

Integrated Flow Management in Packet Transport System

Chang-Ho CHOI****, Whan-Woo KIM**

*Electronics and Telecommunications Research Institute, Korea

**Department of Electronics Engineering, CNU(Chungnam National University), Korea

chhchoi@etri.re.kr, wwkim@cnu.ac.kr

Abstract— Supporting guaranteed QoS is a significant issue in packet transport networks. Since various services require their own quality of service, flow-based traffic management that properly controls network resources when traffic congestion occurs is required for network nodes. In this paper, we propose an integrated flow management scheme to support guaranteed QoS in packet transport systems that comprise packet transport network nodes. The proposed integrated flow is generated using a flow processor to specify each service and is managed by a queue manager to control guaranteed bandwidth before transmitting to the packet transport network. The integrated flow is used to effectively achieve guaranteed quality bandwidth of service management for both Ethernet- and IP-based services. The proposed per-flow QoS guaranteed bandwidth management scheme is demonstrated using a throughput comparison between two test cases on a real system: the first test case is for non-per-flow traffic management, and the other is for per-flow traffic management. The experimental results show that per-flow bandwidth control satisfies the QoS requirements without being affected by violation traffic even if an incoming traffic rate is exceeded as defined by the Committed Information Rate (CIR).

Keywords— Packet Transport, Flow, Ethernet, QoS, Traffic Management

I. INTRODUCTION

One of today's main trends is the migration from existing time division multiplexing (TDM) technology into packet transport technology. TDM technology, such as a Synchronous Optical Network (SONET) or Synchronous Digital Hierarchy (SDH), has been used in legacy transport networks. Despite its high reliability and availability, TDM technology has a limitation in terms of the increasing demands of packet-oriented services.

A packet transport network (PTN) merges legacy and IP-oriented services. However, packet-based services are bursty and do not require dedicated transport bandwidth. While some services, such as best-effort service, have been used without quality of service (QoS) requirements, there are several services that do require specific QoS guarantees. QoS control is the most significant technology in a PTN, and several QoS control mechanisms have been proposed. IntServ [4] and DiffServ [5] are well known mechanisms. However, the IntServ and DiffServ mechanisms have problems with scalability and controllability, respectively.

Methods for providing QoS-guaranteed services in flow-based queue management have been researched over the past decades [1]-[3]. A flow or session is generally, but not always, defined as a 5-tuple IP flow. It can also be defined as an MPLS label for MPLS traffic or as an Ethernet Switched Path (ESP) label for a provider's backbone Ethernet traffic. In a flow-based system, traffic is first classified into the appropriate flow, and traffic control, congestion control, and congestion avoidance then occur at the flow level in an ingress node.

Ethernet technology was initially developed for local area communications. However, as the connections between regionally distributed networks and Internet access have grown in importance, the need for high-speed wideband network services has increased. Metro Ethernet technology was proposed to meet these requirements. Metro Ethernet provides a network configuration scheme for expanding an existing Local Area Network (LAN) into a wide area of a metropolitan size. Currently, Ethernet technology has been expanded from local areas to metro areas, and is aiming to be further expanded into backbone areas. Provider Backbone Bridge-Traffic Engineering (PBB-TE), as standardized in IEEE802.1Qay, was developed to support backbone services using Ethernet technology. Although PBB-TE discriminates Ethernet services by using virtual LAN IDs, accommodating new IP-oriented services is a serious issue.

For these reasons, integrated flow-based traffic management is needed along with provisioned QoS profiles to support QoS guaranteed service in a packet transport network. We focus on a packet transport network based on PBB-TE and propose an integrated flow in the Packet Transport System (PTS) to guarantee the QoS requirement. A packet is assigned to a specific flow that contains QoS profiles, and its color is then metered and marked. The packet is then mapped to a predefined packet transport label (PTL) and stored in a queue where traffic management is performed. If a packet exceeds the bandwidth limitation, it is dropped rather than being sent to the queue. We implement the proposed flow management scheme in a packet transport system and evaluate its traffic management performance on a real system.

