suportados em cada elemento. A topologia da rede de acesso é reduzida a uma árvore, porque se considerou não ser necessário a redundância de percurso e de elementos na rede de acesso. A existência de apenas um percurso possível em direcção a cada utilizador, permite utilizar simplificações dos protocolos de encaminhamento existentes actualmente; o endereçamento no interior da rede de acesso é apenas necessário para se efectuar o encaminhamento dentro da rede: pode ser configurado apenas um endereço IP privado por elemento. Foi analisada a reutilização do PPP (*Point-to-Point Protocol*) nas redes de acesso e verificou-se que este tipo de solução apresenta um conjunto de problemas, dos quais se destaca a necessidade de estabelecimento de túneis para possibilitar o funcionamento do PPP numa rede partilhada.

A arquitectura para redes de acesso IP proposta recorre a protocolos e tecnologias que permitem introduzir funcionalidades de QdS na rede de acesso e integração de serviços e aplicações multimédia. Estes protocolos incluem: o SIP (*Session Initiation Protocol*) que permite estabelecer, gerir e terminar sessões multimédia sobre IP; o COPS (*Common Open Policy Service*) que permite gerir as políticas de QdS, fazendo também uso de protocolos de AAA (*Authentication, Authorization and Accounting*); e o RSVP (*resource ReSerVation Protocol*) que permite reservar recursos para cada sessão. A sincronização entre o estabelecimento de uma chamada e a reserva de recursos pode ser imperativa ou opcional. A introdução recente de uma extensão do RSVP para suportar agregação de fluxos individuais permite estabelecer compromissos entre custo e desempenho da rede de acesso. Assim, nesta arquitectura os elementos de rede têm de suportar as seguintes funções: todos os elementos necessitam de suportar o RSVP e incluir um elemento que implemente as políticas de QdS (designado por PEP - *Policy Enforcement Point*); o BAS contém um servidor *proxy* e tem acesso a um elemento que toma decisões de políticas de QdS (designado por PDP - *Policy Decision Point*), o qual contacta servidores de AAA dos ISPs (*Internet Service Provider*) e um repositório de políticas. Esta proposta de rede de acesso foi implementada em ambiente laboratorial. Os testes efectuados comprovam o correcto inter-funcionamento dos protocolos, e a possibilidade de atribuir QdS diferenciada a cada tipo de serviços.

Na segunda parte da Tese foram estudados dois mecanismos de controle de admissão de fluxos que permitem obter escalabilidade em redes IP. Ambos os mecanismos permitem gerir os recursos da rede sem a necessidade de manutenção do estado de todos os fluxos activos em cada elemento da rede. O mecanismo de *probing*, apresentado no capítulo 7, aceita ou rejeita um novo fluxo, com base na determinação do nível de QdS da rede através da inserção de fluxos de teste. O mecanismo baseado em agregação de reservas individuais, apresentado no capítulo 9, permite determinar se existem ou não recursos disponíveis na rede para aceitar um novo fluxo, com base nas reservas dos agregados.

O mecanismo de *probing* é muito simples porque não requer funcionalidades acrescidas nos elementos do núcleo da rede. Todo o processo de controle de admissão é efectuado nos extremos da comunicação (terminais ou *routers* de entrada e saída das redes de cliente). No entanto, para investigar o estado da rede, é necessário inserir fluxos de teste na rede, e por conseguinte, se esta se encontrar congestionada, estes fluxos vão degradar o nível de QdS que a rede está a fornecer. Além disso, os mecanismos de *probing* introduzem um problema de roubo de recursos em sistemas com múltiplas classes de serviço. Se o sistema tiver uma única classe de serviço, isto é, se não for afectado por roubo de recursos, a utilização dos recursos da rede é muito elevada (a menos do tráfego de *probing* que não conta para utilização), é possível obter um serviço de carga controlada, e é possível diferenciar entre várias classes de serviço.

O mecanismo baseado em agregação de reservas individuais requer funcionalidades de controle de admissão fluxo-a-fluxo e manutenção do estado de cada fluxo nos *routers* fronteira entre domínios (considerando que cada região de agregação é um domínio), e de manutenção do estado de cada agregado nos *routers* do núcleo. Este mecanismo permite implementar qualquer dos modelos de serviço das arquitecturas IntServ ou DiffServ. No entanto, para que a carga de sinalização e a complexidade dos *routers* seja muito pequena, a utilização dos recursos da rede pode ser muito limitada.

O trabalho realizado nesta Tese permitiu atenuar alguns dos problemas inerentes a estes mecanismos. Em relação ao primeiro caso, foi proposto um novo mecanismo de probing, o ?-*probing*, que permite, através da inspecção do estado de congestionamento de todas as classes de serviço, atenuar o roubo de recursos entre classes. Este mecanismo introduz, em simultâneo com os fluxos de *probing*, fluxos de ?-*probing* com largura de banda inferior à do fluxo que pede admissão para averiguar o efeito deste fluxo em todas as

outras classes de serviço. Os resultados obtidos (numéricos, de simulação e experimentais) permitem concluir que o mecanismo de ?-*probing* permite obter simultaneamente uma elevada utilização, diferenciação entre classes de serviço, e um modelo de serviço de carga controlada sem roubo de recursos. A largura de banda dos fluxos de ?-*probing* pode ser várias vezes inferior à largura de banda dos fluxos de *probing*, mas o comprimento dos seus pacotes deve ser semelhante ao comprimento médio dos pacotes de dados. Duas características importantes e limitativas destes mecanismos são a necessidade de um tempo elevado de estabelecimento das sessões e a impossibilidade de atingir perdas muito pequenas numa rede congestionada. As aplicações que utilizam estes mecanismos devem ser tolerantes ao tempo de estabelecimento e a algumas perdas.

Em relação ao mecanismo baseado em agregação apresentado no capítulo 9, foi estudada a hierarquização de domínios, ou seja, a sua divisão em áreas mais pequenas. Esta hierarquização permite aumentar a utilização dos recursos reservados. A configuração de agregados por áreas, em oposição aos agregados extremo-a-extremo no domínio, permite aumentar a quantidade de tráfego no mesmo agregado e aumentar a partilha de recursos. Para estudar estas duas formas de agregação, foram propostos dois modelos analíticos: um primeiro baseado em cadeias de *Markov*, e um segundo caracterizado por ter um tráfego oferecido variante no tempo. Os resultados obtidos com estes dois modelos e com estudos de simulação de eventos discretos, utilizando medidas de tráfego agregado reais, permitem concluir que agregação por área, em comparação com agregação extremo-a-extremo, permite obter simultaneamente uma diminuição da carga de sinalização dos elementos interiores às áreas, e uma diminuição no bloqueio dos fluxos, aumentando a utilização dos recursos reservados. O problema de agregação por área é o aumento da carga de sinalização nos *routers* fronteira entre áreas, pois estes devem efectuar sinalização fluxo-a-fluxo. Consequentemente, o custo, complexidade e capacidade dos elementos fronteira entre áreas aumenta mas, em contrapartida, os requisitos necessários dos elementos interiores diminuem significativamente.

10.2 Sugestões para trabalho futuro

Nesta secção são apresentados os tópicos em aberto, não abordados nesta Tese, e que se consideram importantes como propostas de trabalho futuro.

