MÁRIO EZEQUIEL AUGUSTO

AVALIAÇÃO EXPERIMENTAL DE FERRAMENTAS PARA MEDIÇÃO DE LARGURA DE BANDA

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Mestre em Informática.

Orientadora: Prof.ª Dr.ª Cristina Duarte Murta

CURITIBA

2002

Ministério da Educação
Universidade Federal do Paraná
Mestrado em Informática

PARECER

Nós, abaixo assinados, membros da Banca Examinadora da defesa de Dissertação de Mestrado em Informática, do aluno *Mário Ezequiel Augusto*, avaliamos o trabalho intitulado, ***"Avaliação Experimental de Ferramentas para Medição de Largura de Banda"***, cuja defesa foi realizada no dia 25 de novembro de 2002, às quatorze horas, no anfiteatro A do Setor de Ciências Exatas da Universidade Federal do Paraná. Após a avaliação, decidimos pela aprovação do candidato.

Curitiba, 25 de novembro de 2002.

Profª. Dra. Cristina Duarte Murta
DINF/UFPR

Prof. Dr. José Augusto Suruagy Monteiro
UNIFACS/BA

Prof. Dr. Elias Procópio Duarte Jr.
DINF/UFPR

Agradecimentos

Primeiramente quero agradecer a Deus pelas oportunidades que aparecem na minha vida e pelo sucesso obtido neste Mestrado. Agradeço de coração a minha família que sempre acreditou em mim, principalmente minha esposa Geane Célia pelo carinho e paciência, meu filho Marcelo Henrique, meus pais Gilberto Rigonati e Fátima Sgarbi e meus irmãos Ricardo Sgarbi e Giovana de Fátima.

Agradeço também aquelas pessoas que, direta ou indiretamente, me apoiaram e contribuíram para a realização deste trabalho, principalmente minha orientadora Cristina pela sua dedicação e perfeccionismo, ao diretor da Business Internet, Roberto Melani, por dispensar-me do trabalho para que eu pudusse me dedicar ao Mestrado, à Cláudia Watanabe da Impsat, ao Luis Bona e Martin Kretschek dos laboratórios da UFPR, ao Evandro Regolin do POP-PR e ao Renato Faraco do POP-MG por cederem máquinas em suas redes e fornecerem informações importantes sobre as mesmas, e aos integrantes da banca, Elias Procópio e José Suruagy, pelas contribuições para a dissertação.

Obrigado
Mário E.

Resumo

Medições de desempenho em redes de computadores são fundamentais nas atividades de gerência, diagnóstico de problemas de desempenho, implantação e verificação de contratos de qualidade de serviço, planejamento de capacidade, caracterização do tráfego, dentre outras. A análise das medições pode auxiliar a administrar o uso e a distribuição de largura de banda entre os vários serviços, a planejar novas redes, novos sistemas, ou prever a utilização e necessidades de atualização de um sistema existente. Centenas de ferramentas para medição de desempenho em redes estão disponíveis para várias métricas, diversos tipos de redes e ambientes operacionais. Ferramentas de medição devem apresentar algumas características importantes para que sejam úteis e confiáveis, dentre elas precisão e robustez.

Este trabalho apresenta um estudo sobre técnicas e ferramentas disponíveis para medição de largura de banda, acompanhado de resultados de avaliação experimental de algumas das ferramentas disponíveis para medição de largura de banda. Três métricas para medição de largura de banda foram consideradas: largura de banda de contenção, largura de banda nominal e largura de banda disponível. As ferramentas avaliadas são `clink`, `pathrate`, `bprobe`, `cprobe`, `pchar` e `nettimer`. Os experimentos foram feitos em redes consideradas de produção. As ferramentas foram avaliadas e comparadas de acordo com critérios exatidão das medidas, robustez e tempo de avaliação.

Este trabalho contribui para a identificação de vários problemas relacionados às medições de largura de banda em redes. Os resultados mostram que as medições não são satisfatórias em vários casos. Uma ampla discussão sobre as causas das imprecisões nas medidas é apresentada, assim como possíveis soluções para alguns dos problemas identificados. O trabalho apresenta também algumas técnicas e ferramentas para medição de outras métricas de desempenho, a saber, atraso, variação do atraso e perda de pacotes.

Abstract

Network performance measurement is a crucial activity for network management, for setting and checking service level agreements, performance diagnosis, capacity planning, and traffic characterization. The measurement analysis can help the network manager in the task of bandwidth allocation, the design of new networks and systems, the changes in traffic load and utilization, and determining the need of upgrading the system. Hundreds of tools for network performance measurement are available for the evaluation of many metrics, considering different kinds of networks and operational environments. Measurement tools must have some characteristics to be useful, such as accuracy and robustness. Accuracy is a very important characteristic that the tools should provide.

This dissertation presents a research about techniques and tools available to bandwidth measurement, and results of experimental assessment of some available tools to bandwidth measurement. Three bandwidth metrics are considered: contention bandwidth, nominal bandwidth and available bandwidth. The tools evaluated are `clink`, `pathrate`, `bprobe`,`cprobe`, `pchar` and `nettimer`. The tools were evaluated and compared according to criteria measures accuracy, robustness and time to estimation.

This dissertation helps to point out problems related with bandwidth measurement in networks. The results show that the measurements are not satisfactory in many cases. A wide discussion about the causes of imprecision of the measures is shown, as well as possible solutions to some identified problems. This dissertation also presents some techniques and tools employed to measure other performance metrics such as delay, jitter and packet loss.

Conteúdo

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

Medições de desempenho em redes de computadores são fundamentais nas atividades de gerência de redes e sistemas, diagnóstico de problemas de desempenho, implantação e verificação de contratos de qualidade de serviço, planejamento de capacidade e caracterização de carga, dentre outras. As medições çontribuem para prover informação necessária para a utilização da rede pelas aplicações e serviços. Durante a operação é necessário verificar se uma rede pode atender requisitos de um serviço e avaliar a qualidade do serviço no momento de sua execução. A análise das medições pode auxiliar a administrar o uso e a alocação de largura de banda entre os vários serviços.

A atividade de medição também provê informações essenciais para o planejamento do crescimento da infra-estrutura da rede. Através do acompanhamento do crescimento e das alterações do perfil do tráfego, é possível prever o crescimento da demanda e planejar o aumento da capacidade no prazo adequado, antecipando as necessidades de atualização de um sistema existente.

Outra contribuição importante das medições é no projeto dos sistemas de rede. A partir da análise do crescimento e das características da carga ao longo do tempo podemos propor alterações nos projetos dos dispositivos de rede e servidores. Conhecendo melhor a carga podemos ter sistemas cujos projetos levem em consideração mais aspectos desta carga e que, em conseqüência, operem com melhor desempenho.

Medições de desempenho em redes envolvem a análise de métricas que caracterizam o comportamento tanto da rede física como também dos protocolos, roteamento, tráfego e serviços existentes. Várias métricas podem ser de interesse. As métricas mais comuns são largura de banda, atraso e perda de pacotes. A partir destas métricas podemos descrever uma rede de alto desempenho como uma rede com uma grande largura de banda, um pequeno atraso e pouca perda de pacotes. Um diagnóstico preciso do desempenho de uma rede só é possível com ferramentas adequadas e confiáveis.

Em redes tais como a Internet, composta por várias redes que formam um ambiente

bastante heterogêneo, a medição de desempenho torna-se ainda mais importante. Neste tipo de rede, a capacidade de um enlace ou de um caminho entre dois pontos pode não ser conhecida. Assim, medições fim-a-fim, que requerem a cooperação apenas dos pontos terminais, podem ser a única forma de monitorar um caminho que inclui várias redes.

O aumento da utilização de serviços tais como vídeo conferência, transmissão de programas de TV, voz sobre IP, telemedicina e ensino à distância, entre outros, na Internet, tem aumentado a demanda por recursos como largura de banda e requisitos específicos de tempo de resposta. Estes serviços disputam a banda da rede com vários outros serviços, por exemplo, os oferecidos pelas aplicações *peer-to-peer*. O aumento do tráfego pode gerar atraso nas respostas e perda de dados. Controlar a operação das redes, diagnosticando problemas de desempenho, avaliando a qualidade das conexões, otimizando o uso e a alocação de banda, medindo e acompanhando a dinâmica da rede é cada vez mais uma tarefa necessária no dia a dia da Internet.

O conhecimento da largura de banda ao longo de um caminho pode beneficiar protocolos, serviços e aplicações. A partir desta medida pode-se avaliar a possibilidade de executar uma aplicação ou serviço, prover informação sobre a banda disponível, tarifar provimento de acessos e serviços, auxiliar a seleção dinâmica de servidor, otimizar o uso e a alocação de banda, e melhorar o desempenho dos servidores, dentre outras utilizações.

Este trabalho aborda o problema da medição fim-a-fim de largura de banda em redes IP. Nos últimos anos foram propostas várias técnicas e ferramentas para medição fim-a-fim de largura de banda em redes IP. No entanto, não encontramos nenhum trabalho que fizesse uma avaliação experimental ampla das ferramentas propostas. Tendo em vista a importância da medição de largura de banda na Internet, a utilização ampla desta medida de desempenho, e a necessidade de termos medições precisas e confiáveis, entendemos que uma avaliação das ferramentas quanto a alguns critérios desejáveis em ferramentas de medição, e uma comparação empírica entre as ferramentas disponíveis é de grande interesse para a comunidade.

Assim, este trabalho apresenta um estudo e discussão sobre as técnicas e ferramentas disponíveis para medição de largura de banda, acompanhado de resultados de avaliação experimental de algumas das ferramentas disponíveis para medição de largura de banda. As ferramentas avaliadas são `clink` [DOW 99, DOW 02], `pathrate` [DOV 01], `bprobe` [CAR 96b], `cprobe` [CAR 96b], `pchar` [MAH 99a] e `nettimer` [LAI 00, LAI 01]. Os experimentos foram feitos em redes consideradas de produção, LANs e WANs, que incluem a rede do Departamento de Informática da UFPR, o backbone da RNP e o backbone da Impsat. As ferramentas foram avaliadas e comparadas de acordo com critérios específicos, a saber, exatidão das medidas, robustez e tempo de avaliação.

Este trabalho contribui para a identificação de vários problemas relacionados às medições

de largura de banda em redes. Os resultados mostram que as medições não são satisfatórias em vários casos. Uma ampla discussão sobre as causas das imprecisões nas medidas é apresentada, assim como possíveis soluções para alguns dos problemas identificados. O trabalho apresenta também algumas técnicas e ferramentas para medição de outras métricas de desempenho, a saber, atraso, variação do atraso e perda de pacotes.

Outro aspecto que consideramos bastante relevante neste trabalho é sua contribuição para o domínio da tecnologia de medições em redes. Há várias evidências de que a tecnologia de medições em redes ainda está em estágio inicial de desenvolvimento. O número de publicações sobre este tema cresceu muito na última década, em especial nos últimos três anos. Alguns grupos de trabalho sobre o tema, no âmbito do IETF [IPP 01], foram criados recentemente. Eventos e workshops específicos sobre medições estão sendo propostos e realizados.

Este trabalho serve de apoio às necessidades de avaliação de desempenho em redes, desde a definição das métricas a serem analisadas até a escolha da ferramenta mais adequada para cada caso.

Este trabalho é formado por seis Capítulos dos quais este é o primeiro. O Capítulo 2 aborda os conceitos relacionados ao tópico de medição em redes, tais como métricas e tipos de medição, as técnicas e ferramentas de medição de atraso, variação do atraso e perda de pacotes. Este capítulo apresenta também uma classificação das ferramentas estudadas e os trabalhos relacionados. O Capítulo 3 apresenta as técnicas para medição de largura de banda e descreve as ferramentas escolhidas para a avaliação experimental. O Capítulo 4 descreve a metodologia do trabalho experimental. O Capítulo 5 apresenta os resultados dos experimentos realizados. O Capítulo 6 apresenta a conclusão do trabalho e sugestões de tópicos para trabalhos futuros.

Capítulo 2

Medições em Redes de Computadores

Medição é o processo de encontrar experimentalmente um valor para uma quantidade física, com a ajuda de meios especiais denominados instrumentos de medida [RAB 92]. A medição é sempre um processo experimental, e representa uma propriedade através de um valor quantitativo expresso em unidades de medida.

O desempenho dos sistemas computacionais é expresso e representado por quantidades físicas denominadas métricas [JAI 91]. Uma métrica é uma quantidade mensurável, bem definida qualitativamente e expressa quantitativamente. Também é usada para representar quantidades em outros sistemas, não apenas em sistemas computacionais. As métricas de desempenho mais comuns e de maior interesse em redes de computadores são largura de banda, atraso e perda de pacotes. No entanto, várias outras métricas podem ser definidas e utilizadas [PAX 98b, IPP 01].

Os componentes necessários para qualquer medição são o método de medição e o instrumento de medida. Em redes de computadores, os métodos distintos para medição de uma dada métrica são denominados **técnicas de medição**. Estas técnicas são procedimentos ou metodologias para proceder à medição. Cada implementação de uma técnica gera uma **ferramenta de medição**, que é o instrumento de medida.

Neste capítulo são discutidos, nas seções 2.1 e 2.2, diversos conceitos relativos à medição de desempenho em redes e a definição das métricas e características desejáveis das medições. A seção 2.3 apresenta o contexto histórico e os trabalhos relacionados. A seção 2.4 descreve as técnicas e ferramentas para medição das métricas atraso, variação do atraso e perdas de pacotes. A seção 2.5 apresenta uma proposta de classificação das ferramentas de medição estudadas. Um estudo detalhado das técnicas e ferramentas específicas para medição de largura de banda é apresentado no capítulo 3.

2.1 Metrologia Aplicada às Redes de Computadores

Neste trabalho vamos utilizar a terminologia relativa a medições em redes definida em [PAX 98b] e apresentada a seguir. Um *host* é um computador capaz de comunicar-se com outros computadores utilizando protocolos da Internet, o que inclui os roteadores. Um enlace é uma conexão única no nível de enlace entre dois *hosts*. Um caminho é uma seqüência da forma $< h_0, l_1, h_1, l_2, h_2, ..., l_n, h_n >$ onde $n >= 0$, h_i é um *host*, l_i é um enlace entre h_{i-1} e h_i, e cada $h_1..h_{n-1}$ é um roteador. Um par $< l_i, h_i >$ é denominado um *hop*.

Metrologia é a ciência que estuda os processos de medição, as medidas e os erros associados. Em metrologia, as medições têm sido historicamente classificadas em medições diretas, indiretas e combinadas [RAB 92]. Nas medições diretas o instrumento de medida interage diretamente com o objeto de estudo e indica o valor da medida. No caso da medição indireta, o valor da quantidade mensurável é encontrado a partir de relações de dependência conhecidas entre esta quantidade e outros parâmetros observados ou avaliados e pode ser feita de várias formas [PAX 98b], por exemplo, a projeção de uma métrica a partir de medições em um nível mais baixo ou a estimativa de uma métrica a partir de um conjunto de medidas agregadas são algumas das metodologias aplicáveis. Medições combinadas empregam ambos os métodos. Em muitas situações, a distinção entre medição direta e indireta pode não ser simples.

Uma outra classificação pode ser utilizada para as medições em redes de computadores. Algumas ferramentas de medição em redes transmitem pacotes extras na rede para realizar as medições, enquanto outras apenas inspecionam o tráfego na rede ou seus registros. Desta forma, as ferramentas podem ser divididas em dois grupos: ferramentas de medição ativa e ferramentas de medição passiva.

A medição ativa é realizada através da injeção de pacotes extras na rede, utilizando tráfego de teste. Um ou mais pacotes são enviados pela rede e o resultado da transmissão é medido. A medição pode ser realizada no mesmo *host* que enviou os pacotes ou no *host* destino. Uma desvantagem deste tipo de medição é a adição de tráfego na rede devido aos pacotes transmitidos pelo instrumento de medição.

A medição passiva não gera tráfego extra na rede, apenas captura pacotes que estão trafegando na rede no momento da medição. A captura dos pacotes é feita através de filtros que permitem escolher quais pacotes capturar. Este tipo de medição pode apresentar limitações em termos de quantidade de dados capturados, caso o mecanismo usado não seja rápido o suficiente para capturar todos os pacotes que trafegam pela rede [PAX 97]. Outra desvantagem é a dificuldade de análise em tempo real pois é necessário armazenar um registro dos pacotes capturados e depois realizar a análise. As formas de medição ativa e passiva são utilizadas neste trabalho como um critério para classificação das ferramentas.

Baseando-se nestas duas classificações, as ferramentas podem ser divididas em quatro

grupos. As ferramentas que realizam medição ativa direta injetam tráfego extra na rede e medem diretamente a métrica em questão. Na medição ativa indireta, a ferramenta injeta tráfego, calcula uma métrica e depois infere a métrica em questão baseando-se nestes cálculos. Na medição passiva direta, a ferramenta apenas monitora os pacotes que estão trafegando na rede e com base nestes pacotes calculam a métrica desejada. Na medição passiva indireta, a ferramenta utiliza os pacotes monitorados para calcular uma métrica e, com base nesta métrica, infere a métrica desejada.

Pelo menos três características essenciais a uma metodologia ou técnica de medição podem ser apontadas. A característica mais importante da qualidade da medição é a precisão [RAB 92]. A precisão de qualquer medição particular é determinada pela precisão dos instrumentos de medida empregados, pela técnica de medição empregada e, algumas vezes, pela habilidade do executor da medição. A repetibilidade das medições é outra propriedade importante. Se medições da mesma quantidade são realizadas múltiplas vezes sob as mesmas condições, as medições obtidas devem ser consistentes, isto é, iguais ou próximas por uma diferença aceitável. A reprodutibilidade é a capacidade de gerar resultados de medição confiáveis da mesma medida em condições diferentes, isto é, diferentes locais, ambientes e equipamentos.

Uma medição de uma quantidade cujo valor verdadeiro é A fornece uma estimativa $\tilde{A}$ da quantidade. O erro absoluto da medição é dado por ϵ e pode ser definido como $\epsilon = | \tilde{A} - A |$. No entanto, esta relação não pode ser utilizada para encontrar o erro de uma medição pelo simples fato de que o valor verdadeiro não é conhecido. Se o valor verdadeiro fosse conhecido, então não haveria necessidade da medição. Por esta razão, os erros devem ser estimados utilizando dados indiretos.

Quanto às suas propriedades, os erros podem ser classificados em erros aleatórios e sistemáticos. Um erro é sistemático se ele permanece constante ou muda de maneira regular em experimentos repetidos. O erro aleatório pode ser identificado quando existem diferenças entre os resultados de medições independentes, estas diferenças não podem ser previstas individualmente e as regularidades inerentes à medição só podem ser observadas para um grande número de medições.

As técnicas e ferramentas estudadas neste trabalho fazem medição ativa das métricas de desempenho das redes e baseiam-se no envio de pacotes. As técnicas geralmente são fundamentadas em algum comportamento observado na transmissão de pacotes pela rede ou mesmo a partir de algum conceito. Por exemplo, o distanciamento de pacotes transmitidos juntos por um caminho de rede, observado por Van Jacobson em 1988 [JAC 88], deu origem à técnica *packet-pair*.

Diversos protocolos são utilizados nos procedimentos de medição. A escolha do protocolo é feita de acordo com o projeto da metodologia de medição ou ferramenta. A seguir apre-

sentamos uma breve descrição dos protocolos utilizados pelas ferramentas analisadas neste trabalho. Uma descrição mais detalhada de cada um pode ser encontrada em [COM 00].

O IP (*Internet Protocol*) desempenha a função de roteamento, escolhendo um caminho por onde os dados serão enviados. As ferramentas utilizam o campo TTL (*time to live*) do protocolo IP para obter os endereços dos roteadores intermediários entre dois *hosts*. Diferentemente do protocolo TCP, o IP é não confiável e não orientado a conexão.

O ICMP (*Internet Control Message Protocol*) permite o envio de mensagens de erro ou de controle pela rede. Este protocolo é uma parte integrante do protocolo IP e é mais usado pelas ferramentas com o objetivo de descobrir o caminho que os pacotes percorrem pela rede. Também é o protocolo mais usado quando a ferramenta é executada em apenas um *host*, pois através deste protocolo a ferramenta pode enviar pacotes e obter uma resposta.

