I.J. Wireless and Microwave Technologies, **2011, 2, 47-54**
Published Online April 2011 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijwmt.2011.02.07

Available online at http://www.mecs-press.net/ijwmt

Discussion on IGMP Snooping and Its Sequence in Industrial Applicaiton

Sheng LU, Chuan WEN

School of Computer Science and Information Engineering, Chongqing Technology and Business University, Chongqing, China, 400067

Abstract

This paper has an introduction on a new intelligent controller for industrial Ethernet with IGMP: ICIE (Intelligent Controller for Industrial Ethernet). It proposed a new mechanism to minimize the congestion which is based on the taking an adaptive decision during transferring multicast messages. It also focuses on the sequences analysis on IGMP snooping. It has a further discussion on the steps: receiving IGMP loin message, receiving IGMP leave message, receiving IGMP query message, sending IGMP join message, sending IGMP leave message and network topology change. According to the requirement of end device ring topology, we add a DSA tagged BPDU to indict the target of packet. It is through the external devices solicitation of join messages that allow the IGMP Snooping component to correctly decipher on which ports the downstream listeners are connected.

Index Terms: IGMP; industrial Ethernet; BPDU

1. Introduction

Nowadays, more and more network applications, such as data distribution, distant education and distributed database, work on the multicast communication mode.

The IGMP Snooping service of our product (Intelligent Controller for Industrial Ethernet, ICIE) provides a method to filter multicast traffic to downstream devices. The filtering consists of blocking the multicast traffic on ports to which there are no downstream consumers – a process known as “pruning”. As a down stream device on a port registers for a particular multicast stream – an Ethernet/IP listen only connection for example – the IGMP Snooping component recognizes that a device in the direction of this port is requesting to receive the particular multicast traffic and allows the traffic to flow out of the port. On another port, however, if no downstream

* Corresponding author.
E-mail address: lusheng8815@126.com

device requests the traffic, the IGMP Snooping component will cause the embedded switch to block this multicast traffic to the port. In this manner, ideally, only devices requesting this traffic receive this traffic [1].

The process of a device requesting to start and stop the reception of the multicast streams is accomplished through IGMP join and Leave message requests. The IGMP Snooping component monitors (snoops) these join and leave messages to allow it to know which streams to prune from which ports. This process uses a device performing a manager role to periodically query all devices in the subnet and subsequently cause them to re-join the multicast group of listeners for any stream in which they may be interested. The management role is known as an "IGMP Snooping Querier" (from now on known as the "querier") and it is a service provided by most managed Ethernet switches [2].

In the event that a querier fails, the query messages will stop being sent into the network and the devices will stop renewing their registration for their multicast streams. The IGMP Snooping component will age out the registrations for these streams when no join messages have been seen for the specified age out time. As a result, with the loss of a querier the multicast streams will revert to being flooded in the network.

2. End Device Ring Topology

This section simply analyzes the potential network topologies in which our customers may use the ICIE module and how the ICIE module's architecture facilitates these topologies [3].

The users may use the ICIE module in a dual port end device ring in which each of the device supports enhanced RSTP protocol as shown in Figure 1. In this topology, the ICIE module should act as a root or non-root node based on the users' desires and their application operation behavior; the IGMP Snooping and QOS features should be enabled to improve the performance. The recovery time mechanism is the same as addressed in section IV.

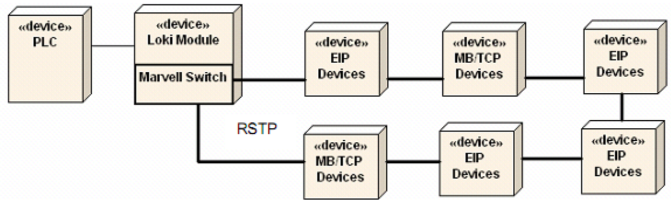

Figure 1. The dual port end device ring network topology with the ICIE module.

ODVA EIP Specification defines an end Device Level Ring (DLR) protocol that provides media redundancy in a ring topology. Our ICIE module does not implement this protocol. But the customers may use the ICIE module in their DLR ring as a non-DLR switch or a non-DLR end device [4].

