On assuring QoS in Ethernet access network

Robert Janowski, Piotr Krawiec and Wojciech Burakowski
Warsaw University of Technology, ul. Nowowiejska 15/19, 00-665 Warsaw, Poland
{ robert; pkrawiec; wojtek}@tele.pw.edu.pl

Abstract

This paper deals with the problem of assuring strict QoS guarantees for the end to end connections that originate from Ethernet access network. It shows that despite high link capacities in some cases Ethernet network might be the reason of QoS deterioration. The primary reason is the lack of appropriate QoS differentiation and traffic isolation mechanisms. The shared buffers and priority schedulers available in most of Ethernet switches appear to be not sufficient to guarantee strict QoS. For these cases new solution is proposed which relies on additional traffic control mechanisms available in other network elements. Only additional mechanism supporting typical functionality of Ethernet switch can provide strict QoS guarantees what was verified in simulations studies.

1. Introduction

Providing strict end-to-end QoS (*Quality of Service*) guarantees in multi-domain heterogeneous network requires assuring QoS in every part of this network: core domains as well as access domains. Since one of the most popular technology used in access is Ethernet, the approach for supporting QoS in Ethernet network is strongly desirable.

Shared Ethernet, which uses CSMA/CD (*Carrier Sense Multiple Access with Collision Detection*) mechanism to determine which device can use transmission medium, cannot guarantee bounded delays because the time needed to cover the network is potentially non-deterministic due to collisions occurring in shared medium. However, the replacement of traditional shared Ethernet by switched Ethernet has solved this issue thanks to eliminating the collision domains. Moreover, the IEEE 802.1Q [1] and 802.1p specifications (which is a part of the IEEE 802.1D [2]) have brought traffic differentiation mechanism to the MAC layer. Taking into account capabilities mentioned above, a lot of effort was put to assure QoS guarantees in Ethernet access network. However, some solutions do not provide strict QoS, as EtheReal [3], which is throughput oriented and only supports best effort traffic, or stochastic approach presented in [4], which only yields average delay performance bound. Other approaches, as [5] and [6] require additional cost of hardware and/or software modification.

Solutions proposed in [7] and [8] do not override the IEEE specifications and rely on standard Ethernet switches with priority scheduling, but require a separate queue for each priority class in order to work properly. Unfortunately, typical Ethernet switches currently offered by vendors have only a common buffer, which is shared by all priority classes. It may lead to the violation of QoS guarantees of high priority traffic in case the whole Ethernet switch buffer is occupied by low priority traffic.

In this paper we present an approach to assure strict QoS guarantees for the end to end connections that originate from Ethernet access network. We assume using only currently widespread Ethernet equipment, with shared buffers and priority scheduler. Moreover, any modification of switch software or hardware is not performed.

The rest of the paper is organized as follows: Section 2 describes the main problem treated in this paper which is assurance of target QoS level in Ethernet access network when switches support only shared buffers. In Section 3 the proposed solution is presented and simulation results are reported. Section 4 summarizes the paper and draws some conclusions.

2. Statement of the problem

2.1 The concept of end-to-end classes of service

One of the acknowledged approaches for assuring QoS for packet transfer over heterogeneous multi-domain network, is by implementing the end to end classes of service (e2e CoSs). Roughly speaking, the e2e CoS corresponds to network capabilities, for transferring the packets belonging to selected connections with assumed QoS guarantees. These QoS guarantees are expressed in terms of IP packet loss ratio (IPLR), IP packet transfer delay (IPTD) and IP packet delay variation (IPDV) metrics as defined in [9]. An e2e CoS is fully dedicated for handling the packets generated by given type of applications as, for example, for VoIP (*Voice over IP*) or FTP (*File Transfer Protocol*) connections.

This approach has been applied in currently running IST EuQoS project [10] where five types of e2e CoSs have been defined, that are differing in QoS objectives according to the requirements coming from the types of applications we study in EuQoS (see Table 1).