O TCP (*Transmission Control Protocol*) é o protocolo de transporte mais comumente utilizado para transmitir dados entre dois *hosts*. A vantagem maior do TCP para as ferramentas é que o protocolo implementa mecanismos de retransmissão, garantindo que os pacotes da medição sejam totalmente transmitidos, resultando em maior confiabilidade para a ferramenta. O TCP também é usado quando a medição envolve algum serviço que usa o TCP, como, por exemplo, o monitoramento de tráfego TCP. Este protocolo é orientado a conexão.

O UDP (*User Datagram Protocol*) é o protocolo de transporte mais usado para transmitir dados entre dois *hosts* quando a eficiência é mais importante que a confiabilidade de entrega dos dados. Por exemplo, para uma ferramenta fazer várias medições gerando tráfego na rede, o UDP irá gerar menos tráfego que o TCP, e os resultados não serão prejudicados se apenas uma pequena parte da amostra de medição for perdida. Embora as unidades de tráfego UDP sejam chamadas de datagramas, neste trabalho foi utilizado o termo pacotes para generalizar os pacotes usados pelas ferramentas, independente do protocolo utilizado. O UDP também é um protocolo não orientado a conexão.

O SNMP (*Simple Network Management Protocol*) é o protocolo padrão de gerenciamento de redes TCP/IP. O SNMP provê às ferramentas informações contidas nos roteadores como, por exemplo, informações de tráfego, roteamento, número e tamanho dos pacotes, utilização de CPU do roteador e pacotes descartados. Com estas informações, um serviço pode armazenar um histórico sobre as características do tráfego e gerar estatísticas sobre este tráfego. Em [AHN 99] os autores descrevem o projeto e implementação de um sistema de gerenciamento para analisar o desempenho da Internet usando a MIB-II do SNMP.

As técnicas e ferramentas de medição em redes, como os demais métodos e instrumentos de medição, devem também apresentar algumas características para que sejam úteis. Reprodutibilidade, pouca intrusão, repetibilidade, precisão e rapidez para a obtenção da medida são as principais características desejáveis [LAI 00].

A característica de reprodutibilidade, algumas vezes referenciada na literatura de redes como robustez, é definida pela habilidade de empregar a técnica ou ferramenta em uma grande variedade de ambientes encontrados na Internet, tais como, poucos ou muitos *hops* entre os dois pontos da medição, enlaces com pouca ou muita carga, diferentes tecnologias de enlace com ou sem fio, diferentes disciplinas de fila e implementações de roteadores.

A intrusão é dada pela quantidade de carga inserida na rede para obter a medição. É desejável que a técnica ou ferramenta insira a menor quantidade de carga possível para proceder a medição. A repetibilidade é também referida como consistência nas medições, e refere-se à habilidade da ferramenta obter os mesmos valores medidos para condições iguais de medição. A precisão é uma característica essencial, pois a confiança no resultado fornecido pela ferramenta deve-se à precisão de sua medição. A rapidez de execução das medidas pode ser desejável em algumas situações em que o dado estimado será utilizado em tempo real para a tomada de alguma decisão, mas pode não ser essencial em outras situações.

Em muitos casos, as ferramentas executam muitas medições repetidas da mesma métrica e fazem análises estatísticas dos resultados desta grande quantidade de valores antes de chegar a um resultado da medição. Métodos estatísticos complexos são utilizados. Embora estes métodos sejam muito importantes para o resultado da medição, eles não serão estudados neste trabalho.

Uma outra característica das técnicas que pode ser muito desejável é a sua escalabilidade em termos de sua aplicabilidade prática para um grande número de caminhos e enlaces. Uma caracterização ampla do desempenho de várias redes que compõem a Internet, por exemplo, só seria possível com um método escalável.

2.2 Definição das Métricas

Para a obtenção de um entendimento comum e preciso do desempenho e da confiabilidade da Internet, métricas de desempenho e confiabilidade devem ser propostas e medidas destas métricas devem ser tomadas. A proposição de métricas deve seguir alguns critérios, já colocados de forma geral na seção anterior, por exemplo, as métricas devem ser bem definidas, e as metodologias devem apresentar as propriedades de precisão, repetibilidade e reprodutibilidade.

Várias métricas são conhecidas e definidas para medição de desempenho e confiabilidade em redes de computadores [PAX 98b, IPP 01, MAH 99b, ALM 99a, ALM 99b, ALM 99c, MAT 01, KOO 02]. Dentre estas métricas podemos citar atraso de ida, atraso de ida e volta, perdas, variação do atraso, reordenação de pacotes, tempo de propagação de um enlace, largura de banda de um enlace, a rota ou caminho entre dois *hosts* de uma rede, o número de *hops* de uma rota. Deve-se ressaltar que qualquer métrica pode ser proposta,

definida e utilizada. Sua utilização ampla dependerá do emprego e da utilidade da medição. A seguir são definidas três métricas relacionadas à largura de banda, além das métricas atraso, variação do atraso e perdas de pacotes.

2.2.1 Largura de Banda e Utilização

Largura de banda é a taxa de transmissão de dados na qual um enlace pode propagar informação [PAX 99, LAI 99, PET 99, CAR 96b, DOV 01]. A largura de banda nominal ou capacidade nominal de um enlace é freqüentemente referida simplesmente como largura de banda. A unidade da medição é dada em bits por segundo (b/s) ou Kb/s, Mb/s e Gb/s abreviando Kilobits, Megabits e Gigabits por segundo, considerando potência de 10, ou seja, 1 Kb representa 1000 bits [PAX 98b].

Três métricas relacionadas à largura de banda podem ser definidas:

- **largura de banda de contenção** (*bottleneck bandwidth*): é a taxa máxima que uma rede pode transmitir dados de um transmissor para um receptor na ausência de qualquer outro tráfego, considerando que o ponto de contenção esteja na largura de banda e não nos roteadores intermediários. Se entre um transmissor e um receptor de dados há enlaces de 64 Kb/s, 2 Mb/s e 1 Mb/s então a largura de banda de contenção é 64 Kb/s. Este valor é razoavelmente estável, considerada a estabilidade das rotas entre dois *hosts*, e também não se altera devido ao aumento ou à diminuição no tráfego, pois esta métrica não considera o tráfego;

- **largura de banda utilizada**: representa a quantidade de dados trafegando por um enlace em um determinado momento. A largura de banda utilizada em um enlace é a soma das bandas utilizadas de todos os fluxos de dados que estão trafegando pelo enlace;

- **largura de banda disponível** (*available bandwidth*): é a taxa máxima na qual um *host* consegue transmitir dados ao longo de um caminho da rede em um certo momento. Este valor varia com o tráfego existente no momento da medição. Se, por exemplo, em um certo enlace é possível transmitir 3 Mb/s e o enlace já apresenta um tráfego de 1,2 Mb/s, então a largura de banda disponível no momento é de 1,8 Mb/s. Em um caminho de rede, a largura de banda disponível é determinada pelo enlace com a menor largura de banda não utilizada.

Estas métricas estão relacionadas. A capacidade de um caminho é determinada pelo enlace com a menor capacidade nominal, que é o enlace de contenção. Ele define o limite superior para a taxa de transmissão entre os dois pontos terminais do caminho. Assim, a medição de largura de banda de contenção é a forma de estimar a capacidade nominal de um enlace ou caminho.

A utilização é definida como a fração da capacidade nominal que está sendo utilizada, ou seja, a razão entre a largura de banda utilizada e a largura de banda nominal. Se, em um enlace com capacidade nominal de 1Mb/s, o tráfego está em 700Kb/s, então a utilização naquele momento é de 70% ou 0,7.

O complemento da utilização indica a quantidade de dados que pode efetivamente ser transmitida em um certo momento, que é a porcentagem da capacidade nominal que deve estar disponível para uma aplicação. A largura de banda disponível e a largura de banda utilizada nunca excedem a largura de banda de contenção [PAX 99].

2.2.2 Atraso

O atraso (*delay*) corresponde ao tempo de transmissão de um *host* de uma rede a um outro *host* da mesma rede ou fora dela [PET 99]. O tempo geralmente é medido em milissegundos (ms). Por exemplo, numa rede local, o atraso de um pacote é geralmente da ordem de alguns décimos de milissegundos. Entre duas cidades de uma mesma região, o atraso pode ser da ordem de algumas dezenas de milissegundos ou mais.

O tempo necessário para transmitir um pacote a um *host* destino e retransmiti-lo de volta à origem é o atraso de ida e volta ou *round-trip time* (RTT). Há uma relação entre tamanho do pacote e atraso. Quanto maior o tamanho do pacote a ser enviado, maior será o tempo gasto para transportá-lo até o destino, embora a diferença seja muito pequena em redes de alta velocidade.

O tempo gasto numa transmissão é a soma de vários tempos, como mostrado na Figura 2.1, extraída de [HEN 95]. O tempo gasto por um *host* transmissor para processar um pacote, que inclui o tempo de processamento de CPU necessário para acrescentar o cabeçalho e colocar o pacote na área de armazenamento de saída (*buffer*) da rede, é chamado de processamento do pacote. Na transmissão de um pacote, o processamento ocorre tanto no transmissor quanto no receptor dos dados. O tempo de propagação inicia no momento em que o primeiro bit sai do transmissor até o momento que este bit chega no receptor, incluindo atrasos devido a repetidores e outros equipamentos no caminho. O tempo de propagação também é chamado de latência. O tempo de transmissão, que é o tempo para um pacote ser transmitido pela rede, é a razão entre o tamanho do pacote em bits e a largura de banda nominal. O atraso de transporte é o tempo gasto para o pacote atravessar a rede, e inclui os tempos de propagação e de transmissão. O atraso total é o atraso no transporte mais os tempos de processamento no transmissor e no receptor.

Desta forma, o atraso total é dado pela seguinte soma:

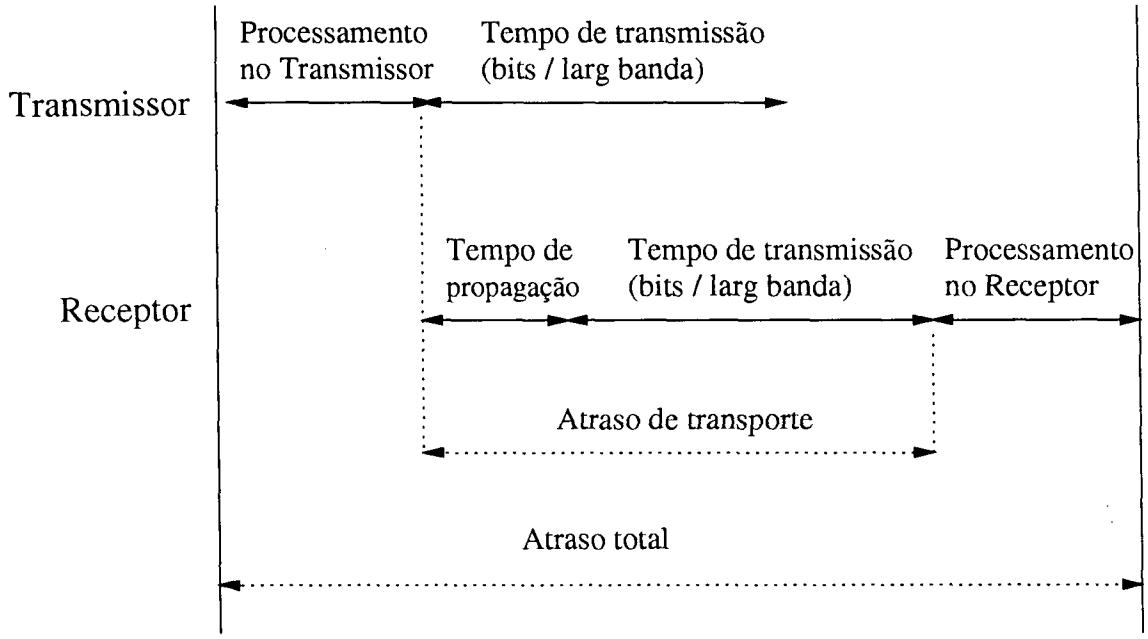

Figura 2.1: Representação dos diversos tempos que compõem o atraso da transmissão de um pacote.

$$atraso\ total = processamento\ no\ transmissor + tempo\ de\ propagação + \frac{tamanho\ mensagem}{largura\ banda} + processamento\ no\ receptor$$

O produto da latência pela largura de banda corresponde à quantidade total de bits que podem trafegar ao mesmo tempo em um enlace.

2.2.3 Variação do Atraso

A variação de atraso corresponde à diferença entre os atrasos na transmissão de pacotes subsequentes. Por exemplo, se dois pacotes são transmitidos, o primeiro pacote chega ao destino com um atraso de 6,0 ms e o segundo pacote chega com um atraso de 6,4 ms, então a variação de atraso neste caso foi de 0,4 ms. O valor da variação de atraso normalmente é considerado baixo quando o atraso máximo não é mais que 10-20% maior que o atraso mínimo [NLA 01].

A variação do atraso é uma métrica derivada do atraso, e pode ser medida a partir do atraso de ida ou do atraso de ida e volta. A unidade para esta métrica é um número real em segundos, positivo, zero ou negativo, ou um número indefinido de segundos [DEM 02]. Esta última unidade é utilizada quando o destino não recebeu um ou os dois pacotes.

Esta é uma métrica muito importante para aplicações em tempo real. Em uma vídeo conferência, por exemplo, as imagens devem ser transmitidas com uma freqüência pouco variável. Se ocorre variação de tempo na entrega dos dados, caracterizando a existência de variação de atraso, podem ocorrer cortes ou congelamento nas imagens.

2.2.4 Perda de Pacotes

Esta métrica corresponde à porcentagem de pacotes perdidos durante uma transmissão. Esta perda pode ser conseqüência, por exemplo, de um roteador sobrecarregado ou de congestionamento no enlace. A perda de pacotes é prejudicial aos serviços da rede pois pode implicar em retransmissão de pacotes perdidos. As perdas decorrem tipicamente de congestionamento, que é agravado pela retransmissão dos pacotes perdidos. A perda de pacotes costuma ser muito menor em redes locais do que em redes de longa distância.

As métricas podem ter importância relativa de acordo com a necessidade da medição e até mesmo em função do próprio resultado da medição. Por exemplo, quando a utilização é baixa (< 30%), o atraso pode ser uma métrica bem importante pois a qualidade de um serviço pode depender deste valor. Por outro lado, quando a utilização é alta (30% a 70%), a variação do atraso e as perdas podem ser medidas bem importantes para avaliar se ainda é possível executar um serviço ou aplicação.

2.3 Contexto Histórico e Trabalhos Relacionados

As primeiras ferramentas para medições em redes surgiram na década de 80. O objetivo era realizar testes simples de conexão e roteamento. A primeira ferramenta foi provavelmente o **ping** [MUR 00], em 1980, que foi incorporado ao sistema operacional UNIX, sendo bastante útil para testes simples de conexão em rede. O **traceroute** [MUR 00] surgiu em 1988 e foi desenvolvido por Van Jacobson.

Das ferramentas de medição de largura de banda estudadas, a primeira a ser desenvolvida foi o **MRTG** [MRT 01] em 1994. O **MRTG** foi desenvolvida por Tobias Oetiker devido à necessidade de monitorar o tráfego de rede de um enlace de 64 Kb/s na De Montfort University.

As ferramentas de medição ativa de largura de banda, ou seja, aquelas que inserem pacotes extras na rede, foram inspiradas no conceito de dispersão de pacotes apresentada por Van Jacobson em 1988 [JAC 88]. Deste conceito surgiu a técnica *packet-pair*. Em 1996 Carter e Crovella apresentaram as ferramentas **bprobe** e **cprobe** [CAR 96b] que utilizam *packet-pair* para estimar, respectivamente, a menor largura de banda existente em um caminho de rede e a menor largura de banda não usada em um caminho de rede. No mesmo ano Vern Paxson desenvolveu **tcpanaly** [PAX 97] também com o objetivo de estimar

a menor largura de banda de um caminho de rede.

Outros pesquisadores desenvolveram técnicas e ferramentas de medição de largura de banda. Em 1997 Van Jacobson desenvolveu a técnica *one-packet* e a ferramenta **pathchar** [JAC 97] que calcula a largura de banda de todos os enlaces no caminho. Baseando-se nesta ferramenta, em 1999 surgiram **pchar** [MAH 99a] desenvolvida por Bruce Mah e **clink** [DOW 99] desenvolvida por Allen Downey. Ainda em 1999, Savage apresenta sua ferramenta de medição de perda de pacotes chamada **sting** [SAV 99].

Em 2000, Lai e Baker desenvolvem a técnica *packet-tailgating* e implementam a ferramenta **nettimer** [LAI 00][LAI 01]. No mesmo ano, Balbinot e Maiko de Andrade implementam a ferramenta **linkstat** [BAL 00] que mede o tráfego de dados em um enlace e é baseada na ferramenta **RRDTool**.

Em 2001, Dovrolis [DOV 01] implementa a ferramenta **pathrate**, que mede a menor largura de banda em um caminho de rede e é baseada na dispersão de pares de pacotes (*packet-pair*) e de séries de pacotes. Em 2002, Dovrolis apresenta a ferramenta **pathload** em [DOV 02]. Várias destas técnicas e ferramentas são discutidas e testadas nesta dissertação. A ferramenta **pathload** não foi analisada porque não estava disponível quando foi realizada a experimentação.

Alguns autores estudam o comportamento da rede e dos protocolos sem implementar uma ferramenta para análise. Em [BOL 93] Jean Bolot realiza medições de atraso e perda de pacotes utilizando pequenos pacotes UDP em intervalos regulares. Bolot observou que o tráfego Internet é formado por uma quantidade pequena de pacotes grandes, que geram a maioria do volume em bytes, e uma quantidade maior de pacotes menores. Alguns autores como Van Jacobson [JAC 88] e Srinivasan Keshav [KES 91a][KES 91b] estudam como medições de desempenho podem melhorar o controle de fluxo e congestionamento nas redes e apresentam abordagens para melhorar o fluxo de dados baseadas em taxa de transferência.

Um tópico relacionado à medição de desempenho é a implementação de infra-estruturas, ambientes e estratégias que permitam estabelecer e cumprir requisitos específicos de qualidade para os serviços executados pelas redes. Por exemplo, a alocação de largura de banda na Internet atual segue o modelo *best-effort*, que significa que a rede fará a transmissão de dados da melhor maneira possível, sem, no entanto, dar nenhuma garantia quanto à taxa de serviço ou latência da transmissão. Neste modelo, a largura de banda é compartilhada entre todos os usuários, praticamente sem esquema de prioridade.

Há um grande esforço para implantar redes de computadores que possam estabelecer e cumprir metas específicas de desempenho e confiabilidade. Serviços tais como teleconferência, ensino à distância, voz sobre IP, e transmissões diversas de áudio e vídeo, além de diversos outros, apresentam requisitos específicos de transmissão para que sejam operados de forma útil e transparente para o ususário. Para prover qualidade de serviço, existem

alguns mecanismos propostos tais como IntServ e DiffServ [MUR 00]. Estes mecanismos podem fornecer à aplicação garantias de vazão, atraso, variação de atraso e taxa de perdas limitada. Uma desvantagem destes mecanismos é que sua implantação exige mudanças nas redes de hoje, por exemplo, mudanças do sistema operacional de alguns roteadores, para que se adequem aos novos protocolos exigidos pelos serviços. Em [KES 98] o autor analisa a gerência de redes como uma forma de melhor fornecer garantias de qualidade de serviço. Um ambiente para descrever e gerenciar políticas de qualidade de serviço foi proposto em [GRA 01].

Paralelamente com as pesquisas de técnicas de medição e controle de qualidade de serviços, estão os projetos de medição de larga escala, que são projetos que visam o desenvolvimento de métodos de análise, padronização de métricas, infraestruturas de medição e também a criação de novas ferramentas de medição e avaliação de desempenho.

