Schemes to Measure Available Bandwidth and Link Capacity with Ternary Search and Compound Probe for Packet Networks

Khondaker M. Salehin and Roberto Rojas-Cessa
Networking Research Laboratory
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
Newark, New Jersey 07102–1982
Email: {kms29, rojas}@njit.edu

***Abstract*—Accurate measurement of network parameters such as available bandwidth (ABW) and link capacity are needed for analyzing network performance. Active measurement is an attractive approach as it has the advantage of controllability and flexibility for performing network measurement. However, it can affect both the data traffic and the measurement process itself, affecting the accuracy of the measurement if significant amount of probe traffic is injected into the network. Furthermore, measurement must be completed in short time to effectively monitor the network state. In this paper, we propose two measurement schemes: one for measuring ABW, and the other for measuring per-hop link capacities of an end-to-end path. The ABW proposed scheme performs measurement in short period of time and with small amount of probe traffic and it achieves accuracy comparable to that of IGI and Pathload. The proposed link-capacity scheme provides immunity to cross traffic. We present ns-2 simulation results of the ABW and link-capacity measurement schemes to show their performance.**

***Index Terms*—Link capacity, available bandwidth, Network measurement, active network measurement.**

I. INTRODUCTION

Accurate measurement of network parameter supports traffic engineering to ensure efficient use of network resources. End-to-end available-bandwidth (ABW) and link-capacity are two important parameters to determine network condition for service providers to efficiently manage resources and services. In brief, ABW is the difference between the link-capacity and the traffic load over a path at an instance of time, whereas (physical) link-capacity is the maximum data transfer capability of a link at every instance of time. Several ABW and link-capacity measurement schemes have been proposed [1]-[11]. All these measurement schemes can be coarsely divided into passive and active methods. Passive measurement methods use flowing data traffic through a measuring node to estimate network characteristics. Here, measurement is only possible with the administrative control and the existence of network traffic on the link (or node) under measurement interest. Multi Router Traffic Grapher (MRTG) [1] is an example of a passive measurement tool. On the other hand, active measurement methods [2]-[11] proactively sends probe packets from a source node towards a destination node to estimate various network parameters. In these methods, different probe packet characteristics, such as probe size, number of probes, and inter-probe gaps, etc. determine the features of the measurement process. However, active measurement schemes add extra traffic into the network, which may affect legitimate data traffic and hurt measurement accuracy during ABW measurement [4]. In case of link-capacity measurement, congested network may decrease the accuracy of the measurement process [10],[11].

In this paper, we propose two schemes, one for ABW measurement and the other for link-capacity measurement. The proposed ABW measurement scheme uses a ternary search algorithm [12], which converges in a short measurement time and with a small number of trails. The link-capacity measurement scheme uses a probe structure, called compound probe, that shows immunity to heavy load of cross traffic. This paper presents a performance study of both schemes in terms of probe load, and measurement accuracy.

The remainder of this paper is organized as follows. After presenting related work in Section II, we introduce the algorithm of the proposed schemes to measure ABW and link capacity in Section III. We present our simulation results in Section IV. We present our conclusions in Section V .

II. RELATED WORKS

Controllability of the active measurement process (e.g., rate and size of probe packets) makes it an attractive approach. Active measurement schemes send probe packets either as pairs or probe trains on a source-destination path and check the modification that the probe pairs or trains undergo to collect link-state information. These schemes can be further categorized into either probe-gap or probe-rate models. Probe-gap model schemes measure gap dispersion (i,e., change) value of a probe packet pair at the destination whereas probe-rate model schemes compare the transmission rates of probe packets at source and destination for measurement. We discuss some of the most popular schemes to measure ABW and link capacity and their probe models.