Table 1. End-to-end CoS defined in EuQoS

e2e EuQoS CoS	QoS parameters			If CAC is performed
	IPLR	IPTD [ms]	IPDV [ms]	
Telephony	10^{-3}	100	50	Yes
RT Interactive	10^{-3}	100	50	Yes
MM Streaming	10^{-3}	1000	-	Yes
High Throughput Data	10^{-3}	1000	-	Yes
Standard (Best Effort)	N/a	N/a	N/a	No

For implementing these e2e CoSs, adequate CAC (*Connection Admission Control*) algorithms were designed to limit QoS traffic and QoS mechanisms like schedulers, shapers, policers etc., available in network elements (IP routers, WiFi access points, LAN switches etc.) were used. Only for Standard CoS there is neither CAC function performed nor the QoS parameters are guaranteed as this CoS is intended to provide similar service as Best Effort network. However, Standard CoS has been guarantied a minimum bandwidth which is forced with nonzero weight assignment in WFQ scheduler and in this way protected against starvation.

Implementation of e2e CoSs encounters different obstacles when considering each of possible access network technologies i.e. WiFi, UMTS, xDSL or Ethernet. In this paper we focus on the problem of Ethernet access network. It is commonly believed that the Ethernet access networks are over-provisioned. However, the word “over provisioning” is not precisely defined and mainly used to qualitatively characterized capabilities of a given element saying that it negligibly contributes to the performance deterioration. Furthermore, a careful examination of QoS assurance in Ethernet network reveals the existence of some problems.

In Ethernet the primary mechanism to differentiate traffic is Priority Scheduler, practically available in almost every switch. Taking into account 802.1p standard which defines 8 priority classes in Ethernet network, we propose the following mapping between end-to-end EuQoS CoSs and these priority classes (see Table 2).

Table 2. Mapping between end-to-end EuQoS CoS and Ethernet priority classes

e2e EuQoS CoS	Ethernet priority class	802.1p priority value
	Network Management	7 (highest)
Telephony,RT Interactive	Voice	6
	Video	5
MM Streaming, High Throughput Data	Controlled Load	4
	Excellent Effort	3
Standard (Best Effort)	Best Effort	0
	Undefined	2
	Background	1

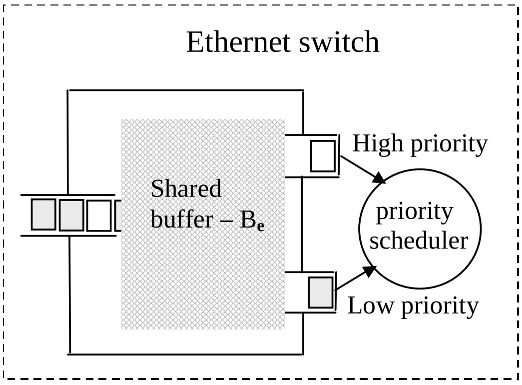

Figure 1. Architecture of shared buffer in Ethernet switches.

2.2 Ethernet switches with shared buffers

The mapping shown in Table 2 implies that the traffic from Standard CoS is served with the lowest priority regarding other end-to-end EuQoS CoSs. Unfortunately, typical Ethernet switches do not support per class buffer but only a common one, which is shared by all CoSs including Standard CoS. Although this buffer is quite large – usually thousands of packets, the fact it is common poses the main problem [11]. Since the traffic submitted to Standard CoS is by definition not controlled by CAC function, it is very likely that it overloads the network. When it fully occupies Ethernet buffer then it may deteriorate the performance of traffic from other CoSs. This deterioration will impact only IPLR metric and is inherently due to the lack of separation between buffering space for packets from different CoSs. When the packets from Standard CoS occupy too much buffer space, the arriving packets from other CoSs are dropped due to the lack of room. Despite this, IPTD or IPDV metrics will never be influenced since once a packet enters shared buffer it is scheduled to the transmission according to priority rule and hence traffic from Standard CoS cannot delay packets from other CoSs. The mechanism of shared buffer combined with priority scheduling has been symbolically shown in Figure 1 by using model of Drop Tail queue at the entrance to the Ethernet buffer but Priority Scheduler at the exit.

