ABSTRACT

PROACTIVE METHODS FOR MEASUREMENT OF AVAILABLE BANDWIDTH AND LINK CAPACITY

by
Khondaker Musfakus Salehin

With the continuous expansion of network infrastructure and deployment of applications sensitive to Quality-of-Service, network measurement plays a major role in both network planning and management. Accurate measurement of various network parameters, e.g., available bandwidth, link capacity, delay, packet loss and jitter, provides a positive impact for effective traffic engineering, Quality-of-Service (QoS) routing, optimization of end-to-end transport performance, and link capacity planning. For network measurement, there exists several proactive estimation tools based on either probe-gap model or probe-rate model that estimate path related attributes. Most of these tools that have been implemented can measure tight-link capacity (smallest available bandwidth) and/or narrow-link capacity (smallest link capacity) between a source node and destination node along a particular path. However, network measurement also has negative impacts on the cross traffic that induces extra queuing delay and packet loss for the legitimate data traffic that results in a significant intrusion and in degree of erroneous estimation. In this thesis, a combined measurement tool for measuring both available bandwidth and link capacity using a combination of probing packets and ICMP probe packets is been proposed. A study of the proposed tool is presented. The proposed schemes provide acceptable accuracy, low overhead, and avoid over-estimation.

PROACTIVE METHODS FOR MEASUREMENT OF AVAILABLE BANDWIDTH AND LINK CAPACITY

by
Khondaker Musfakus Salehin

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Telecommunications

Department of Electrical and Computer Engineering

January 2008

Blank Page

APPROVAL PAGE

PROACTIVE METHODS FOR MEASUREMENT OF AVAILABLE BANDWIDTH AND LINK CAPACITY

Khondaker Musfakus Salehin

Dr. Roberto Rojas-Cessa, Thesis Advisor — Date
Associate Professor, Department of Electrical and Computer Engineering, New Jersey Institute of Technology

Dr. Nirwan Ansari, Committee Member — Date
Professor, Department of Electrical and Computer Engineering, New Jersey Institute of Technology

Dr. Sotirios G. Ziavras, Committee Member — Date
Professor, Department of Electrical and Computer Engineering, New Jersey Institute of Technology

BIOGRAPHICAL SKETCH

Author: Khondaker Musfakus Salehin

Degree: Master of Science

Date: January 2008

Undergraduate and Graduate Education:

- Master of Science in Telecommunications,
  New Jersey Institute of Technology, Newark, New Jersey, 2008

- Bachelor of Science in Computer Science and Engineering,
  Ahsanullah University of Science and Technology, Dhaka, Bangladesh, 2004

Major: Telecommunications

To my beloved parents

ACKNOWLEDGMENT

I would like to express my deepest appreciation to Dr. Roberto Rojas-Cessa, who not only served as my thesis supervisor, providing valuable and countless resources, insight, and intuition, but also constantly gave me support, encouragement, and had believe in me. He is more a friend than an advisor to me. Special thanks are given to Dr. Nirwan Ansari and Dr. Sotirios G. Ziavras for actively participating in my committee.

All of my fellow graduate students in Dr. Roberto Rojas-Cessa's research group, and my friends Iyamideh Coker and Danish Ahmed Syed are deserving recognition for their support. I also thank Haider Ali Bhuiyan of University of Georgia and Dr. Muhammand Haq of McNeese State University for their constant guidance since after I am here in USA for academic pursuit.

Finally, I would like to express my gratitude to Allah, my parents, and my sister who have supported me in every walk of my life for my success and happiness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Data network is expanding very fast each year as its importance in today's world is increasing day by day. New applications are being deployed on the existing network infrastructure without making any change on the remote physical connections over which local service providers have no control. With the emergence of new technology and their stringent necessities, new applications tend to be faster, more bandwidth hungry and more sensitive to delay and loss (audio and video streaming). These demands must be attained to provide optimal service to end users. Here, optimal service is defined in terms of throughput, reliability of performance, fault tolerance and satisfaction of service agreements along the source-destination path. Such criteria now govern the success of current network technology and network related business.

With the emergence of new applications, new technologies have emerged to ensure optimal network performance. Traffic Engineering statistically ensures that the end users get the most out of deployed service by modeling, designing, configuring, and managing control mechanisms of the service providers' network resources. Overlay networks configure their routing tables based on their overlay link bandwidth to maximize the user satisfaction. For some applications, Quality-of-Service (QoS) can also provide defined priorities to different end users in order to ensure desired performance in accordance with Service Level Agreement (SLA). Thus, the success of all these technologies is strictly dependant on the reliable information of the network parameters mentioned above. According to [8], unused link capacity (i.e., available bandwidth) is the one of the key factors for applications to dynamically form network connections, for service providers to lease links to new customers for efficient end-to-end admission control and intelligent routing in the intermediate nodes, and also for data carriers to upgrade carrier link capacity.

1.1 Classification of Network Measurement Methodologies

Various network measurement tools have been proposed so far to measure link capacity, available bandwidth, and delay separately, except for Pathchar [9], which is capable of measuring link capacity, available bandwidth and delay together by statistically analyzing the collected information during a probing period. All these estimation techniques can be categorized in many different ways according to [1],[12],[18]. Per-hop capacity vs. end-to-end capacity measurement methods and packet pair vs. packet train methods are the two categories that distinguish measurement tools into two different branches. But, it is the active and passive measurement methods that can narrow down all of the proposed methodologies into two distinct measurement classes regardless of their probing packet characteristic and the number of link they can measure along a particular source-destination path.

a) *Active measurement:* Active probing methods proactively shoot packets towards a destination node from a source node to measure various network characteristics. Nodes that inject packets into the network require to determine some probe packet related parameters e.g., probe length, the number of probes to inject, the number of iterations to do the same measurement. This measurement traffic is characterized by link capacity, unused link capacity, cross traffic (data traffic) density, queue length of the intermediate nodes, number of intermediate nodes, and the number of cross traffic generators on the specific path. Destination nodes respond with similar type of measurement probes to the source node that eventually infers network conditions which is controlled by either probe sending and receiving time or probe timestamps.

Active probing methods have further been categorized into several other classes. Depending on how the change of probes in this methodologies are decided, active probing can further be divided into two major methodologies: Delay dased and dispersion based methodologies. Delay based methodologies are usually used to measure per-hop link capacity, queuing delay, and they are implemented through Round Trip Time (RTT) or One Way Delay (OWD) responses from the remote destination nodes. Pathchar [9] and cing [11] are the two tools that use RTT response methodology for estimating link capacity, available bandwidth, delay and queuing delay between a source and destination nodes respectively.

Dispersion based methodologies are used to measure per-path based measurement that estimates narrow-link (link with the smallest link capacity) and tight-link (link with the smallest unused link capacity) capacities along a specific path. The source node, after receiving the response from its peer entity (destination node), compares the rate of change of the probing packets departure and arrival to infer the cross traffic and/or link capacity impact on the measurement probes. Except for link capacity and available bandwidth estimation, dispersion based methodologies can also be used per-path delay and queuing delay and the later is closely related with available bandwidth estimation. For example, both Initial Gap Increasing (IGI) [2] and Spruce [6] use dispersion based probing methodology while Pathload [4] uses delay based probing methodology for tight-link capacity estimation. Pathrate [5] and Bprobe [8] are two narrow-link estimation tools that use dispersion based probing methodology.

b) *Passive measurement:* Passive measurement techniques make use of the ongoing data traffic flows through a measuring node to estimate the network characteristics. Measurement probe generation is not required for such methodology. This method counts on the traffic packet size, interface capacity, packet flags, and timestamps of the passing cross traffic to estimate network related parameters independently. Such techniques can characterize network parameters associated to any node and over which an administrator has full control. Moreover, passive measurement is only possible with the availability of network traffic on the link that is measured. In addition, high infrastructural performance is necessary as the measurement nodes must have higher CPU capability and memory capacity to estimate a secondary activity other than packet routing. End-to-end measurement is not possible in passive measurement.

Multi Router Traffic Grapher (MRTG) [24] is an example of passive measurement tool that uses Simple Network Management Protocol (SNMP) to monitor and measure router traffic on associated links. In order to collect network related information, MRTG requests performance data using SNMP to the connected routers. An SNMP enabled router checks the Object Identifier (OID) to reply the request. MRTG receives SNMP encapsulated raw data and logs them with previously recoded data based for every connected router on a client machine. The final output of network performance is shown in graphs providing an idea of the network characteristics for a specific period of time. MRTG usually generates a traffic graph of every five minutes of interval. It provides maximum, average, and current values of link utilization measurement.

1.2 Challenges of Network Measurement

Various difficulties in network measurement have made this field of research more complex in achieving optimal results. From the previous section, we find that different measurement methodologies require different types of infrastructural support, implementation, and deployment assistance. In recent years, diverse network technology, increasing amount of traffic, QoS demand along others have put measurement issues as well as its accuracy into a challenge in order get better idea about network characteristics. Some of these challenging facts are mentioned below:

a) *Overload over legitimate traffic:* Internet links and intermediate nodes (i.e. routers and switches) carry various types of traffic to better manage the flow of legitimate application traffic over network for what intranets are built. Although network control (ICMP and various TCP packets) and management (various SNMP packets) traffic are necessary to ensure optimal operation and performance of connection and connectionless oriented application, this traffic adds no input for the final throughput estimation of any of such applications. For these reasons, bandwidth hungry and delay/loss sensitive applications adopt the User Datagram Protocol (UDP) rather than Transmission Control Protocol (TCP) to increase overall throughput performance.

This gives a better insight about the overhead that can be caused by active measurement probes. The basis of any proactive measurement scheme is to inject a number of probe packets for accurate measurement. Hence, such schemes have higher potential of grabbing a large amount of network resources during the probing period. For example, Pathload, which is an active available bandwidth measurement tool, usually generates around 2.5MB to 10MB of probing traffic for each measurement. This amount of non-legitimate data (non-application data) can badly degrade overall performance of any types of application traffic over a short period of time. Probing overload is one of the most significant challenges that comes first for network measurement.

b) *Network Technology:* With the steep increase of Internet traffic, new network technologies are emerging every year to accommodate bandwidth hungry services and their applications. No measurement scheme design is possible without considering various aspects of different network technologies. Maximum probe length, probe rate, the processing time of intermediate nodes of these technologies play a major role in devising new measurement schemes. For example, Pathload cannot adopt a probe length smaller than 96 bytes in order to make it operable over AAL5-ATM links, which avoid zero-padding to the probing packets. Similar consideration must be taken into account for other measurement tools.

c) Deployment of measurement methodology: Infrastructural support for probing methods has a direct impact on the deployment of a measurement scheme on any particular path. Decision for pursuing either an active or passive measurement approach is dependent on the applications and other network related issues such as network resource availability and their controllability. Considering all these aspects, adopting either a proactive or passive measurement is a big question. On the other hand, it is not possible to have a self standing scheme that can measure under all parameters in any circumstances. Deployment issues are another major trade-off issue for characterizing network parameters and hence for implementing estimation tools.

d) *Accuracy:* Accuracy is associated with the degree of the probing load. A bulky probe generating estimation tool that requires a large number of probes to figure out actual network conditions can very likely overestimate any of the network parameters, while a very strict control over the probing load can underestimate the same. Clock synchronization, measurement resolution, and estimation period can also directly impact the accuracy of any measurement tool.

Classification of measurement schemes and their challenges has been presented in the previous sections. In the following chapters, two different types of proactive measurement schemes (i.e., probe-gap and probe-rate model, and our proposed scheme) for available bandwidth and link capacity estimation based on these active measurement models will be emphasized.