The rest of this paper is organized as follows. In Section II, we briefly describe the PTN used. In Section III, we propose integrated flow management in a PTS, which is used to

effectively perform guaranteed quality of service for bandwidth management. Section IV summarizes our experiments and results. Finally, we offer some concluding remarks in section V.

II. PACKET TRANSPORT NETWORK BASED ON PBB-TE

Figure 1 shows a simple packet transport network architecture based on Provider Backbone Bridge-Traffic Engineering (PBB-TE); its standardization was finished in 2009 as IEEE802.1Qay. PBB-TE promises to combine traditional SONET/SDH connection capabilities and operational features with the lower costs and packet-oriented advantages of a carrier-grade Ethernet. PBB-TE is based on the existing IEEE802.1ah standard for Provider Backbone Bridging (PBB). PBB-TE disables the Ethernet spanning tree and address learning, and uses an external management plane for determining and deploying traffic-engineered paths [9].

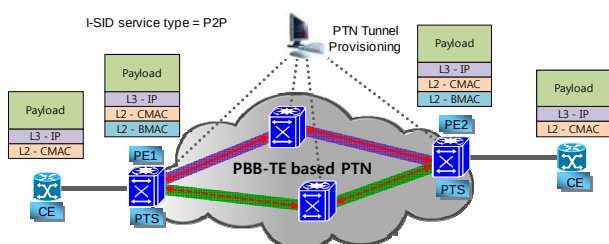


Figure 1. Packet transport network architecture based on PBB-TE.

Figure 2 shows the MAC-in-MAC frame format specified in IEEE802.1Qay PBB-TE. An Ethernet Switched Path (ESP) is defined to provide provisioned traffic-engineered unidirectional connectivity among two or more Provider Edge Switches, and is identified by the Backbone Destination MAC Address (B-DA), Backbone Source MAC Address (B-SA), and Backbone VLAN ID (B-VID). An ESP is point-to-point or point-to-multipoint.

A Packet Transport Label (PTL) is commonly used in a Packet Transport Network (PTN) and is equivalent to an ESP in IEEE802.1Qay.

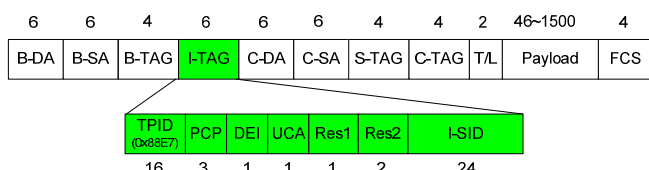


Figure 2. MAC-in-MAC frame format

III. INTEGRATED FLOW MANAGEMENT IN PTS

A. Classification and Flow Generation

In a PTS system, a flow is defined as an integrated flow from an L2 flow generated by L2 information, and from an L3 flow generated by L3 information. Figure 3 shows an algorithm used to generate an integrated flow at the flow processor. The flow processor is responsible for generating and managing the integrated flow, as shown in Figure 4.