À data da escrita desta Tese, a rede de acesso IP com suporte de QdS encontra-se em fase de implementação em laboratório. A interacção entre os protocolos SIP e RSVP foi implementada com sucesso. No entanto, para replicar a rede proposta, é necessário implementar um servidor de políticas e as funções de AAA baseadas nas políticas de QdS do utilizador/serviço. A configuração dos elementos de rede e o *download* das políticas de QdS para estes elementos, de uma forma automática através do COPS, é de extrema importância para comprovar o funcionamento da rede. Um outro ponto importante para trabalho futuro é a extensão para IPv6 da arquitectura proposta nesta Tese.

Em relação aos mecanismos de *probing*, o trabalho futuro proposto centra-se na realização de um estudo teórico mais abrangente, ou seja, na derivação de expressões teóricas para a utilização, para o roubo de recursos e para o rácio de perdas em função das características dos fluxos de dados, dos fluxos de *probing* e de ?-*probing*, do tempo de *probing*, e das características do sistema a analisar. Em relação à implementação laboratorial, encontra-se em falta a implementação dos mecanismos de *probing* com outros algoritmos de escalonamento, para averiguar os compromissos entre roubo de recursos e utilização dos recursos da rede. Finalmente, a implementação realizada apenas considerou a taxa de perdas para efeitos de decisão, sendo interessante implementar mecanismos de *probing* que se baseiem na marcação/remarcação de pacotes para detectar o nível de QdS da rede.

Em relação à agregação de reservas individuais, o trabalho proposto para investigação futura incide num estudo de optimização para poder determinar em que pontos da rede se deve agregar e desagregar o tráfego e para determinar o valor óptimo do *bulk* de actualização do agregado. Deste modo, poderão ser potenciados os recursos de rede, limitando o aumento da capacidade de processamento (e custo) dos elementos de rede.

ANEXO A

MÉTODOS DE CÁLCULO DE PROBABILIDADE DE BLOQUEIO

Neste anexo são descritos os métodos de cálculo da probabilidade de bloqueio das chamadas de uma sessão de tráfego pertencente a um dado serviço. Estes métodos são aplicados a casos específicos de dimensionamento de redes de acesso ao nível da chamada apresentados no capítulo 5. O método directo de *Knapsack* [Kaufman81] (secção A.2) considera que o sistema é constituído apenas por uma ligação. Os três métodos seguintes consideram uma rede com várias ligações e com uma topologia em árvore (secção A.3) ou com uma topologia genérica (secções A.4 e A.5).

A.1 *Knapsack* estocástico

Alguns dos métodos descritos neste anexo baseiam-se no *Knapsack* estocástico [Ross95]. O *Knapsack* estocástico é um recurso com C unidades, ao qual chegam objectos de K classes. Os objectos da classe k distinguem-se pelo seu tamanho, b_k, pela sua taxa de chegada, $?_k$, e pelo seu tempo médio de permanência no sistema exponencialmente distribuído, $1/?_k$. Assume-se que as chegadas se dão de acordo com um processo de *Poisson*. Ao longo deste anexo, os objectos da classe k serão denominados de sessões de tráfego k.

Considerando que n_k é o número de sessões de tráfego k no *Knapsack*, a quantidade total de recursos utilizados é dada por $\mathbf{b}.\mathbf{n}$, em que $\mathbf{b} := (b_1,...,b_K)$, $\mathbf{n} := (n_1,...,n_K)$ e

$$\mathbf{b}.\mathbf{n} := \sum_{k=1}^{K} b_k n_k \tag{A-1}$$

Neste caso, a probabilidade de bloqueio de uma sessão de tráfego k é dada por:

$$B_k = 1 - \frac{\sum_{\mathbf{n} \in \Omega_k} \prod_{j=1}^{K} \rho_j^{n_j} / n_j!}{\sum_{\mathbf{n} \in \Omega} \prod_{j=1}^{K} \rho_j^{n_j} / n_j!}, k = 1, \ldots, K \tag{A-2}$$

em que $\Omega_k = \Omega / \Omega_k$, e Ω e Ω_k são conjuntos que representam espaços de estados definidos da seguinte forma:

$$\Omega := \{\mathbf{n} \in \mathbb{N}^K : \mathbf{b}.\mathbf{n} \leq C\} \tag{A-3}$$

$$\Omega_k := \{\mathbf{n} \in \Omega : \mathbf{b}.\mathbf{n} > C - b_k\}, k = 1, \ldots, K \tag{A-4}$$

em que $\mathbb{N}$ é o conjunto de inteiros não negativos.

Quando os valores de C e K são significativos, os espaços de estados Ω e Ω_k são muito grandes, tornando impraticável o cálculo directo dos somatórios. Na secção seguinte é apresentado um método recursivo para determinar estas probabilidades de bloqueio.

A.2 Método directo de *Knapsack* aplicado a múltiplos serviços

O método directo de *Knapsack* [Kaufman81] determina o valor exacto da probabilidade de bloqueio, num sistema constituído apenas por um recurso com múltiplas sessões de tráfego, através de um algoritmo recursivo.

A probabilidade de bloqueio de chamadas pertencentes a sessões de tráfego k é dada por

$$B_k = \sum_{c=C-b_k+1}^{C} q(\rho_k, b_k, c), k = 1, \ldots, K \tag{A-5}$$

em que C é a capacidade do *Knapsack*, e $q(\rho_k, b_k, c)$ é a probabilidade de ocupação do *Knapsack* dada pela seguinte expressão:

$$q(\rho_k, b_k, c) = \frac{g(\rho_k, b_k, c)}{\sum_{c=0}^{C} g(\rho_k, b_k, c)}, k = 1, \ldots, K \tag{A-6}$$

com $g(\rho_k, b_k, c)$ obtida recursivamente pela seguinte expressão:

$$g(?_k, b_k, c) = \frac{1}{c} \sum_{k=1}^{K} b_k \, ?_k \, g(?_k, b_k, c - b_k) \qquad \text{(A-7)}$$

As condições iniciais de $g(?_k, b_k, c)$ são as seguintes: $g(?_k, b_k, c) = 1$ para $c = 0$, e $g(?_k, b_k, c) = 0$ para $c < 0$.

A.3 Algoritmo de convolução baseado no *Knapsack* estocástico num sistema de serviço único

O algoritmo de convolução baseado no *Knapsack* estocástico [Tsang90] permite calcular o valor exacto da probabilidade de bloqueio de cada sessão de tráfego numa rede com topologia em árvore. A rede consiste em K ligações de acesso com v_k canais e uma ligação comum com V canais (Figura A-1). Supõe-se que existem K sessões de tráfego, em que cada chamada da sessão k requer um canal na ligação de acesso k e um canal na ligação comum. Supõe-se que as chamadas da sessão de tráfego k chegam ao sistema de acordo com um processo de *Poisson*.