CHAPTER 2

PROACTIVE MEASUREMENT SCHEMES

Even though active measurement schemes induce some extra load and an eventual unwanted intrusion on legitimate Internet traffic (application traffic), passive measurement schemes have not been proved to be the sole solution to end-to-end link estimation. Deployment of a new service to a remote user must be done over some the network territory that ISPs do not have any level of control. To ensure QoS connectivity and to comply with the Service Level Agreement (SLA), some degree of network forecasting capability is necessary for operational sustenance. In order to conform with such necessity, many proactive measurement tools have been proposed [2]-[9],[11].

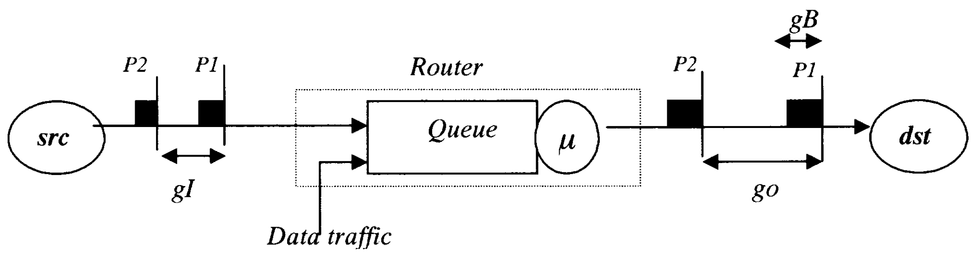
Amongst the various proactive schemes that have been proposed so far, most of them are capable of evaluating available bandwidth and link capacity measurement. Available bandwidth estimation tools such as IGI, Spruce, Train of Packet Pair (TOPP) [7], and link capacity estimation tools such as Pathrate, Bprobe, and Nettimer [1] are some of the examples of recent measurement tools. Although queuing delay, one way delay (OWD) [3], and jitter are some other networking parameters that require future in-depth interest for estimation, proactive methods for evaluating used and unused bandwidth capacity is the main focus of this work where various aspects of their feasibility of implementation and applicability have been evaluated. In this regards, the basic working principles of proactive measurement processes on the basis of probing mechanism, a comparative study of these schemes, and a brief review of some proposed tools is given in the following sections.

2.1 Probe-gap Model vs. Probe-rate Model

Proactive estimation tools inject probing packets, either as probe pair or probe train, on a source-destination path and check changes in the probes to collect link-state related information. A simple elaboration of both probe-gap model and probe-rate model is given below.

2.1.1 Probe-gap Model

Probe-gap model is generally used to measure tight-link capacity between end-to-end nodes by checking the significant dispersion of the initial gap of a probe pair by the time the pair reaches the destination to characterize unused link capacity. A probing train consists of multiple probe pairs that is injected into the network to get an average final estimation over a large period of time. IGI and Spurce are two examples of the available bandwidth measurement tools that use such probing mechanism. The principle measurement mechanism of gap model requires single hop assumption and a particular queuing region operation for successful estimation.



gI = Initial probe pair gap

Bc = Data traffic load

$P1$= 1st probe of packet pair of L Bytes

$P2$ = 2nd probe of packet pair of L Bytes

go = Output probe pair gap

gB = bottleneck gap (time required to process a probing packets by narrow-link)

Q = queue length

μ = Processing delay

Bo = Bottleneck capacity (link between *Router* and *dst* node)

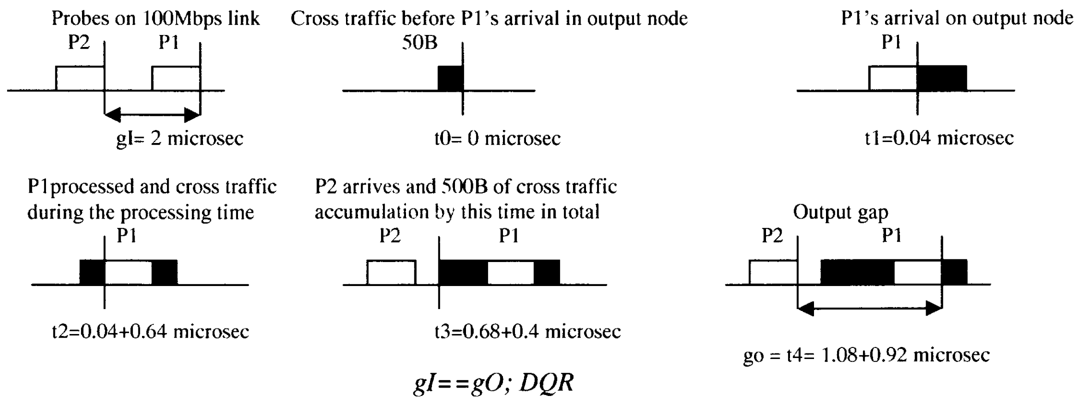
Figure 2.1 Single hop model for IGI and PTR estimation methodology.

Figure 2.1 shows a simple model of single hop architecture with some related link parameters and an intermediate node (e.g., a router/switch). In probe-gap model, a source to destination path with narrow-link and tight-link capacities on the same link (i.e., narrow-link is equivalent to tight-link) is usually referred as single hop model. A source node (*src*) shoots fixed sized (*L* bytes) probe packet pairs with a specific initial gap, *gI*, to the destination node (*dst*). The packets experience various types of delay (propagation, transmission, and queuing delay) while propagating to the *dst* node that contribute to the final dispersion, *go*, at the destination. Basically, the bottleneck link's (*Bo* and both the narrow-link and tight-link of the path) transmission capacity and amount of cross traffic (*Bc*) produce an impact on probe pairs' final gap changes along with some other factors such as queue length (*Q*). The parameter *gB* reflects the bottleneck's transmission capacity and the gap difference *(go-gI)* reflects the cross traffic load. Achieving an increased output gap, *go*, defines the validation for the gap model's efficiency.

An increase in the probe pair gap at the destination is the major estimation requirement for higher accuracy in the gap model and it is only achieved by generating probes into a specific queuing region called Joint Queuing Region (JQR). In short, the availability of probe packets of a probe pair at the same time in any queue before transmitting the first probe of the same pair onto the associated output link can be defined as JQR operation for the particular queue. *Bo* is directly dominated by the bottleneck link and the probe pair should not have an initial gap that would surpass the tight-link's capacity to get an increased output gap. The output gap increment is feasible only when *gI* corresponds to a probing rate smaller than the tight link capacity. The joint queuing region operation is the only way to maintain such specification. Theoretically, a pair of packets operate in the JQR if the second packet of the pair arrives at the node (queue) before the departure of the pervious packet. If not, the packets are called to be operating in the Disjoint Queuing Region (DQR). Let us refer to the following two scenarios for further elaboration.

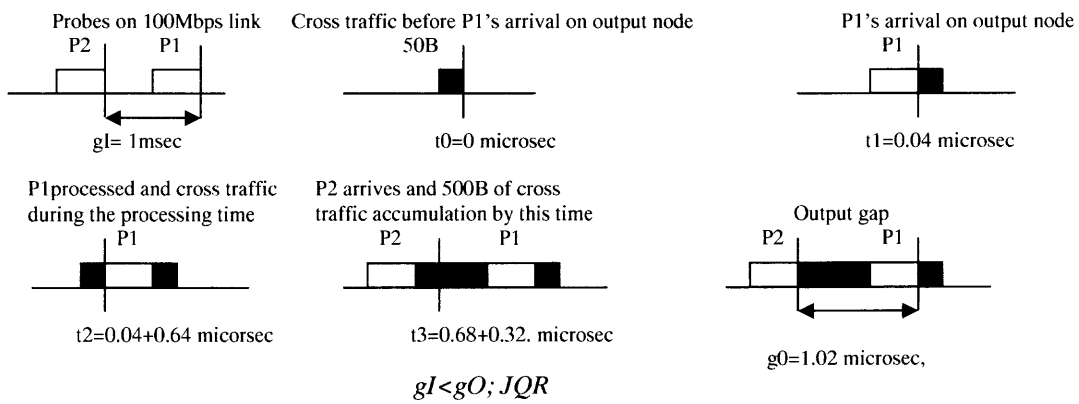
Initial Gap = 0.1msec and 0.15msec, Packet length =1200B, Input link=100Mbps, Output link=10Mbps

1st Case, Disjoint Queuing Region:



(a) Disjoint Queuing Region.

2nd Case, Joint Queuing Region:



(b) Joint Queuing Region.

Figure 2.2 (a) DQR and (b) JQR illustration.

The first scenario (i.e., Figure 2.2(a)) gives an example of disjoint queuing region operation. Here, the queue is empty upon *P2*'s arrival in the queue (Step 5) and the resulting output gap comes out smaller than that of the initial probing phase. Eventually, 500 bytes of cross traffic have no impact on the final gap.

In the second scenario (i.e., Figure 2.2(b)), the output gap at the destination is larger than the initial gap and the queue is not empty upon the arrival of packet *P2*. Clearly, the cross traffic directly contributes to the addition dispersion in the probing pair. If we had different cross traffic load the output gap would vary according to processing time of the the data packet for the same initial gap. This shows a linear relationship between *gI* and *Bc*. This is the reason why probe gap model's operation and accuracy depend on queue's JQR conformation.

$$Output\ probe\ pair\ gap,\ go = gB + \left(\frac{Bc * gI}{Bo}\right) \qquad (1)$$

$$Available\ Bandwidth = Bc - Bc\left(\frac{go}{gI}\right) \qquad (2)$$

Equations 1 and 2 estimate JQR and available bandwidth respectively, where the knowledge about bottleneck capacity, *Bc*, in these equations is required in advance. Knowledge of bottleneck capacity ensures that the output gap changes proportionately to the cross traffic load on tight link for a specified initial gap and hence the scheme is potential enough for correct measurement.

2.1.2 Probe-rate Model

Probe-rate model is also used to measure available bandwidth of an end-to-end network link. Probes are sent to a destination node and transmission rate plays a key for measuring link related capacities. Generally source node sends out a long train of probe packets with an initial rate and adjusts it during each iteration before reaching to a transmission rate that is equal to the receiving rate at the destination node of a particular path. As this model's probe transmission rate determination is not closely regulated by any link related criteria (i.e., queuing region) like gap model, starting probe rate is determined only by two simple parameters: probe length and inter probe gap timing (i.e., the transmission rate determined as $R = L/T$ bps where R = transmission rate, L = packet size, and T = time period). The same process continues to iterate by adjusting T to achieve different probing rate without further constraints, but such a scheme has to conform to data link layer's packet size specifications (i.e., minimum and maximum packet length for measurement accuracy and deployment). Such restriction basically defines the lowest possible available bandwidth estimation range of such scheme. Still, any effect on probing train caused by the cross traffic can be independently measured without any major constrains such as prior knowledge of bottleneck capacity.

Pathload and TOPP are two of the rate model based tools that measure unused link capacity of an end-to-end path. Unlike the gap model, the rate model is neither dependent on JQR operation nor its working principle depends on single hop model. The background principal of such methodology is that the packets in the probing train experiences an increasing delay at intermediate queues for relatively smaller unused capacity and eventually probe rate at the destination node becomes smaller than the initial transmission rate at the source.

For example, Self-Loading Period Streams (SLoPS) [3] is a probe rate model that injects K L-byte packets with an initial constant rate, R, that measures the OWD of the probing train at the destination node in order to measure unused link capacity. OWD is the time required by the destination node for processing a packet. If there are more packets waiting in the queue to be processed, the incoming packets have to wait longer, which results in relatively higher delays. In such way, probing trains with higher rate induces increasing delays for all the probe packets of the train at the destination. Otherwise, a non-increasing trend of OWD for probe packets would be experienced. In case of increasing OWD, the tight link capacity is not fast enough to transmit packets accordingly to the probe rate. Thus, it builds a large queue in its node and eventually negatively affects the probe rate at the destination end of the path.