Figure 2 shows an EIP DLR ring network topology in which the ICIE module involves in the ring as a non-DLR switch. In this topology, the ICIE switch should allow the user to disable the IGMP snooping on the ring ports and enable 802.3 tagged so that the DLR control message frames can go through the ICIE switch; the ICIE switch should allow the user to disable MAC learning; the QOS feature should be enabled to improve the performance [5].

Figure 3 shows an EIP DLR ring network topology in which the ICIE module involves in the ring as a non-DLR end EIP device. In this topology, the ICIE switch works as normal.

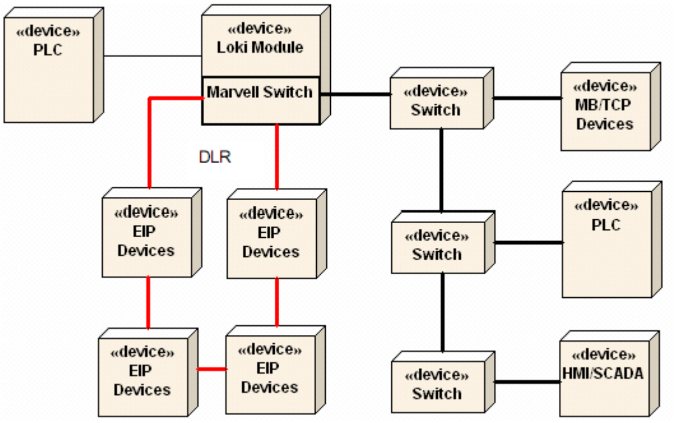

Figure 2. The EIP end device DLR ring topology with a ICIE as a NON-DLR switch.

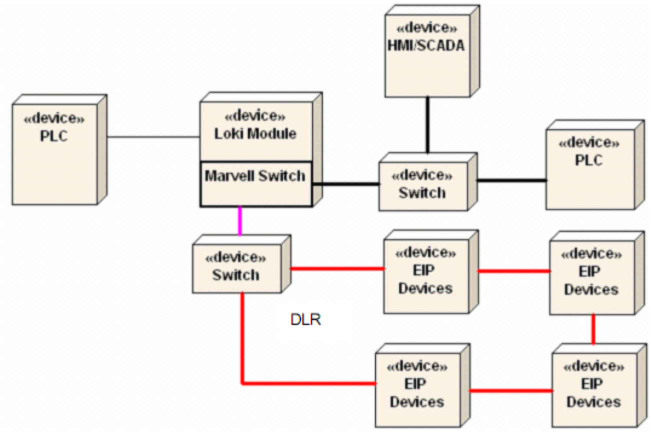

Figure 3. The EIP end device DLR ring topology with a ICIE as a NON-DLR end device.

3. DSA Tagged BPDU

The Ethernet switch chip selected for ICIE is from the Marvell Link Street family of integrated networking devices with part number of 88E6165.

Multicast Congestion control can be handle by router or end to end entity which may be done by source based approach as well as receiver driven approach. Source based approach is not much efficient and can not handle heterogeneous receivers. Receiver driven approach based on concept that all active decision is taken by receiver. This approach use layered transmission technique in which incoming stream is divided into different layer depend on the QoS requirements [6].

A. 88E6165 overview

The Marvell 88E6165 device embedded in ICIE is a single-chip 6 port Gigabit Ethernet switch with five integrated Gigabit Ethernet transceivers. RSTP, as a loop-prevention protocol or algorithm, heavily relies on this device to run in ICIE. To support RSTP protocol, this device provides the following two special features:

1) Support of the Marvell Distributed Switching Architecture (DSA) for RSTP and CPU-directed packet processing

2) Port States & BPDU handling for RSTP

B. BPDU and Ether Type DSA Tag

BPDU stands for Bridge Protocol Data Unit. Three types of BPDUs are defined in the reference document, which are Configuration BPDU for STP compatibility, Topology Change Notification BPDU (TCN BPDU) for STP compatibility and Rapid Spanning Tree BPDU (RST BPDU) for RSTP.

