Scheduling queues in the Ethernet switch, considering the waiting time of frames

Evgeniy Kizilov[1], Nikolas Konnov[1], Kirill Nikishin[1], Dmitry Pashchenko[1,a], and Dmitriy Trokoz[1]

[1]*Penza State University, Departmen of Computer Engineering, 40, Krasnaya street, Penza, 440026, Russia*

Abstract. Authors proposes an algorithm to scheduling queues with temporal selection of frames in the Ethernet switches with QoS support, which is based on the waiting time of frames in the queues of different classes. Evaluation of the effectiveness of this algorithm compared to the classical cyclic algorithms by simulation with hierarchical temporal coloured Petri nets using CPN Tools package was conducte.

1 Introduction

A necessary requirement to modern networks is the support quality of service (QoS), which enables various network applications to coexist on the same network, without consuming the bandwidth of other applications. Among the mechanisms to ensure quality (marking and traffic classification, Scheduling queues, congestion control, bandwidth control) is the most important mechanism for Scheduling queues, i.e. the selection rules of the queue to read data which directly determine the frame delay, its variation (jitter) and the probability of their loss [1,2].

In the intelligent Ethernet switches with QoS support for queuing the cyclic algorithms are widely used that provide traffic prioritization and allocation of guaranteed bandwidth to individual classes: Weighted Round Robin (WRR) and Deficit Round Robin (DRR). In these algorithms for each traffic class its own queue is allocated in the buffer of the output port. Queues are serviced in order of ascending of their numbers, if there is the absence of frames in some queue, the algorithm switches to another queue.

The amount of data sent from each queue during one session of servicing, is set by its assigned weight which determines allocated for the corresponding traffic class bandwidth of an output port.

In the algorithm WRR weight specifies the maximum number of frames transmitted from the queue in one session, and in the algorithm DRR weight specifies number of bytes of data. Difference between DRR and WRR algorithm is that if a situation arises when the band of queue has not been developed, but it is not sufficient for transmission of the next packet, then the problem frame packet is transmitted, but the amount of information which exceeded the allocated band is subtracted from the newly allocated for this queue bandwidth in the next cycle of servicing. This allows you

to select more accurately the desired band with a significant spread of sizes of frames in queues [3].

An advantage of cyclic algorithms is an exception of the monopolization of channels and the ease of implementation, and the disadvantage of cyclic algorithms is the uncontrolled size of jitter of switched frames from the spread waiting time of frames in different queues, which is caused by the deterministic order of service. Besides, the delay of frames of a certain class of QoS and the size of the corresponding queue depend on the allocated bandwidth for a specific class. It leads to increasing of delay of frames and congestion of queues in channels with small assigned bandwidth when there is the pulsating character of the traffic, as the round robin algorithms do not effectively redistribute unused bandwidth underutilized channels to extend the overloaded bandwidth

2 The principle of queuing with time selection of frames

To solve this problem, the authors propose a new algorithm queue control, called "Time Selection Service" (TSS), which is based on the waiting time of frames in the queues of different classes.

The TSS algorithm works in such a way:
- each frame sent in the appropriate class QoS queue, is supplied the attribute, marking time of setting in queue;
- after the end of the servicing the next queue the controller of switch analyzes the attributes of frames which are in the queues and selects for the further servicing of queue with its smallest value, i.e., the queue with the most "old" frame at the moment;
- from the analysis of queues the last serviced one is eliminated (except when all other queues are empty);

---

[a] Corresponding author: dmitry.pashchenko@gmail.com

- to regulate the bandwidth of each traffic class the standard methods of weighting (WTSS) or deficit (DTSS) are used.

3 Modeling by Petri nets

Evaluating the effectiveness of the TSS-algorithms compared with WRR and DRR algorithms is implemented by their simulation modeling in CPN Tools package, using the mathematical apparatus of hierarchical temporal coloured Petri nets. Freeware CPN Tools package was selected as tool system, which due to a rich set of extensions allows simulating various aspects of the behavior of telecommunication systems [4].

The structure of the Petri net modeling TSS-algorithms is built on the principles defined in [5] and in many respects is similar to the previously developed networks, modeling the WRR and DRR algorithms [6,7]. Within developing of the model the methodology of top-down design was used. Model Petri net of the switch on the top level of the hierarchy is shown on Figure 1, it actually follows the structure of the switch [5] and includes:

- Subnet "Bufer", which simulates a frame buffer directed to the corresponding output port of the switch. This subnet actually implements QoS mechanisms, so it will be considered further.

- Subnet "Outport", which simulates the switch ports that provide access to the Ethernet communication lines. This subnet generates the delay serialization during the transmitting of frames, the duration of which depends on the size of the preamble and interval of frames.