O projeto WAWM (*Wide Area Web Measurement*) [BAR 99] propõe a utilização de uma infraestrutura distribuída pela Internet para estudar o desempenho da Web. Esta infraestrutura permite medições simultâneas de desempenho de cliente Web, desempenho de rede e desempenho de servidor Web. A associação CAIDA (*Cooperative Association for Internet Data Analysis*) [CAI 01] aponta problemas de medição e desempenho de tráfego Internet e problemas de comunicação e cooperação entre provedores de serviços Internet. O projeto NIMI (*National Internet Measurement Infrastructure*) [NIM 01, PAX 98a] tem como objetivo criar uma infra-estrutura de medição para a Internet. O grupo de trabalho IPPM (*Internet Protocol Performance Metrics*) [IPP 01] do IETF visa desenvolver um conjunto de métricas e procedimentos de medição que possam ser aplicados para avaliar a qualidade, o desempenho e a confiabilidade dos serviços de entrega de dados na Internet. O laboratório NLANR (*National Laboratory for Applied Network Research*) [NLA 01] visa criar uma infraestrutura de análise de rede para fornecer um melhor entendimento dos modelos de serviços e métricas da Internet. Este projeto coloca disponível em sua página Web centenas de ferramentas de medição e monitoramento de redes. O projeto IPMA (*Internet Performance Measurement and Analysis*) [IPM 01] tem como objetivo estudar o desempenho de redes e seus protocolos em redes locais e de longa distância. O projeto Surveyor [SUR 01] realiza medições de atraso e perda entre sites remotos.

Várias ações e iniciativas mostram que esta área de pesquisa é recente, além das próprias publicações. Novos grupos de trabalho sobre o tema são criados, por exemplo, o IMRG (*Internet Measurement Research Group*) [IMR 02] e o GT-QoS da RNP [RNP 02]. Estes grupos têm o objetivo de prover um fórum de discussões sobre medições na Internet, aumentar a interação entre operadores, desenvolvedores e pesquisadores, e fomentar o desenvolvimento de novas técnicas e ferramentas para medição. Além disso, *workshops* específicos sobre medições têm sido realizados, por exemplo, o PAM (*Passive and Active Measurement*

Workshop) [PAM 01].

2.4 Ferramentas para Medições em Redes IP

As ferramentas apresentadas nesta seção estão divididas de acordo com a métrica calculada.

2.4.1 Ferramentas para Avaliação de Atraso e Variação de Atraso

A medição de atraso em redes é feita através do envio de pacotes para um *host* destino denominado receptor. A hora da transmissão é registrada. Do momento em que os pacotes chegam no receptor, é subtraída a hora da transmissão. Este é o atraso de ida.

O receptor pode ser o mesmo *host* que enviou os pacotes. Neste caso, os pacotes são transmitidos até um outro *host* e mede-se o tempo de ida e volta. Ferramentas mais simples de medição de atraso transmitem pacotes ICMP com pedido de eco (ECHO_REQUEST) e recebem a resposta (ECHO_RESPONSE).

Medições mais precisas requerem que a ferramenta seja executada em dois *hosts* distintos. Neste caso, pacotes TCP ou UDP são transmitidos com o registro do momento da transmissão. Ao chegarem ao *host* receptor, subtrai-se o tempo de transmissão. Para esta técnica fornecer resultados precisos, os relógios dos dois *hosts* devem estar sincronizados através de um sistema de sincronização de relógio.

O método é o mesmo para o cálculo da variação de atraso. Porém os dois *hosts* não precisam estar sincronizados pois é necessário medir apenas a diferença de tempo entre as chegadas.

A seguir são descritas algumas ferramentas que medem atraso e variação de atraso.

ping

O **ping** [MUR 00] é uma ferramenta simples, usada para medições de curto período de tempo. Geralmente é usada para verificar se um *host* pode ser alcançado pela rede e medir o tempo de ida e volta dos pacotes até este *host*. **ping** é um comando padrão no UNIX e disponível também na plataforma Windows.

ping repetidamente mede o atraso de ida e volta (RTT) até a máquina ou roteador destino enviando pacotes ICMP ECHO_REQUEST e recebendo pacotes ICMP ECHO_RESPONSE. A cada resposta obtida, o atraso é impresso na tela.

Na primeira linha de saída do resultado, **ping** mostra o endereço IP do destino e a quantidade de bytes de dados a ser enviada. Nas próximas linhas mostra o valor do TTL (campo Time-To-Live do protocolo IP) que voltou nos pacotes e o atraso medido em milissegundos. Após terminar de medir o atraso, opcionalmente o **ping** informa a quantidade de pacotes

transmitidos, recebidos e porcentagem de perda. Na última linha mostra o menor tempo, tempo médio e maior tempo de atraso de ida e volta medido.

Entre as vantagens da ferramenta **ping** podemos citar a rapidez do resultado. No primeiro segundo de execução, o primeiro atraso já começa a aparecer. Outra vantagem é a facilidade de uso pois basta executar a ferramenta em linha de comando. Além disso, a ferramenta não exige instalação extra ao sistema operacional.

Entre as desvantagens da ferramenta **ping** estão o uso do protocolo ICMP, o qual tem pouca prioridade em alguns roteadores. Isso pode gerar resultados não muito confiáveis. Outra desvantagem é que realiza medição ativa. Pelo fato da ferramenta transmitir dados pela rede, este tipo de medição gera tráfego na rede, embora seja pouco, apenas algumas dezenas de bytes por segundo.

Uma limitação desta ferramenta é que ela depende da implementação dos roteadores intermediários: alguns roteadores bloqueiam pacotes ICMP, com isso podemos ter a falsa impressão de que o roteador não está ligado ou o *host* destino não pode ser alcançado.

traceroute

A ferramenta **traceroute** [MUR 00] é usada para indicar o caminho que os pacotes IP percorrem do *host* que executa a ferramenta **traceroute** até um *host* destino através da rede, indicando o atraso de ida e volta dos pacotes para cada *host* intermediário.

A ferramenta **traceroute** tenta indicar a rota que um pacote IP seguiria até determinado *host* na Internet usando o campo TTL do protocolo IP para obter uma resposta ICMP TIME_EXCEEDED de cada *host* intermediário entre origem e destino. Primeiramente o TTL é inicializado em 1 e acrescido de 1 a cada *host* medido, mandando três pacotes para cada um. A ferramenta continua medindo cada *host* até receber uma mensagem ICMP "port unreachable", significando que a máquina destino foi alcançada. Para isso, os pacotes são endereçados para uma porta destino sem uso.

Na primeira linha do resultado da execução, **traceroute** imprime o endereço IP do *host* destino, o número máximo de *hosts* intermediários e o tamanho dos pacotes enviados. Em seguida, a ferramenta apresenta os resultados para cada *host* medido: nome do *host* e/ou endereço IP, e os atrasos entre origem e destino (3 por padrão).

traceroute fornece as mesmas vantagens, desvantagens e limitações descritas para a ferramenta **ping**.

mtr

mtr [MTR 01] é uma ferramenta que é uma combinação do **ping** e do **traceroute**. Quando **mtr** é iniciado, este traça o caminho entre o *host* que executa **mtr** e o *host* destino (funcionalidade do **traceroute**). Após determinar o endereço de cada *host* intermediário, **mtr**

manda pedidos ICMP ECHO repetidamente para cada um a fim de medir o atraso e a taxa de perda (funcionalidade do **ping**). **mtr** imprime o resultado das medições à medida em que recebe as respostas ICMP ECHO. O acesso a esta ferramenta ocorre apenas através da linha de comando.

As vantagens, desvantagens e limitações são as mesmas das ferramentas **ping** e **traceroute**, com uma vantagem adicional, pelo fato de executar as duas ferramentas em apenas uma.

iperf

O principal objetivo desta ferramenta [NLA 01] é ajudar no ajuste de conexões TCP em um caminho de rede específico. O principal ajuste é o tamanho da janela TCP. **iperf** mede a largura de banda, perda de pacotes, atraso, variação de atraso e MTU. Usa tanto o TCP (Transmission Control Protocol) quanto o UDP (User Datagram Protocol) para fazer as medições. É necessário executar a ferramenta em dois *hosts*, um servidor e um cliente.

Entre as vantagens da ferramenta **iperf** estão o ajuste fino no tamanho da janela TCP, que pode aumentar a largura de banda atingida nas conexões. Outra vantagem é que a variação de atraso medida ajuda na verificação de qualidade para os serviços de tempo real.

Uma desvantagem do **iperf** é a necessidade de ser executada nos dois *hosts* da conexão. Outra desvantagem é a geração de tráfego na rede pois também realiza medição ativa.

Uma limitação no uso da ferramenta **iperf** é o uso do tipo de serviço que nem sempre é possível utilizar. Isso é devido ao fato de que alguns roteadores ignoram este campo.

tcptrace

tcptrace [TCP 01] é uma ferramenta usada para gerar estatísticas detalhadas sobre conexões TCP tendo como entrada arquivos *dump*. Estes arquivos são resultados das execuções de ferramentas como o **tcpdump**. Esta ferramenta pode ser executada no FreeBSD, NetBSD, Linux, Darwin/OSX (Mac) e Tru64 (Alpha).

tcptrace é executada em linha de comando e usa a ferramenta gráfica **xplot** para criar os gráficos [XPL 01].

tcpanaly

tcpanaly foi desenvolvida por Vern Paxson [PAX 97] e analisa *traces* de pacotes capturados na rede através da ferramenta **tcpdump**. **tcpanaly** tem informações de várias implementações TCP (família Tahoe, Reno e outras), com isso pode identificar a implementação TCP usada através do *trace*. Se a ferramenta encontra inconsistência entre as informações no *trace* e a implementação, ela realiza o diagnóstico de valores de tamanho de janela de

congestionamento e MSS, entre outros. Tal diagnóstico da ferramenta ajuda a determinar como a implementação TCP se comporta.

tcpanaly tem a vantagem de não gerar tráfego extra na rede, além de não precisar da implementação do TCP a ser analisada, apenas um *trace* obtido em um ponto da rede é necessário. Desta forma, pode-se analisar implementações TCP cujo código fonte não é disponível, onde não seria possível recorrer à instrumentação ou manipulação da implementação.

A desvantagem é que **tcpanaly** deve ser modificado a cada implementação TCP que surge, a fim de poder identificá-la posteriomente. Outra desvantagem é o fato da ferramenta estar sujeita aos erros de filtragem do **tcpdump** como, por exemplo, pacotes não capturados, pacotes duplicados (se o **tcpdump** for executado no mesmo *host* que gera o tráfego), reordenação dos pacotes capturados e imprecisão do tempo em que os pacotes foram capturados.

Devido ao fato da ferramenta **tcpanaly** analisar *traces* apenas, os resultados ficam limitados às informações geradas pelo **tcpdump**.

2.4.2 Ferramentas para Avaliação de Perda de Pacotes

As ferramentas mais simples de medição de perda de pacotes transmitem pacotes ICMP para um *host* destino e aguardam uma resposta destes pacotes. Se a resposta não chega em um certo tempo, a ferramenta considera que o pacote foi perdido. Outras ferramentas usam a mesma técnica, porém transmitindo pacotes UDP. Técnicas mais sofisticadas diferenciam perda de pacotes nos dois sentidos, ida e volta.

sting

sting, desenvolvida por Stefan Savage [SAV 99], mede a taxa de perda de pacotes entre dois *hosts*. Esta ferramenta diferencia a taxa de perda dos pacotes nos dois sentidos, ou seja, do transmissor para o receptor e do receptor para o transmissor.

Ao executar a ferramenta o usuário informa o número de pacotes a serem transmitidos, o *host* e a porta de destino. Como resultados a ferramenta informa a quantidade de pacotes enviada e recebida, e as taxas de perda em ambos os sentidos.

A técnica utilizada por Stefan Savage [SAV 99] é dividida em duas fases. Na primeira fase o *host* que executa a ferramenta transmite uma série de pacotes TCP em seqüência para o *host* receptor e vai contabilizando os *acks* recebidos. Na segunda fase o transmissor envia um pacote com um número de sequência seguinte ao último transmitido. Se o receptor responde com um *ack* deste pacote, então não houve perda. Senão, o receptor responde com um *ack* indicando um pacote perdido. Para cada *ack* de pacote perdido, o transmissor retransmite o pacote e contabiliza a perda. O transmissor repete o último passo até que todos os pacotes

transmitidos na primeira fase sejam confirmados pelo receptor. No final deste processo o transmissor sabe a quantidade de pacotes recebidos no receptor e conseqüentemente a taxa de pacotes perdidos. Para saber a taxa de perda no sentido inverso, o transmissor considera que a quantidade de *acks* enviada pelo receptor é igual à quantidade de pacotes recebidos. A taxa de perda no sentido inverso é igual à razão entre o número de *acks* recebidos no transmissor na primeira fase e o número de *acks* enviados pelo *host* receptor.

sting apresenta algumas vantagens com relação às outras ferramentas do gênero. Esta ferramenta mede a taxa de perda nos dois sentidos (ida e volta) independentemente, possibilitando uma melhor investigação do comportamento da rede. Embora diferencie as taxas de perda, a ferramenta não precisa ser executada em ambos os *hosts*. Outra vantagem é que não depende do protocolo ICMP.

Uma desvantagem da ferramenta é a adição de tráfego na rede. Como não faz monitoramento do tráfego da rede, **sting** fica limitada a calcular a taxa de perda apenas dos pacotes gerados pela ferramenta.

2.4.3 Ferramentas para Captura de Pacotes

Algumas ferramentas não geram medidas como resultado mas são utilizadas para armazenar pacotes que trafegaram em um determinado trecho de rede, registrando-os em arquivos que, posteriormente, são analisados ou utilizados como entrada para outras ferramentas de análise como, por exemplo, **tcptrace** e **tcpanaly**.

Embora existam outras ferramentas com esta finalidade [NLA 01], a única ferramenta estudada neste trabalho foi o **tcpdump**.

tcpdump

tcpdump [NLA 01] é executada em linha de comando. A ferramenta imprime o cabeçalho dos pacotes que trafegam por uma determinada interface de rede, que são selecionados por um filtro especificado quando a ferramenta é iniciada. Este filtro é usado para escolher os pacotes que devem ser capturados.

Uma vantagem do **tcpdump** é que os pacotes capturados podem ser armazenados e usados como carga real em simulações de rede, ou utilizados para uma análise mais detalhada do tráfego incluindo o conteúdo de dados.

A ferramenta não gera estatísticas para uma análise de desempenho, apenas captura os pacotes da rede. Para gerar estatísticas é necessário outra ferramenta como **tcptrace**, por exemplo.

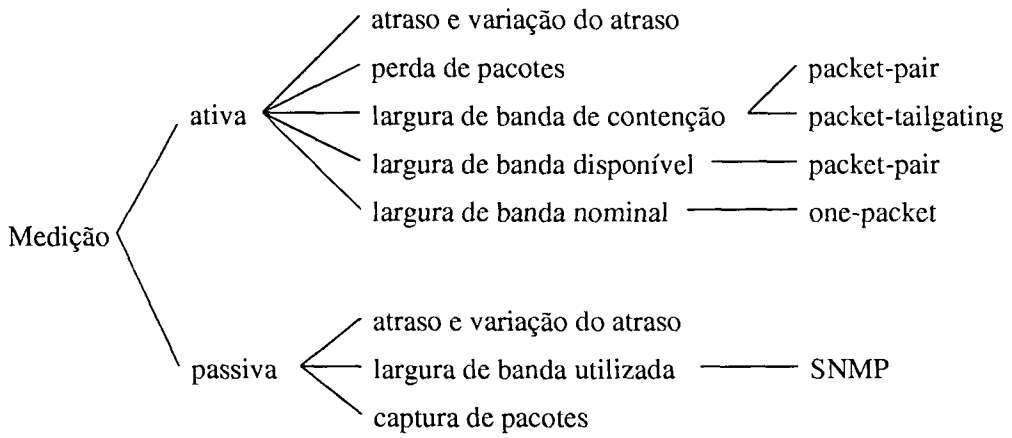

Figura 2.2: Classificação das métricas e técnicas de medição de desempenho em redes.

2.5 Classificação das Ferramentas

Há um grande número de ferramentas disponíveis e todas com características diferentes. Apesar da diversidade de métricas, técnicas e ferramentas no contexto de avaliação de desempenho de redes, não encontramos uma metodologia para classificação destas ferramentas. Após o estudo de diversas técnicas e ferramentas, identificamos alguns critérios que podem ser utilizados para classificá-las. A classificação proposta neste trabalho está baseada em três critérios, criando três níveis de divisão. O primeiro critério é o tipo da medição, portanto, num primeiro nível, as ferramentas estão divididas entre ferramentas de medição ativa e de medição passiva. Entre as ferramentas de medição passiva estão incluídas as de monitoramento de rede que apenas capturam o tráfego e geram um registro dos pacotes. O segundo critério é a métrica medida, portanto, num segundo nível, as ferramentas estão divididas de acordo com a métrica calculada e o terceiro critério divide as ferramentas de acordo com a técnica utilizada para a medição.

A classificação quanto às técnicas utilizadas na medição foi feita apenas para a métrica largura de banda. A Figura 2.2 apresenta as métricas e técnicas divididas de acordo com a classificação proposta.

As ferramentas de avaliação conhecidas podem então ser classificadas de acordo com a classificação proposta. A Tabela 2.1 apresenta as ferramentas descritas neste trabalho, classificadas de acordo com o modelo de classificação apresentado na Figura 2.2.

As ferramentas utilizam vários protocolos. Assim, outra forma de classificá-las seria utilizando o critério de protocolo utilizado. A Tabela 2.2 apresenta uma classificação para as ferramentas em função da métrica utilizada e do protocolo.

Métrica	Medição Ativa	Medição Passiva
Atraso	ping, traceroute, mtr, iperf, pathchar, pchar, clink	tcptrace, tcpanaly
Variação de atraso	iperf	
Perdas	ping, mtr, iperf, sting	
Largura banda contenção	bprobe, pathrate, nettimer	
Largura banda utilizada		MRTG, linkstat, tcptrace
Largura banda disponível	iperf, cprobe	
Largura banda nominal	pathchar, pchar, clink	

Tabela 2.1: Ferramentas divididas segundo a métrica e tipo de medição.

Métrica	ICMP	TCP	SNMP	UDP
Atraso	ping, mtr, traceroute, pathchar, pchar, clink	iperf, tcptrace		iperf, pathchar, pchar, clink
Variação de atraso		iperf		
Perdas	ping, mtr	iperf, sting		iperf
Largura de banda de contenção	bprobe	nettimer		pathrate
Largura de banda utilizada		tcptrace	MRTG, linkstat	
Largura de banda disponível	cprobe	iperf		iperf
Largura de banda nominal	pathchar, pchar, clink			pathchar, pchar, clink

Tabela 2.2: Ferramentas divididas segundo o protocolo usado.

Capítulo 3

Técnicas e Ferramentas para Medição de Largura de Banda

Neste capítulo estão descritas as técnicas para avaliação de largura de banda e algumas ferramentas que implementam estas técnicas. A largura de banda tem um papel importante nas aplicações de rede, principalmente na Internet. O conhecimento da largura de banda pode trazer benefícios para qualquer aplicação que realiza transferências de dados através de uma rede, pois permite a definição de parâmetros de velocidade de transferência e a escolha de caminhos de rede, entre outros. Para fornecer medidas precisas de largura de banda os pesquisadores propõem várias técnicas e metodologias de medição e implementam ferramentas baseadas nestas técnicas.

3.1 Técnicas para Avaliação de Largura de Banda

Nesta seção estão apresentadas as técnicas de medição de largura de banda. Vários pesquisadores têm estudado técnicas para medir largura de banda em redes de computadores. O desenvolvimento de técnicas cada vez mais eficientes tem como objetivo oferecer medidas mais confiáveis e rápidas da largura de banda para os serviços, para os administradores de redes, para dimensionar e configurar novos serviços e para realizar o planejamento de capacidade de enlaces.

Entre as técnicas de medição de largura de banda conhecidas estão a técnica *one-packet* usada por V. Jacobson [JAC 97], Allen Downey [DOW 99] e Bruce Mah [MAH 99a], a técnica *packet-pair* usada por Carter e Crovella [CAR 96b], Vern Paxson [PAX 99], Lai e Baker [LAI 99], Srinivasan Keshav [KES 91a][KES 91b] e Constantinos Dovrolis [DOV 01] e a técnica *multi-packet* [LAI 00]. Lai e Baker também descrevem a técnica *packet-tailgating* em [LAI 00].

