Scheme for Measuring Queueing Delay of a Router using Probe-Gap Model: The Single-Hop Case

Khondaker M. Salehin and Roberto Rojas-Cessa

***Abstract*—In this letter, we propose a probe-gap scheme to measure queueing delay of a router in a single-hop path. The proposed scheme is based on estimating the change of an intra-probe gap between a pair of probing packets after they traverse the router under test. The proposed scheme can be applied to online routers and is sensitive to small queueing delays. Our simulation results show that this scheme consistently achieves high measurement accuracy under different levels of queueing delays.**

***Index Terms*—Queueing delay, active measurement, packet pair, intra-probe gap, single hop, clock synchronization.**

I. Introduction

The delay a packet experiences in a router consists of 1) processing delay, 2) queueing delay, and 3) transmission delay. The processing delay is the time a router takes to forward a packet from its input queue to its output queue, including destination lookup (IP lookup [1]), packet-header updates, etc. The queueing delay is the time that a packet waits in a router after arriving at the router queue until it starts transmission through to the output link. Queueing delay is a dynamic network parameter which is proportional to the number (or transmission delays) of packets that are already waiting in the output queue at a given instance of time.

Figure 1 shows examples of two different queueing scenarios at a router, node i. Node i has multiple input links (L_1, ..., L_n) and one output link (L_o), with link capacities l_1, ..., l_o, respectively. The processing time of node i to forward a packet from the input queues (Q_1, ..., Q_n) to the output queue (Q_o) is constant. Figure 1(a) shows that node i has only one traffic flow between L_n and L_o (F_n), where the flow rate is smaller than or equal to l_n; therefore, packet P_1 from F_n does not experience additional waiting time at Q_o before transmission through L_o. In Figure 1(b), node i carries two traffic flows, one between L_1 and L_o (F_1) and the other between L_n and L_o (F_n), where the combined traffic rate of F_1 and F_n is larger than l_o. In this scenario, Q_o receives P_1 from F_n while packet C_1 from F_1 is about to be transmitted through L_o; therefore, P_1 will experience a queueing delay equivalent to the transmission time of C_1 over L_o, or $\frac{s_c}{l_o}$, where s_c is the size of C_1.

Measurement of queueing delay of routers is key for characterizing the level of congestion in a network [2]. Precise knowledge of queueing delay allows a service provider to determine and maintain Quality-of-service (QoS) and service-level agreements, respectively, for various delay-sensitive applications in the network. For example, accurate one-way delay (OWD) measurements of end-to-end paths is necessary to attain high accuracy in triangulation-based IP geolocation [3], [4]. Because queueing delays increase OWD of a path (e.g., 10 μs of queueing delay at a router may vary the estimated geographic distance by 2 km, considering the end-to-end delay vs. physical distance mapping over a optical-fiber link [4]), it has the potential to degrade the geolocation accuracy.

The authors are with Networking Research Laboratory, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA. Email: {kms29, rojas}@njit.edu.

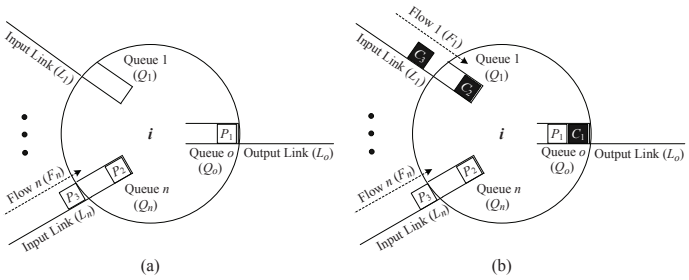


Fig. 1. Queue length at the output queue, Q_o, of a router, node i, in presence of (a) one traffic flow and (b) two traffic flows at different instances of time.

Queueing-delay measurement is complex because it may require Internet control message protocol (ICMP) packets [5], [6], infrastructural support from the Internet [6], [7], clock synchronization [8], and complex processing of the sampled data set [7], [9] to obtain high accuracy. In this letter, we propose an active scheme to measure queueing delay of a router using a pair of user datagram protocol (UDP) packets that does not require infrastructural support or clock synchronization. The scheme is simple and it measures queueing delay by determining the change in the intra-probe gap of packet pairs (i.e., the separation between the last bits of the packets in each pair) at the router under test. The scheme is designed to measure queueing delay in a single-hop scenario. It is sensitive to small queueing delays and can be applied to online routers.