- IGI is an ABW measurement scheme based on probe-gap model [4]. It measures ABW of a path by sending out probe packets and determining the dispersion of probe packet pair gaps due to the interference of cross-traffic available on the path during measurement. It ensures proportional dispersion of probe packets with respect to cross-traffic by ensuring joint queuing region (JQR) operation of the probe packet pairs on the smallest link-capacity of the path. It is called JQR operation when the first packet of a packet pair is available in the output queue of a node while the second packet arrives in the same queue. Note that IGI actually measures cross-traffic through gap model scheme and the ABW value is obtained by deducting this value from the smallest link-capacity of the path.
- Pathload is an ABW measurement scheme based on the probe-rate model [7]. It iteratively sends out probe packets towards destination to determine the maximum probing rate over the measurement path until the sending rate at source and the receiving rate at destination are equal. Pathload uses the binary search algorithm [13] that adjusts the probing rate at each iteration. During the probing process, it compares the one way delays of the probe packets at destination instead of their inter-probe gaps. Pathload shows that probe packets have non-increasing delays when the probing rates at source and destination are less than the ABW of the measurement path.
- Bprobe is a link-capacity measurement scheme based on the probe-gap model [10]. It sends out a train of Internet Control Messaging Protocol (ICMP) ECHO packets to destination node and waits for the replies to measure the inter-arrival times of consecutive probe packets. Because the dispersion between a probe pair is inversely proportional to the smallest link-capacity of a path, probe pair dispersion with JQR operation with no intervening cross traffic can represent the actual link capacity. Because the probe packets are is not guaranteed to operate in JQR, it is vulnerable to cross-traffic. Bprobe applies either an intersection-set or a union-set operation to filter out measurement errors in order to determine the representative output gap between the probes. The link capacity is measured from dividing the probe packet size by the statistically processed dispersion value.
- Pathrate is a link-capacity measurement scheme based on the probe-gap model [11]. It measures link-capacity through two probing phases using User Datagram Protocol (UDP) packets to estimate the representative dispersion value of the probe packets over the measurement path. The first phase provides some candidate capacity values, and the second phase provides the minimum capacity of the measurement link. Based on this information, the final link capacity is estimated through statistical processing of the first and second phase data, which is considered free of both cross traffic and probe-packet related errors.

The existing ABW and link-capacity measurement schemes may induce a significant amount of probe load and might require long measurement time. The link-capacity measurement schemes also use complex statistical data processing techniques to avoid congestion and probe-packet related errors during the measurement. Still, the performance of existing link-capacity measurement schemes are poor under heavy cross traffic [10],[11].

III. PROPOSED MEASUREMENT SCHEMES

A. ABW Measurement Scheme

The proposed ABW measurement scheme is based on probe-rate model which iteratively sends probe trains (i.e., sequence of probe packets) at specific rate over the end-to-end path and estimates changes in the of probe train's transmission time at the destination to estimate the ABW capacity. Here, the transmission time of a probe train at the destination is calculated from the time difference of the first and last probe packets. Change in transmission time of the probe train is affected by the smallest ABW of the path during the measurement period. If the probe transmission rate is lower than or equal to the ABW, the probe train does not experience changes in its transmission time. We adopted the ternary search algorithm in the proposed scheme to find the smallest ABW. This search algorithm sends two probe trains with two different transmission rates, during each iteration. These two probing rate values are one third and two thirds of the difference between the minimum and maximum probe rate limits of the ternary search algorithm. This search algorithm then speeds up the search time.

1) Measurement Algorithm: Figure 1 shows a flow chart that describes the ABW measurement scheme. In this scheme, the source node shoots two probe trains towards the destination at two different transmission rates, i.e., *g-rate* and *h-rate*. The source node computes the cumulative gaps of the trains at the destination, i.e., *g_outgap* and *h_outgap*, upon receiving the replies of the probes, and compares them with the initial cumulative gap at the source, i.e., *g_ingap* and *h_ingap*, respectively. The source updates the probing rates in the following iterations based on the change of *g_increment* (i.e., *g_increment = g_outgap - g_ingap*) and *h_increment* (i.e., *h_outgap - h_ingap*). The measurement process terminates by providing an ABW value when either the *g_increment* or *h_increment* rate has a change between 1% and 1.5% or when the *g-rate* and *h-rate* difference is smaller than or equal to 1.5 Mb/s. Here the selected minimum measurable ABW is 0.5 Mb/s. This minimum is selected according to the complexity to comply with the termination condition of the search scheme and in function of the expected link capacity. We chose to adopt 2.5% of the link capacity (e.g. 20 Mbp/s) for the considered simulations examples. The maximum value is defined by the smallest ABW of the path.

