

1 Introdução

1.1. Qualidade de Serviço na Internet

Várias técnicas baseadas na utilização do protocolo IP (*Internet Protocol*) estão sendo desenvolvidas para transportar integralmente, em uma única rede, os serviços de voz, dados e vídeo. O IP é parte fundamental da rede Internet e junto com o TCP (*Transmission Control Protocol*) e o UDP (*User Datagram Protocol*) são responsáveis pelo funcionamento de inúmeras aplicações. O protocolo IP é definido na camada 3 (nível de rede) do modelo OSI. Cada datagrama IP transporta informações suficientes que possibilitam à rede encaminhar os pacotes para o destino correto, sem necessidade de qualquer configuração adicional na rede. Para entregar os pacotes o provedor de serviço utiliza o serviço de melhor esforço (*Best Effort*), que provê simplesmente a conectividade básica sem garantias. Em situações de congestionamento, que ocorre quando os recursos da rede tornam-se escassos, o provedor de serviço pode retardar ou até mesmo descartar pacotes. Além disso, os pacotes podem ser entregues fora de ordem e algumas vezes o mesmo pacote pode ser entregue mais de uma vez. Sendo assim, os protocolos ou aplicações que agem em camadas superiores ao nível 3 devem estar cientes dessas possíveis falhas. Apesar de não oferecer garantias, o serviço de melhor esforço é apropriado para uma variedade de aplicações de rede, tais como, correio eletrônico e transferência de arquivos, pois essas adaptam-se facilmente às condições da rede.

O impressionante crescimento da Internet e a flexibilidade das redes IP comutadas por pacote têm acelerado a convergência das comunicações de dados, voz e vídeo dentro de uma arquitetura unicamente baseada em IP. Não é difícil prever que a infra-estrutura futura deve ser capaz de interconectar redes heterogêneas (fixa e móvel, baixa e grande largura de faixa) e suportar uma variedade de aplicações. Além de transações de dados como comércio eletrônico fluxos de áudio e vídeo têm se tornado cada vez mais freqüente na

Internet com a utilização de softwares como Netmeeting e CuSeeMe para vídeo conferência. Entretanto, como visto anteriormente, a Internet oferece apenas o serviço de “melhor esforço” sem qualquer garantia de entrega dos pacotes. Para atender as exigências das várias aplicações disponíveis atualmente na Internet é necessário prover Qualidade de Serviço (QoS). Qualidade de serviço é um termo subjetivo, que caracteriza o nível de satisfação usuário. Por exemplo, a qualidade da voz em uma conversação. Na prática, QoS é definida através de vários parâmetros como: vazão, que corresponde a taxa efetiva de transferência de dados, retardo, que é o tempo gasto pelos dados para ir da fonte ao destino, e *jitter* ou variação do retardo. Aplicações diferentes requerem diferentes níveis de QoS, pois podem ser sensíveis a diferentes parâmetros. Por exemplo, a transferência de um arquivo muito grande precisa de alta largura de faixa, mas não é sensível ao retardo e à sua variação. Em contrapartida, para a transmissão de vídeo conferência é necessária alta largura de faixa e as informações são altamente sensíveis ao retardo e sua variação. Desta forma, prover QoS significa controlar estes parâmetros para cada aplicação a fim de otimizar os recursos da rede e oferecer diferentes níveis de serviço.

Grandes esforços estão sendo realizados para prover QoS em redes IP. Duas propostas apresentadas pelo IETF (*Internet Engineering Task Force*) [33] vêm recebendo destaque: os Serviços Integrados (*IntServ*) [1] e os Serviços Diferenciados (*DiffServ*) [2].

Os Serviços Integrados (*IntServ*) provêm qualidade de serviço através da reserva de recursos por fluxo realizada pelo protocolo de sinalização RSVP (*Resource Reservation Protocol*) [4]. Entretanto, essa arquitetura requer a manutenção de estados por fluxos individuais nos roteadores inviabilizando a sua utilização de forma escalável.

A arquitetura de Serviços Diferenciados (*DiffServ*) [2] tem sido apresentada como uma solução alternativa aos Serviços Integrados (*IntServ*) [1] devido a sua escalabilidade. A arquitetura DiffServ realiza a marcação e o policiamento de pacotes nos roteadores de entrada da rede e provê tratamento diferenciado com base em agregados de tráfego. A existência de um byte no cabeçalho de cada pacote IP, o ToS (*Type of Service*), torna possível a marcação da prioridade ou nível de serviço que cada pacote requer. Os fluxos provenientes das aplicações interligados aos roteadores de entrada são agregados em classes de tráfego pré-definidas (mesmo ToS). Isso minimiza o processamento gerado nos roteadores da rede, pois as informações não são tratadas como um fluxo simples (*IntServ*), mas como um agregado de fluxos.