Thus, the problem of assuring appropriate performance of e2e EuQoS CoSs in Ethernet network is mainly the problem of controlling traffic from Standard CoS and preventing it from occupying too much buffer space so that to avoid packet losses.

2.3 Reference scenarios of Ethernet access network

Ethernet switches, allow to build networks of different size in tree-like topology. This is achieved by cascading as many Ethernet switches as required, typically not more than 3 to 4 levels. We distinguish between two representative scenarios: SOHO (*Small Office Home Office*) and Enterprise as depicted in Figure 2. In both cases Ethernet access network is connected to the Internet through a router. These scenarios, except the level of aggregation and internal distribution of traffic, differ in the capacities of output links from switches and the router. In SOHO scenario the capacity of the router output link is approximately 10 times smaller than the capacity of the output link from the preceding switch. On the contrary, in Enterprise scenario the capacity of the router output link might be either smaller (case A in Figure 2) or equal (case B in Figure 2) to the capacity of the switch output link. This strongly effects the solution for providing QoS in Ethernet access network.

2.4 Evaluation of the capabilities of Ethernet access network to assure target level of QoS

To quantify the capability of Ethernet access network (Ethernet switch) to assure target level of QoS we performed simulation studies using ns-2 platform [13] in the network scenario depicted in Figure 3.

In this scenario we assume the cascade of Ethernet switch and edge router. Furthermore, we assume that the edge router supports WFQ (*Weighted Fair Queuing*) scheduler with two classes (class #1 and #2) and each of them is dedicated separate buffer B_1, B_2 respectively. This means that the router provides separation between CoSs and by limiting the traffic load to some threshold value (by means of CAC) it is able to guarantee the target values of QoS parameters for each end-to-end CoS according to the requirements from Table 1.

The Ethernet switch supports priority scheduler and a buffer B_e shared by all CoSs. There are only two types of traffic: one representing traffic from Telephony CoS for which we must guarantee target QoS level (i.e. IPLR value equal 10^{-3}) and the other which represents traffic from Standard CoS. The first type of traffic is modelled as a Poisson stream with 200B long packets served with high priority in Ethernet switch and by class #1 in the router. The other is modelled as a number (N) of TCP greedy connections served with low priority in Ethernet switch and by class #2 in the router.

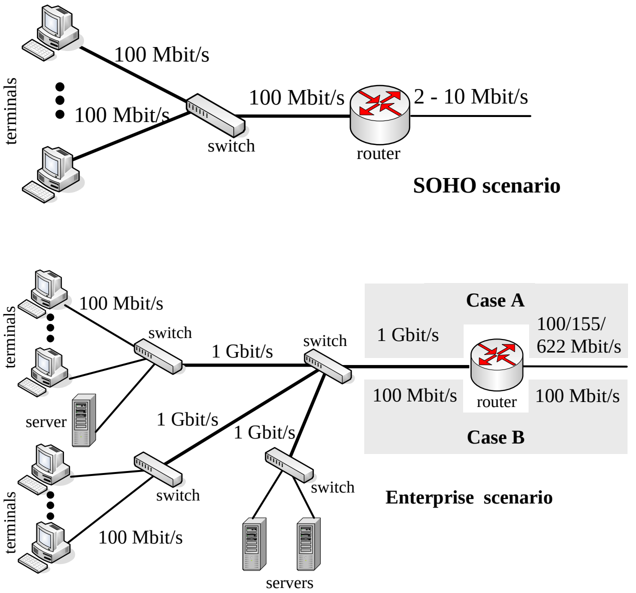

Figure 2. Representative scenarios for Ethernet access network

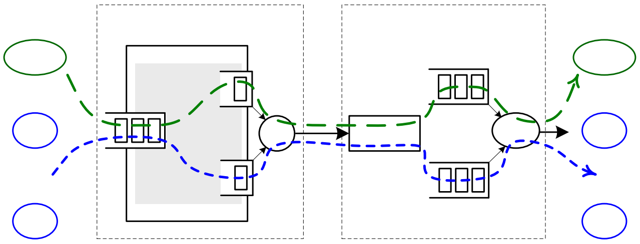

Figure 3. Simulation scenario used in evaluation of capabilities of Ethernet switch to assure QoS.

