PBB-TE 기반의 패킷전송시스템에서 멀티캐스트 서비스와 계층적 QoS 구현

정회원 이 원 경*, 최 창 호*, 김 선 미*

Point-to-Multipoint Services and Hierarchical QoS on PBB-TE System

Wonkyoung Lee*, Chang-Ho Choi*, Sun-Me Kim* *Regular Members*

ABSTRACT

We have proposed a solution to multicast services and an advanced quality of service (QoS) mechanism on a packet transport system (PTS) based on PBB-TE. The point-to-multipoint (PtMP) connection in the PBB-TE system have been realized by grouping point-to-point (PtP) PTL trunks and mapping a BSI onto the PtP PTL trunks using a multicast backbone destination address. To provide end-to-end QoS of the PtMP services, the hierarchical QoS scheme for backbone service instances and connection-oriented paths has been implemented in the PTS. For providing different capabilities for service selection and priority selection, the PTS offers to customers three basic types of the port-based, C-tagged, and S-tagged service interface defined by the IEEE 802.1ah. To offer to customers different capabilities of the layer 3 applications and services, moreover, an IP-flow service interface have been added. In order to evaluate traffic performance for PtMP services in the PTS, the PtMP throughputs for the link capacity of 1 Gbps at the four service interfaces were measured in the leaves of the ingress edge node, the transit node, and the egress edge node. The throughputs were about 96 % because the B-MAC overhead of 22 bytes occupies 4 % of the 512-byte packet. The QoS performance is ability to guarantee an application or a user a required bandwidth, and could be evaluated by the accuracy of policing or shaping. The accuracy of the policing scheme and the accuracy of the shaping scheme were 99 % and 99.3 % respectively.

Key Words : Point-to-multipoint, PBB-TE, hierarchical QoS, MAC-in-MAC, policing, shaping

Ⅰ. Introduction

Due to rapid growth of demand for bandwidth in today's network, the network has been evolved into simpler and more efficient structure. In this situation, the core network or the backbone network has been replacing SDH/SONET platforms with packet transport platforms[1]. The packet transport technology such as provider backbone bridge – traffic engineering (PBB-TE) and MPLS transport profile (MPLS-TP) is getting the spotlight as a key point of the next generation network. Provider backbone bridge – traffic engineering (PBB-TE) defined by IEEE

802.1Qay[2] is representative carrier Ethernet transport technology that extends well-known and widely distributed Ethernet services to core of the public network while maintaining simplicity, flexibility, and cost effectiveness of the Ethernet service[3]. The PBB-TE adds transport hierarchy of MAC-in-MAC encapsulation to Ethernet frames and provides traffic engineering for connection-oriented paths and protection switching within 50 ms.

In the next generation network, the PBB-TE technology should provide multicast video streaming services and support traffic engineering for end-to-end label switched paths. There have been no proper solutions to multicast services on packet transport platforms based on PBB-TE so far[4] since the PBB-TE technology does not allow MAC learning, spanning tree protocol, and broadcast of unknown frame for providing deterministic, protected, and connection-oriented trunks and services. Moreover, it has not been easy to classify layer 3 applications and services due to layer 3 service transparency of the carrier Ethernet transport. In this study, we propose a solution to multicast services and IP flow awareness that have been weak points of PBB-TE technology. Moreover, we propose advanced QoS mechanism based on the hierarchical QoS layer. We have implemented a packet transport system (PTS) based on the PBB-TE. The PTS provides multicast services, end-to-end QoS, and IP flow awareness. In order to emulate PtMP services in the PTS, we have measured traffic throughputs for the link capacity of 1 Gbps at various service interfaces. To evaluate the performance of the advanced quality of service (QoS) mechanism in the PTS, accuracies of the policing and shaping schemes have been measured.
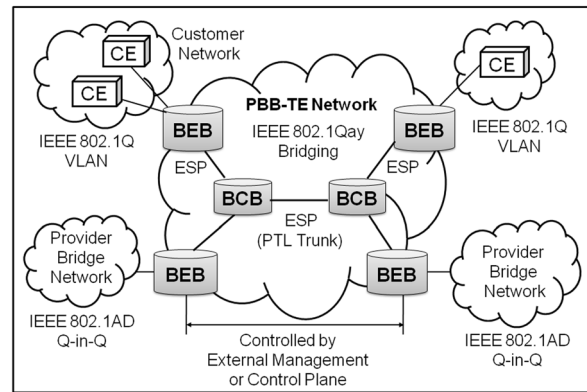
Ⅱ. Packet Transport System based on PBB-TE

