Packet Delay Evaluation Using M/G/WFQ Model

BRUNONAS DEKERIS, LINA NARBUTAITE, TOMAS ADOMKUS
Department of Telecommunications
Kaunas University of Technology
Studentu st. 50, Kaunas
Lithuania

Abstract: - An actual problem is how to evaluate the quality of IP network. Differentiated Services is a promising architecture for the next generation Internet due to its highly flexible, scalable and interoperable design. In Differentiated Services, scheduling disciplines play an important role in achieving service differentiation. In this paper two queuing models are pending: M/G/WFQ and M/G/1/p3. Thus we are considering, which model can better support quality of services with the lowest difference of delay between data packets.

Key-Words: IP network, Differentiated Services, Quality of Service, Class of Service, Queuing discipline, M/G/1/WFQ, M/G/1/p5.

1 Introduction

IP networks are currently evolving from their original architecture, capable of supporting a single Best Effort (BE) type of service, towards to a new advanced architecture characterized by the capability of differentiating a multiplicity of Classes of Services (CoS) providing different levels of performances (IP Differentiated Services scenario) [1]. This evolution enables the utilization of IP networks to offer a wide variety of highly valuable services, creating new business opportunity for Telecom operators and accelerating the process of renovation of network infrastructures [2]. In order to be successful, however, this process requires a greater capability of controlling network performances with respect to the former structure. As a consequence, there is an increasing interest towards to the definition and the implementation of techniques exploitable to meet the desired level of performance under different operational conditions. This new field of activities is usually referred to as a traffic engineering for IP networks [3].

An issue that still needs more research in the Differentiated Services (DiffServ) architecture is how to evaluate the flow level quality requirements and fairness issues [4], e.g., packet loss probabilities and fairness in bandwidth allocation, which can be achieved by means of the packet level control mechanisms. Functioning of various DiffServ PHB groups, including Assured Forwarding (AF), Expedited Forwarding (EF) and Best Effort Forwarding (BF) has been analyzed with analytical models in [5] and with simulation models in [6].

A guaranteed, or at least expected, throughput is the quality of service (QoS) feature of interest to most applications, and that tagging packets should be used to provide such a guarantee [7, 8, 9]. However, current proposals for DiffServ architectures do not quantify the service they would provide applications [10, 11, 5].

In order to support DiffServ in IP networks, different queuing disciplines are in use [12]. In this paper we consider absolutely new queuing disciplines: M/G/1/p3 and M/G/WFQ, based on M/G/1. In the second one queuing discipline, there are integrated M/G/1 queuing discipline and WFQ queue scheduling discipline. Weighted Fair Queuing (WFQ) was developed independently in 1989 by Lixia Zhang and by Alan Demers, Srinivasant Keshav, and Scott Shenke. WFQ is the basis for a class of queue scheduling disciplines. WFQ can provide strong upper-bound, end-to-end delay performance guarantees and supports the fair distribution of bandwidth for variable-length packets by approximating a generalized processor sharing (GPS) system.

The rest of the paper is organized as follows. In Section 2, we analyze and obtain analytic expressions for queuing models M/G/1/p3 and M/G/WFQ. In Section 3 we use expressions from Section 2, as well as simulations to illustrate the benefits and shortcomings of these models. Section 4 concludes the paper.

2 Analysis of queuing models M/G/WFQ and M/G/1/p3

We are analyzing two queuing models. The first one is, proposed by us, M/G/WFQ model, based on M/G/1 queuing model with several queues for each incoming priority traffic and WFQ queue scheduling discipline. WFQ queue scheduling discipline does provide fairness protection across classes, so all classes of the traffic receives assured access to the output link [12]. And the second one is M/G/1/p3 model, based on M/G/1 queuing model with one queue for all incoming traffic, without reference which priority class it has. Both these models can be realized in the DiffServ router and support proper quality of service. Our goal is to analyze, which model can better support less difference between mean delay of the data packet with high and low priority of class. So we do not take into consideration packet loss, in other word the length of buffer in our models is infinity.