PTR [2] is another implementation of the rate model for available bandwidth where the probing rate is characterized by cumulative dispersion of the probing train instead of its relative OWD. But the basis of measurement is same as Pathload. The source node keeps on sending probing trains with different rate until the probes' cumulative gap at the destination equals to that of source end i.e. non-increasing tendency of probe train. The final estimation is done by considering the amount of probe traffic captured during the arrival of the first and last probe of the probing train at the destination end. The PTR equation for available bandwidth:

$$Available\ bandwidth = \frac{\left(probe\_length * (probing\_train\_length - 1)\right)}{probing\_train\_gap\_at\_destination} \quad (3)$$

2.2 Comparative Study of Probe-gap Model and Probe-rate Model

This section covers some measurement aspects and the infrastructural condition that give a comparative study of these two proactive estimation processes in regard to performance and accuracy.

a) *Schematic dependencies:* Prior bottleneck capacity knowledge and Layer-2 packet length dependencies are the major restricting factors for performing gap model and rate model measurement respectively. Gap model is also limited by the Layer-2 Maximum Transmission Unit (MTU) limitation [19] but this limitation does not have a drastic effect on the measurement methodology as long as the queuing region operation is satisfied. Knowledge of accurate tight-link capacity knowledge and strict JQR operation are two crucial restricting factors for recruiting gap model. Hence, the rate model seems to be more flexible for performing measurement without stringent restrictions for better accuracy.

b) *Probing load:* As mentioned earlier, the rate model requires to send relatively larger amount of probe traffic because of its iterative mechanism of probing links. Note that, Internet traffic consists of many TCP connections that change their throughput rate based on link congestion, delay, and packet loss. Large amount of non-legitimate traffic is a good source for choking link bandwidth. So, Pathload, a probe rate model, has been found to be overestimating measurement in [6]. Over-estimated measurement due to probe load is not always dominant in gap model unless bottleneck bandwidth knowledge for adjusting probe rate is inaccurate. Moreover, controlled probeing train gap management is also capable of eliminating such phenomenon [6].

c) *Hardware induced limitations:* Packets travel through the network nodes before reaching their destination. These packets experience delays due to other packets queuing, scheduling speed, and link capacity restriction. But another source of delay is the processing time required by the nodes from receiving a packet before putting it to the outgoing link. According to Figure 2.3 below, even a back-to-back packet train experiences some dispersion at the output link even if there exist no cross traffic. The gaps in between packets (*P1, P2, P3*) on the output link can be in the range between 10-40 microseconds depending on the node's (router or switch) operational level (i.e., kernel level or user level [1],[6]).

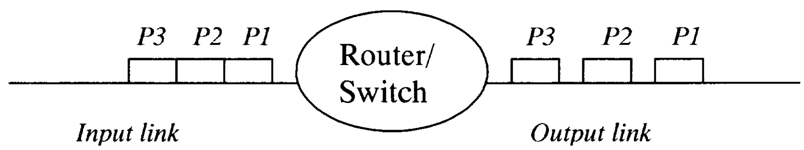


Figure 2.3 Back-to-back packet dispersion at a node's output link.

For gap model, this source of dispersion is not considered at the destination node. The final probe pair gap is assumed to be only induced by the cross traffic that intervenes the initial probe gap given that the probes are operating in JQR. At JQR, the output gap at the destination is theoretically proportional to cross traffic in between probes.

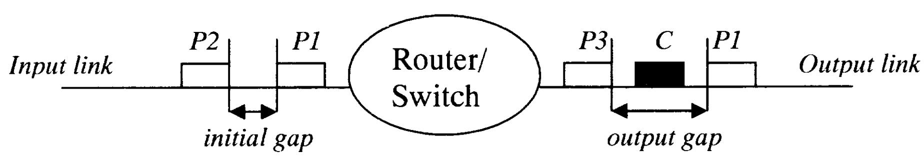


Figure 2.4 JQR operation of gap-model probes with probe packet *P1* & *P1* and cross traffic *C* at the output link.

Figure 2.4 illustrates the ignored hidden gap between *P1* and *P2* at the output link induced by a data packet *C* from the processing time required by the associated node. Even though JQR operation assumes to accurately measure cross traffic rate, probe-gap model literally overestimates the cross traffic load calculation due to such unavoidable extra gap, induced by the legitimate traffic flow. Moreover, the margin of such error is even higher for high load of cross traffic. The gap between probes (*P1 and P2*) proportionately increases in the order of *(n+1)* node's packet processing time* for each data packet found within a probe pair, where *n* refers to number of intervening data packets in within a probe. Hence, gap model's function is flawed and there is no feasible way to know the number of cross traffic packet that intrudes in between probes to correct this infrastructural error.

On the other hand, the probe-rate model is not affected by node's packet processing time. The rate model compares the relative increase in probe gaps at both source and destination rather than measuring the number of cross traffic packets for concluding the estimation process.

Considering the above, the probe-rate model proves to be a better candidate over probe-gap model for measuring unused link capacity. Moreover, with rate model, under suitable conditions, per-hop link estimation is also feasible. This constitutes an additional advantage for this network measurement methodology.

2.3 A Gap Model Tool for Available Bandwidth Estimation: Initial Gap Increasing

IGI sends out a long sequence of probing train to the destination with an initial gap conforming to JQR operation and monitors the average difference between the input (initial/source) gap and the output (output) gap for an equivalent value. Note that, IGI can also be considered as a rate model because it iteratively changes the probe transmission rate until finding a cumulative difference of the transmission and receiving rate that becomes zero. The point when the probe rate at both the sender and receiver nodes is equivalent is defined as turning point in IGI. IGI starts with a small initial gap value and keeps on increasing the same in order to get a transmission rate that matches with the probe receiving rate at the destination. The tool then uses the generic gap model equations (i.e., Equation 1 and 2) to finally estimate the unused link capacity for a particular end-to-end path.

IGI Algorithm:

```
/* initialization */
probe_num = 60
packet_size = 700B
Bo = Bottleneck capacity of the path
gB = 700B/Bo
initial_probing_gap = gB/2
gap_step = gB/8
cumulative_src_gap = probe_number * initial_probe_gap;

/* iterative algorithm of searching turning point (i.e transmission rate == receiving rate)
*/
while (cumulative_src_gap != cumulative_dst_gap) {
        initial_probing_gap += gap_step;
        cumulative_src_gap = probe_num*initial_probing_gap;
        SEND NEW PROBE TRAIN WITH NEW INITAIL RATE;
        MEASURE cumulative_dst_gap;
}

/* Available bandwidth estimation using IGI formula*/
MEASRUE inc_gap_sum (cumulative_dst_gap – cumulative_src_gap)
cross_traffic_load = Bo*inc_gap_sum/cumulative_dst_gap;
available_bw = Bo-cross_traffic_load;
```

Figure 2.5 The IGI algorithm.

In IGI, bottleneck capacity, *Bo*, is measured using a link capacity estimation tool, e.g., bprobe, nettimer and pathrate. Thus, its accuracy also depends a lot on the performance of these tools as the initial probe transmission rate (*initial_probing_gap*) as well as gap increment rate (*gap_step*) is closely related with accurate link capacity knowledge.

According to the IGI algorithm, suppose, out of sixty probe packets, M probe gaps experience positive dispersion, while N and K probe gaps experience zero and negative dispersion respectively regarding the initial probing gap at the turning point. $M+N+K$ constitute the total number of probe gaps (i.e., 59). Then IGI uses the following gap model equation to measure the cross traffic load for the particular time period of its estimation:

$$\frac{Bo\sum_{i=1}^{M}(g_i^{+} = gB)}{\sum_{i=1}^{M} g_i^{+} + \sum_{i=1}^{N} g_i^{=} + \sum_{i=1}^{K} g_i^{-}} \quad (4)$$

In Equation (4), the positive sign ('+') over *gi* indicates the increased initial gap while equal ('=') and negative ('-') sign illustrates zero and negative increase of initial gap at the destination respectively. Finally, IGI terminates its measurement procedure by deducting this traffic rate from the bottleneck link capacity measuring the available bandwidth of a source-destination path.

2.4 A Rate Model Tool for Available Bandwidth Estimation: Pathload

Unlike IGI & PTR, Pathload does not require prior knowledge of the bottleneck link capacity for estimation process. In pathload, a sender process shoots probing packets and adjusts the probing rate regarding the processing delay (i.e., OWD) of the probes to measure a representative available bandwidth of a path. Pathload searches for a range of available bandwidth where the probing train shows an insignificant non-increasing delay tendency over the path's tight link capacity.

SLoPS is the background technology being implemented in pathload. In SLoPS, the sender checks for relative delay variation of all probing packets at the destination for measurement. During the estimation period, if the transmission rate of the probing train is higher than the bottleneck capacity, the queue will be stacked with additional traffic and delay will ensue in exponential order for the following probes to arrive at that link. Otherwise, the increase of delay from the first to the last probing packet will not be significant enough to experience congestion at that particular link. Here, the sender timestamps each of the probe packets on the probing train prior to their transmission to the destination node. The node at the destination marks the arrival of these probes upon reception. The sender finally computes the OWD from these two timestamp differences. Assume, ti and ai are the two timestamps, then OWD = $ai - ti$. This value shows the delay tendency of the probing train and governs the next probing rate until measurement termination.

Pathload Algorithm:

```
Initially, Gmax=Gmin=Rmin=0

/* Non-increasing trend */
If R(n) < A
        Rmin = R(n)
        If(Gmin>0)
                R(n+1) = Gmin+Rmin/2
        Else
                R(n+1) = Rmax+Rmin/2

/* Increasing trend */
If R(n) > A
        Rmax = R(n)
        If(Gmax>0)
                R(n+1) = Gmax+Rmin/2
        Else
                R(n+1) = Rmax+Rmin/2

/* Grey region */
If Gmax <= R(n)
        Gmax = R(n)
        R(n+1)= Rmax+Gmax/2
Else (Gmin > R(n))
        Gmin=R(n)
        R(n+1)= Gmin+Rmin/2

/* Termination condition */
(Rmax – Rmin <=  ω ) || (((Rmax - Gmax <= χ)&&(Gmin – Rmin<= χ ))
```

Figure 2.6 Pathload's rate adjustment algorithm.

Figure 2.6 shows the iterative rate adjustment algorithm of pathload which is based on binary search technique. The algorithm starts with some initial values (e.g., *Rmax ,Rmin, Gmax* and *Gmin*) and iterates until it converges to one of the three termination conditions available for final available bandwidth calculation. Here ω and χ are the two resolution values named max-min rate boundary and grey-region boundary respectively. Upon termination, pathload gives a [*Rmax, Rmin*] range value of available bandwidth rather than a single value like most of the estimation tools.

Interestingly, instead of always expecting increasing and non-increasing delay trend for probes, pathload evaluates a complex delay tendency, which is neither increasing nor non-increasing for a particular probing period. This complex phenomenon is characterized by grey-region in Pathload. Pathload formulates such scenario through *Gmax, Gmin,* and χ parameters in its rate adjust algorithm. *Gmax* and *Gmin* represents the grey-region window, and χ one of the termination conditions of pathload in case of indeterminist OWD tendency.

As for grey-region elaboration, suppose, no probe packet till the 60-th probe of pathload's 100-packet long probing train, experiences significant increasing OWD tendency at the destination. At such point, the relative OWD for the 60-th packet is insignificant and approximately similar to that of its preceding probe packets. But the OWD tendency drastically alters for the following packets characterizing extra delay in the intermediate nodes for the probes and it continues till the last probe packet of the probing train. The relative OWD for the last packet (100-th packet) becomes larger than the first probe's OWD of the train due to the relative queuing delay for all the packets following the 60-th probe packet. At this point, it is not possible to clearly identify whether the probing rate is equivalent to or higher than the current available bandwidth. Note that, a single probing train experienced both non-increase and increasing delay tendency over a single probing period. This scenario does not satisfy any of the PCT test and PDT tests which determines the dominating OWD tendency for a probing train at the destination node. According to [3], this phenomenon is considered as the grey-region in SLoPS methodology.

