Global Interoperability Trial for Next-Generation Photonic IP Network
– GMPLS/OUNI Protocols and Applications –

Daisaku Shimazaki (0), (1), Eiji Oki (0), (1), Itaru Nishioka (0), (2), Shinya Kano (0), (3), Akira Nagata (0), (3),
Mikako Nanba (0), (4), Kazumasa Morita (0), (4), Motohumi Tanabe (0), (5),
Kazuhiro Kusama (0), (6), Pinai Linwong (0), (6), Tohru Hirai (0), (7), Kohei Okazaki (0), (8)
(0) Photonic Internet Lab.
(1) NTT Network Service Systems Laboratories, NTT Corporation
3-9-11 Midori-cho, Musashino-shi, Tokyo, 180-8585 Japan
Email: shimazaki.daisaku@lab.ntt.co.jp
(2) System Platforms Research Laboratories, NEC Corporation
(3) Photonics Networking Laboratories, Fujitsu Laboratories Ltd.
(4) Network Products Division, Furukawa Electric Co., Ltd.
(5) Information Technology R&D Center, Mitsubishi Electric Corporation
(6) Systems Development Laboratory, Hitachi Ltd.
(7) Carrier Network Systems Division, Hitachi Communication Technologies, Ltd.
(8) Faculty of Science and Technology, Keio University

***Abstract*— We confirm three GMPLS technologies, GMPLS/MPLS inter-working, OUNI, and the multi-CoS recovery mechanism, using the MPLS/GMPLS multi-region network connected between PIL site in Japan and ISOCORE site in USA. The combination provides wideband video data transmission service and auto route selection according to content data volume.**

***Index Terms*— MPLS, GMPLS, Interoperability, End-to-End Recovery, End-to-End Path Control, OUNI, Inter-working**

I. Introduction

A. Internet and GMPLS

The number of Internet users in Japan exceeds 77 million and number of the broadband contracts has reached 18 million [1]. The broadband service fee in Japan is now the lowest in the world and this has driven its growth [2]. We can predict that the number of broadband access lines will continue to grow. A research agency is saying that FTTH service is growing rapidly and will overtake ADSL service in terms of the number of contracts in 2006. The Internet traffic volume is rapidly increasing with the number of broadband access lines. The average traffic volume was about 342 Gbps in November 2004 [3]. The backbone network must offer much higher throughputs because the rate of growth in traffic is about 200 - 300 % per year. The current traffic is about 50% P2P data since the P2P application readily consumes a lot of bandwidth. There are also other type services such as video streaming and IP telephony that also demand real-time data transfer. Broadband real-time applications such as Video-phone will soon be widely used.

The next-generation backbone networks should consist of IP routers with IP packet switching capability and optical cross-connects (OXCs); wavelength path switching will be used to reduce IP packet switching loads. These networks will be controlled by Generalized Multi-Protocol Label Switching (GMPLS) [4]. GMPLS is being developed in the Internet Engineering Task Force (IETF). It is an extended network control protocol of MPLS. While MPLS was originally developed to control packet-based networks, GMPLS controls several packet and non-packet based networks, such as IP packet, Layer 2 frames (such as ATM, Frame Relay, and Ethernet), time-division multiplexing (TDM), wavelength, and optical fiber layers. The GMPLS suite of protocols is expected to support new capabilities and functionalities for an automatically switched optical network (ASON) as defined by the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) [5]. ASON provides the dynamic setup of optical connections, and fast and efficient restoration mechanisms and solutions for automatic topology discovery and network inventory. GMPLS will reduce network operation cost and service offering delay.

B. GMPLS Interoperability Test in Photonic Internet Lab.

The Photonic Internet Lab. (PIL) was founded Autumn 2002 by 7 companies (6 venders and 1 service provider) for realizing new photonic network control protocols based on photonic technologies for managed networks [6], [7].

PIL has conducted GMPLS inter-operability tests on routing and signaling, and has made several demonstrations, PIL workshop 2003 [8], Gigabit Network Symposium 2004 [9], Supercomm 2004, and MPLS 2004. These interoperability tests and demonstrations have confirmed the interoperability of the GMPLS protocol suite (Open Shortest Path First (OSPF), ReSerVation Protocol (RSVP) and etc.)

