

Congestion control mechanisms of IEEE 802.1Qau for reducing the load of Metro Ethernet networks

Maha BOUAZIZ¹, Héli KAANICHE², and Lotfi KAMOUN¹

¹ Laboratory of Electronics and Information Technology LETI-ENIS, Sfax-TUNISIA

² CRISTAL Laboratory, ENSI, University of Manouba, 2010 Manouba-Tunis, TUNISIA

Maha.bouaziz@gnet.tn, Heli.kaaniche@tunet.tn, Lotfi.Kamoun@isecs.rnu.tn

Abstract: In the past few years, the data-processing network has become an essential resource in all fields. Metro Ethernet is a new type of data-processing network, which is commonly deployed for metropolitan areas. Yet, despite the importance of this network, it continues to be afflicted by the issue of congestion, caused by insufficiency of resources or by the bandwidth difference between one network and another. Therefore, to ensure an effective use of this resource, Metro Ethernet networks must be properly controlled to avoid the loss of data and to maintain good performance.

Recently, several congestion control mechanisms, which follow the standard IEEE 802.1Qau [WGP 08], have been proposed, including the Ethernet Congestion Manager (ECM), Quantized Congestion Notification (QCN), Quantized Ethernet Congestion Manager (QECM), and QCN Serial HAI. A comparative study and evaluation of performance among these mechanisms was made, and it was noted that QCN Serial HAI was the most efficient mechanism [ABD 07b]. However, this mechanism uses an enormous number of notification messages, which overloads the network.

Our objective in this paper is to suggest ways of reducing the load of networks.

Key words: Metro Ethernet, congestion, control of congestion, IEEE 802.1Qau, load of network, number of notification.

Introduction

Since their appearance, the applications of communication networks have been growing rapidly. The increase in the number of users, in the traffic, and in the need for new services stimulated the development of new technologies and the spread of networks with high bandwidth, such as the Metro Ethernet networks. But this growth has caused problems of congestion on the level of the nodes of the network [DAV 08]. Indeed, the congestion occurs when a network has an insufficiency of resources (busy bandwidth, storage capacity of switches, etc.), as it can be caused by the difference of the busy bandwidth from one network to another.

The congestion causes loss of data circulating in the network, thus reducing the quality of services. To ensure an effective use of the available resources, networks must be supported by effective congestion control mechanisms.

In this project, we were interested in suggesting new approaches for the congestion control mechanisms following the standard IEEE 802.1Qau for Metro Ethernet networks. Our aim was to help reduce the load of networks caused by the large number of notification messages.

The remainder of this paper is organized as follows: The first section presents the main principle of congestion control mechanisms. The second and third sections present our suggested solution to reducing the load of the network, focusing on the benefits of the new approach. Our final section deals with conclusions and perspectives.

1. Related work for congestion control in Metro Ethernet networks

The commonly used congestion control mechanisms follow the standard IEEE 802.1Qau [WGP 08], [ITG 07]. They are based on three functions [RAJ 07]. The first is the detection of congestion. The second is the principle of notification, i.e., to inform the source to react before the production of the congestion in order to be able to adjust the rate of transfer according to the availability of the constant resources. The third entity is the reaction by the source according to the notification received.

Mechanisms are based on two entities [HUG 07]. The first is the congestion point (CP), which is an entity integrated into a switch. The CP takes a frame sample among the entering frames at each interval, and then generates a notification message addressed to the source of the frame sample. This generated message contains information concerning the state of the queue of the switch. The second entity is the reaction point (RP), an entity integrated into a rate limiter related to the source. It receives messages of notification and adjusts its transfer rate according to the received information from the CP [ITG 07].

The queue of the switch contains three thresholds: the equilibrium threshold (Q_{eq}), the delicate threshold (Q_{mc}), and the severe threshold (Q_{sc}). The main goal of the congestion control mechanism is to keep the occupation in the queue on the level of the equilibrium threshold. This is carried out by adjusting the rate of transfer of the source using a calculated value of feedback F_b [DAV 07a] [DAV 07b].

The operations of previous congestion control mechanisms are ECM (represented in [DAV 05], [YI 07]), QCN (represented in [DAV 07c], [GUE 07], [RON 07], [ABD 07a], and [ABD 07c]) and QECM (represented in [GUS 07]).

The most efficient mechanism is the QCN Serial HAI [BER 08b], [ABD 07b] in terms of its performance criteria (stability, fairness, scalability).