II. Related Work

Existing schemes for measuring queueing delay can be divided into two methods: a) active and b) passive. Active schemes proactively inject probing packets in the path under test for measuring queueing delay of an end-to-end path, e.g., Pathchar [5], cing [6], network tomography based schemes [7], [10], etc. Pathchar measures link capacity and queueing delay of an end-to-end path using linear-regression model on the round-trip times (RTTs) of ICMP packets. It is considered network intrusive (i.e., the probing packets themselves can affect the existing traffic flow(s) in the path under test) and requires long measurement time [11]. cing measures queueing delay using pairs of ICMP packets. Even though this scheme is simple, its feasibility of deployment and measurement accuracy heavily depends on collecting timestamps for every

probing packet from the routers in the path under test [6]. Network-tomography based schemes are complex and they use unicast or multicast packets for measuring queueing delay. These schemes require large infrastructural support from networks as they can measure queueing delay only if multiple end-to-end paths share links amongst them [7], [10].

Passive schemes do not generate probing packets, they rather utilize the existing data traffic in the router under test for queueing-delay measurement, e.g., instrumentation-based [8], [12], [13] and analysis-based [9] schemes. Instrumentation-based schemes measure queueing delay of a router by instrumenting its input and output links using specialized packet-capture cards [8] or kernel processes [12], [13], and a packet monitor [8], [12]. These schemes may require a reliable timing source (e.g., global-positioning system) and synchronization between the packet-capture cards [8]. The analysis-based scheme applies a stochastic model on the captured data traffic for characterizing the queueing delay of a router. However, this scheme is complex and requires prior knowledge of the traffic load, number of traffic flows, packet loss, etc., of the router under test [9].

III. Proposed Scheme

The proposed scheme uses a packet-pair structure consisting of a small heading packet (P_h) and a large trailing packet (P_t) to measure queueing delay of a router from intra-probe gap measurement between the last bits of P_h and P_t, as shown in Figure 2. The proposed scheme uses UDP packets in its probing structure. Therefore, its deployment feasibility in the Internet is high. Because the scheme determines queueing delay from the change in the intra-probe gap, it does not require clock synchronization during measurement.

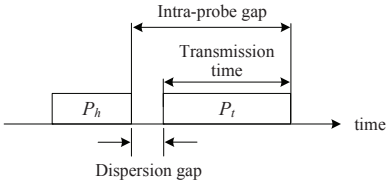


Fig. 2. A packet-pair structure consisting of a small heading packet, P_h and a large trailing packet, P_t.

The basic principle of the proposed scheme is that a non-zero dispersion gap between P_h and P_t (i.e., the separation between the last bit of P_h and the first bit of P_t, as Figure 2 shows) may decrease in the presence of cross-traffic packet(s) ahead of P_h. This phenomenon would reflect the size of queue in a route with a change (i.e., decrement) in the intra-probe gap, transmission time of P_t plus dispersion gap, between the probing packets. When a small P_h ahead of a large P_t is used in the packet pair, a non-zero dispersion gap can be created at the output link of a router under test [14]–[16], in order to determine a change in the intra-probe gap between the probing packets when there is queueing. For example, in a node i, if the transmission time of P_h on the output link (L_{i+1}) is smaller than the transmission time of P_t on the input link (L_i), the probing packets experience a non-zero dispersion δ_i, or:

$$\delta_i = \left(\frac{s_t}{l_i} - \frac{s_h}{l_{i+1}}\right) \quad (1)$$

where $\delta_i > 0$, s_h and s_t are the sizes of P_h and P_t, respectively, such that $s_h \leq s_t$, and l_i and l_{i+1} are the link capacities of L_i and L_{i+1}, respectively. If the queueing delays of P_h and P_t at node i are $t_{Q(h)}$ and $t_{Q(t)}$, respectively, the queueing delay at node i can be determined from the change in the intra-probe gap (Δ) as:

$$\Delta = \left(\frac{s_t}{l_{i+1}} + \delta_i\right) - \left[\left(\frac{s_t}{l_{i+1}} + \delta_i\right) - \left(t_{Q(h)} - t_{Q(t)}\right)\right] \quad (2)$$

where $\left(\frac{s_t}{l_{i+1}} + \delta_i\right)$ is the intra-probe gap between P_h and P_t, $\left(\frac{s_t}{l_{i+1}} + \delta_i\right) - \left(t_{Q(h)} - t_{Q(t)}\right)$ is the affected intra-probe gap due to queuing delay at node i, $\delta_i \geq t_{Q(h)}$, and $t_{Q(t)}$ is expected to be 0.

