

IP PACKET QUEUING DISCIPLINES AS BASIC PART OF QOS ASSURANCE WITHIN THE NETWORK

Saša Klampfer, Jože Mohorko, Žarko Čučej

University of Maribor, Faculty of Electrical Engineering and Computer Science,
Maribor, Slovenia

Key words: waiting queues, queuing mechanisms, delay, dropped traffic, queuing, quality of service

Abstract: This paper represents results of study how different queuing methods influences on service quality as basic part of QoS system in wired networks. We briefly introduces six queuing methods, where two of them are high sophisticated (WFQ – Weighted Fair Queuing and CBWFQ – Class Based Weighted Fair Queuing) used in Cisco's routers, the third one (CQ – Custom Queuing) is used as basic method for comparison with all others sophisticated methods. All mentioned queuing principles are tested in OPNET Modeler simulation tool. During many simulation runs we get the best results with combined PQ-CBWFQ method which is also presenter of a low latency queuing. The satisfying results are also obtained with CBWFQ method where high priority traffic undisturbed travel through the network, but majority low-priority traffic must wait. This is also the main reason why we obtain highest delays within network in case of CBWFQ method. When are present only two different traffic flows, WFQ is the best choice concerning delays and dropped traffic. All other queuing disciplines don't fulfill our expectations.

Mehanizmi paketnega uvrščanja kot osnovni pogoj za zagotavljanje kvalitete storitev v omrežju

Ključne besede: čakalne vrste, mehanizmi uvrščanja, zakasnitve, zavržen promet, uvrščanje, kvaliteta storitev

Izvleček: V članku predstavljamo rezultate študije, ki se nanaša na opazovanje vpliva posameznih metod uvrščanja paketov v čakalne vrste na kakovost storitev v omrežjih. Mehanizmi uvrščanja paketov v čakalne vrste ob enem tudi predstavljajo ključen del sistema QoS. Na kratko predstavljamo šest različnih metod uvrščanja, med katerimi sta dve visoko sofisticirani (WFQ – utežnostno pravično uvrščanje in CBWFQ – utežnostno pravično uvrščanje na osnovi razredov) uporabljeni tudi v Cisco-vih usmerjevalnikih. CQ (ang. Custom Queuing) je uporabljena kot osnovna metoda uvrščanja in nam v zaključni fazi predstavlja referenco pri primerjavi rezultatov naprednejših mehanizmov uvrščanja z osnovnimi načini uvrščanja, kamor spada CQ. Vse čakalne mehanizme smo preizkusili v simulacijskem orodju OPNET Modeler. Tekom številnih simulacij smo najboljše rezultate prejeli s kombinirano metodo PQ-CBWFQ, ki je ob enem predstavnik čakalnih vrst z najnižjimi zakasnitvami. Zadovoljive rezultate smo prav tako prejeli z CBWFQ metodo, kjer visoko-prioritetni podatkovni tokovi nemotene tečejo skozi omrežje, vendar pri tem morajo nižje-prioritetni tokovi, ki predstavljajo glavnino vsega prometa, čakati. To je ob enem tudi glavni vzrok za dobljene visoke zakasnitve znotraj omrežja v primeru uporabe CBWFQ načina uvrščanja. Študija je pokazala, da v primeru prisotnosti samo dveh različnih podatkovnih tokov, zadovolji naše potrebe že WFQ metoda uvrščanja, predvsem v smislu zavrženega prometa. Vse ostale metode uvrščanja naših pričakovanj niso izpolnile.

1. Introduction

Demand for internet quality of service (QoS) in wired networks has been growing rapidly over recent years, especially in networks of large companies and organizations, with limited wired link bandwidth, where are in use many different applications, such as VoIP, video on demand, video conferencing, FTP sessions, web browsing, E-mail etc. Such organizations often have internet link which also serves hundreds or even thousands of internal network computers, named nodes. In those cases we must be able to provide the sufficient bandwidth and low latency for high priority applications, for example VoIP. Without distinguishing between the different priority applications we cannot be able to provide service quality on the desired level. Recent, strong growth of bandwidth hungry peer-to-peer applications has made this task especially difficult and complex. One of the important parts of QoS service are also queuing methods which manages traffic flows in the normal operation conditions and also in the situations when the network becomes overloaded. Our goal, in this research, is to explore how each of mentioned queuing meth-