B. Link-capacity Measurement Scheme

The proposed link-capacity measurement scheme uses a probe train of compound probes separated by an inter-

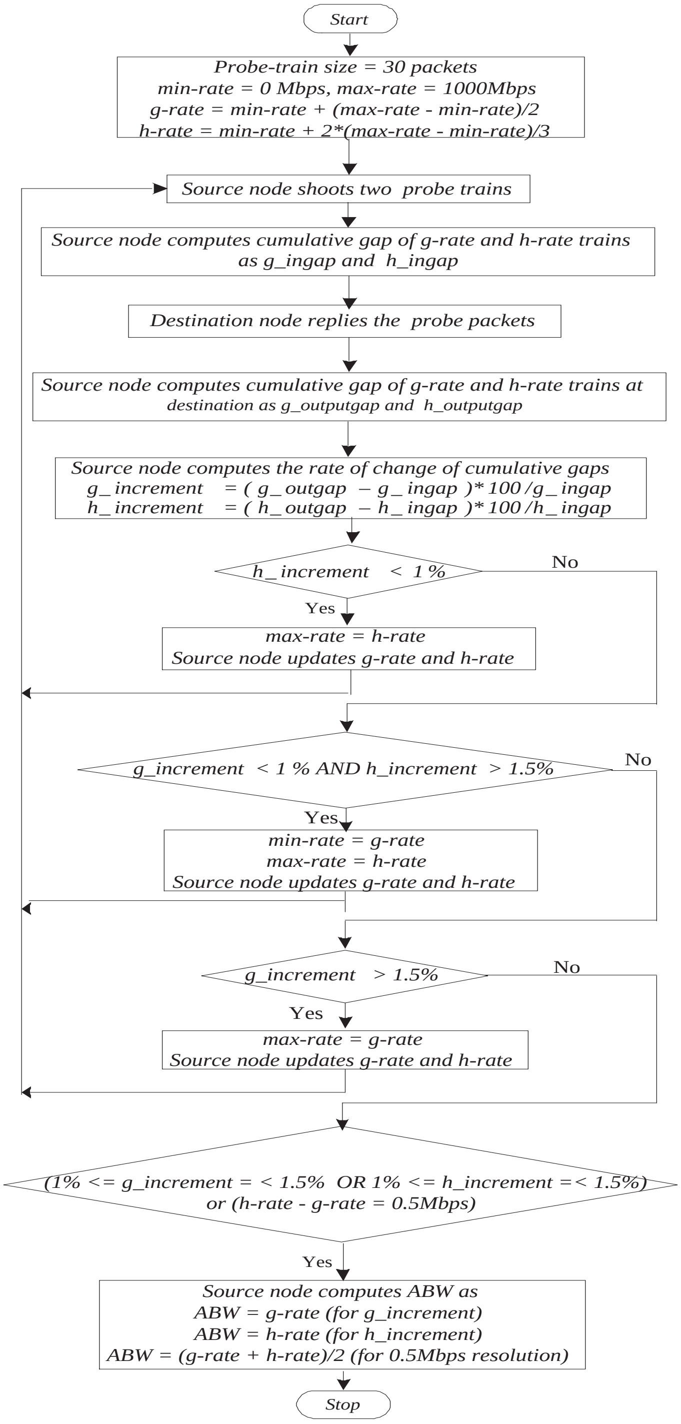


Fig. 1. Proposed ABW measurement algorithm.

compound-probe gap, which is determined by the ABW of the source-destination path. The proposed probe structure avoids dependencies on complex statistical processing of data in link-capacity measurement by ensuring JQR operation of its probe packets. The proposed scheme measures link capacity in a hop-by-hop manner and estimates the smallest link-capacity of the path from the measured per-hop link capacities. Figure 2 shows the structure of the proposed compound probe.

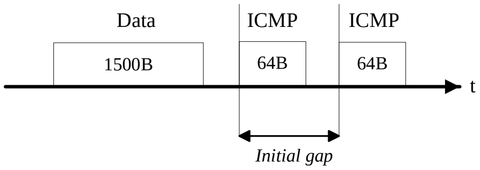


Fig. 2. Structure of a compound probe.