Figure 3 illustrates the intra-probe gaps between P_h and P_t with and without cross traffic at node i. In Figure 3(a), P_h and P_t arrive at node i from L_i back to back (i.e., with a zero-dispersion gap) when the node does not carry any other traffic; however, the probing packets experience a non-zero δ_i in the intra-probe gap (G_i) on L_{i+1} because of the difference between their transmission times on L_{i+1} and L_i, respectively. Figure 3(b) shows the effect of cross traffic at node i on G_i at L_{i+1}, which produces a smaller than the expected intra-probe gap. A magnified view of node i, Figure 3(c), shows that interference by a cross-traffic packet, C, ahead of P_h at the output queue, Q_{i+1}, delays the transmission of P_h through L_{i+1}, and this additional delay decreases the expected G_i on the output link to reflect the queuing delay at node i.

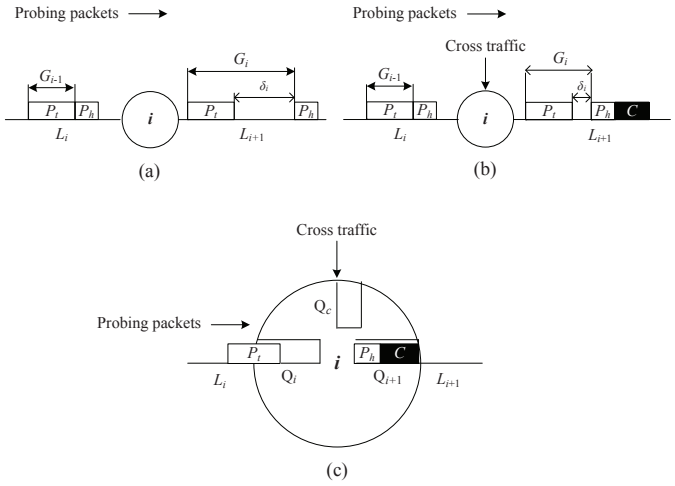


Fig. 3. Intra-probe gap between P_h and P_t (a) without and (b) with cross-traffic interference at node i, and (c) magnified view of node i when a cross-traffic packet, C, is queued before P_h at the output queue Q_{i+1}.

To measure queueing delay, the proposed scheme uses a pre-defined dispersion gap between P_h and P_t considering that the input- and output-link capacities of a router under test is known. Link capacities of a router can be known a priori through physical access or through estimation using existing capacity-measurement schemes [5], [11], [14], [15], [17]. The detailed steps of the proposed scheme to measure queueing delay at node i are presented below:

1. Identify the link capacities l_i and l_{i+1} of node i.
2. Determine the packet sizes s_h and s_t such that (1) holds.

3. Estimate the expected intra-probe gap between P_h and P_t at node i, $E[G_i]$, as:
$$E[G_i] = \frac{s_t}{l_{i+1}} + \delta_i \quad (3)$$
4. Send P_h and P_t using s_h and s_t, respectively, with a zero-dispersion gap between the probing packets to node i from L_i.
5. Time stamp P_h and P_t after receiving them on L_{i+1} and measure the intra-probe gap $M[G_i]$, where $M[G_i] \geq \frac{s_t}{l_{i+1}}$.
6. Determine the change in the intra-probe gap in reference to $E[G_i]$ as:
$$\Delta = E[G_i] - M[G_i], \quad (4)$$
such that $\Delta > 0$ as it corresponds to the decrement of the intra-probe gap, whereas $\Delta \leq 0$ corresponds to either no change or increment in the intra-probe gap between P_h and P_t.
7. Discard Δ if the corresponding $M[G_i] = \frac{s_t}{l_{i+1}}$ since this condition suggests that the queueing delay at node i could be larger than δ_i that is used in step 3.
8. Repeat steps 4 to 6 for $n > 1$ times and determine Δ in each iteration.
9. Calculate the mean and standard deviation of the sampled set, μ_Δ and σ, respectively, to provide a range of estimated queueing delay at node i as:
$$\text{Range of queueing delay} = \mu_\Delta \pm \sigma \quad (5)$$