The main objective of this simulation test was to verify if end-to-end EuQoS CoSs were still supported when the traffic originated from Ethernet access network. In other words we wanted to check if the target IPLR=10^{-3} was assured for Telephony CoS when it was loaded up to the limit determined by CAC algorithm performed in router and additionally the traffic from Standard CoS was present. The CAC load limit ($\rho_{Telephony}$) was found from the algorithm described in [12]. For the following set of parameters: B_1=10 packets, IPLR=10^{-3} it gave $\rho_{Telephony}$=0.714. Using this result, and assuming that a fraction (w_1) of the output capacity C_2 is dedicated to Telephony CoS, we obtain the following rate limit of Poisson traffic:

$$Rate_{Telephony} = w_1 \times C_2 \times \rho_{Telephony} \qquad (1)$$

The remaining parameters of the model have the following values:

- router buffer for the Standard CoS is B_2=100 packets,
- shared buffer at the Ethernet switch is B_e=2000 packets.

All results were represented with 95% confidence intervals calculated for 12 time repeated simulation tests unless these intervals are negligible.

Table 3. Simulation results of QoS metrics for Telephony CoS in SOHO scenario ($C_2<C_1$). The values of remaining parameters are the following: w_1=0.5, N=100.

C_1 [Mbps]	C_2 [Mbps]	IPLR	IPTD [ms]	IPDV [ms]
100	20	$4.2\text{x}\ 10^{-4} \pm 0.4\text{x}10^{-4}$	1.1	1.9
100	10	$3.4\text{x}\ 10^{-4} \pm 0.5\text{x}10^{-4}$	2.1	3.8
100	8	$3.3\text{x}\ 10^{-4} \pm 0.5\text{x}10^{-4}$	2.6	4.7
100	2	$2.0\text{x}10^{-4} \pm 0.8\text{x}10^{-4}$	10.3	18.6±0.4

Table 4. Simulation results of QoS metrics for Telephony CoS in Enterprise scenario, case A ($C_2<C_1$) and case B ($C_1=C_2$). The values of remaining parameters are the following: w_1=0.1, N=100.

C_1 [Mbps]	C_2 [Mbps]	IPLR	IPTD [ms]	IPDV [ms]
1000	100	$5.5\text{x}10^{-4} \pm 0.6\text{x}10^{-4}$	0.4	1.6
1000	622	$5.3\text{x}10^{-4} \pm 0.3\text{x}10^{-4}$	0.1	0.3
100	100	$4.9\text{x}10^{-3} \pm 0.1\text{x}10^{-3}$	0.4	0.4

Results from Table 3 and Table 4 show that in case when $C_2<C_1$ (SOHO scenario or Enterprise scenario case A), the assumed QoS level is assured for Telephony CoS. The values of IPLR, IPTD and IPDV metrics stay below the target ones. The observed increase in values of delay metrics (IPTD and IPDV) in Table 3 is only due to lower values of capacity C_2. The appropriate performance of Telephony CoS is guaranteed by the fact that relatively low capacity C_2, as compared to C_1, throttles TCP traffic carried by Standard CoS. In consequence this traffic cannot occupy too much buffer space in Ethernet switch and deteriorate the performance of traffic from Telephony CoS.

On the contrary, the simulation results obtained for Enterprise scenario case B i.e. $C_2=C_1$ (see Table 4) reveal the situation where at the Ethernet access network the QoS level of the traffic from Telephony CoS is deteriorated. In this case the problem is evident since IPLR value is 5 times higher than the target one. The primary reason is that the output link from Ethernet switch creates the main bottleneck while its lacks appropriate mechanisms to provide CoS isolation and QoS assurance.

3. Proposed solution

In the previous section we have shown that in Enterprise scenario the Ethernet access network is not able to assure the target level of QoS for Telephony CoS in the presence of traffic served by Standard CoS, at least in terms of IPLR metric.

