1
Introdução

1.1
Qualidade de serviço na Internet

Ao longo das duas últimas décadas do século passado, a Internet assumiu uma posição de destaque entre os meios de comunicação e consequentemente, passou a abrigar uma diversidade de aplicações para as quais não tinha sido projetada, como home banking, comércio eletrônico, ensino à distância, rádios, TV's, etc. Devido a isso, tem sido necessário buscar novas tecnologias para a comunicação entre computadores, que permitam um nível de desempenho aceitável para as diversas aplicações.

Parte destas tecnologias se destina a prover Qualidade de Serviço (QoS) na Internet, que tem sido um objetivo cada vez mais em evidência, dado o nível de importância que a Internet assumiu.

O desenvolvimento de tecnologias para integração de serviços e oferecimento de QoS em redes de comunicações começou com as redes ISDN(Integrated Services Digital Network), inicialmente em banda estreita. Posteriormente surgiram as redes de serviços integrados de banda larga ou B-ISDN (*Broadband ISDN*), cuja principal tecnologia é o ATM (*Asynchronous Transfer Mode*), que possibilita, pela sua característica de transmissão assíncrona de dados, um melhor aproveitamento dos recursos disponíveis além de apresentar uma estrutura projetada para fornecer suporte a QoS (ver Apêndice B).

Os princípios da B-ISDN podem ser resumidos nos seguintes itens:

- Negociações entre o usuário e a rede estabelecendo, da parte do usuário, as características do seu tráfego e, da parte da rede, as garantias de desempenho; esta negociação é formalizada em um contrato de serviço (SLA - *Service Level Agreement*).
- Provisão de recursos pela rede atendendo às garantias contratadas.

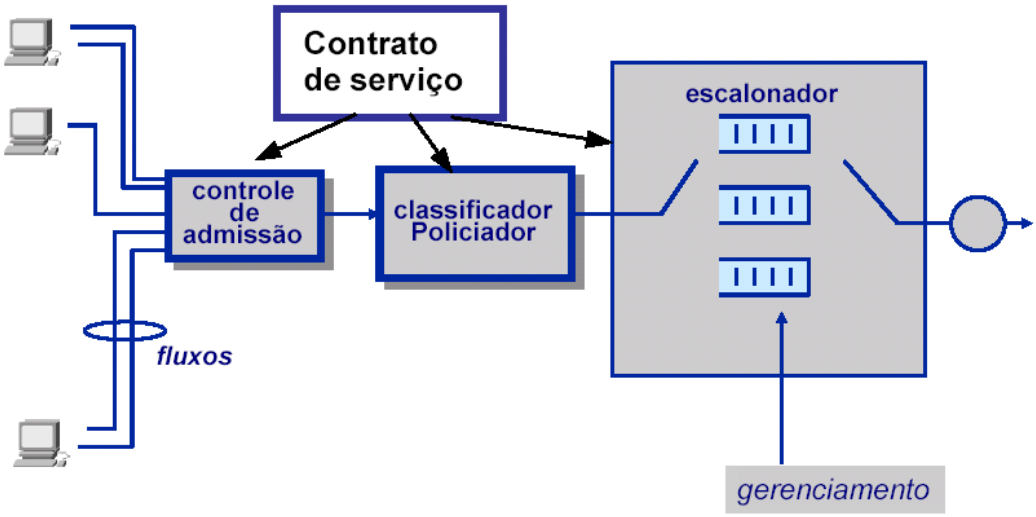

Figure 1.1: Mecanismos de controle de tráfego

- Controle de admissão de novas conexões condicionado à disponibilidade de recursos.
- Controle exercido pela rede para verificar a fidelidade do tráfego aos parâmetros estabelecidos no contrato. Este controle, denominado policiamento, está geralmente associado a mecanismos de marcação, ou seja, identificação da classe de tráfego através de algum campo do cabeçalho do pacote. Como consequência do policiamento, o pacote pode ser encaminhado, descartado, encaminhado com restrições ou retido temporariamente. Neste último caso ocorre o processo denominado moldagem de tráfego (*traffic shaping*).
- Controle exercido pela rede nos seus diversos nós para identificar o tipo de tráfego e dar tratamento diferenciado, de acordo com sua classe de serviço. O tratamento diferenciado é feito através do processo denominado escalonamento (ou agendamento), priorizando pacotes de classe superior no encaminhamento ao nó seguinte.