IV. Simulation Results

We evaluated the performance of the proposed scheme in a simulation environment using ns-2 [18]. Figure 4 shows the simulation setup of a single-hop topology consisting of a router (node 1) connected between src (node 0) and dst (node 2). The capacities of the input and output links (L_1 and L_2, respectively) of node 1 are 100 Mb/s. The single-hop path carries four cross-traffic flows between node 1 and node 2 over L_2. The traffic flows are used to build queueing delay at node 1, which is of interest in our evaluation. In this simulation, we did not evaluate the performance of our scheme against other existing active schemes, e.g., cing, because they either use ICMP packets or require infrastructural support from the networks, which are not feasible to emulate in a simulation environment.

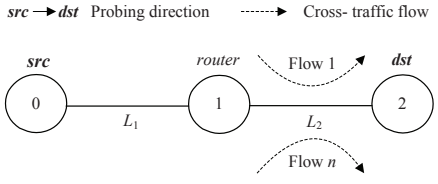


Fig. 4. A single-hop path with n cross-traffic flows, where $n = 4$.

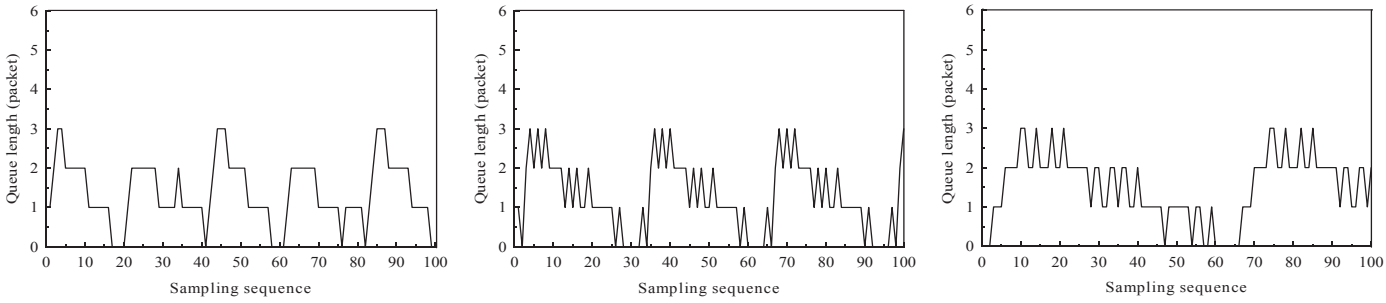
Table I shows the expected dispersion gaps between P_h and P_t (the last column) when different s_t sizes are used with $s_h = 64$ bytes considering the link capacities (i.e., 100 Mb/s) at node 1. We used 64-bytes packet in P_h as it is the smallest frame size in the Ethernet standard. Because the maximum decrease in the length of the intra-probe gap between P_h and P_t is determined by the expected dispersion between P_h and P_t, the values in the last column of Table I correspond to the upper bounds of the queueing delays that can be measured at node 1 using $s_h = 64$ bytes and $s_t = \{800, 1000, 1200, 1500\}$ bytes for P_h and P_t. We measured three different queueing delays, i.e., 11.21 μs (6.86 μs $\pm$ 4.35 μs), 17.36 μs (10.16 μs $\pm$ 7.2 μs), and 35.04 μs (21.44 μs $\pm$ 13.6 μs), at node 1 to evaluate the performance of the proposed scheme that are within the upper bounds of the queueing delays, as stated in Table I. The three different queueing delays are generated at node 1 by using 64-, 100-, 200-byte packets in each cross-traffic flow, respectively. The queue-length pattern (the first 100 samples for every 5 μs time interval) at node 1 for the above stated queueing delays are shown in Figure 5.