ods influences on dropped packets and delays in network when we use high priority applications in cohabitation with all other low level priority applications. The effectiveness of each queuing method will be introduced in continuation. Queuing plays an important part within whole QoS system. This area presents issue for many research processes. For example, /1/ describes influence of different queuing methods on the common utilization network characteristic, meanwhile /7/ represents providing QoS system for multimodal systems in wired networks. As wired networks, are in last few years also very important wireless networks where is also need upon quality of service. Such solutions are described in /6/. Congestion control with use of queuing mechanisms for Byte-Stream networks is described in /10/. Regarding well spread VoIP traffic within wired and also wireless networks, are many research procedures regarding QoS concentrated into already mentioned area.

Mostly used queuing methods are detail described in second and third section, meanwhile fourth section represents OPNET modeler simulation tool, which is used for proving

our assumptions, and decisions. Fifth section gives a description of wired network simulation structure, in sixth section are presented simulation results meanwhile seventh section gives a short conclusion with decisions and ascertainties, afterwards follow directions and points of departure for further work and researches in this area.

2. Queue

Wide-area networks (WAN) consist from wide segments where each of them is again combination of smaller segments – local-area networks (LAN). Similar topology is at local area networks level, where each segment is linked to network stations using basic network components such as routers, switches, bridges. Each of mentioned elements contain queuing mechanisms which carries each individual arrived packet according to priority number, entered into the type of service (ToS) field of IP packet header. Most sophisticated solutions on the market contain more than one queues and each of them is intend to provide place for different traffic priority flows, respectively each of them belongs to specific priority class.

Firstly, the queuing mechanisms will be explained on the simplest example of single queue, which is known as FIFO (first in, first out). FIFO queue is data structure which acts as settled list. On one side data comes into such list, and on the other side they come out in the same sequence as they came in. Operations which are executed upon waiting queue are: element insertion, element removal, which element is placed at the start place of queue, occupies places and free places in the queue /1/.

3. Queuing mechanisms

Traffic link congestions are today one of the most bothering things because they affect on quality of service of other applications presented in network. One way for handling congestions, represents queuing algorithms, which sorts traffic according with queuing methods and predefined priority levels applied by ToS field. Cisco is one of the most renowned companies for network equipment. Their networking algorithms /8/ contain the following queuing mechanisms which are capable to handle with congestions in network:

- FIFO queuing
- Priority queuing (PQ)
- Custom queuing (CQ)
- Weighted Fair Queuing (WFQ)
- Class Based Weighted Fair Queuing (CBWFQ)
- Multiclass weighted round robin (MWRR)
- Deficit weighted round robin (DWRR)

Each algorithm is designed to solve the specific problem concerned with the network traffic. Each solution can affect on network performances in sense of link throughput, delays, dropped traffic /1, 2/ etc.

Custom queuing (CQ) task is to share link bandwidth between active applications in the network respectively between the organizations. In such circumstances the bandwidth must be divided proportionally between applications and end users in sense that congestions cannot appear. CQ mechanism provides satisfactory bandwidth at congestion point and at the same time protects specific traffic with constant bandwidth amount, meanwhile remains bandwidth CQ leaves to other active data traffic in the network. Traffic is managed with assigning weight amount respectively to place in queue for each class of packets. In that case we have 17 waiting queues – classes or more which are mutual equivalent. In each waiting queue class has annotation which tells us, how much bandwidth should be provided for transferring data into the output link. Afterwards, algorithm arrange messages into one of seventeen available queues, where queue with index 0 is intended for storing system messages, such as keepalive messages and many other signal messages. Queue empty procedure executes upon principle weight largeness, obtained from priorities, what means, that message with high priority amount has a largest weight comparison with message with lower priority amount. Router manage waiting queues from index one to index 16 in round robin manner. Figure 1 show custom queuing mechanism with up to seventeen waiting queues served in round robin manner.