The notion of pathload estimation process is to first, detect the increasing delay tendency of the OWD metric and second, adjust the transmission rate representing the delay trend to converge at an available bandwidth range specificed by the available bandwidth resolution and grey region resolution.

Relative OWD tendency of a probing train in a single iteration is deduced from two statistical estimation techniques called pairwise comparison test (PCT) and pairwise difference test (PDT). These two statistics measure the relative OWD in two different ways for each probing train of every iteration. According to pathload paper, a pre-processing step precedes these statistical testing by dividing the OWD array $\{D^1, D^2, ..., D^k\}$ into *Γ= sqrt(k)* number of sub-groups. Afterwards, the median of these subgroups are computed to get a final OWD array $\{\hat{D}^1, \hat{D}^2, ..., \hat{D}^\Gamma\}$. This whole pre-process eliminates the outline errors of the estimation process.

The pairwise comparison test (PCT) measures the proportion of probing packets that shows increasing delay trend using the following equation:

$$S_{PCT} = \frac{\sum_{k=2}^{\Gamma} I(\hat{D}^k > \hat{D}^{k-1})}{\Gamma - 1} \qquad (5)$$

Here, *I(X)* is '1' if the condition holds, otherwise '0' and has the range [0,1]. SPCT = 0.5 when OWDs are independent (i.e., there is no strict relative tendency to define whether the probe train has increasing or decreasing trend). Pathload sets this value to larger than 0.66 and samller to 0.54 for increasing and non-increasing trends, respectively, otherwise the train is declared ambiguous (i.e., the probe train is operating at the grey-region).

The pairwise difference test (PDT) measures the strength of relative OWDs of the probes in a train (i.e., how frequently the OWDs vary during their transmission). The smaller this variation of relative delays is the better the cumulative increase in OWDs is observed at the receiver. The PDT equation is

$$S_{PDT} = \frac{\hat{D}^{\Gamma} - \hat{D}^{1}}{\sum_{k=2}^{\Gamma} \left| \hat{D}^{k} - \hat{D}^{k-1} \right|} \qquad (6)$$

Here, the PDT varies around [-1,1] and has value of '0' if the delay tendency of the probing trans is independent (i.e., not co-related). Pathload sets this value to larger than 0.55 and smaller than 0.45 for increasing and non-increasing trend, respectively, otherwise the probing train is declared ambiguous.

Usually, an increasing or non-increasing tendency of PCT and PDT values with an ambiguous probe trend of any of the two statistical tests determine the definite increasing and non-increasing trend of a probe train. Otherwise, the probe train is characterized ambiguous. In pathload, the sender shoots one hundred probe packets in each probing train and each train is sent twelve times before modifying the probe transmission rate in the following iteration.

2.5 A Link Capacity Measurement Tool: Bprobe

Bprobe proactively sends out a sequence of ICMP ECHO packets from a source to a destination and waits for replies to measure the inter-arrival time of the consecutive probing packets. It assumes that the dispersion between a probe pair is inversely proportional to narrow-link capacity where probe pair dispersion with JQR operation, no packet loss, no intervening cross traffic, and no downstream congestion can represent the desired link capacity.

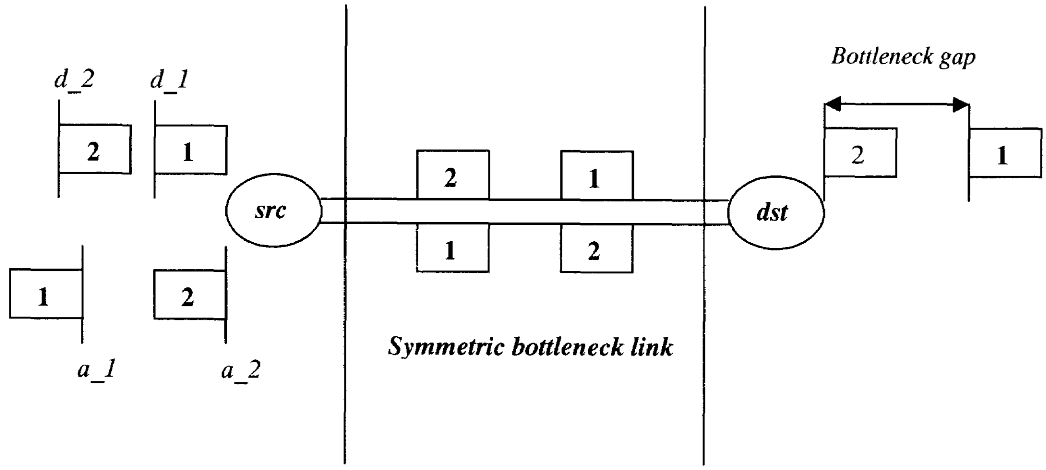


Figure 2.7 Bprobe's probe packet flow where the inter-probe dispersion is shaped by the bottleneck link capacity of the source-destination path.

Figure 2.7 shows a single hop topology for link capacity measurement with inter-probe transmission gap of *d_2 - d_1* sec and arrival gap of *a_2 - a_1* sec. The inter-arrival gap corresponds to the transmission time for the second ICMP packet of the pair that gives an estimate of the narrow-link capacity along the path. Given a probe size of P, the bottleneck capacity is calculated as

$$\textit{Bottleneck bandwidth, } B = \frac{P}{(a\_2 - a\_1)} \textit{ bytes/sec} \qquad (7)$$

Usually, bprobe sends a large number of ICMP ECHO probes with varying sizes and larger packet size in the successive probing iterations. It then filters out the inter-arrival time of the probe pairs in an effort to discard the inaccuracy caused by queuing failure (i.e., probes operating in DQR), cross traffic intervention, packet loss, and downstream congestion along its two way transmission from source to destination and destination to source, respectively, in order to measure the link capacity. Probe packet size for bprobe can be of 124 to 8000 bytes depending on the link technology. Bprobe sends 10 equal-sized packets from source to destination for probing in each iteration.

In bprobe, the probe pair's error-interval (i.e., probe pair dispersion caused by links other than narrow-link along a path) is eliminated by one of the two set operations: intersection filtering and union filtering. Since bprobe probes a source-destination path a number of times, it uses the output gaps of each of the 10-packet train to find out the most representing gap interval of the narrow-link bandwidth using these simple set operations. Here, each probing iteration is considered as a sample set in filtering process.

The intersection filtering method performs an intersection-set operation on all sample sets of measurement to determine out the most frequently occurring output gap (i.e., gap dispersion) at the destination. According to intersection-set operation, the intersected value corresponds to the most representing gap dispersion contributed by the bottleneck link that primarily shapes the output gap value filtering out the error-intervals induced by the links other than the narrowest link along the path. On the other hand, the union filtering method performs a union-set operation on the sample sets and takes the midpoint of the output gap value that occurs most frequently in the resulting probe set. Union filtering method is used in the current bprobe version.

CHAPTER 3

PROPOSED PROACTIVE MEASUREMENT SCHEMES

Two new estimation schemes for measuring available bandwidth and link capacity along a source-destination path are proposed in this chapter. The proposed available bandwidth scheme measures the tight-link capacity while the link capacity scheme computes the narrow-link bandwidth along with intermediate link capacities under certain limitations. Both probe-rate model and probe-gap model have been adopted to measure these two path related parameters. The following sections describe these two schemes.

3.1 Available Bandwidth Measurement Scheme

The background principle of the available bandwidth scheme is to shoot a fixed-length probing train of packets from a source node to a destination node with specific transmission rates and to check the rate of change of gaps at the destination to infer packet transmission capacity. The change of cumulative train length is computed from the cumulative gap difference of the probe pairs at the source and destination ends. The cumulative gap difference is the time space between the arrivals of the first/last bit of first probe to the first/last bit of the last probe in the probing train respectively. Change in cumulative gap is governed by the tight-link's capacity along the path, which is the smallest during the period of estimation. If the probe transmission rate is lower than or equal to the available bandwidth, no change in the train gap occurs, otherwise a change in the probing train gap is expected at the destination point. As in [2],[4], an iterative probing methodology is used in the proposed scheme to terminate the estimation upon reaching a rate with no significant change, equaling the available bandwidth.

```
Pseudocode for available bandwidth estimation:

/* Initial value */
        Probe train length = 30 packets
        initial minimum transmission rate = 0 Mbps (minrate)
        initial maximum probe transmission rate = 100,000Mbps (maxrate)
        initial g-probing rate =  minimum rate+(maximum rate-minmum rate)/3
        initial h-probing rate = minimum rate+2*(maximum rate-minmum rate)/3

/* Intermediate calculation */
        g_ingapdifference = gap of the probe train at the source end for g-probing
        g_outgapdifference =  gap of the probe train at the destination end for g-probing
        h_ingapdifference = gap of the probe train at the source end for h-probing
        h_outgapdifference =  gap of the probe train at the destination end for h-probing
        g_increase =g_ outgapdifference –g_ ingapdifference
        h_increase =h_ outgapdifference –h_ ingapdifference
        g_ increment =g_ increase*100/g_ingapdifference
        h_ increment = h_increase*100/h_ingapdifference

/* Iterative rate adjustment */
        If( (g_increment is less than 1%) &&(h_increment is less than 1%))
              minimum rate = h-probing rate
              g-probing rate =  minimum rate+(maximum rate-minmum rate)/3
              h-probing rate = minimum rate+2*(maximum rate-minmum rate)/3;

        If( (g_increment is less than 1%) && (h_increment is greater than 1.5%))
              minimum rate =g-probing rate
              maximum rate = h-probing rate
              g-probing rate =  minimum rate+(maximum rate-minmum rate)/3
              h-probing rate = minimum rate+2*(maximum rate-minmum rate)/3;

        if(g_ increment is greater than 1.5%)
              maximum rate = g-probing rate
              g-probing rate =  minimum rate+(maximum rate-minmum rate)/3
              h-probing rate = minimum rate+2*(maximum rate-minmum rate)/3;

/* Termination Condition */
        If ( g_increment or h_increment is in between 1% to 1.5% )
              available bandwidth = g-probing rate (for g_increment)
              available bandwidth = h-probing rate (for h_increment)

        If (h-probing  rate – g-probing  rate = 0.5Mbps)
              available bandwidth = (g-probing rate+h-probing rate)/2
```

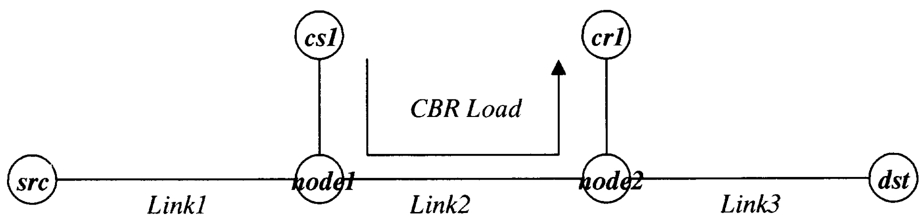
Figure 3.1 Proposed algorithm for measurement of available bandwidth.