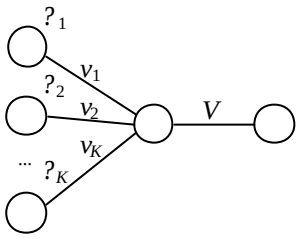


Figura A-1 : Rede com topologia em árvore

Assumindo todos estes pressupostos, a probabilidade de uma chamada pertencente a uma sessão de tráfego k ser bloqueada é obtida pela expressão:

$$B_k = G(V, v_k, ?) = \frac{g(V, ?) - g(v_k, ?) \cdot \sum_{v=0}^{V-v_k-1} g(k, v)}{\sum_{v=0}^{V} g(v)}, \; k = 1, \ldots, K \qquad \text{(A-8)}$$

em que

$$g_k(v, ?) = \begin{cases} ?^v_k / v!, & v = 0, \ldots, v_k \\ 0, & v > v_k \end{cases}, \; k = 1, \ldots, K \qquad \text{(A-9)}$$

$$g_{(k)}(v) = \rho_1 * \ldots * \rho_{k-1} * \rho_{k+1} * \ldots * \rho_K \quad (v) \tag{A-10}$$

$$\rho_k = [g_k(0), g_k(1), \ldots, g_k(V)], \; k = 1, \ldots, K \tag{A-11}$$

O símbolo $*$ representa o operador de convolução, $g_{(k)}(v)$ representa a convolução de todos os ρ_h excepto $h=k$, e $\rho_1 * \ldots * \rho_{k-1} * \rho_{k+1} * \ldots * \rho_K \quad (v)$ é o v-ésimo elemento resultante do vector de convolução.

A.4 Aproximação de carga reduzida num sistema de serviço único

Numa rede com uma topologia genérica diferente da topologia em árvore das redes de acesso, pode ser necessário utilizar métodos de cálculo aproximado das probabilidades de bloqueio. Um método possível é a aproximação de carga reduzida. Esta aproximação assume que o bloqueio é independente de ligação para ligação, e que o tráfego oferecido a uma ligação é reduzido de acordo com a sobrecarga sofrida pelas sessões que atravessam essa ligação, nas outras ligações atravessadas pelas mesmas sessões.

Considere-se que as chamadas pertencentes a uma sessão de tráfego k referente a um mesmo serviço ($b_k = 1$ para todas as sessões) chegam ao sistema de acordo com um processo de *Poisson*. Partindo do pressuposto que o bloqueio é independente de ligação para ligação, a intensidade de tráfego oferecido à ligação j pela sessão de tráfego k é aproximada por

$$\rho_k \prod_{h \in \pi_k, h \neq j} (1 - L_h) \tag{A-12}$$

em que π_k é o conjunto de ligações atravessadas pelo tráfego da sessão k, e L_h é a probabilidade de bloqueio de uma chamada na ligação h.

Assumindo que as sessões são independentes e que cada ligação pode ser partilhada por mais do que uma sessão de tráfego, a intensidade de tráfego total oferecido à ligação j é dada por

$$\sum_{k \in \sigma_j} \rho_k \prod_{h \in \pi_k, h \neq j} (1 - L_h) \tag{A-13}$$

em que σ_j é o conjunto das sessões de tráfego que atravessam a ligação j.

O valor aproximado da probabilidade de bloqueio de uma chamada na ligação j é dado pela fórmula de *Erlang B*:

$$L_j = ER\left(k \in \Omega_j\right) \sum_{h \in \Omega_k - j} \left(1 - L_h, v_j\right), \quad j = 1, 2, \ldots, J \tag{A-14}$$

em que v_j é o número de canais da ligação *j*. Este conjunto de *J* equações não lineares com *J* incógnitas pode ser resolvido pelo método de repetições sucessivas [Ross95].

O valor aproximado da probabilidade de bloqueio de uma chamada pertencente à sessão de tráfego *k* é obtida pela seguinte expressão:

$$B_k = 1 - \prod_{j \in \Omega_k} \left(1 - L_j\right), \quad k = 1, 2, \ldots, K \tag{A-15}$$

A.5 Aproximação de carga reduzida de *Knapsack* num sistema multi-serviço

Neste método descrito em [Chung93] considera-se uma rede com topologia genérica e com múltiplos serviços. Um sistema com múltiplos serviços representa, neste caso, um sistema em que as larguras de banda das sessões de tráfego dos diferentes serviços podem ser diferentes. Nesta aproximação assume-se, tal como no caso de um serviço único, que o bloqueio das chamadas pertencentes a uma sessão de tráfego *k* na ligação *j* é independente em cada ligação. Supõe-se que as chamadas da sessão *k* chegam ao sistema de acordo com um processo de *Poisson.*

Considerando que L_{jk} é o valor aproximado da probabilidade de bloqueio das chamadas da sessão *k* na ligação *j*, e atendendo ao pressuposto de independência do bloqueio das ligações assumido, a intensidade de tráfego da sessão *k* oferecido à ligação *j* é dada por

$$\rho_k \prod_{h \in \Omega_k - j} \left(1 - L_{hk}\right) \tag{A-16}$$

O valor das probabilidades de bloqueio L_{jk} são dadas por:

$$L_{jk} = Q_k\left(\rho_k \prod_{h \in \Omega_k - j} \left(1 - L_{hk}\right), v_j, b_k\right), \quad j = 1, \ldots, J, k = 1, \ldots, K \tag{A-17}$$

em que

$$Q_k\left(\rho_k, v_j, b_k\right) = 1 - \sum_{v=0}^{v_j - b_k} q\left(\rho_k, b_k, v\right) \tag{A-18}$$

é a probabilidade de bloqueio de uma sessão de tráfego k no *Knapsack*, e $q(?_k, b_k, v)$ é a probabilidade de ocupação do *Knapsack* definida na secção A.1.

Finalmente, o valor da probabilidade de bloqueio das chamadas da sessão de tráfego k é aproximado por

$$B_k \approx 1 - \prod_{j \in \mathcal{A}_k} (1 - L_{jk}) \qquad \text{(A-19)}$$

O sistema de equações obtido, em que L_{jk} são as incógnitas, pode ser também resolvido através do método de repetições sucessivas apresentado em [Ross95].

LISTA DE ACRÓNIMOS

AAA	*Authentication, Authorization and Accounting*
AAL	*ATM Adaptation Layer*
ABR	*Available Bit Rate*
ABR	*Area Border Router*
ACTS	*Advanced Communications Technology and Services*
ADSL	*Asymmetrical Digital Subscriber Line*
AF	*Assured Forwarding*
AS	*Access Server*
ATC	ATM *Transfer Capabilities*
ATM	*Asynchronous Transfer Mode*
AVP	*Attribute Value Pair*
BA	*Behavior Aggregate*
BAS	*Broadband Access Server*
BL	Banda Larga
BB	*Bandwidth Brokers*
BBL	*BroadBand Loop*
BGP	*Border Gateway Protocol*
BGRP	*Border Gateway Reservation Protocol*
BHCA	*Busy Hour Call Attempt*
BTI	*Broadband Trial Integration*
CAC	*Call Admission Control* – Controle de Admissão de Chamadas
CATV	*Community Antenna TeleVision*