1) Compound Probe: Figure 2 shows a compound probe, which consists of a 1500B data (UDP) packet followed by two 64B ICMP packets. The *initial gap* (i.e., intra-probe gap between Data and ICMP or two ICMP packets) value is equivalent to the average packet processing time of network nodes (e.g., routers usually require 40 microseconds of packet processing time [15]), which is considered a constant parameter [4],[14]. With small *initial gap* value and 1500B data packet, the compound probe ensures that the two 64B packets operate in JQR at every intermediate node along a source-destination path. The *initial gap* dispersion in the compound probe occurs at every node, and it is proportional to the node's input link capacity.

2) Measurement Algorithm: Figure 3 shows that the source node sends a 30-packet train of compound probes with 40 microseconds of initial gaps toward each intermediate node until it reaches the destination node. For link-capacity measurement, the source node uses the ICMP timestamp replies to compute the *initial gap* dispersion between the ICMP packets over the input link of the destination node (i.e., *outputgap*). The minimum *outputgap* of the 30 compound probe packets is used to measure link-capacity of each hop. Unlike existing schemes, the proposed scheme does not require rigorous statistical processing to eliminate measurement errors due to the robustness of compound probe structure.

3) Immunity to Heavy Loads of Cross Traffic: We discuss the immunity to cross traffic issue of the proposed scheme as below.

- It has been claimed that cross-traffic can hardly interfere a probe pair when it has an inter-probe gap in the order of tens of microseconds. Thus, 40 microseconds of *initial gap* in the compound probe has a small probability to catch cross-traffic packet(s).
- The proposed scheme samples a link with multiple compound probes for measurement. The capacity calculation is performed by taking the smallest dispersion value (i.e., *min_outputgap*) at the destination node.

IV. SIMULATION RESULTS

The proposed schemes along with IGI and Pathload were simulated in the ns2 network simulator [16] using cross traffic, which was modeled as constant bit rate (CBR) traffic. Figure 4.(a) shows the single-hop scenario used in our simulation. The proposed schemes were also tested in two multiple-hop scenarios as shown in Figures 4.(b) and 4.(c). Previous works

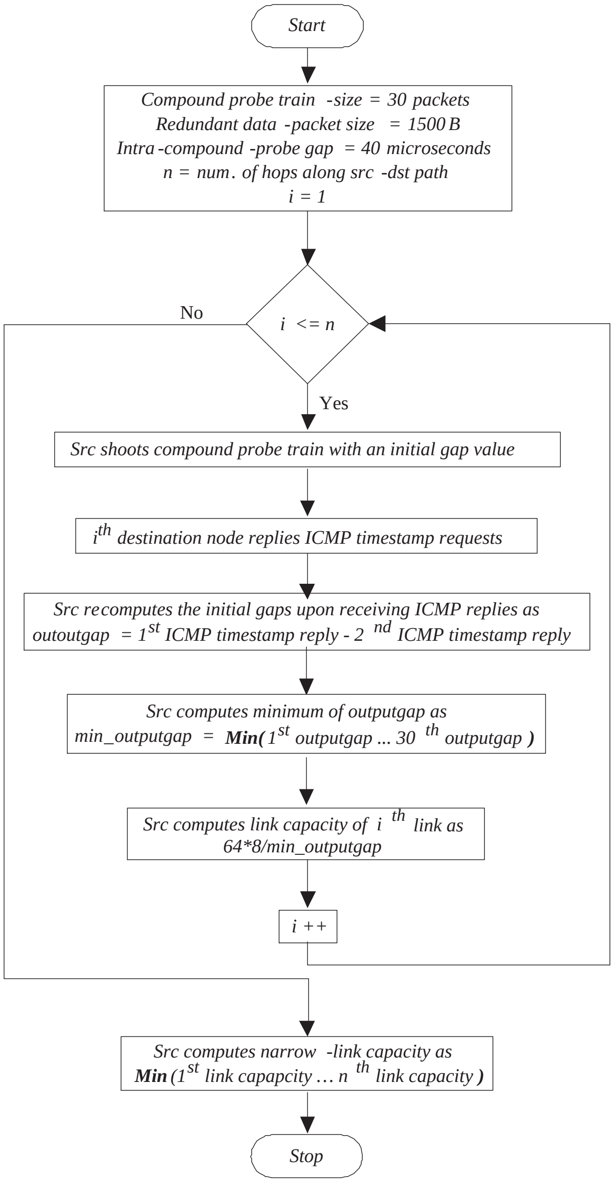
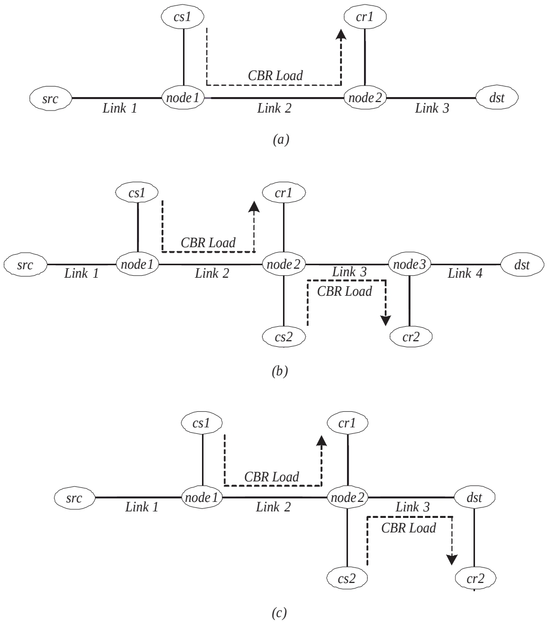