2.1 M/G/WFQ Stochastic Analysis

Let us consider a single-server EFTF WFQ system fed by multiple Poisson streams with arrival rates $\lambda_1, \lambda_2, ..., \lambda_K$ as shown in Figure 1. The buffers corresponding to different flows are infinite in length and the packets in each of those buffers are served in the order they arrive. We use M_k^i to denote the length (in bits) of the *i*th data packet arrival at the *k*th buffer, $k \in K$.

We use new approach $T_k^i = M_k^i / C$ to denote the service time (in seconds) of the *i*th data packet arrival at the *k*th buffer, where *C* is the output link capacity. The random variables T_k^i from the multiple Poisson streams are identically distributed, mutually independent, and independent of the arrival times. Such variables T_k^i can assume any general distribution. We denote the mean service time of arriving data packets by $\overline{T} = E[T_k^i] = 1/\mu$, where μ is the mean service rate. The second moment of the service time is denoted by $\overline{T^2}$. For convenience, we refer to the total arrival rate at the WFQ system by $\lambda = \sum_{k \in K} \lambda_k$.

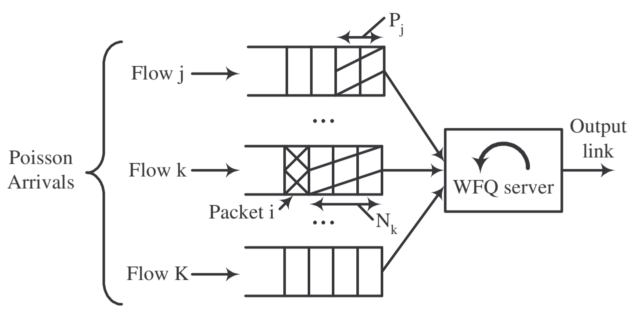

Figure 1. Arrivals at the Weighted Fair Queuing System.

The utilization of each connection *k* is denoted by $\rho_k = \lambda_k \overline{T} = \lambda_k / \mu$, while the utilization of the output link is given by $\rho = \lambda \overline{T}$. We also previously defined $\rho'_k = \lambda_k \overline{M_k} / r_k$. In this analysis we maintain $\rho'_k < 1$ for all $k \in K$. This ensure that we maintain $\rho < 1$, which keeps the system from being overloaded (in an average sense).

We denote the *mean* packet waiting time in queue *k* by W_k, and the *expected* number of packets in such a queue (*not* including any packet that may be in service) by N_k. We assume *ergodicity* of the queueing system (which is true provided, when $\rho < 1$) and note that, in our system, the values of N_k and W_k seen by an outside observer at a random time are the same as seen by an arriving customer. This is due to the Poisson character of the arrival process, which implies that the occupancy distribution upon arrival is typical [13].

Now, let us consider the *i*th data packet arrival at the *k*th queue of the WFQ system. This packet must wait in queue for a mean residual time *R* until the end of the current packet transmission and must also wait for the transmission of the mean number of packets N_k currently in the *k*th queue ahead of it. In addition, the *i*th packet must also wait for the transmission of all packets in the system (not in queue *k*) with timestamps (virtual finish times), that are smaller than the timestamp assigned to the *i*th packet. The mean number of such packets in each queue *j* in the system is denoted by P_j (see Figure 1). Thus, the mean queuing delay (waiting time in queue *k*) for the *i*th packet is given by,

$$W_k = R + \frac{1}{\mu}\left(N_k + \sum_{j \in J} P_j\right), \quad (1)$$

where we used the fact that buffer occupancy is independent of individual packet service times, and we also used *J* to represent the set of all flows supported by the scheduler except flow *k*. In other words, $J = \{j \in K : j \neq k\}$. We can evaluate the mean residual time *R* by a graphical argument as in [13] to obtain $R = \lambda \overline{T^2}/2$. Using Little's law, $N_k = \lambda_k W_k$, we get,

$$W_k = \frac{1}{2}\overline{T^2}\lambda + \rho_k W_k + \frac{1}{\mu}\sum_{j \in J} P_j \quad (2)$$