CBQ	*Class-Based Queuing*
CBR	*Constant Bit Rate*
CDV	*Cell Delay Variation*
CDVT	CDV *Tolerance*
CER	*Cell Error Ratio*
CJVC	*Core-Jitter-Virtual Clock*
CLP	*Cell Loss Priority*
CLR	*Cell Loss Ratio*
CM	*Cable Modem*
CMTS	*Cable Modem Termination System*
CMR	*Cell Misinsertion Rate*
COPS	*Common Open Policy Service*
CTD	*Cell Transfer Delay*
DAVIC	*Digital AudioVIsual Council*
DBR	*Domain Border Router*
DHCP	*Dynamic Host Configuration Protocol*
DiffServ	*Differentiated Services* – Diferenciação de Serviços
DNS	*Domain Name System*
DOCSIS	*Data Over Cable System Interface Specification*
DPS	*Dynamic Packet State*
DRR	*Deficit Round Robin*
DSCP	DiffServ *CodePoint*
DSL	*Digital Subscriber Line*
DSLAM	*Digital Subscriber Line Access Multiplexer*
DVB	*Digital Video Broadcasting*
ECN	*Explicit Congestion Notification*
EF	*Expedited Forwarding*
ETS	*European Telecommunications Standard*
FDD	*Frequency Division Multiplexing*
FIFO	*First In First Out*
FITL	*Fiber Into The Loop*
FQ	*Fair Queuing*

FSAN	*Full Services Access Network*
FTP	*File Transfer Protocol*
FTTB	*Fiber To The Building*
FTTC	*Fiber To The Curb*
FTTCab	*Fiber To The Cabinet*
FTTH	*Fiber To The Home*
GCRA	*Generic Cell Rate Algorithm*
GFC	*Generic Flow Control*
GoS	*Grade of Service* – Grau de Serviço
GPS	*Generalized Processor Sharing*
HDLC	*High-Level Data Link Control*
HDSL	*High data rate* DSL
HEC	*Header Error Control*
HFC	*Hybrid Fiber-Coax*
ICMP	*Internet Control Message Protocol*
IETF	*Internet Engineering Task Force*
IntServ	*Integrated Services* – Integração de Serviços
IP	*Internet Protocol*
IPv4	*Internet Protocol version 4*
IPv6	*Internet Protocol version 6*
ISIS	*Intermediate System to Intermediate System Intra Domain Routing Exchange Protocol*
ISP	*Internet Service Provider*
ITU	*International Telecommunication Union*
ITU-T	ITU – *Telecommunication Standardization Sector*
IWU	*Inter-Working Unit*
JVC	*Jitter-Virtual Clock*
L2TP	*Layer 2 Tunneling Protocol*
LAN	*Local Area Network*
LDAP	*Lightweight Directory Access Protocol*
LMDS	*Local Multipoint Distribution Service*
LSP	*Label Switching Path*

LSR	*Label Switching Router*
MAC	*Media Access Control*
MAC	*Medium Access Control*
MBAC	*Measurement-Based Admission Control*
MBS	*Maximum Burst Size*
MCD	*Mean Call Duration*
MCR	*Minimum Cell Rate*
MF	*Multi-Field*
MMDP	*Markov Modulated Deterministic Process*
MPEG	*Moving Picture Experts Group*
MPLS	*Multi-Protocol Label Switching*
NAP	*Network Access Provider*
NAS	*Network Access Server*
NAT	*Network Address Translation*
NNI	*Network to Network Interface*
NRM	*Network Resource Management*
nrt-VBR	*non real time* VBR
ns	*network simulator*
NT	*Network Terminator*
ODN	*Optical Distribution Network*
OLT	*Optical Line Termination*
ONT	*Optical Network Termination*
ONU	*Optical Network Unit*
OSI	*Open Systems Interconnection*
OSPF	*Open Shortest Path First*
PBR	*Peak Bit Rate*
PCR	*Peak Cell Rate*
PDP	*Policy Decision Point*
PDU	*Packet Data Unit*
PEP	*Policy Enforcement Point*
PGPS	*Packet Generalized Processor Sharing*
PHB	*Per Hop Behavior*

PON	*Passive Optical Network*
PoP	*Point of Presence*
POTS	*Plain Old Telephony System*
PPP	*Point-to-Point Protocol*
PPPoE	*Point-to-Point Protocol over Ethernet*
PSTN	*Public System Telephone Network*
PT	*Payload Type*
QdS	Qualidade de Serviço
RADIUS	*Remote Access Dial In User Service*
RCC	*Return Control Channel*
RED	*Random Early Detection*
RDIS	Rede Digital com Integração de Serviços
RIO	RED *with In and Out*
RIP	*Routing Internet Protocol*
RR	*Round Robin*
RSVP	*resource ReSerVation Protocol*
rt-VBR	*real time* VBR
SA	*Security Association*
SCFQ	*Self-Clocking Fair Queuing*
SCM	*Sub-Carrier Multiplexing*
SCMA	*Sub-Carrier Multiple Access*
SCORE	*Scalable Core*
SCR	*Sustainable Cell Rate*
SDH	*Synchronous Digital Hierarchy*
SDL	*Simple Data Link*
SDP	*Session Description Protocol*
SECBR	*Severely Errored Cell Block Ratio*
SIP	*Session Initiation Protocol*
SLA	*Service Level Agreement*
SNMP	*Simple Network Management Protocol*
SOHO	*Small Office Home Office*
SP	*Service Provider*

STM	*Synchronous Transmission Module*
SVC	*Switched Virtual Connection*
TA	*Terminal Adapter*
TCA	*Traffic Conditioning Agreement*
TCP	*Transmission Control Protocol*
TDM	*Time Division Multiplexing*
TDMA	*Time Division Multiple Access*
TMForum	*TeleManagement Forum*
ToS	*Type of Service*
UBR	*Unspecified Bit Rate*
UDP	*User Datagram Protocol*
UNI	*User to Network Interface*
UPC	*Usage Parameter Control*
URL	*Uniform Resource Locator*
VBR	*Variable Bit Rate*
VC	*Virtual Clock*
VC	*Virtual Channel*
VCC	*Virtual Channel Connection*
VCI	*Virtual Channel Identifier*
VC-4	*Virtual Container*
VDSL	*Very high data rate* DSL
VoD	*Video on Demand*
VoIP	*Voice over IP*
VP	*Virtual Path*
VPC	*Virtual Path Connection*
VPI	*Virtual Path Identifier*
WFQ	*Weighted Fair Queuing*
WRR	*Weighted Round Robin*
WWW	*World Wide Web*

GLOSSÁRIO

C	Capacidade de uma ligação
c_k	Taxa mínima garantida a uma classe k
K	Número de sessões de tráfego
b_k	Largura de banda de cada chamada pertencente à sessão de tráfego k
$?_k$	Taxa de chegada das chamadas pertencentes à sessão de tráfego k
$??_k$	Tempo médio de permanência no sistema de uma chamada pertencente à sessão de tráfego k
$?_k$	Intensidade de tráfego pertencente à sessão k
B_k	Probabilidade de bloqueio de uma chamada/fluxo pertencente à sessão de tráfego k
L_j	Probabilidade de bloqueio da ligação j
v_j	Número de canais da ligação j
L_{jk}	Probabilidade de bloqueio de uma chamada pertencente a uma sessão de tráfego k na ligação j
M	Número de ONUs
N	Número de utilizadores em cada ONU
S	Número de serviços
?	Conjunto de serviços
$?$	Percentagem de tráfego gerado numa PON que tem como destino os utilizadores da mesma PON?