De forma geral, as técnicas consistem em enviar um ou mais pacotes pela rede e medir

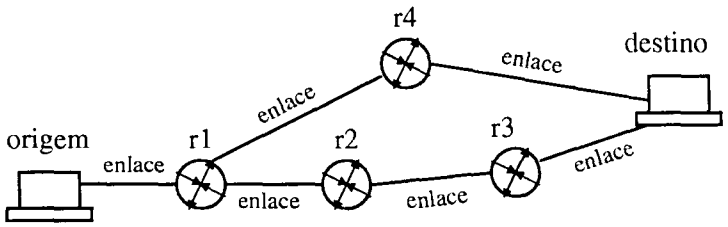

Figura 3.1: Representação do modelo de rede.

o tempo necessário para o pacote chegar até um *host* destino. O modelo de rede segue o modelo da Internet, onde existem um ou mais caminhos entre dois *hosts*, origem e destino. Os caminhos são formados por um ou mais enlaces, cuja capacidade de transmissão de dados varia de enlace para enlace. Entre cada par de enlaces adjacentes existe um roteador, cuja função é examinar o endereço de destino de cada pacote e repassá-lo ao próximo roteador. Na Figura 3.1 estão representados dois *hosts*, `origem` e `destino`, e dois caminhos de rede distintos entre eles, `origem`-r1-r2-r3-`destino` e `origem`-r1-r4-`destino`, onde r1, r2, r3 e r4 são roteadores intermediários.

3.1.1 Técnica *One-Packet*

A técnica *one-packet* é assim denominada em [LAI 00] e descrita em [LAI 00] e [DOW 99]. Esta técnica consiste em enviar vários pacotes de tamanhos variados e fazer medições individuais para cada pacote e enlace. Para medir cada enlace, o campo TTL do cabeçalho IP é primeiramente inicializado em 1 e, posteriormente, acrescido de 1 a cada enlace medido. Para cada pacote transmitido, o tempo até o transmissor receber uma resposta ICMP de erro é medido e, através de análise estatística, são calculados a latência e a largura de banda de cada enlace do caminho.

A análise é baseada no modelo de rede apresentado na Figura 3.2, extraída de [DOW 99]. Antes de um pacote sair do nodo n-1, este espera em uma fila (q1) até ser realmente colocado no enlace. O tempo de transmissão do pacote no enlace é uma função linear do tamanho do pacote e o atraso de transporte é dado por $latência + (tamanho\_pacote/largura\_banda)$. No nodo n, o pacote espera em fila (q2) novamente até que o roteador o processe e gere uma mensagem de erro. O pacote de erro espera em fila no nodo n (q3), então retorna ao nodo n-1 com um tempo de transporte igual a $latência + (tamanho\_pacote\_erro/largura\_banda)$, onde *tamanho_pacote_erro* é o tamanho de um pacote de erro ICMP de 56 bytes. Quando o pacote de erro chega no nodo n-1, este novamente espera na fila (q4) até ser processado. Portanto, o RTT do pacote é dado por:

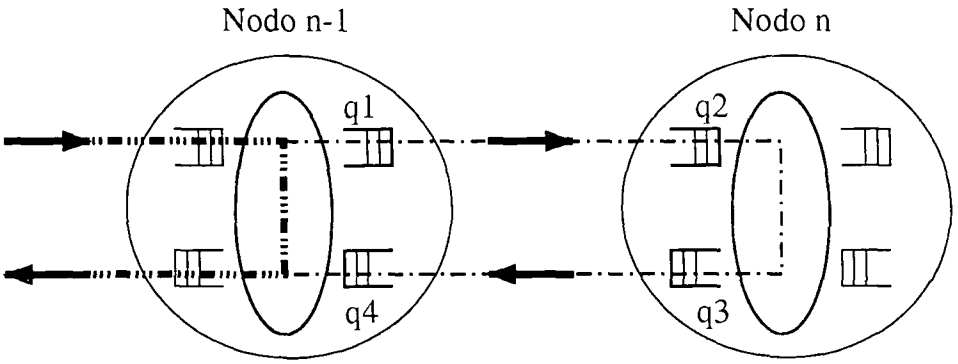

Figura 3.2: Representação das filas em *hosts* subseqüentes.

$$RTT = q1 + (latência + tamanho\_pacote/largura\_banda) + q2 + processamento + q3 + (latência + tamanho\_pacote\_erro/largura\_banda) + q4$$

onde q1, q2, q3 e q4 são os tempos em fila e *processamento* é o tempo de processamento do pacote transmitido e geração e recepção do pacote de erro no nodo n.

Para simplificar este cálculo, a técnica *one-packet* baseia-se em três pressupostos: o tamanho do pacote de erro é pequeno o suficiente, de forma que $tamanho\_pacote\_erro/largura\_banda$ é insignificante; o tempo de processamento no nodo n é insignificante; e, se um grande número de medições for realizado, ocorrerá pelo menos uma medição com tempos de fila próximos de zero. Eliminando os tempos de fila e processamento obtemos a equação base para a técnica:

$$RTT = \left(latência + \frac{tamanho\_pacote}{largura\_banda}\right) + latência \tag{3.1}$$

A latência é calculada utilizando-se regressão linear com base nos tempos de ida e volta (RTT). Conhecendo o RTT, a latência do enlace e o tamanho do pacote transmitido, obtém-se a largura de banda através da fórmula 3.1.

Na Figura 3.3, extraída de [LAI 00], está representada a transmissão de um pacote através de dois enlaces, mostrando os tempos de transmissão e latência. Os valores do tamanho de pacote, largura de banda e latência são, neste exemplo: $s^0 = 6000$ bits, $b_0 = 2$ Mb/s, $d_0 = 2$ ms, $b_1 = 3$ Mb/s e $d_1 = 5$ ms. As variáveis estão definidas na Tabela 3.1. Neste exemplo, temos um pacote de 6000 bits que começa a ser transmitido no enlace 0 no tempo t_0^0. Após o tempo de transmissão e propagação (latência), o pacote se encontra no roteador anterior ao enlace 1 no tempo t_1^0. Neste momento este roteador inicia a transmissão

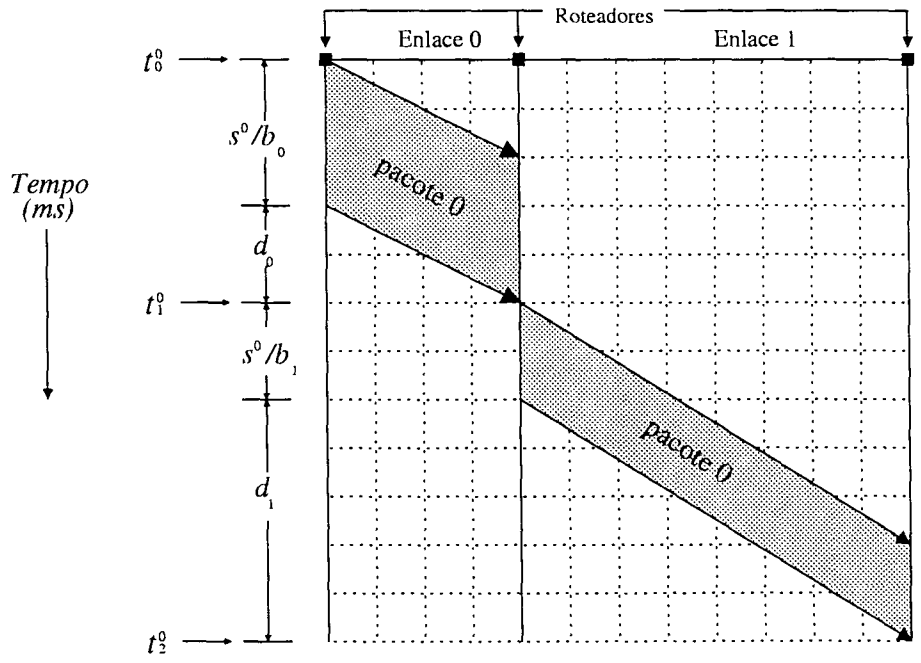

Figura 3.3: Representação do tempo de transporte na técnica *One-Packet*.

do pacote no enlace 1. Novamente após o tempo de transmissão e propagação, o pacote se encontra no próximo roteador no tempo t_2^0.

Variável	*Unidade*	*Definição*
d_l	segundos	latência no enlace l
b_l	b/s	largura de banda do enlace l
s^k	bits	tamanho do pacote k
t_l^k	segundos	tempo em que o pacote k chega completamente no enlace l
q_l^k	segundos	tempo que o pacote k fica enfileirado no enlace l

Tabela 3.1: Variáveis usadas nas equações da seção 3.1.

A técnica *one-packet* é baseada em alguns pressupostos:

1. o tempo de transmissão é proporcional ao tamanho do pacote;
2. os roteadores seguem a política de *store-and-forward* (recebe todo o pacote antes de começar a repassá-lo ao próximo roteador);
3. os enlaces são formados por um único canal;
4. alguns pacotes da medição não são enfileirados ao longo do caminho;
5. as respostas ICMP são prontamente retornadas;

6. o uso do campo TTL não identifica nodos invisíveis, que não tem um endereço IP e são *store-and-forward*, porém estes nodos podem gerar atrasos na medição, diminuindo a precisão.

Estes pressupostos nem sempre são verdadeiros e é necessário acrescentar alguns comentários a respeito de alguns deles. Com relação à política *store-and-forward*, como a técnica leva em consideração o tempo de transmissão e o tempo de propagação, se o roteador não seguisse a política de *store-and-forward* e o pacote de retorno ICMP fosse transmitido antes do término da chegada do pacote da medição, o atraso medido no enlace seria menor e a largura de banda resultante seria superestimada. O pressuposto de que os enlaces são formados por um único canal nem sempre é verdadeiro pois há enlaces formados por vários canais. O pressuposto de que as respostas ICMP são prontamente retornadas nem sempre é verdadeiro pois alguns roteadores diminuem a prioridade de processamento destes pacotes.

A técnica fica limitada pelo pressuposto de que os enlaces são formados por apenas um canal. Se um enlace de 128Kb/s é composto por dois canais de 64Kb/s, então a largura de banda informada pela técnica será de apenas 64Kb/s. Outro fator que limita esta técnica é o pressuposto de que outro tráfego não causa o enfileiramento dos pacotes. Isto acontece porque não é possível saber quando pacotes de outros serviços foram transmitidos no enlace em questão. Na Internet este pressuposto quase sempre é falso [LAI 00].

As ferramentas que utilizam esta técnica são **pathchar**, **clink** e **pchar**.

3.1.2 Técnica *Packet-Pair*

Diferentemente da técnica *one-packet*, que calcula a largura de banda de cada um dos enlaces, a técnica *packet-pair* calcula apenas a largura de banda de contenção em um caminho (*bottleneck bandwidth*). Esta técnica baseia-se no conceito de dispersão de pacotes apresentado em 1988 por Van Jacobson [JAC 88]. A idéia fundamental da dispersão de pacotes é que, se dois pacotes são transmitidos juntos por um caminho de rede e se enfileiram juntos no enlace de contenção, ao passarem por este enlace eles serão separados pelo atraso de transporte do primeiro pacote neste enlace. Quando as confirmações de recebimento (*acknowledgment* ou simplesmente *ack*) dos dois pacotes chegam no *host* transmissor, o tempo entre as chegadas é o tempo resultante da dispersão no enlace de contenção. Este tempo é utilizado para calcular a largura de banda de contenção.

A Figura 3.4, adaptada de [JAC 88], é uma representação do conceito de dispersão de pacotes e mostra como dois pacotes transmitidos são separados após o enlace de contenção. Estão representados três enlaces onde o mais estreito é o enlace de contenção. As setas indicam a direção em que os pacotes são transmitidos. A direção horizontal representa o tempo decorrido. Os dois pacotes de mesmo tamanho são transmitidos um atrás do outro. Como o tamanho dos pacotes continua o mesmo durante a transmissão até o *host* receptor,

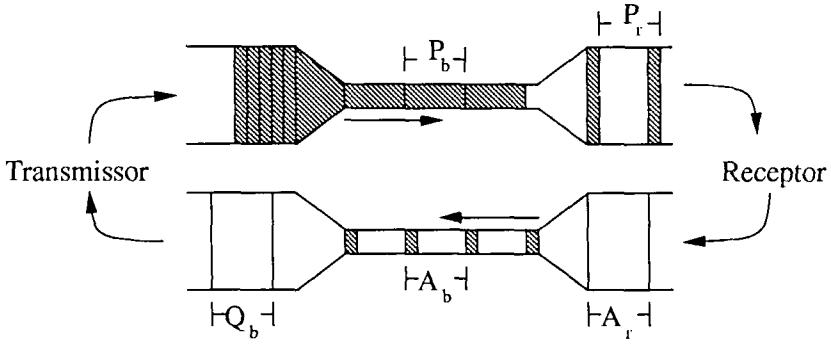

Figura 3.4: Representação da dispersão de pacotes após um enlace de contenção.

ao passarem pelo enlace de contenção os pacotes são "esticados" em tempo. O tempo P_b é o tempo de transmissão do pacote pela banda de contenção. Quando os pacotes chegam no último enlace, o tempo entre as chegadas continua o mesmo, ou seja, $P_r = P_b$, pois pressupõe-se que nenhum outro tráfego é transmitido entre os dois pacotes.

Considerando que os dois pacotes tenham o mesmo tempo de processamento no *host* receptor, então o tempo entre os *acks* continua o mesmo, ou seja, $A_r = P_r$. Como o tamanho de um pacote *ack* é menor que o do pacote que foi transmitido, o tempo para transmiti-lo não vai aumentar, conseqüentemente o tempo entre o par de pacotes também não. Desta forma, o tempo entre as chegadas dos *acks* Q_b é igual a P_b ($Q_b = A_b = A_r = P_r = P_b$). Embora Van Jacobson tenha definido que o tempo entre as chegadas comece no momento em que o primeiro bit do primeiro pacote chega ao transmissor e termine no momento em que o primeiro bit do segundo pacote chega ao mesmo transmissor, os artigos mais atuais [CAR 96b][DOV 01][LAI 00][LAI 01] definem que o tempo entre as chegadas começa com a chegada do último bit do primeiro pacote e termina com a chegada do último bit do segundo pacote.

Denotando a largura de banda de contenção como sendo β e o tamanho dos pacotes transmitidos como b bits, a largura de banda de contenção é calculada por:

$$Q_b = \frac{b}{\beta} \tag{3.2}$$

Esta técnica é baseada em alguns pressupostos que nem sempre são reais nas redes atuais e acabam limitando a confiabilidade dos resultados:

1. os pacotes enviados não são reordenados no caminho de rede, ou seja, o primeiro pacote a ser transmitido é o primeiro a chegar ao *host* destino e o primeiro a retornar ao transmissor;

2. o caminho de rede por onde trafegam os pacotes não muda com freqüência;

3. a largura de banda de contenção é a mesma em ambas as direções (ida e volta);

4. os dois pacotes são transmitidos com um intervalo de tempo suficientemente pequeno tal que sejam enfileirados juntos no enlace de contenção, impedindo que outro tráfego seja enfileirado entre os pacotes da medição;

5. os dois pacotes da medição se enfileiram juntos no enlace de contenção e não em um enlace posterior.

Com relação ao primeiro pressuposto, se houvesse reordenação dos pacotes, o tempo entre as chegadas seria negativo, dependendo da implementação da ferramenta, consequentemente impróprio para o cálculo da largura de banda. O pressuposto de que o caminho de rede não muda com freqüência é necessário porque, como as ferramentas não transmitem apenas um par de pacotes, o resultado da medição se tornaria impreciso se o tempo entre chegadas de um par de pacotes fornecesse a largura de banda de um enlace e outro par de pacotes fornecesse a largura de banda de outro enlace, pois se o caminho de rede mudar durante a medição, a banda de contenção também pode mudar. Com relação ao pressuposto de que a banda de contenção é a mesma em ambas as direções, se a banda de contenção no caminho de volta for menor que a banda de contenção no caminho de ida, o tempo entre as chegadas dos dois pacotes refletirá a banda de contenção do caminho de volta, que é menor. Se a banda de contenção no caminho de volta for maior que a banda de contenção no caminho de ida, não haverá problema pois a dispersão do par de pacotes não aumentaria. O pressuposto de que outro tréfego não se enfileire entre o par de pacotes pode não acontecer, principalmente em redes congestionadas. Se outro tráfego se enfileirar entre o par de pacotes, a largura de banda resultante será subestimada pois o tempo entre as chegadas aumentaria. Com relação ao último pressuposto, se o par de pacotes da medição se enfileirar em um enlace posterior ao de contenção, a largura de banda resultante refletirá a menor largura de banda posterior ao enlace de contenção.

As ferramentas que utilizam esta técnica são **bprobe**, **cprobe** e **pathrate**.

3.1.3 Técnica *Multi-Packet*

Diferentemente das técnicas *one-packet* e *packet-pair* descritas anteriormente, a técnica *multi-packet* mede o atraso de todos os pacotes em um certo fluxo de dados. Esta técnica também leva em consideração os tempos que os pacotes ficam enfileirados ao longo do caminho.

A Figura 3.5, adaptada de [LAI 00], mostra a transmissão de três pacotes através dos enlaces *l-1* e *l* e os tempos de transmissão e fila. Primeiramente o pacote $k-1$ chega no enlace l e, como não há fila, o roteador inicia sua transmissão pelo enlace. Quando o pacote

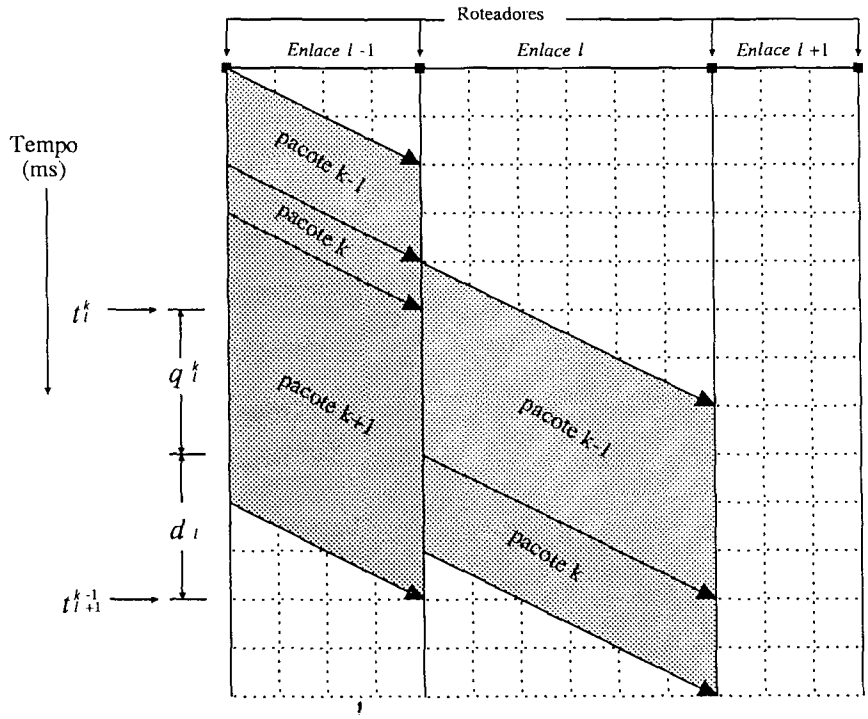

Figura 3.5: Representação do tempo de transporte na técnica *multi-packet*.

k chega ao enlace l, este fica enfileirado pois o pacote $k-1$ ainda está sendo enviado. O pacote k apenas começa a ser transmitido quando o pacote $k-1$ termina de ser transmitido.

A seguinte equação é usada para o cálculo da largura de banda. As variáveis estão definidas na Tabela 3.1:

$$t_l^k = t_0^k + \sum_{i=0}^{l-1} \left(\frac{s^k}{b_i} + d_i + q_i^k \right) \tag{3.3}$$

Esta equação representa o tempo em que o pacote k chega no enlace l, que é a soma do tempo em que o pacote foi transmitido mais a soma da latência, do tempo de transmissão e dos atrasos nas filas de todos os enlaces anteriores.