When packets have been entered into a classifier to generate an L2 flow, the Ether type, Customer Destination MAC Address (C-DA), and VLAN ID are parsed. After parsing these parameters, the packet is sent to one of the lookup modules. Before classification of its Ether type, the port-based configuration is checked. If the port-based configuration is set, the packet is sent to the port-based lookup module regardless of its Ether type. Otherwise, if its Ether type is 0x8100 (customer VLAN tagged), the packet is sent to the Customer VLAN (C-VLAN) aware lookup module. Or, if its Ether type is 0x0800 (untagged), before sending it to a C-VLAN aware lookup module, a default VLAN ID is assigned to the packet. In addition, if its Ether type is 0x88a8 (service VLAN tagged), this packet is sent to a Service VLAN (S-VLAN) aware lookup module. Once the packet is sent to a C/S-VLAN aware lookup module, an L2 flow, defined as ETH_flow, is generated through a lookup using each VLAN ID and Destination MAC address. Finally, after L2 flow generation, the flow processor checks additional lookups to determine whether an L3 flow is configured or not. If an additional lookup has been set, the flow processor lookups the L3 table to generate an L3 flow, defined as IP_flow, and generates the final integrated flow using the previously generated ETH_flow and IP_flow. Otherwise, if an additional lookup has not been set, the final integrated flow is generated using only ETH_flow. The final integrated flow will be managed in an integrated flow table.

```

If (port based is set) then
    Lookup port based table
else
    If (ether type is 0x0800) then
        Assign default C-VID
        Lookup C-VLAN aware table
    end if
    If (ether type is 0x8100) then
        Lookup C-VLAN aware table
    end if
    If (ether type is 0x88a8) then
        Lookup S-VLAN aware table
    end if
end if
Generate and allocate flow and flow ID according to L2

If (additional lookup is set) then
    Lookup L3 table using information of L3
    Generate and allocate flow and flow ID according to L3
    Generate final integrated flow and flow ID by using L2 flow
    and L3 flow
end if

Lookup integrated flow table by using the integrated flow ID
If (lookup is succeeds) then
    Get a QoS profile and statistics counter index
    Update integrated flow table
else
    Add new flow in integrated flow table

```

Figure 3. Algorithm used to generate an integrated flow

The flow processor has an aging mechanism for deleting old entries from the integrated flow table. Each time an entry is matched, the entry is refreshed. The aging process deletes

all entries that have not been refreshed during a defined time period.

B. Metering and Marking

Metering refers to measuring a flow's traffic. Metering results determine whether a packet, which belongs to a flow, is within the expected traffic parameters for the flow, or whether it violates these parameters. Traffic rate and traffic burst sizes are parameters that are typically metered. Metering standardized in IETF [6], [7], MEF [8] uses token bucket schemes. A token bucket counter of up to 1 M is assigned for metering in the proposed system. After being metered, packets are marked to reflect their metering results. Marking is usually done for the next hop in the network to prevent congestion.

C. Traffic Management

A Weighted Random Early Discard (WRED) mechanism is used to perform traffic policing and congestion avoidance. The WRED mechanism chooses whether to drop or forward a packet. The WRED profile is a set of parameters, associating each color with a drop function, where each drop function describes the congestion level at which a drop decision should be made. When a packet is forwarded to the output queue, it is assigned a flow ID. Each flow ID is associated with a WRED profile in the output queue. The WRED mechanism selects a drop function from the profile according to the packet's color. Shapers at each level of the hierarchical traffic manager allow per-flow maximum bandwidth guarantees without packet drops while smoothing bursty traffic flows.

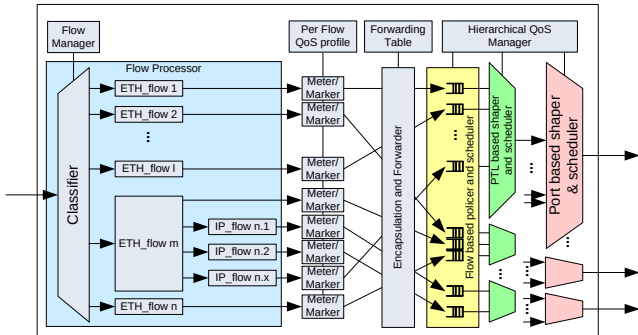


Figure 4. Conceptual architecture of flow-based traffic management in PTS

Figure 4 illustrates the conceptual architecture of flow-based traffic management in a PTS. Once a packet has entered the flow processor, per-flow traffic management procedures are followed. First, the packet is classified by the flow manager and an integrated flow is assigned, which was described in the previous paragraph. Next, the packet is metered and marked according to the QoS profiles. A packet with a flow ID is then mapped to 16K queues, and the WRED mechanism drops a packet when it exceeds the bandwidth limitation. Each of four queues from the 16K queues is mapped to one of 4K PTL queues. Queues within the same priority level have Weighted Fair Queuing (WFQ) enforced among them. Each PTL queue contains a dual leaky bucket shaper to manage packets that violate the traffic parameters. After being mapped to a PTL queue, the packet is mapped to

one of 32 logical ports. Finally, the packet is transmitted through a physical interface. Each PTL queue is assigned one of four priority levels, and the priority levels are propagated from the flow to the physical port to enable service level guarantees at all levels.