In the proposed scheme, the ternary search algorithm [21] for iterative rate adjustment of probing traffic rate has been implemented, where it sends two probing trains with different transmission rates (e.g., g-probing rate and h-probing rate) in each iteration. These two probing rates simultaneously probe the path where these values are taken as one-third and two-third of the *minimum* and *maximum* rate range for all instances, respectively. The ternary search algorithm is faster than the binary search algorithm used in pathload, considering the number of iterations required and the amount of probe load generated during the measurement period. For example, according to the pseudocode in Figure 3.1, when both the *g-probing* rate and *h-probing rate* have increased delay tendency, the probing window for the next iteration narrows down to *minimum rate* to the *g-probing rate* range that constitutes only one-third of the whole of *minimum* and *maximum* rate window. Same condition occurs for decreasing delay tendency of the two probing rates. Hence, ternary search has the higher possibility of finding the available bandwidth that with smaller number of iterations and eventually less non-legitimate traffic load.

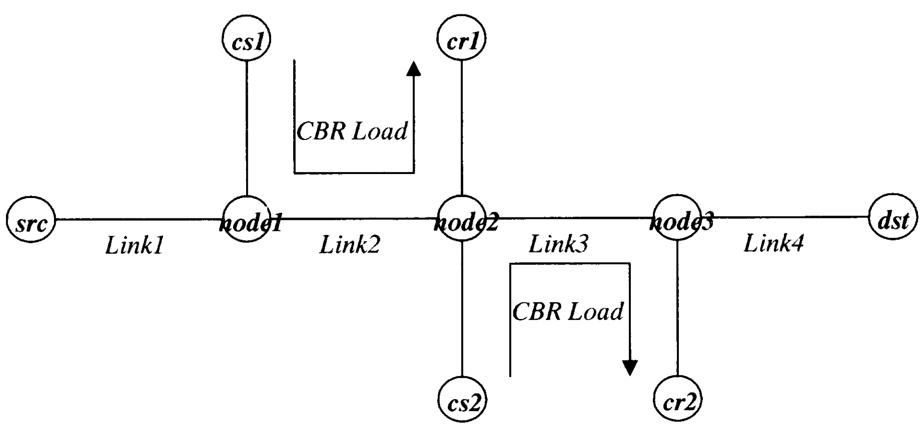
The proposed measurement scheme is an improvement over both IGI and pathload methodologies concerning the probe rate adjustment and the way it detects probes' delay tendency techniques that are being used in those two tools. The proposed scheme has adopted the simple rate of cumulative gap change model for probes' delay tendency detection from IGI. This scheme only compares the probing train's gap values at the source and destination for adjusting the next transmission rate. On the other hand, pathload uses OWD detection methodology, which requires higher degree of computation complexity. Moreover, this scheme does not require prior knowledge of any link information (i.e., narrow-link capacity) like pathload for computing the available bandwidth of a particular path. The basic principle of the scheme is to devise a comparatively low load inducing and relatively simple measurement methodology.

3.1.1 Simulation Results

Available bandwidth measurement scheme has been tested in simulation environment using the ns2 simulator [22]. Following are the network topologies that were used for various experiments to verify the proposed measurement schemes:

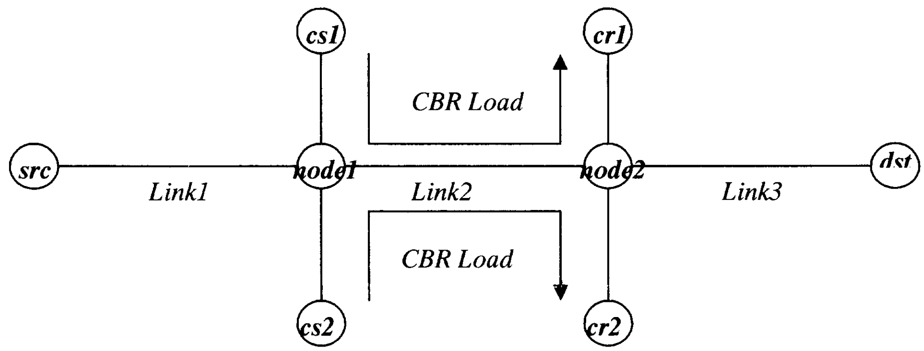


(a) Single-hop network topology.



(b) Multiple-hop network topology.

Figure 3.2 (a) Single-hop and (b) multiple-hop network topologies with single cross traffic.



(c) Single-hop network topology with double source cross traffic load.

Figure 3.2 (c) Single-hop network topology with multiple source cross traffic load (Continued).

3.1.1.1 Probing Train Length for Available Bandwidth Measurement. Probing train length for pathload and IGI are 60 and 100 respectively. As both of them iteratively adjusts their probing rate, the amount of probe load during estimation period is closely related to probing train length parameter. Here, the effect of probing train length over available bandwidth measurement accuracy have been investigated using single hop cross traffic load simulation using the binary search algorithm [20], which has been implemented in pathload for its rate adjustment algorithm. Table 3.1 shows the simulation results of four different probing train lengths.

Table 3.1 Available Bandwidth (AB) Measurement using Binary Search and Single-hop Load for Various Probing Train Lengths

Binary search for AB measurement with single-hop CBR load						
Probe packet size: 800, Termination resolution: 0.15Mbps, Iteration: 100						
			Measured AB(Mbps)			
Link(1,2,3) (Mbps)	**Load (Link2)**	**Actual AB (Mbps)**	**trainlength_ 20**	**trainlength_ 30**	**trainlength_ 60**	**trainlength_ 100**
20	19	1	1.25	1.25	1.25	1.25
20	18	2	2.26	2.26	2.26	2.5
20	17	3	3.2	3.24	3.28	3.75
20	16	4	4.2	4.26	4.21	5
20	15	5	5.33	5.25	5.27	5.62
20	14	6	6.18	6.23	6.25	6.25
20	13	7	7.19	7.26	7.27	7.5
20	12	8	8.26	8.44	8.28	8.75
20	11	9	9.18	9.14	9.23	9.38
20	10	10	10.11	10.22	10.22	10.71
20	9	11	11.4	11.14	11.22	11.4
20	8	12	12.18	12.12	12.24	12.81
20	6	14	14.29	14.23	14.19	14.91
20	5	15	14.05	15.27	15.21	15.62
20	4	16	16.31	16.19	16.24	16.36
20	2	18	15.63	17.21	18.2	18.43
20	1	19	18.64	18.9	19.13	19.84

As in Table 3.1, measurement results for probing train lengths of 30 and 60 seem to have almost similar accuracy, but 100-probe packets in a single probing train overestimates the measurement comparing to the other used train lengths. The reason for such output is due to probe packet's higher rate of intervention in CBR traffic along the source-destination path. Regarding these results, the scheme shoots 30 probing packets in every probing iteration for simulating the proposed measurement scheme in ns-2 considering less intrusion rate and faster measurement time.

As for probe length, a length of 800 bytes for probe packet has been adopted. According to [2], probe sizes of around 700 bytes to 1500 bytes are good enough to give accurate measurement result for available bandwidth. Note that, probe packet's length dependency is extensively elaborated in [14].

3.1.1.2 Simulation Results of the Proposed Measurement Scheme. Table 3.2 shows the simulation results of the proposed available bandwidth measurement scheme that uses 800 bytes in a probe packet, 30-packet probing train, and ternary search algorithm as some of its process parameters. Each of the following results are obtained with 25 iterations performed by the proposed scheme for each particular available bandwidth measurement. In this table, the average error for single hop cross load is way below 10% and this accuracy sustains for 90% load of the link's capacity. Error rate for 1Mbps and 0 Mbps of unused link capacity is steep but it is usual to have such erratic measurement due to measurement's algorithmic termination condition, and extreme cross traffic load condition. Moreover, the average error rate of all the existing measurement tools is over 20% which validates the significance of our proposed scheme.

Table 3.2 Available Bandwidth (AB) Measurement of a Source-destination Path with Single-hop Single Source CBR Traffic Load

Ternary search for AB measurement with single-hop CBR load			
Probe packet size: 800B, Trainlength:30, Iteration:25, Termination Resolution: 0.5Mbps			
Link (1,2,3) (Mbps)	Load (Link2) (Mbps)	Actual AB (Mbps)	Measured AB (Mbps)
20	**1**	19	18.37
20	**2**	18	17.48
20	**3**	17	16.56
20	**4**	16	15.87
20	**5**	15	14.69
20	**6**	14	13.67
20	**7**	13	12.76
20	**8**	12	11.55
20	**9**	11	10.86
20	**10**	10	9.67
20	**11**	9	8.93
20	**12**	8	8.02
20	**13**	7	6.91
20	**14**	6	5.72
20	**15**	5	5.01
20	**16**	4	3.95
20	**17**	3	3.13
20	**18**	2	2.07
20	**19**	1	1.20

Table 3.3 shows the measurement results for multi-hop cross load simulation, Figure 3.2(b), where two CBR traffic sources generate data packets over the source-destination path. Different combinations of cross load have been tested here to verify the accuracy of the scheme. This scenario is complex and is comparable to the realistic network scenario. Still the measurement results conform to an error rate of below 10% in every cases. As an example, for 16 Mbps of available bandwidth, the estimated output has been found to be 14.88 Mbps, where the error-rate is 7%.

Table 3.3 Available Bandwidth (AB) Measurement of a Source-destination Path with Multiple-hop CBR Traffic Load

Ternary search for AB measurement with multiple-hop (2 CBR sources) random load				
Probe packet size: 800B, Trainlength:30, Termination resolution: 0.5Mbps, Iteration: 25				
Link (1,2,3,4) (Mbps)	**Load (Link2) (Mbps)**	**Load (Link3) (Mbps)**	**Actual AB (Mbps)**	**Measured AB (Mbps)**
20	0.5	0.5	19.5	18.32
20	1	1	19	17.35
20	1	2	18	16.76
20	1	3	17	16.37
20	2	3	17	15.92
20	4	2	16	15.25
20	4	3	16	14.88
20	5	3	15	14.04
20	5	4	15	14.09
20	6	4	14	13.45
20	5	6	14	12.79
20	5	7	13	12.81
20	10	3	10	9.80
20	4	10	10	9.72
20	11	5	9	8.81
20	13	12	7	6.59
20	14	10	6	5.95
20	15	4	5	5.01
20	16	10	4	4.17
20	17	15	3	3.20

A number of single-hop cross traffic experiment with multiple CBR traffic sources has also been tested to validate the robustness of the scheme. Figure 3.2(c) shows 2 CBR traffic generators which provide 2 data flows on *link 2*, which is been used to simulate our desired network scenario. Again, according to the measurement results in Table 3.4, the average measurement accuracy seems to be consistent and similar to that of the previous to network topologies.

Table 3.4 Available Bandwidth (AB) Measurement of a Source-destination Path with a Single-hop and Double Source CBR Traffic Load

Ternary search for AB measurement with multi source (2 CBR sources) random load			
Probe packet size: 800B, Trainlength:30, Termination resolution: 0.5Mbps, Iteration: 25			
Link (1,2,3) (Mbps)	**Load (Link2) (Mbps)**	**Actual AB (Mbps)**	**Measured AB (Mbps)**
20	0.5+0.5	19	18.61
20	1+1	18	16.93
20	1+2	17	15.75
20	1+3	16	15.13
20	2+3	15	14.00
20	4+2	14	13.03
20	4+3	13	12.61
20	5+3	12	11.65
20	5+4	11	10.46
20	6+4	10	9.70
20	5+6	9	8.74
20	5+7	8	8.04
20	10+3	7	7.01
20	4+10	6	5.92
20	7+8	5	5.28
20	14+2	4	4.04
20	10+7	3	3.13
20	15+3	2	2.09
20	15+4	1	1.30

Figure 3.3 shows a measurement comparison of pathload's ns2 implementation [23] and the proposed scheme on a single-hop cross traffic topology (e.g., Figure 3.2(a)) with 200Mbps of link capacity. Here, the x-axis denotes the cross traffic load during measurement and the y-axis denotes the estimated available bandwidth for the corresponding load by the above mentioned schemes. Although, pathload seems to give more accurate output than the proposed scheme, the error rate in the later case is never above 3% of the actual available bandwidth. Moreover, the ns2 implementation for pathload requires further study as it gives below 1% error rate for such experimentation. However, considering the performance of the proposed scheme, it can be assumed as a consistent methodology for measuring available bandwidth.