1.2. Objetivo

Dentre as duas propostas descritas anteriormente, os serviços diferenciados são uma alternativa viável para a provisão de diferentes níveis de serviço. Como consequência, vários estudos relacionados à estrutura da arquitetura foram realizados. Devido ao tratamento diferenciado do tráfego ser proporcionado pela marcação dos pacotes, vários estudos propuseram diferentes marcadores [11, 14, 16, 18, 19, 21, 22, 24, 25] e avaliaram os seus desempenhos. Nas propostas apresentadas nos artigos citados acima, os pacotes podem, usualmente, ser marcados de verde, amarelo ou vermelho. Essas cores indicam o nível de serviço que cada pacote requer. Pacotes marcados de verde têm prioridade de processamento, pacotes vermelhos serão descartados primeiramente em caso de congestionamento e os pacotes amarelos receberão um tratamento intermediário. Os resultados mostram problemas de justiça na alocação de largura de faixa entre fluxos que pertencem a um mesmo agregado, e na alocação do excesso de largura de faixa entre agregados.

Poucos estudos estão relacionados ao compartilhamento da largura de faixa excedente (largura de faixa não contratada), por isso este trabalho propõe um marcador de tráfego que distribui de maneira mais justa a largura de faixa excedente entre agregados para um enlace com carga de ocupação variando entre 20% e 60%. Como ponto de partida, várias soluções propostas [13, 15, 26, 28] foram analisadas e avaliadas. Uma solução denominada *IswTCM (A New Aggregate Marker to Improve Fairness in DiffServ)* [26], apresenta em seus resultados que o número de pacotes marcados de amarelo são um importante fator para se atingir justiça na alocação da largura de faixa excedente. Partindo desta premissa, é proposto modificações no algoritmo de marcação *IswTCM* de forma a atingir uma alocação mais justa da largura de faixa excedente.

O algoritmo proposto neste trabalho é comparado com as soluções propostas em [13, 15, 26] em relação à justiça no compartilhamento da largura de faixa excedente, em diferentes cenários de rede onde há mistura de fontes de tráfego como FTP/TCP, CBR/UDP, Exponencial. A mistura de diferentes tipos de tráfego visa avaliar o desempenho dos marcadores em uma situação mais real, pois o tráfego existente na Internet é caracterizado pela heterogeneidade.

Como dito anteriormente, em situações de escassez de recursos os pacotes que trafegam na rede podem ser descartados. Quando isso ocorre, os protocolos que atuam na camada superior ao IP devem reagir de modo a minimizar o impacto gerado sobre as aplicações. O protocolo TCP reage através da utilização de mecanismos de controle de congestionamento. As diferentes maneiras de controlar o congestionamento apresentadas pelo TCP motivou a análise de comportamento dos marcadores diante das diferentes versões do protocolo TCP (*Reno*, *New Reno*, *Sack*).

1.3. Organização do texto

Os oito capítulos que compõem esta dissertação estão organizados da seguinte forma :

O capítulo 2 descreve os parâmetros essenciais em uma rede capaz de prover QoS, e os requisitos das principais aplicações utilizadas na Internet.

No capítulo 3 são abordadas as duas arquiteturas propostas no IETF para provimento de QoS em redes IP: IntServ e DiffServ.

No capítulo 4, são apresentados os componentes básicos de um condicionador de tráfego, enfatizando o bloco de marcação de pacotes fundamental para este trabalho. Neste capítulo são descritos vários marcadores de tráfego.

O capítulo 5 apresenta os mecanismos utilizados pelo TCP para controle de congestionamento. Primeiramente, são apresentados os quatro algoritmos básicos de controle de congestionamento e em seguida as diferentes implementações do protocolo TCP utilizadas atualmente.

No capítulo 6 são apresentados os trabalhos relacionados à marcação de tráfego em redes Diffserv. Neste capítulo está descrito o algoritmo proposto nesta tese, sendo esta a principal contribuição deste trabalho.

No capítulo 7 estão descritos a topologia de simulação utilizada neste trabalho, a métrica de desempenho utilizada, e os resultados das simulações.

Finalmente, o capítulo 8 apresenta as conclusões sobre este trabalho e sugestões para trabalhos futuros.