$?^s$	Penetração do serviço *s*
b^s	Largura de banda de cada chamada do serviço *s*
$?^s$	Taxa de chegada das chamadas de um serviço *s*
$???^s$	Tempo médio de permanência no sistema de uma chamada pertencente ao serviço *s*
$?^s$	Intensidade de tráfego de um serviço *s*
?	Conjunto das origens interiores à PON
?	Representação da OLT
?	Representação do exterior
$?^s_{lm}$	Intensidade do tráfego gerado em *l* com destino em *m*, pertencente ao serviço *s*
B^s_{lm}	Probabilidade de bloqueio das sessões de tráfego pertencentes ao serviço *s*, com origem em *l* e destino em *m*
L^s_{ij}	Probabilidade de bloqueio no VPC com origem em *i* e destino em *j*, que transporta sessões de tráfego do serviço *s*
$L^s_{ij,lm}$	Probabilidade de bloqueio da sessão de tráfego do serviço *s*, com origem em *l* e destino em *m*, no VPC com origem em *i* e destino em *j*
v^s_{lm}	Número de VCCs do VPC atravessado pela sessão de tráfego do serviço *s*, com origem em *l* e destino em *m*
v^s_{ij}	Número de VCCs do VPC com origem em *i* e destino em *j*, atravessado pelas sessões do serviço *s*
V_{ij}	Número de VCCs do conjunto dos VPCs nas ligações com origem em *i* e destino em *j*
c_{ij}	Capacidades dos VPCs com origem em *i* e destino em *j*
c_{lm}	Capacidade dos VPCs atravessados pelas sessões de tráfego com origem em *l* e destino em *m*
C_{ij}	Capacidade das ligações com origem em *i* e destino em *j*
t	Tempo
b	Taxa de um *leaky bucket*

p	Tamanho máximo de cada rajada de pacotes num *leaky bucket*
F_i^k	Instante de partida do pacote k na fila de espera i
l_i^k	Comprimento do pacote k que atravessa a fila de espera i
a	Número de nós atravessados por um pacote
$?_k$	Limiar de perdas da classe/sessão k
d_k	Limiar de atraso da classe/sessão k
U_k	Largura de banda utilizada na classe/sessão k
?	Conjunto das filas de espera
? '	Conjunto das filas de espera não vazias
$?_k$	Peso associado à classe/sessão k num algoritmo de escalonamento WFQ
r_k	Largura de banda do tráfego oferecido à classe/sessão k
$?_k$	Rácio de perdas da classe/sessão k
$?$	Probabilidade de perdas de pacotes no sistema
n_k	Número de fluxos da classe/sessão k no sistema
$\mathbf{b}$	$\mathbf{b} ? ?b_1,..., b_K ?$
$\mathbf{n}$	$\mathbf{n} ? ?n_1,..., n_K ?$
?	Espaço de estados obtido num sistema
$? _{AE}$	Espaço de estados obtido num sistema com algoritmo de escalonamento *AE*
$? _{AE}$	Espaço de transições obtido num sistema com algoritmo de escalonamento *AE*
$?_{\mathbf{n}}$	Probabilidade estacionária do estado $\mathbf{n}$
U	Utilização do sistema
p_{st}^{AE}	Probabilidade de roubo de recursos num sistema com algoritmo de escalonamento *AE*
$\mathbb{Z}^K$	Conjunto dos K conjuntos de inteiros não negativos
$?$	Parâmetro de *shape* de uma distribuição de Pareto
? ?	Conjunto dos pares de ABRs no domínio
J	Número de pares de ABRs no domínio
C_j	Capacidade do par de ABRs j

?	Conjunto de todos os agregados
H	Número de agregados
r_h	Largura de banda do agregado h
$?_h$	Percurso descrito pelos pares de ABRs que os agregados atravessam
a_h	Número de *routers* atravessados pelo agregado h
?	Conjunto das sessões de tráfego
K	Número de sessões de tráfego
$?_k$	Percurso descrito pelos agregados que a sessão k atravessa
$?_k$	Percurso descrito pelos pares de ABRs que a sessão k atravessa
r_k	Largura de banda agregada da sessão k
a_k	Número de *routers* atravessados pelo fluxo pertencente à sessão k
q_h	Largura de banda do *bulk* de actualização do agregado h
$?_h$	Conjunto de sessões que atravessam o agregado h
$?_j$	Conjunto dos agregados que atravessam o par de ABRs j
H_j	Número de agregados que atravessam o par de ABRs j
$\mathbf{n}_k^?$	Estado atingido a partir de $\mathbf{n}$ através do aumento em uma unidade do número de fluxos da sessão k
?	Conjunto de transições permitidas
?	Conjunto de transições proibidas
$I_h^{?\mathbf{n},\mathbf{n}_k^??}$	Função indicadora de actualização ou não da reserva do agregado h aquando da transição do estado $\mathbf{n}$ para $\mathbf{n}_k^?$
$?_A^B$	Taxa de mensagens de sinalização num sistema com agregação, em que A representa os *routers* que processam as mensagens e B representa o tipo de mensagens consideradas
$?_{A,ref}^B$	Taxa de mensagens de sinalização num sistema com sinalização fluxo-a-fluxo, em que A representa os *routers* que processam as mensagens e B representa o tipo de mensagens consideradas
G_A^B	Ganhos de sinalização entre $?_{A,ref}^B$ e $?_A^B$
$?_k$	Espaço de estados que podem admitir pelo menos mais um fluxo da sessão k

f_k	Largura de banda média do tráfego agregado da sessão k
e_k	Amplitude da sinusóide do tráfego agregado da sessão k
T	Período da sinusóide do tráfego agregado da sessão k
$?_k$	Fase aleatória uniformemente distribuída do tráfego agregado da sessão k
$??$	Duração do intervalo de tempo de actualização da largura de banda do agregado
x_h	Intervalo de tempo de actualização da largura de banda do agregado h
P_k	Probabilidade de sobrecarga de uma sessão k
T_j	Probabilidade de sobrecarga no par de ABRs j
r_{h,x_h}	Largura de banda máxima oferecida no intervalo x_h ao agregado h
$\hat{r}_j$	Carga reduzida oferecida ao par de ABRs j
$\hat{r}_{j,x_h}$	Largura de banda que se pretende reservar no intervalo de tempo x_h e no par de ABRs j

REFERÊNCIAS

[AForum10] ATM Forum, af-uni-0010.002, "ATM User-Network Interface Specification V3.1".

[AForum56] ATM Forum, af-tm-0056.000, "Traffic Management 4.0".

[Além01] R. Além e M. Carmo, "Mecanismos de *Probing* para Controlo de Admissão de Chamadas em Redes IP com QoS", Relatório de projecto em Eng. Electrónica e de Telecomunicações sob a orientação do Prof. Associado Rui Valadas com a colaboração da Eng. Susana Sargento e Eng. Victor Marques, Universidade de Aveiro, Setembro 2001.

[Andersen97] N. Andersen et al, "BBL: A Full-Service Access Network for Residential and Small Business Users", *IEEE Communications Magazine*, vol. 35 pp.88-93, 1997.

[Azcorra98] A. Azcorra et al., "Broadband Trial Integration", In *Proceedings of IDC'98*, Setembro 1998.

[Baker01] F. Baker, C. Iturralde, F. Faucheur e B. Davie, "Aggregation of RSVP for IPv4 and IPv6 Reservations", IETF RFC 3175, Setembro 2001.