- Subnet "SetTrafGenerator" that simulates the flow of frames belonging to different QoS classes and following through the time intervals distributed according by a given law [8].

-The position of "Outport", containing the tokens that imitate the frames of switched traffic.
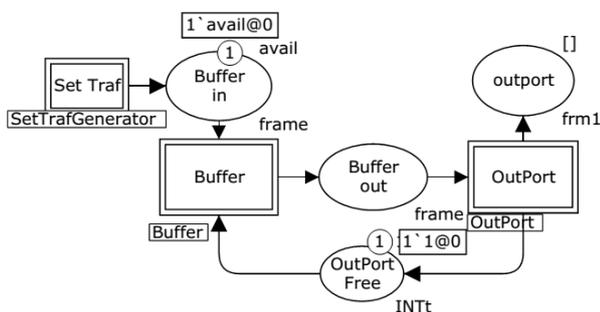

Fig. 1. General model

To simulate the communication protocol of network devices it is very important to effectively use the apparatus of color tokens. Network traffic is modeled by a sequence of tokens received at the input ports and leaved from the output ports, colours of which form a tuple frm of fields corresponding to the frame format Ethernet 802.1 q/p [9]: source and destination addresses, the field tag with information about belonging to a specific priority level, a field indicating the length of the payload of the frame, as well as servicing time field the parish and the outcome of the frame (used to collect statistics). The variety of colors frame is also used, it is the bearer, which may be either busy sending a frame frm, or free avail, that allows to detect and handle events related to the presence or absence of the frame. Additionally, the tokens are introduced, they reflect the colour characteristics of the utility model (queue size, amount of transferred data, the attribute of queue cleaning et al.).

The model of "Buffer" (fig.2) serves queue of 4 QoS classes and contains the following elements:

- "Buffer in" and "Buffer out" - input and output positions of subnet;

- "Classifier" – subnet of classifier, which analyzes the token of input frame and sends it to one of four queues, simultaneously sending to "Count time" the token the colour of which is set in equal to the value of model time of arrival of the token frame (attribute);

-"Queue1", ... "Queue4" - subnets of queues containing frame tokens, waiting for servicing. Subnet is also formed in the positions "indicator1", ... "indicator4" the colour of the token that reflects the time of admission to queue of the frame located at the head of the appropriate queue, and in the positions of the "Size Queue1", ... "Size Queue14" the colour of the token indicating the size of the queue is formed;

- "Scan" – is subnet, which implements the TSS-algorithm of queuing, forming the token for one of the positions "Start selection1", ... "Start selection4";

- "bend1", ... "bend4" – is subnet, allocating bandwidth for the traffic of each queue;

-"Number Queue" - position that contains the number of the queue that was last serviced;

="Holder" – position the token of which synchronizes the operation of the subnet "Scan" and "bend";

-"OutPortFree" - position reflecting the state of the output port.

After classification the next incoming frame token is transmitted to one of the queues "Queue", each of which contains two FIFO buffer for token of frames and their attributes. Let's note that to describe the queue the language CPN ML is used, which allows to represent the queue by one token, which is a list of frames. In the initial marking the list is empty.

The size of the buffers can be defined in frames or in bytes, depending on the method of bandwidth control. Work of the buffers during writing and reading is synchronized, so that the head token of queue of attributes is always placed in the position "Indicator". The subnet "Queue" is also counting the current size of the queue displayed in the token position "Size Queue", and is resetting the tokens when the queue is full.

Reading from the queues is performed on arrival of the authorizing token at the position "Start selection" from the subnet "Scan", which selects the served queue as follows:

- when there is the releasing of an output port, it writes the corresponding position "OutPortFree" the authorizing token, if there is the token then a subnet "Scan" analyzes the colour of the token in "Holder";

- if it is 0, it means that the session of service the previous queue, the number of which defines the colour of the token at position "Number Queue", is finished, the

subnet "Scan" comparing attribute values in the positions "indicator" defines the queue attribute of which has the smallest value and sends it to position "Start selection4" of the tokens. Thus if there are multiple active queues it prevents duplicate service of previously served queue, which will eliminate the monopolization of the output port of any queue.

- if token color is 1, the subnet "Scan" sends a permissive reading in position "Start selection4" of the queue, the number of which specifies the colour of the token at position "Number Queue", so the session of servicing of selected before queue is continued.

Subnet "Bend" implements the allocation of bandwidth for each queue by setting a quota of data transferred in one session of service. For the WTSS algorithm weight set for queues 1, 2, 3 and 4 that sets a quota in the frames and for the DTSS algorithm sizes of a quota are set to 1500, 3000, 4500 and 6000 byte.