More recently, PIL called for companies and universities all over the world to conduct an inter-operability test on GMPLS and Optical User Network Interface (OUNI) [17], and 14 companies and 1 university accepted the challenge. This test has done under the joint auspices of PIL and ISOCORE [10], an interoperability site in the USA. A part of this test was publicly demonstrated at iPOP2005 [11].

This most recent inter-operability test was designed considered the following items.

- Cooperation with existing networks and applications
- Effect of GMPLS-based recovery on existing networks
- Efficient use of GMPLS network resources

The inter-operability test directly examined the following items.

- GMPLS/MPLS inter-working
- Handling GMPLS LSP establish request from user network
- GMPLS-based recovery
- Application-triggered GMPLS signaling

We focus on the iPOP 2005 demonstration held in February, 2005.

II. GMPLS NETWORK COORDINATED OPERATION

A. *Multi-layer Network*

GMPLS is a hierarchical network and lower layer Label Switched Paths (LSPs) are treated as logical links called Forwarding Adjacency (FA) in the higher layers. FA is advertised by the routing protocol, OSPF. In the higher layer, the logical topology consists of FA which is used for route calculation and routing.

B. *MPLS/GMPLS Inter-working*

1) MPLS/GMPLS Signaling: When GMPLS is introduced into the backbone network of a carrier, existing MPLS networks are connected to the GMPLS network. To establish an end-to-end path from one MPLS network to another MPLS network via the GMPLS network, GMPLS must co-operate with the MPLS networks [12], [13]. There are two signaling schemes that permit GMPLS and MPLS cooperation.

- Pre-provisioned signaling:
  In this scheme, the GMPLS LSP is established in advance. The MPLS LSP is established via the existing GMPLS LSP. If the GMPLS LSP is advertised to other networks as an FA link, an MPLS node can recognize the routes available in the GMPLS network. An MPLS node can recognize the GMPLS LSP as a link and establish an MPLS LSP via GMPLS using strict signaling. GMPLS edge nodes simply forward the MPLS signaling message to GMPLS LSP by referring to the label table. When GMPLS LSP is not advertised to the MPLS networks, an MPLS node can use loose signaling to establish an LSP via the GMPLS network if IP reachability is advertised to the MPLS networks. Figure 1 shows the pre-provisioned signaling sequence.
  This scheme has the disadvantage that GMPLS LSPs must be established in advance by manual setting.
- Triggered signaling:
  In this scheme, the GMPLS edge node selects the route and establishes a GMPLS LSP upon receipt of MPLS LSP signaling. An MPLS node can establish an LSP via the GMPLS network by using loose signaling or specifying the destination node because each MPLS node receives reachability information to other MPLS networks. GMPLS edge nodes select the route to forward the MPLS LSP upon receipt of MPLS signaling. If there is no route, GMPLS nodes establish a new GMPLS LSP and associate the MPLS LSP with the route. Figure 2 shows the triggered signaling sequence.
  This scheme demands a new function that ensures that GMPLS signaling cooperates with MPLS LSP establishment. The GMPLS network does not need LSPs to be pre-provisioned and can automatically establish LSPs as required.

We validated and confirmed these two schemes in the iPOP 2005 demonstration.