Fig. 1. The PBB-TE network

As shown in Fig. 1, a PBB-TE network comprises a set of backbone edge bridges (BEBs) and backbone core bridges (BCBs) that are connected by Ethernet tunnels referred as Ethernet switched paths (ESPs)[2]. Backbone edge bridges are responsible for adding transport hierarchy to customer frames in ingress edge nodes and restoring customer frames by removing the transport hierarchy in egress edge nodes. Backbone core bridges in transit nodes are responsible for swapping transport label or backbone VLAN ID (B-VID). Each ESP as a connection-oriented path is identified by the triplet of a backbone source address (B-SA), a backbone destination address (B-DA), and a B-VID. The ESP is also called as a packet transport layer (PTL) trunk and is provisioned by an external management system.
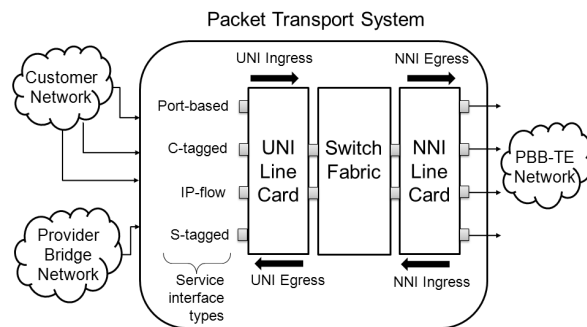

Fig. 2. The structure of the PBB-TE system

Figure 2 is the structure of the packet transport system based on the PBB-TE that we have implemented. The PTS consists of user network interface (UNI) line cards, network to network

interface (NNI) line cards, and a switch fabric that connects between UNI line cards and NNI line cards. In order to offer to customers one or more types of service interfaces, each providing different capabilities for service selection, priority selection, and service access protection, the PTS has three types of customer service interfaces called the Port-based, C-tagged, S-tagged[5]. To offer to customers different capabilities of the layer 3 applications and services, moreover, we have added a type called the IP-flow as the customer service interface. The port-based service interfaces that attach to VLAN-unaware bridges (802.1D bridges), routers or end-stations, classify a backbone service instance (BSI) by just an input port. The C-tagged service interfaces that attach to C-VLAN bridges (802.1Q bridges), classify a BSI by an input port and a C-VLAN identifier. The S-tagged service interfaces that attach to a provider bridged network, map a service instance identified by an S-VID to a backbone service instance identified by an backbone service instance identifier (I-SID). IP-flow service interface subdivides layer 2 flows using 5-tuple information of layer 3 such as a destination IP address, a source IP address, a destination port, a source port, and a protocol. If a port of the UNI line card receives a customer frame, a backbone service instance is allocated to the customer frame and then a PTL trunk is created in the NNI line card. The B-SA and the B-DA are MAC addresses assigned to NNI line cards.

Ⅲ. Implementation of Bidirectional Point-to-Multipoint Services

Figure 3 shows a point-to-multipoint (PtMP) service that is the virtual Ethernet communication between one root that distributes multicast video stream and multiple leaves that receive multicast video stream simultaneously. In the PtMP communication, the B-DA is a multicast backbone MAC address that determines output ports on which replicated packets are transmitted, and the

B-SA is a unicast backbone MAC address. In the NNI egress and the transit egress, the MAC-in-MAC frames of IEEE 802.1ah are replicated as many as the number of output ports. The PtMP communication in the backward direction is transmission from single leaf to one root. Packet forwarding in the backward direction is the same with that of the point-to-point connection.
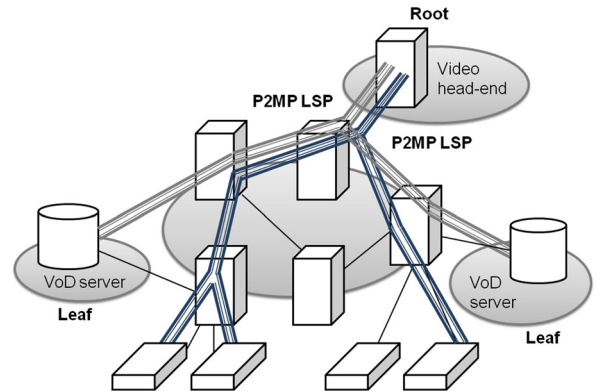