TABLE I
Probing-packet Size vs. Dispersion Gap at node 1

Probing-packet sizes		Link capacity		Dispersion gap
s_h (bytes)	s_t (bytes)	l_1 (Mb/s)	l_2 (Mb/s)	δ_1 (μs)
64	800	100	100	58.88
64	1000	100	100	74.88
64	1200	100	100	90.88
64	1500	100	100	114.88

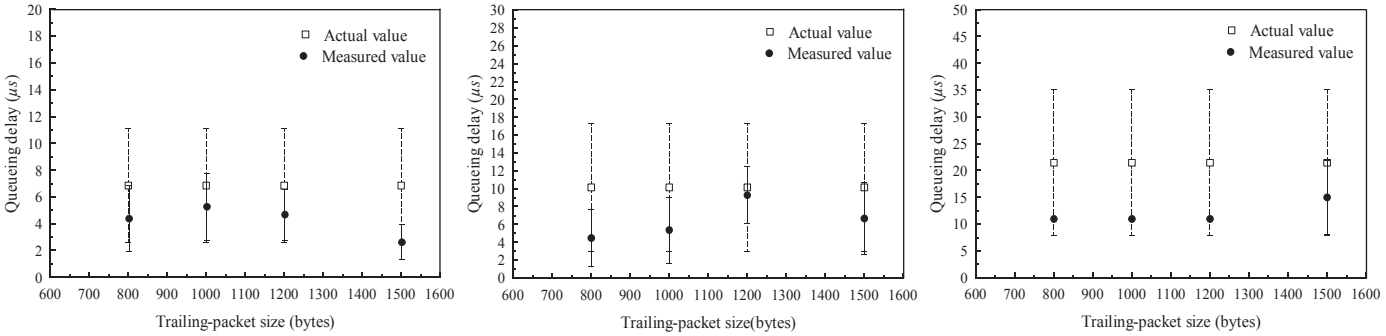
We evaluate the performance of the proposed scheme by sending 1000 pairs of probing packets between src and dst, where each pair is separated by a maximum of 100 μs. Figure 6 shows the simulation results of the proposed scheme when the upper bounds of the generated queueing delays at node 1 are: (a) 11.21 μs (6.86 μs $\pm$ 4.35 μs), (b) 17.36 μs (10.16 μs $\pm$ 7.2 μs), and (c) 35.04 μs (21.44 μs $\pm$ 13.6 μs). In Figure 6, the hollow boxes and solid circles are the mean of the actual queueing delays and measured queueing delays by the proposed scheme, respectively. The standard deviation of the actual and measured queueing delays are shown with the dashed and solid whiskers, respectively. We repeated the above stated evaluation of the scheme using different number of packet pairs, e.g., 30, 500, etc.; however, the 1000-pair test provided us the best possible results.

The simulation results in Figure 6 show that the proposed scheme achieves high accuracy for the tested queueing delays. For example, the range of the measured queueing delay, illustrated by the upper and lower whiskers from each mean value, overlap with range of the actual queueing delays at the router under test in every instance. We also calculated the error in the measured queueing delay of every case presented in Figure 6, where the error is the ratio between the length of the range of measured queueing delay that is outside the range of the actual value and the total length of the range of measured queueing delay. In majority of the cases, the accuracy of measurement is high (i.e., 100%). Otherwise, the errors are 14 and 50% when $s_t = 800$ and 1500 bytes in Figure 6(a), respectively, and 27, 18, and 4% when $s_t = 800$, 1000, and 1500 bytes in Figure 6(b), respectively. Considering the dynamic changes of the queueing delays and the errors in the measured values, the proposed scheme consistently achieves high measurement accuracy for different level of queueing delays. The measured values in Figure 6 also show that when the sizes of the probing packets create a larger dispersion in the packet pair, the queueing delay can be measured with high accuracy.



(a) 64-byte cross-traffic packet (b) 100-byte cross-traffic packet (c) 200-byte cross-traffic packet

Fig. 5. The output-queue length pattern at node 1 when the traffic-packet sizes of each cross-traffic flow are (a) 64 (b) 100, and (c) 200 bytes for generating 6.86 μs $\pm$ 4.35 μs, 10.16 μs $\pm$ 7.2 μs, and 21.44 μs $\pm$ 13.6 μs, of queueing delays, respectively.



(a) Queueing delay: 6.86 μs $\pm$ 4.35 μs (b) Queueing delay: 10.16 μs $\pm$ 7.2 μs (c) Queueing delay: 21.44 μs $\pm$ 13.6 μs

Fig. 6. Simulation results of the proposed scheme under different levels of queueing delays: (a) 6.86 μs $\pm$ 4.35 μs, (b) 10.16 μs $\pm$ 7.2 μs, and (c) 21.44 μs $\pm$ 13.6 μs at node 1 of the single-hop path.