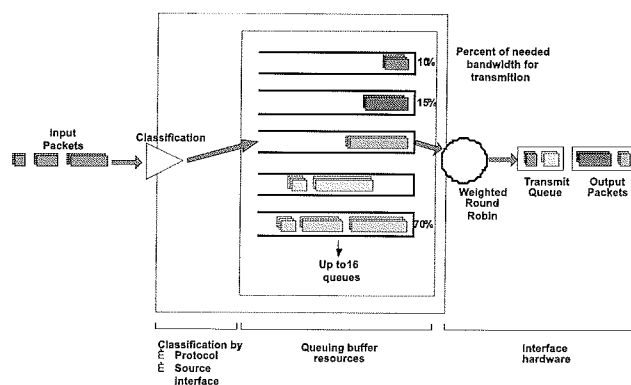


Fig. 1. Custom queuing (CQ) discipline

CQ is statically configured /11/ as many other mechanisms, such as PQ etc. Statically means that queuing mechanism is not capable automatically change router settings regarding to dynamical changes within the network. This is the reason why CQ is not capable to automatically accommodate to dynamical changes in network /4/.

Weighted fair queuing (WFQ). In situations where we want to provide constant response time and keep delays in required range without assigning excessive bandwidth, is optimal solution weighted fair queuing. WFQ is algorithm/mechanism which introduces bit-wise fairness, and it allows, that is each queue fair served. Fairness is provided using mechanism which counts bytes. For the simplest example we observe two queues of the same length. First of them contains 100 packets and second one contains only 50 packets. In that case will WFQ work on the follow-

ing manner; firstly it will take two packets from first queue then it takes one packet from second queue and then again two from first and one from second queue etc.

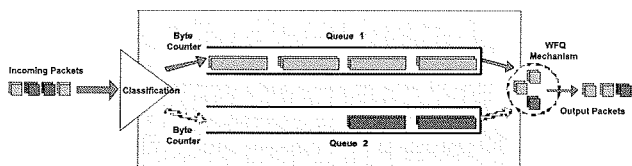


Fig. 2. Weighted fair queuing (WFQ) discipline

Such algorithm makes possible service fairness for each participating queue. Low level priority traffic flows undisturbed travel through network, what is a good compromise for each participating traffic flow in the network. WFQ also have other benefits such as minimizing configuring costs, because it is capable automatically accommodate to the dynamic network changes. Because of its good properties is in use on the majority of the serial E1/T1 interfaces. Weight is, also in this case, defined with IP priority amount (ToS field of IP packet). For IP priorities are in use settings in range from 0 (best effort) to 5 (IP quality speech), meanwhile settings 6 and 7 are reserved. Algorithm then uses these data to figure out, how many additional services it must provide/predict for each individual queue. It is capable to use each of available bandwidth for low priority traffic flows even if is presented high-priority traffic flow which don't use all bandwidth. In basis this principle is opposite comparison with time-division multiplexing (TDM), which reserves all available bandwidth and lives it unused if traffic flow is not presented. WFQ operates with IP priority settings as well with resource reservation protocol (RSVP) and also is capable to manage round trip delay problem. Such queuing clearly improves algorithms such as SNA, logical link control (LCC) and transmission control protocol (TCP), and at the same time it is capable accelerating flows and removing congestions in the network. Results became more predictable during whole routing path, meanwhile delays can be multiple reduced/decreased comparison with other queuing disciplines (CQ, PQ, FIFO) /8/.

Multiclass weighted round robin (MWRR) bases on a best effort connection scheduling discipline. It represents some kind of simplest emulation of generalized processor sharing discipline. Comparison, while generalized processor sharing (GPS) /12/ serves infinitesimal amount of data from each non empty connection, MWRR serves a number of packets for each connection which is not empty, where is this number (denoted with x) defined with the following equation

$$x = \text{normalised} \left(\frac{\text{Weight}}{\text{Mean_Packet_Size}} \right) \quad (1)$$

That the MWRR emulation of generalized processor sharing discipline can be correctly in cases, when packets can take different sizes, WRR server must know a source's mean packet size in advance. In practice, source's mean

packet size can be unpredictable, and this is the reason why MWRR server doesn't predict/allocate bandwidth fairly. Additionally there is also the problem that algorithm is fair only over time scales longer than a round time. If the time scale is shorter than round time then some connections may get more services than others. If a connection has a small weight or respectively if number of connections is large, then these scenario leads to long periods of unfairness /3, 5/.