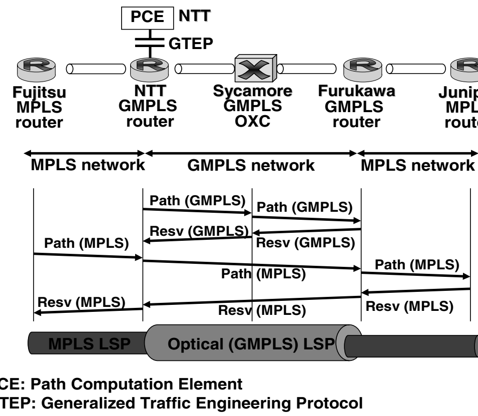

Fig. 1. Pre-provisioned signaling sequence

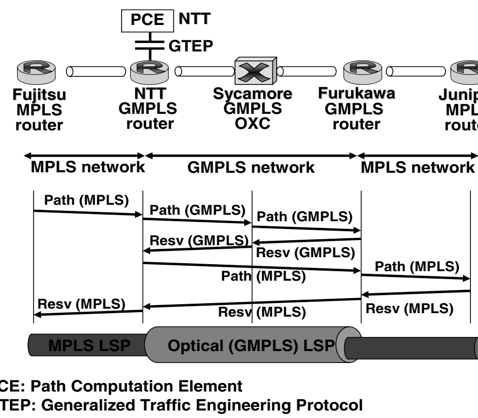

Fig. 2. Triggered signaling sequence

2) Path Computation Client (PCC) – Path Computation Element (PCE) Communication: GMPLS will be introduced to a carrier's network. GMPLS backbone network topology may not be advertised to the MPLS user networks. In this case, MPLS nodes can not determine a route across the GMPLS network. Moreover, different carriers will have different route selection policies. The solution to these difficulties is to separate the PCC function from PCE and create a dedicated protocol that suits the carrier's policy, for LSP establishment to realize PCC-to-PCE communication. Some Internet-Drafts [14], [15] about this theme are discussed in IETF PCE working group. PIL has proposed the Generalized Traffic Engineering Protocol (GTEP) [16]. GTEP has two functions.

- Link State Database (LSDB) synchronization function
  Each GMPLS node advertises network topology by the GMPLS routing protocol and synchronizes the LSDB. Multilayer routes can be determined using the database. If PCE is to calculate the route, PCC must notify it of the appropriate LSDB via the PCC-PCE interface.
- Request / Response to Route calculation When PCE receives a route request from PCC, PCE performs a multilayer route calculation. If the lower layer LSP required by a higher layer LSP is not established, PCC requests its setup. The PCC responds to the LSP setup request after

the lower layer LSP has been established. After the lower layer LSP is established, PCE sends a route response to PCC and the higher layer LSP is established after the lower layer LSP. Figure 3 shows the PCC-PCE interface and sequence.

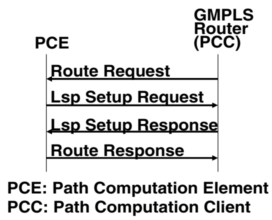

Fig. 3. Generalized Traffic Engineering Protocol

3) Implementation of Control Plane: The MPLS network has an in-band control plane. The GMPLS network, on the other hand, has an out-band control plane that is separated from the data plane. We should consider the difference in C-plane implementation between GMPLS and MPLS. There are two approaches to routing signaling messages.

- Tunneling:
  In this scheme, the MPLS signaling message is forwarded to the other MPLS network via a GMPLS LSP. In this case, the GMPLS LSP is advertised to the MPLS networks by the routing protocol.
- Contiguous:
  In this approach, the MPLS signaling message is forwarded to the other MPLS network via the GMPLS C-plane. The GMPLS edge node must translate the MPLS signaling message into a GMPLS signaling message.

We employed the tunneling approach in the iPOP 2005 demonstration.

C. Co-operation with User Network (OIF-UNI)

There are two requirements when user networks are connected to a carrier's network.

- Carrier does not want the network information to be advertised to the user networks.
- The carrier and user have separate addressing spaces.

Setting a UNI interface between the carrier and user networks satisfies both requirements. The user can establish connections between Customer Edges (CE) without concern for the complexity of the GMPLS multi-layer network. The carrier is not affected by the user's addressing space and does not need to consider addressing conflicts between users. This network architecture is called the overlay mode. In this demonstration, some nodes implemented OIF UNI (OUNI) [17].

OUNI does not define the carrier's (GMPLS) network behavior when it receives OUNI signaling. The Provider Edge (PE) establishes the connection across the Interior Network-to-Network Interface (I-NNI). Figure 4 shows the OUNI signaling sequence.