Fig. 3. Proposed link-capacity measurement algorithm.

[17]-[19] showed that the infrequent change of traffic load (i.e., constant rate of traffic load) over Internet paths for a long interval makes network measurement possible over long periods of time. Moreover, in [20], IGI and Pathload implementations adopt the CBR traffic model. We follow this practice in this paper.

A. ABW Results

1) Accuracy of ABW Measurement: Table I shows the ABW measurement results of the proposed scheme for the multiple-hop scenario in Figure 4.(b) with different combinations of cross-traffic load values. For the proposed scheme, the error[1] is below 10%, even under high load condition.



src = Measurement source
cs1 = Cross-traffic1 source
cs2 = Cross-traffic2 source
dst = Measurement destination
cr1 = Cross-traffic1 destination
cr2 = Cross-traffic2 destination

Fig. 4. (a) Single-hop topology, (b) and (c) multiple-hop topologies.

TABLE I
ABW MEASUREMENT USING A MULTIPLE-HOP TOPOLOGY.

Ternary search, Topology: 6.(b), Probe packet size: 800B, Train size: 30, Iterations: 25

Link (1,2,3,4) (Mb/s)	Load (Link2) (Mb/s)	Load (Link3) (Mb/s)	Actual ABW (Mb/s)	Measured ABW (Mb/s)
20	1	1	19	17.0
20	1	2	18	16.5
20	2	3	17	16.0
20	4	3	16	15.0
20	5	3	15	14.0
20	6	4	14	13.0
20	5	7	13	13.0
20	10	3	10	10.0
20	11	5	9	9.0
20	13	12	7	7.0
20	14	10	6	6.0
20	15	4	5	5.0
20	16	10	4	4.0

Figure 5 shows an accuracy comparison graph of IGI, Pathload, and the proposed scheme on the single-hop scenario in Figure 4.(a) with 200 Mb/s of link capacity for each link. The graph contains average values from 25 measurements for each ABW point. The simulation results for the proposed scheme show an error equal to or smaller than 3% in every case. Even though IGI and Pathload have higher accuracy (i.e., smaller error) than the proposed scheme, the accuracy of the proposed scheme is comparable to that of IGI and Pathload.

2) Measurement Time and Probe Load: We investigated the measurement time and injected load of the proposed scheme and compared them to those of IGI, which is considered as

---

[1] Error is the percentage error in reference to the actual value, which is calculated as the ratio of the difference between the actual value and the measured value, and the actual value.