Deficit weighted round robin discipline (DWRR) is some kind of a modified ordinary weighted round robin scheduling. It is capable handling variable size packets without knowing their main size in advance. A DWRR scheduler associates each connection with a deficit counter initialized to 0. Such scheduler visits each connection in turn and tries to serve one quantum of bits from each visited connection. For example, the packet at the head of the queue is served only, if it isn't larger than the quantum size. Opposite, if it is larger, the quantum is added to the connections deficit counter. If the scheduler visits a connection such that the sum of the connection's deficit counter and the quantum is larger than or equal to the size of the packet at the head of the queue, then the packet at the head of the queue is served and the deficit counter is reduced by a packet size.

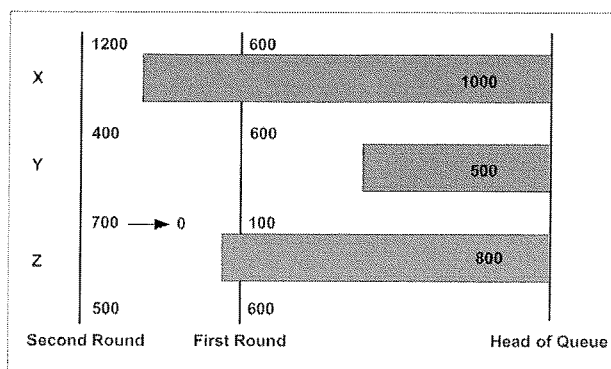


Fig. 3. Deficit weighted round robin queuing discipline

In the above figure one packet from each of connections X, Y and Z is queued at the DWRR scheduler. The quantum size is 600 units. In the first round Y's packet is served, and deficit counters for X and Z increase to 600. In the second round, X's packet is served and its deficit counter is set to 400. Z's packet is also served and its deficit counter is set to 500. Y's counter is cleared, because it has no packets to send /3, 6, 7/.

Class based weighted fair queuing (CBWFQ) represents the newest scheduling mechanism intended for handling with congestions' and at the same time provides better flexibility. We can use it when want to provide minimal required amount of bandwidth to specific application (in our case VoIP). For that case, network administrator must provide classes with predefined bandwidth amounts, where one class is for example intended for video conferencing

application, another to VoIP application and so on. Instead of waiting queue assurance for each individual traffic flow, CBWFQ determines different traffic flows. For each such class is assured minimal bandwidth. We take a look at example, where could majority lower priority traffic flow take off an important part of bandwidth from highest priority traffic flow in case of WFQ use. Typical example is video transmission, which needs half of T1 connection bandwidth / 8/. Sufficient link bandwidth can be assured using WFQ mechanism, but only when present are only two traffic data flows. In cases when more than two traffic flows appear, video session will suffers for bandwidth because WFQ mechanism works on fairness principle. For example, if nine additional traffic flows make demands upon same bandwidth, video session gets only 1/10 of whole bandwidth, and that is not enough. Even, if we put IP priority level 5 into ToS field of IP packet header, circumstances will not change. In that case, video conference gets only 6/15 of bandwidth, and that is still not enough, because mechanism must provide half of all available bandwidth on T1 link. With use of CBWFQ mechanism we can assure required bandwidth. Network administrator just needs to define for example video conferencing class and install video session into that class, the same principal can be used for all other applications which need specific amount of bandwidth. Such classes are served with flow based WFQ algorithm, which allocates remain bandwidth to other active applications in the network.

4. Opnet modeler simulation tool

OPNET Modeler is a leading simulation tool used for modeling, simulations and analysis of communication networks, devices, protocols and applications. Modeler is a graphically oriented simulation tool, which uses project, node and process editors for building communication models. The project editor offers graphic topological description, whilst in the meantime the node editor is used for describing protocols. The process editor is an upgrade of C language, which uses a finite-state machine for algorithm and protocol descriptions. OPNET includes many standard communication models for constructing wired, radio, optical and satellite communication structures. Tool offers also 2D and 3D animation to accompany changes in the network structure /9/.

5. Simulation scenario for queuing disciplines evaluation on VoIP application

Test simulation network architecture is a copy of the real private network which belongs to private company. Our main goal is oriented to improve network performances.