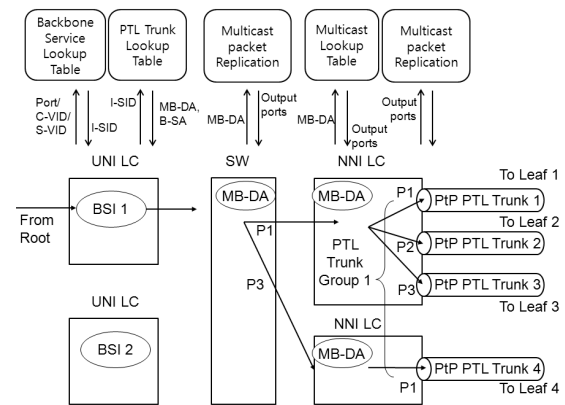Fig. 3. A point-to-multipoint service

Fig. 4. The proposed method for the PtMP service in the PTS system

Figure 4 shows the proposed method for the PtMP service in the PTS system based on the PBB-TE. A backbone service instance is allocated to a customer frame by the backbone service lookup table in the UNI LC according customer service interfaces. PtP PTL trunks connected with each leaf node are created in the NNI LC. A multicast B-DA (MB-DA) maps the backbone service instance (BSI) with PtP PTL trunks. In the switch fabric (SW), the MAC-in-MAC frames are copied as many as the number of output

435

slots. In the NNI LC, the MAC-in-MAC frames are replicated as many as the number of output ports in a slot. By grouping PtP PTL trunks and mapping a BSI onto the PtP PTL trunks using a multicast B-DA (MB-DA), PtMP connection in the PBB-TE system could be realized.

IV. Implementation of End-to-End QoS

The PTL trunks that built up PtMP connection are traffic engineered ESPs. Differentiated QoS according service policies should be applied while providing multicast services and connection-oriented paths. To apply differentiated QoS policies for both customer and backbone services in a cost-effective way, we have proposed the advanced QoS mechanism based on hierarchical QoS layer[5]. QoS layer for customer services are focused on sophisticated bandwidth control and QoS layer for backbone services are focused on minimizing data loss.
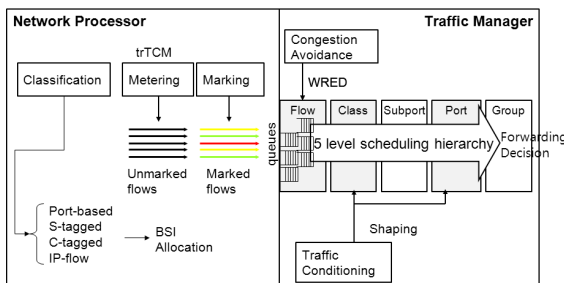

Fig. 5. Advanced QoS mechanism based on hierarchical QoS

Figure 5 describes the advanced QoS mechanism implemented in the PTS. For sophisticated bandwidth control for customer services, policing method has been used as traffic conditioning mechanism for the BSI. The policing method meters packets based on a policing profile of a committed information rate (CIR), an excess information rate (EIR), a committed burst size (CBS), and an excess burst size (EBS) by using the two rate three color marker (trTCM) algorithm specified by RFC 2698[6]. And then, packets are marked with one of three colors according to the metering results of confirm,