D. GMPLS Co-operates with User Applications

A broadband user application may directly request the GMPLS network to setup a GMPLS LSP. When an application needs real-time data transfer, a GMPLS LSP that satisfies the requested bandwidth should be established for the holding time specified by the application. When the application transfers a large content and does not need real-time data transfer, a

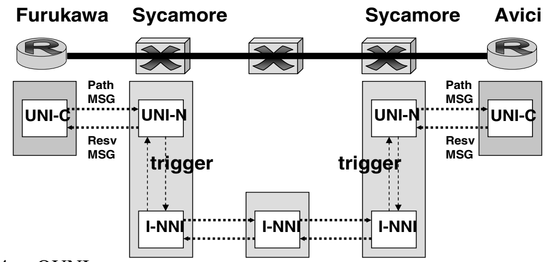

Fig. 4. OUNI

wideband GMPLS LSP is established for a short period as triggered by the user application and the content is transferred before the LSP is torn down.

Keio University is developing a system that selects the route according to the required content [18]. This system consists of a proxy server and content server. The content server also has route selection function. It selects appropriate route depending on the type of required content. I describe sequence of the system.

- Client requests a content to content server via proxy.
- Proxy server establishes a narrow-band GMPLS LSP to the content server and uses it to send the content request message to the content server.
- The content server establishes the appropriate GMPLS LSP for transferring the content and sends it to the client via the proxy server. The type of GMPLS LSP and its route depend on the type and size of content. For example, a wideband LSC LSP will be established when the content is a large movie file.
- GMPLS LSP is deleted when content transfer is completed.

III. GMPLS-based Recovery [19]

A. End-to-End Recovery

GMPLS end-to-end recovery supports multiple classes of service (CoS) [20]. One attribute related to CoS in optical networks is the downtime permitted by customer applications. GMPLS has the potential to enable various recovery types independent of network topologies or equipment. GMPLS-based end-to-end recovery, LSP recovery in GMPLS terminology, provides sufficient functionality to support multiple CoS. Four types of end-to-end recovery are currently defined in GMPLS [6], (1) dedicated LSP protection, (2) pre-planned LSP re-routing, (3) full LSP re-routing and (4) unprotected. Table I describes the characteristics of each type. These types of recovery enable CoS to be provided either singularly or in combination. For example, preplanned LSP re-routing or full LSP re-routing provides a relatively economical service class, but does not guarantee fault tolerance. On the other hand, the combination of dedicated LSP protection and full LSP re-routing offers the most reliable and expensive service class. In the interoperability test, we have verified the service class using pre-planned LSP re-routing with shared mode, which achieves a balance between recovery time and network resource usability.

B. Consistency Issues

GMPLS-based end-to-end recovery procedures are being defined at IETF [21], but we have to agree on the further consistency among vendors, to ensure successful interoperability. PIL have developed an implementation agreement (IA). Followings are some important items discussed in the IA.

Recovery Type	Recovery Time	Fault Tolerance	Required Resources
(1)	$<$ 50 ms	High	2
(2)	$<$ seconds	High	$<$ 2 (1:1)
(2)	$<$ seconds	Fair	$<$ 1.7 (shared)
(3)	$<$ minutes	Best-effort	1
(4)	–	No tolerance	1

TABLE I
RECOVERY TYPE CHARACTERISTICS

1) Recovery Coordination between Layers: In multi-layer networks, it is essential to discuss how recovery is coordinated among the layers. GMPLS-based recovery can restore paths on any layer (packet, TDM, wavelength, fiber). If the recovery is performed in multiple layers simultaneously, contention between the upper layer and the lower layer will occur. To avoid this contention, there are two alternative ways. One is to use the holding timer mechanism [22], and the other is to use the FA protection mechanism [23]. In the former mechanism, the upper layer initiates recovery after the configured timer is expired. In the latter mechanism, recovery paths are established exclusively among the layers. In PIL IA, we focused on the FA protection mechanism illustrated in Fig. 2. We defined an FA as a redundant virtual link, which consists of a pair of working and backup paths created in the lower layer. We also defined a rule for assigning a link protection type to the FA, which enable the upper layer to establish recovery paths exclusively. Table II shows the mapping table between the recovery types and the link protection type. The FA is advertised as a TE (Traffic Engineering) link, and used as a resource for setting up upper layer paths. These upper layer paths can be recovered from a failure by executing the recovery procedure on the FA.