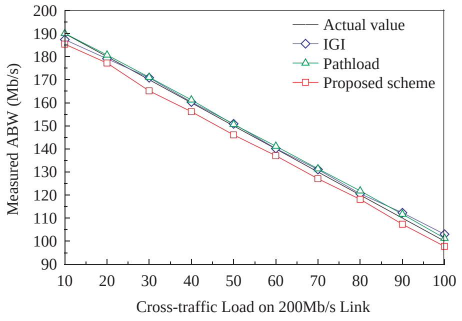


Fig. 5. Accuracy comparison of different ABW measurement schemes.

the fastest existing ABW measurement scheme [4]. For a fair comparison, both schemes used 60 800B probe packets per train and the simulation is ended after one measurement of similar termination conditions (i.e., IGI and the proposed scheme terminate with 1% and from 1% to 1.5% variations, respectively). According to Figure 6, the average measurement time of the proposed scheme is 1.36 seconds with a standard deviation of 0.35 seconds. On the other hand, IGI has an average measurement time of 1.55 seconds and standard deviation of 0.32 seconds.

Figure 6.(b) shows the comparison of the number of measurement probes used by the proposed scheme and IGI in the same experiment. The graph shows that the proposed scheme generates a smaller number of probe packets than IGI to achieve comparable accuracy. In fact, the proposed scheme requires at least 984 fewer probe packets than IGI for every case. For example, the maximum difference is 1464 fewer probe packets for 60Mb/s of ABW scenario by the proposed scheme.

Figure 6.(c) shows the number of iterations required for different amount of ABW measurements in the proposed scheme. This figure can be used to describe the fluctuation in the measurement time and probe load of the proposed scheme as shown in Figures 6.(a) and 6.(b). According to Figure 1, two probe trains with two different probing rates are sent during each iteration before obtaining the ABW of the measurement path. So, if the proposed scheme requires 7 iterations (e.g, ABW 80Mb/s) instead of 6 iterations (e.g., ABW 70Mb/s) for the ABW measurement, additional time and number of probe packets associated with the two probe trains are also added to the measurement time and the probe count, respectively. As the number of iterations required for ABW measurements varies with different ABW capacity of the path, the fluctuations in measurement time and probe count in Figures 6.(a) and 6.(b) are expected.

B. Link-capacity Results

1) Accuracy of Link-capacity Measurement: Table II shows the results for the multiple-hop scenario in Figure 4.(c). The results show that the proposed scheme has high accuracy for all links (e.g., 15 Mb/s and 150 Mb/s, respectively) under high

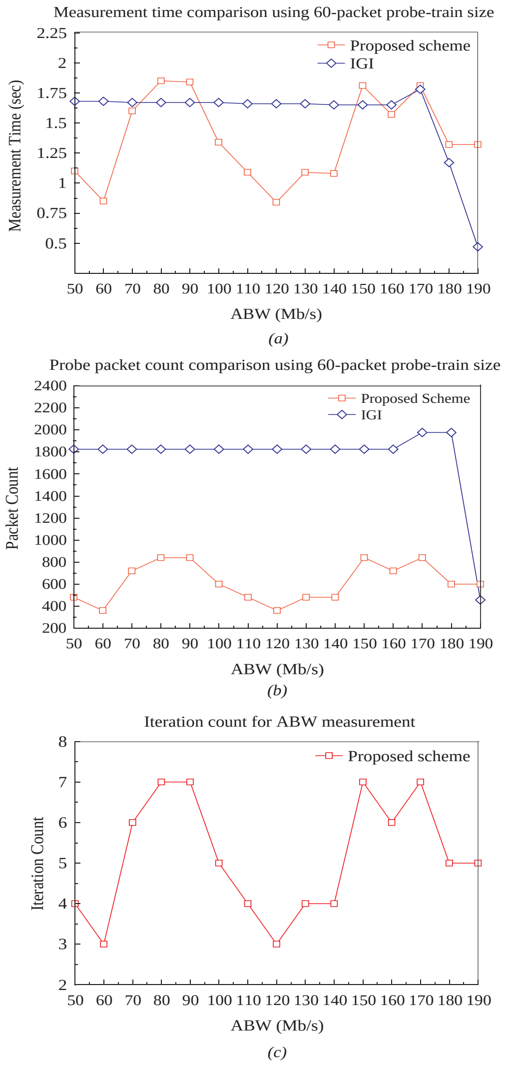


Fig. 6. (a) Measurement time of the proposed scheme and IGI, (b) probe load of the proposed scheme and IGI, and (c) iteration count of the proposed scheme.

cross-traffic loads (e.g., 18 Mb/s over a 20 Mb/s link). In every case, the measurement error is smaller than 1%. The measurement results with 16Mb/s and 18Mb/s load scenarios also show that the amount of cross-traffic does not affect the capacity measurement processes.