O tempo que um pacote k fica enfileirado no roteador anterior ao enlace l $\left(q_l^k\right)$ é o tempo desde sua chegada no roteador $\left(t_l^k\right)$ até o instante em que este começa a ser transmitido. Este instante, por sua vez, é o tempo em que o pacote anterior (k - 1) chega no próximo roteador $\left(t_{l+1}^{k-1}\right)$ menos a latência daquele enlace (d_l). Portanto, q_l^k é dado por:

$$q_l^k = max\left(0\ ,\ t_{l+1}^{k-1} - d_l - t_l^k\right) \tag{3.4}$$

Se o pacote k chega no roteador anterior ao enlace l após o roteador terminar de processar o pacote *k-1*, o tempo de fila calculado será um valor negativo. Para evitar que o tempo de fila calculado seja negativo, é usada a função max(). Combinando as duas últimas equações,

a equação de atraso para a técnica *multi-packet* é:

$$t_l^k = t_0^k + \sum_{i=0}^{l-1} \left(\frac{s^k}{b_i} + d_i + max\left(0, t_{i+1}^{k-1} - d_i - t_i^k\right) \right) \tag{3.5}$$

Para explicar melhor o tempo de fila, podemos visualizar a Figura 3.5 e atribuir os seguintes valores: $s^{k-1} = 4000$ bits, $s^k = 2000$ bits, $s^{k+1} = 12000$ bits, $b_{l-1} = 2$ Mb/s, $d_{l-1} = 2$ ms, $b_l = 1$ Mb/s, $d_l = 3$ ms. O pacote k fica enfileirado no enlace l por $q_l^k = max\left(0, t_{l+1}^{k-1} - d_l - t_l^k\right) = max(0, 11 - 3 - 5) = 3$ ms.

Assim como nas técnicas anteriores, esta também é baseada em alguns pressupostos:

1. pressupõe que pacotes de outros fluxos não se enfileiram no fluxo que está sendo medido;

2. pressupõe que o tempo de transmissão é proporcional ao tamanho do pacote;

3. pressupõe que os roteadores intermediários são *store-and-forward.*

Como foi visto anteriormente, a presença de pacotes de outros fluxos subestimaria o resultado da medição. Para resolver este problema, também são usados os atrasos mínimos. O pressuposto de que o tempo de transmissão é proporcional ao tamanho do pacote permite obter a latência do enlace através de regressão linear utilizando os valores mínimos dos atrasos.

Esta técnica não está implementada em nenhuma das ferramentas avaliadas neste trabalho mas é a base para a técnica *packet-tailgating* descrita a seguir.

3.1.4 Técnica *Packet-Tailgating*

Esta técnica foi desenvolvida a partir da técnica *multi-packet* [LAI 00]. Nesta técnica são transmitidos dois pacotes juntos. A técnica pressupõe que o primeiro pacote não se enfileira até alcançar seu destino e o segundo pacote se enfileira atrás do primeiro até o enlace de contenção e em nenhum enlace posterior a este.

Primeiramente os autores modificam a equação 3.5 da técnica *multi-packet* para calcular a largura de banda do enlace l_q, o qual apresenta fila. Esta largura de banda é denominada b_{l_q}. A seguinte equação representa o tempo que o pacote k demora para ser transmitido até um determinado enlace n (variáveis definidas na Tabela 3.1):

$$t_n^k = t_0^k + \sum_{i=0}^{n-1} \left(\frac{s^k}{b_i} + d_i + max\left(0, t_{i+1}^{k-1} - d_i - t_i^k\right) \right)$$

Assumindo fila apenas no enlace l_q, o tempo t_n^k é dividido entre o tempo até o enlace l_q, o tempo no enlace l_q e o tempo após o enlace l_q:

$$t_n^k = \left[t_0^k + \sum_{i=0}^{l_q-1} \left(\frac{s^k}{b_i} + d_i \right) \right] + \left[\frac{s^k}{b_{l_q}} + t_{l_{q+1}}^{k-1} - t_{l_q}^k \right] + \left[\sum_{i=l_{q+1}}^{n-1} \left(\frac{s^k}{b_i} + d_i \right) \right]$$

Simplificando a equação acima, chega-se à equação usada pela técnica [LAI 00]:

$$b_{l_q} = \frac{s^{k-1}}{\left(t_n^k + \frac{s^k - s^{k-1}}{b^{l_q-1}} - \frac{s^k}{b^{n-1}} - t_0^{k-1} - d^{n-1} \right)} \tag{3.6}$$

Assim, pode-se calcular a largura de banda no enlace onde ocorre fila b_{l_q} através do tamanho de dois pacotes s^{k-1} e s^k, o tempo de chegada do segundo pacote t_n^k, o tempo de transmissão do primeiro pacote t_0^{k-1}, a largura de banda de enlaces anteriores b^{l_q-1} e b^{n-1} e a latência de todos os enlaces anteriores d^{n-1}.

A técnica é dividida em duas fases. A primeira fase é denominada *sigma* e a segunda fase é denominada *tailgating*. Na fase *sigma* a técnica calcula os valores b^{n-1} e d^{n-1} da equação 3.6. Para isso, transmite pacotes de diferentes tamanhos e usa regressão linear com base nos atrasos mínimos de cada tamanho.

Na fase *tailgating* são calculadas as outras variáveis da equação 3.6. Para isso, é transmitido um pacote com o maior tamanho possível (geralmente 1500 bytes) com o campo TTL do protocolo IP programado para expirar no enlace l_q, seguido por um pacote com o menor tamanho possível. O pacote menor tem um tempo de transmissão menor que o pacote maior. Quando estes dois pacotes são transmitidos, o pacote menor (chamado *tailgater*) se enfileira continuamente atrás do pacote maior (chamado *tailgated*) devido ao seu tamanho menor. Devido ao valor do TTL no pacote maior, este é transmitido até o enlace l_q, e o pacote menor continua a ser transmitido sem atraso em fila até o *host* destino.

Na fase *tailgating* também é usada a regressão linear usando os menores valores de atraso do pacote menor para estimar a latência. Primeiramente o TTL é inicializado com o valor 1 e incrementado até alcançar o destino, assim as larguras de banda de cada enlace vão sendo calculadas.

A única ferramenta que utiliza esta técnica é **nettimer**.

3.2 Discussão sobre as Técnicas de Medição de Largura de Banda

Cada técnica apresentada na seção anterior tem alguma particularidade. Com relação à metrica estimada, por exemplo, as técnicas *one-packet*, *multi-packet* e *packet-tailgating* calculam a largura de banda de todos os enlaces ao longo de um caminho enquanto que a técnica *packet-pair* calcula apenas a largura de banda de contenção.

O tempo para executar a medição varia de acordo com a metodologia adotada pela ferramenta. Quanto maior a quantidade de pacotes transmitidos e maior o tempo entre as transmissões, maior é o tempo de execução. O tempo varia também de acordo com a precisão da estimativa calculada pela ferramenta.

Todas as técnicas baseiam-se em alguns pressupostos. Alguns pressupostos, caso não sejam verdadeiros, aumentam ou diminuem o atraso medido e consequentemente diminuem ou aumentam a largura de banda resultante, como ocorre nos pressupostos de que o tempo de processamento nos destinos e o tempo de transmissão de pacotes de erro ICMP sejam insignificantes e de que alguns pacotes da medição não se enfileirem. Um pressuposto comum a todas as técnicas é que os roteadores são *store-and-forward.* Caso não fossem, a largura de banda resultante poderia ser maior que a real como explicado nos pressupostos da técnica *one-packet.* Outra situação é a presença de tráfego e fila nos roteadores, o que ocasiona aumento do atraso medido e aumento do tempo entre as chegadas dos pacotes da medição. Outro pressuposto é sobre a reordenação dos pacotes da medição que, caso ocorra, resulta em tempo entre chegadas negativo. O pressuposto de que não ocorre mudança de rota no caminho de rede, caso ocorra, pode ocasionar mudança do enlace de contenção. Outro pressuposto diz respeito à presença de redes assimétricas, ou seja, o caminho de ida é diferente do caminho de volta do pacote, pois neste caso o enlace de contenção no caminho de ida pode ser diferente do enlace de contenção no caminho de volta.

Estas técnicas surgiram nos últimos cinco anos. Todas apresentam alguma desvantagem ou limitação que as impede de fornecer resultados mais exatos ou mais rápidos. Recentemente os pesquisadores têm estudado fatores de rede como comportamento de atraso e congestionamento para desenvolver técnicas melhores. Isto mostra que ainda há trabalho a ser realizado nesta área e as técnicas existentes ainda precisam evoluir.

3.3 Ferramentas para Avaliação de Largura de Banda

Esta seção apresenta as ferramentas estudadas usadas para medir largura de banda, sendo que algumas foram testadas neste trabalho e outras não. Para cada ferramenta apresentada são analisadas as vantagens, desvantagens, a técnica usada, qual o uso mais indicado para a ferramenta e limitações das mesmas. As ferramentas estão apresentadas de acordo com a métrica estimada: largura de banda de contenção, nominal, disponível e utilizada.

3.3.1 Largura de Banda de Contenção

bprobe

Esta ferramenta foi desenvolvida por Robert Carter e Mark Crovella [CAR 96b]. **bprobe** calcula a largura de banda do enlace de contenção de um caminho de rede.

A técnica usada pela ferramenta é a *packet-pair*. **bprobe** transmite pacotes ICMP ECHO consecutivos e registra os tempos entre as chegadas das respostas ICMP. Uma das vantagens desta ferramenta é que a medição é rápida o suficiente para ser usada em aplicações de tempo real tais como navegadores e clientes ftp [CAR 96b]. Apesar de introduzir tráfego na rede, a quantidade de tráfego gerado não é alta, considerando que a transmissão de dois pacotes é suficiente para se ter uma medida da largura de banda. Uma terceira vantagem de **bprobe** é que apenas é necessário executar a ferramenta em um *host*, não sendo necessária a execução no *host* destino ou em *hosts* intermediários.

A ferramenta **bprobe** apresenta algumas desvantagens advindas da técnica *packet-pair*, na qual é baseada. Em redes assimétricas, ou seja, o caminho de ida é diferente do caminho de volta, o resultado pode não ser confiável pois a largura de banda de contenção pode não ser a mesma. Outra desvantagem é o fato da ferramenta utilizar o tempo de envio e retorno dos pacotes (RTT), pois o resultado é menos exato que se fosse medido apenas o tempo de envio.

As limitações são aquelas decorrentes da técnica *packet-pair*. Porém, para cada problema a ferramenta apresenta uma solução. Com relação à falha de enfileiramento, **bprobe** transmite vários pares de pacotes para garantir que os pacotes enviados sejam enfileirados no ponto de contenção, variando seu tamanho. Pacotes maiores têm mais chance de se enfileirar.

Outra limitação é a presença de tráfego entre o par de pacotes. Para evitar que outros pacotes se enfileirem entre o par de pacotes, a ferramenta transmite um grande número de pacotes, aumentando a possibilidade de um par chegar junto no ponto de contenção.

Uma terceira limitação é aquela relacionada com a perda de pacotes, quanto maior o pacote transmitido, maior a chance de perda deste pacote. Para evitar este problema a ferramenta transmite pacotes de vários tamanhos e vários pacotes do mesmo tamanho.

pathrate

A ferramenta **pathrate** foi desenvolvida por Constantinos Dovrolis e é apresentada em [DOV 01]. A ferramenta utiliza o protocolo UDP para fazer as medições e uma conexão TCP para controle dos pacotes transmitidos e recebidos. Além disso, **pathrate** utiliza o atraso apenas de ida do pacote, por isso deve ser executada nos dois fins do caminho de rede, onde os *hosts* utilizados são denominados transmissor e receptor. A ferramenta utiliza, além de pares de pacotes da técnica *packet-pair*, trilhas de pacotes.

A execução da ferramenta é dividida em duas fases. Na primeira fase, chamada *packet-pair probing*, um grande número de pares de pacotes são transmitidos do *host* transmissor até o *host* receptor. Como resultado, obtém-se uma distribuição da largura de banda com uma ou mais modas, cada uma indicando uma possível largura de banda de contenção. Isto

acontece porque os pacotes podem se enfileirar em outros enlaces que não o de contenção. Nesta fase a ferramenta calcula o tamanho do *bin*, ou seja, o tamanho do intervalo da distribuição, que também será a resolução final da largura de banda de contenção, pois o resultado da ferramenta não é um valor discreto de largura de banda e sim um intervalo.

Se a distribuição resultante apresentar apenas uma moda, está é considerada a largura de banda resultante (com pouco tráfego no caminho de rede, a largura de banda é medida com mais facilidade [DOV 01]). Se a distribuição apresentar várias modas, a ferramenta executa a segunda fase, chamada *packet train probing*. Nesta fase **pathrate** transmite séries cada vez maiores de pacotes até a distribuição resultante de largura de banda convergir para apenas uma moda, chamada de ADR (*Asymptotic Dispersion Rate*). A largura de banda de contenção calculada é a primeira moda da primeira fase que é maior que o valor ADR.

pathrate tem a vantagem de ser robusta o suficiente para medir com a mesma exatidão enlaces com pouco ou muito tráfego. Outra vantagem é que o resultado não é prejudicado por assimetria na rede (largura de banda de contenção é diferente nos dois sentidos) devido ao fato da ferramenta medir o atraso somente de ida do pacote.

A ferramenta apresenta pelo menos três desvantagens:

1. devido à necessidade de ser executada nos dois *hosts*, a medição pode ser dificultada;
2. a geração de tráfego é da ordem de dezenas de Kb/s;
3. embora use atraso somente de ida (o que resulta numa maior precisão), o resultado não é um valor discreto e sim um intervalo, por exemplo, de 28 a 29 Mpbs para *bin* de 1Mb/s. O valor do *bin* é calculado pela ferramenta, não podendo ser especificada como parâmetro para a ferramenta.

nettimer

A ferramenta **nettimer** foi elaborada por Kevin Lai e Mary Baker [LAI 00][LAI 01], e tem o objetivo de estimar a largura de banda de contenção de um caminho de rede (*bottleneck bandwidth*). Para realizar os cálculos, a ferramenta realiza medição passiva monitorando o tráfego corrente e medição ativa usando a técnica *packet-tailgating* descrita anteriormente.

Uma vantagem da ferramenta é o tempo de resposta. Devido à agilidade da técnica usada, **nettimer** executa rapidamente as medições.

Outra vantagem é a robustez e a flexibilidade. Segundo os autores, **nettimer** pode detectar enlaces formados por vários canais, não depende da velocidade de entrega de pacotes ICMP dos roteadores, não depende de confirmação de pacotes (ACK), rede assimétrica, pode ser executada somente em um *host* ou nos dois fins (transmissor e receptor dos pacotes).

Uma desvantagem da ferramenta, decorrente da técnica usada, é a dificuldade de medir um enlace de alta velocidade após um enlace de baixa velocidade. A presença de fila em

qualquer *host* pelo caminho de rede perturba a medição de todos os enlaces.

Uma limitação da ferramenta, segundo os autores, é que exige recurso computacional no transmissor para que os pacotes sejam transmitidos rapidamente no primeiro enlace com o objetivo de se enfileirarem. A configuração de recurso mínimo de *hardware* não é estipulada pelos autores.

3.3.2 Largura de Banda Nominal

pathchar

A ferramenta **pathchar**, desenvolvida por Van Jacobson [JAC 97], estima a latência e a largura de banda nominal de cada enlace ao longo de um caminho Internet. Esta ferramenta é baseada na técnica *one-packet* de medição de largura de banda.

pathchar utiliza o campo TTL do pacote IP para identificar cada *host* no caminho e medir a largura de banda de cada enlace. **pathchar** transmite uma série de pacotes de tamanhos variados. Para cada pacote transmitido, a ferramenta aguarda uma resposta ICMP de erro e mede o tempo de retorno deste pacote de erro (RTT). Através de análise estatística destas medições, **pathchar** infere o atraso e largura de banda de cada enlace, a distribuição dos tempos em fila e a probabilidade de perda de pacotes.

A ferramenta **pathchar** baseia-se em alguns pressupostos:

- o tamanho do pacote de erro ICMP é pequeno e, por isso, seu tempo de retorno é considerado desprezível;
- o tempo de processamento dos pacotes nos *hosts* é desprezível;
- os pacotes com os menores tempos de ida e volta (RTT), dentre os RTTs medidos para uma quantidade considerável de pacotes, não sofreram atrasos nas filas.

Através de parâmetros passados ao **pathchar**, podemos executar a ferramenta informando o TTL máximo, tamanho mínimo e máximo dos pacotes a serem transmitidos, o número de pacotes de cada tamanho e o número de bytes incrementados entre os tamanhos dos pacotes a serem usados na medição, além da opção de não resolver nomes dos *hosts*. A configuração padrão de **pathchar** transmite 32 pacotes de cada tamanho, variando o tamanho de 64 bytes ao maior tamanho possível (MTU), com incrementos de 32 bytes ao tamanho.

Além de medir as três métricas mais importantes em desempenho de redes (atraso, largura de banda e taxa de perda), **pathchar** realiza medições e apresenta os resultados para todos os enlaces individualmente.

Uma desvantagem é o tempo necessário para se obter os resultados pois a ferramenta transmite algumas centenas de pacotes por enlace e cada pacote só é transmitido quando a

resposta do pacote anterior chega, aumentando o tempo de execução.

Devido ao sistema utilizado para as medições, **pathchar** apresenta algumas limitações:

- a ferramenta não fornece resultados exatos para enlaces com mais de um canal devido às causas descritas na técnica *one-packet*;
- quando o caminho de rede é assimétrico (caminho de ida é diferente do caminho de volta), os resultados não são precisos pois a ferramenta utiliza o RTT;
- a ferramenta apresenta imprecisão nos resultados caso o caminho mude durante a execução da ferramenta pois **pathchar** utiliza medições dos enlaces anteriores para calcular os próximos enlaces.

clink

clink é uma ferramenta baseada no **pathchar**. **clink** foi escrita por Allen Downey e é apresentada em [DOW 99] e [DOW 02]. Esta ferramenta mede a latência e a largura de banda de cada enlace.

A interface e a técnica são baseadas na ferramenta **pathchar**. **clink** utiliza a técnica *one-packet*, enviando pacotes UDP e incrementando o campo TTL para medir o tempo de ida e volta dos pacotes para cada enlace do caminho de rede.

Apesar de usar a mesma técnica que a ferramenta **pathchar**, Downey descreve que **clink** não transmite sempre a mesma quantidade de pacotes para calcular a largura de banda em cada enlace e sim, usa métodos estatísticos para reduzir a quantidade de medições necessárias. Quando **clink** detecta a convergência de um valor para a largura de banda durante o procedimento de medição, a ferramenta finaliza o envio de pacotes.

clink apresenta as mesmas vantagens, desvantagens e limitações da ferramenta **pathchar**, com a vantagem a mais que é a utilização de menos pacotes, consequentemente gerando menos tráfego na rede e diminuindo o tempo de execução.

pchar

A ferramenta **pchar** foi desenvolvida por Bruce A. Mah [MAH 99a]. Esta ferramenta é uma reimplementação do **pathchar**, utiliza a mesma técnica (*one-packet*) e apresenta as mesmas vantagens, desvantagens e limitações que **pathchar**.

A ferramenta **pchar** tem sido atualizada no decorrer do tempo. As atualizações, que incluem suporte ao protocolo SNMP e IPv6, podem ser acompanhadas no endereço Web do autor [MAH 01].

3.3.3 Largura de Banda Disponível

cprobe

Assim como **bprobe**, **cprobe** também foi desenvolvida por Carter e Crovella [CAR 96b]. **cprobe** foi projetada para estimar a largura de banda disponível no enlace de contenção. Os autores também propõem uma seleção dinâmica de servidor baseada nestas duas ferramentas em [CAR 96a].

Para o cálculo da largura de banda disponível, primeiramente a ferramenta calcula a largura de banda de contenção através do algoritmo da ferramenta **bprobe**. Em seguida transmite uma sequência de pacotes ICMP ECHO a uma taxa superior à banda de contenção até o *host* destino e registra o tempo entre o recebimento do primeiro pacote e do último quando estes voltam ao *host* transmissor. A largura de banda disponível é resultado da razão entre a quantidade de bytes transmitidos e a diferença de tempo entre a chegada do primeiro pacote e do último. Quanto maior o tráfego no enlace de contenção, menor será a largura de banda disponível e maior será o tempo de transmissão dos pacotes.

A ferramenta **cprobe** apresenta as mesmas vantagens da ferramenta **bprobe** citada anteriormente. Eventualmente o *host* transmissor pode atrasar o processamento dos pacotes em decorrência de efeitos do sistema operacional, por exemplo. Esta é uma desvantagem do **cprobe**, pois este atraso resulta em um cálculo errôneo da largura de banda estimada, diminuindo-a. Outra desvantagem é que, como utiliza a ferramenta **bprobe** para calcular a largura de banda no enlace de contenção, ela fica sujeita aos erros de medição decorrentes desta ferramenta.