violate, and exceed. Different actions such as transmit, drop, and remark are taken depending on the color of the marked packet. Policing accuracy is important because it is an index indicating how minutely to control bandwidth for customer services. Traffic conditioning mechanism for the PTL trunk has used shaping method to minimize data loss in the PTL trunk of connection-oriented path. The shaping method buffers packets instead of dropping them in the case of congestion. Therefore, shaping accuracy is an index indicating how well to guarantee the reserved bandwidths of PTL trunks without data loss. Traffic manager has scheduling hierarchy of 5 levels. The weighted random early discarding (WRED) algorithm, specified by RFC 2309[7], takes charge of actions for the marked packets. The policing process of a BSI is applied to one out of 64k flow queues of level 4 and the shaping of a PTL trunk is applied to one out of 4k tunnel queues of level 3 in the traffic manager. When a network operator configures the differentiated QoS bandwidth per customer service, the traffic policing is applied at the network processor and the traffic manager in the UNI linecard that attaches the customer service. The traffic shaping is applied at the traffic manager in the NNI linecard that attaches the backbone service when the network operator configures the differentiated QoS bandwidth per backbone service.

V. Experiments and Results

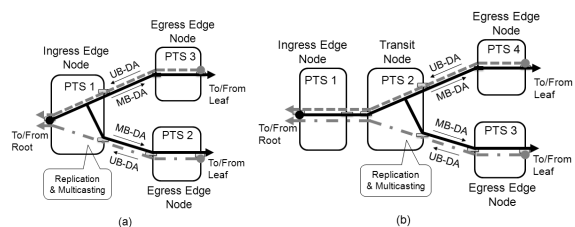### V.1. Emulation of Bidirectional PtMP Services in the PBB-TE System

Fig. 6. Examples of bidirectional point-to-multipoint services using the packet transport systems
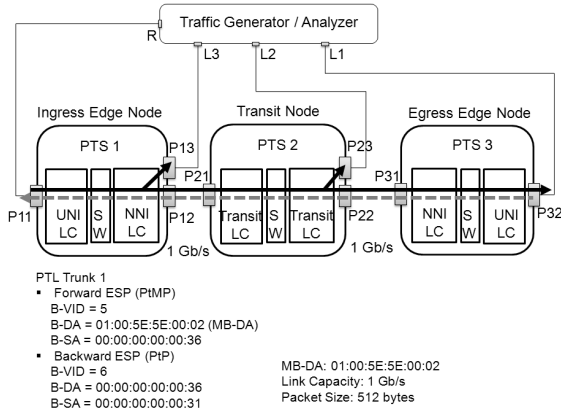
436

Fig. 7. Experimental configuration to measure throughput and QoS performance for the bidirectional PtMP of the PTS

Figure 6 (a) and (b) show examples of bidirectional point-to-multipoint services using the packet transport systems. In the Fig. 6(a), packets are replicated and multicasted in the ingress edge node. In the Fig. 6(b), packets are replicated and multicasted in the transit node. Figure 7 describes the test configuration for the bidirectional PtMP services considering both Fig. 6(a) and Fig. 6(b). Packets transmitted from the traffic generator were inputted into the P11 port of the PTS 1. And then, packets were duplicated in the NNI LC of the PTS 1 and transmitted to the P12 and P13 ports simultaneously. Packets transmitted from the P12 port of the PTS 1 were inputted into the P21 port of the PTS 2 and were duplicated as many as output ports configured by the multicast B-DA of 01:00:5E:5E:00:02 in the transit LC of the PTS 2 and were output to the P22 and P23 ports simultaneously. Packets transmitted from the P22 port of the PTS 2 were inputted into the P31 port of the PTS 3, and then, packets in which B-DA, B-SA, and B-VID were removed, transmitted to the P32 port of the PTS 3.

Customer packets at the rate of about 960 Mbps except inter-packet gap and preamble were transmitted from the R port of the traffic generator to the P11 port of the PTS 1 with the four service interfaces. The customer packets were untagged frames, C-tagged frames with the C-VID of 0x64, and S-tagged frames with the S-VID of 0xC8 according to the service interface types of

port-based, C-tagged, IP-flow, and S-tagged. Packets were received simultaneously at the L3, L2, and L1 ports of the traffic analyzer connected with the P13 port of the PTS 1, the P23 port of the PTS 2, and the P32 port of the PTS 3.

Figure 8 described the frames captured at the L3, L2, and L1 ports. The frames output from the P13 and the P23 ports were MAC-in-MAC frames encapsulated with backbone VLAN tags (B-TAGs), on the other hand, the frames transmitted from P32 port of the PTS 3 were the customer frames that B-TAG was removed in the egress edge node. The backward transmission means that packets are forwarded from one leaf to one root and it is not point-to-multipoint connection but point-to-point connection. When customer packets at the rate about 960 Mbps transmitted from the L1 port of the traffic generator to the P32 port of the PTS 3, packets were received at the only R port of the traffic analyzer connected with the P11 port of the PTS 1.