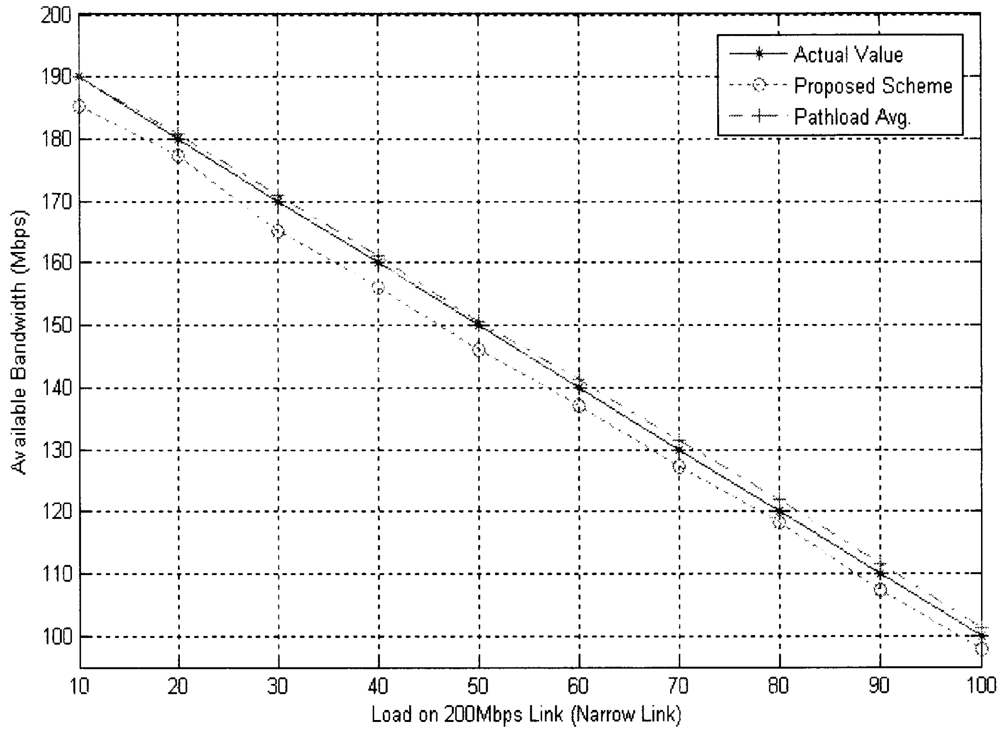


Figure 3.3 Available bandwidth measurement on single-hop cross traffic topology: Pathload vs. Proposed Scheme.

3.2 Link Capacity Measurement Scheme

Unlike pathrate and bprobe, the proposed scheme efficiently uses the ICMP timestamp requesting packets to measure link capacity between every intermediate nodes down to the destination before the final estimation of the narrow-link capacity of a source-destination path. If there are *(n-2)* intermediate nodes on the end-to-end path, the sender first shoots probe packets to the first intermediate node (*1st* node) and iteratively reaches destination node (*nth* node*)* before terminating the final measurement. Thus, this technique is considered as a hop-by-hop measurement technique, which is different than those of pathrate and bprobe. Note that, cing measures internal queuing delays of all intermediate nodes along a path in hop-by-hop manner.

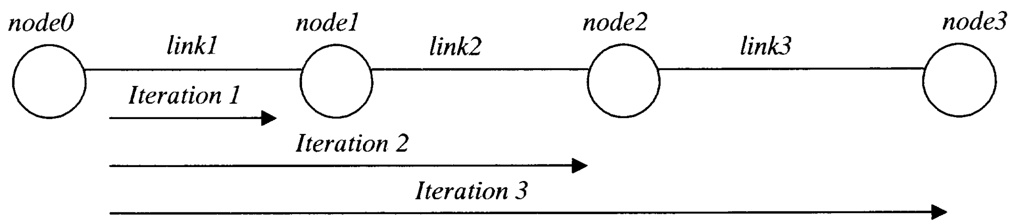


Figure 3.4 Iterative methodology for measuring narrow and intermediate link capacities.

For link capacity estimation, the proposed scheme shoots compound probe packets, which are completely different from that of the other probe models used in various tools for network measurement. The scheme uses a long probing train of multiple compound probes separated by enough inter-probe gap that respects the available bandwidth of the source destination path. The following figure shows the structure of our newly devised compound probe.

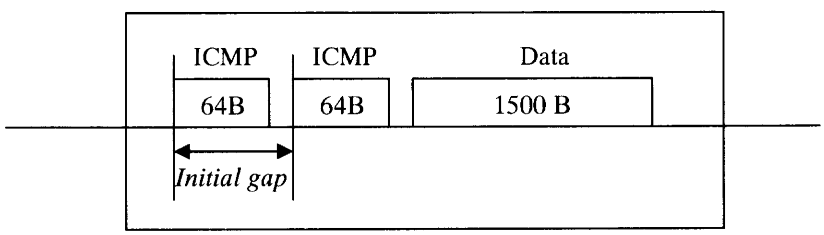


Figure 3.5 One compound probe.

Figure 3.5 shows the structure of a compound probe packet. Here, there are two ICMP timestamp requests that follow a redundant 1500B data packet. These three packets are considered as a single compound probe packet. The intra-compound-probe gap (i.e., the gap between the last and first bits of any of the two probes in the compound probe structure) is taken to be 40 microseconds. This is the maximum time required by a real Internet node (e.g., router or switch) to process an incoming packet before putting it onto an output link. The reason behind having 40 microseconds of intra-compound-probe gap is to avoid any short of dispersion in between these packets that can be caused by the cross traffic at the destination. With this time value the scheme has the potential to avoid an error in the *initial gap* interval that is critically eliminated by union set operation in bprobe discussed in Chapter 2. With this precise intra-compound-probe gap and *initial gap*, it is ensured that back-to-back probe packets are received at every destination. Interestingly, the 1500B data packet is discarded using appropriate TTL value and the scheme only counts on the ICMP replies for link capacity measurement. The proposed scheme has adopted a default probing train length of 30 packets for the link capacity measurement.

Schematically, the sending node shoots a probing train and wait for the ICMP timestamp replies at the destination where the data packets get discarded. The main idea behind is to send 1500 bytes data packets to ensure that the two ICMP timestamp request packets operate in JRQ.

For link capacity estimation, it is essential to get the actual time gap value that is required to process specified number of packets (e.g., probe packets) by the narrow link. Here, the compound probe structure has been designed to efficiently measure the exact dispersion between the last bits of the two ICMP probes at the destination nodes that represents particular link capacity. The compound probe scheme eliminates the possibility of stretching the initial gap on the narrow-link when there is no cross traffic to force JQR operation. This assumption also holds for all the intermediate links along the path under certain circumstances given that the output link at the destination nodes are not fast enough to process the 1500B data packets before the two following two ICMP probe reaches the queue. Simulation results are provided in the following sections to verify this claim.

3.2.1 Link Capacity Measurement Steps

Step1. The source node sends a compound probe train to the first intermediate node requesting timestamp for measurement. The initial intra-compound-probe gap is taken as 40 microseconds.

Step2. The destination node replies the timestamp request by providing its own timestamp values and drops the 1500 byes data packet of every compound probe of the train.

Step3. The source node computes the difference (output gap value) of the two returned timestamp values of each compound probe packets and takes the minimum output gap value for link capacity measurement.

Step4. The source computes the link capacity dividing the ICMP probe length (64 bytes) by the minimum output gap value determined in Step 3.

Step5. The source repeats Steps 1 to 4 for each intermediate node along the path until it reaches the destination node.

Step6. Finally, the source node calculates the narrow-link capacity by taking the minimum of all the link capacities along the source-destination path measured in Steps 1 to 5.

Figure 3.6 shows the pseudocode for link capacity measurement.

```
Pseudocode of link capacity measurement:

Probing train length = 30 compound packets

For n nodes along a path from source to destination

 i=0
While(i != n) {
     Source shoots compound probe to i-th destination node
              For 10 times
                       outputgap = min(timestamps reply of the second ICMP packet - timestamp
                       reply of the first ICMP packet)
              Link capacity of i-th node = 64*8/outputgap
     Updates destination node to i++
}

Narrow-link bandwidth = min(i-th link capacity, i+1-th link capacity, ..., n-th link capacity)
```

Figure 3.6 Pseudocode for link capacity measurement.

3.2.2 Simulation Results

This section provides the results of ns2 simulation experiments to validate the proposed link capacity measurement scheme. The same single-hop topology as in Figure 3.2(a) is used for single hop cross traffic measurement. Also, for link capacity with no load measurement, similar topology as in Figure 3.2(a) is used without cross load on *Link2.* Multiple-hop cross traffic load measurement (Table 3.8) was simulated with the following topology in Figure3.7.

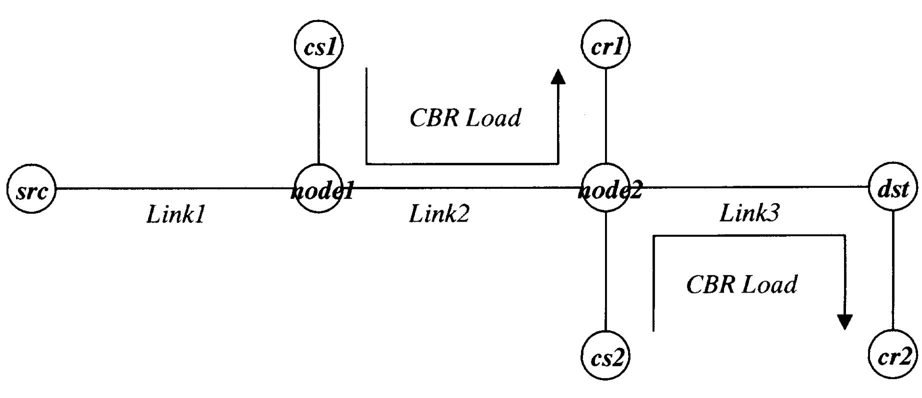


Figure 3.7 Multiple-hop cross traffic network topology.

3.2.2.1 Link Capacity Measurement Results with No Load and Load. Table 3.5 shows the simulation results for hop-by-hop link capacity estimation for different link capacities of the 2nd link (*Link2*), i.e. the link between *node1* and *node3*, of the single hop topology when there is no cross traffic available on any of the links along the source-destination path. Note that, all these results show the average of 10 iterations for each link's capacity measurement. According to the simulation results, we have almost accurate link capacity measurement for all the links along the path that matches

Table 3.5 Hop-by-hop Link Capacity Measurement Results of a 3-hop Source-destination Link

Link capacity in Mbps with no cross traffic					
Compound probe size: 64+64+1500 Bytes, Iteration: 10					
Actual Link Capacity			**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	**20**	20	19.99	20	20
20	**19**	20	19.99	19	20
20	**18**	20	19.99	17.99	20
20	**17**	20	19.99	16.99	20
20	**16**	20	19.99	15.99	19.99
20	**15**	20	19.99	15	20
20	**14**	20	19.99	13.99	19.99
20	**13**	20	19.99	13	19.99
20	**12**	20	19.99	12	19.99
20	**11**	20	20	10.99	20
20	**10**	20	19.99	9.99	19.99
20	**9**	20	19.99	8.99	19.99
20	**8**	20	19.99	7.99	19.99
20	**7**	20	19.99	6.99	19.99
20	**6**	20	19.99	5.99	19.99
20	**5**	20	19.99	4.99	19.99
20	**4**	20	19.99	3.99	19.99
20	**3**	20	19.99	2.99	20
20	**2**	20	19.99	1.99	20
20	**1**	20	20	0.99	20

with values of each of the links. Even for the difficult scenario when the 2nd link has 1Mbps of link capacity, the output has similar accuracy to that of the other combinations of the link capacities in the similar network topology.

