Controlled Bandwidth Borrowing with Extended RSVP-TE to Maximize Bandwidth Utilization

Chul-Kim, Young-Tak Kim ***Regular Members***

ABSTRACT

Multiprotocol Label Switching (MPLS) has been developed as a key technology to enhance the reliability, manageability and overall quality of service of core IP networks with connection-oriented tunnel LSP and traffic engineering such as constraint-based routing, explicit routing, and restoration. In this paper, we propose a controll bandwidth borrowing scheme that maximizes the utilization of tunnel LSPs or physical links by an extension to the RSVP-TE label distribution protocol.

MPLS-based core switching network and VPN services rely on the establishment of connection-oriented tunneled LSPs that are configured or predefined by network management systems. The mechanism of network management system varies from (i) a relatively static LSP establishment accounting, to (ii) a dynamic QoS routing mechanisms. With the use of hierarchical LSPs, the extra bandwidth that is unused by the trunk (outer) LSPs should be fully allocated to their constituent end-to-end user traffic (inner) LSPs in order to maximize their utilization.

In order to find out the unused extra bandwidth in tunnel LSP or physical link and redistribute these resources to constituent LSPs, we expend the functionality of RSVP-TE and the found unused extra bandwidth is redistributed with a weight-based recursive redistribution scheme. By the extended RSVP-TE and proposed recursive redistributed scheme, we could achieve the instantaneous maximized utilization of tunnel LSP or physical link suffering from the potential under-utilization problem and guarantee the end-to-end QoS requirements. With the proposed scheme, network manager can manage more effectively the extra available bandwidth of hierarchical LSPs and maximize the instantaneous utilization of the tunneled LSP resources.

Keyword: MPLS, RSVP-TE, Bandwidth Borrowing, Traffic Engineering.

I. Introduction

Multiprotocl Label Switching (MPLS) is emerging as a key technology to enhance the reliability, manageability and overall quality of service of core IP network[1]. The use of connection-oriented, constraint-based routing, and bandwidth reserved LSPs in the core routers enables more flexible and predictable traffic engineering[2].

Tunnel LSPs are established among the core routers to support specific traffic engineering and other goals, such as VPN service provisioning. The typical application of Tunnel LSP can be found in [3] where tunnel LSPs are configured to deliver several LSPs (VC LSPs) in which various layer 2 protocols' PDU are transported. This concept is adapted to L2VPN[4]. Typically, such tunnel LSPs are configured by the management system and are prescribed by Service Level Agreements (SLAs) between the network and its peers or customers.

In fixed-bandwidth circuit switched network, the under-utilization of circuit has been the major problem. Since MPLS LSP can be regarded as the connection-oriented virtual circuit, it also has the same potential problem, especially when the established LSP is kept with a strict bandwidth

---

영남대학교 대학원 정보통신공학과 차세대네트워킹 연구실(goldfe@nownuri.net, ytkim@yu.ac.kr)
논문번호: 030220-0526, 접수일자: 2003년 5월 27일

reservation regardless of the actual LSP utilization. To solve this problem several schemes have been proposed such as bandwidth overbooking, bandwidth borrowing, and bandwidth sharing, and so on. Bandwidth overbooking[5] in MPLS network can be performed in off-line or on-line. This scheme requires the information of the traffic demand prediction, the real-time link state information and the link state routing. Since the bandwidth overbooking always has the risk of coincident peak bandwidth requests from multiple LSPs, the possible malicious user-behavior, and the inaccurate prediction of traffic demands, it is not a good approach of traffic engineering for the connection-oriented tunnel LSP. It may also suffer from the potential under-utilization.

Bandwidth sharing [6][7](e.g., CBQ) has considered about the distribution of resource to constituent flows. This scheme can be used in DiffServ-aware MPLS network in order to provide a high utilization of E-LSP with flexible bandwidth sharing among the constituent micro flows [8].