A Figura 1.1 apresenta um esquema ilustrativo identificando os mecanismos básicos de controle de tráfego descritos acima.

O ATM foi uma solução adotada sobretudo em redes públicas e redes privadas, com administração centralizada pelas operadoras. No entanto, a Internet é um ambiente extremamente heterogêneo, um grande sistema de comunicação formado por inúmeros subsistemas menores. Os subsistemas possuem escopos que podem variar desde uma comunidade local a subsistemas de alcance global ([4]), constituindo uma grande diversidade de tecnologias e modelos de serviço em uso. Essa multiplicidade de modelos dificulta a adoção de uma plataforma comum de provisão de Qualidade de Serviço, pois a garantia

fim-a-fim de parâmetros de transmissão implica que todos os subsistemas entre o usuário e a aplicação adotem um sistema homogêneo de gerenciamento de QoS, o que nem sempre ocorre. Devido a essa dificuldade, o nível de serviço oferecido na Internet é ainda de melhor esforço (*Best Effort*), sem diferenciação no tratamento dos pacotes e consequentemente, sem qualquer garantia real de desempenho.

Na atual geração da Internet, todos os pacotes IP, independentemente da aplicação a que pertençam, são processados de forma similar, sem especificação ou garantia de qualidade. A arquitetura de rede, baseada nos protocolos TCP/IP, impõe sérias limitações a um tráfego com restrições de tempo real, que considera atraso máximo dos pacotes fim-a-fim. O tráfego oriundo de uma aplicação sem requisitos de atraso, como e-mail, por exemplo, é tratado da mesma maneira que o tráfego de uma aplicação de tempo real, como video-conferência. Este problema agrava-se quando a rede está congestionada, pois os protocolos de congestionamento presentes no protocolo IP descartam pacotes indiscriminadamente.

Para que a Internet adquira uma estrutura de maior confiabilidade, com níveis de serviços e garantias fim-a-fim (garantias de parâmetros como por exemplo, taxa de perda máxima, atraso máximo de pacotes, etc) que possibilitem que as novas aplicações (transmissão de áudio e vídeo de alta qualidade, por exemplo) sejam satisfatoriamente suportadas, é preciso modificar a sua arquitetura de serviços, com tratamentos diferenciados para cada tipo de tráfego. Muitas propostas têm sido discutidas nesse sentido, visando desenvolver uma arquitetura que alie robustez e escalabilidade.

Há alguns anos, o IETF (*Internet Engineering Task Force*) apresentou a proposta da arquitetura de Serviços Integrados (*Integrated Services*) ou IntServ [21, 22, 23]. Segundo essa proposta, haveria um rigoroso controle de cada fluxo que trafegasse na rede através de protocolos de sinalização (particularmente, o **RSVP**), para que cada usuário recebesse um serviço dentro dos parâmetros de qualidade que foram acordados no momento da sua entrada na rede. Devido à grande carga de sinalização que circulava, esse modelo possibilitava garantias de QoS severas, ao custo de reserva de recursos e monitoração individualizada de cada fluxo que circulasse na rede, o que gerava um sério problema de escalabilidade, que crescia exponencialmente em complexidade, tornando inviável a sua implantação como arquitetura padrão na Internet.

Para contornar esse problema, propôs-se a arquitetura de Serviços Diferenciados (*Differentiated Services*) ou DiffServ [9, 10, 11], que traz o conceito de tratamento de fluxos agregados, e não individualizados. Ou seja, os pa-

cotes IP são marcados em seu cabeçalho como pertencentes a uma determinada classe e os nós do domínio DiffServ fazem o encaminhamento segundo os níveis de serviço pré-determinados para cada classe (*Per-hop Behavior* - PHB). Os agregados de fluxos formam grandes conjuntos chamados BA (*Behavior Aggregate*), aos quais é realizada a provisão de serviços sem reserva de recursos. Um BA é definido formalmente como "uma coleção de pacotes com o mesmo DSCP (*DiffServ Code Point*) que está cruzando um enlace em uma determinada direção em um domínio DS"[10]. Todos os pacotes de um BA recebem o mesmo tratamento de encaminhamento (PHB) nos roteadores por onde passam. A vantagem da agregação de fluxos em BA's é mais significativa em redes de grande porte, nas quais o número de usuários é elevado, tornando o controle de fluxos uma tarefa bastante complexa.