The highest level on Figure 4 represents network server architecture offered from internet service provider (ISP). Servers' subnet consists of five Intel servers where each

of them offers own profile, such as; web profile (web server), VoIP, E-mail, FTP and video profile. These servers are connected over 16 port switch through wired link onto private company's router. Company network consists of four LAN segments including different kind of users. On the west wing of the company are VoIP users which represent technical support to their customers. At south wing of company is placed conferencing room where employees make meetings with other outland partners. Here are two places intended for 2 simultaneous sessions. At north wing is a smallest office with only 10 employees which presents the development part of company, and they uses different applications needed for their work. For example, they searching information on the web; call their suppliers, mutual exchange files over FTP and so on. Remain west wing includes fifty disloyal employees which surfs on the net (web) during work time, downloading files etc (heavy browsing).

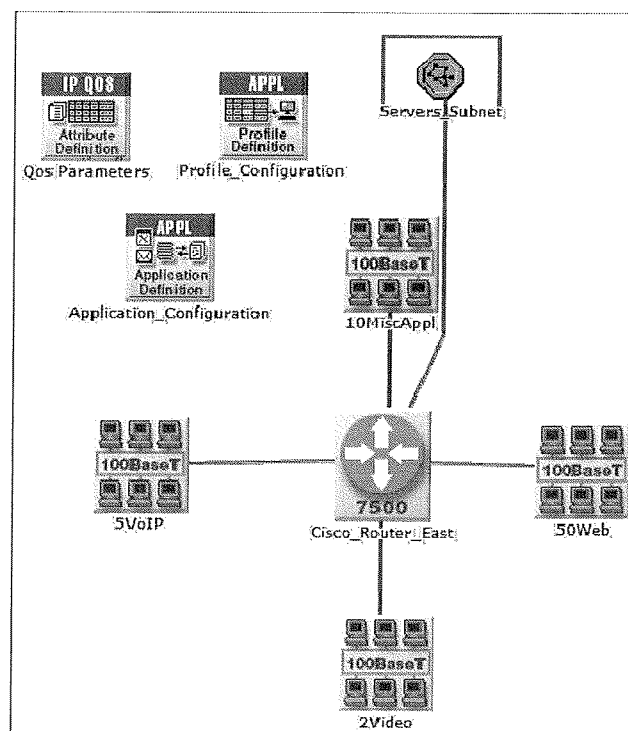


Fig. 4. Tested wired network architecture which is a copy of a real network

The highest level on Figure 4 represents network server architecture offered from internet service provider (ISP). Servers' subnet consists of five Intel servers where each of them offers own profile, such as; web profile (web server), VoIP, E-mail, FTP and video profile. These servers are connected over 16 port switch through wired link onto private company's router. Company network consists of four LAN segments including different kind of users. On the west wing of the company are VoIP users which represent technical support to their customers. At south wing of company is placed conferencing room where employees make meetings with other outland partners. Here are two places intended for 2 simultaneous sessions. At north wing is a

smallest office with only 10 employees which presents the development part of company, and they uses different applications needed for their work. For example, they searching information on the web; call their suppliers, mutual exchange files over FTP and so on. Remain west wing includes fifty disloyal employees which surfs on the net (web) during work time, downloading files etc (heavy browsing).

Each wing of company is connected over 100BaseT link to Cisco 7500 router. Further, this router is connected with ISP servers switch over a wired (VDSL2) 10Mbit/s ISPs' link. Connections between servers and switch are also 100BaseT. Wired link in this case represents bottleneck, so for that case we have to involve QoS system and their different queuing disciplines.

Table 1. Used applications and number of their clients

Application	Number of users
Heavy web browsing	50 clients
FTP	4 clients
Video conferencing	7 clients
VoIP	5 clients
E-mail	1 client

We create six scenarios; in the first scenario we tested custom queuing discipline which represents a basic for comparison with WFQ (second), MWRR (third), DWRR (fourth), CBWFQ (fifth) and with combined hybrid PQ-CBWFQ (sixth scenario) queuing disciplines. Network topology in all scenarios remain the same, differences are only in used queuing disciplines. By comparison of simulations' results for different scenarios we tried to prove, how each of queuing discipline serves used network applications.