The motivation of this paper is to address the potential under-utilization problem of hierarchical LSPs by periodic discovery of unused bandwidth of physical link or tunnel LSPs and redistribution of the discovered unused bandwidth among their constituent flows. In case of hierarchical LSPs, this implies recursive bandwidth redistribution across multiple levels of encapsulated LSPs. In order to increase the utilization of the bandwidth of hierarchical LSPs and physical link, we propose an extension to RSVP-TE [3] that enables a controlled redistribution of unused available bandwidth and a weight-based recursive bandwidth redistribution scheme. The proposed scheme is carried out only for unused extra available bandwidth during temporarily period (e.g., every 1sec). We present simulation results of the proposed scheme to demonstrate its usage to achieve the maximum utilization with controlled manner.

This paper is organized as follows: Section II describes the proposed Dynamic Distribution of extra bandwidth scheme and weight-based recursive bandwidth redistribution scheme. Section III presents the simulation results that demonstrate the enhancement of QoS, tunnel utilization and network utilization that can be achieved using the proposed mechanism. Then we conclude in section IV.

II. Dynamic Distribution of Extra Bandwidth

2.1 Related Works

In order to support MPLS tunnel establishment RSVP is extended for providing label binding and label distribution functionality. RSVP-TE signaling protocol uses downstream-on-demand label distribution. A request to bind labels with a specific LSP tunnel is initiated by an ingress LSR through the RSVP Path message. When an egress LSR receives the Path message, labels are allocated downstream and distributed by means of the RSVP Resv message. For these purpose LABEL_REQUEST and LABEL object are extended [9].

The one of most attractive features of MPLS is a constraint-based routing. In order to support constraint-based routing, OSPF is extended to provide additional link information. The OSPF-TE defines several sub-TLVs such as maximum bandwidth, maximum reservable bandwidth, and unreserved bandwidth [10]. The maximum bandwidth specifies the maximum bandwidth that can be used on the link. The maximum reservable bandwidth specifies the maximum bandwidth that may be reserved on the link. This may be greater than the maximum bandwidth in which case the link may be oversubscribed. The Unreserved bandwidth specifies the amount of bandwidth not yet reserved at each priority levels [10]. A network management policy that is combined with additional information provided by OSPF-TE can establish LSPs that have constraint-based routes.

2.2 Dynamic Distribution of Extra Bandwidth

In order to detect the unused bandwidth along the tunnel LSP and physical link and redistribute the found unused bandwidth, we extend RSVP-TE signaling protocol. The extended functions are Resource Discovery (RD), Resource Redistribution (RR), and Rollback (RB). Each function has own message object. These objects can be encompassed in Path/Resv message or can be generated their own message. Because of scalability issue of RSVP-TE, we define new messages for each function. The proposed scheme may be one of the management functions to maximize the utilization of tunnel LSP or physical link. The proposed scheme can be initiated by network operators or activated automatically by the policy-based management system such as performance monitoring function with MPLS OAM [13].

2.2.1 Resource Discovery (RD) Phase

In RD phase, the ingress LSR of the tunnel LSP generates the RD message. The RD message is defined as follows:

RD Message::=
<RSVP-TE Common Header> <Session Object>
<RSVP HOP> <Resource Discovery Descriptor>
Resource Discovery Descriptor::=
<SENDER_TEMPLATE> <Discover_TSPEC>

Resource Discovery Descriptor consists of SENDER_TEMPLATE and Discover_TSPEC. The Discover_TSPEC object stores the discovered minimum extra bandwidth of a tunnel LSP at each LSR along the path.

When a core LSR receives a RD message, it checks the unused extra bandwidth of the given tunnel LSP or physical link. If the locally available LSP bandwidth is less than the received Discover_TSPEC, it updates the Discover_TSPEC parameter with the locally available bandwidth and forwards the RD message to next LSR. This procedure is repeated until the egress LSR.

Through the RD procedure, the minimum unused extra bandwidth along the tunnel LSP or physical link can be found.