For proposing a solution to alleviate the problem stated in section 2, we follow the afore mentioned strategy for implementing e2e CoSs which assumes the usage of the QoS mechanisms available in the network elements. Since in an ordinary Ethernet switch these mechanisms are of limited capability, we will try to make use of the mechanisms available in other network elements.

Our solution is based on the assumption that the traffic from Standard CoS uses TCP protocol. Since TCP traffic is self-regulating, any action undertaken on the traffic from Standard CoS at any node along its end-to-end path, will also affect its behaviour at the Ethernet switch. In consequence it is possible to control traffic from Standard CoS at the first router which follows the Ethernet switch (see Figure 2). This element was chosen for two reasons. First, routers support wide range of mechanisms for traffic conditioning i.e. policing, shaping, different types of scheduling. Secondly, the Standard CoS traffic should be throttled as closely to its source as possible to avoid unnecessary congestion at other network elements located further along the end-to-end path. As the regulating mechanism we propose to use a shaper with appropriately set shaping rate C_S and the buffer size B_S. The model of the Ethernet switch and router in tandem with this additional traffic control mechanism is depicted in the Figure 4.

The values of shaper parameters C_S and B_S play crucial role in controlling TCP traffic carried by Standard CoS. Shaper rate C_S throttles TCP sources by constraining the rates they can attain. Taking into account TCP behaviour in slow start phase [14], these sources can temporarily send packets with the rate equal to $2C_S$ at most. Since the shaping rate is C_S a queue builds up in the shaper until it reaches B_S size. Then the packet losses occur and consequently, TCP sources respond to them by decreasing their sending rates.

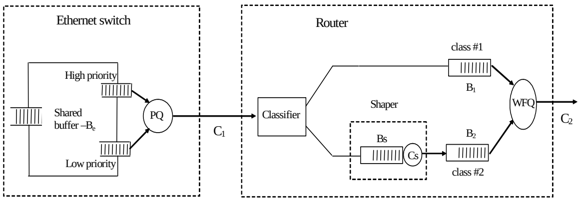

Figure 4. Model of Ethernet switch and router in tandem with additional shaping capability.

During this initial period until the first packet loss, also a queue in the Ethernet switch may build up depending on the mutual relation between the capacities C_1, C_S and the rate of Telephony CoS. If capacity C_S meets the following condition:

$$C_s > \frac{C_1 - Rate_{Telephony}}{2} \qquad (2)$$

the queue also builds up in the switch. Otherwise, even during the slow start phase the TCP sources send packets with the rate lower than $C_1 - Rate_{Telephony}$ and there is no chance for the queue to build. After the slow start phase, which ends with the first packet loss, the sending rate of TCP sources oscillates according to the self-regulating feature of TCP. During this "after slow start" period the mean queue size in the switch is practically zero (only simultaneous packet arrivals can build small queues) if the condition (2) is not met. Otherwise, the queue size may be quite large since it is dependent on the number of TCP sources.

3.1 Tuning parameters of the shaper

Regarding the scenarios shown in Figure 2, it should be noted that in case of SOHO there is no need for such additional mechanism. Since the capacity C_2 is lower than the half of capacity $C_1 - Rate_{Telephony}$, TCP sources can never overload the buffer in Ethernet switch. The shaping is required only in case of Enterprise scenario when capacities C_1, C_2 are of similar values e.g. equal. In such scenario, to be definitely sure that the traffic from Standard CoS can never occupy so much buffer space in the Ethernet switch to cause the packet losses of traffic from other CoSs, C_S should be set to the value lower than the half of $C_1 - Rate_{Telephony}$. Then the choice of B_S value has no importance. However, such solution has one drawback. Since at any node supporting classes of service, Standard CoS has been guaranteed the minimum rate and the ability to utilize all the available capacity (e.g. by serving it with WFQ scheduler), the low value of C_S may preclude it from achieving this objective.

The other solution is to set the C_S value closer to C_1. Then the value of B_S should be carefully tuned so that to avoid too high queue in the Ethernet switch.