In addition to the calculation of transferred data while promoting in the "Buffer out" of each frame, subnet sets the colour of the token in the position "Number Queue", equal to the number of the served queue, and the colour of the token in the position "Holder" is set to 1 during transmitting of all frames except the last one in the servicing session, which is zero colour of this token.
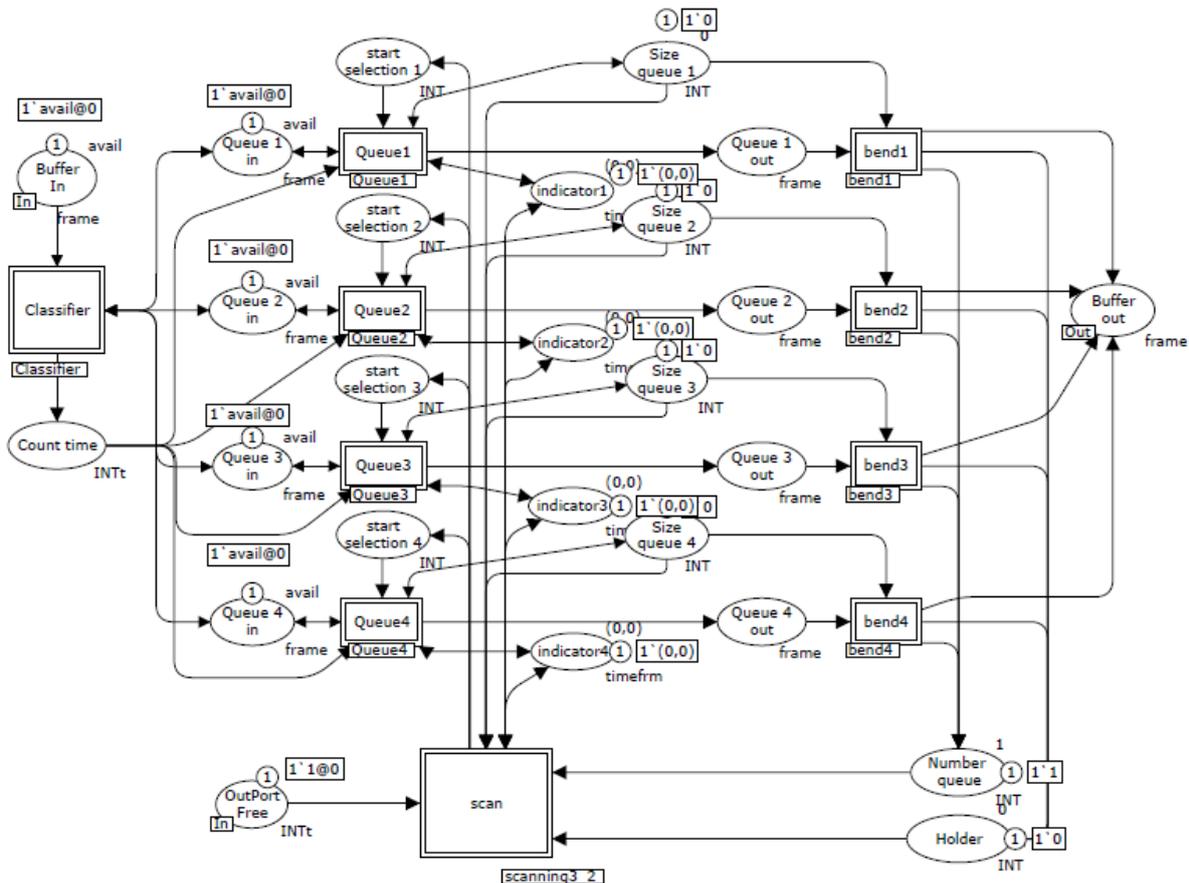

Fig.2. Subnet of the Buffer of Output Port

4. Discussing of modeling results

The proposed model allows to investigate the probabilistic-temporal characteristics (average delay, jitter, percentage of packets lost) of mixed traffic, switched to the Ethernet.

As an example of using of the proposed model, the results of the evaluation of the distribution function of the jitter of frames with different QoS classes with algorithms queuing WRR, DRR [6,7] and WTSS, DTSS are shown below.

The input traffic is modeled with using of special Petri nets [8] and represents a combination of stream of frames belonging with equal probability to 4 QoS classes and following through the time intervals distributed exponentially with an average value of the delivery period, and quasi-regular steam with rare frames of the maximum duration. Model time is expressed by clock intervals TI, which is equal to 100 ns for Ethernet or 10 ns for Fast Ethernet. The random frame length has the following distribution corresponding to real networks [10]: 25% of all frames have a minimum and maximum length (512 and 12176, respectively), length of remaining ones is evenly distributed in the remaining range.