The BPDUs are special switch management data frames used for inter-bridge communications in order to run RSTP protocol algorithm. Below is how the BPDUs should be managed in ICIE based on the RSTP requirement and single Marvell switch device 88E6165 used [7].

A BPDU frame entering an external port needs to go to the CPU only for processing even though it is a multicast frame (DA=01:80:C2:00:00:00) and the CPU needs to know the Source Port information for this BPDU to be able to run the RSTP algorithm [8].

After the BPDU is processed, the CPU may need to send the updated BPDU to the switch device for transmission and, for this case, the switch device must be told though which port the BPDU egresses out [9].

TABLE I. 4 TYPES OF DSA TAG MODE

DSA Tag mode	Description
Forward	For normal frames
To_CPU	For MGMT (management) or control frames that needs to go to the CPU. Used by RSTP to receive BPDUs.
From_CPU	For MGMT (management) or control frames that come from the CPU. Used by RSTP to transmit BPDUs through the CPU port.
To_Sniffer	For MGMT (management) or control frames used for chip-to-chip communication.

BPDU frames need to be tunneled through blocked ports. If an external port is in the blocked state, all frames are discarded except for frames with a DA address that is considered a MGMT address (DA=01:80:C2:00:00:00 for RSTP). The CPU port on the switch device must be configured in a DSA Tag mode (FrameMode in Port Control Register– offset 0x04). The BPDU’s DA (01:80:C2:00:00:00) must be loaded into the switch’s ATU as a static entry associated with the CPU port [10].

There are four major standard DSA Tag mode frame types as described in the table I: 4 types of DSA Tag mode. It can be also in Distributed Switch Architecture (DSA) Ports of the reference document.

An alternate DSA Tag mode, called Ether type DSA Tag is supported and shown in Figure 4. Ether type DSA encapsulates the standard DSA Tags, described above, after a programmable Ether type. The Ether type DSA mode is optimized for switch to CPU interconnections for the reasons haven’t been discussed in this paper for pages limitation [11].

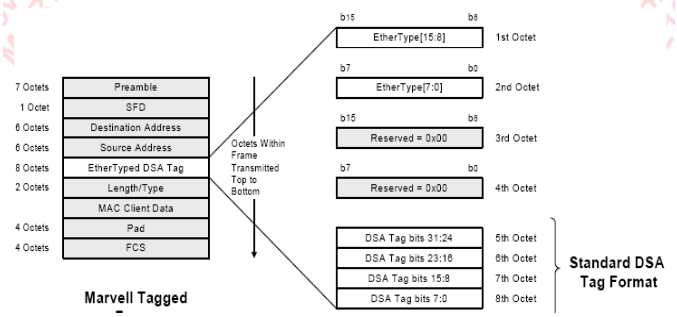

Figure 4. Ether Type DSA Format

4. Actions Sequence

In this version of the ICIE module architecture, the IGMP Snooping feature of the Marvell switch must be always enabled so that the Marvell switch always passes the IGMP packets to the ICIE module (CPU) on top of its other responsibilities [12].

In the ICIE module, the NetMux component (task) is responsible for the identification of the IGMP packets (query, join report, and leave report); the IGMP component inside the TCP/IP Stack subsystem is responsible for managing the IGMP messages as a host; the Switch Management subsystem (IGMP Snooping Task) is responsible for determining whether the received IGMP packets need passing to the IGMP component or not and performing the IGMP snooping operations (updating the Marvel 88e6165 switch multicast table and sending IGMP queries) [13].

C. Receiving IGMP Join Message

This section analyzes the IGMP snooping behavior of the ICIE module when receiving an IGMP join message from another host in the network. In this case, the IGMP snooping task should forward it through all the router ports. The IGMP snooping task should create a new entry for the receiving port and the multicast group pair in the multicast filter table and start an age timer if no forwarding entry exists for the reported group in the multicast filter table. If the entries for the reported group already exists but without the receiving port included, the IGMP snooping task should add the receiving port and start its age timer. If the entry for the reported group and port pair exists, the IGMP snooping task should restart its age timer.