However, this mechanism uses an enormous number of notification messages, which might overload the network. To this end, we are interested in finding a possible solution to this problem.

2. Techniques used to reduce the number of notifications for improving the network's load

To reduce the network load in terms of the number of notifications circulating in the network, the proposed approach consists in modifying the principle of the two basic entities of the congestion control mechanism: detection and notification. We also seek to eliminate useless notifications.

2.1. New principle of detection and generation of the notification

Among the deficiencies of previously used mechanisms, the notification messages sent at each sampling interval were sometimes useless yet treated. Our proposal is to carry out detection and generation of messages only in the case of network need, based on the thresholds of the queue of the congested switch. That is to say, these two principles work only when the size of frames occupation in the queue reaches a threshold. At this point, the

notification takes the values of information for congestion control. In a second step, these messages are sent to the source to react and recover the proper functioning of the network.

The proposed solution adds two other thresholds (Q_t and Q_m) between the sensitivity threshold (Q_{mc}) and the equilibrium threshold (Q_{eq}), as illustrated in Figure 1. Thus, detection is based on these five thresholds, to check the state of congestion of the queue and generate a new notification.

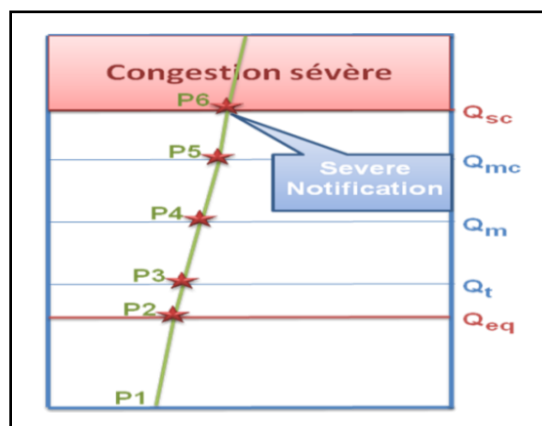


Figure 1. Representation of different thresholds of the queue switch

In addition, to reduce the number of notification messages, the proposed switch does not transmit a notification whenever it reaches a certain threshold. It waits for a time (ΔT) before issuing a new notification at the level of the same threshold (Figure 2), except for the severe threshold (Q_{sc}) at which notification messages are sent without any delay to avoid the risk of data loss.

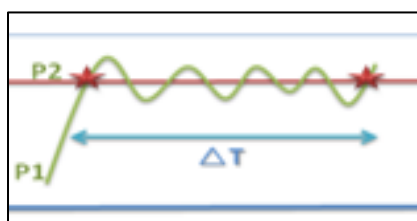


Figure 2. The expectation of ΔT between two successive notifications

2.2. Notification of the appropriate source

According to previous versions of congestion control mechanisms in line with the IEEE 802.1Qau, a notification message is issued to each sampling interval. This notification is also transmitted to the source, which sends a frame when the interval is reached (the source of the sample).

Previously, the choice of the notification of the source had not been well studied. Indeed, notification may be sent to a source of low flow that does not cause congestion. In such cases, notification is sent unnecessarily and the switch is forced to send additional notifications to solve the problem.

In addition, the number of notification messages sent by the congested switch is high, as shown in Figure 3; at $t = 5s$, 5824 notifications are sent. This figure is achieved in a scenario of three traffic sending sources to a recipient via a single switch for the "QCN Serial HAI" mechanism.

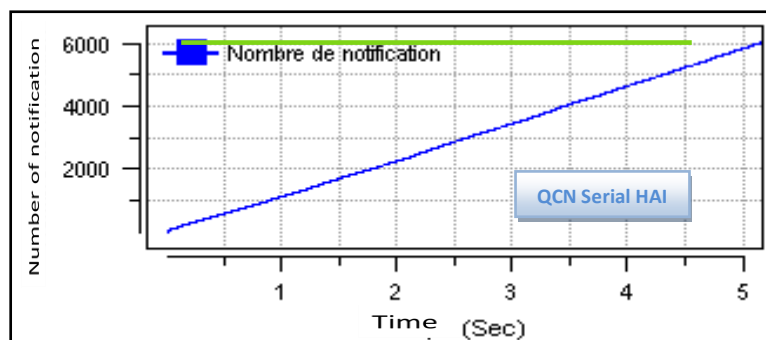


Figure 3. Growth in the number of notifications sent by the congested switch for mechanism "Serial QCN HAI"

In order to reduce the number of notification messages flowing through the network, the proposed switch is expected to notify at the threshold level the source which really causes the congestion. This is the source with the highest transfer rate. Consequently, the addresses of different sources of frames inserted between two successive thresholds are stored in a register, and the switch selects the proper source for notification.