3.3.4 Largura de Banda Utilizada

MRTG

MRTG (*Multi Router Traffic Grapher*) é uma ferramenta de monitoração de tráfego de dados em enlaces de rede. A ferramenta usa o protocolo SNMP (*Simple Network Management Protocol*) para ler informações de tráfego dos roteadores e *switches*. Com estas informações armazenadas, o MRTG gera estatísticas de tráfego e apresenta páginas HTML contendo gráficos da largura de banda utilizada nos enlaces.

Os gráficos gerados pelo MRTG informam a utilização diária, dos últimos sete dias, das últimas quatro semanas e dos últimos doze meses. Os dados apresentados incluem largura de banda máxima utilizada, média e atual. Como estas páginas ficam disponíveis via Web, pode-se criar senhas para o acesso remoto.

Algumas vantagens desta ferramenta:

- dados precisos: a banda utilizada não é calculada levando em consideração pressupostos como em outras ferramentas, e sim lida diretamente das variáveis SNMP;

- embora tenha estatísticas dos últimos 12 meses, os arquivos de registro de utilização não crescem, pois não acumulam todas as informações obtidas, apenas uma média das informações mais antigas;
- facilidade de acesso e interpretação dos gráficos gerados;
- MRTG pode ser executado tanto em plataforma UNIX como em Windows;
- o código é aberto, ou seja, qualquer pessoa tem acesso aos programas fontes;
- permite ainda obter informações de quaisquer outras variáveis SNMP disponíveis nos roteadores;
- como utiliza SNMP, pode ser aplicada em qualquer rede com SNMP.

Entre as desvantagens encontram-se:

- os gráficos não são gerados no momento em que são acessados e sim, a cada 5 minutos por padrão;
- a banda utilizada informada no gráfico não é a banda medida no momento em que o gráfico foi gerado e sim, a média de banda utilizada dos últimos 5 minutos;
- a ferramenta não mede a largura de banda nominal do enlace, apenas a utilizada. A largura de banda nominal do enlace é fornecida nos arquivos de configuração;
- como o SNMP utiliza o protocolo UDP, o MRTG pode ficar sem a informação atualizada em alguns momentos pois a requisição de informação ao roteador pode se perder na rede devido aos fatores já vistos na métrica "perda de pacotes". Não tendo a informação atualizada, o MRTG repete a última medida obtida, podendo gerar uma linha reta horizontal durante alguns períodos, gerando assim, informação errada sobre a utilização do enlace.

Uma limitação desta ferramenta é que ela obtém as informações de desempenho das variáveis SNMP. Portanto as métricas estão restritas às informações fornecidas por estas variáveis. Outra limitação é que a utilização da ferramenta é restrita aos administradores das redes. Assim, só é possível obter dados através desta ferramenta sob a autorização dos administradores de cada rede.

linkstat

Esta ferramenta foi desenvolvida na Universidade do Vale do Rio dos Sinos (UNISINOS) em São Leopoldo, RS. **linkstat** é apresentada em [BAL 00]. Assim como a ferramenta **MRTG**,

linkstat é utilizada para estimar a largura de banda utilizada em enlaces da Internet, a partir de informações obtidas dos roteadores através do protocolo SNMP.

linkstat utiliza a ferramenta RRDtool [CAI 01] de auxílio à manutenção de dados medidos e geração de gráficos. A ferramenta é composta de dois módulos: o primeiro captura as informações de tráfego e o outro recebe as consultas dos usuários e apresenta os gráficos através de página Web.

Algumas vantagens da ferramenta são:

- flexibilidade de visualização do tráfego: o usuário escolhe duração da medição, tamanho do gráfico, mostrar ou não legendas e picos de tráfego;
- facilidade de uso: arquivos de configuração fáceis de serem alterados e a interface com o usuário é simples;
- uso de recursos computacionais sob demanda: diferente do **MRTG** onde os gráficos são atualizados a cada cinco minutos (por padrão), no **linkstat** são gerados apenas quando um usuário executa a ferramenta.

linkstat também utiliza o protocolo UDP e pode ficar sem informação atualizada caso as requisições de informação se percam, resultando em linha reta horizontal no gráfico.

Três limitações são observadas nesta ferramenta: o tempo mínimo de atualização dos dados é de um minuto, as informações de tráfego também ficam limitadas às informações contidas nas variáveis SNMP e seu acesso também depende da autorização dos administradores de rede.

Capítulo 4

Projeto dos Experimentos e Resultados

Neste capítulo apresentamos o projeto dos experimentos para as medições de largura de banda, juntamente com a descrição dos caminhos de rede e velocidades dos enlaces. Apresentamos também os resultados dos experimentos de medição de largura de banda nominal, banda de contenção e banda disponível. No final do capítulo apresentamos ainda uma descrição de alguns fatores que limitam a precisão dos resultados das ferramentas.

4.1 Projeto dos Experimentos

Nesta seção apresentamos o projeto dos experimentos para as medições de largura de banda. A avaliação experimental consiste em executar algumas ferramentas a partir de um único *host* origem, com destino a outros diferentes *hosts*, e comparar os resultados obtidos com as capacidades nominais dos enlaces. Nesta seção estão descritos os caminhos de rede incluídos nos experimentos com os respectivos *hosts* e roteadores intermediários, a velocidade dos enlaces e justificativas para a escolha destes caminhos.

4.1.1 O Ambiente Experimental

Os experimentos foram projetados para avaliar as ferramentas de medição ativa de largura de banda de contenção, de banda disponível e de banda nominal através de execuções destas ferramentas em uma rede operacional de uso comum. Os resultados obtidos foram comparados com as capacidades nominais e também comparados entre si.

Os experimentos foram realizados a partir da rede local do Departamento de Informática da Universidade Federal do Paraná incluindo a rede local e a rede de longa distância entre o POP-PR (Ponto de Presença da RNP no Paraná), POP-MG (Ponto de Presença da RNP em Minas Gerais), a rede interna da Impsat de Curitiba e rede do provedor de acesso

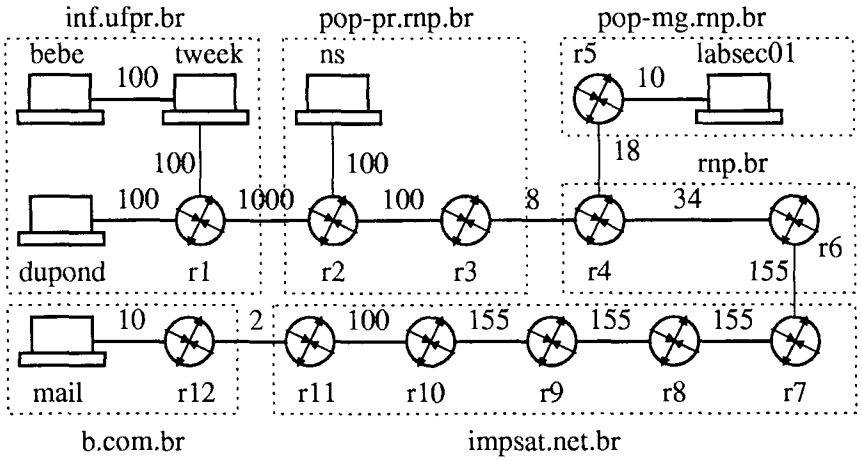

Figura 4.1: Mapa das redes envolvidas nos experimentos.

Business Internet, abrangendo enlaces que variam de 2 Mb/s a 1 Gb/s. O mapa da Figura 4.1 apresenta os computadores utilizados nos experimentos e as conexões de rede entre eles, incluindo os enlaces e os roteadores. As velocidades dos enlaces desta figura estão representadas em Mb/s e os nomes dos roteadores denominados no mapa como r1 a r12 estão listados na Tabela 4.1.

Como pode ser observado na Figura 4.1, os caminhos são de várias distâncias medidas em número de *hops*. O caminho mais curto a partir do computador `tweek` é o caminho `tweek` ↔ `bebe`, de apenas um *hop*. O segundo caminho é o `tweek` ↔ `dupond` de dois *hops*, cujo roteador intermediário r1 é o *gateway* do Departamento de Informática da UFPR. Estes três computadores e o roteador r1 estão conectados através de *switchs* em uma rede local Fast Ethernet. O caminho de três *hops* `tweek` ↔ `ns` tem um enlace de 1 Gb/s ligando o prédio do Departamento de Informática ao POP-PR. A largura de banda de contenção destes três primeiros caminhos é de 100 Mb/s. O segundo caminho mais longo, `tweek` ↔ `labsec01`, tem um comprimento de seis *hops* e a banda de contenção é o enlace de 8 Mb/s que conecta o POP-PR à rede RNP em São Paulo, onde está localizado o roteador r4. Um enlace com velocidade próxima ao de contenção neste último caminho é o enlace de 10 Mb/s da rede local onde está conectado o computador `labsec01`. O caminho mais longo é o `tweek` ↔ `mail`, que é formado por enlaces do Departamento de Informática da UFPR, do POP-PR, da rede RNP, e vários enlaces ATM e *Fast Ethernet* internos da rede da Impsat, além do enlace da rede *Ethernet* do provedor Business Internet. A banda de contenção deste último caminho está no enlace de 2 Mb/s que o provedor contrata da Impsat.

Os caminhos de rede foram escolhidos de forma a incluir LANs e WANs, o que implicou uma grande variação de distância física, variação de velocidades e variação de atrasos. Outro fator que influenciou a escolha foi a boa vontade dos administradores das redes em instalar

Roteador	*Nome do Roteador*	*Endereço IP*
r1	ruivao.inf.ufpr.br	200.17.212.119
r2	gw6611.ufpr.br	200.17.209.33
r3	bb2.pop-pr.rnp.br	200.19.74.20
r4	sp-atm104.bb3.rnp.br	200.143.254.190
r5	mg-atm108.bb3.rnp.br	200.143.254.153
r6	optix-hssil-0.bb3.rnp.br	200.143.254.97
r7	200.185.4.38	200.185.4.38
r8	optiglobe-a1-3-1-core01spo.impsat.net.br	200.186.138.33
r9	coa-g10-0-core01spo.impsat.net.br	200.186.145.105
r10	cta-a4-0-1-core03cta.impsat.net.br	200.186.176.26
r11	cta-f2-1-0-core01cta.impsat.net.br	200.196.88.1
r12	e09230.impsat.com.br	200.186.249.230

Tabela 4.1: Roteadores da Figura 4.1.

computadores próprios, total ou parcialmente dedicados aos experimentos, além de fornecerem informações essenciais ao trabalho tais como, por exemplo, configurações de hardware e software dos equipamentos e capacidades nominais dos enlaces.

Os experimentos foram realizados a partir de um microcomputador do Laboratório de Redes e Sistemas Distribuídos da Universidade Federal do Paraná cujo nome é `tweek`. As configurações deste microcomputador e dos microcomputadores alvo estão descritas na Tabela 4.2.

As ferramentas utilizadas nos experimentos foram **bprobe**, **cprobe**, **pchar**, **clink**, **nettimer** e **pathrate**, todas de medição ativa fim-a-fim, ou seja, ferramentas que geram tráfego na rede e que não precisam ser executadas nos nodos intermediários. Estas ferramentas foram escolhidas pela disponibilidade do código fonte, pela existência de documentação explicando seus fundamentos e por serem ferramentas amplamente referenciadas na literatura. Deve-se ressaltar que existem dezenas de ferramentas disponíveis na Internet [NLA 01]. A ferramenta **pathchar** não foi avaliada por já estar ultrapassada pelas ferramentas **clink** e **pchar** que usam a mesma técnica e também por não estar disponível seu código fonte. A ferramenta **iperf** não foi avaliada pelo fato de não termos encontrado referências bibliográficas sobre a mesma. As ferramentas de medição de largura de banda utilizada não foram experimentadas porque não encontramos ferramentas de medição ativa nem fim-a-fim, apenas ferramentas que monitoram o tráfego nos roteadores.

4.1.2 Metodologia e Validação dos Resultados

Foram realizadas cinco execuções de cada ferramenta. Este número de execuções mostrou-se satisfatório pois foi suficiente para observar a variação dos resultados e a comparação entre as ferramentas. Um número maior de execuções dificultaria o trabalho devido a mudanças no roteamento, principalmente nos caminhos `tweek` ↔ `labsec01` e `tweek` ↔ `mail` que são mais longos. Além disso, as ferramentas são projetadas para fornecerem um resultado satisfatório em uma única execução. Assim, o nosso objetivo com a execução repetida da ferramenta é verificar a característica de repetibilidade. Para as mesmas condições, em execuções repetidas, a ferramenta deve fornecer os mesmos resultados ou, pelo menos, resultados bem parecidos.

Os horários de realização dos experimentos foram bem variados, incluindo manhãs, tardes e noites, de segunda-feira a domingo. Com a variação do horário, acreditamos que testamos as ferramentas com diferentes intensidades de tráfego. Os experimentos foram realizados entre os dias 5 de março de 2002 e 12 de setembro de 2002.

Host	Hardware	Sistema Operacional
`tweek`	IBM Pentium III 800Mhz 192MB RAM placa de rede Accton Tecnology Corporation SMC2-1211TX 100Mb/s	Debian kernel 2.4.12-686
`bebe`	IBM Pentium III 800Mhz 192MB RAM placa de rede Accton Tecnology Corporation SMC2-1211TX 100Mb/s	Debian kernel 2.4.17
`dupond`	AMD Athlon Dual 1312Mhz 2GB RAM placa de rede Tulip 100Mb/s	Linux Woody kernel 2.4.12
`ns`	Power PC R6000 2x233Mhz 256MB RAM placa de rede 3Com 3C905-TX-IBM Fast Etherlink XL NIC 10/100	AIX 4.3.2
`labsec01`	Pentium 166 64MB RAM placa de rede Realtek RTL-8029(AS)	Red Hat 7.2 kernel 2.4.7-10
`mail`	AMD-K6 350Mhz 128MB RAM placa de rede Realtek 8029 10Mb/s	Slackware 8.0 kernel 2.2.19

Tabela 4.2: Microcomputadores utilizados nos experimentos.

Para a validação dos resultados foram utilizadas as velocidades nominais dos enlaces, fornecidas pelos administradores de rede dos *backbones*, por mapas de rede da RNP e gráficos de utilização de enlace. Estes dados estão apresentados na Figura 4.1.

Este ambiente experimental variado nos permite testar a precisão, a robustez e o tempo

de execução das ferramentas.

4.2 Resultados

Nesta seção apresentamos e discutimos os resultados dos experimentos. A seção está dividida em subseções de acordo com a métrica avaliada. A subseção 4.2.1 apresenta os resultados dos experimentos de medição de largura de banda nominal. A subseção 4.2.2 mostra os resultados dos experimentos de medição de largura de banda de contenção. A subseção 4.2.3 trata dos resultados da medição de largura de banda disponível e a subseção 4.3 discute os problemas relacionados às medições realizadas, suas possíveis causas e soluções.

Os resultados estão organizados em tabelas para cada caminho medido. Para as ferramentas que estimam a largura de banda nominal foram colocados os valores para todos os enlaces. O tempo de execução, no formato horas:minutos:segundos, para todos os experimentos, também foi registrado. As situações nas quais não foi possível executar a ferramenta ou quando a ferramenta, após a execução, não forneceu nenhum resultado, são indicadas por um traço (-). Os enlaces indicados nas tabelas são os apresentados na Figura 4.1.

4.2.1 Largura de Banda Nominal

Os experimentos para a métrica largura de banda nominal foram feitos com as ferramentas **clink** e **pchar**. Ambas estimam a capacidade nominal de cada enlace. Foram realizados experimentos em todos os caminhos apresentados na Figura 4.1 a partir da máquina `tweek`, ou seja, nos caminhos `tweek` ↔ `bebe`, `tweek` ↔ `dupond`, `tweek` ↔ `ns`, `tweek` ↔ `labsec01` e `tweek` ↔ `mail`.

Na Tabela 4.3 estão apresentados os resultados das ferramentas **clink** e **pchar** no caminho `tweek` ↔ `bebe`. A segunda coluna apresenta a largura de banda nominal para o enlace especificado na primeira linha, seguido de cinco resultados de medições, dados em Mb/s. A última coluna apresenta o tempo de execução para cada experimento. A largura de banda obtida por ambas as ferramentas é bem inferior à largura de banda nominal, sendo os resultados de **pchar** melhores que os resultados de **clink**. O tempo de execução foi estável para todos os experimentos realizados. As execuções de **pchar** foram mais longas.

Na Tabela 4.4 estão apresentados os resultados das ferramentas **clink** e **pchar** no caminho `tweek` ↔ `dupond`. A segunda e terceira colunas apresentam a largura de banda nominal para o enlace especificado na primeira linha, seguido de cinco resultados de medições em Mb/s. A última coluna apresenta o tempo de execução para cada experimento. A imprecisão nos resultados e a estabilidade dos tempos de execução de ambas as ferramentas se repetem neste caminho. Os dois enlaces neste caminho têm largura de banda de 100 Mb/s. **clink** subestima a medida no primeiro enlace e a superestima no segundo enlace. Os hífens

	tweek - bebe	tempo
banda nominal	100	
clink	16,947	00:12:23
	21,794	00:12:19
	21,770	00:12:19
	20,673	00:12:20
	21,255	00:12:18
pchar	21,876	00:19:33
	28,081	00:19:36
	28,127	00:19:36
	26,701	00:19:34
	26,385	00:19:34

Tabela 4.3: Medições do **clink** e **pchar** no caminho `tweek - bebe`.

mostrados nos resultados de **pchar** indicam que a ferramenta obteve valores negativos para a largura de banda no segundo enlace (como explicado no manual da ferramenta). Os tempos de execução da ferramenta **clink** foram praticamente os mesmos no caminho `tweek` $\leftrightarrow$ `bebe`, apenas as execuções de **pchar** foram mais longas.

Na Tabela 4.5 estão apresentados os resultados das ferramentas **clink** e **pchar** no caminho `tweek` $\leftrightarrow$ `ns`. A segunda, terceira e quarta colunas apresentam a largura de banda nominal para o enlace especificado na primeira linha, seguido de cinco resultados de medições em Mb/s. A última coluna apresenta o tempo de execução para cada experimento. A largura de banda de contenção está em destaque em negrito. Nota-se praticamente os mesmos resultados no primeiro enlace do caminho, `tweek - rl`, mostrado na Tabela 4.4. Para os demais enlaces, de 1000 Mb/s e o seguinte, a ferramenta **clink** apresentou resultados negativos, o que é incorreto, e **pchar** não apresentou resultados. Os resultados indicam que as ferramentas são inadequadas para medir enlaces com maior largura de banda. Este fato é confirmado pela literatura [DOW 99][JAC 97]. Os tempos de execução do **clink** mostram variação, sendo significativamente menores do que os tempos dos experimentos anteriores. Os tempos de execução do **pchar** continuam estáveis.

Na Tabela 4.6 estão apresentados os resultados da ferramenta **clink** no caminho `tweek` $\leftrightarrow$ `labsec01`. A primeira e segunda colunas apresentam, respectivamente, o enlace medido e a banda nominal em Mb/s. A última linha apresenta o tempo de execução para cada experimento. Os resultados para o enlace de contenção (8 Mb/s) estão em negrito. Observa-se, mais uma vez, a imprecisão para larguras de banda maiores. As medidas mais precisas foram obtidas no enlace de contenção. Os tempos de execução neste caminho, mais longo

	tweek - rl	rl - dupond	tempo
banda nominal	100	100	
clink	17,902	148,217	00:12:33
	17,862	150,593	00:12:32
	17,648	191,620	00:12:32
	14,700	100,334	00:12:26
	14,607	99,189	00:12:34
pchar	17,878	-	00:25:45
	17,774	-	00:25:41
	17,814	-	00:25:42
	14,760	-	00:25:41
	14,450	-	00:25:42

Tabela 4.4: Resultados **clink** e **pchar** no caminho `tweek - dupond`.

que os anteriores, mostram-se também variados.