6. Simulation results

As we mentioned before, we collected delay statistics from six different queuing disciplines scenarios (CQ, WFQ, MWRR, DWRR, CBWFQ and PQ-CBWFQ) for two different active applications (VoIP and HTTP) in the network and with different applied priorities by ToS field of IP packet header. For that case we defined VoIP traffic flows between clients where such flows represents high priority traffic, meanwhile HTTP traffic represents low priority flow, based on best effort type of service. In our scenarios we have 82.09% users which use lower priority HTTP traffic and only 17.91% users which use high priority VoIP application. Considering on Figure 5, means, that only 17.91% of users takes a whole majority part of bandwidth, so lowest priority HTTP traffic which represents majority of all traffic must wait. Because such majority traffic must wait, delays rapidly increase what is clearly seen on a Figure 5. Evident is, that VoIP traffic has lowest delays in comparison with HTTP traffic. Best results we obtain with custom queuing method which ensures required bandwidth on possible congestion points and serves all traffic fairly. After CQ

queuing scheme follows WFQ, DWRR, MWRR, CBWFQ and PQ-CBWFQ. WFQ, DWRR, MWRR and CBWFQ reaches worse delay results, because of fairness queuing discipline. Similar results are obtained also in HTTP case. IF CBWFQ scheme is in use then high-priority traffic will be ensured with fix amount of bandwidth defined from administrator side. For example, administrators, using CBWFQ, define 9Mbit/s for VoIP, so in that case remains only 1Mbit/s for all other majority applications, so majority low level traffic will affect on rapidly delay increasing, as is shown on figure 6.

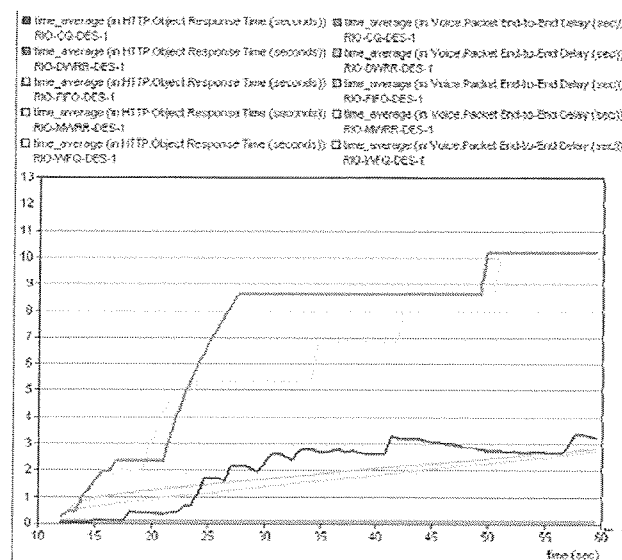


Fig. 5. Used different queuing disciplines upon VoIP and HTTP traffic

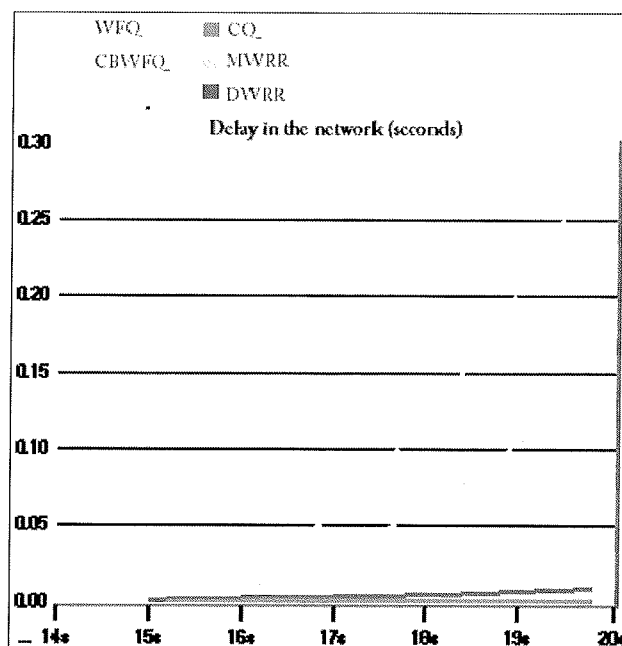


Fig. 6. Time average global delay in the network

Figure 7 shows the amount of VoIP dropped packets, when using different queuing schemes. As we mentioned above, best results in that case we obtain with CBWFQ method which have fixed guaranteed amount of bandwidth. Then follow WFQ, DWRR, MWRR and CQ queuing scheme. The opposite situation is when we take in to the consideration delays. There selection of CBWFQ introduces biggest delay, because majority low level traffic must wait.

