Analysis of Efficiency of Protocols, Guaranteeing QoS on VoIP Technologies

Nikolai V. Penev[1] and Vassil M. Kadrev[2]

Abstract – **In the packet switching networks of TCP\IP stack of protocols the main problem is to guarantee the quality of services for services provided in real-time. The inherent particularity of these networks is their aptitude to overloading, which brings about essential distortion of packetized voice facilities. On the base of the developed mathematical model, different varieties of protocols of standby have been analyzed. The evaluation of improving the quality of services has been made on one hand, and on the other hand the reducing efficiency of the network has been estimated.**

Keywords – **IP, QoS, Traffic**

I. Introduction

In the Internet layer of the network a mechanism for inverse request in the case of finding mistakes of the accepted information is not provided. Possibility for inverse request is realized in the TCP layer that is inadmissible for voice and video services. It brings the necessity to introduce mechanisms to ensure the quality of services with the realization of VoIP. The essential parameters of quality of services, which affect and which are guaranteed are: end-to-end delay, jitter, probability of the loss of the package on IP layer because of a mistake. These mechanisms present protocols of standby, anti-jitter buffers and algorithms for correcting the mistakes in the receiver without an inverse request.

With the occurrence of the idea about building a digital network with integrated services on the base of IP, the necessity of using methods to ensure the quality of services (QoS) has appeared. The switching units, terminals and centers for operation and control in the IP network are specialized computers that can perform the algorithms of precise information processing, as the mentioned criteria.

Most generally, QoS contains three criteria of quality: delay, jitter and losses of packets. The various kinds of services are critical to various parameters and have different criteria. With its developing as a network with integrated services (approach ISA - Integrated Services Architecture of IETF – Internet Engineering Task Force), the IP network, which has occurred on the purpose of transmitting data without warranties of the particular packets delivery to the receiver, has already involved mechanisms (protocols) guaranteeing the quality of services critical to the mentioned criteria of QoS.

[1]Nikolai V. Penev is with the War Academy, 82 E. Georgiev, Sofia, Bulgaria, E-mail: penevnv@yahoo.com

[2]Vassil M. Kadrev is with the Higher School of Transport, 158 Geo Milev, 1754 Sofia, Bulgaria, E-mail: kadrev @ internet-bg.net

II. Analyses of QoS in IP Network

The mechanisms of providing guaranteed delay of the traffic critical to this criterion, which are known at present, are three: RSVP (Resource Reservation Protocol), MPLS (Multi-Protocol Label Switching) and DS (Differentiated Services). These are mechanisms without which services such as telephone, video, multimedia and interactive data exchange could not be offered in the IP integrated environment with a guaranteed norm of delay. The problems of the jitter in the IP network, toward which the telephone service is particularly critical, can be solved by anti-jitter buffers in the routers, from which it can be read at a constant rate. With a sufficient bit rate of the lines and switch centers, the routers in the IP network, this mechanism is effective, as the increase of the delay with a big volume of anti-jitter buffers has to be taken into consideration. The TCP layer in the IP network guarantees a sure information delivery with which great delay of the feedback and re-transmitting the data mistaken have been introduced. It could result in exceeding the admissible norms of information delay. That can be avoided by introducing FEC (Forward Error Correction) mechanisms performing marking of the mistaken bits by finding noise-resistant code and correlation data analysis in one or a number of packets aiming at approximate restoration of the mistaken bits. Due to the continuous character and the even changes of the telephone and video signals in time FEC guarantees small losses (sound and picture distortion) without introducing additional delay for that purpose. The experience accumulated with the operation of IP networks with mixed integrated traffic shows that the anti-jitter and FEC mechanisms are not the main part for the great “end-to-end” information delay. In order to meet the norms of the delay of the telephone and video services, effective mechanisms for mixed traffic control in the IP network are necessary.

IETF defines two main mechanisms of the IP traffic control guaranteeing QoS (delay) by the priorities of the traffic flows of the various services. They are the mechanisms of integrated services ISA and of DS. Both mechanisms use field TS 8 bits in the title of IP-packet for specifying the priority. In the field of options data contained can give a possibility to introduce a dynamic priority (moment of packet generating, time of packet life).

The main purpose of ISA mechanism is to recognize packages of the priority service flow (requiring little delay) and transmit them without waiting in queues aiming at delay not exceeding the admissible one. This mechanism is specified in protocol RSVP where two classes of integrated services are defined: GS – Guaranteed Services and CL – Controlled Load.

