Simulation technique for available bandwidth estimation

T.G. Sultanov
Samara State Aerospace University
Moskovskoe sh., 34,
Samara, 443086, Russia
E-mail: tursul@rambler.ru

A.M. Sukhov
Samara State Aerospace University
Moskovskoe sh., 34,
Samara, 443086, Russia
E-mail: amskh@yandex.ru

Abstract—**This paper proposes a method for measuring the available bandwidth, based on testing network packets of various sizes (variable packet size method, VPS). The range of applicability of the model has been found, it is based on the accuracy of the measurements of packet delays. We have also derived a formula for the upper limit of the bandwidth. Computer simulations have been performed and the relationship between the measurement error of the available bandwidth and the number of measurements has been found. Experimental verification with the use of the RIPE test box measuring system has shown that this method has advantages over existing measurement techniques. The *Pathload* utility has been chosen as an alternative technique of measurement, and to ensure that reliable result statistics have been withdrawn by the SNMP agent directly from the router.**

Keywords-**Available bandwidth, RIPE test box, packet size, end-to-end delay, variable delay component.**

I. INTRODUCTION

Various real-time applications on the Internet, especially the transmission of audio and video information, have become more and more popular. The major factors defining the quality of the service are the quality of the equipment (the codec and a video server) and the available bandwidth of the Internet link. ISP providers should provide the required bandwidth for voice and video applications to guarantee the provision of the demanded services in the global network.

In this paper, the network path is defined as the sequence of links (hops) which forward packets from the sender to the receiver. There are various definitions of throughput metrics, but we will use the approaches accepted in [5], [9], [10].

Two bandwidth metrics that are commonly associated with a path are its capacity C and its available bandwidth B_{av} (see Fig 1). The *capacity C* is the maximum IP-layer throughput that the path can provide to a flow, when there is no competing traffic load (cross traffic). The *available bandwidth B_{av}*, on the other hand, is the maximum IP-layer throughput that the path can provide to a flow, given the path's current cross traffic load. The link with the minimum transmission rate determines the capacity of the path, while the link with the minimum unused capacity limits the available bandwidth. Moreover, measuring the available bandwidth is important for providing information to network applications on how to control their incoming and outgoing traffic and fairly share the network bandwidth.

Another related throughput metric is the *Bulk-Transfer-Capacity* (BTC). The BTC of a path in a certain time period is the throughput of bulk TCP transfer when the transfer is only
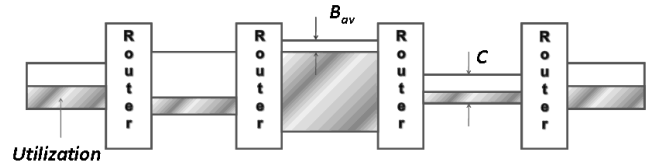

Fig. 1. Illustration of throughput metrics

limited by the network resources and not by limitations at the end-systems. The intuitive definition of BTC is the expected long-term average data rate (bits per second) of a single ideal TCP implementation over the path in question.

In order to construct a perfect picture of a global network (for monitoring and bottlenecks troubleshooting) and develop the standards describing new applications, a modern measuring infrastructure should be installed. In this paper we describe the usage of a RIPE test box measurement system which is widely used [7].

According to [7], this system doesn't measure the available bandwidth, but it collects the numerical values that characterize key network parameters such as packet delay, jitter, routing path, etc.

In this paper we attempt to provide a simple, universal model that allow us to estimate the available bandwidth based on data received from a RIPE test box measurement infrastructure (see our previous work [4]). The method is based on the *variable packet size* (VPS) method and was used in [6]. This method allows us to estimate the network capacity of hop i by using the connection between the round-trip time (RTT) and the packet size W.

II. THE MODEL AND ITS APPLICABILITY

The well-known expression for the throughput metric describing the relation between a network delay and the packet size is Little's Law [13]:

$$B_{av} = W/D, \tag{1}$$

where B_{av} is the available bandwidth, W is the size of the transmitted packet and D is the network packet delay (one way delay). This formula is ideal for calculating the bandwidth between two points on the network that are connected without any routing devices. In the general case, the delay value is caused by constant network factors as propagation delay, transmission delay, per-packet router processing time, etc. [9].