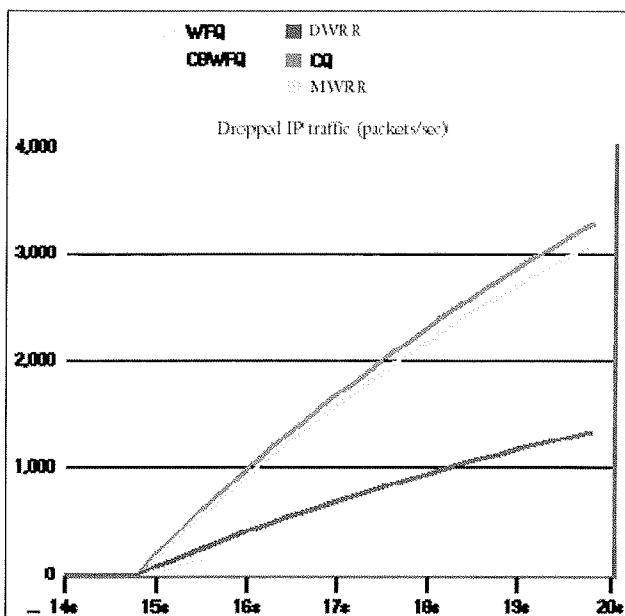


Fig. 7. Tested wired network architecture which is a copy of real network

On the Figure 8 we will show, how combined queuing method PQCBWFQ improves the delays comparison with before presented delays, shown on figure 6.

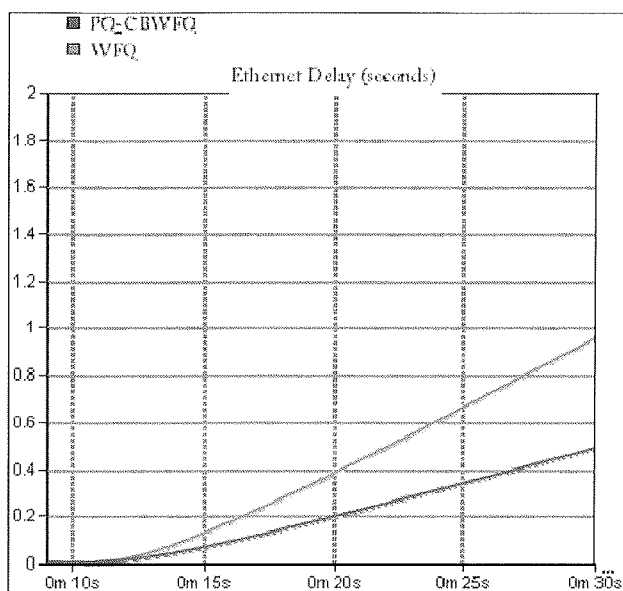


Fig. 8. Ethernet delay (seconds) for combined PQCBWFQ method comparison with WFQ

As we can see on Figure 8, PQ-CBWFQ delay is smallest than WFQ delay, but on figure 6 has ordinary CBWFQ method bigger delay than WFQ, observed in whole ethernet segment. Such combinations can perceivable improves network performances. Similar effect that is shown as on Figure 8 can be seen also in Figure 9 for VoIP delay.