GS is a protocol for forwarding packets with a reserved traffic capacity that results in little usage of the network resources. Thus GS guarantees little delay and absence of losses (telephony, video, multimedia). It should be outlined as a protocol disadvantage that it throws off the packets after the deadline of their life and arrived after the admissible time of delay.

CL mechanism allows the router-realizing RSVP to process the service package flow as ordinary IP datagrams (Best Effort), to control the service packages and to increase their priority in the network with the increase of the package delay (decrease of the life left) up to the moment, i.e. to decrease their stay in the lines. In the latest versions of CL, with the impossibility to guarantee the admissible delay of the packets, additional sessions for the same service are established, thus the delay sharply drops down to be included in the norms.

The second mechanism of control and guaranteeing the QoS of IP-traffic is DS based on the so-called PHB (Per Hop Behavior) routing. It includes as separate mechanisms EF (Expedited Forwarding) and AF (Assured Forwarding). With this mechanism a number of classes of packets with various admissible delay are defined (according to the services) for which a respective buffer space and rate of forwarding are reserved. In each class of packets there are a number of priorities as the packets of higher priority are kept and the ones of lower priority are thrown off with network overloading. It is necessary to apply DS with using wide-banded services and it depends on the rates of forwarding. The mechanism of detecting and preventing the conjunctions RED (Random Early Detection) is also maintained. It detects the random arrived packets on the base of the analysis of the title part of TCP datagram and throws them off if they are not addressed to the ports of addresses that prevail in the queue of different priorities. Thus the buffers (queues) with different priority in the router are kept semi-filled up with packets that also decrease their delay.

III. Model, Quantitative Ratios and Results

The presence of multi-priority flow of incoming asking for service is common for all mechanisms of the IP traffic control. For an IP network with integrated services including subscribers and transit devices (routers), the delay of the package "end-to-end" for a homogeneous flow of packages (for each service) is determined as T_w:

$$T_w = 2t_{sl} + \sum_{i=1}^{d} (t_w + t_{pr} + t_{tr}) - t_{tr} \ , \qquad (1)$$

where t_{sl} – average time for transfer by subscriber line; d – average number of hops; t_w – average waiting time in the queue of a node; t_{pr} – average processing time of the packets on the router; t_{tr} – average time for transfer by trunk.

The components of T_w: t_{sl} and t_{tr}, which use DSL technology can be assumed as constants and only t_w and $d$remain to influence on T_w.

The paper presents a model GS mechanism of RSVP protocol examined with which RSVP is provided according to the strict requirements to work in real time a guaranteed frequency band, as well as little delay of "end-to-end" packets and absence of packets loss as a result of their arrangement in queues. Due to this, with serving one and the same GS flow, each router in the network (which requires separate functions for that in network control) has to distribute the frequency band and the respective buffer space according to the priority of entering packets of various services. As a great number of traffic flows along the lines of great traffic capacity enter the router, it can be assumed that the input traffic to the router is punch [5]. The particular router is a system of mass service with waiting and service discipline with priorities. Under these conditions T_w can be determined for each service. It is known that teletraffic system M/D/1 is characterized by average time of waiting twice less than that of system M/M/1. The examinations will be carried out using teletraffic system M/M/1 and stipulating that the results with determining the average time of waiting for a system of mass service M/M/1 are the upper limit for the average time of waiting as M/D/1 is in practice. It is possible to determine the dependencies of service quality by the rates of lines, the size of packets and the number of transit sections. The determination of the results of the priority service influence on the quality of service for a service device can be also used for the network as a whole. The position of each package in the queue of the service device will be a variable function of time having in mind the possibility that a packet of higher priority can enter the queue. On one hand, the system of priorities can be defined according to the priority: if it is absolute (fixed) or depends on a given function, and on the other hand – if processing of the packet served is interrupted at the moment of the entrance of a packet of higher priority and later is restored (at the moment of breaking). For a priority system of service (of P priorities as p = 1, 2, 3, ... P) with a fixed priority of serving with the entrance of a packet of a higher priority there is given (with $0 \leqslant \rho < 1$, i.e. absence of losses)[2]:

$$T_{wp} = \begin{cases} \dfrac{\frac{\rho_p}{\mu_p} + \sum\limits_{i=p+1}^{P} \rho_i \cdot \left(\frac{1}{\mu_p} + \frac{1}{\mu_i} \right) + \sum\limits_{i=p+1}^{P} \rho_i \, T_i}{1 - \sum\limits_{i=p}^{P} \rho_i}, & \text{for } p \geqslant j \\ \infty, & \text{for } p < j \end{cases} , \qquad (2)$$

where j is minimum and integer and $\sum\limits_{i=j}^{P} \rho_i < 1$.