[Basilier01] H. Basilier et al., "AAA Requirements for IP Telephony/Multimedia", IETF *Internet Draft* <draft-calhoun-sip-aaa-reqs- 03.txt>, October 2001.

[Benameur01] N. Benameur, S. Fredj, S. Oueslati-Boulahia e J. Roberts, "Integrated Admission Control for Streaming and Elastic Traffic", In *Proceedings of QofIS 2001*, Coimbra, Portugal, Setembro 2001.

[Bernet98] Y. Bernet, R. Yavatkar, P. Ford, F. Baker e L. Zhang, "A *Frame*work for End-to-End QoS Combining RSVP/IntServ and Differentiated Services", *IETF Internet Draft*, Março 1998.

[Bianchi00] G. Bianchi, A. Capone e C. Petrioli, "Throughput Analysis of End-to-End Measurement-Based Admission Control in IP", In *Proceedings of IEEE INFOCOM 2000*, Março 2000.

[Bound01] J. Bound et al., "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", IETF *Internet Draft* < draft-ietf-dhc-dhcpv6-21.txt>, 2001.

[Boyle00] J. Boyle et al., "The COPS (Common Open Policy Service) Protocol", IETF RFC 2748, 2000.

[Braden94] R. Braden, D. Clark e S. Shenker, "Integrated Services in the Internet Architecture: An Overview", RFC 1633, *IETF*, Junho 1994.

[Braden 97] R. Braden, L. Zhang, S. Berson, S. Herzog e S. Jamin, "Resource Reservation Protocol (RSVP) – Version 1 Functional Specification, RFC 2205, *IETF*, Setembro 1997.

[Breslau00a] L. Breslau, S. Jamin e S. Shenker, "Comments on the Performance of Measurement-Based Admission Control", In *Proceedings of IEEE INFOCOM 2000*, Março 2000.

[Breslau00b] L. Breslau, E. Knightly, S. Shenker, I. Stoica e H. Zhang, "Endpoint Admission Control: Architectural Issues and Performance", In *Proceedings of ACM SIGCOMM 2000*, Agosto 2000.

[Calhoun01] P. Calhoun et al., "DIAMETER Base Protocol". IETF *Internet Draft* <draft-ietf-aaa-diameter-02.txt>, 2001.

[Carlson98] M. Carlson, W. Weiss, S. Blake, Z. Wang, D. Black e E. Davies, "An Architecture for Differentiated Services", RFC 2475, *IETF*, Dezembro 1998.

[Case90] J. Case et al., "Simple Network Management Protocol (SNMP)", IETF RFC 1157, Maio 1990.

[CCITT92a] CCITT Rec. I.311, "B-ISDN General Network Aspects", Junho 1992.

[CCITT92b] CCITT Rec. I.361, "B-ISDN ATM Layer Specification", Junho 1992.

[Cetinkaya00] C. Cetinkaya e E. Knightly, "Egress Admission Control", In *Proceedings of IEEE INFOCOM 2000*, Março 2000.

[Chung93] S. Chung and K. Ross, "Reduced Load Approximations for Multirate Loss Networks", *IEEE Transactions on Communications*, vol. 41, pp.1222-1231, 1993.

[Clark98] D. Clark e W. Fang, "Explicit Allocation of Best Effort Packet Delivery Service", *IEEE/ACM Transactions on Networking*, 6(4), 1998.

[DAVIC95] DAVIC, DAVIC 1.0 Specification Revision3.0, 1995.

[Deering98] S. Deeringe e R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", IETF RFC 2460, 1998.

[Droms97] R. Droms, "Dynamic Host Configuration Protocol", IETF RFC 2131, 1997.

[Elek00] V. Elek, G. Karlsson, e R. Ronngren, "Admis sion Control Based on End-to-End Measurements", In *Proceedings of IEEE INFOCOM 2000,* Março 2000.

[Elwalid93] A. Elwalid e D. Mitra, "Effective Bandwith of General Markovian Traffic Sources and Admission Control of High Speed Networks", *IEEE/ACM Transactions on Networking*, Nº 3, pp. 329-343, Junho 1993.

[ETS-V5.2] "V5.2 Interface to Support the Access Network", European Telecommunication Standard (ETS) 300 347-1.

[Figueira95] N. Figueira e J. Pasquale, "An Upper Bound on Delay for the Virtual Clock Service Discipline", *IEEE/ACM Transactions on Networking*, 3(4), Abril 1995.

[Floyd93] S. Floyd e V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", *IEEE/ACM Transactions on Networking*, Agosto 1993.

[Floyd99] S. Floyd e K. Ramakrishnan, "A Proposal to Add Explicit Congestion Notification (ECN) to IP", *IEEE/ACM Transactions on Networking*, Maio 1999.

[Fu01] H. Fu e E. Knightly, "Aggregation and Scalable QoS: A Performance Study", In *Proceedings of IWQoS '01*, Junho 2001.

[Gall91] D. Le Gall, "*MPEG:* A Video Compression Standard for Multimedia Applications", *Communications of the ACM*, vol. 34,n. 4, 1991.

[Georgiadis96] L. Georgiadis, R. Guerin, V. Peris e K. Sivarajan, "Effective Network QoS Provisioning Based on Per Node Traffic Shaping", *IEEE/ACM Transactions on Networking*, 4(4), pp. 482-501, Agosto 1996.

[Gibbens99] R. Gibbens e F. Kelly, "Distributed Connection Acceptance Control for a Connectionless Network", In *Proceedings of ITC'99*, Junho 1999.

[Golestani94] S. Golestani, "A Self-Clocked Fair Queuing Scheme for Broadband Applications", In *Proceedings of IEEE INFOCOM'94*, Junho 1994.

[Handley98] M. Handley e V. Jacobson, "SDP: Session Description Protocol", IETF RFC 2327, 1998.

[Handley99] M. Handley et al., "SIP: Session Initiation Protocol", IETF RFC 2543, 1999.

[Heinanen99] J. Heinanen, F. Baker, W. Weiss e J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, *IETF*, Junho 1999.

[Huish96] P. Huish et al., "VDSL Copper Transport System", *in Proceedings of the Full Services Access Networks Conference*, Londres, Junho 1996.

[IETF-int] Endereço *Internet* do organismo IETF, www.ietf.org.

[Int-DOCSIS] Página *Internet* do DOCSIS, www.docsis.org.

[Int-LMDS] Página Internet do LMDS, http://www.lmdswireless.com.

[Int-NLANR] Página *Internet* da NLANR, http://moat.nlanr.net/Traces/Kiwitraces/auck2.html.

[Int-Qbone] Página do *Internet2 QoS Working Group*, http://www.internet2.edu/QoS/wg.

[Int-Vidconf] Endereço *Internet*: http://www.onvoy.com/pdf/videoconferencing.pdf.

[ITU-T] Endereço *Internet*: http://www.itu.org.

[ITU-T91] ITU-T, I.321, "B-ISDN Protocol Reference Model and its Application", Abril 1991.

[ITU-T93a] ITU-T, I.371, "Traffic Control and Congestion Control in B-ISDN", Março 1993.

[ITU-T93b] ITU-T, I.356, "B-ISDN ATM Layer Cell Transfer Performance", Outubro 1993.