Na Tabela 4.7 estão apresentados os resultados da ferramenta **pchar** no caminho `tweek` $\leftrightarrow$ `labsec01`. A primeira e segunda colunas apresentam, respectivamente, o enlace medido e a banda nominal em Mb/s. A última linha apresenta o tempo de execução para cada experimento. A largura de banda de contenção está em negrito. Novamente observa-se a imprecisão das medições para larguras de banda maiores. Nota-se os melhores resultados nos enlaces com menor largura de banda, que são o enlace de contenção e o último enlace.

Na Tabela 4.8 estão apresentados os resultados da ferramenta **clink** no caminho **tweek** $\leftrightarrow$ `mail`. A primeira e segunda colunas apresentam, respectivamente, o enlace medido e a banda nominal em Mb/s. A última linha apresenta o tempo de execução para cada experimento. A largura de banda de contenção está em negrito. Novamente os melhores resultados foram os que mediram os enlaces menores, de 8Mb/s e 2Mb/s. Quanto aos tempos de execução, observa-se que a ferramenta demorou entre uma hora e meia e quase quatro horas para obter o resultado da medição.

Na Tabela 4.9 estão apresentados os resultados da ferramenta **pchar** no caminho `tweek` $\leftrightarrow$ `mail`. A primeira e segunda colunas apresentam, respectivamente, o enlace medido e a banda nominal em Mb/s. A última linha apresenta o tempo de execução para cada experimento. A largura de banda de contenção está em negrito. Novamente os resultados das larguras de banda menores, ou seja, até 10Mb/s, foram mais precisos. Os tempos de execução variaram um pouco mas foram mais estáveis que os tempos da ferramenta **clink**.

As ferramentas obtiveram resultados parecidos quanto à precisão, provavelmente devido ao fato de que ambas são implementações da mesma técnica. A precisão foi ruim quando

	tweek - r1	r1 - r2	r2 - ns	tempo
banda nominal	100	1000	100	
clink	**17,597**	-214,256	-171,480	00:00:45
	18,024	-315,874	-215,305	00:01:24
	17,589	-2.110,410	-113,451	00:00:52
	14,750	-259,791	-539,786	00:01:19
	14,623	-189,303	-550,242	00:01:01
pchar	**18,013**	-	-	00:18:26
	17,734	-	-	00:18:26
	17,840	-	-	00:18:25
	14,725	-	-	00:18:25
	14,456	-	-	00:18:28

Tabela 4.5: Resultados **clink** e **pchar** no caminho `tweek` - `ns`.

as ferramentas mediram enlaces com 100Mb/s ou mais. Em alguns enlaces o resultado foi negativo, por exemplo, no enlace r1 $\leftrightarrow$ r2 de 1 Gb/s apresentado na Figura 4.1. Uma explicação para isso é o fato de que a velocidade para um pacote trafegar e voltar num certo enlace pode ser menor que o tempo de resposta do roteador anterior a este enlace. Observamos que os RTTs medidos no caminho `tweek` $\leftrightarrow$ r2 foram, na maioria, menores que os RTTs medidos no caminho anterior `tweek` $\leftrightarrow$ r1.

Comparando **pchar** e **clink**, a primeira apresentou resultados mais próximos da capacidade nominal, embora ambos os resultados tenham sido ruins em todos os experimentos.

Com relação ao tempo de execução, ambas as ferramentas demoraram mais para executar do que as outras, da ordem de dezenas de minutos e, em algumas execuções, algumas horas, sendo que **clink** obteve os resultados, em geral, em menos tempo, na ordem de dezenas de minutos nos caminhos `tweek` $\leftrightarrow$ `bebe`, `tweek` $\leftrightarrow$ `dupond` e `tweek` $\leftrightarrow$ `labsec01`, e na ordem de um minuto no caminho `tweek` $\leftrightarrow$ `ns`. Apenas no caminho mais longo, `tweek` $\leftrightarrow$ `mail`, a ferramenta **pchar** obteve os resultados em menos tempo que **clink**, entre duas e três horas de execução.

O número de pacotes transmitidos por **clink** e **pchar** também é maior do que o número de pacotes utilizados pelas demais ferramentas testadas neste trabalho. **clink** e **pchar** enviam, para cada enlace, uma quantidade maior de pacotes para garantir que pelo menos uma pequena quantidade de pacotes transmitidos não sofra atraso devido a fila nos roteadores. Esta quantidade é fixa e da ordem de centenas de pacotes para cada enlace, quando apenas um endereço IP destino responde aos pacotes em um mesmo enlace.

Em relação à execução dos experimentos, a única dificuldade observada nestas duas

enlace	banda	resultado				
	nominal	1	2	3	4	5
tweek - r1	100	17,786	17,705	14,190	14,678	14,755
r1 - r2	1000	-250,821	-265,853	-145,086	-190,277	-216,911
r2 - r3	100	-1.016,086	-382,983	-1.720,581	-719,096	6.538,732
r3 - r4	**8**	**5,739**	**6,623**	**7,243**	**3,775**	**6,860**
r4 - r5	18	-0,324	18,927	15,929	2,032	16,262
r5 - labsec01	10	0,313	3,550	3,714	6,255	3,565
tempo		01:21:57	00:18:08	00:20:42	01:00:38	00:19:18

Tabela 4.6: Resultados **clink** no caminho `tweek - labsec01`.

enlace	banda	resultado				
	nominal	1	2	3	4	5
tweek - r1	100	17,945	17,893	14,508	14,937	14,721
r1 - r2	1000	-	-	-	-	-
r2 - r3	100	-	-	-	1.729,600	-
r3 - r4	**8**	**6,732**	**6,678**	**6,828**	**0**	**2,402**
r4 - r5	18	20,802	57,406	36,378	0	1,554
r5 - labsec01	10	5,798	4,592	5,103	0,393	0
tempo		00:51:06	00:51:19	00:51:38	01:05:04	01:01:00

Tabela 4.7: Resultados **pchar** no caminho `tweek - labsec01`.

ferramentas foi o tempo longo de execução, o que inviabiliza seu uso se a finalidade for usar o resultado em tempo real.

4.2.2 Largura de Banda de Contenção

Para a medição de largura de banda de contenção foram testadas as ferramentas **bprobe**, **nettimer** e **pathrate**.

Na Tabela 4.10 estão apresentados os resultados das ferramentas **nettimer**, **pathrate** e **bprobe** no caminho `tweek` ↔ `bebe`. A segunda coluna apresenta cinco resultados de medições em b/s. **pathrate** apresenta os resultados na forma de um intervalo que deverá conter o valor estimado. A terceira coluna apresenta o tempo de execução para cada experimento. Os resultados da ferramenta **nettimer** indicam uma largura de banda de contenção superior à instalada no caminho. **pathrate** e **bprobe** apresentam resultados satisfatórios. A ferramenta **pathrate** alcançou todos os resultados já na fase inicial da execução. **netti-**

enlace	banda nominal	resultado				
		1	2	3	4	5
tweek - r1	100	14,732	14,737	14,686	14,643	14,713
r1 - r2	1000	-255,003	-217,372	-5.341,729	-216,702	-216,080
r2 - r3	100	-1.102,109	-716,394	-132,463	-309,515	-792,793
r3 - r4	8	-47,098	1,193	5,367	4,942	1,183
r4 - r6	34	-1,376	-0,926	21,655	-33,546	-1,771
r6 - r7	155	0,575	-23,872	-1,288	-32,612	-8,251
r7 - r8	155	-1,068	-11,481	-12,747	10,560	18,802
r8 - r9	155	0,920	2,381	2,014	-16,504	7,338
r9 - r10	155	-0,635	2,921	-18,592	-4,143	-50,999
r10 - r11	100	1,431	4,272	1,211	11,251	-2,288
r11 - r12	**2**	**0,872**	**1,910**	**14,299**	**-0,734**	**1,118**
r12 - mail	10	-1,054	2,911	1,210	0,398	32,509
tempo		03:31:07	03:45:36	03:21:02	02:00:18	01:32:40

Tabela 4.8: Resultados do **clink** no caminho `tweek - mail`.

mer e **bprobe** procederam a medição em um segundo, enquanto **pathrate** demorou cerca de um minuto e quinze segundos.

Na Tabela 4.11 estão apresentados os resultados das ferramentas **pathrate** e **bprobe** no caminho `tweek` ↔ dupond. A segunda coluna apresenta cinco resultados de medições em b/s. A terceira coluna apresenta o tempo de execução para cada experimento. Não foi possível obter qualquer resultado de **nettimer** devido ao bloqueio no *firewall*. Os resultados de **pathrate** estão bem próximos da banda nominal. Além disso, nas cinco execuções, a ferramenta alcançou os resultados na fase inicial. Os resultados de **bprobe** indicam largura de banda de contenção bastante inferior à banda nominal instalada no enlace.

Na Tabela 4.12 estão apresentados apenas os resultados da ferramenta **bprobe** no caminho `tweek` ↔ `ns`. A segunda coluna apresenta cinco resultados de medições em b/s. A terceira coluna apresenta o tempo de execução para cada experimento. Embora a banda de contenção seja 100 Mb/s, os cinco resultados de **bprobe** mostram uma banda de contenção inferior a 10 Mb/s. Não foi possível obter qualquer resultado de **nettimer** devido ao bloqueio no *firewall*. Nas execuções de **pathrate** a ferramenta iniciava a transmissão de pacotes e depois finalizava a execução, indicando um grande número de perdas.

Na Tabela 4.13 estão apresentados os resultados das ferramentas **nettimer**, **pathrate** e **bprobe** no caminho `tweek` ↔ `labsec01`. A segunda coluna apresenta cinco resultados de medições em b/s. A terceira coluna apresenta o tempo de execução para cada experimento.

enlace	banda nominal	resultado				
		1	2	3	4	5
tweek - r1	100	14,605	14,483	14,576	14,199	14,103
r1 - r2	1000	-	-	-	-	-
r2 - r3	100	-	-	995,692	-	-
r3 - r4	8	0	3,776	1,445	22,237	-
r4 - r6	34	1,233	-	0	9,535	0,988
r6 - r7	155	0	1,604	0,793	4,886	4,886
r7 - r8	155	0,602	3,440	-	1,016	1,016
r8 - r9	155	-	0	-	0	0
r9 - r10	155	2,490	1,927	1,329	0	0
r10 - r11	100	4,881	3,095	-	1,435	1,311
r11 - r12	**2**	**3,324**	**1,537**	**1,593**	**1,202**	**0**
r12 - mail	10	-	-	-	4,623	1,135
tempo		02:45:16	02:43:24	02:15:09	02:02:34	02:06:49

Tabela 4.9: Resultados do **pchar** no caminho `tweek - mail`.

Nos resultados da ferramenta **nettimer** há uma concentração no valor 29 Mb/s, sendo que a banda de contenção é de 8 Mb/s. Dos cinco resultados de **pathrate**, os dois primeiros foram alcançados na fase inicial e os outros três nas fases seguintes de execução. Nota-se que os dois últimos resultados tendem a uma banda de contenção de 10 Mb/s. Este resultado provavelmente é influência do último enlace do caminho que é de 10 Mb/s. Como foi visto anteriormente, os pacotes transmitidos na técnica *packet-pair* podem ter se enfileirado no enlace de 10 Mb/s, posterior ao enlace de contenção. Os resultados de **bprobe** apresentam certa variação, com a maioria dos resultados um pouco acima do esperado.

Na Tabela 4.14 estão apresentados os resultados das ferramentas **nettimer**, **pathrate** e **bprobe** no caminho `tweek` $\leftrightarrow$ `mail`. A banda de contenção neste caminho é de 2 Mb/s. A segunda coluna apresenta cinco resultados de medições em b/s. A terceira coluna apresenta o tempo de execução para cada experimento. Nota-se novamente a instabilidade dos resultados da ferramenta **nettimer**. **pathrate** obteve os resultados na fase inicial na primeira, terceira e quinta execuções e novamente obteve a melhor estabilidade e precisão dos resultados. No entanto, **pathrate** chegou a demorar mais de uma hora e meia para realizar a medição. **bprobe** também obteve resultados bons embora com certa variação.

De forma geral, os resultados da ferramenta **nettimer** apresentaram pouca confiabilidade. A banda de contenção variou entre 92 Mb/s e 173 Mb/s em um enlace de 100 Mb/s (caminho `tweek` $\leftrightarrow$ `bebe`) e entre 0,75 Mb/s e 30 Mb/s para um enlace de contenção de

ferramenta	banda de contenção	tempo
nettimer	92.419.208	00:00:01
	133.333.333	00:00:01
	129.032.258	00:00:01
	166.666.666	00:00:01
	173.913.043	00:00:01
pathrate	93M a 98M	00:01:14
	94M a 99M	00:01:13
	94M a 99M	00:01:13
	95M a 100M	00:01:14
	95M a 100M	00:01:13
bprobe	93.303.200	00:00:01
	95.370.300	00:00:01
	95.250.400	00:00:01
	93.324.900	00:00:01
	94.293.900	00:00:01

Tabela 4.10: Medidas de banda de contenção no caminho `tweek - bebe`.

2Mb/s (caminho `tweek` ↔ `mail`).

As ferramentas **pathrate** e **bprobe** obtiveram os resultados mais próximos das larguras de banda de contenção, sendo que **pathrate** apresentou as medições mais precisas. Apesar da melhor precisão destas duas ferramentas, os resultados de **bprobe** no caminho `tweek` ↔ `dupond` e no caminho `tweek` ↔ `ns` ficaram abaixo de 10% do valor nominal da banda de contenção, que é de 100 Mb/s. A largura de banda de contenção obtida no primeiro caminho ficou entre 10 Mb/s e 22,5 Mb/s e entre 3 Mb/s e 6,6 Mb/s no caminho `tweek` ↔ `ns`.

O tempo de execução das ferramentas **bprobe** e **nettimer** foi da ordem de alguns segundos. O **pathrate** demorou entre 1 e 2 minutos quando a ferramenta alcançava um bom resultado na fase inicial, geralmente em horários de menor tráfego, e demorou algumas dezenas de minutos quando a ferramenta executava as duas fases, chegando a demorar até pouco mais de uma hora e meia.

O número de pacotes enviados por **nettimer** variou entre oito e dezesseis pacotes, segundo a saída da ferramenta no momento da execução. O número de pacotes enviados por **bprobe** não é informado pela ferramenta quando esta é executada mas, segundo os autores [CAR 96b], a ferramenta envia dez pacotes de 124 bytes cada um, e continua enviando pacotes cada vez maiores até que o tamanho chegue ao tamanho máximo de pacote permitido

ferramenta	banda de contenção	tempo
pathrate	95M a 100M	00:01:16
	98M a 103M	00:01:13
	97M a 102M	00:01:13
	97M a 102M	00:01:17
	95M a 101M	00:01:17
bprobe	22.518.500	00:00:01
	10.015.600	00:00:01
	15.599.900	00:00:01
	22.265.300	00:00:01
	22.044.100	00:00:01

Tabela 4.11: Medidas de banda de contenção no caminho `tweek` - `dupond`.

ferramenta	banda de contenção	tempo
bprobe	6.689.200	00:00:01
	4.845.070	00:00:01
	4.436.370	00:00:01
	4.568.580	00:00:01
	6.535.070	00:00:01

Tabela 4.12: Medidas de banda de contenção no caminho `tweek` - `ns`.

pela conexão. O número de pacotes enviados por **pathrate** depende se a ferramenta consegue um resultado na fase inicial ou precisa executar as outras duas fases (*packet pair probing* e *packet train probing*, transmitindo pacotes durante algumas dezenas de minutos). Como exemplo do número de pacotes enviados por **pathrate**, podemos citar os experimentos no caminho `tweek` ↔ `bebe` no qual foram enviados 1080 pacotes e a ferramenta executou apenas a fase inicial, e o segundo experimento apresentado na Tabela 4.14, no caminho `tweek` ↔ `mail`, no qual foram enviados 3918 pacotes e a ferramenta executou as outras duas fases. Diferentemente de **clink** e **pchar**, estas três ferramentas transmitem quantidades variadas de pacotes em cada execução.

A ferramenta **nettimer** apresentou algumas dificuldades de execução, começando com a compilação. Não foi possível compilar **nettimer** em nenhum dos microcomputadores. Foi preciso usar uma versão pré-compilada. Também não foi possível compilar **bprobe** e **cprobe** porque são ferramentas originalmente desenvolvidas para serem executadas em computadores Silicon Graphics. Foi usada uma versão para PC desenvolvida pelo Prof.

ferramenta	banda de contenção	tempo
nettimer	29.851.411	00:00:01
	29.061.195	00:00:01
	29.465.909	00:00:01
	29.662.146	00:00:01
	10.653.399	00:00:01
pathrate	6,7M a 7,0M	00:01:22
	6,8M a 7,2M	00:01:21
	6,7M a 7,0M	00:26:29
	9,9M a 10,3M	00:50:57
	9,8M a 10,1M	00:29:54
bprobe	7.965.690	00:00:01
	9.834.780	00:00:01
	10.475.700	00:00:01
	8.413.960	00:00:01
	8.264.240	00:00:01

Tabela 4.13: Medidas de banda de contenção no caminho `tweek - labsec01`.

Chow [CHO 01] da University of Colorado.

A execução de **nettimer** e **pathrate** no caminho `tweek` ↔ `dupond` e `tweek` ↔ `ns` foi dificultada devido ao *firewall* da rede da UFPR, que bloqueava as portas necessárias às ferramentas. Como **pathrate** usa apenas duas portas, estas foram liberadas no *firewall*. As portas necessárias à ferramenta **nettimer** não puderam ser liberadas, por questões de segurança, porque a ferramenta usa um grande número de portas.

A ferramenta **pathrate** deve ser executada nos dois fins do caminho, enquanto as demais são executadas a partir de um único *host*. Por outro lado, **pathrate** não necessita de execução como usuário administrador e obtém resultados mais precisos.

4.2.3 Largura de Banda Disponível

A única ferramenta testada que mede a largura de banda disponível é **cprobe**. A validação do **cprobe** só foi possível no caminho `tweek` ↔ `mail`, onde está disponível a ferramenta MRTG, com as estatísticas de tráfego do enlace. Para validar o resultado nos outros caminhos seria necessário ter acesso aos gráficos de tráfego gerados pelo MRTG nos enlaces.

Na Tabela 4.15 estão apresentados os resultados das execuções da ferramenta nos caminhos `tweek` ↔ `bebe`, `tweek` ↔ `dupond`, `tweek` ↔ `ns` e `tweek` ↔ `labsec01`. A primeira coluna indica o *host* destino e nas outras colunas estão apresentados os resultados de cinco

ferramenta	banda de contenção	tempo
nettimer	30.651.790	00:00:01
	1.955.352	00:00:01
	756.856	00:00:01
	10.738.738	00:00:01
	3.655.278	00:00:01
pathrate	1,9M a 2,0M	00:01:21
	1,9M a 2,0M	01:38:44
	1,9M a 2,0M	00:01:53
	1,9M a 2,0M	00:52:52
	1,9M a 2,0M	00:01:22
bprobe	1.843.380	00:01:01
	1.743.560	00:01:01
	1.794.530	00:01:01
	1.846.380	00:00:31
	1.559.260	00:00:01

Tabela 4.14: Medidas de banda de contenção no caminho `tweek` - `mail`.

execuções da ferramenta **cprobe**. Nas execuções com destino ao *host* `labsec01` a ferramenta mostrou mensagem de tempo excedido e não obteve qualquer resultado válido. Os resultados até os *hosts* `bebe`, `dupond` e `ns` estão dentro do intervalo possível, considerando que o valor máximo para a largura de banda disponível é a largura de banda de contenção. Apenas o primeiro resultado até o *host* `dupond` ficou fora do intervalo possível.

Como o valor do tráfego mostrado no gráfico MRTG é a média dos últimos cinco minutos, no caminho `tweek` ↔ `mail` a ferramenta **cprobe** foi executada cinco vezes em cada experimento, sendo uma vez a cada minuto. Desta forma foi comparada a média dos cinco resultados com o gráfico MRTG, evitando considerar apenas um valor isolado. Na Tabela 4.16 pode ser observado o resultado de cada uma das cinco execuções de cada experimento até o *host* `mail`, a média das cinco execuções de cada experimento e o valor da banda disponível segundo o gráfico MRTG durante os cinco minutos em que a ferramenta foi executada. A banda disponível, extraída do gráfico MRTG, é a banda utilizada informada no gráfico subtraída da banda nominal. Quatro dos vinte e cinco resultados estão acima da banda nominal, ou sèja, fora do intervalo possível.