Type of services [4] are: voice, video, interactive data (data 1), download (data 2). The parameters of the arriving traffic and of the data processing [4] are shown in Table 1.

Average servicing time $1/\mu$ depends from average size of the packet l and average send rate c. Average value of input traffic ρ depend from average input rate of the packets λ and average servicing time $1/\mu$=const. For input traffic

$$\sum_{i=1}^{4} \rho_i = \rho < 1. \qquad (3)$$

Table 1. Parameters of servicing

Service	Video	Voice	Interactive data	Download
Average servicing time, $1/\mu = const$	$\frac{1}{\mu_1} = \frac{l_1}{c_1}$	$\frac{1}{\mu_2} = \frac{l_2}{c_2}$	$\frac{1}{\mu_3} = \frac{l_3}{c_3}$	$\frac{1}{\mu_4} = \frac{l_4}{c_4}$
Input traffic, ρ	$\rho_1 = \frac{\lambda_1}{\mu_1}$	$\rho_2 = \frac{\lambda_2}{\mu_2}$	$\rho_3 = \frac{\lambda_3}{\mu_3}$	$\rho_4 = \frac{\lambda_4}{\mu_4}$
Priority of servicing, p	4	3	2	1

Priority of servicing increases with p. Send rate of subscriber line is $c_{sl} = 8$ Mbps, send rate of trunk is $c_{tr} = 40$ Mbps, average number of hops is $d = 4$, average processing time on the router t_{pr} is approximately 0.

Values of output data for determining of the servicing parameters [4] are shown in Table 2.

Table 2. Type of traffic sources and parameters of servicing

	Traffic			
	Video	Voice	Interactive data	Download
Average norm of delay, ms	60	160	600	2000
Average norm of delay for one hope, ms	15	40	150	500
Average length of the block, kbit	burst traffic	burst traffic	400	4000
Average length of the packet, kbit	8	20	200	400
Average time for transfer by trunk, ms	0,05	0,5	5	10
Average time for transfer by subscriber line, ms	0,25	2,5	25	50
Total average transfer time end-to-end for d = 4, ms	0,7	7	70	140
Average admissible waiting time in the queue of a node, ms	14.825	38.25	130	410

The average admissible waiting time in the queue of one processing device is the difference between the average admissible delay for each service and the total average transmission time end-to-end with d sections.

Table 3. Structure of traffic sources

Percentage of the arriving traffic, %	Video	Voice	Interactive data	Download
Relation 1	10	5	15	70
Relation 2	20	10	35	35
Relation 3	49	1	35	15
Delay, ms	w4	w3	w2	w1

The results obtained for the delay values depending on the input traffic ρ and at concrete proportion of the input traffic

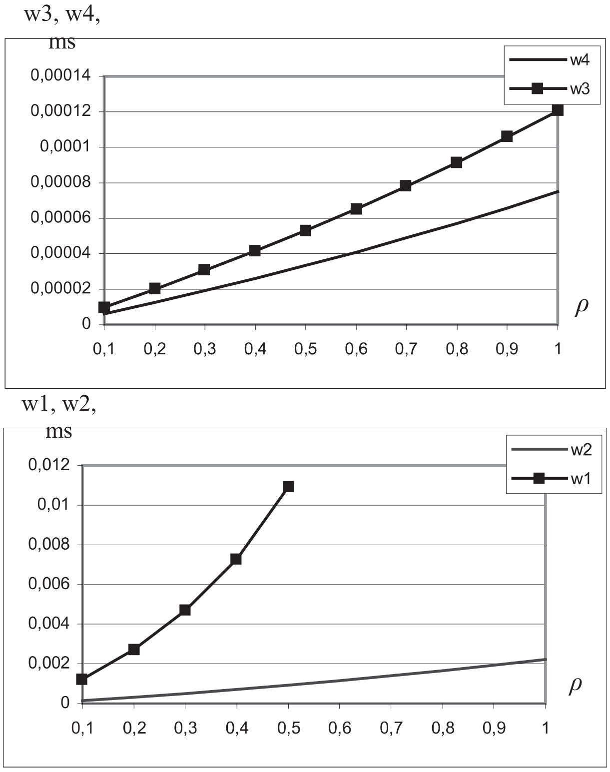


Fig. 1. Delays for relation 1 (Table 3) of input traffic for: video (w4), voice (w3), interactive data (w2), download (w1)