Table 3.6 Hop-by-hop Link Capacity Measurement Results for Descending and Ascending Order Link Capacities of a 3-hop Source-destination Path

Link capacity in Mbps with no cross traffic					
Compound probe size: 64+64+1500 Bytes, Iteration: 10					
Actual Link Capacity			**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	12	6	19.99	12	5.99
20	18	14	19.99	17.99	13.99
20	18	15	19.99	17.99	15
20	19	18	19.99	19	17.99
30	20	17	30	20	16.99
20	25	20	19.99	24.99	20
20	25	2	19.99	24.99	1.99
100	150	90	99.99	149.99	89.99
15	18	14	14.99	17.99	13.99
15	18	15	14.99	17.99	15
20	16	16	19.99	15.99	15.99
20	16	17	19.99	15.99	16.99
20	16	18	19.99	15.99	17.99
30	12	12	30	12	12
30	12	18	30	12	17.99

Table 3.6 shows similar simulation topology with no cross traffic load but having more realistic topological scenarios. Here, the proposed scheme is simulated for various link capacities along a particular path that has both decreasing and increasing link bandwidth down towards the destination node. In the first seven rows, the link bandwidths have decreasing link bandwidths and the measurements, once again, show almost accurate estimation for all instances. As an example, for link capacities of 100Mbps (*Link1*), 150Mbps (*Link2*) and 90Mbps (*Link3*) our scheme is providing 99.99Mbps (*Link*1), 149.99Mbps (*Link2*), and 89.99Mbps (*Link3*) of link bandwidth.

Table 3.7 and 3.8 show simulation results with cross traffic load on different intermediate links along the path. In Table 3.7, the CBR cross traffic load is in the range of 2Mbps and 13Mbps on *link2* in every instance. Here, the simulation scenario has equivalent link capacity topology and also single point narrow-link (i.e. *link2*). In every cases, the measurement outputs provides desired accuracy.

Table 3.7 Hop-by-hop Link Capacity Measurement Results with Cross Traffic Load on the 2nd Link (*Link2*) of a 3-hop Source-destination Path

Link capacity in Mbps single-hop cross traffic (CBR)						
Compound probe size: 64+64+1500 Bytes, Iteration: 10						
Actual Link Capacity				**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	Load at *Link2* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	20	20	**2**	19.99	19.99	19.99
20	20	20	**4**	19.99	19.99	19.99
20	20	20	**6**	19.99	20	20
20	20	20	**8**	19.99	20	20
20	20	20	**10**	19.99	19.99	19.99
20	20	20	**12**	20	20	20
20	15	20	**13**	20	15	20
20	15	20	**12**	20	15	20
20	15	20	**11**	20	15	20
20	15	20	**10**	20	15	20
20	15	20	**9**	19.99	14.99	19.99
20	15	20	**8**	19.99	14.99	19.99

Again, Table 3.8 shows measurement results with two cross traffic sources which characterizes multi-hop network path (e.g., Figure 3.7) in real life scenario. Two traffic CBR generators have been used in this simulation experiments to generate legitimate data load on *link2* and *link3* respectively in order to design a difficult network condition. Even with high link capacity (e.g., 150 Mbps) and high cross traffic load (e.g., 12Mbps over 20Mbps link) the proposed scheme gives satisfactory results as above.

Table 3.8 Hop-by-hop Link Capacity Measurement Results with Cross Traffic on the 2nd Link (*link2*) and 3rd Link (*link3*) of a 3-hop Source-destination Path

Link capacity in Mbps multiple-hop cross traffic (CBR)							
Compound probe size: 64+64+1500 Bytes, Iteration: 10							
Actual Capacity					**Measured Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	Load at *Link2* (Mbps)	Load at *Link3* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	15	20	**10**	**10**	19.99	14.99	20
20	15	20	**12**	**12**	19.99	14.99	20
20	15	20	**14**	**14**	20	15	20
20	20	20	**10**	**10**	19.99	19.99	20
20	20	20	**12**	**12**	19.99	20	20
20	20	20	**14**	**14**	19.99	20	20
20	20	20	**16**	**16**	20	20	20
20	20	20	**18**	**18**	20	20	20
20	150	20	**6**	**6**	19.99	149.99	20
20	150	20	**8**	**8**	19.99	149.99	20
20	150	20	**10**	**10**	19.99	149.99	20
20	150	20	**12**	**12**	19.99	149.99	20

3.2.2.2. Impact of 1500B Data Packet in Link Capacity Measurement. Section 3.2 describes the notion of having a data packet in our compound probe structure. This particular probing component forces the two ICMP packets to operate in JQR for accurate measurement. Otherwise, this packet is not used in numerical calculation of link measurement. This phenomenon is also been described in [13] and [16]. In addition to that, a separate experiment with two ICMP packets without the 1500B packet in front in a compound probe has been tested in simulation. As desired, the final estimation shows low accuracy for both hop-by-hop link capacity and narrow-link capacity estimation for a 3-hop source-destination path topology according to Table 3.9. This result eventually validates the necessity of having a wedging probe component in front of the two ICMP packets in order to ensure accurate link measurement. Moreover, the accuracy of the results in Table 3.5-3.8 further verifies such a design idea.

Table 3.9 Hop-by-hop Link Capacity Measurements and Narrow-link Estimation Using 2 ICMP Packets

Link capacity in Mbps with no cross traffic					
Compound probe size: 64+64 Bytes, Iteration: 10					
Actual Link Capacity			**Measured Link Capacity**		
Link1(Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	20	20	12.79	13.51	13.52
20	18	20	12.79	13.31	13.32
20	16	20	12.79	13.11	13.11
20	14	20	12.79	12.91	12.91
20	12	20	12.79	11.99	11.99
20	10	20	12.79	10	10
20	8	20	12.79	7.99	7.99
20	6	20	12.79	6	6
20	4	20	12.79	3.99	3.99
20	2	20	12.79	1.99	1.99

3.2.2.3. Data Packet Length Dependency in Link Capacity Measurement. Because of various possible link capacities in Internet, 1500B of redundant data packet has a limitation for measuring link bandwidth along a source-destination path. Such limitation is closely dependent on the nodes' packet processing time describe earlier in this section. For the proposed scheme, 40 microseconds is been considered for intra-compound–probe gap for link measurement which is equivalent to the maximum packet processing time required by a router or switch. The redundant data packet forces the first ICMP packet to wait in the output queue for JQR operation at least for 40 microseconds before the second ICMP is available to be sent immediately afterwards. Otherwise, the dispersion caused by the first probe's processing time provides error-interval in probe gaps that doesn't conform with the actual link capacity and eventually the measurement output shows low accuracy. So, a high link capacity that processes 1500B of redundant packet earlier than the intra-compound-probe gap would prove to be infeasible for measurement. As such, determining measurement threshold limit for a single 1500 B packet is essential.

Table 3.10 Link Capacity Measurement Dependency over Compound Probe's Data Packet with a Length of 1500 Bytes

Link capacity and data packet (in-front of ICMP packets) length dependency					
Compound probe size: 64+64+1500 Bytes, Iteration: 10					
Measurement done with cross traffic on the path					
Actual Link Capacity			**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	20	50	19.99	20	50
20	40	50	19.99	40	49.99
20	80	50	19.99	79.99	49.99
20	100	50	19.99	99.99	49.99
20	150	50	19.99	149.99	49.99
20	200	50	19.99	200	49.99
20	250	50	19.99	250	49.99
20	300	50	19.99	300	49.99
20	350	50	19.99	350	49.99
20	400	50	19.99	399.99	49.99
20	425	50	19.99	425	49.99
20	450	50	19.99	449.99	49.99
20	460	50	19.99	460	49.99
20	465	50	19.99	464.99	49.99
20	467	50	19.99	467	49.99
20	468	50	19.999999	468	49.99
20	**469**	50	19.999999	**463.78**	49.99
20	**470**	50	19.999999	**445.11**	49.99
20	**475**	50	19.999999	**373.64**	49.99
20	**480**	50	19.999999	**341.76**	49.99
20	**500**	50	19.999999	**256.09**	49.99

Table 3.10 is the measurement results for a 1500B of redundant data packet and the link bandwidth threshold for this packet length seems to be 468Mbps. Beyond this bandwidth capacity, results for *link2* capacity measurement is erroneous and measurement continues to degrade further with increasing bandwidth. In such case, the error-interval increases proportionally to the output link's capacity that eventually destroys the conformity required for JQR operation. Table 3.11 is another set of measurement results for one 1000B of redundant data packet where the threshold link capacity is 312Mbps.

Table 3.11 Link Capacity Measurement Dependency over Compound Probe's Data Packet with a Length of 1000 Bytes

Link capacity and Data packet (in-front of ICMP packets) length dependency					
Compound probe size: 64+64+1000 Bytes, Iteration: 10					
Measurement done with cross traffic load on the path					
Actual Link Capacity			**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	20	50	19.99	19.99	49.99
20	40	50	19.99	40	49.99
20	80	50	19.99	79.99	49.99
20	100	50	19.99	99.999	49.99
20	150	50	19.99	149.99	49.99
20	200	50	19.99	200	49.99
20	250	50	19.99	250	49.99
20	300	50	19.99	300	49.99
20	310	50	19.99	309.99	49.99
20	312	50	19.99	312	49.99
20	**313**	50	19.99	**307.65**	49.99
20	**314**	50	19.99	**298.66**	49.99
20	**315**	50	19.99	**293.99**	49.99
20	**320**	50	19.99	**258.90**	49.99
20	**330**	50	19.99	**215.71**	49.99
20	**340**	50	19.99	**182.52**	49.99
20	**370**	50	19.99	**142.92**	49.99

Experiment results in Table 3.10 and 3.11 is been achieved in ns2 simulation where packet processing time for ns2 nodes is different than that of real routers or switches. Given that the *initial gap* in between compound probes are 40 microseconds and the redundant probe length is 1500B, the link capacity estimation threshold for the same packet length can be defined as

$$\text{Threshold link capacity} = \text{Packet\_length}*8/\text{initial gap} \quad (8)$$

$$= 1500*8/0.00004$$

$$= 300{,}000{,}000 \text{ bytes/sec}$$

In the simulated results, ns2 simulator processes packets faster (i.e., 25 microseconds) than real routes or switches and thus does not follow the theoretical threshold bandwidth calculation. Note that, this theoretical estimation of 300Mbps for

1500B is a rough calculation to define the upper range of link bandwidth estimation in the proposed scheme and it requires further investigation through Internet experimentations.

3.3 Integration of the Two Measurement Schemes

The proposed measurement scheme estimates two different network parameters using two different probing phases. For available bandwidth estimation, it uses 30 packets long probing train that consists of simple 800-bytes packet as single probing component. Again, for link capacity estimation, the scheme uses a complex probe packet structure for each 30 packets long probing train consisting of two ICMP timestamp request packets with a 1500B data packet ahead as single compound probe component. In order to implement these schemes together in a single tool, experiments were conducted by performing available bandwidth estimation before link capacity measurement and vice-versa. In both cases, the results seem to be similar to having no effect on one another as far as the measurement is concerned. But it is preferable to perform available bandwidth estimation in the first phase rather than link capacity estimation for reducing the affect of probing traffic on the legitimate traffic flow during network estimation.