The only unknown quantities in (2) now are the values of P_j. To find these values, let us consider a single queue $j \in J$ in the system. Assume that this queue has a connection potential $\nu_j(a_k^i)$ at the time a_k^i of packet's *i* arrival. Assume also that the connection potential of queue *k* was $\nu_j(a_k^i)$ at that time. Noting that the P_jth packet in the *j*th queue should have a timestamp that is smaller than the *i*th packet (in queue *k*) so that the WFQ scheduler can serve it first, we get:

$$E[\nu_j(a_k^i)] + P_j \frac{\overline{M_j}}{r_j} \le E[\nu_k(a_k^i)] + (N_k+1)\frac{\overline{M_k}}{r_k}.$$

Also, packet $M_j + 1$ in the jth queue should have a timestamp that is larger than the ith packet in queue k. Hence, we can write, that

$$E[\nu_j(a_k^i)] + (M_j+1)\frac{\overline{M_j}}{r_j} \ge E[\nu_k(a_k^i)] + (N_k+1)\frac{\overline{M_k}}{r_k}$$

Rearranging this expression, we get the following upper and lower bounds on M_j , respectively,

$$P_j \le \min\left((N_k+1)\frac{r_j}{r_k} + \overline{\delta_{kj}}\frac{r_j}{\overline{\overline{T}}C}, N_j \right),\ j \in J \tag{3}$$

$$P_j \ge \min\left(\max\left((N_k+1)\frac{r_j}{r_k} + \overline{\delta_{kj}}\frac{r_j}{\overline{\overline{T}}C} - 1, 0 \right), N_j \right), \tag{4}$$

$j \in J$

where $\delta_{kj} = \nu_k - \nu_j$ is the difference in the connection potentials of flows k and j. Lets take a note, that in (4) we limited the lower bound on P_j to a minimum value of zero. Otherwise, such a lower bound may turn out to be negative, which happens when N_k is small and the conversion factor is $r_j / r_k << 1$. Such a negative value of P_j is not practically acceptable. Also lets take a note, that we have limited the P_j upper and lower bounds in (3) and (4) to a maximum value equal to the mean number of packets N_j in the jth buffer, which is another practical limit we have to maintain.

2.1.1 The Upper Bound on Mean Packet Delay

The challenge we face in trying to solve for the upper bound on mean packet delay is that we need to choose the minimum of two quantities in (3) for each flow $j \in J$ before being able to substitute it into (2). Such a decision cannot be made without prior knowledge of the actual N_j (or W_j) values, which are the unknowns we are seeking to find.

In order to avoid such a problem we noticed, that as far as the upper bound on mean packet delay is concerned, using the first expression on the right hand side of (3) instead of the minimum does not actually affect the correctness of the upper bound on P_j , although it might slightly weaken its tightness. Since such an expression is not dependent upon the value of N_j , we can drastically simplify the derivation process, which gives us the following upper bound on mean packet delay, which is based on (2), (3) and the upper bound on $\overline{\delta_{kj}}$,

$$W_k \le \frac{\frac{1}{2}\overline{T^2}\lambda + \sum_{j\in J}\left[\frac{1}{\mu}\frac{r_j}{r_k} + \rho'_j \overline{\psi_j}\frac{r_j}{C}\right]}{1 - \rho_k \frac{\sum_{j\in K} r_j}{r_k}} \tag{5}$$

where W_k is mean packet delay; $\overline{T^2}$ – service time; λ – arrival rate; μ – mean service rate; r – conversion factor; ρ – utilization of the output link; C – output link capacity and fairness bound $\overline{\psi_j} = \overline{L_j} / r_j$.