The IGMP snooping task should not pass the IGMP join message to the IGMP component (representing a host) to avoid the suppressing of the IGMP join message from the ICIE module to respond a query. This is one of the IGMP snooping functionalities. In IGMP V1 and V2 [14], if the host receives an IGMP report from another host in the same group to respond a query before it sends the report, it will not send its own IGMP report. The suppression mechanism in IGMP V1 and V2 prevents the switch from knowing whether there are hosts still interested in the reported group from the host ports.

D. Receiving IGMP Leave Message

This section analyzes the IGMP snooping behavior of the ICIE module when receiving an IGMP leave message from another host in the network. In this case, the IGMP snooping task needs to search the switch multicast table to see if the receiving port and the multicast group pair already exists. If no, the IGMP snooping task should discard the IGMP Leave message. If an entry for the receiving port and the multicast group pair already exists, the IGMP snooping task should forward the IGMP Leave message through all the router ports; the IGMP snooping task should not remove the entry immediately and restart the aging timer instead before it figures out that these are no hosts on the receiving port which are interested in the reported group; the IGMP snooping task should send an IGMP group specific query through the receiving port to see if there are still other hosts which are interested in the reported group attached to this receiving port. If no report received before the aging timer expires, the IGMP snooping task should remove the entry from the multicast table [15].

E. Receiving IGMP Query Message

This section analyzes the IGMP snooping behavior of the ICIE module when receiving an IGMP query message. In this case, the IGMP snooping task should forward the IGMP query through all the ports except the receiving port. The IGMP snooping task also needs to pass the IGMP query to the IGMP component because it represents the ICIE module host.

The IGMP snooping task also needs to search the router list table (if the switch does not have one, the Switch Management subsystem needs to maintain a router list table?) to see if the receiving port already exists. If yes, reset the aging timer. If no, the IGMP snooping task should create a new entry for the receiving port in the router list [16].

F. Sending IGMP Join Message

The IGMP component (representing the ICIE module host) will send the IGMP join message in two circumstances:

● Upon receiving an IGMP query, the IGMP component sends an IGMP join message to respond the query. This case is already addressed in the section above.

● When some application component in the ICIE module system wants to join an multicast group, the IGMP component sends an IGMP join message to the multicast router (or switch) to indicate its interest in the multicast data of the joined multicast group.

The NetMux task should distinguish these two cases. In the second case, the NetMux task should pass the IGMP join message to the IGMP snooping task. The IGMP snooping task should create a new entry for the CPU port and the multicast group pair in the multicast filter table and start its aging timer. At the same time, the IGMP snooping task should forward the IGMP join message through all the router ports.

G. Sending IGMP Leave Message

This section analyzes the IGMP snooping behavior of the ICIE module when sending an IGMP leave message. In this case, the NetMux task should pass the IGMP leave message to the IGMP snooping task. The IGMP snooping task should remove the entry for the CPU port and the multicast group pair in the multicast filter table. At the same time, the IGMP snooping task should forward the IGMP leave message through all the router ports.

H. Network Topology Change

The ICIE module figures out the network topology change in three circumstances:

●when the RSTP task receives a network Topology Change Notification(TCN) message.

●when RSTP task figures out the network Topology Change by comparing its stored network topology information with the received Configuration Message.

● when RSTP task changes the switch port state.

This section analyzes the IGMP snooping behavior of the ICIE module when the network topology change occurs. In this case, the RSTP task should inform the IGMP Snooping task. The IGMP Snooping task should send the IGMP general query message through each new forwarding port. Then the IGMP Snooping task needs to set up the multicast filter entries for each new forwarding port based on the responses received to the general query. For the existing forwarding ports, the IGMP Snooping task should not change their existing multicast filter entries in the multicast table so that the known multicast traffics will still be forwarded to those ports only. During the set up period of the new multicast filter entries in the multicast table for the new forwarding ports, the switch still floods the unknown multicast traffics.