According to [1], Little's Law could be modified with D^{fixed}:

$$B_{av} = W/(D - D^{fixed}),\qquad(2)$$

where D^{fixed} is the *minimum fixed delay* for the packet size W. The difference between the delays D and D^{fixed} is the *variable delay component* d_{var}. In paper [3] it was shown that the variable delay is exponentially distributed.

Choi [2] and Hohn [12] showed that the minimum fixed delay component $D^{fixed}(W)$ for the packet size W is a linear (or affine) function of its size:

$$D^{fixed}(W) = W \sum_{i=1}^{h} 1/C_i + \sum_{i=1}^{h} \delta_i,\qquad(3)$$

where C_i is the capacity of the appropriate link and δ_i is the propagation delay. To prove this assumption, the authors experimentally found the minimum fixed delays for packets of the same size for three different routes and constructed a function of the dependence of a delay on the packet size W.

In order to eliminate the minimum fixed delay $D^{fixed}(W)$ from Eq. (2) we suggest testing the network link with packets of different sizes [1], so that the packet size varies up to the maximum possible value without possible router fragmentation. Then Eq. (2) could be modified to a suitable form for measuring the available bandwidth:

$$B_{av} = \frac{W_2 - W_1}{D_2 - D_1}\qquad(4)$$

This method allows us to find a way to eliminate the measurement limitations of the variable delay component d_{var}. The variable delay component is the cause of rather large measurement errors in other methods, which will be described in the last section of this paper.

The proposed model is quite simple, but it's still difficult to find an accurate measuring infrastructure. The first problem concerns the applicability of the model, i.e., what range of throughput metrics can be measured using this method. The second issue is the number of measurements (groups of packets) needed to achieve a given accuracy.

The first problem could be solved by using the measurement error (based on the delay accuracy measurement):

$$\eta = \frac{\Delta B}{B} = \frac{2\Delta D}{D_2 - D_1},\qquad(5)$$

where η is the relative error of measurement of the available bandwidth, ΔB is the absolute error of measurement of the available bandwidth, and ΔD is the precision of measuring the the packet delay.

With this expression we can easily find an upper bound $\bar{B}$ for the available bandwidth:

$$\bar{B} = \frac{W_2 - W_1}{2\Delta D}\eta\qquad(6)$$

Thus, with the RIPE test box, which allows for finding the delay to within 2 microseconds, $\Delta D = 2 \cdot 10^{-6}$ second precision, we can measure the available bandwidth up to the upper bound $\bar{B} = 300$ *Mbps* with a relative error of

$\eta = 10\%$. Moreover, if we use the standard utility *ping*, with a relative error $\eta = 25\%$ and a precision of 1 millisecond $\Delta D = 10^{-3}$ second, we could get results for the network available bandwidth up to 1.5 *Mbps*.

III. EXPERIMENTAL COMPARISON OF DIFFERENT METHODS

In this part of the paper we would like to compare different methods of measuring the available bandwidth, presenting the results obtained from experiments.

The experiment was divided into three stages. In the first stage we used a RIPE test box measurement system with two different packet sizes. The number of measurement systems in the global measurement infrastructure reaches 80, these points cover the major world centres of the Internet, reaching their highest density in Europe. The measurement error of packet the delay is 2–12 μs [14]. In order to prepare the experiments, three test boxes were installed in Moscow, Samara and Rostov on Don in Russia from 2006 to 2008, supported by RFBR grant 06-07-89074. For further analysis, we collected several data sets containing up to 3000 data results in different directions, including Samara–Amsterdam (tt01.ripe.net–tt143.ripe.net). Based on these data, we calculated the available bandwidth and the dependence of the measurement error on the number of measurements (see Fig 3).

The second stage compared the data obtained with our method with the results of traditional methods of throughput measurement. *Pathload* was selected as a tool that implements a traditional method of measurement [10]. This software is considered to be one of the best tools to assess the available bandwidth. *Pathload* uses self-loading periodic streams (SLoPS). It is based on a client–server architecture, which is a disadvantage, since you want to install the utility on both hosts. The advantage of *Pathload* is that it does not require root privileges, as the utility sends only UDP packets.

