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—The paper proposes a method for measuring available bandwidth, based on testing network packets of various sizes (Variable Packet Size method, VPS). The boundaries of applicability of the model have been found, which are based on the accuracy of measurements of packet delays, also we have derived a formula of measuring the upper limit of bandwidth. The computer simulation has been performed and relationship between the measurement error of available bandwidth and the number of measurements has been found. Experimental verification with the use of RIPE Test Box measuring system has shown that the suggested method has advantages over existing measurement techniques. *Pathload* utility has been chosen as an alternative technique of measurement, and to ensure reliable results statistics by SNMP agent has been withdrawn directly from the router.

Index Terms—Available bandwidth, RIPE Test Box, packet size, end-to-end delay, variable delay component.

I. INTRODUCTION

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

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

Two bandwidth metrics that are commonly associated with a path are the capacity C and the 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 available bandwidth is important to provide information to network applications on how to control their incoming and outgoing traffic and fairly share the network bandwidth.
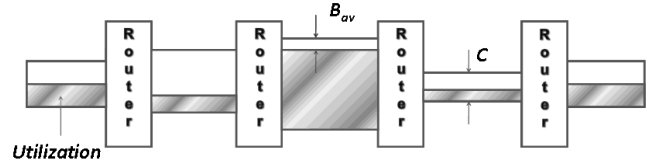

Fig. 1. Illustration of throughput metrics

Another related throughput metric is the *Bulk-Transfer-Capacity* (BTC). BTC of a path in a certain time period is the throughput of a bulk TCP transfer, when the transfer is only 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 (monitoring and bottlenecks troubleshooting) and develop the standards describing new applications, modern measuring infrastructure should be installed. In this paper we are describing usage of 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, which characterize key network parameters such as packet delay, jitter, routing path, etc.

In this paper we attempt to provide universal and simple model that allow us to estimate available bandwidth based on received data from RIPE Test Box measurement infrastructure. The method is based on *Variable Packet Size* (VPS) method and was used in [6]. This method allows us to estimate network capacity of a hop i by using connection between the Round-Trip Time (RTT) and packet size W.

II. THE MODEL AND ITS APPLICABILITY

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

$$B_{av} = W/D, \qquad (1)$$

where B_{av} is available bandwidth, W is the size of transmitted packet and D is 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 general case 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 *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 variable delay is exponentially distributed.

Choi [2], Hohn [12] showed that minimum fixed delay component $D^{fixed}(W)$ for the packet size W is an 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 capacity of appropriate link and δ_i is propagation delay. To prove this assumption authors experimentally found the minimum fixed delays for packets of the same size for three different routes and constructed function of dependence of a delay from the packet size W.

In order to eliminate minimum fixed delay $D^{fixed}(W)$ from Eqn. (2) we are suggesting to test network link with packets of different sizes [1], so that the packet size varied at the maximum possible value without possible router fragmentation. Then Eqn. (2) could be modified to a suitable form for measuring 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 variable delay component d_{var}. The variable delay component is the cause rather large measurement errors of other methods, which will be described in the last section of this paper.

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

First problem could be solved by using measurement error (based on delay accuracy measurement):

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

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

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

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

Thus, with the RIPE Test Box, which allows to find the delay to within 2 microseconds $\Delta D = 2 \cdot 10^{-6}$ second precision, we can measure the available bandwidth to the upper bound $\bar{B} = 300$ *Mbps* with relative error $\eta = 10\%$. Moreover, if we were using standard utility *ping*, with a relative error $\eta = 25\%$ and precision of 1 millisecond $\Delta D = 10^{-3}$ second, we could get results for network available bandwidth up to 1.5 *Mbps*.

III. EXPERIMENTAL COMPARISON OF DIFFERENT METHODS

In this part of the paper we would like to show results based on comparing different methods of measuring available bandwidth by using obtained results of experiments.