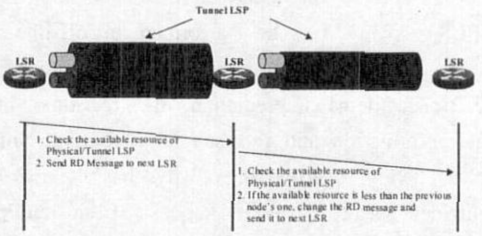


Fig. 1. Resource Discovery Phase Diagram

Fig. 1 depicts the procedure of RD phase. In Fig. 1, since two tunnel LSPs have a different extra-bandwidth, at the end of RD phase, we would find out the minimum extra available bandwidth depends on the unused bandwidth of the second tunnel.

Several bottleneck links might be prevailed in high-speed network. These bottleneck links must be considered in RD phase. In RD phase, routers that receive the RD message check the extra bandwidth of the considered physical link or tunnel LSP. If the RD message were arrived at the router that has a bottleneck link, the found extra bandwidth would be zero. This value will be updated into the Discover_TSPEC object in the RD message. If there are several bottleneck links, the Discover_TSPEC object in the RD message keeps the minimum extra bandwidth of the given physical link or tunnel LSP. Then, this information is delivered to egress LSR and triggers RR phase. In RR phase, the RR message that is transmitted toward ingress LSR keeps the minimum extra bandwidth found in the RD phase. According to the found minimum extra bandwidth value, each router redistributes bandwidth. Because the minimum extra bandwidth of the given physical link or tunnel LSP is detected through the RD phase locally, the minimum bandwidth of link or path that has several bottlenecks can be easily found.

2.2.2 Resource Redistribution (RR) Phase

When an egress LSR receives a RD message, the RD message triggers Resource Redistribution phase. In RR phase, the egress LSR generates RR message based on the received RR message and forwards it toward the ingress LSR. The RR message is defined as follows:

```
RR Message::=
<RSVP-TE Common Header> <Session Object>
<RSVP-HOP><STYLE>
<Resource Redistribution Descriptor List>
Resource Redistribution Descriptor::=
<RR Descriptor list>
RR Descriptor::=
<FLOWSPEC> <FILTER SPEC>
```

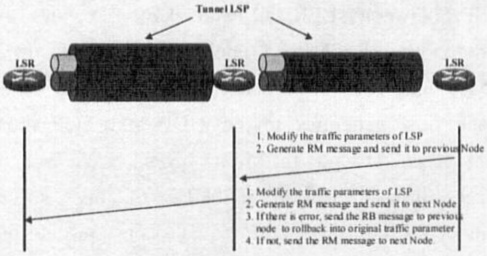


Fig. 2. Resource Redistribution Phase Diagram

RR Descriptor contains flow specification object that represents the minimum available bandwidth and filter specification object that indicates the source LSRs. When a core LSR receives the RR message, it retrieves the minimum available bandwidth from FLOWSPEC object. Then, it redistributes a real resource to the constituent LSPs according to the proposed weight-based recursive bandwidth redistribution scheme. Fig. 2 shows the RR phase.

2.2.3. Rollback (RB) Phase

When an error occurs either in RD phase or in RR phase, a Rollback message is created and delivered to the ingress LSR and the egress LSR to release the recently allocated extra bandwidth and to abort the RD. These errors can be a resource allocation error, new LSP establishment during RD or RR phase, and so on. In order to recover the previous state of the LSP, the previous traffic parameters, such as Peak Data Rate (PDR), Peak Burst Size (PBS), Committed Data Rate (CDR), Committed Burst Size (CBS) are retained. When a LSR receives the RB message, it restores the previous traffic parameters. The RB message is defined as follows:

```
Rollback Message ::=
<RSVP-TE Common Header><Session>
<RSVP-HOP> <Rollback Descriptor list>
Rollback Descriptor ::=
<SENDER_TEMPLATE>
```

In case a new LSP establishment is requested, the redistributed bandwidth might be released along the path of new LSP. If the available bandwidth after redistribution is enough to accommodate new LSP's request, the redistributed bandwidth dose not need to be released. However, when there is not enough available bandwidth for the newly requested LSP, the redistributed bandwidth should be released by RB phase before the establishment of new LSP.