The results of measurements by *pathload* are displayed as a range of values rather than as a single value. The mid-range corresponds to the average throughput, and the band shows the change in the available bandwidth during the measurements.

The third stage involved the comparison of data obtained in the first and second stages with data directly from the router SSAU that serves the narrowest part of the network.

The experiment between points tt143.ripe.net (Samara State Aerospace University) and tt146.ripe.net (FREENet, Moscow) consists of three parts:

1) Measuring the available bandwidth by testing pairs of packets of different sizes using the measuring system RIPE test box (packet sizes of 100 and 1100 bytes);
2) Measurements of the available bandwidth using the utility *pathload*;
3) Measuring the available bandwidth by MRTG on the router SSAU which serves the narrowest routing hop (see Fig 2).

It is worth noting that all three inspections should be conducted simultaneously in order to maximize the reliability of the statistics. The structure of the measuring system RIPE
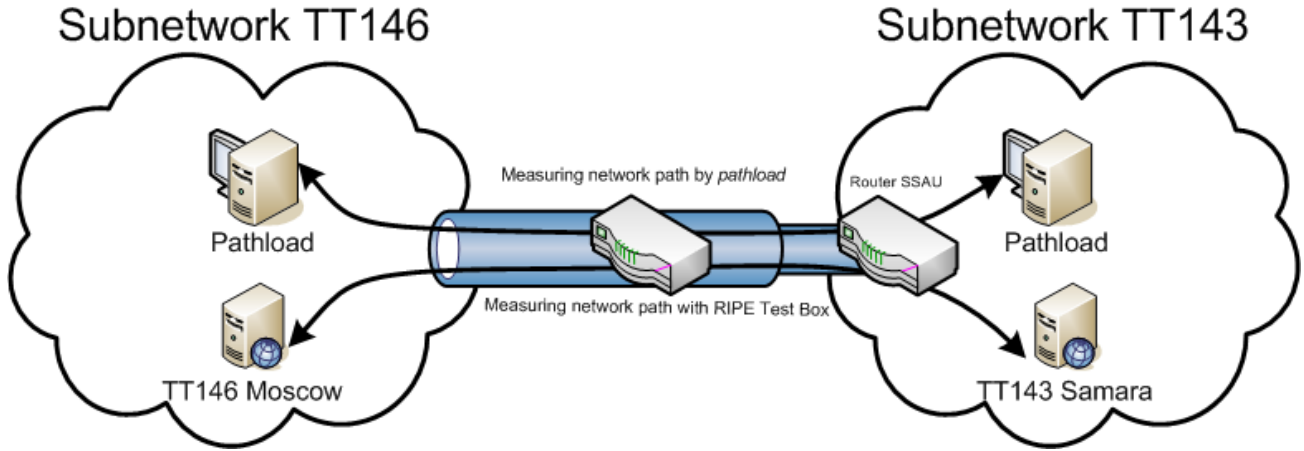
Fig. 2.   Scheme of the experiment

test box meets all the requirements of our method—it allows of changing the size of the probe packet and of finding high-precision delay.

By default, the test packet size is 100 bytes. There are special settings that allow adding testing packets of up to 1500 bytes to the desired frequency. In our case it is reasonable to add a packet size of 1100 bytes. It should be noted that testing of these packets does not begin until the next day after sending a special request.

In order to gain access to the test results it is necessary to apply for remote access (*telnet*) to the RIPE test box on port 9142. The data includes information about the desired delay packets of different sizes. In order to extract the data it is necessary to identify the packet on the receiving and transmitting sides.

First, to explore the sender's side:

SNDP 9 1263374005 -h tt01.ripe.net -p 6000 -n 1024 -s	1353080538
SNDP 9 1263374005 -h tt146.ripe.net -p 6000 -n 100 -s	**1353080554**
SNDP 9 1263374005 -h tt103.ripe.net -p 6000 -n 100 -s	1353080590

TABLE I
THE DATA OF SENDING BOX

The last value in line is the serial number of the packet. It should be remembered for getting a packet already on the receiving side of the channel. Below is a sample line on the receiving side.