On the other hand, using the improved upper bound on $\overline{\delta_{kj}}$ expression (5) transforms into,

$$W_k \le \frac{\frac{1}{2}\overline{T^2}\lambda + \sum_{j\in J}\left[\frac{1}{\mu}\frac{r_j}{r_k} + \left(\rho'_j \overline{\psi_j} - \rho'_k \overline{\psi_k}\right)\frac{r_j}{C}\right]}{1 - \rho_k \frac{\sum_{j\in K} r_j}{r_k}} \tag{6}$$

2.1.2. The Lower Bound on Mean Packet Delay

In order to find the lower bound on the mean packet delay W_k requires a similar approach to that of finding the upper bound. For the lower bound, however, we cannot just substitute the first expression in (4) instead of the required minimum since this is mathematically incorrect. However, we can still derive a simple equation for the lower bound similar to that of (5) by setting the minimum in (4) to zero all the time, which gives the following simple lower bound on mean packet delay:

$$W_k \ge \frac{\frac{1}{2}\overline{T^2}\lambda}{1-\rho_k} \tag{7}$$

where W_k is mean packet delay; $\overline{T^2}$ – service time; λ – arrival rate; ρ_k – utilization of the output link.

Obviously, this is not the best possible lower bound on mean packet delay. However, this lower bound is reasonably tight in almost all practical cases.

Now, using the upper and lower bounds on mean packet delay in (5), (6) and (7) we can derive the corresponding upper and lower bounds on mean buffer occupancy using Little's law, which states that $N_k = \lambda_k W_k$.

2.1.3. Properties of the Delay Bounds

An important observation we can make about the M/G/WFQ delay bounds derived earlier in Section 3 is that both the upper and lower bounds increase in inverse proportion to $1-\rho_k$ (or $1-\rho'_k$). This means that the mean packet delay in our system is expected to be dramatically increased as the utilization factor approaches unity, or at least that would be the behavior of the delay bounds in such a condition. Comparing this to an M/G/1 queuing system, we notice the same exact behavior for the mean packet delay versus utilization. This behavior is actually a general characteristic of almost any queuing system one might encounter.

Now let us consider which operating condition would result in tighter bounds on mean packet delay. It is easy to see from (5) and (7) that the difference between the upper and lower bounds is mainly dependent on a summation factor including the fairness bounds $\rho'_j\overline{\psi_j}$ for all $j \in J$. This means that tighter bounds are expected in the following situations: 1) when the number of flows K supported by the scheduler is smaller, 2) when the fairness bound of the WFQ system is tighter and 3) when the load on queue k, measured by ρ_k, is smaller. In Section 3 we propose simulation results of this model.

2.2. M/G/1/p3 Stochastic Analysis

This model has infinite queue and, in our case, it can serve the traffic with three priorities: Assured Forwarding (AF), Expedited Forwarding (EF) and Best-Effort Forwarding (BF), for example it can be video, voice and data streams. In other cases this model can work up to five priorities data streams. So, all three data streams pass into one infinite queue, then they take a service and fall into output link (see Figure 2). Following our usual notation, we assume that the arrivals come from a Poisson process with rates λ_1, …, λ_n, and that the mean service time is [14]:

$$\mu = \frac{1}{T}, \qquad (8)$$

where

$$T = \frac{M}{C}, \qquad (9)$$

and M is length (in bits) of the data packet and C – the output link capacity.

The total arrival rate is

$$\lambda = \sum_{n=3}^{n} \lambda_n, \qquad (10)$$

where n is priority class and, in our case, it can vary from 1 to 3.

Expected number ρ being served at any time is:

$$\rho = \frac{\lambda}{\mu}, \ \rho < 1 \qquad (11)$$

Figure 2. Arrivals at the M/G/1 system.

The basic parameters for M/G/1/p3 model we know from [14]. The additional coefficients we can express in this way.

For priority class 1:

$$A_1 = \frac{\lambda_1}{\mu}, \ \lambda_1 > 0.$$

For priority classes with lower priority:

$$A_n = A_m + \frac{\lambda_n}{\mu}, \ \lambda_n > 0,$$

where m is class with higher priority and, in our case, it varies from 1 to 3.

$$C = \frac{\lambda}{2} \cdot V + \frac{1}{\mu^2},$$

where V is a variance of the service time.

Mean delay of the data packet for every priority class can be calculated as follows.