Ao contrário do procedimento realizado na arquitetura IntServ, em que as especificações dos serviços foram plenamente detalhadas, no DiffServ o IETF decidiu não padronizar os serviços, mas apenas definir duas classes de encaminhamento: o PHB EF (*Expedited Forwarding* [35]), com garantias de banda, atraso e jitter, e o PHB AF (*Assured Forwarding* [34]), com a diferenciação de serviços baseada em níveis de precedência para descarte de pacotes.

A adoção da arquitetura DiffServ traz dois aspectos principais que devem ser observados. Por um lado, ganha-se em escalabilidade devido ao tratamento de BA's ao invés de fluxos. Entretanto, as garantias obtíveis não são (a princípio) estritas, como no caso IntServ, visto que não há controle individualizado dos fluxos. No modelo DiffServ, trabalha-se com garantias estatísticas. Em geral, podemos considerar que há um compromisso (Figura 1.2) entre a qualidade das garantias oferecidas e a eficiência da utilização de recursos ([36]) em qualquer sistema de alocação de usuários com Qualidade de serviço. O desafio da implantação de QoS na Internet está em otimizar essa relação.

1.2 Motivação

Uma vez estabelecidas as arquiteturas, é necessário desenvolver os algoritmos e os protocolos para a sua implementação. Muitos esforços têm sido direcionados para a implementação dos mecanismos de Qualidade de Serviço do modelo DiffServ.

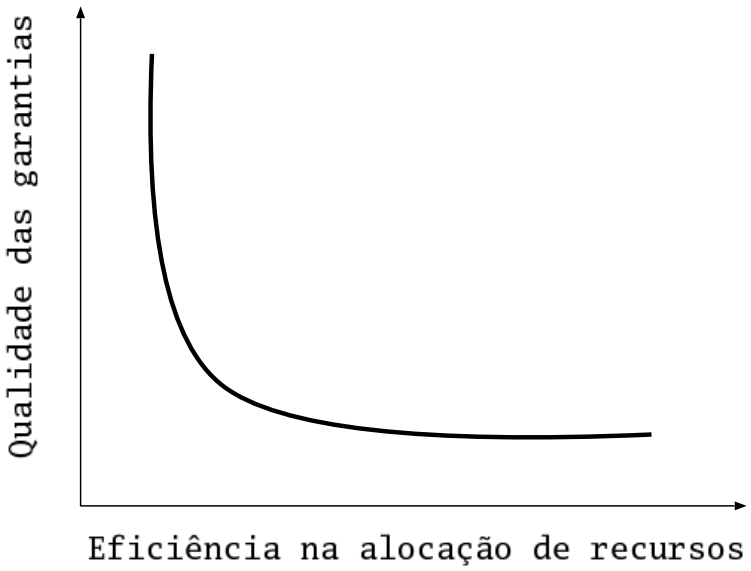

Figure 1.2: Qualidade X Eficiência

A grande dificuldade em prover QoS sob a arquitetura de Serviços Diferenciados é que esse modelo trabalha com garantias estatísticas, sem reserva de recursos aos usuários. Para obter-se garantias fim-a-fim, torna-se necessária a adoção de determinadas políticas nas extremidades da rede. Entre estas, uma de particular relevância é a disciplina de Controle de Admissão. A definição da arquitetura DiffServ não inclui esse mecanismo, mas considerando-se a sua grande importância para a obtenção e manutenção de QoS, diversas propostas de Controle de Admissão têm surgido, voltadas especificamente para a arquitetura de Serviços Diferenciados. Os mecanismos de Controle de Admissão têm o objetivo de regular a entrada de usuários na rede, evitando, assim, a alocação excessiva de usuários, que poderia levar à degradação do nível de serviço.

Em uma arquitetura como o IntServ, em que há reserva de recursos, pode-se realizar a implementação de mecanismos de controle de admissão de maneira simples, dado que, sabendo a "quantidade" de recursos reservada para os usuários em atividade comparada com a "quantidade" de recursos disponíveis, a decisão de admissão é imediata. Porém, no caso do DiffServ, a situação é bem mais complexa, pois não há conhecimento da "quantidade" de recursos que os usuários ativos necessitam, principalmente em se tratando de aplicações Internet, cujo tráfego é normalmente caracterizado por alta explosividade. Para superar esses obstáculos, os algoritmos buscam caminhos alternativos, de estimação e monitoração do tráfego, visando alcançar os objetivos de justiça, equidade e disponibilidade.