IV. EXPERIMENTS AND RESULTS

A. Test Environment

Figure 5 shows the experimental setup used to evaluate the performance of the proposed integrated flow-based traffic management in real time. The experimental setup comprises a packet generator and three integrated flow-based PTSs. The link between the packet generator and PTS has 10Gbps capacity for each line card. Four flows were set up in the packet generator for comparing the bandwidth control during periods of congestion.

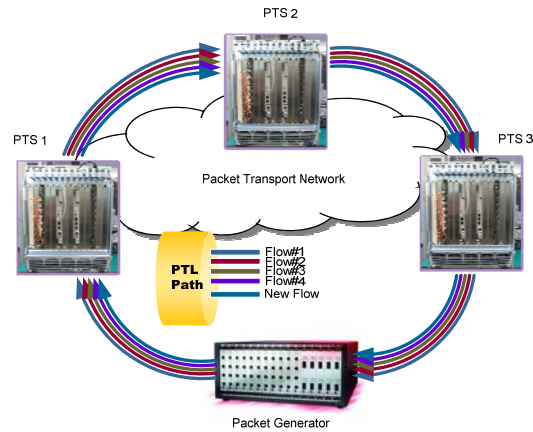


Figure 5. Experimental setup used to evaluate the performance of the integrated flow-based traffic management system.

B. Flow based Bandwidth control

Four flows were assigned to predefined flow IDs that have their own QoS parameters. Table 1 shows the QoS profiles for each flow and PTL. Flows 1, 2 and 4 are integrated flows whose IP-based services were specified as VoD traffic, while the Ethernet service for flow 3 was specified as best-effort traffic. Even though the flows were classified into different services, these flows can be mapped to a PTL since a PTL is defined as a provisioned transport path between two nodes and can contain one service type or more than two service types. This PTL was identified using the B-DA, B-SA, and B-VID, as described in section II.

The Committed Information Rates (CIRs) for the flows were set to 100, 200, 50, and 150Mbps for the first, second, third, and fourth flows, respectively. To support efficient bandwidth utilization, the Excess Information Rates (EIRs) for the flows were set to 300, 300, 500, and 200Mbps from the first to fourth flows, respectively.

A PTL that contains all four flows is limited to a bandwidth of 520Mbps. Although the total sum of the CIRs of the flows mapped to the PTL is 500Mbps, we have set the PTL bandwidth to 520Mbps to compensate for encapsulation overhead. As described in section II, a PBB-TE-based PTN

requires B-MAC header encapsulation at the ingress node, and B-MAC header decapsulation at the egress node. Since all packet lengths were fixed at 512bytes and the encapsulation overhead is 22bytes, 4% of the bandwidth was added into the PTL QoS parameters.

TABLE 1. QOS PROFILES FOR EACH FLOW AND PTL

Flow Index	Flow ID	CIR (Mbps)	EIR (Mbps)	PTL Index	PTL B/W (Mbps)
1	0x004001	100	300	1	520
2	0x000801	200	300	1	520
3	0x000C00	50	500	1	520
4	0x001001	150	200	1	520

Figure 6 shows the transmission duration of all flows, which were transmitted at different time intervals. The x axis represents the testing time, and the y axis represents the transmission rate in Mbps. First, flow 3 was transmitted at 500Mbps at the 3rd second mark, and flow 2 was then transmitted at 300Mbps at the 7th second mark. Next, flow 1 was transmitted at 300Mbps at the 13th second mark, and finally, flow 4 was transmitted at 200Mbps at the 16th second mark. Except for flow 3, which was transmitted for 30 seconds, the other flows were transmitted for 15 seconds.