O tempo de execução do **cprobe** foi de um segundo para todas as execuções da ferramenta, mesmo nos caminhos mais longos e com atrasos maiores como nos caminhos `tweek` ↔ `labsec01` e `tweek` ↔ `mail`.

destino	resultado				
	1	2	3	4	5
bebe	90.254.828	89.161.896	88.671.352	88.436.728	88.723.936
dupond	149.799.328	18.110.962	18.021.096	17.996.252	17.959.752
ns	4.419.215	4.122.076	4.610.634	4.629.258	4.716.148
labsec01	-	-	-	-	-

Tabela 4.15: Medidas de banda disponível até os *hosts* `bebe`, `dupond`, `ns` e `labsec01`.

	experimento				
	1	2	3	4	5
resultados cprobe	1.101.731	960.949	5.403.882	904.024	1.601.478
	1.128.657	1.351.234	711.087	5.511.281	1.447.804
	1.294.164	711.478	1.055.999	1.554.116	1.468.380
	1.733.706	2.446.698	1.537.784	1.350.478	1.599.806
	1.166.309	1.093.902	3.628.785	968.927	1.646.086
média	1.284.913	1.312.852	2.467.507	2.064.965	1.552.710
banda disp. MRTG	1.933.700	1.767.700	1.964.100	1.942.300	1.977.100

Tabela 4.16: Medidas de banda disponível no caminho `tweek - mail`.

4.3 Limitações das Ferramentas

Nesta seção estão apresentadas algumas limitações impostas às ferramentas devido a fatores como o comportamento da transmissão de pacotes nas redes e roteadores, o atraso devido ao processamento fim-a-fim, a metodologia usada pela ferramenta, entre outros.

Um dos fatores que podem limitar a precisão das medições nas ferramentas estudadas é a necessidade de recebimento de resposta para os pacotes transmitidos. Algumas ferramentas utilizam respostas ICMP para medir o atraso de ida e volta de pacotes. Esta metodologia pode gerar resultados imprecisos pois os roteadores presentes no caminho de rede provavelmente têm diferentes tempos de processamento e de resposta para estes pacotes. Além disso, alguns roteadores podem ser programados para dar prioridade baixa aos pacotes de resposta ICMP [SAV 99], o que aumenta o tempo de resposta e diminui a estimativa da largura de banda.

O tráfego de dados de outros fluxos no caminho de rede pode prejudicar os resultados da medição, de pelo menos duas formas, quando a técnica utilizada é a *packet-pair*. A primeira é quando tráfego extra é processado entre os dois pacotes da medição, o que faz

com que a dispersão dos pacotes seja maior, diminuindo a largura de banda resultante. Outra forma é quando o primeiro pacote do par encontra fila num roteador após o enlace de contenção. Isso pode fazer com que o segundo pacote do par alcance o primeiro. Neste caso a largura de banda de contenção resultante será uma largura de banda medida após o enlace de contenção.

Nas medições fim-a-fim, o processamento de outros aplicativos nos microcomputadores de origem e destino aumenta a carga nos processadores e pode atrasar o processamento dos pacotes e o registro da hora da chegada dos pacotes e, consequentemente, alterar a largura de banda resultante.

Em redes assimétricas, a dispersão dos pacotes pode ser alterada quando a ferramenta mede o tempo de ida e volta dos pacotes, e a banda de contenção no caminho de ida é diferente da banda de contenção no caminho de volta, principalmente se na volta a banda de contenção é menor.

Em relação à resolução do relógio, quanto maior a resolução do relógio do microcomputador receptor de pacotes, ou seja, quanto menor o tempo entre as atualizações deste relógio, melhor será a precisão do resultado. Por exemplo, para um relógio com uma resolução de dez milisegundos, se um pacote chega três milisegundos após a última atualização do relógio, a ferramenta contará sete ms a mais no tempo de chegada do pacote, tornando o resultado menos preciso (diminuindo a largura de banda resultante), pois a ferramenta não consegue obter uma diferença de tempo menor que dez milisegundos.

A escolha do tamanho dos pacotes enviados usando pares ou trilhas de pacotes é importante. Pacotes grandes resultam em dispersão maior, o que é melhor para medir devido ao problema da resolução do relógio do computador, e é mais robusto quanto a tráfego extra entre os pacotes da medição. Por outro lado, pacotes pequenos diminuem a probabilidade de ocorrer tráfego extra entre os pacotes da medição pois o tempo entre os dois pacotes é menor.

Para o uso de pacotes grandes, é necessário garantir que o pacote não seja maior que o tamanho máximo permitido (MTU) nos enlaces. Se o pacote da medição for maior, este será fragmentado. Uma conseqüência deste fato é que um roteador pode responder ao pacote assim que o primeiro fragmento chega ao roteador, ou seja, antes que o pacote inteiro da medição chegue. Isso resulta em um menor atraso medido, superestimando a largura de banda.

Algumas omissões na medição do atraso afetam o cálculo da latência. Considerar que os tempos de processamento na origem e destino e o tempo de transmissão de pacotes de erro ICMP são insignificantes causa subestimação da largura de banda pois a ferramenta considera que estes atrasos são resultantes do tempo de transmissão.

Para algumas metodologias, como a usada nas ferramentas que utilizam a técnica *one-*

packet, há diferença entre transmitir pacotes de tamanhos mais variados ou mais pacotes de mesmo tamanho. A vantagem de utilizar mais pacotes do mesmo tamanho é que há uma maior possibilidade da ferramenta obter amostras com RTT mínimo, ou seja, sem fila no caminho. Por outro lado, transmitindo pacotes de tamanhos mais variados, obtém-se mais valores para realizar a regressão linear para descobrir a latência. Estudos mostram que, neste caso, a escolha em transmitir pacotes de tamanhos mais variados é melhor [DOW 99].

Quanto mais rápido for o microcomputador que está medindo a dispersão dos pacotes, menor será o valor da dispersão que este pode medir e, conseqüentemente, maior será a largura de banda que pode ser medida. A banda máxima que pode ser medida é a razão entre o tamanho do pacote e a dispersão mínima. A dispersão mínima é determinada pelo tempo para receber um pacote no sistema operacional, mover o pacote para o espaço do usuário, registrar a hora exata da chegada do pacote e executar qualquer outra operação da ferramenta antes de esperar pelo próximo pacote. Por exemplo, em um microcomputador Pentium-II com sistema operacional Free-BSD 3.2, a dispersão mínima é da ordem trinta a quarenta μs [DOV 01]. Denotando a dispersão mínima no receptor como Δ_m, a largura de banda máxima que pode ser medida para um tamanho de pacote L é de $\frac{L}{\Delta_m}$, então para $\Delta_m = 40$ μs e $L = 1500$ *bytes*, a banda máxima é de 300 Mb/s.

Alguns enlaces são formados por vários enlaces menores. A metodologia usada pela ferramenta para medir a largura de banda deve detectar os enlaces multi-canais. Por exemplo, se um enlace multi-canal não for detectado e os pacotes da medição forem transmitidos através de um mesmo canal, o resultado da ferramenta poderá refletir apenas a banda deste canal.

Alguns equipamentos como *switches* não têm um endereço IP e não decrementam o TTL do pacote IP. Porém, realizam processamento nos pacotes e consequentemente aumentam o tempo de transmissão dos mesmos. Este atraso implicará em uma largura de banda resultante menor que a real.

Medição de atraso em enlaces com maior velocidade de transmissão são mais sensíveis a outros tráfegos. Por exemplo, se um pacote de outro tráfego provoca um aumento de três milisegundos em uma medição, para um enlace lento, com um atraso de, por exemplo, trinta milisegundos, o erro provocado pelo ruído é de 10%. Para um enlace mais rápido com um atraso de, por exemplo, doze milisegundos, o erro provocado pelo ruído é de 25%.

Capítulo 5

Conclusões

Medições fim-a-fim de largura de banda são essenciais em redes como a Internet, pois podem ser a única forma de conhecer a capacidade máxima e a capacidade instantânea de um caminho. Este trabalho apresenta um estudo sobre algumas ferramentas utilizadas na avaliação de desempenho de redes IP, especificamente para as métricas atraso, variação de atraso, perda de pacotes e largura de banda.

As ferramentas `clink`, `pathrate`, `bprobe`, `cprobe`, `pchar` e `nettimer` para medição ativa fim-a-fim de largura de banda foram testadas em trechos diversos de algumas redes. Os testes demonstraram as características de precisão, repetibilidade e robustez das ferramentas. A avaliação simultânea e no mesmo ambiente de teste destas ferramentas é uma contribuição original deste trabalho.

Os resultados dos experimentos realizados neste trabalho indicam que a ferramenta que produziu os melhores resultados foi **pathrate**, que obteve bons resultados em três dos cinco caminhos medidos. Um resultado da estimativa de um caminho foi considerado bom quando a ferramenta produziu, em pelo menos três das cinco medições realizadas, resultados com erro igual ou inferior a 10% da medida em relação à capacidade instalada conhecida. Os resultados de execuções que não conseguiram estes índices de precisão foram considerados insatisfatórios.

Seguindo este critério, a segunda melhor ferramenta foi **bprobe** com dois resultados bons. Embora esta ferramenta tenha alcançado resultados nos cinco caminhos, três foram considerados insatisfatórios.

Com relação às ferramentas **clink**, **pchar** e **nettimer**, todos os resultados foram considerados insatisfatórios. Quanto à ferramenta **cprobe**, os únicos resultados que puderam ser validados foram considerados insatisfatórios. Para esta última ferramenta foi comparada a média das cinco execuções com o valor esperado.

A ferramenta **pathrate**, embora tenha sido considerada a mais precisa, apresentou alguns tempos de medição muito longos, o que inviabiliza a utilização da ferramenta em

situações que exigem resposta em tempo real, por exemplo, para escolha dinâmica de servidor.

Nenhuma das ferramentas testadas apresentou simultaneamente confiabilidade e rapidez nas medições. Algumas ferramentas não são robustas com relação à mudança grande de largura de banda em enlaces adjacentes. Por este motivo, alguns valores negativos de largura de banda foram obtidos. A partir destas observações podemos concluir que as ferramentas ainda precisam ser melhoradas, ou novas ferramentas devem ser propostas.

Este trabalho contribui também para o domínio da tecnologia de medição de desempenho em redes. Uma continuação importante deste trabalho é a proposta de uma nova ferramenta que apresente bons resultados quanto à precisão, confiabilidade, repetibilidade e tempo de execução. A ampliação dos experimentos para contemplar outras ferramentas de medição de largura de banda e também outras métricas de desempenho é uma extensão natural deste trabalho. Outra proposta a ser avaliada é a utilização de mecanismos de sincronização para a melhoria de precisão na medição de atraso como, por exemplo, sistemas GPS e CDMA [GEO 01]. Acreditamos que a medição com utilização destes mecanismos elimina o problema do sincronismo dos relógios, revelando os problemas específicos das técnicas e ferramentas de medição.

Bibliografia

[AHN 99] AHN, S. J. et al. Design and Implementation of a Web-based Internet Performance Management System using SNMP MIB-II. **International Journal of Network Management**, [S.l.], v.9, p.309–321, 1999.

[ALM 99a] ALMES, G.; KALIDINDI, S.; ZEKAUSKAS, M. A One-way Delay Metric for IPPM. RFC2679, setembro, 1999.

[ALM 99b] ALMES, G.; KALIDINDI, S.; ZEKAUSKAS, M. A One-way Packet Loss Metric for IPPM. RFC2680, setembro, 1999.

[ALM 99c] ALMES, G.; KALIDINDI, S.; ZEKAUSKAS, M. A Round-trip Delay Metric for IPPM. RFC2681, setembro, 1999.

[BAL 00] BALBINOT, L. F.; ANDRADE, M. de. Uma Ferramenta Flexível para a Medição de Tráfego Baseada no RRDtool. **II Workshop RNP2**, Belo Horizonte, MG, v.1, 2000.

[BAR 99] BARFORD, P.; CROVELLA, M. Measuring Web Performance in the Wide Area. **ACM SIGCOMM**, Cambridge, Massachusetts, v.27, p.37–48, 1999.

[BOL 93] BOLOT, J. End-to-End Packet Delay and Loss Behavior in the Internet. **Proceedings of ACM SIGCOMM**, San Francisco, CA, v.1, p.289–298, agosto, 1993.

[CAI 01] CAIDA. CAIDA - Cooperative Association for Internet Data Analysis. Disponível em <http://www.caida.org>, setembro, 2001.

[CAR 96a] CARTER, R. L.; CROVELLA, M. E. Dynamic Server Selection using Bandwidth Probing in Wide-Area Networks. Boston University, 1996. Relatório TécnicoBU-CS-96-007.

[CAR 96b] CARTER, R. L.; CROVELLA, M. E. Measuring Bottleneck Link Speed in Packed-Switched Networks. **Performance Evaluation**, Lausanne, Suíça, v.27 e 28, p.297–318, outubro, 1996.

[CHO 01] CHOW, C. E. Home page do Prof. C. Edward Chow. Disponível em <http://cs.uccs.edu/~chow>, junho, 2001.

[COM 00] COMER, D. E. **Interligação em Rede com TCP/IP**, v.1. Editora Campus, 2000.

[DEM 02] DEMICHELIS, C. IP Packet Delay Variation Metric for IPPM. Trabalho em Progresso - Internet Draft, agosto, 2002.

[DOV 01] DOVROLIS, C.; RAMANATHAN, P.; MOORE, D. What do Packet Dispersion Techniques Measure? **Proceedings of IEEE INFOCOM**, Anchorage, Alaska, v.1, p.905–914, abril, 2001.

[DOV 02] DOVROLIS, C.; JAIN, M. Pathload, a Measurement Tool for End-to-end Available Bandwidth. **Passive and Active Measurements Workshop**, Fort Collins, Colorado, v.1, p.14–25, março, 2002.

[DOW 99] DOWNEY, A. B. Using Pathchar to Estimate Internet Link Characteristics. **Proceedings of ACM SIGCOMM**, Cambridge, Massachusetts, v.1, 1999.

[DOW 02] DOWNEY, A. B. clink: a Tool for Estimating Internet Link Characteristics. Disponível em <http://rocky.wellesley.edu/downey/clink>, agosto, 2002.

[GEO 01] GEORGATOS, F. et al. Providing Active Measurements as a Regular Service for ISP's. **PAM2001**, Amsterdan, Holanda, v.1, 2001.

[GRA 01] GRANVILLE, L. Z. **Gerenciamento Integrado de QoS em Redes de Computadores.** UFRGS - Instituto de Informática, setembro, 2001. Tese de Doutorado.

[HEN 95] HENNESSY, J. L.; PATTERSON, D. **Computer Architetures: A Quantitative Approach**. Morgan Kaufmann Publishers, 1995. p.562-573.

[IMR 02] Internet Measurement Research Group. Disponível em <http://imrg.grc.nasa.gov/imrg/>, 2002. Chair Vern Paxson.

[IPM 01] IPMA. IPMA - Internet Performance Measurement and Analysis. Disponível em <http://nic.merit.edu/ipma>, setembro, 2001.

[IPP 01] IPPM. IPPM - Internet Protocol Performance Metrics. Disponível em <http://www.advanced.org/IPPM>, setembro, 2001.

[JAC 88] JACOBSON, V. Congestion Avoidance and Control. **Proceedings of ACM SIGCOMM**, Stanford, CA, v.1, p.314–329, agosto, 1988.

[JAC 97] JACOBSON, V. Pathchar - a Tool to Infer Characteristics of Internet Paths. Mathematical Sciences Research Institute, 1997. Relatório técnico.

[JAI 91] JAIN, R. **The Art of Computer System Performance Analysis.** John Wiley e Sons, 1991.

[KES 91a] KESHAV, S. A Control-Theoretic Approach to Flow Control. **Proceedings of ACM SIGCOMM**, Zürich, Suíça, v.1, p.3–15, setembro, 1991.

[KES 91b] KESHAV, S. Packet-Pair Flow Control. Disponível em <http://www.cs.cornell.edu/skeshav/doc/94/2-17.ps>, 1991.

[KES 98] KESHAV, S.; SHARMA, R. Achieving Quality of Service through Network Performance Management. **Proceedings of NOSSDAV**, Cambridge, Reino Unido, v.1, p.1, julho, 1998.

[KOO 02] KOODLI, R.; RAVIKANTH, R. One-way Loss Pattern Sample Metrics. RFC3357, agosto, 2002.

[LAI 99] LAI, K.; BAKER, M. Measuring Bandwidth. **Proceedings of IEEE INFOCOM**, New York, NY, v.1, p.235–245, março, 1999.

[LAI 00] LAI, K.; BAKER, M. Measuring Link Bandwidths using a Deterministic Model of Packet Delay. **Proceedings of ACM SIGCOMM**, Stockholm, Suécia, v.1, p.283–294, 2000.

[LAI 01] LAI, K.; BAKER, M. Nettimer: A Tool for Measuring Bottleneck Link Bandwidth. **Proceeding of the 3rd USENIX Symposium on Internet Technologies and Systems**, Boston, Massachusetts, v.1, p.123–134, março, 2001.

[MAH 99a] MAH, B. A. Pchar: Child of Pathchar. **DOE NGI Testbed Workshop**, Berkeley, CA, v.1, julho, 1999.

[MAH 99b] MAHDAVI, J.; PAXSON, V. IPPM Metrics for Measuring Connectivity. RFC2678, setembro, 1999.

[MAH 01] MAH, B. A. Disponível em <http://www.employees.org/~bmah/Software/pchar>, setembro, 2001.

[MAT 01] MATHIS, M.; ALLMAN, M. A Framework for Defining Empirical Bulk Transfer Capacity Metrics. RFC3148, julho, 2001.

[MRT 01] MRTG - Multi Router Traffic Grapher. Disponível em <http://people.ee.ethz.ch/~oetiker>, maio, 2001.

[MTR 01] MTR Site Oficial. Disponível em <http://www.bitwizard.nl/mtr>, maio, 2001.

[MUR 00] MURHAMMER, M. et al. **TCP/IP Tutorial e Técnico**. Makron Books, 2000. p.487-513.

[NIM 01] NIMI. NIMI - National Internet Measurement Infrastructure. Disponível em <http://www.ncne.nlanr.net/nimi>, setembro, 2001.

[NLA 01] NLANR. NLANR - National Laboratory for Applied Network Research. Disponível em <http://dast.nlanr.net/Guides/GettingStarted>, maio, 2001.

[PAM 01] Passive and Active Measurement Workshop. Disponível em <http://www.ripe.net/pam2001>, 2001.

[PAX 97] PAXSON, V. Automated Packet Trace Analysis of TCP Implementations. **Proceedings of ACM SIGCOMM**, Cannes, French Riviera, França, v.1, setembro, 1997.

[PAX 98a] PAXSON, V. et al. An Architecture for Large-Scale Internet Measurement. **IEEE Communications Magazine**, [S.l.], v.36, p.48–54, agosto, 1998.

[PAX 98b] PAXSON, V. et al. Framework for IP Performance Metrics. RFC2330, maio, 1998.

[PAX 99] PAXSON, V. End-to-End Internet Packet Dynamics. **IEEE/ACM Transactions on Networking**, [S.l.], v.7, p.277–292, 1999.

[PET 99] PETERSON, L. L.; DAVIE, B. S. **Computer Networks, a Systems Approach**. Morgan Kaufmann, 1999. p.10,11,18-29,426-430.

[RAB 92] RABINOVICH, S. **Measurement Errors - Theory and Practice**. American Institute of Physics, 1992.

[RNP 02] Rede Nacional de Pesquisa. Disponível em <http://www.rnp.br>, 2002.

[SAV 99] SAVAGE, S. Sting: a TCP-based Network Measurement Tool. **Proceedings of the 1999 USENIX Symposium on Internet Technologies and Systems**, Boulder, Colorado, v.1, p.71–79, 1999.

[SUR 01] SURVEYOR. SURVEYOR Project. Disponível em <http://www.advanced.org/surveyor>, setembro, 2001.

[TCP 01] TCPTRACE Site Oficial. Disponível em <http://www.tcptrace.org>, maio, 2001.

[XPL 01] xplot Tool. ftp://mercury.lcs.mit.edu/pub/shep, maio, 2001.