The experiment was divided in three stages. In the first stage we have used RIPE Test Box measurement system with two different packet sizes. Number of measurement systems in global measurement infrastructure reaches 80, these points are covering major Internet world's centers of the Internet, reaching highest density in Europe. The measurement error of packet delay is 2-12 μs [14]. In order to prepare the experiments, three Test Boxes have been installed in Moscow, Samara and Rostov on Don in Russia during 2006-2008 years in support of 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 available bandwidth and dependence of measurement error on the number of measurements (see Fig 3).

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

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

The third stage involves the comparison of data obtained in the first and second stages with the data directly from the router SSAU which 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 size of 100 and 1100 bytes);
2) Measurements of available bandwidth using the utility *pathload*;
3) Measuring the available bandwidth by MRTG on the router SSAU which serves the narrowest hop of routing
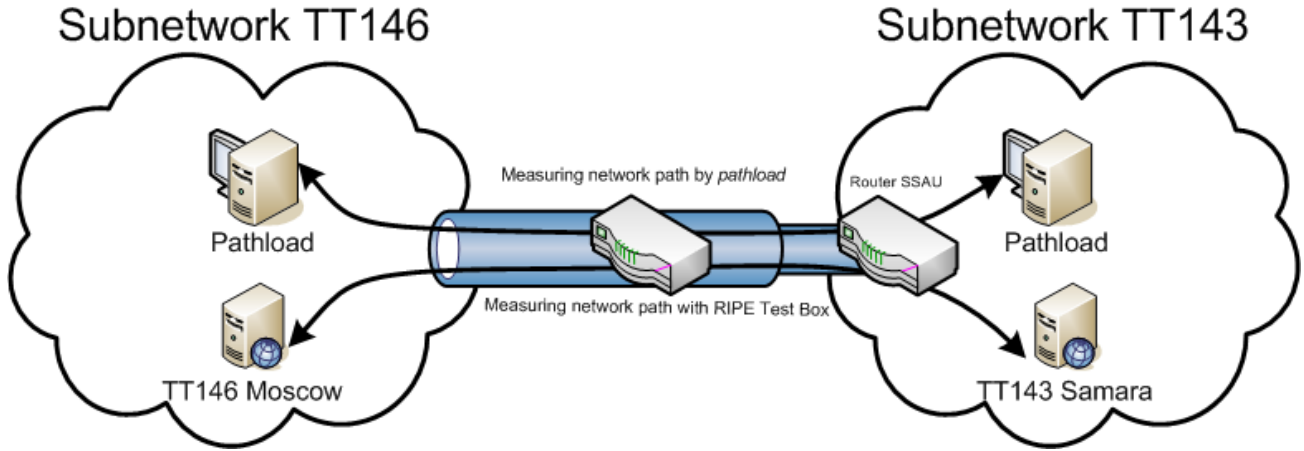
Fig. 2.   Scheme of the experiment

(see Fig 2).

It is worth noting that all the three inspections should be conducted simultaneously in order to maximize the reliability of the statistics. The structure of the measuring system RIPE Test Box meets all the requirements of our method - it allows to change the size of the probe packet and find 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 receiving and transmitting sides.

First, it should to explore 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 packet. It should to be remembered to get a packet already on the receiving side of the channel. Below is a sample line on the receiving side.

For a given number of packet is easy to find the packet delay. In this case it is 0.009001 sec. The following packet 1353091581 is the size of 1100 bytes and the delay is 0.027033 seconds. Thus, the difference is 0.018032 seconds. Other values are processed packet delay similar.