Tendo em vista esse cenário, vários projetos europeus, terminados ou em curso, têm discutido soluções de QoS para a implantação de serviços premium

em larga escala na Internet. Estes projetos envolvem questões como Intserv, Diffserv, negociadores de banda e operação integrada. Dentre os diversos aspectos que têm sido objeto de estudo, os mecanismos de controle de admissão têm recebido especial atenção dada a sua extrema importância no framework de provisão de QoS. Podemos destacar alguns desses projetos:

- **TEQUILA** (*Traffic Engineering for Quality of Service in the Internet, at Large Scale*): Projeto de pesquisa e incentivo ao desenvolvimento, implementação e validação de definições de serviço para a obtenção de Qualidade de Serviço fim-a-fim na Internet (especificamente, no modelo DiffServ). O Projeto Tequila ([46]) foi iniciado em janeiro de 2000, e consiste da parceira de diversas empresas de telecomunicações (Alcatel, France Telecom, etc) e instituições de ensino (University College London, National Technical University of Athens, etc).
- **COPS DRA** (*COPS Diffserv Resource Allocation*): Projeto de pesquisa do CoRiTeL consortium ([8]), e visa o desenvolvimento de servidor e clientes COPS. CoRiTeL é um consórcio de pesquisa na área de telecomunicações e tecnologias relacionadas. Foi fundado em 1994 e é composto de quatro parceiros: o laboratório da Ericsson na Itália, a Universidade de Roma La Sapienza, a Universidade de Bolonha e a Universidade de Salerno.
- **CADENUS** (*Creation and Deployment of End-UserServices in Premium IP Networks*): Proposta de um *framework* para a criação, configuração e provisionamento de serviços ao usuário final com QoS. O projeto CADENUS ([7]) constitui uma proposta de arquitetura mediadora entre o usuário e o provedor de serviços, realizando as tarefas relacionadas à manutenção de Qualidade de Serviço.
- **Outros projetos:** Géant ([16]) e Sequin ([38])

Aproveitando essa motivação, e considerando a extensão do tema e a necessidade de avaliação e integração de diferentes tipos de procedimentos, o trabalho aqui apresentado trata da implementação de algoritmos de Controle de Admissão propostos para o modelo DiffServ. Buscamos verificar sua estabilidade e aderência às recomendações mais recentes do IETF e, quando necessário, implementamos extensões e soluções amplas, capazes de atender às necessidades de um esquema sólido de provisão de QoS.

Para esta análise comparativa, foi utilizado o software *Network Simulator* (*ns*), de uso público [32], desenvolvido em programa de pesquisa internacional com participação da University of Southern California e Information Sciences Institute.

As técnicas de Controle de Admissão implementadas são propostas recentes ([17] e [47]) e foram escolhidas por representarem abordagens originais ao problema do Controle de Admissão sem sinalização explícita. Ao longo dos trabalhos de implementação e testes, foram necessárias algumas modificações em relação às propostas originais (particularmente no algoritmo proposto em [47]) para que estas se adequassem à arquitetura DiffServ. As modificações e os procedimentos para validação dos programas são expostos ao longo da apresentação de resultados, no Capítulo 4.

1.3 Organização da Dissertação

Esta dissertação é dividida da seguinte maneira: no Capítulo 2 são apresentados conceitos básicos sobre controle de admissão e alguns dos mecanismos mais utilizados.

O Capítulo 3 traz uma descrição do panorama da implantação de QoS na arquitetura de Serviços Diferenciados. Nesse Capítulo, são abordadas resumidamente as principais características da Arquitetura DiffServ, apresenta-se um breve histórico das técnicas de Controle de Admissão propostas para diversas tecnologias e, enfim, são examinadas as características particulares dos mecanismos de Controle de Admissão que têm sido propostos para aplicação à Arquitetura de Serviços Diferenciados.

No Capítulo 4, são apresentados dois desses algoritmos de Controle de Admissão. Esses métodos são comparados através dos resultados de simulações obtidos com a implementação dos algoritmos no simulador *ns*.

Finalmente, no Capítulo 5 são apresentadas conclusões do trabalho e algumas propostas para trabalhos futuros.