Figure 4 represents the same scenario discussed previously but using the new approach by sending notification to the source with the highest rate among those recorded in the register.

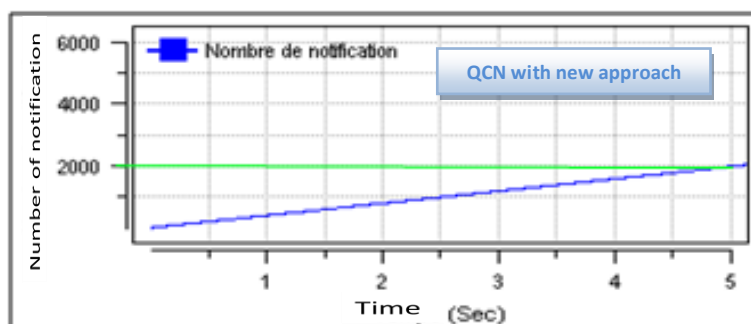


Figure 4. Growth in the number of notifications sent by the switch for the new approach

This new method improves the network load. In fact, the number of notifications sent by the congested switch is reduced. At $t = 5$ seconds, the number of notification messages decreases from 5824 messages in the case of "QCN Serial HAI" to 1981 messages when using the new approach (Figure 4).

2.3. Elimination of unnecessary notifications

During severe congestion, a source that receives a notification ECM (0.0), would stop the transfer of frames during a specified time. During the downtime, it rejects all other frames of notification. In such a case, the network will be burdened by unnecessary notifications that would be rejected. In the new approach we propose to avoid sending notifications to a source that has been a pattern of severe congestion while at rest, in order to further reduce the number of notifications circulating in the network and improve its load.

When a switch sends a frame of severe notification toward a source, and it wants to send another notification to another source, then it must check if the time separating the two notifications is higher than the rest period. If the check is positive, it sends it; otherwise, it does not.

3. Improved fairness in the new approach

According to previous versions of congestion control mechanisms of the IEEE 802.1Qau, since the notification is sent to the source of the sample, the fairness isn't well maintained. For example, for QCN Serial HAI in Figure 5, at $t = 1.39$ s source B has a rate equal to 70Mbit/s. In contrast, for other sources, the rate is about 20Mbit/s.

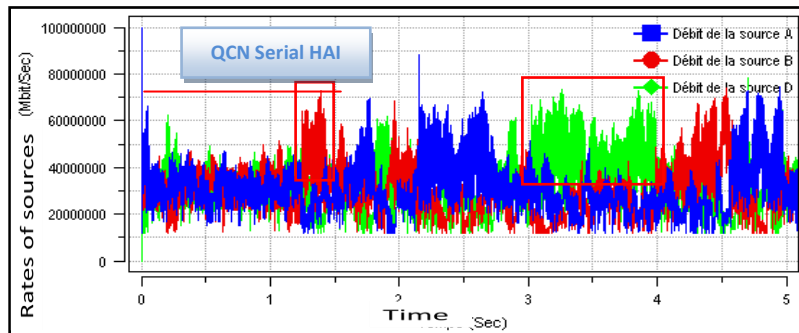


Figure 5. *The variation rates of sources for the mechanism QCN Serial HAI*

The technique used in the notification for the adequate source (which consists in the choice of the source with the highest transfer rate among those registered in the sources register) is to improve the fairness of the transfer rates of different sources. Figure 6 shows that although the flows of springs are almost flat around 33Mbit/s, the problem of fairness is solved. The figure illustrates a scenario of three traffic sources transmitting to a recipient via a single switch.