[ITU-T93c] ITU-T, I.362, "B-ISDN ATM Adaptation Layer (AAL) Functional Description", Março 1993.

[ITU-T93d] ITU-T, I.363, "B-ISDN ATM Adaptation Layer (AAL) Specification", Março 1993.

[ITU-T95] ITU-T, I.361, "B-ISDN ATM Layer Specification", Outubro 1995.

[ITU-T98] ITU-T Rec. G.983.1, "Broadband Optical Access Systems Based on Passive Optical Networks (PON)", 1998.

[ITU-T99] ITU-T Rec. G.983.2, "ONT Management and Control Interface Specification for ATM-PON", 1999.

[ITU-T01] ITU-T Rec. G.983.3, "Broadband Optical Access Systems with Increased Service Capability by Wavelength Allocation", 2001.

[Jacobson99] V. Jacobson, K. Nichols, K. Poduri, "An Expedited Forwarding PHB", RFC 2598, *IETF*, Junho 1999.

[Katz01] D. Katz e D. Yeung, "Traffic Engineering Extensions to OSPF", IETF *Internet Draft* < draft-bitar-rao-ospf-diffserv-mpls-01.txt >, Julho 2001.

[Kaufman81] J. Kaufman, "Blocking in a shared resource environment", *IEEE Transactions on Communications*, 10, pp.1474-1481, 1981.

[Kelly86] F. Kelly, "Blocking Probabilities in Large Circuit-Switched Networks", *Advanced on Applied Probabilities*, vol. 18, pp.473-505, 1986.

[Kelly00] F. Kelly, P. Key e S. Zachary, "Distributed Admission Control", *IEEE Journal on Selected Areas in Communications*, Vol. 18, pp. 2617-2628, 2000.

[Kelly01] T. Kelly, "An ECN Probe-Based Connection Acceptance Control", Computer Communication Review, ACM SIGCOMM, pp. 14-25, Julho 2001.

[Keshav00] S. Keshav, "An Engineering Approach to Computer Networking – ATM Networks, the Internet and the Telephone Network", Addison-Wesley, 2000.

[Kesidis93] G. Kesidis, J. Walrand e C. Chang, "Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources", *IEEE/ACM Transactions on Networking*, Nº 4, Agosto 1993.

[Knightly99] E. Knightly e N. Shroff, "Admission Control for Statistical QoS: Theory and Practice", *IEEE Network*, Vol. 13(2), pp. 20-29, Março 1999.

[Kompella01] K. Kompella e Y. Rekhter, "LSP Hierarchy with MPLS TE", IETF *Internet Draft* <draft-ietf-mpls-lsp-hierarchy-03.txt>, Novembro 2001.

[Labourdett92] J. Labourdette and G. Hart, "Blocking Probabilities in Multitraffic Loss Systems: Insensivity, Asymptotic Behavior, and Approximations", *IEEE Transactions on Communications*, **40**, pp.1355-1366, 1992.

[Maeda01] Y. Maeda et al., "FSAN OAN-WG and Future Issues for Broadband Optical Access Networks", *IEEE Communications Magazine*, Dezembro 2001.

[Malkin98] G. Malkin, "RIP Version 2" IETF RFC 2453, 1998.

[Mamakos99] L. Mamakos et al., "A Method for Transmitting PPP Over Ethernet (PPPoE)", IETF RFC 2516, 1999.

[Marshall01] W. Marshall et al., "SIP Extensions for Resource Management", IETF *Internet Draft* <draft-ietf-sip-manyfolks-resource-02.txt>, 2001.

[Menendez97] R. Menendez et al., "Full Services Access Networks: Systems Engineering/Arquitecture (SE/A)", *in Proceedings of VIII International Workshop on Optical/Hybrid Access Networks*, Atlanta, GA, Março 1997.

[Mocci94] U. Mocci and C. Scoglio, "Traffic Clustering Rules in ATM Networks", *in Proceedings of the GLOBECOM'94*, pp. 783-787, 1994.

[Neilson99] R. Neilson, J. Wheeler, F. Reichmeyer e S. Hares, "A Discussion of Bandwidth Broker Requirements for Internet2 Qbone Deployment", Internet Qbone BB Advisory Council, Version 0.7, Agosto 1999.

[Nichols97] K. Nichols, V. Jacobson e L. Zhang, "A Two-Bit Differentiated Services Architecture for the Internet", *IETF Internet Draft*, Novembro 1997.

[Nichols98] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, *IETF*, Dezembro 1998.

[Nilsson99] A. Nilsson, M. Perry, A. Gersht and V. Iversen, "On Multi-rate Erlang-B Computations", in *Proceedings of ITC*, 1999.

[NS-2End] Endereço *Internet*: http://www.isi.edu/nsnam/ns/index.html

[Pan00] P. Pan, E. Hahne e H. Schulzrinne, "BGRP: A Tree-based Aggregation Protocol for Inter-Domain Reservations", *Journal of Communications and Networks*, 2(2), pp. 157-167, Junho 2000.

[Parekh92a] A. Parekh, “A Generalized Processor Sharing Approach to Flow Control in Integrated Services Network”, Pd.D. Thesis, Massachussets Institute of Technology, 1992.

[Parekh92b] A. Parekh e G. Gallager, “A Generalized Processor Sharing Approach to Flow Control – The Single Node Case”, In *Proceedings of IEEE INFOCOM’92*, Abril 1992.

[PONB97] G.PONB – Draft D, “ATM PON Specification”, Abril 1997.

[Postel94] J. Postel, “Domain Name System Structure and Delegation”, IETF RFC 1591, 1994.

[Qiu99] J. Qiu e E. Knightly, “Inter-Class Resource Sharing Using Statistical Service Envelopes”, In *Proceedings of IEEE INFOCOM’99*, Março 1999.

[QoSForum98] QoSForum, “Introduction to QoS Policies”, *White Paper*, 1998.

[Quayle97] J. Quayle et al, “Full Services Access Networks Requirements Specification”, *in Proceedings of VIII International Workshop on Optical/Hybrid Access Networks*, Atlanta, GA, Março 1997.

[Rekhter95] Y. Rekhter e T. Li, "A Border Gateway Protocol 4 (BGP-4)", IETF RFC 1771, Março 1995.

[Ricken98] C. Ricken, “DAVIC Cable Modem”, DAVIC 1.5 Specification, 1998.

[Rigney00] C. Rigney et al., Remote Authentication Dial In User Service (RADIUS)”, IETF RFC 2865, 2000.

[Rosenberg01] J.Rosenberg and H. Schulzrinne, “Reliability of Provisional Responses in SIP”, IETF *Internet Draft* <draft-ietf-sip-100rel-04.txt>, 2001.

[Ross95] K. Ross, “Multiservice Loss Models for Broadband Telecommunication Networks”, *Springer*, 1995.

[Ross97] K. Ross, “Introduction to probability models”, Academic Press, 1997.

[Salgado02] Roger Salgado et al., “A Demonstrator of an IP-based Access Network for Broadband Multimedia Services”, In *Proceedings of the 10th International Conference on Telecommunication Systems, Modeling and Analysis*, Monterey (Califórnia , USA), Outubro de 2002.

[Santiago02] Carlos Santiago, “Algoritmos de Escalonamento para Provisão de Qualidade de Serviço em Redes IP”, Dissertação de Mestrado,

Departamento de Electrónica e de Telecomunicações, Universidade de Aveiro, submetida, 2002.