For a given number of packets, it is easy to find the packet delay. In this case it is 0.009001 sec. The following packet 1353091581 has a size of 1100 bytes and the delay is

RCDP 12 2 89.186.245.200 55730 193.233.1.69 6000	1263374005.779364
0.009001 0X2107 0X2107 **1353080554** 0.000001 0.000001	
RCDP 12 2 200.19.119.120 57513 193.233.1.69 6000	1263374005.905792
0.160090 0X2107 0X2107 1353080554 0.000003 0.000001	

TABLE II
THE DATA OF RECEIVING BOX

0.027033 seconds. Thus, the difference is 0.018032 seconds. Other values are processed with similar packet delays.

The mean value $D_2 - D_1$ should be used in Eq. (4), so its necessary to average several values, going consistently. In the present experiment, the averaged difference $D_{av}(1100) - D_{av}(100)$ amounted to 0.000815 seconds in the direction $tt143 \rightarrow tt146$. Then the available bandwidth is

$$B_{av}(tt143 \rightarrow tt146) = \frac{8 \times 1000}{0.000815} = 9.8 \ Mbps$$

The average difference in the direction $tt146 \rightarrow tt143$ was 0.001869 seconds. Then the available bandwidth will be:

$$B_{av}(tt146 \rightarrow tt143) = \frac{8 \times 1000}{0.001869} = 4.28 \ Mbps$$

Measuring with the *pathload* utility had periodic troubles, even though all necessary ports had been opened. In the direction of measurement $tt146 \rightarrow tt143$ the program did not get any results despite all our attempts. It was idle and filled the channel chain packets. The *pathload* results give a large spread of values, clearly beyond the capacity of the investigated channel. The other measurements with utilities *pathChirp* and *IGI* were also unsuccessful. These programs gave errors and refused to measure the available bandwidth.

Therefore, it was decided to compare the results with data obtained directly from the router. The *Traceroute* utility determines the bottleneck of a route between SSAU and Institute of Organic Chemistry at the Russian Academy of Sciences. It was an external SSAU router whose bandwidth was limited to 30 Mbps. The SNMP agent collected statistics of the border router SSAU.

All data are presented in Table III indicating the time of the experiment. Table III shows that the results obtained by our method are in a good agreement with the router data, while the *pathload* measurements differ. The study of the statistical type of delay [3] provides an answer to the question why this is happening. The dispersion of measurements results speaks of the presence of a variable part of delay d_{var}. This utility uses self-loading periodic streams (SLoPS) like most others. This method consists in the generation of packets

N	Date	Direction	Available bandwidth (data of RIPE test box)	Available bandwidth (data of *pathload*)	Data from router
1	13.01.2010	tt143→tt146	10.0 ± 2.2 Mbps	21.9 ± 14.2 Mbps	12.1 ± 2.5 Mbps
2	13.01.2010	tt146→tt143	4.4 ± 1.2 Mbps		7.8 ± 3.8 Mbps
3	23.01.2010	tt143→tt146	20.3 ± 5.1 Mbps	41.2 ± 14.0 Mbps	18.7 ± 1.1 Mbps
4	23.01.2010	tt146→tt143	9.3 ± 2.7 Mbps		11.3 ± 2.6 Mbps
5	06.02.2010	tt143→tt146	9.2 ± 1.4 Mbps	67 ± 14 Mbps	12.0 ± 2.0 Mbps
6	06.02.2010	tt146→tt143	3.5 ± 1.2 Mbps		4.5 ± 2.0 Mbps

TABLE III
COMPARATIVE ANALYSIS OF MEASUREMENT RESULTS

chain with redundant frequency when time packet delivery will significantly increase due to long queues at the routers. In this case, the transmitter starts to reduce the frequency of packets generation until the queue disappears. Next, the process will be repeated for as long as the average frequency of packets generation will not approach the available bandwidth. The main disadvantage of this technique is that the measurements are unreliable because they have not considered the influence of the variable part of delay. This is the reason for the fantastic figure of 90 Mbps given by *pathload* for channel with a 30 Mbps capacity.