2.3 Bandwidth Distribution Scheme

In order to efficiently distribute the unused bandwidth into constituent LSPs of tunnel LSPs and physical links, we propose a weight-based recursive distribution scheme. We assume that each inner LSP of a given tunnel LSP is assigned a weight value in accordance with the network management policy or Service Level Agreement (SLA).

Let us denote the extra bandwidth of a tunnel LSP or physical link as $extraAvailableBW$ and the number of tunnel LSPs as k. Then, the redistributed bandwidth of each Tunnel LSP denoted as $availableBW\_LSP_i$ is determined as follows:

$$availableBW\_LSP_i = \frac{w_i}{\sum_k w_k} \times extraAvailableBW$$

Each tunnel LSP is given the amount of $availableBW\_LSP_i$. If the tunnel LSP has m inner LSPs, $availableBW\_LSP_i$ is redistributed recursively to each LSP as follows:

$$availableBW\_LSP_{ij} = \frac{w_{ij}}{\sum_m w_{im}} \times availableBW\_LSP_i$$

w_i is the weight value of tunnel LSP j and the weight value of the i^{th} internal LSP is w_{ij}. The network management system has to assign the weight value of each tunnel LSP. Fig 3 depicts the assumed environment of hierarchical LSPs with assigned weights and outline the recursive bandwidth redistribution scheme.

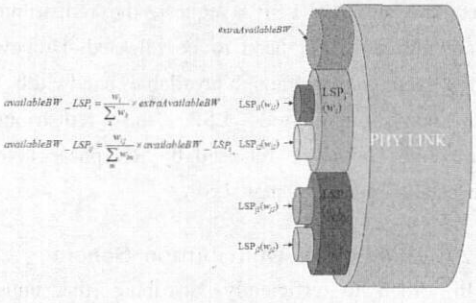


Fig. 3. Recursive Bandwidth Redistribution

Ⅲ. Simulation

3.1 Test Network Configuration

We demonstrated the usefulness of the proposed scheme of dynamic redistribution of extra available bandwidth using NIST GLASS (GMPLS Lightwave Agile Switching Simulator) network simulator [14]. Fig. 4 shows the test network configuration with 9 LSRs, 4 label edge routers (LERs), and 14 hosts. Each link among LSRs has 20Mbps capacity and 0.5msec propagation delay, while the links between hosts and LERs have 5Mbps bandwidth and 1msec propagation delay. To configure a bottleneck link, we reduce the bandwidth of link between LSR X and LSR Y to 10Mbps. In order to deliver user traffic we establish three LSPs that have 4Mbps, and one LSP between LER C and LER D that has bandwidth of 2Mbps. Each LSP is set up at 25sec and each host starts transmission at 50sec. Each host generates mixed UDP and TCP traffic that pass through the four LSPs. We set the target link utilization as 80% of link so we cannot establish more LSP or re-distribute bandwidth more than that. Every LSR performs a dual-token bucket traffic policing, and core LSRs use the active queue management such as RED.