Fig. 8. Frames received in the three leaves

In order to evaluate traffic performance for PtMP services in the PTS, we have measured PtMP throughputs at output ports of the ingress edge node, the transit node, and the egress edge node. Each input traffic rate for port-based, C-tagged, S-tagged, and IP-flow service interfaces was 1 Gbps including IPG and preamble. To exclude impacts of policing and shaping, policing and shaping set to be off. Table 1 shows experimental results of the throughput for the

bidirectional PtMP service. Throughputs for 512-byte packet and physical link capacity of 1 Gbps were measured 100 % at the two ports of the ingress edge node and the transit node. In the egress edge node, however, 4 % of the traffic dropped. It attributed that the B-MAC overhead of 22 bytes occupies 4 % of the 512-byte packet. Figure 9 explains the impact of the B-MAC overhead that occupying the packet on the traffic throughput according to packet sizes. The larger the packet size is, the smaller the percentage of the B-MAC overhead is, and then, the fewer packets drops. In bidirectional PtMP connection, we have measured the throughput of 96 % after the egress edge node and the throughputs of 100 % after the ingress edge node and the transit node at four service interfaces.
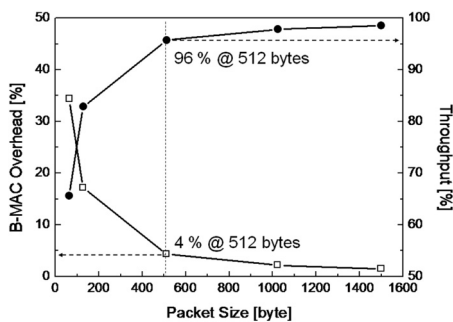

Fig. 9. Impact of B-MAC overhead on the traffic throughput according to packet sizes

Table 1. Experimental results for throughput of the bidirectional PtMP services

Service Interface		Traffic Generator/Analyzer			
		R port [Mbps]	L1 port [Mbps]	L2 port [Mbps]	L3 port [Mbps]
Port -based	TX	963.3	962.3		
	RX	924.1	924.1	963.8	963.9
	Through -put	96 %	96 %	100 %	100 %
C -tagged	TX	962.3	962.3		
	RX	923.6	923.6	963.6	963.6
	Through -put	96 %	96 %	100 %	100 %
IP-flow	TX	962.3	962.3		
	RX	924.1	924.1	963.8	963.9
	Through -put	96 %	96 %	100 %	100 %
S -tagged	TX	962.3	962.3		
	RX	924.1	924.1	963.8	963.9
	Through -put	96 %	96 %	100 %	100 %

V.2. Performance Measurement of Hierarchical QoS

Quality of service is the ability to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow. For example, a required bit rate, delay, jitter, packet dropping probability and/or bit error rate may be guaranteed. Quality of service guarantees are important if the network capacity is insufficient, especially for real-time streaming multimedia applications such as voice over IP, online games and IP-TV, since these often require fixed bit rate and are delay sensitive, and in networks where the capacity is a limited resource. The QoS performance means the ability to guarantee a application or a user a required bandwidth, and could be evaluated by the accuracy of policing or shaping. In order to evaluate the performance of the end-to-end QoS of the PTS, we have measured policing accuracy and shaping accuracy for BSI and PTL trunk at C-tagged service interface.