IV. THE REQUIRED NUMBER OF MEASUREMENTS

The main disadvantage of most modern tools is a large spread of values of the available bandwidth. Measurement mechanisms of throughput utilities do not take into account the effect of the variable part of the delay. Unfortunately, not all utilities have compensation mechanisms for the random component's delay.

Any method that gives accurate results should contain a mechanism for smoothing the effect of d_{var}. In order to understand the effect of the variable part on the measurement results, we turn to the following experiment. The series of measurements were made between RIPE test boxes tt01.ripe.net (Amsterdam, Holland) and tt143.ripe.net (Samara State Aerospace University, Russian Federation). About 3000 values of of the delay for packet sizes of 100 and 1024 bytes in both directions were received. Using the proposed new method, values of the available bandwidth were calculated for cases where the averaging was performed on 20, 50 and 100 pairs of values. In Fig. 3 the available bandwidth calculated for various conditions of averaging is represented.

Apparently, beatings of the calculated available bandwidth remain critical at 20 averaged values. At 50 it is less noticeable, and at 100 the values in the curve are almost equalized. There was a clear correlation between the number of measurements and the variation of the calculated available bandwidth. The beats are caused by the variable part of the delay; its role is reduced as the number of measurements increases.

In this section the necessary number of measurements is calculated using two methods: from the experimental data of the RIPE test box, and by simulation knowing the distribution type for network delay.

Based on data obtained from boxes tt01 and tt143 we computed the standard deviations (SD) $\sigma_n(B)$ of the available bandwidth.

The data are presented in Table IV and graphically depicted in Fig. 4.

Fig. 4 shows that it is necessary to take at least 50 measurements (the delay difference for 50 pairs of packets). In this case, the calculated value exceeds twice the capacity of SD, i.e., $B \geq 2\sigma_n(B)$.

A more accurate result can be obtained using the generating functions for describing the delay packets. In paper [3] it is shown that the delay distribution is described by an exponential law and the following generating function can be used for delay emulation:

$$D = D_{min} + W/B - (1/\lambda)ln(1 - F(D, W)), \quad (7)$$

where $\lambda = 1/(D_{av} - D_{min})$. The function $F(W, D)$ is a standard random number generator in the interval $[0; 1)$.

Knowledge of the generating function allows of calculating the tabulated values of η_n^T from Eq. 5. Earlier standard deviation $\sigma_n^T(D_2 - D_1)$ for the delay difference is found taking $\lambda^T = 1000 s^{-1}$. The calculation will hold for the following values: $\Delta W^T = W_2 - W_1 = 1000$ *bytes*, $B^T = 10$ *Mbps*, which corresponds to $D_2^T - D_1^T = 8 \cdot 10^{-4}s$.

For the $\sigma_n^T(D_2 - D_1)$ values from Table V, values of η_n^T could be found (see Table VI).

Number n of measurements,	5	10	20	30	50	100	200
Measurement error, η_n^T (%)	82.6	61.1	44.2	35.5	24.4	13.9	9.4

TABLE VI
DEPENDENCE OF ERROR ON THE NUMBER OF MEASUREMENTS

During the real experiment, the measured quantities λ^{exp}, $D_2^{exp} - D_1^{exp}$, and B^{exp} take arbitrary values, but correction factors can easily calculate the required number of measurements:

$$\eta_n^T = k(D_2 - D_1) \cdot k(\lambda) \cdot \eta_n^{exp}, \quad (8)$$

where $k(\lambda) = \lambda^{exp}/\lambda^T$, and $k(D_2 - D_1) = (D_2^{exp} - D_1^{exp})/(D_2^T - D_1^T)$.