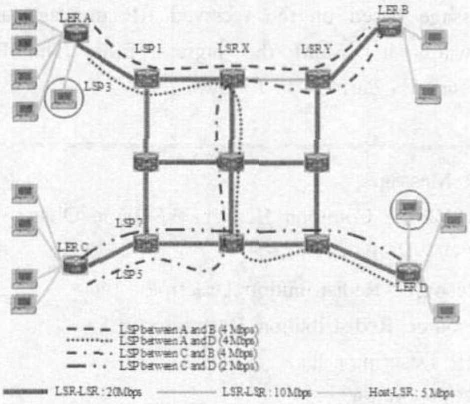


Fig. 4.Test Network Configuration

3.2 Extra Available Bandwidth Redistribution

We demonstrate the proposed scheme's behavior under the given test network. Since there is a bottleneck link between LSR X and LSR Y, the LSP 1 and LSP 5 cannot take the redistributed bandwidth. However, LSP 3 and LSP 7 can get the redistributed bandwidth and achieve high utilization. Fig. 5 depicts the throughput of LSP 1 and LSP 3. Since LSP 1 passes through the bottleneck link between LSR X and LSR Y, there is no extra bandwidth to be redistributed. However, LSP 3 takes extra bandwidth through the proposed scheme. We assign the weight value *each LSP to get the extra bandwidth according to* their weight value.

Fig. 6 depicts the throughputs of LSP 5 and LSP 7. Since LSP 5 goes through the bottleneck link, it also cannot take the extra-bandwidth. However, LSP 7 gets extra-bandwidth through the proposed scheme. Since we assign a different weight value to LSP 7, it gets less extra-bandwidth.

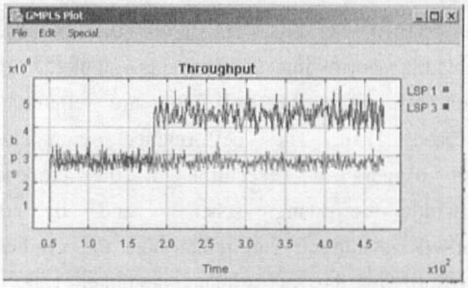


Fig. 5. Throughput LSP 1 and LSP 3

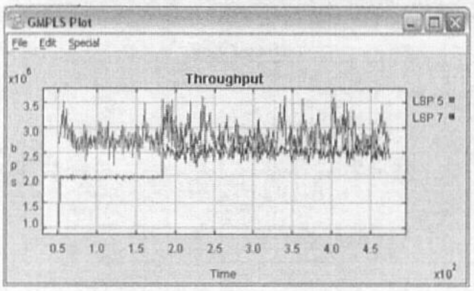


Fig. 6. Throughput LSP 5 and LSP 7

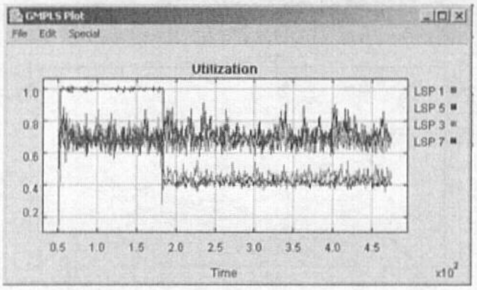


Fig. 7. Utilization of Four LSPs

Through this simulation, we can see that the proposed controlled bandwidth redistribution scheme dynamically distribute extra available bandwidth to the LSPs that require more bandwidth temporarily and maximize the LSP and link utilization. Fig. 7 depicts the utilization of four LSPs. While LSP 1 and LSP 5 do not get extra bandwidth, LSP 3 and LSP 7 get extra bandwidth and utilize their bandwidth

3.3 Rollback and establish new LSP

In this simulation, we verify the Rollback procedure that is invoked in two cases: an RR error and a Rollback to previous state according to network policy such as a new LSP establishment.

Let's assume that after the proposed scheme is deployed, one user demands a new LSP establishment. In this case, if there is enough bandwidth to accommodate the new LSP along the path, we don't rollback the LSPs that have received extra bandwidth. Otherwise, we should rollback and start LSP setup procedure.

In order to simulate this situation, we add a new client and a server to LER A and LER D, respectively (circled nodes in Fig. 4). After the proposed scheme is deployed, a new LSP (LSP 9) setup request is invoked along the same path with LSP 3. New LSP is established at 300sec and released at 400sec. Because LSP 3 and LSP 7 have extra bandwidth through the proposed scheme, we can verify the rollback procedure.