Expected time in the system:

$$W_n = W_{qn} + \frac{1}{\mu}, \ \lambda_n > 0. \qquad (12)$$

Expected waiting time in the queue for priority class 1:

$$W_{qn} = \frac{C}{(1-A_n)}. \qquad (13)$$

Expected waiting time in the queue for priority classes with lower priority:

$$W_{qn} = \frac{C}{(1-A_m)\cdot(1-A_n)}, \qquad (14)$$

where m is class with higher priority and, in our case, it varies from 1 to 3.

This model is flexible and it can be used to calculate mean delay of the data packet for two, three, four or five priority classes.

3. Modeling results

Several simulations are performed to validate the results of the M/G/WFQ and M/G/1/p3 analysis presented in Section 2. In this experiment, a WFQ server in the M/G/WFQ model and server in M/G/1/p3 model with a total output link capacity C = 1,1 Mbit/s supports three incoming Poisson streams with different priorities. The reserved rates for the different connections with different priorities in M/G/WFQ model are: r_j = 0,5 Mbit/s; r_k = 0,4 Mbit/s; r_K = 0,2 Mbit/s. The incoming sources rates in M/G/WFQ and M/G/1/p3 models are: $\lambda_j = \lambda_k = \lambda_K$ and they varies between 0,1 and 6 erlangs (erl.). The mean service rate μ varies between 6 and 20. Using expressions (6) and (7) for M/G/WFQ model we obtain upper and lower bounds of packet delay in the system. In Figure 3 the mean delay of the packets in the system is shown only for packets with high and low priority. Using expression (12) for M/G/1/p3 model we obtained the mean delay of the packets in the system. In Figure 3 the mean delay of the packets is shown only for packets with high and low priority, too.

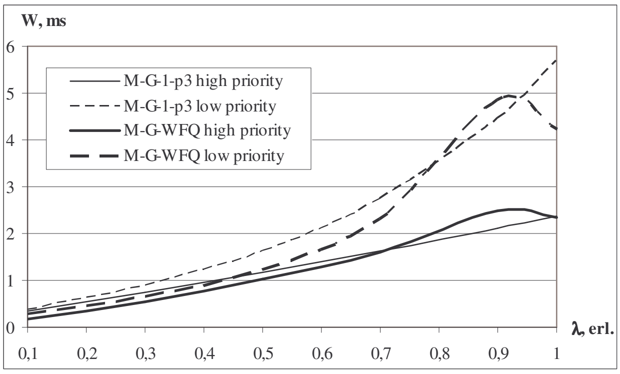

Figure 3. The mean delay of packets in the system, when $\mu = 6$.

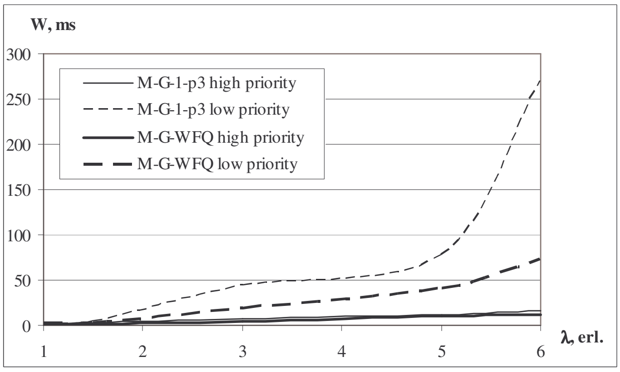

Figure 4. The mean delay of packets in the system, when $\mu = 20$.

In Figures 3 and 4 we can observe very interesting dependence, when the arrival rate is about 0,9 erl., mean delay of the data packets with high and low priority of class in the system with M/G/WFQ queuing discipline is higher, than in the system with M/G/1/p3 queuing discipline. But, when the arrival rate increases and reaches 6 erl., mean delay of the data packets with low priority of class in the system with M/G/1/p3 queuing discipline becomes extremely high, comparison with system with M/G/WFQ queuing discipline. It was not in vain, because the algorithm of WFQ queue scheduling discipline using resources of processor and, when arrival rate is low, mean delay of the data packet is higher, than in the system without this discipline. But, when arrival rate is high WFQ queue scheduling discipline with its fair distribution of bandwidth for all priorities of rate, provides smaller mean delay both, and data packets with high priority of class, and especially data packets with low priority of class.