The mean value $D_2 - D_1$ should be used in Eqn. (4), so its necessary to average several values, going consistently. In the present experiment, the averaged difference

```
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

$D_{av}(1100) - D_{av}(100)$ amounted to 0.000815 seconds in the direction $tt143 \rightarrow tt146$. Then the available bandwidth can be calculated as:

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

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 of *pathload* utility was with periodically troubles, even though that had been opened all necessary ports. In the direction of measurement $tt146 \rightarrow tt143$ the program has not get any results despite all our attempts. It is 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. Programs give errors and refused to measure the available bandwidth.

Therefore, it was decided to compare the results obtained by different methods with data obtained directly from the router. *Traceroute* utility determines "bottleneck" of route path between SSAU and Institute of Organic Chemistry at the Russian Academy of Sciences. It was an external SSAU router which bandwidth was limited up to 30 *Mbps*. SNMP agent collects statistic 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 and router data are in a good agreement, while the *pathload* measurements differ. The study of the statistical

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

TABLE III

COMPARATIVE ANALYSIS OF MEASUREMENT RESULTS

type of delay [3] provides an answer to the question why this is happening. The dispersion of measurements results speaks 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 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 it 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 unreliable measurements because they have not considered the influence of the variable part of delay. This is the reason for fantastic 90 *Mbps pathload* result 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 available bandwidth. Measurement mechanisms of throughput utilities do not take into account the effect of the variable part of the delay. Unfortunately, in all developed utilities compensations mechanism for the random component's delay isn't provided.

Any method that gives accurate results should contain mechanism for smoothing the impact 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 have been made between RIPE Test Boxes: tt01.ripe.net (Amsterdam, Holland) and tt143.ripe.net (Samara State Aerospace University, Russian Federation). It was received about 3000 values of delay packet size of 100 and 1024 bytes in both directions. Using the presented method quantities of available bandwidth have been calculated for cases where the averaging is performed on 20, 50 and 100 pairs of values. On Fig. 3 the schedule of the available bandwidth calculated for various conditions of averaging is represented.

Apparently from the schedule, beatings of the calculated available bandwidth remain critical at 20 averaged values. At 50 it is less noticeable, and at 100 the values the curve is almost equalized. There was a clear correlation between the number of measurements and 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.