Substituting into Eq. 8 the values of the coefficients $k(D_2 - D_1)$, $k(\lambda)$ and the desired accuracy of measurements η^{exp}, we compare the obtained values with the tabulated η_n^T and find
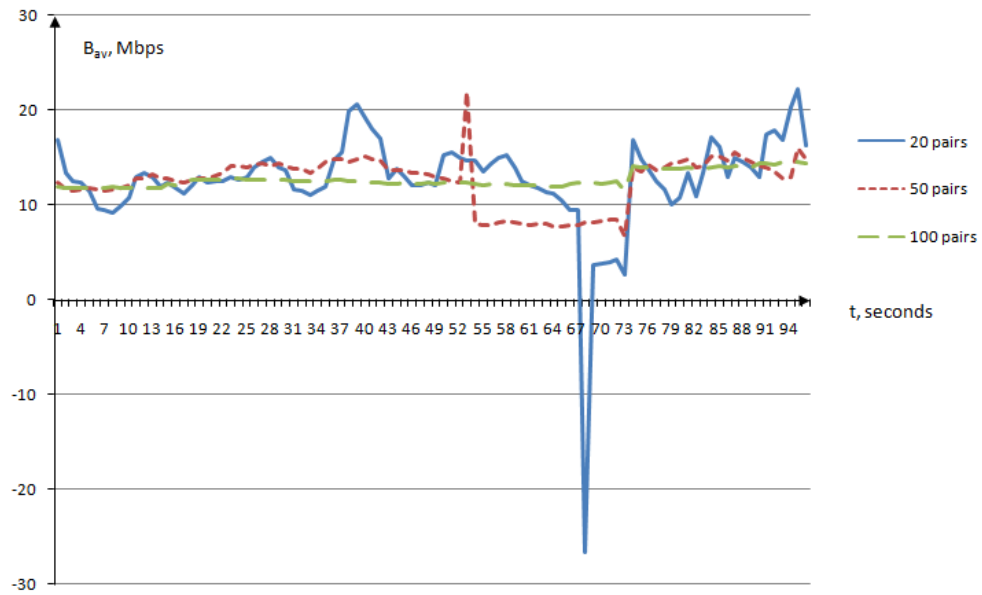
Fig. 3.   Dependence of available bandwidth on the number of measurements

Number of measurements, n	5	10	20	30	40	50	70	100	200	300
Standard deviations, $\sigma_n(B)$ (Mbps)	22.2	14.9	10.2	8.3	7.3	6.7	5.7	4.9	2.9	2.3
The average value of available bandwidth, B_{av} (Mbps)	13.1									

TABLE IV
DEPENDENCE OF SD ON THE NUMBER OF MEASUREMENTS
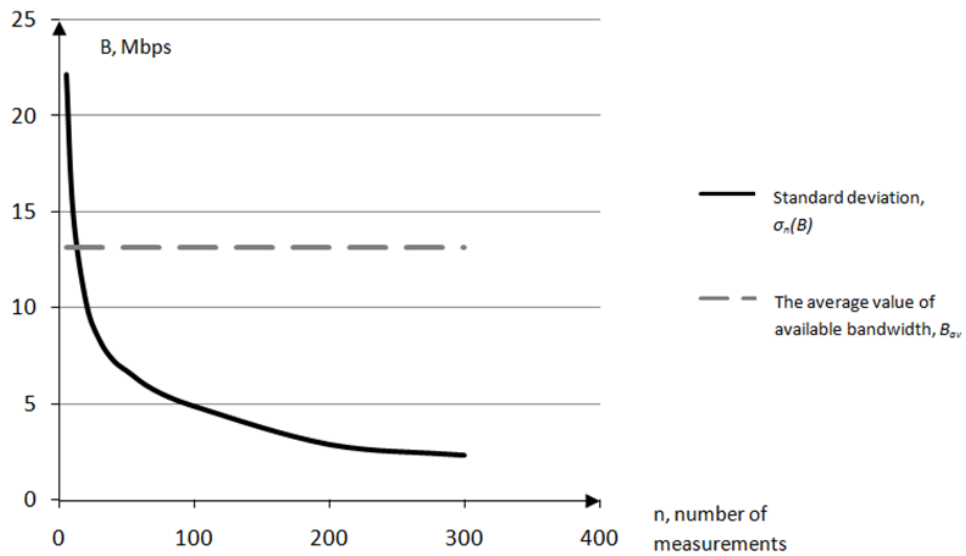
Fig. 4.   Dependence of SD on the number of measurements

Number of measurements, n	5	10	20	30	50	100	200
Standard deviations, $\sigma_n^T(D_2 - D_1)$ (ms)	0.661	0.489	0.354	0.284	0.195	0.111	0.075

TABLE V
EMULATION OF DEPENDENCE SD ON THE NUMBER OF MEASUREMENTS ($\lambda^T = 1000\ s^{-1}$)

the number of measurements n required to achieve a given error.