Noticeable difference in the statistical characteristics of the jitter frames RR- and TSS-algorithms is occurred when the load factor of the channel is greater than 0.5. The following data for the load factor of the channel is equal to 0.8. The frames have uniform QoS labels from 0 to 7. 4 classes of traffic are formed, the appropriate queues are allocated for them, and the following values of the bandwidth are set:

- queue with QoS 0-1 – 10% of the output bandwidth;

- queue with QoS 2-3 – 20% of the output bandwidth;
- queue with QoS 4-5 – 30% of the output bandwidth;
- queue with QoS 6-7 – 40% of the output bandwidth;

As well, the generated traffic of class QoS 0-1 exceeds the allocated bandwidth in 2 times, QoS 2-3 is agreed with the allocated bandwidth and QoS 4-5 and 6-7 – underutilize the allocated bandwidth.

The average values of frame delay for each class, received during the simulation are presented in Table 1, show that the TSS-algorithms reallocate unused bandwidth, in order to reduce the delay in overloaded queues at the expense of the most underutilized ones.

Table 1. Average values of frame delay (jitter)

QoS class	Average delay (TI)			
	WRR	DRR	WTSS	DTSS
0-1	55601	45788	45544	28279
2-3	27631	28429	26554	32967
4-5	20688	25676	24634	30710
6-7	21817	26018	29094	33807
All	31434,3	31477,8	31456,5	31440,8
Variation of average delay	34913	20112	20910	5528

Using algorithms posravneniyu EYY QC reduces the variation of the average delay for the different classes of traffic qo 1.68 to 3.64 times.

The fact that the redistribution of unused bandwidth graphs illustrate the probability distribution functions of jitter for classes QoS 0-1 (overloaded), and QoS 6-7 (underused) are shown in Fig. 3.
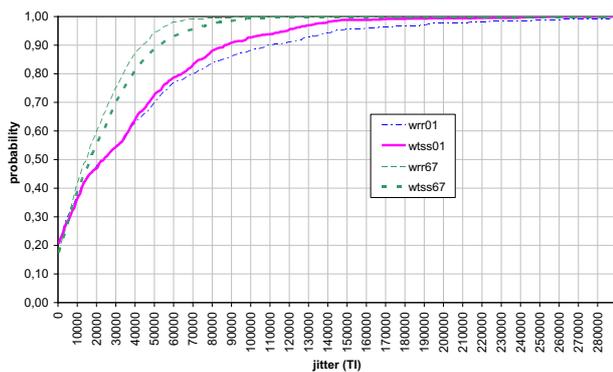

Fig.3. The probability distribution functions of jitter for classes QoS 0-1 QoS 6-7 и scheduling algorithms WRR and WTSS .

5.Conclusions

Thus, it can be considered that TSS-algorithms provide a range of delays in the promotion of traffic of different classes, as well as more comfortable conditions for the transferring of traffic with different classes in case of imbalance of bandwidth allocated to them, at the expense of more fair and effective redistribution of unused resource.

Acknowledgements

References

1. N. Olifer, V. Olifer. *Computer Networks: Principles, Technologies and Protocols for Network Design,* Wiley E-Text, (2007)
2. Miguel Barreiros, Peter Lundqvist *QOS-Enabled Networks: Tools and Foundations*. (2011)
3. S. Vegesna. *IP Quality of Service.* Cisco Press*,(* 2001)
4. K. Jensen, L. Kristensen. *Coloured Petri Nets: modelling and validation of concurrent systems.* (2009)
5. D.A. Zaitsev. Mathematics and Computers in Simulation. 65, 245-249 (2004)
6. V.B. Mehanov, E.A. Kizilov. *Novyie informatsionnyie tehnologii i* sistemyi, **1**. 67-73 (Penza, PGU, 2010)
7. E. Kizilov, N., Konnov, D Pashchenko, D. Trokoz, *2015 5th International Workshop on Computer Science and Engineering: Information Processing and Control Engineering,(WCSE 2015-IPCE), 185-*190. (Moscow, 2015)
8. K.I. Nikishin, N.N. Konnov, A.L. Domnin. *Sovremennyie problemyi kompyuternyih nauk (SPKN-2013).* 120-123 (Penza, PGU, 2013)
9. E. Niclas. IEEE 802.1 P,Q - QoS on the MAC level, http://www.tml.tkk.fi/Opinnot/Tik110.551/1999/papers/08IEEE802.1QosInMAC/qos.html
10. R. Sinha, C. Papadopoulos, J. Heidemann. Internet Packet Size Distributions: Some Observations. Technical Report ISI-TR-2007-643, USC/Information Sciences Institute, (2007). http://netweb.usc.edu/ rsinha/pkt-sizes