[Sargento98] S. Sargento, A. Sousa e R. Valadas, ACTS 0038 – BroadbandLoop – *Deliverable* 1.2.7 – "Real Traffic Sources: Observation and Modelling", 1998.

[Sargento99a] Susana Sargento, Amaro Sousa e Rui Valadas, "Dimensioning Methodologies for the Broadband Loop Access Network", *in Proceedings of the NOC'99 (Networks and Optical Communications)*, pp. 85-93, Junho 1999.

[Sargento99b] Susana Sargento, Rui Valadas e Amaro Sousa, "Dimensioning Methodologies for the BBL (BroadbandLoop) and FSAN (Full Services Access Networks) Access Networks", *In Proceedings of BAC'99* (*Broadband Access Conference*), pp. 184-193, Outubro 1999.

[Sargento01a] Susana Sargento, Rui Valadas, Edward Knightly, "Call Admission Control in IP networks with QoS support", *In Proceedings of ConfTele'2001,* Figueira da Foz (Portugal), Abril 2001, pp. 567-571.

[Sargento01b] Susana Sargento, Rui Valadas, Jorge Gonçalves, Henrique Sousa, "IP Access Networks with QoS Support", *In Proceedings of SPIE ITCOM 2001 – "Technologies, Protocols, and Services for Next-Generation Internet"*, Denver (Colorado, USA), pp. 91-101, Agosto 2001.

[Sargento01c] Susana Sargento, Rui Valadas, Edward Knightly, "Resource Stealing in Endpoint Controlled Multi-class Networks", Artigo convidado *In Proceedings of IWDC 2001* (*International Workshop on Digital Communications*), Taormina (Itália), Setembro 2001.

[Sargento02a] S. Sargento et al., "An Experimental Study of Probing-Based Admission Control for DiffServ Architectures", In *Proceedings of Networking 2002,* Pisa (Itália), Maio 2002.

[Sargento02b] Susana Sargento e Rui Valadas, "Performance of Hierarchical Aggregation in Differentiated Services Networks", In *Proceedings of 10th International Conference on Telecommunication Systems, Modeling and Analysis*, Monterey (Califórnia , USA), Outubro 2002.

[Sargento02c] Susana Sargento e Rui Valadas, "Tradeoffs Between Signaling and Resource Utilization in DiffServ Networks with Hierarchical Aggregation", In *Proceedings of IDMS/PROMS'2002* (*Joint International Workshop on Interactive Distributed Multimedia Systems / Protocols for Multimedia Systems*), Coimbra (Portugal), Novembro 2002.

[Schelén98] O. Schelén e S. Pink, "Aggregation Resource Reservations over Multiple Routing Domains", In *Proceedings of IWQoS'98*, Napa, CA, Maio 1998.

[Schmitt99] J. Schmitt et al., "Aggregation of Guaranteed Service Flows", In *Proceedings of IWQoS'99*, Maio 1999.

[Schulzrinne99] H. Schulzrinne et al., "Interaction of Call Setup and Resource Reservation Protocols in Internet Telephony", Relatório técnico, 1999.

[Schulzrin02a] H. Schulzrinne, "DHCP Option for SIP Servers", IETF *Internet Draft* <draft-ietf-sip-dhcp-06.txt>, Março 2002.

[Shenker97] S. Shenker, C. Partridge e R. Guerin, Specification of Guaranteed Quality of Service", RFC 2212, *IETF*, Setembro 1997.

[Shreedhar95] M. Shreedhar e G. Varghese, "Efficient Fair Queuing Using Deficit Round Robin", In *Proceedings of ACM SIGCOMM'95*, Setembro 1995.

[Shroff98] N. Shroff e M. Schwartz, "Improved Loss Calculations at an ATM Multiplexer", *IEEE/ACM Transactions on Networking*, Vol. 6(4), pp. 411-422, Agosto 1998.

[Siebenhaar95] R. Siebenhaar, "Multiservice Call Blocking Approximations for Virtual Path Based ATM Networks with CBR and VBR Traffic", *in Proceedings of INFOCOM*, 1995.

[Sinnreich01] H. Sinnreich et al., "AAA usage for IP Telephony with QoS", IETF *Internet Draft* <draft-sinnreich-aaa-interdomain-sip-qos-osp-00.txt>, 2001.

[Simpson94] W. Simpson, "The Point-to-Point Protocol (PPP)", IETF RFC 1661, 1994.

[Skelly93] P. Skelly, M. Schwartz e S. Dixit, "A Histogram-based Model for Video Traffic Behavior in an ATM Multiplexer", *IEEE/ACM Transactions on Networking*, Vol. 1(4), pp. 446-459, Agosto 1993.

[Smit01] H. Smit e T. Li, "IS-IS Extensions for Traffic Engineering", IETF *Internet Draft* < draft-ietf-isis-traffic-04.txt >, Agosto 2001.

[Srisuresh01] P. Srisuresh e K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", IETF RFC 3022, 2001.

[Stoica98] I. Stoica, S. Shenker e H. Zhang, "Core-Stateless Fair Queuing: A Scalable Architecture to Approximate Fair Bandwidth Allocations in High Speed Networks", In *Proceedings of ACM SIGCOMM'98*, Agosto 1998.

[Stoica99] I. Stoica e H. Zhang, "Providing Guaranteed Services Without Per Flow Management", In *Proceedings of ACM SIGCOMM'99*, Setembro 1999.

[TMForum97] "Performance Reporting Definitions Document", *Network Management Forum, Issue* 1.0, Abril 1997.

[Townsley99] W. Townsley et al, "Layer Two Tunneling Protocol L2TP", IETF RFC 2661, 1999.

[Tsang90] D. Tsang e K. Ross, "Algorithms to Determine Exact Blocking Probabilities for Multirate Tree Networks", *IEEE Transactions on Communications, vol.* 38, pp.1266-1271, 1990.

[Valadas98] R. Valadas, "Dimensioning and Resource Management of ATM Networks", *in Proceedings of the 7th IFIP/ICCC Intl. Conf. on Information Networks and Data Communications*, Aveiro, Portugal, pp. 209-220, 1998.

[Wahl97] M. Wahl et al., "Lightweight Directory Access Protocol (v3)", IETF RFC 2251, 1997.

[Walrand98] J. Walrand, Comunicação Privada, 1998.

[Wang01] Z. Wang, "Internet QoS: Architectures and Mechanisms for Quality of Service", Morgan Kaufmann Publishers, 2001.

[Wroclaw97] J. Wroclawski, "Specification of the Controlled-Load Network Element Service", RFC 2211, *IETF*, Setembro 1997.

[Yavatkar00] R. Yavatkar, "A Framework for Policy-Admission Control", IETF RFC 2753, 2000.

[Zhang90] L. Zhang, "Virtual Clock: A New Traffic Control Algorithm for Packet Switching Networks", In *Proceedings of ACM SIGCOMM'90*, pp. 19-29, Setembro 1990.

[Zhang93] L. Zhang, S. Deering, D. Estrin, S. Shenker e D. Zappala, "RSVP: A New Resource ReSerVation Protocol", *IEEE Network*, pp. 8-18, Setembro 1993.