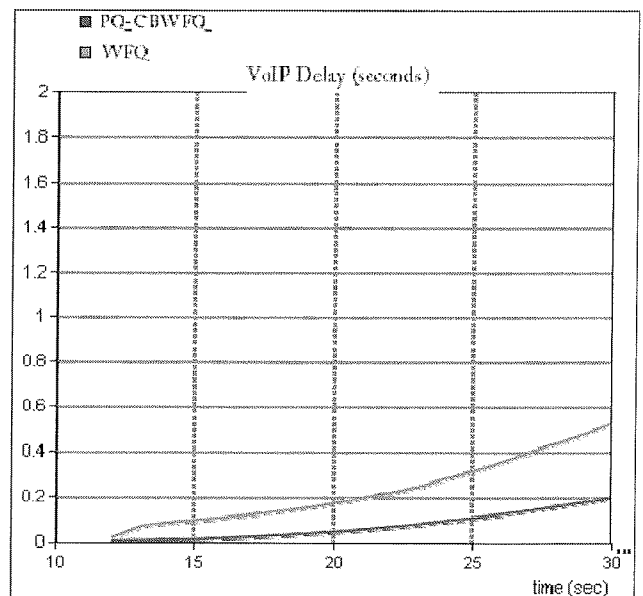


Fig. 9. VoIP delay (seconds) for combined PQCBWFQ method comparison with WFQ

Using combined queuing method the delay is also reduced in comparison with ordinary WFQ queuing for VoIP traffic. Delay in VoIP application plays an important role on quality of perception. The smaller it is, the better voice quality can be offered.

7. Conclusion

During many simulation runs and graph analysis we can say that queuing policy discipline influences on quality of service for the network applications. In many cases CQ queuing discipline was the best choice, in the case, when we have only two traffic flows WFQ has been the best choice, but looking from other perspective, when we must handle multiple traffic flows CBWFQ was the best solution. Such method (CBWFQ) has also disadvantages; in our case we defined only one class with bandwidth amount 9Mbit/s reserved for VoIP, rest of bandwidth belongs to majority low priority HTTP traffic. In that point, majority traffic hasn't enough bandwidth and must wait, what influences on common delay. This is the main reason why CBWFQ has highest average delay in network. No matter to that delay, VoIP delay is constant during simulation because of ensured bandwidth by defined class. Looking from other side, if we want fairness queuing discipline which serves fair all applications then we use WFQ or CQ mechanism, but if we want only that highest priority traffic flows pass through the network, we should use priority queuing PQ.

Delays in CBWFQ case can be reduced using PQ-CBWFQ queuing scheme. This is a combination of a priority PQ and a CBWFQ queuing mechanisms. PQ-CBWFQ is a queuing method from family that offers smallest delays. Such queuing allows that specific flow class defined with IP priorities is served as strict priority queue. Highest priority class is served before all other priority classes. Such combination and their improvement from view of Ethernet delays are shown on Figure 9. Our simulations show that we must look for solutions also in using combined queuing methods. All other available combinations represent a challenge for further researches on that area.

References

- /1/ S. Klampfer, J. Mohorko, and Č. Žarko, "Vpliv različnih načinov uvrščanja na karakteristiko prepustnosti omrežja," ERK 2007
- /2/ M. Fras, S. Klampfer, Ž. Čučej, "Impact of P2P traffic to the IP communication network's performance," IWSSIP 2008
- /3/ T. Subash and S. IndiraGandhi, "Performance Analysis of Scheduling Disciplines in Optical Networks," MADRAS Institute of Technology, Anna University
- /4/ Jernej Kozak: Podatkovne strukture in Algoritmi. Društvo MFA SRS, Ljubljana.
- /5/ Lary L. Peterson and Bruce S. Davie, Computer Networks, Edition 3, San Francisco 2003
- /6/ Stefan Bucheli, "Compensation Modeling for QoS Support on a Wireless Network," Master degree thesis
- /7/ Kian Meng Yap, Alan Marshall, Wai Yu, "Providing QoS for Multimodal System Traffic Flows in Distributed Haptic Virtual Environments," Queen's University Belfast
- /8/ Internetworking Technology Handbook – Quality of Service (QoS), Cisco Systems
- /9/ OPNET Modeler Technical Documentat
- /10/ Samuel P. Morgan, "Queuing Disciplines and Passive Congestion Control in Byte-Stream Networks," IEEE Transactions On Communications
- /11/ http://www.cisco.com/en/US/docs/ios/12_0/qos/configuration/guide/qccq.html
- /12/ Maria Johanna Gerarda van Uiter, "Generalized Processor Sharing Queues," Masterdam 2003

Saša Klampfer, Jože Mohorko, Žarko Čučej
University of Maribor, Faculty of Electrical Engineering
and Computer Science, Maribor, Slovenia
sasa.klampfer@uni-mb.si

Prispelo (Arrived): 28.10.2008 Sprejeto (Accepted): 09.06.2009