Figure 10 describes QoS bandwidth configuration to measure policing accuracy of the BSI in the PtMP connection. Output traffic rate is controlled by only CIR and EIR of the BSI 1 because BSI 1 is smaller than input traffic rate. Traffic rate transmitted to the PTL trunk 1 is smaller than CIR of the PTL trunk 1, therefore, the traffic doesn't be limited by shaping of the PTL trunk. The traffic generator sent traffic with rate of 500 Mbps from R port of the traffic generator to the P11 port of the PTS 1 and measured output traffic rates at the P13 port of the PTS 1, the P23 port of the PTS 2, and the P32 port of the PTS 3. Using measured and expected output traffic rates, the policing accuracy of the BSI was calculated.
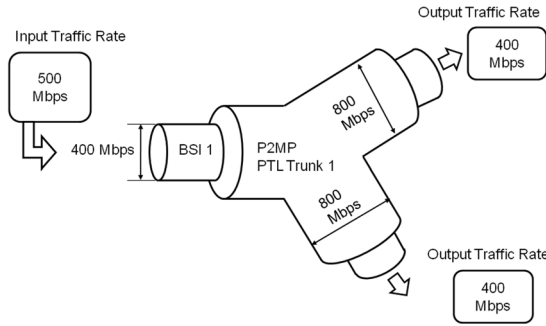
438

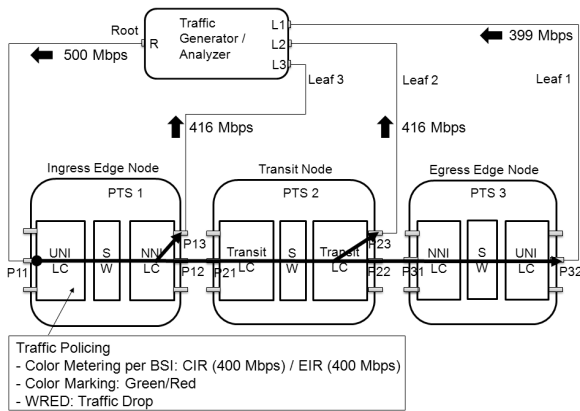Fig. 10. QoS bandwidth configuration for BSI QoS test

Fig. 11. Expected output traffic rates in the experiment for measuring the policing accuracy

Figure 11 explains the expected rate of output traffic. If the P11 port of the PTS 1 receives C-tagged frames at the rate of 500 Mbps, IPG of 12 bytes and preamble of 8 bytes within the frames are removed and then network processor (NP) of the UNI LC receives the traffic at the rate of 481 Mbps. The NP in the UNI LC of the PTS 1 performs color metering and marking process using the policing profile that is shown in Fig. 11. Traffic up to 400 Mbps is marked with green color and the rest of 81 Mbps is marked with red color. B-MAC overhead of 22 bytes is inserted in the color-marked frames, therefore, the traffic rate marked with green color is 416 Mbps and the traffic rate marked with red color is 84.45 Mbps. Only green packets are forwarded to P13 and P12 ports of the PTS 1 and red packets are dropped by WRED algorithm of the traffic manager in the PTS 1. PTS 2 swaps the MPLS label. As a result, the L3 and L2 ports of the traffic analyzer will receive packets at the rate of

416 Mbps. PTS 3 is the egress edge node so B-TAG decapsulation is required. L1 port of the traffic analyzer will receive packets at the rate of 399 Mbps because of removal of B-MAC overhead and restoration of customer frames.

$$Policing/Shaping\ Accuracy[\%]$$
$$= \left[1 - \left| \frac{R_{\mathrm{expected}} - R_{measured}}{R_{\mathrm{expected}}} \right| \right] \times 100 \quad (2)$$

where $R_{expected}$ and $R_{measured}$ is expected rate and measured rate of output traffic, respectively. Table 2 describes expected and measured rates of output traffic. Policing accuracy is defined as the ratio for difference between the expected rate and the measured rate of output traffic as shown in Eq. 1. Policing accuracies for a BSI in the three leaves were about 99 % in the PtMP connection.

Table 2. Experimental results for policing accuracy

Traffic Analyzer	Expected Rate	Measured Rate	Policing Accuracy (Eq. 1)
L1 Port	416 Mbps	420 Mbps	98.92 %
L2 Port	416 Mbps	420 Mbps	98.92 %
L3 Port	399 Mbps	403 Mbps	98.97 %