5. Conclusions and Perspective

As one of the most important elements of streaming architecture is control the network traffic. Network traffic evolves issues like rate control also called as flow control or congestion control. It is very crucial to resolve congestion state to maintain the flow of streams. It is important to control multicast packets for Ethernet switch. Congestion becomes more important at the multicast scenario where the entire receiver may have capability to adapt different bandwidth. In this paper, we discussed a new mechanism to minimize the congestion which is based on the taking an adaptive decision during transferring multicast messages. Proposed approach is that a device requesting to start and stop the reception of the multicast streams is accomplished through IGMP join and Leave message requests. It also focuses on the sequence study on IGMP snooping. It has a further discussion on four actions: receiving IGMP loin message, receiving IGMP leave message, receiving IGMP query message, sending IGMP join message, sending IGMP leave message and network topology change. The IGMP Snooping component monitors (snoops) these join and leave messages to allow it to know which streams to prune from which ports. This process uses a device performing a manager role to periodically query all devices in the subnet and subsequently cause them to re-join the multicast group of listeners for any stream in which they may be interested.

References

[1] Postel,J.:Internet protocol,Request for Comments 791,September, 1981.

[2] Qian Zhang,Quji Guo,Qiang Ni,Wenwu Zhu,and Ya-Qin Zhang.: Source adaptive multi-layered multicast algorithms for realtime video distribution, IEEE/ACM Transactions on Networking,vol.8,no.6. pp.720-733,2006.

[3] S.McCanne,M.Vetterli,and V.Jacobson.: Low-complexity video coding for receiver driven layered multicast,IEEE Journal on Selected Areas in Communications, vol.15, no.6, pg.982-1001,1997.

[4] Satish Kumar,Pavlin Radoslavov,David Thaler,Cengiz Alaettinoglu, Deborah Estrin,Mark Handley.: The MASCBGMP Architecture for Inter-domain Multicast Routing,in ACM SIGCOMM, April 1998,pp. 93 to 104.

[5] Stian Johansen,Anna N.Kim,Andrew Perkis.: “Quality Incentive Assisted Congestion Control for Receiver-Driven Multicast” IEEE Communications Society ICC 2007.

[6] Deering.S.:Multicasting Routing in Internetwork and Extended LANs, SIGCOMM Summer 1988 Proceeding,Aug 1988.

[7] Distance Vector Multicast Routing Protocol,Request for Comments 1075,November 1988.

[8] J.Byers,M.Frumin,et al.,"FLID-DL:congestion control for layered multicast",in Proc.NGC2000,Palo Alto,USA,pp.71-81,Nov.2000

[9] J.C.Bolot,T.Turletti,and I.Wakeman.:Scalable feedback control for multicast video distribution in the internet, Conference of the Special Interest Group on Data Communication ACM,pages 58-67,SIGCOMM’ 1994.

[10] J.C.Bennett,H.Zhang.:"Hierarchical packet fair queuing algorithms", IEEE/ACM Trans. on Networking,Vol.5(5), pp.675-689,1997.

[11] Johanson,M.:Scalable Video Conferencing Using Subband Transform Coding and Layered Multicast Transmission, International Conference on Signal Processing Applications and Technology(ICSPAT’99), Orlando,Florida,Nov.1-4,1999.

[12] Karan Singh,Rama Shankar Yadav,Manisha Manjul,Rainu Dhir.: Bandwidth Delay Quality parameter Based Multicast Congestion Control”published in International Conference on Advanced Computing and

Communication(ADCOM 08)at Department of Information Technology, MIT, Anna University,Chennai,2008

[13] Kimura J.,Tobagi,F.A.,Pulido,J.-M.and Emstad,P.J.: Perceived Quality and Bandwidth Characterization of Layered MPEG-2 Video Encoding,Proceedings of the SPIE International Symposium on Voice, Video and Data Communications,Boston,Sept.1999.

[14] Kwon,G.I.and Byers,J.:Smooth multirate multicast congestion control, Proceedings of IEEE Infocom,vol.2,pp.1022-1032,March 2003.

[15] L.Vicisano,L.Rizzo,and J.Crowcroft.:TCPlike congestion control for layeredmulticast data transfer, in Proc. Conference on Computer Communications,pg.996-1003,March 1998.

[16] Legout A.Legout and E.W.Biersack.:Pathological behaviors for RLM and RLC,International Conference on Network and Operating System Support for Digital Audio and Video,Chapel Hill,NC,USA,pg.164-172,June 2000.