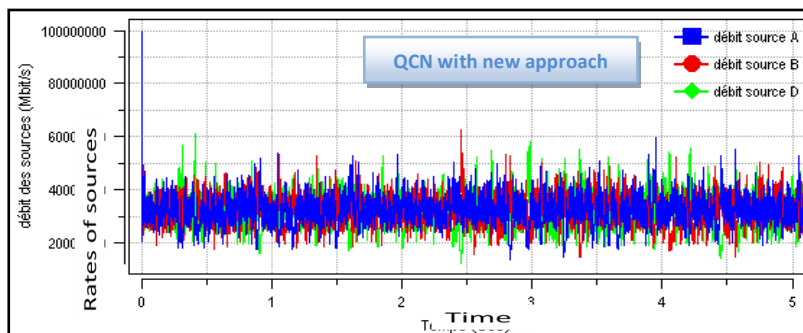


Figure 6. *The variation rates of sources for the new approach*

4. Conclusions and perspectives

In this paper, we noted that the mechanism QCN Serial HAI is the most efficient among the various mechanisms of IEEE 802.1Qau. Its performance in networks is improved compared to mechanisms such as ECM, QCN, and QECM. Yet, in this mechanism fairness was not well maintained, and the overloaded networks which are saturated by an enormous number of notification messages for adjusting transfer rate, remained unsolved.

We propose a new approach, which is designed to act against congestion and which provides better performance in networks than the QCN Serial HAI mechanism. The principle of detection and notification is modified. The proposed approach relies on thresholds to verify the state of the waiting queue of a congested switch (i.e., the check and the notification are executed when the curve of occupation frame in the queue reaches a certain threshold, instead of executing it at each interval automatically). So, the number of notifications

circulating in the network is decreased and the network load seems to be improved. Therefore, the fairness and the load of the network are better maintained.

As a perspective, the adaptation of QCN Serial HAI mechanism for enormous and complex topologies in order to reduce the rate of lost frames would be highly recommended. Such a recommendation is due to the failure of this mechanism in the elimination of data frames losses.

REFERENCES

[ABD 07a] Abdul Kabbani, Rong Pan, Balaji Prabhakar, Ashvin Laxmikantha, Mick Seaman “*QCN: transience, equilibrium, implementation*”, Meeting.

[ABD 07b] Abdul Kabbani, Rong Pan, Balaji Prabhakar, Mick Seaman “QCN: Algorithm for P- code”, Meeting, Décembre 2007.

[ABD 07c] Abdul Kabbani, Rong Pan, Balaji Prabhakar, Mick Seaman “QCN: Stably improving transient Response” Meeting, Novembre 2007.

[DAV 05] Davide Bergamasco (Cisco), “*Data Center Ethernet Congestion Management: Backward Congestion Notification*”, IEEE 802.1 Interim Meeting, Berlin Germany, May 2005.

[DAV 07a] Davide Bergamasco (Cisco), “*Ethernet Congestion Manager*”, initial Draft, May 2007.

[DAV 07b] Davide Bergamasco (Cisco), “*Ethernet Congestion Manager*”, IEEE 802 Plenary Meeting, Orlando, FL, Mars 2007.

[DAV 07c] Davide Bergamasco, “*QCN: Problems, Solutions, and a Convergence Proposal*”, IEEE 802.1 Inetrim Meeting, Stockholm, Sweden, 2007.

[DAV 08] Dave Allan, “*Metro Ethernet Convergence Toolkit*” Nortel, February 2008.

[GUE 07] Guenter Roeck, Teak Technologies, “*Congestion Management Protocols: More QCN Simulation Results*”, Meeting, July 2007.

[GUS 07] M. Gusat, C. Minkenberg, G. Roeck, “*Closed Loop CM w/ Probing QECCM-(S)P*”, Meeting October 2007.

[HUG 07] Hugh Barrass (Cisco), “*Quality of service for unbounded data streams*” Joint ITU-T/IEEE Workshop, June 2007.

[ITG 07] Interworking Task Group, “*Congestion Management*”, draft standard for local and metropolitan area networks, IEEE P802.1au/D0.2 August 2007.

[RAJ 07] Raj Jain, Jinjing Jiang, Chakchai So-In, “*Congestion Notification for Congestion Notification for Data Center Ethernet Networks: Key Principles*” IEEE 802.1au Meeting, San Francisco, July 2007.

[RON 07] Rong Pan, Balaji Prabhakar, Ashvin Laxmikantha, “*QCN : Quantized Congestion Notification*”, Meeting.

[WGP 08] IEEE 802.1au Work Group Public Documents Repository, <http://www.ieee802.org/1/pages/802.1au.html>.

[YI 07] Yi Lu, Rong Pan, Balaji Prabhakar, Davide Bergamasco, Valentina Alaria, Andrea Baldini, “*Congestion control in networks with no congestion drops*”, article, 2007.