A new LSP setup request is delivered to ingress LSR which checks whether there is enough bandwidth to accommodate it. Because of the proposed scheme, LSP 3 and LSP 7 are allowed to use the extra-bandwidth, and LSP 9 cannot be established by current situation. Therefore, the Rollback procedure must be invoked and the given extra bandwidth should be released. Then the new LSP can be established.

Fig. 8 and Fig. 9 depicts the throughput of LSP 1, LSP 3, LSP 5, LSP 7, and LSP 9. The allocated extra bandwidth of LSP 3 and LSP 7 is released and new LSP is established. However, this rollback procedure does not affect other LSPs that have different route compared with the new LSP. We can verify this fact by no bandwidth change in LSP 1 and LSP 5. Through the rollback procedure, we can dynamically release the allocated extra bandwidth according to the current network status and policy.

3.4 Simulation Experiments Under *MPLS-based Virtual Private LAN* Service (VPLS)

In [15], we had proposed the performance management functionality for the guaranteed end-to-end QoS provisioning on MPLS-based VPLS and VPLS OAM function. Based on OAM function and performance monitoring, we adjusted the bandwidth of LSP for VPLS for supporting performance tuning functionality by the proposed bandwidth borrowing mechanism. We had configured the test network and assumed the performance degradation criteria as shown in Table 1.

Table. 1. Example Criteria of Severly degraded Performance

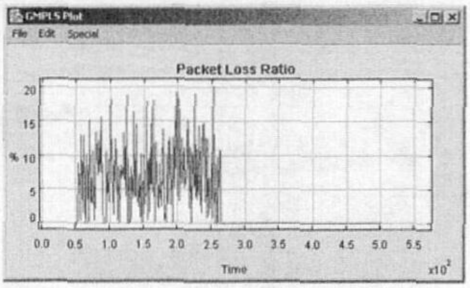
Traffic/Qos Parameter	Threshold of severe performance degradation
End-to-end Delay	More than 120% of agreed end-to-end delay limit
Jitter	More than 200% of agreed jitter limit
Packet Loss	More than 20% of transmitted data
Available Bandwidth	Less than 80% of CDR

The performance degradation is reported from the egress LSR to the ingress LSR using the VPLS OAM functions then, the ingress LSR *performs the proposed* bandwidth borrowing function to guarantee the end-to-end QoS. Table 2 shows the detection and notification of performance degradation [15].

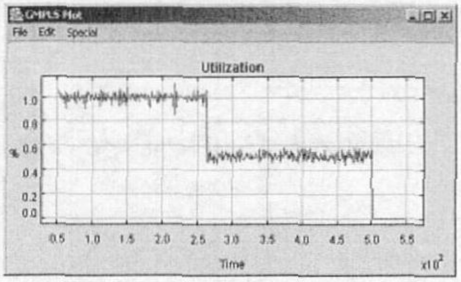
Table 2. Detection and Notification of Performance Degradation

OAM Degradation Report
LSP ID: 500
Ingress LER: 220
Egress LER: 110
Detection Time: 253.0903
Packet Loss Warning: 20.19579395
Node(220) Module(pmProc)

As shown in Fig. 10, after the proposed bandwidth borrowing scheme is applied, the packet loss ratio and utilization are remarkably decreased [15]. The performance degradation report triggers the execution of the proposed bandwidth borrowing scheme, and by the proposed scheme we can redistribute the available bandwidth to the performance-degraded LSP to guarantee the promised end-to-end QoS.



(a) Reduced Packet Loss Ratio after bandwidth borrowing



(b) Reduced Utilization after bandwidth borrowing

Fig. 10. Result of Bandwidth Borrowing

From this result, we can apply the proposed bandwidth borrowing scheme not only to address the potential under-utilization problem but also to re-adjust the LSP for performance tuning.