3.2 Evaluation of the QoS assurance

To verify our approach we performed simulation tests in both scenarios using ns-2 platform [13]. For this purpose we use the same traffic scenario and values of parameters as in section 2.4. The only difference is the presence of the shaping mechanism in the router.

Table 5. Simulation results of QoS metrics for Telephony CoS in Enterprise scenario, case B; B_1=10 [pkts], ρ_1=0.714 , w_1=0.1, B_2=100 [pkts], N=100, C_1=100 Mbps, C_2=100 Mbps, B_e=2000 [pkts]

Shaper parameters		Simulated QoS metrics		
B_S [pkts]	C_S [Mbps]	IPLR	IPTD [ms]	IPDV [ms]
1000	40	0	0.2	0.4
1500	40	0	0.2	0.4
1000	90	$9.9\text{x}10^{-3} \pm 2\text{x}10^{-3}$	0.5	1.7
1500	90	$8.6\text{x}10^{-3} \pm 1.3\text{x}10^{-3}$	0.5	1.7

The results show that if shaper rate C_S does not meet the condition from equation (2), e.g. C_S=40 Mbps, the traffic carried by Standard CoS has no chance to build high queues in Ethernet switch. Thus, it cannot deteriorate the performance of the traffic carried by Telephony CoS independently of the buffer size B_S in shaper.

Different situation is encountered when the shaper rate C_S meets the condition from equation (2), e.g. C_S=90 Mbps. Then the appropriate size of the shaper buffer B_S is crucial. If B_S is too high, TCP sources can chance to extremely inflate their congestion windows and even occupy the whole space in Ethernet buffer. This case is shown in the two last rows of Table 5, where the resulted IPLR value is 10 times higher than the target one. In order to avoid performance deterioration of Telephony CoS, the buffer size B_e in Ethernet switch should be at least so long to accommodate the maximum queue size provoked by TCP traffic from Standard CoS and still have the space equal to the size of the buffer B_1 dedicated to Telephony traffic in the edge router. This means that before the Ethernet switch buffer fills up to the value B_e-B_1, which is at time T_{FullBe}, the buffer B_S at the shaper must overflow so that the occurring packet losses could slow down the rates of TCP sources and in this way avoid filling to much space of Ethernet switch buffer. Time T_{FullBe} is calculated as the ratio of the target buffer occupancy B_e-B_1 and the filling up rate of TCP traffic. This rate is equal to the difference between the input rate ($2C_S$) and the drainage rate equal to $C_1 - Rate_{Telephony}$:

$$T_{FullBe} = \frac{(B_e - B_1) \times 1500 \times 8}{2C_s - (C_1 - Rate_{Telephony})} \qquad (3)$$

In order to overflow the shaper buffer B_S before filling the Ethernet switch buffer to B_e-B_1 value, the time to fill up the shaper buffer T_{FullBs}:

$$T_{FullBs} = \frac{B_S \times 1500 \times 8}{(C_1 - Rate_{Telephony} - C_S)} \qquad (4)$$

must be shorter than the time required to fill up Ethernet buffer - T_{FullBe}. Comparison of these two times gives the following condition for the maximum value of the shaper buffer B_s which assures the target QoS level of Telephony CoS traffic:

$$B_S < \frac{(B_e - B_1)(C_1 - Rate_{Telephony} - C_S)}{2C_S - C_1 + Rate_{Telephony}} \qquad (5)$$

Below we provide the simulation results for the same scenario which was evaluated in Table 5. This time we show that the choice of the shaper buffer B_S according to formula (5) lets to assure the target IPLR value for Telephony CoS.