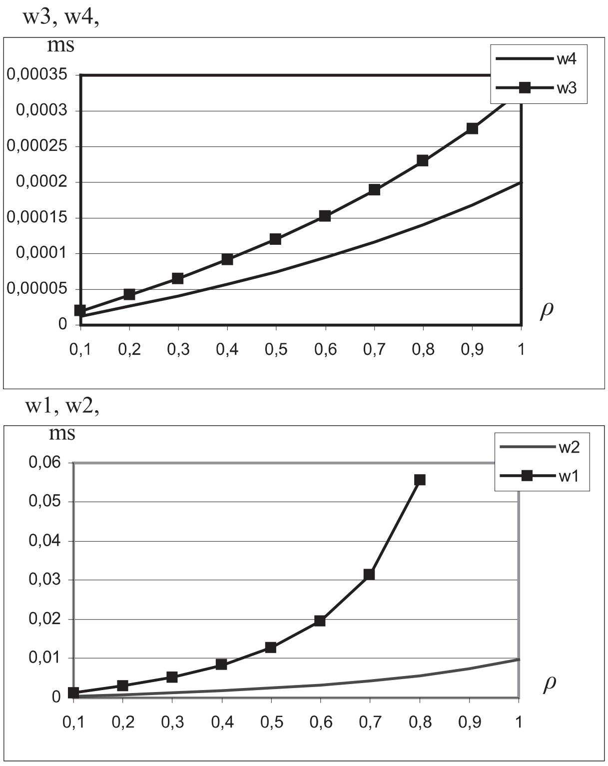
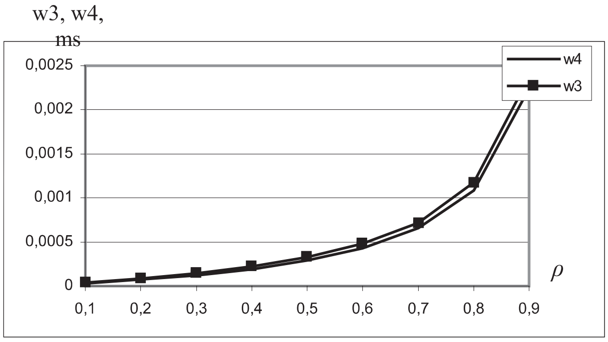


Fig. 2. Delays for relation 2 (Table 3) of input traffic for: video (w4), voice (w3), interactive data (w2), download (w1)

from the separated services with priority processing (Table 3) are shown on Figs 1, 2 and 3.



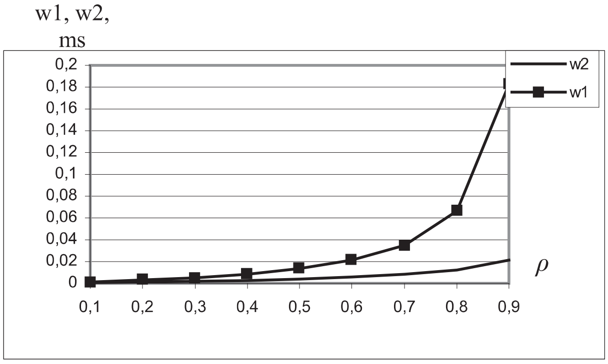


Fig. 3. Delays for relation 3 (Table 3) of input traffic for: video (w4), voice (w3), interactive data (w2), download (w1)

IV. Conclusion

From the results obtained it can be seen that with the given parameters of the lines, devices and the traffic, the delay which the separate services receive is within the norms for a wide range of change for the values of these parameters: the size and the proportions of the arriving traffic (loading of the router) and the size of the packets. This is true when there are faster lines and approximately all the norm of delay may be used for waiting in queues.

By increasing the length of the packets for the traffic with high priority (voice, video), average waiting time sharply decreases. This results quickly not within the norms of delay, for loading of the router over 0,7.

The research is made for unlimited number of nodes as their connection is 30.

The results for mechanism CL of RSVP can be obtained in a similar way with the dynamic priorities of queue servicing.

References

[1] Kleinrock L. Theory of queuing systems – I, Moscow, Mashinostroenie, 1979.

[2] Kleinrock L. Communication Nets (stochastic message flow and delay). Mcgraw-Hill.

[3] Models of teletrafic theory in telecommunications and electronics. Proceedings - A. D. Charkevich, V. A. Garmash. Moscow, Science, 1985.

[4] Parvanova N. Broadband ISDN. NIIS-CENTI, Sofia, 1997.

[5] Alcatel Telecommunications Review, No. 2, 1999.

[6] Gantchev I. Computer Networks and Communications. Plovdiv., University of Plovdiv, 1999.

[7] Boyanov K. and all. Computers networks and Internet. Sofia, CLPOI-BAS, 1998.

[8] www.ietf.org.