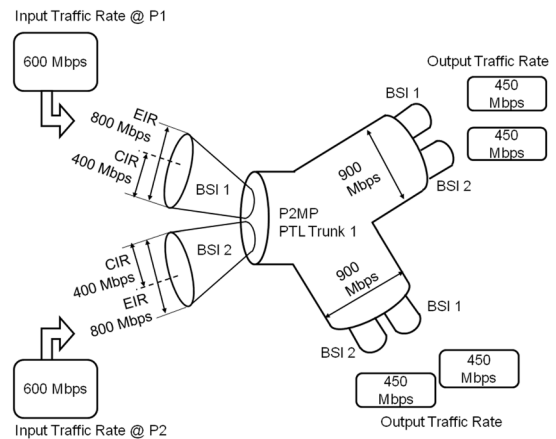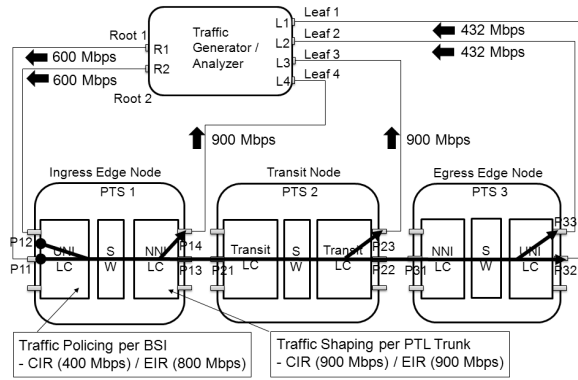Fig 12. QoS bandwidth configuration for the QoS test of a PTL trunk

Fig. 13. Expected output traffic rates in the experiment for measuring the shaping accuracy

Figure 12 shows QoS bandwidth configuration to measure shaping accuracy of a PTL trunk in PtMP connection. Two BSIs with CIR of 400 Mbps and EIR of 800 Mbps respectively were mapped to one PTL trunk with CIR of 900 Mbps and EIR of 900 Mbps. Output traffic rate is controlled by only CIR and EIR of the PTL trunk 1. The traffic doesn't be limited by policing because EIRs of two BSIs are larger than input traffic rate and congestions don't happen. Figure 13 explains the expected output traffic rates. If P11 and P12 ports of the PTS 1 receive C-tagged frames at the rate of 600 Mbps respectively, IPG of 12 bytes and preamble of 8 bytes within the frames are removed and then network processor (NP) of the UNI LC in the PTS 1 receives the traffic at the rate of 577 Mbps respectively. After color metering and marking process in the UNI LC of the PTS 1, traffic up to 400 Mbps is marked with green color and the rest of 177 Mbps is marked with yellow color. B-MAC overhead of 22 bytes is inserted in the color-marked frames, therefore, the traffic rate marked with green color is 416 Mbps and the traffic rate marked with yellow color is 184 Mbps. Both green packets and yellow packets are entirely forwarded because the sum of the green and yellow packets is smaller than the EIR of BSI 1. The shaping profile has CIR of 900 Mbps and EIR of 900 Mbps. Therefore, all green packets are forwarded but yellow packets are forwarded up to the rate of 68 Mbps. The rest of

yellow packets are dropped at the rate of 233 Mbps. As a result, the L3 and L4 ports of the traffic analyzer will receive packets at the rate of 900 Mbps. After removal of B-MAC overhead in the egress edge node of the PTS 3, The L1 and L2 ports of the traffic analyzer will receive traffic at each rate of 432 Mbps. The shaping accuracy like the policing accuracy is defined as the ratio for the difference between the expected rate and the measured rate of output traffic. Shaping accuracies for PTL trunk were measured about 99.3 % as shown in Table 3.

Table 3. Experimental results for policing accuracy

Traffic Analyzer	Expected Rate	Measured Rate	Shaping Accuracy (Eq.1)
L3 port	900 Mbps	906 Mbps	99.3 %
L2 port	900 Mbps	906 Mbps	99.3 %

Ⅵ. Conclusions

The PBB-TE technology is the carrier Ethernet transport technology that provides connection-oriented Ethernet, end-to-end QoS, and robust OAM. However, there have been no proper solutions for multicast services in the PBB-TE technology since the PBB-TE technology does not allow MAC learning, spanning tree protocol, and broadcast of unknown frame for providing deterministic, protected, and connection-oriented trunks and services. Moreover, it has not been easy to classify layer 3 applications and services due to layer 3 service transparency of the carrier Ethernet Transport. In this study, we have proposed a solution to multicast services and advanced QoS mechanism based on the hierarchical QoS layer. And, We have added IP flow awareness function to offer to customers different capabilities of the layer 3 applications and services.