V. CONCLUSION

In this paper we found a way to measure the available bandwidth by using the data that could be collected by the RIPE test box about packet delays. This method consists of comparing the average end-to-end delays for packets of different sizes: from this, we can calculate the available bandwidth.

We carried out a further study of the model and found the limits of its applicability, which depend on the accuracy of the measurements of the packet delays. The experimental results were obtained both from our method and the alternative ones. The benchmark tool was the utility *pathload*. This paper shows that the accuracy of calculations of the available bandwidth depends on the variable delay component, d_{var}.

Experiments and computer simulations, using the generating function of the delay, were conducted. They have shown that achieving a given error requires to average a large number of measurements. We found a relationship between the accuracy and the number of measurements to ensure the required level of accuracy.

In the future we plan to implement this method within the mechanism of the RIPE test box measurement infrastructure.

ACKNOWLEDGMENT

REFERENCES

[1] Platonov A.P., Sidelnikov D.I., Strizhov M.V., and Sukhov A.M., Estimation of available bandwidth and measurement infrastructure for Russian segment of Internet, Telecommunications, 2009, 1, pp. 11–16.

[2] Choi, B.-Y., Moon, S., Zhang, Z.-L., Papagiannaki, K., and Diot, C., Analysis of point-to-point packet delay in an operational network. In: Infocom 2004, Hong Kong, pp. 1797–1807 (2004).

[3] Sukhov, A., and Kuznetsova, N., What type of distribution for packet delay in a global network should be used in the control theory? 2009; arXiv: 0907.4468.

[4] Sultanov, T., Sukhov, A., Strizhov, M. and Polukarov, D., Throughput metrics and packet delay in TCP/IP networks. RIPE59 Meeting, Lisbon, 2009 arXiv: 0907.3710.

[5] Dovrolis C., Ramanathan P., and Moore D., Packet-dispersion techniques and a capacity-estimation methodology, IEEE/ACM Transactions on Networking, v.12, n.6, Dec. 2004, pp. 963–977.

[6] Downey A.B., Using Pathchar to estimate internet link characteristics. In: Proc. ACM SICCOMM, Sep. 1999, pp. 222–223.

[7] RIPE Test Box, http://ripe.net/projects/ttm/

[8] Jacobson, V., Congestion avoidance and control. In: Proceedings of SIGCOMM 88 (Stanford, CA, Aug. 1988), ACM.

[9] Prasad R.S., Dovrolis C., and Mah, B.A., The effect of layer-2 store-and-forward devices on per-hop capacity estimation. In: Proc. IEEE INFOCOM, Mar. 2003, pp. 2090–2100.

[10] Jain, M., and Dovrolis, K., End-to-end estimation of the available bandwidth variation range. In: SIGMETRICS'05, Banff, Alberta, Canada (2005).

[11] Crovella, M.E., and Carter, R.L., Dynamic server selection in the Internet. In: Proc. of the Third IEEE Workshop on the Architecture and Implementation of High Performance Communication Subsystems (1995).

[12] Hohn, N., Veitch, D., Papagiannaki, K., and Diot, C., Bridging router performance and queuing theory, Proc. ACM SIGMETRICS, New York, Jun. 2004.

[13] Kleinrock, L., Queueing Systems, vol. II. New York: Wiley, 1976.

[14] Georgatos, F., Gruber, F., Karrenberg, D., Santcroos, M., Susanj, A., Uijterwaal, H., and Wilhelm R., Providing active measurements as a regular service for ISPs. In: PAM2001.

[15] Padhye, J., Firoiu, V., Towsley, D., and Kurose, J., Modeling TCP throughput: A simple model and its empirical validation. In: Proc. SIGCOMM Symp. Communications Architectures and Protocols, pp. 304–314 (1998).