In the second experiment, using the same initial data and mathematical expressions, as in the first experiment, we proved which factor of difference is between the mean delay of the data packets in the system with high and low priority of class (see Figures 5 and 6).

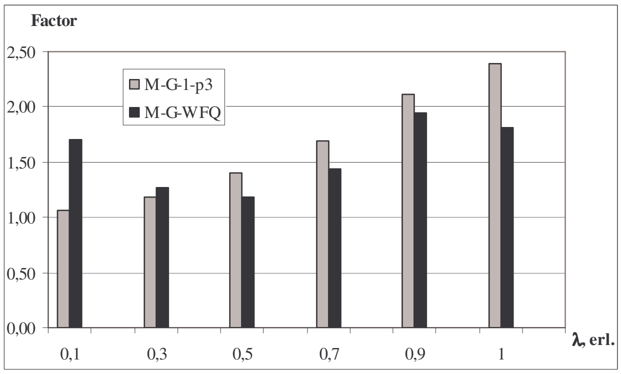

Figure 5. Factor of difference between the mean delays of the packets in the system with high and low priority, when $\mu = 6$.

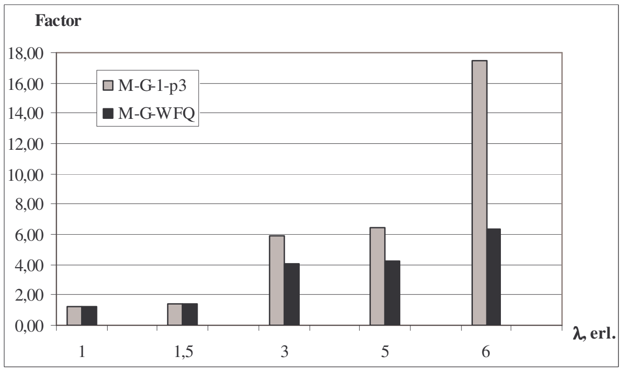

Figure 6. Factor of difference between the mean delays of the packets in the system with high and low priority, when $\mu = 20$.

In this experiment we observe very similar tendency of mean delay of the data packets in the systems with M/G/1/p3 and M/G/WFQ queuing disciplines, as it was in the first experiment. Only in this experiment, we computed the factor of difference between the mean delay of the data packets with high and low priority of class in the systems with M/G/1/p3 and M/G/WFQ queuing disciplines. The model with WFQ queue scheduling discipline and the data packets with high and low priority of class serves more flexible and the difference between them of the mean delay is smaller, than in the system with M/G/1/p3 queuing model.

In the last of our experiment, a WFQ server in the M/G/WFQ model and server in M/G/1/p3 model with a total output link capacity C = 1,1 Mbit/s supports three incoming Poisson streams with different priorities. The reserved rates for the different connections with different priorities in

M/G/WFQ model are: r_j = 0,5 Mbit/s; r_k = 0,4 Mbit/s; r_K = 0,2 Mbit/s. The incoming source rates in M/G/WFQ and M/G/1/p3 models are: $\lambda_j = \lambda_k = \lambda_K = 0{,}5$ erl. The mean packet length M varies between 100 and 1000 bits per packet. Using expressions (6) and (7) for M/G/WFQ model we obtain upper and lower bounds of packet delay in the system. In Figure 7 the mean delay of the packets in the system is shown only for packets with high and low priority. And using expression (12) for M/G/1/p3 model we obtained the mean delay of the packets in the system. In Figure 7 the mean delay of the packets in the system is shown only for packets with high and low priority, too.