Table 6. Simulation results of QoS metrics for Telephony CoS in Enterprise scenario, case B; B_1=10 [pkts], ρ_1=0.714 , w_1=0.1, B_2=100 [pkts], N=100, C_1=100 Mbps, C_2=100 Mbps, B_e=2000 [pkts]

Shaper parameters		Simulated QoS metrics		
B_S [pkts]	C_S [Mbps]	IPLR	IPTD [ms]	IPDV [ms]
64	90	$4.1 \times 10^{-4} \pm 0.4 \times 10^{-4}$	0.5	1.7

4. Conclusions

In this paper we have shown that in some scenarios Ethernet access network might be a bottleneck in providing strict QoS guarantees. Because of limited QoS differentiation mechanisms available in Ethernet switches the problem of QoS assurance could only be solved by introducing additional traffic control mechanism at the neighbouring router. For this purpose the shaping mechanism was proposed. Furthermore, it was shown that although the values of shaper parameters do not have to be especially tailored, their choice is not trivial.

From the deployment point of view this solution requires only activating and proper configuration of the shaping mechanism available in a router so it doesn't have to be taken into account during the network designing phase. It is more flexible. If we decide to connect an Ethernet LAN to the edge router in Enterprise scenario then the shaping should be activated. Otherwise, e.g. connecting WiFi or xDSL doesn't imply the need for configuring the shaper.

Our solution focused on the problem of non-local connections which originate from the Ethernet network but terminate at the remote network. We have not addressed the problem of internal traffic in Ethernet access network which might be of significant volume in case when the Enterprise network is large e.g. consisting of many terminals and servers. We consider this issue as an important topic for our future research since solving it will provide complete solution for strict QoS assurance in the environment with a limited support of QoS differentiation mechanisms and diverse types of connections (local and distant).

5. References

[1] IEEE Standard for Local and metropolitan area networks. *Virtual Bridged Local Area Networks*. IEEE Std 802.1Q – 2003

[2] IEEE Standard for Local and metropolitan area networks. *Media Access Control (MAC) Bridges*. IEEE Std 802.1D – 2004.

[3] S. Varadarajan, "Experiences with EtheReal: a fault-tolerant real-time Ethernet switch", Proc. of 8th IEEE International Conference on Emerging Technologies and Factory Automation, vol. 1, 2001, pp. 183-194

[4] B. Y. Choi, S. Song, N. Birch, and J. Huang, "Probabilistic approach to switched Ethernet for real-time control applications", Proc. of 7th International Conference on Real-Time Computing Systems and Applications (RTCSA'2000), Dec. 2000

[5] Feng-Li Lian, Yang-Chung Tu, Chun-Wei Li, "Ethernet Switch Controller Design for Real-Time Control Applications", Proceedings of the 2004 IEEE International Conference, 2-4 Sept. 2004 Page(s):464 - 467 Vol.1

[6] Hoai Hoang; Jonsson, M.; Hagstrom, U.; Kallerdahl, A.; "Switched Real-Time Ethernet with Earliest Deadline First Scheduling – Protocols and Traffic Handling", Proceedings of the IPDPS 2002, 15-19 April 2002 Page(s): 94 – 99

[7] Xing Fan, Magnus Jonsson, "Guaranteed Real-Time Services over Standard Switched Ethernet", in Proceedings of the IEEE Conference on Local Computer Networks 30th Anniversary (LCN,05), 15-17 Nov. 2005 Page(s): 3 pp

[8] Jiming Chen, Zhi Wang, Youxian Sun, "Switch Real-Time Industrial Ethernet with Mixed Scheduling Policy", IECON 02, Volume 3, 5-8 Nov. 2002 Page(s):2317 - 2321 vol.3

[9] ITU-T Recommendation Y.1541, "Network performance objectives for IP-based services".

[10] IST EuQoS project homepage www.euqos.eu

[11] P. Pedreiras, R. Leite, and L. Almeida, "Characterizing the Real-Time Behavior of Prioritized Switched-Ethernet", In *2nd RTLIA*, Porto, Portugal, June 2003.

[12] H.Tarasiuk, R. Janowski, W. Burakowski "Admissible traffic load of real time class of service for inter-domain peers", in Proceedings of ICNS&ICAS'2005 Conference, Papeete, Thaiti, October 2005, IEEE press.

[13] www.isi.edu/nsnam/ns

[14] R. Stevens "TCP/IP Illustrated" vol. I., Addison-Wesley Publishing Co, 1994