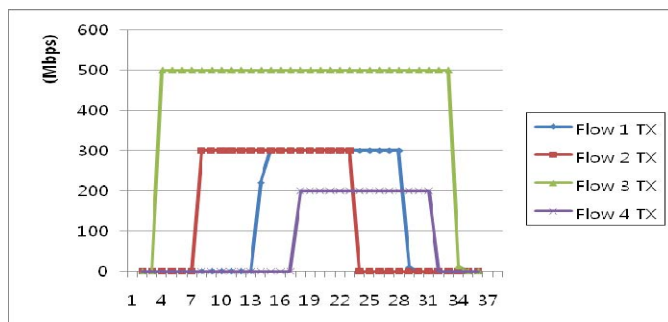
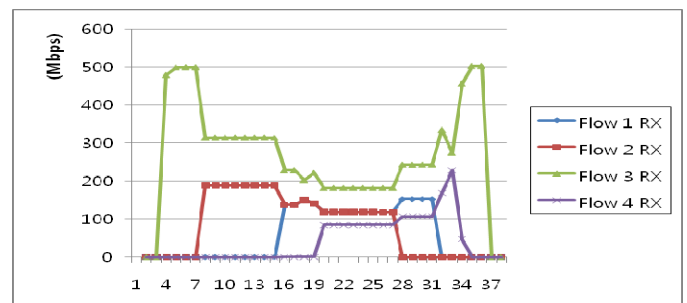


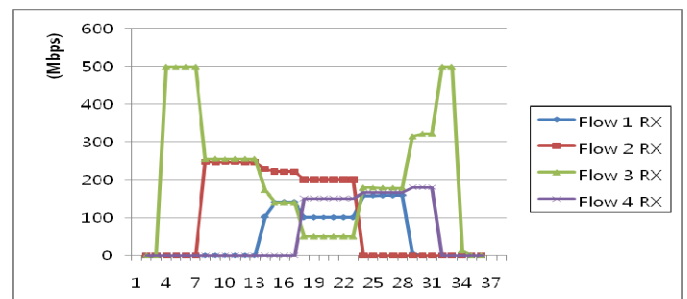
Figure 6. Transmitted duration of each flow

Figure 7 shows the experimental results of bandwidth control (a) without per-flow traffic management and (b) with per-flow traffic management. For both cases in figure 7, flow 3 is received at 500Mbps, which is the same speed it was sent at during the simulation time interval between the 3rd and 7th second marks. Because there was no other traffic, flow 3 could receive the maximum amount of traffic within the EIR. However, as shown in figure 7 (a), after flow 2 was added into the PTL, the reception rate of flow 3 was rapidly reduced to about 300Mbps due to congestion in the PTL. Since the PTL was rate-limited to 520Mbps, exceeding traffic amounts were dropped in each flow. Moreover, the drop rate was maximized during the time interval in which all of the flows were transmitted simultaneously. In figure 7 (a), flow 1 was set to CIR 100Mbps, and was thus supposed to be guaranteed 100Mbps traffic; however, the received traffic was only about 120Mbps during the time interval between the 19th and 28th seconds. Also, flow 2 was set to guarantee 200Mbps, but the received traffic was only at about 120Mbps within the same time interval despite the CIR of flow 2 being higher than that of flow 1. In this case, there was no flow management, and the