V. Conclusions

We proposed an active scheme to measure queueing delay of routers in a single-hop scenario using a packet-pair model. The proposed scheme uses UDP packets in pairs and identifies the change in the intra-probe gap for measuring queueing delay at the router under test; therefore, it does not require infrastructural support and clock synchronization for measurement. The scheme is simple and can be applied to online routers. Simulation results for three different levels of queueing delays show that the proposed scheme is sensitive to small queueing delay and achieves high measurement accuracy.

References

[1] R. Rojas-Cessa, L. Ramesh, Z. Dong, L. Cai, and N. Ansari, "Parallel search trie-based scheme for fast IP lookup," in *IEEE GLOBECOM*, 2007, pp. 210–214.

[2] B. Forouzan, *TCP/IP Protocol Suit, ch. 12*. NJ, USA: McGraw Hall, 2006.

[3] C. Bovy, H. Mertodimedjo, G. Hooghiemstra, H. Uijterwaal, and P. V. Mieghem, "Analysis of end-to-end delay measurements in Internet," in *Proc. of PAM*, CO, USA, 2002, pp. 1–8.

[4] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, "Constraint-based geolocation of Internet hosts," *IEEE/ACM ToN*, vol. 14, no. 6, pp. 1219–1232, 2006.

[5] V. Jacobson. Pathchar - A tool to infer characteristics of Internet paths. [Online]. Available: ftp://ftp.ee.lbl.gov/pathchar/msri-talk.pdf.

[6] K. Anagnostakis, M. Greenwald, and R. Ryger, "cing: measuring network-internal delays using only existing infrastructure," in *IEEE INFOCOM*, CA, USA, 2003, pp. 2112–2121.

[7] N. Duffield, J. Horowitz, F. L. Presti, and D. Towsley, "Network delay tomography from end-to-end unicast measurements," in *Proc. of IWDC*, London, UK, 2001, pp. 576–595.

[8] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, and C. Diot, "Measurement and analysis of single-hop delay on an IP backbone network," *IEEE JSAC*, vol. 21, no. 6, pp. 908–921, 2003.

[9] M. Garetto and D. Towsley, "Modeling, simulation and measurements of queuing delay under long-tail Internet traffic," in *Proc. of ACM SIGMETRICS*, NY, USA, 2003, pp. 47–57.

[10] A. Adams, T. Bu, T. Friedman, J. Horowitz, D. Towsley, R. Caceres, N. Duffield, F. L. Presti, S. Moon, and V. Paxson, "The use of end-to-end multicast measurements for characterizing internal network behavior," *IEEE Commun. Soc. Mag.*, vol. 38, no. 5, pp. 152–159, 2000.

[11] A. Downey, "Using pathchar to estimate Internet link characteristics," in *Proc. of ACM SIGCOMM*, MA, USA, 1999, pp. 241–250.

[12] L. Angrisani, G. Ventre, L. Peluso, and A. Tedesco, "Measurement of processing and queuing delays introduced by an open-source router in a single-hop network," *IEEE TIM*, vol. 55, no. 4, pp. 1065–1076, 2006.

[13] K. Nichols and V. Jacobson, "Controlling queue delay," *Commun. of the ACM*, vol. 55, no. 7, pp. 42–50, 2012.

[14] K. Lai and M. Baker, "Measuring link bandwidths using a deterministic model of packet delay," in *Proc. of ACM SIGCOMM*, Stockholm, Sweden, 2000, pp. 283–294.

[15] K. Harfoush, A. Bestavros, and J. Byers, "Measuring bottleneck bandwidth of targeted path segments," in *Proc. of IEEE INFOCOM*, CA, USA, 2003, pp. 2079–2089.

[16] K. Salehin, R. Rojas-Cessa, C. Lin, Z. Dong, and T. Kijkanjanarat, "Scheme to measure packet processing time of a remote host through estimation of end-link capacity," *IEEE TC*, vol. 99, no. PrePrints, p. 1, 2013.

[17] K. Salehin and R. Rojas-Cessa, "A combined methodology for measurement of available bandwidth and link capacity in wired packet networks," *IET Commun.*, vol. 4, no. 2, pp. 240–252, 2010.

[18] The network simulator – ns-2. [Online]. Available: http://www.isi.edu/nsnam/ns/.