Ⅳ. Conclusion

In this paper, we proposed a controlled bandwidth borrowing mechanism in MPLS network and proposed extension to RSVP-TE for this purpose. Through the simulation, we could verify the applicability of the proposed scheme that eliminates the risk of bandwidth overbooking and gives a controlled bandwidth borrowing mechanism, which maximizes the bandwidth utilization according to the network status. This scheme is also applicable in MPLS-VPN network to allow the network operator to use bandwidth more efficiently in a controlled manner.

REFERENCES

[1] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture," RFC 3031, Jan. 2001.

[2] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, J. McMaus, "Requirement for Traffic Engineering Over MPLS," RFC 2702, Sept. 1999.

[3] IETF draft, "Transport of Layer 2 Frames Over MPLS," <draft-martini-12circuit-trans-mpls-10.txt>, Aug. 2002.

[4] IETF draft, "L2VPN Framwork," <draft-ietf-ppvpn-12-framework-02.txt>, Jan. 2003.

[5] IETF draft, "The notion of overbooking and Its Application to IP/MPLS Traffic Engineering," <draft-cchen-te-overbooking-01.txt>, Informational.

[6] Sally Floyd and Van Jacobson, "Link-sharing and Resource Management Models for Packet Network," IEEE/ACM Trans. on Networking, Vol. 3, No. 4, Aug. 1995.

[7] L. Massoulie and J. Roberts, "Bandwidth sharing: Objectives and Algorithm," in Proc. IEEE INFOCOM 1999, March 1999.

[8] RFC 3270, "Multiprotocol Label Switching (MPLS) Support of Differentiated Service," May 2002.

[9] D. Awduce, L. Berger, T. Li, V. Srinivasan, G. Swallow, "RSVT-TE: Extension to RSVP for LSP Tunnels," Dec. 2001

[10] IETF draft, "Traffic Engineering Extension to OSPF Version 2," <draft-katz-yeung-ospf-traffic-10.txt>, Dec. 2003.

[11] IETF draft, "Virtual Private LAN Service," <draft-kompella-ppvpn-vpls-01.txt>,May 2003.

[12] IETF draft, "Architecture and Model for VPLS," <draft-augustyn-vpls-arch-00.txt>, Nov. 2001.

[13] IETF draft, "A Framework for MPLS User Plane OAM,"<draft-allan-mpls-oam-frmwk-03.txt>, April 2003.

[14] NIST GMPLS Lightpath Agile Switching Simulator, "http://dns.antd.nist.gov/glass/," Dec. 2002.

[15] Seong-Woo Kim, Chul Kim, Young-Tak Kim, "Performance Management and Analysis for Guaranteed End-to-End Qos Provisioning on MPLS-based Virtual Private LAN Service (VPLS)," KICS Journal, Vol. 28, No. 2B. pp.144~156.

Chul Kim Regular Member

He received the B.E. degree in Electronics Engineering from Yeungam University in 1998, and M.E. degree in Information & Communication Engineering in 2000. He is a Ph.D. candidate of Information & Communication Engineering in Yeungam University. He has been a guest researcher of National Institute of Standard and Technology(NIST) in Maryland in U.S.A from Feb. 2002 to Oct. 2003. His current research focuses on NGI, MPLS, DiffServ, and Traffic Engineering.

Young-Tak Kim Regular Member

He received the B.E. degree in Electronics Engineering from Yeungam University in 1984, and the M.E degree and the Ph.D degree in Electrical Engineering from KAIST in 1986 and 1990, respectively. He has been a member of technical staff, section chief of Transport network architecture Lab., in Korean Telecom from 1990 to 1994. He has been a guest researcher of National Institute of Standard and Technology (NIST) in Maryland in U.S.A during Feb. 2001 ~ Jan. 2002. Currently, he is a full professor of School of Electrical Engineering & Computer Science (EECS) in Yeungam University. His current research interest includes Broadband networking, traffic engineering and network management of next generation Internet