probability of a rate drop was only decided by incoming packets. As a result, traffic in the PTL has no bandwidth guarantee per flow. On the other hand, for the proposed integrated flow-based traffic management system, a drop in traffic only occurs on a yellow packet that exceeds the CIR. In figure 7 (b), after flow 2 was added into the PTL, the received rate of flow 3 was different than in figure 7 (a). Because flow 2 was configured to guarantee a bandwidth of 200Mbps, traffic that satisfies the QoS profile is metered “green,” while excess traffic is metered “yellow.” When congestion occurs in the PTL, a green packet is guaranteed, but a yellow packet is dropped rather than being sent to the queue. This bandwidth control with per-flow management is more efficient for a time duration in which all of the flows are transmitted simultaneously. During the time interval between the 17th and 23rd second marks, the total sum of the transmitted rates for the flows was 1.3Gbps. Despite this, each flow guarantees its own QoS requirement. During this time interval, all of the flows were perfectly guaranteed within their own CIR without any degradation in performance.



(a)



(b)

Figure 7. Bandwidth control (a) without and (b) with per-flow traffic management

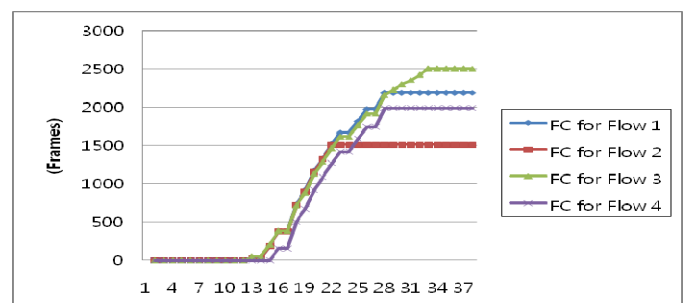


Figure 8. Accumulated flow control packets under congestion

When congestion occurs in a PTL without per-flow traffic management, the framer sends a flow control frame to the client sending the traffic. Figure 8 shows the accumulated flow control packets when congestion occurs in a PTL.

V. CONCLUSIONS

In this paper, we discussed how to effectively control traffic when congestion occurs in a PTN. Flow is commonly used in an IP network to provide fine grade QoS management. However, supporting an IP-based service in an Ethernet-based transport network has many limitations in terms of the QoS and flow management. To solve these problems, we proposed an integrated flow management scheme in a PTN to support IP-oriented services with guaranteed QoS and experimentally evaluated the scheme in a real test environment. QoS-guaranteed traffic was perfectly received without performance degradation when per-flow traffic management was applied within the PTS.

REFERENCES

- [1] N. Yamagaki, H. Tode, and K. Murakami, "DMFQ: Hardware Design of Flow-Based Queue Management Scheme for Improving the Fairness," *IEICE Trans.Comm.*, vol.E88-B, no. 4, Apr. 2005, pp. 1413-1423.
- [2] D. Yamamoto, H. Tode, T. Masaki, and K. MuraKami, "Design and Empirical Evaluation of Control Scheme for End-to-End Delay Stabilization and Packet Loss Improvement in Broadband IP Network," *IEEE ICCCN*, TP9, Hawaii, USA, Aug. 2007.
- [3] Z. Cao and Z. Wang, "Flow Identification for Supporting Per-Flow Queuing," *Computer Comm. and Networks*, Oct. 2000, pp. 88-89.
- [4] R. Braden et al, *Integrated Services in the Internet Architecture: An Overview*, IETF RFC 1633, June 1994.
- [5] S. Blake et al, *An Architecture for Differentiated Services*, IETF RFC 2475, Dec. 1998
- [6] J. Heinanen and R. Guerin, *A single Rate Three Color Marker*, IETF RFC 2697, Sep. 1999
- [7] J. Heinanen and R. Guerin, *A Two Rate Three Color Marker*, IETF RFC 2697, Sep. 1999
- [8] Metro Ethernet Forum, MEF 5, *Traffic Management Specification: Phase 1*, May 2004.
- [9] IEEE Std 802.1Qay – 2009, *IEEE Standard for Local and Metropolitan Area Networks--Virtual Bridged Local Area Networks--Amendment: Provider Backbone Bridge Traffic Engineering*, June 2009