2) Scalability of Measurement: The proposed scheme has an upper limit for capacity measurement determined by the maximum packet processing time of 40 microseconds that has been considered as *initial gap* in the compound probe.

A single-link topology was used to determine dependency on the redundant data-packet. Table III shows the measurement results obtained by using compound probes with the 1500B redundant packet with and without cross-traffic loads. In both load conditions, the measurement process shows low accuracy for link capacities beyond 300 Mb/s. This limitation value follows the theoretical upper measurement limit determined by the use of a 1500B redundant packet and an intra-probe

TABLE II

HOP-BY-HOP LINK-CAPACITY MEASUREMENT USING A MULTIPLE-HOP TOPOLOGY.

Topology: 6.(c), Compound probe size: (64+64+1500)B, Iterations: 10									
Actual Capacity (Mb/s)			Load (Mb/s)	Measured Link Capacity (Mb/s)			Error (%)		
Link1	Link2	Link3	Link2+Link3	Link1	Link2	Link3	Link1	Link2	Link3
20	15	20	10+10	19.99	14.99	20.00	0.050	0.066	0
20	15	20	12+12	19.99	14.99	20.00	0.050	0.066	0
20	15	20	14+14	20.00	15.00	20.00	0	0	0
20	20	20	10+10	19.99	19.99	20.00	0.050	0.050	0
20	20	20	12+12	19.99	20.00	20.00	0.050	0	0
20	20	20	14+14	19.99	20.00	20.00	0.050	0	0
20	20	20	16+16	20.00	20.00	20.00	0	0	0
20	20	20	18+18	20.00	20.00	20.00	0	0	0
20	150	20	6+6	19.99	149.99	20.00	0.050	0.006	0
20	150	20	8+8	19.99	149.99	20.00	0.050	0.006	0
20	150	20	10+10	19.99	149.99	20.00	0.050	0.006	0
20	150	20	12+12	19.99	149.99	20.00	0.050	0.006	0

TABLE III

MAXIMUM MEASURABLE LINK CAPACITY UNDER DIFFERENT LOAD CONDITIONS.

Link-capacity measurement limitation test with 1500B redundant data packet						
No-load Scenario				Load Scenario		
Actual value (Mb/s)	Measured value (Mb/s)	Error (%)		Actual value Mb/s	Measured value Mb/s	Error (%)
200	200	0		200	200	0
295	294.99	0.003		295	294.99	0.003
296	295.99	0.003		296	295.99	0.003
297	297	0		297	297	0
298	297.99	0.003		298	297.99	0.003
299	299	0		299	299	0
299.5	299.49	0.003		299.5	299.49	0.003
300	299.99	0.003		300	299.99	0.003
300.5	289.2	3.760		300.5	293.72	2.250
301	279.18	7.249		301	294.45	2.176
302	261.18	13.516		302	285.67	5.407
303	245.46	18.990		303	279.98	7.597
304	231.61	23.812		304	267.8	11.907
305	219.32	28.809		305	262.16	14.045

gaps of 40 microseconds. Therefore, the upper limit of the link-capacity measurement, called the Maximum Measurable Capacity (MMC), for the proposed scheme with a probe packet size P is:

$$MMC = \frac{P(bits)}{Initial\ gap\ (sec)} = \frac{1500B * 8}{0.00004\ sec} = 300Mb/s \tag{1}$$

V. CONCLUSIONS

In this paper, we have proposed a scheme for the measurement of ABW and another for the measurement of link capacity. The proposed ABW measurement scheme adopts a ternary search algorithm for ABW measurement. This scheme was simulated in ns2 and compared to IGI and Pathload. The accuracy of the proposed scheme is comparable to that of the other two existing schemes.