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

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

Data are presented in Table IV and graphically depicted in Fig. 4.

Figure 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 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 calculating the tabulated values of η_n^T from Eqn. 5. Earlier standard deviation $\sigma_n^T(D_2 - D_1)$ for the delay difference is found taking $\lambda^T = 1000s^{-1}$. 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 experimental 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:

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

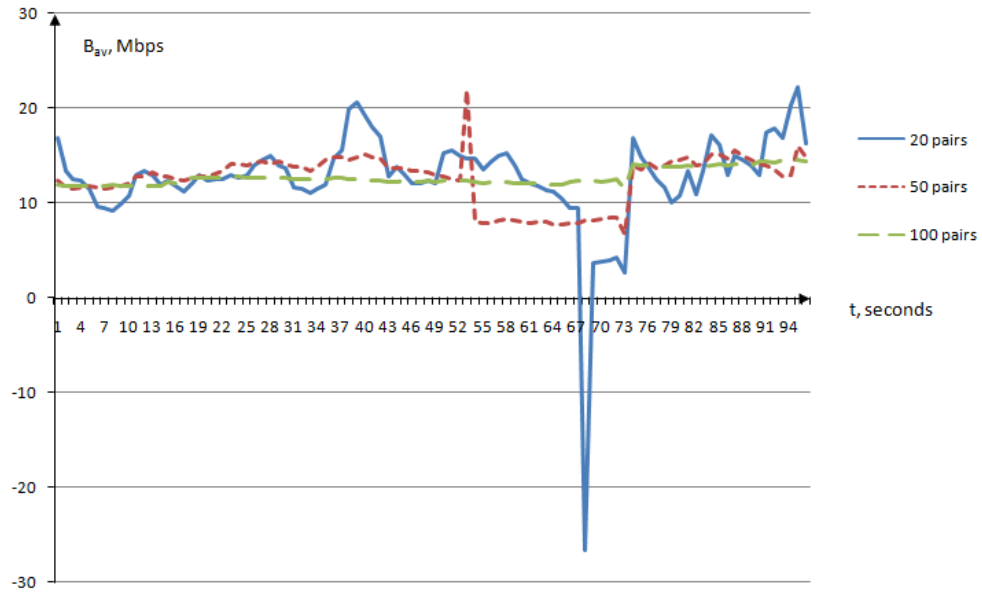where $k(\lambda) = \lambda^{exp}/\lambda^T$, and $k(D_2 - D_1) = (D_2^{exp} - D_1^{exp})/(D_2^T - D_1^T)$.

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
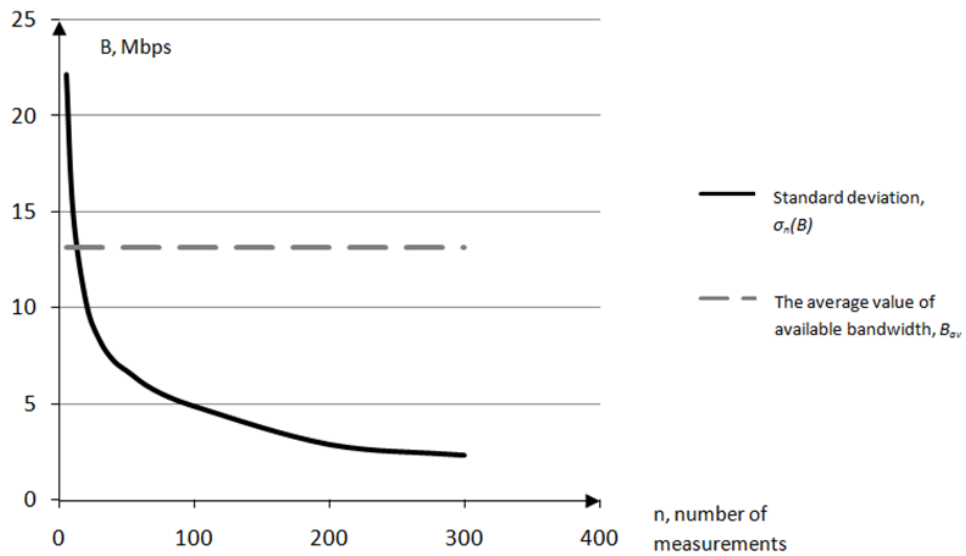
Fig. 4.   Dependence of SD on the number of measurements

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 ($\lambda^T = 1000\ s^{-1}$)

Substituting in Eqn. 8 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 the number of measurements n required to achieve a given error.

V. CONCLUSION

In this paper we found a way to measure available bandwidth by data of delays that could be collected by RIPE Test Box. This method consists in the fact that comparing the average end-to-end delays for packet of different sizes, 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 measurement delays. The experiment results were obtained with our method and the alternative. The benchmark tool has been selected utility *pathload*. The paper shows that the accuracy of calculations available bandwidth depends on variable delay component, d_{var}.

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

In the future we plan to implement this method in the mechanism of the measurement infrastructure RIPE Test Box.

ACKNOWLEDGMENT

REFERENCES

[1] Platonov A.P., Sidelnikov D.I., Strizhov M.V., 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] A.M. Sukhov, N Kuznetsova, What type of distribution for packet delay in a global network should be used in the control theory? 2009; arXiv: 0907.4468.

[4] Padhye, J., Firoiu, V., Towsley, D., Kurose, J.: Modeling TCP Throughput: A Simple Model and its Empirical Validation. In: Proc. SIGCOMM Symp. Communications Architectures and Protocols, pp. 304-314 (1998).

[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, December 2004, p. 963-977.

[6] Downey A.B., Using Pathchar to estimate internet link characteristics, in Proc. ACM SICCOMM, Sept. 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 B. A. 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., Dovrolis, K.: End-to-end Estimation of the Available Bandwidth Variation Range. In: SIGMETRICS'05, Ban, 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] N. Hohn, D. Veitch, K. Papagiannaki and C. Diot, Bridging Router Performance And Queuing Theory, Proc. ACM SIGMETRICS, New York, USA, Jun 2004.

[13] Kleinrock, L. Queueing Systems, vol. II. John Wiley & Sons, 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 ISP's. In: PAM2001.