Table 3.12 shows the measurement results for link bandwidth estimation where the link capacity measurement has been performed after the available bandwidth estimation. According to the following results, the link capacity results seem to be indifferent to the intra-compound-probe gap that should be defined by the measured available bandwidth in the first estimation phase. The accuracy of links' estimated bandwidth is similar to other link capacity estimations verified in the previous simulation results in Tables 3.7 and 3.8. In short, the accuracy of link capacity measurement is not dependent on the intra-compound-probe gap but rather on the JQR operation of the ICMP probes. Hence, no negative effect in measurement accuracy can be found in any of the cases in Table 3.12 where the compound probe rate is not agreeing with the available bandwidth rate of the source-destination path. In every cases, the output is accurate as desired.

Table 3.12 (a) Hop-by-hop Link Capacity Measurement Results with Compound Probe Rate Lower than the Actual Available Bandwidth

Link Capacity measurement with a compound probe rate lower than AB capacity								
Compound probe size:64+64+1500 Bytes, Iteration: 10								
Actual Link Capacity						**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	**Load at Link2** (Mbps)	**Actual AB** (Mbps)	**Comp. Probe Rate** (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	20	20	12	**8**	**4**	19.99	20	19.99
20	19	20	12	**7**	**4**	19.99	18.99	19.99
20	18	20	12	**6**	**4**	19.99	18	19.99
20	17	20	12	**5**	**4**	19.99	16.99	19.99
20	16	20	12	**4**	**4**	19.99	15.99	19.99
20	15	20	12	**3**	**4**	19.99	15	19.99
20	14	20	12	**2**	**4**	19.99	14	19.99
20	19	18	12	**6**	**4**	19.99	18.99	18
20	18	17	12	**5**	**4**	19.99	18	17
20	16	12	5	**7**	**4**	19.99	15.99	11.99
20	8	5	0.5	**4.5**	**4**	20	7.99	5
18	19	20	12	**6**	**4**	18	18.99	19.99
17	18	20	12	**5**	**4**	17	18	19.99
12	16	20	5	**7**	**4**	11.99	15.99	19.99
5	6	20	0.5	**4.5**	**4**	4.99	6	19.99

Table 3.12 (b) Hop-by-hop Link Capacity Measurement Results with Compound Probe Rate Higher than the Actual Available Bandwidth (Continued)

Link Capacity measurement with a compound probe rate higher than AB capacity								
Compound probe size: 64+64+1500 Bytes, Iteration: 10								
Actual Link Capacity						**Measured Link Capacity**		
Link1 (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)	**Load at** *Link2* (Mbps)	**Actual AB** (Mbps)	**Comp. Probe Rate** (Mbps)	*Link1* (Mbps)	*Link2* (Mbps)	*Link3* (Mbps)
20	20	20	12	**8**	**10**	19.99	20	19.99
20	19	20	12	**7**	**10**	19.99	18.99	19.99
20	18	20	12	**6**	**10**	19.99	18	19.99
20	17	20	12	**5**	**10**	19.99	16.99	19.99
20	16	20	12	**4**	**10**	19.99	15.99	19.99
20	15	20	12	**3**	**10**	19.99	15	19.99
20	14	20	12	**2**	**10**	19.99	14	19.99
20	19	18	12	**6**	**10**	19.99	18.99	18
20	18	17	12	**5**	**10**	19.99	18	17
20	16	12	5	**7**	**10**	19.99	15.99	11.99
20	8	5	0.5	**4.5**	**10**	20	7.99	5
18	19	20	12	**6**	**10**	18	18.99	19.99
17	18	20	12	**5**	**10**	17	18	19.99
12	16	20	5	**7**	**10**	11.99	15.99	19.99
5	8	20	0.5	**4.5**	**10**	5	7.99	19.99

The following table shows the cumulative rate of the legitimate data traffic (CBR) along the path during link capacity estimation with a compound probe rate beyond the current available bandwidth.

Table 3.13 Effect of Compound Probe Traffic over CBR Data Traffic along a Path when the Compound Probe Rate is Higher than Current Available Bandwidth

Link Capacity Measurement with a compound probe rate higher than AB capacity						
Compound probe size: 64+64+1500 Bytes, Iteration: 10						
Actual Link Capacity						**Measured CBR**
***Link1* (Mbps)**	***Link2* (Mbps)**	***Link3* (Mbps)**	**Load on *Link2* (Mbps)**	**Actual AB (Mbps)**	**Comp. Probe Rate (Mbps)**	**Load on *Link2* (Mbps) during Link Estimation**
20	20	20	12	**8**	**10**	12.024724
20	19	20	12	**7**	**10**	12.043223
20	18	20	12	**6**	**10**	12.075971
20	17	20	12	**5**	**10**	12.88868
20	16	20	12	**4**	**10**	12.171401
20	15	20	12	**3**	**10**	12.247375
20	14	20	12	**2**	**10**	12.219376
20	13	20	12	**1**	**10**	12.010212
20	12	20	12	**0**	**10**	11.636407

In this experiment, the measured rate of CBR data traffic is been calculated at the receiver's node for every 50 packets that the receiver node receives during the probing period. The last column in above table contains the cumulative CBR rate at the termination of every link bandwidth estimation. Here, the variation in the received constant data rate (CBR) provides a picture of the probe packets' effect in the legitimate traffic's rate during the measurement period. From the above results, the probe packets negatively affects the CBR traffic flow and such phenomenon, if not controlled, can have significant negative impact on TCP connections as they respond to flow and congestion control mechanisms that depend on the current network conditions. In order to avoid such occurrence over TCP flows, and eventual packet delays and packet drops in the intermediate queues, it is recommended to perform available bandwidth before link capacity estimation. It ensures optimal intra-compound-probe gap in the probing train for network friendly measurement at the later phase of the proposed scheme.

The following figure shows the schematic pseudocode of the combined available bandwidth and link capacity measurement scheme.

Pseudocode of available bandwidth and link bandwidth:

Iteratively probe the source-destination path using 30-packet probing train
    Measure available bandwidth

Compute intra-compound-probe gap from available bandwidth

Probe the first intermediate link along the source-destination path
    Measure the link capacity
Repeat link estimation until reaching the destination node of the path using compound probe train

Measure the narrow-link capacity

Figure 3.8 Pseudocode for combined available bandwidth and link capacity measurement scheme.

CHAPTER 4

CONCLUSION

Network measurement is a challenging job considering the current growth of networks and applications. Here, a new mechanism for available bandwidth and link capacity measurement has been proposed along a source-destination path. This scheme includes a different rate adjustment algorithm for available bandwidth estimation after careful investigation of other measurement schemes. Some initial experiments of various measurement associated parameters such as probe train length, probing load, and data flow in ns2 simulator are presented. Moreover, a new link capacity estimation technique has been proposed with a new designed probe packet structure, which has been defined as compound probe that has the potential to measure both hop-by-hop link bandwidths and narrow-link capacity regardless of the cross traffic load on the links. A relation between redundant data packet length of the compound probe structure for measuring highest possible link capacity estimation is also shown in this work. Finally, these two schemes have been implemented in ns2 with acceptable results. But still, it requires further validation in real network scenarios. Verification of the proposed scheme using similar experiments in Internet and calculation of computational complexity for the proposed algorithms would be the next steps for further schematic validation and accuracy comparison with the existing measurement tools.

REFERENCES

[1] R.S. Prasad, M. Murray, C. Dovrolis and K. Claffy, "Bandwidth estimation: Metrics, measurement techniques, and tools," *IEEE Network*, Nov.-Dec. 2003, vol. 17, pp. 27-35.

[2] N. Hu, and P. Steenkiste, "Evaluation and characterization of available bandwidth probing techniques," *IEEE JSAC Special Issue in Internet and WWW Measurement, Mapping, and Modeling*, vol. 21, no. 6, Aug. 2003.

[3] M. Jain and C. Dovrolis, "End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput," *in Proc. ACM SIGCOMM Symp.*, vol. 11, no. 4, Aug. 2003, pp. 295-308.

[4] ---, "Pathload: A measurement tool for end-to-end available bandwidth," *in Proc. Passive and Active Measurement*, Mar. 2002, pp. 14-25.

[5] C. Dovrolis, P. Ramanathan, and D. Moore, "What do packet dispersion techniques measure?" *in Proc. of IEEE INFOCOM*, Apr. 2001, pp. 905-914.

[6] J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," *in Proc. ACM Internet Measurement Conference*, Oct. 2003.

[7] B. Melander, M. Bjorkmann, and P. Gunningberg, "A new end-to-end probing and analysis method for estimating bandwidth bottlenecks," *in Proc. IEEE GLOBECOM-Global Internet Symp.*, Nov. 2000, pp. 415-420.

[8] R. Carter and M. Corvella, "Measuring bottleneck link speed in packet-switched networks," *in Proc. Performance Evaluation*, vol. 27-28, Mar. 1996, pp. 297-318.

[9] V. Jacobson, "Pathchar – A tool to infer characteristics of Internet paths," *ftp://ftp.ee.lbl.gov/pathchar/*, Apr. 1997.

[10] A. B. Downey, "Using Pathchar to estimate Internet link characteristics," *in Proc. ACM SIGCOMM*, Aug. 1999, pp. 241-250.

[11] K.G. Anagnostakis, M. Greenwald, and R. S. Ryger, "cing: Measuring network-internal delays using only existing infrastructure," *in Proc. IEEE INFOCOM*, Mar.-Apr. 2003, pp. 2112-2121.

[12] K. G. Anagnostakis, and M. B. Greenwald, "Direct measurement vs. indirect inference for determining network internal delays," *in Proc. Performance Evaluation*, vol. 49, Sept. 2002, pp. 165-177.

[13] A. Pasztor, and D. Veitch, "Active probing using packet quartets," *in Proc. ACM Internet Measurement Conference*, Nov. 2002, pp. 293-305.

[14] ---, "The packet size dependence of packet pair like methods," *in Proc. EEE/IFIP Int. Workshop Quality of Service*, May 2002, pp. 204-213.

[15] V. Paxson, "Measurement and analysis of end-to-end Internet dynamics," PhD dissertation, Comput. Sci. Div., U.C. Berkeley, Berkeley, CA, May 1999.

[16] J. C. Bolot, "End-to-end packet delay and loss in the Internet," *in Proc. ACM SIGCOMM Symp. Communications Architectures Protocols*, Sept. 1993, pp. 289-298.

[17] N. Hu, L. Li, Z. M. Mao, P. Steenkiste, and J. Wang, "Locating Internet bottlenecks: Algorithms, measurements, and implications," *in Proc. SIGCOMM*, Aug.-Sept. 2004.

[18] A. Johnsson, "On the comparison of packet pair and packet train measurements," in Proc. *Swedish National Computer Networking Workshop*, Sept. 2003.

[19] B. Choi, S. Moon, Z. Zhang, K. Papagiannaki, and C. Doit, "Analysis of point-to-point packet delay in an operational network," *in Proc. IEEE INFOCOM*, Mar. 2004.

[20] P. E. Black, ed., U.S. National Institute of Standards and Technology, "Binary search," *http://www.nist.gov/dads/HTML/binarySearch.html*, May 14, 2007.

[21] W.H. Hesselink, "Ternary search," Dept. of Mathematics and Computing Science, Rijksuniversiteit Groningen, Netherlands, Jul. 2, 2003.

[22] "The Network Simulator - ns-2," *http://www.isi.edu/nsnam/ns*, Jul. 10, 2007.

[23] A. Shriram, "Analysis of Available Bandwidth Measurement Techniques," *http://www.cs.unc.edu/~jasleen/research/ab-estimation/*, Sept. 2007.

[24] T. Oetiker and D. Rand, "MRTG: Multi router traffic grapher," *http://people.ee.ethz.ch/~oetiker/webtools/mrtg/*, Jun. 2007.