We have implemented the packet transport system based on PBB-TE, which provides the point-to-multipoint services, the hierarchical QoS,

440

and the IP flow awareness. The hierarchical QoS for the backbone service instance and the connection-oriented path has been implemented in the PTS. Traffic throughputs for the packet of 512 bytes and the link capacity of 1 Gbps have been measured at port-based, C-tagged, S-tagged, and IP-flow service interfaces. The throughputs of leaves with MAC-in-MAC encapsulation (IEEE 802.1ah) in the ingress edge node and the transit node were measured 100 %. After removing the MAC-in-MAC encapsulation in the egress edge node, 4 % of the traffic dropped. It attributed that the B-MAC overhead of 22 bytes occupies 4 % of the 512-byte packet. The larger the packet size was, the smaller the percentage of the B-MAC overhead was. Therefore the larger the packet size is, the impact of the B-MAC overhead of the PBB-TE technology on the traffic throughput will be reduced. The policing scheme was applied to the BSIs of the UNI side, whereas the shaping scheme was applied to the PTL trunks of the NNI side. In the point-to-multipoint service, the policing accuracy for the backbone service instance was measured almost 99 % and the shaping accuracy for the connection-oriented path was about 99.3 %. The network processor meters packets not by bit unit but by packet unit. Therefore, there could be the gap of 1 % between the expected results and the measured results for policing and shaping. This study supplements solutions for multicast and L3 level services of the PBB-TE technology, therefore the packet transport system based on PBB-TE could support broadband services such as IPTV or multicast video streaming.

References

[1] B. Pratt, "Multi-Layer Switch -ing in Packet Optical Transport System," OFC/ NFOEC 2008, pp. 1-6, 2008.

[2] Virtual Bridged Local Area Networks – Amendment: Provider Backbone Bridge Traffic Engineering, IEEE Std. P802.1Qay, 2008.

[3] K. Fouli and M. Maier, "The Road to Carrier-Grade Ethernet," IEEE Communications Magazine, 2009.

[4] Claus G. Gruber and Achim Autenrieth, "Carrier Ethernet Transport in Metro and Core Networks," 13th International Telecommunications Network Strategy and Planning Symposium, 2008.

[5] Virtual Bridged Local Area Networks – Amendment: Provider Backbone Bridge Traffic Engineering, IEEE Std. P802.1ah/D4.2, 2008.

[6] Wonkyoung Lee, and et al., "Implementation of hierarchical QoS mechanism on PBB-TE system," 9th COIN, 2010.

[7] J. Heinanen and R. Guerin, "RFC 2698 – A Two Rate Three Color Marker," IETF Std. RFC 2698, 1999.

[8] B. Braden, and et al., "RFC 2309 – Recommendations on Queue Management and Congestion Avoidance in the Internet," IETF Std. RFC 2309, 1998.

이 원 경 (Wonkyoung Lee)　　　　정회원
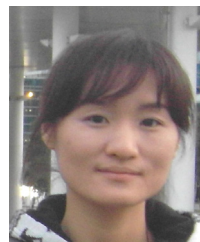
1999년 2월 부산대학교 전자공학과 졸업

2001년 2월 광주과학기술원 정보통신공학과 석사 졸업

2001년 2월~현재 한국전자통신연구원 재직

<관심분야> 광통신 공학, 광소자

최 창 호 (Chang-Ho Choi)　　　　정회원

1998년 2월 전북대학교 자원공학과 학사 졸업

2000년 8월 전북대학교 정보통신공학과 석사

2007년 3월 ~ 현재 충남대학교 전자·전파·정보통신공학과 박사과정

2000년 8월~ 현재 한국전자통신연구원 연구원

<관심분야> 캐리어 이더넷, Packet Transport Network, QoS, Packet processing

김 선 미 (Sun-Me Kim)   정회원
1991년 2월 충남대학교 전산학과 졸업
1993년 2월 포항공대 대학원 전산학과 석사 졸업
1993년 2월~ 현재 한국전자통신연구원 재직
<관심분야> 네트워크 인프라/시스템, 패킷 통신 프
　　로토콜