The link-capacity measurement scheme uses a probe-packet structure, named compound probe, which consists of three probe packets with inter-gaps equivalent to the average packet processing time of network nodes. This probe structure allows the accurate measurement of per-hop link capacities, regardless of the cross-traffic condition of the end-to-end path. The proposed schemes show high accuracy for measuring link capacity and ABW with a short convergence time and low probe load. The two schemes can be used to evaluate link states in single- and multiple-hop paths for packet networks.

REFERENCES

[1] Tobi Oetiker's MRTG - The Multi Router Traffic Grapher. [Available Online]. http://people.ee.ethz.ch/ oetiker/webtools/mrtg, accessed: December, 2007.

[2] V. Jacobson, "Pathchar - A tool to infer characteristics of Internet paths". [Available online]. ftp://ftp.ee.lbl.gov/pathchar/msri-talk.pdf, accessed July, 2007.

[3] A. Downey, "Using Pathchar to estimate Internet link characteristics," in Proceedings of ACM SIGCOMM, Cambridge, USA, August 1999, pp. 241–250.

[4] 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, 2003, 21, (6), pp. 879–864.

[5] B. Melander, M. Bjorkmann, and P. Gunningberg, "A new end-to-end probing and analysis method for estimating bandwidth bottlenecks," in Proceedings of IEEE GLOBECOMM - Global Internet Symp., San Francisco, USA, November 2002, pp. 415–420.

[6] J. Strauss, D. Katabi, and J. Kaashoek, "A measurement study of available bandwidth estimation tool," in Proceedings of ACM SIGCOMM Internet Measurement Conference, Miami, USA, October 2003, pp. 39–44.

[7] M. Jain, and C. Dovrolis, "Pathload: A measurement tool for end-to-end available bandwidth," in Proceedings of Passive and Active Measurement, Colorado, USA, March 2002, pp. 14-25.

[8] —, "End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput," Computer Communication Review, 2002, 32, (4), pp. 295–308.

[9] N. Hu, L. Li, Z. Mao, P. Steekiste, and J. Wang, "Locating Internet bottlenecks: Algorithms, measurements, and implications," Computer Communication Review, 2004, 34, (4), pp. 41–54.

[10] R. Carter, and M. Crovella, "Measuring bottleneck link speed in packet-switched networks," Performance Evaluation, 1996, 27-28, pp. 297–318.

[11] C. Dovrolis, R. Ramanathan, and D. Moore, "What do packet dispersion techniques measure?," in Proceedings of IEEE INFOCOM, Anchorage, USA, 2001, pp. 905–914.

[12] W. H. Hesselink, "Ternary Search." [Available online]. http://www.cs.rug.nl/ wim/pub/whh303.pdf, accessed: July, 2007.

[13] Binary Search. [Available online]. http://www.nist.gov/dads/HTML/ binarySearch.html, accessed: July, 2007.

[14] B. De Schepper, B. Steyart, S. Wittevrongel, M. Moeneclaey, and H. Bruneel, "Constant hardware delay in integrated switching elements with multiserver output queues," IEE Proc. Communications, 2006, 153, pp. 664–670.

[15] K. Papagiannaki, S. Moon, C. Farleigh, P. Thiran, and C. Diot, "Measurement and analysis of single-hop delay on an IP backbone network," IEEE JSAC, 21(6):908–921, 2003.

[16] The network simulator-ns2. [Available online]. http://www.isi.edu/nsnam/ns, accessed: July, 2007.

[17] H. Balakrishnan, S. Seshan, M. Stemm, and R.H. Katz, "Analyzing stability in wire-area network performance," in Proceeding of ACM SIGMETRICS, June 1997, pp. 2–12.

[18] V. Paxson, "End-to-end Internet packet dynamics," IEEE/ACM Transaction on Networking, June 1999, pp. 277–292.

[19] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker, "On the consistency of Internet path properties," in Proceedings of ACM SIGCOMM Internet Measurement Workshop, November 2001, pp. 197–211.

[20] [Online]. Available: http://www.cs.unc.edu/ jasleen/research/ab-estimation, accessed September, 2007.

[21] A. Pasztor, and D. Veitch, "The packet size dependence of packet pair like methods," in Proceedings of IEEE/IFIP International Workshop Quality of Service, Miami, USA, May 2002, 204–213.