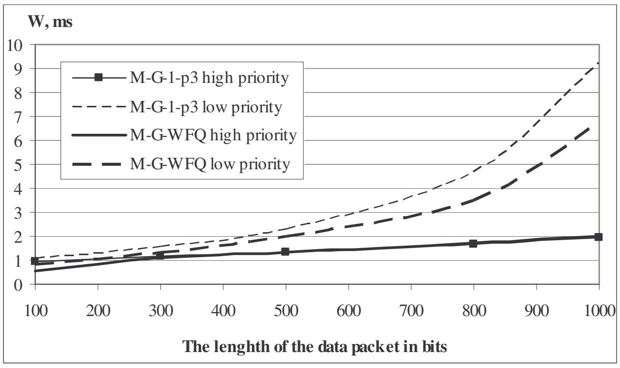

Figure 7. Delay of the traffic in M/G/1/p3 and M/G/WFQ models, vs. the length of the data packet, when λ = 0,5 and C = 1,1 Mbit/s.

In this experiment we observe the same tendency of the mean delay of the data packets with high and low priority of class, as it was in the first and second experiments. Practically there is no difference between the mean delay of the data packets with high priority of class in both systems. But long data packets with low priority of class in the system with M/G/1/p3 queuing discipline have higher mean delay, than it was in the system with M/G/WFQ queuing system.

4. Conclusions

- Experiments, described in this paper, confirm, that more advantageous and flexible is M/G/WFQ queuing model. But in real system the realization of the M/G/WFQ queuing discipline is more complicated.
- In design a real system the load of a system have to be taken into account. Thus for a load, which is about 1 erl., it is preferable to apply M/G/1/p3 queuing discipline, but in the case of variable load or higher than 1 erl., the M/G/WFQ queuing discipline would be more acceptable.
- Mathematical models, developed in the paper, evaluate only delay of the data packets and do not take into consideration packets loss, but can be extended easily, incorporating variable K, with restrictive size of a buffer.

References

[1] Clark, D., Fang, W., Explicit Allocation of Best Effort Packet Delivery Service, *Internet draft*, Sept. 1997, pp. 27-32.

[2] Chen, S., and Nahrstedt, K., An overview of quality-of-service routing for the next generation high-speed networks: Problems and solutions. *IEEE Network, Special Issue on Transmission and Distribution of Digital Video*, November/December 1998, pp. 22-35.

[3] Bala, K., et al., Congestion Control for High-speed Packet Switched Networks, *Proceedings of Infocom '90*, April 1990, pp. 520-526.

[4] Armitage, G., Quality of services in IP networks.-Indianapolis: *MTP*, 2001, pp. 309-317.

[5] Dekeris, B., Adomkus, T., Evaluation the Quality of IP Network Using Differentiated Services// Telecommunications and Electronics vol. 3. – *Scientific Proceedings of Riga Technical University*, 2003, pp. 11-16.

[6] Dekeris, B., Narbutaite, L., Adomkus, T., Analysis of the Differentiated Services rates scheduling disciplines// *Electronic and electrotechnics*.-Kaunas: Technology, No. 6 (55), 2004, pp. 34-39.

[7] Demers, A., Keshav, S., and Shenker, S., Analysis and Simulation of a Fair Queueing Algorithm, *Journal of Internetworking Research and Experience*, October 1990, pp. 3–26.

[8] Nichols, K., Jacobson, V., Zhang, L., A Two-Bit Differentiated Services Architecture for the Internet, November 1997, pp. 35-45.

[9] Roberts, J., Mocci, U., Virtamo, O. (eds), Broadband Network Teletrafic, Lecture Notes in Computer Science, vol. 1155, Springer, 1996, pp. 17-21.

[10] Dekeris, B., Rindzevicius, R., Narbutaite, L., Adomkus, T., Evaluation of the quality of Differentiated Services// *Electronic and electrotechnics*.-Kaunas: Technology, No. 6 (48), 2003, pp. 67-72.

[11] Dovrolis, C., Stiliadis, D., Ramanathan, P., Proportional differentiated services: Delay differentiation and packet scheduling// *Proc. ACMSIGCOMM,* Aug. 1999, pp. 109-119.

[12] Semeria, C., Supporting differentiated service classes: queue scheduling disciplines//*Juniper networks*, 2001, pp. 11-14.

[13] Bertsekas, D., P., and Gallager, R., DataNetworks, *Prentice Hall*, Second Edition, 1991, pp. 36-42.

[14] Virtamo, J., Queuing theory, 2001, pp. 8-34.