Recovery Type	Link Protection type
(1)Dedicated LSP protection	Dedicated 1+1
(2)Pre-planned LSP	Dedicated 1:1
re-routing	Shared
(3) Full LSP re-routing	Not Assigned
(4) Unprotected	Unprotected

TABLE II
MAPPING OF RECOVERY TYPE/LINK PROTECTION TYPE

2) Monitoring Configuration: Current GMPLS standard dose not specify when we should configure cross-connects and start fault monitoring in the path setup/deletion procedure, so this is an implementation matter left to each vendor. However, it is very important to achieve consistency in configuring cross-connects and fault monitoring because inappropriate configuration timing is likely to cause unanticipated errors. This issue is most obvious in the case of SDH (TDM) paths establishing through transparent OXC nodes. Generally, transparent OXC nodes do not launch optical signals until the wavelength path is established. Under this condition, the SDH XC nodes detect loss-of-signal (LOS) alarms. If failure monitoring is started during the path setup procedure, the SDH XC nodes are at risk of detecting these alarms and initiating restoration. To avoid such unanticipated error cases, we use tworoundtrip RSVP signaling. The first signaling establishes the path without monitoring configuration. After completing the first signaling, the second signaling provides an ADMIN_STATUS object to trigger the start of monitoring at each node.

3) Failure Notification: Failure information can be notified to edge nodes on either the data plane or the control plane. The former is SONET/SDH AIS (Alarm Indication Signal), and the latter is the GMPLS NOTIFY message. Each type of notification has both advantages and disadvantages. The use of AIS is more scalable than that of NOTIFY message since it never suffers from congestion of control messages. On the other hand, the use of NOTIFY messages makes it easier to interoperate between different types of nodes. Our test environment contains various types of nodes, so we decided the use of the RSVP NOTIFY message.

IV. PUBLIC DEMONSTRATION

A. Network Configuration

Figure 5 shows data plane network topology that was demonstrated in iPOP2005. The data plane in Japan consisted of Gigabit Ethernet, SONET/SDH (OC192, OC48, etc.) and the interfaces were connected by optical fibers and metal cables. As mentioned above (see II-B.3), the GMPLS control plane is separated from the data plane. Since GMPLS will be introduced to the backbone network, the distance between GMPLS nodes will be several hundreds of km and the control plane should be designed considering the expected transmission delay. In this demonstration, the control plane was implemented via the JGN II network, an open test bed network environment for research and development [24]. We divided the GMPLS nodes to two groups, GMPLS 1 and GMPLS 2, and connected both types to the JGN II network. GMPLS 1 nodes use a NTT router. GMPLS 2 nodes use other routers. A signaling message from GMPLS 1 to GMPLS 2 was transferred via 2,600 km, Tokyo – Osaka – Okayama – Fukuoka – Kanazawa – Osaka – Tokyo, and the delay was about 10 milliseconds. The control packets sent over this long distance were influenced by the delay. This environment provides a realistic assessment of long distance backbone networks.

We also established a network between Japan and ISOCORE in the USA via the Internet using the Ethernet over IP technique. The test equipment in Japan and the USA site shared a common control plane and a virtual data plane implemented using 1000BASE-SX/100BASE-TX media converters and speed converters.

This network was designed to validate the GMPLS and OUNI protocol suite mentioned in II and III.

B. Inter-operability Demonstration

We confirmed the establishment of LSC and MPLS LSPs. LSC LSP, called FA # 1, was established from one NEC node to another NEC node via HITACHI and FUJITSU nodes. The MPLS LSP was established from a Juniper node to another Juniper node via FA # 1 using the pre-provisioned approach. Additionally, we confirmed GMPLS-based protection. Working and protection LSC LSPs were established. The protection LSP's route was NEC – FUJITSU – NEC. The LSC LSP was switched to the protection LSP and no packet loss in the MPLS LSP was detected when link failure was forced to confirm GMPLS-based protection.

FSC LSP, called FA # 2, was triggered by a MPLS signaling message in the triggered signaling approach. The FSC LSP's route was NTT – SYCAMORE – FURUKAWA. The MPLS LSP's route was FUJITSU – FA # 2 – FURUKAWA. We succeeded in transferring streaming video data that needed 30 Mbps bandwidth via the MPLS LSP.

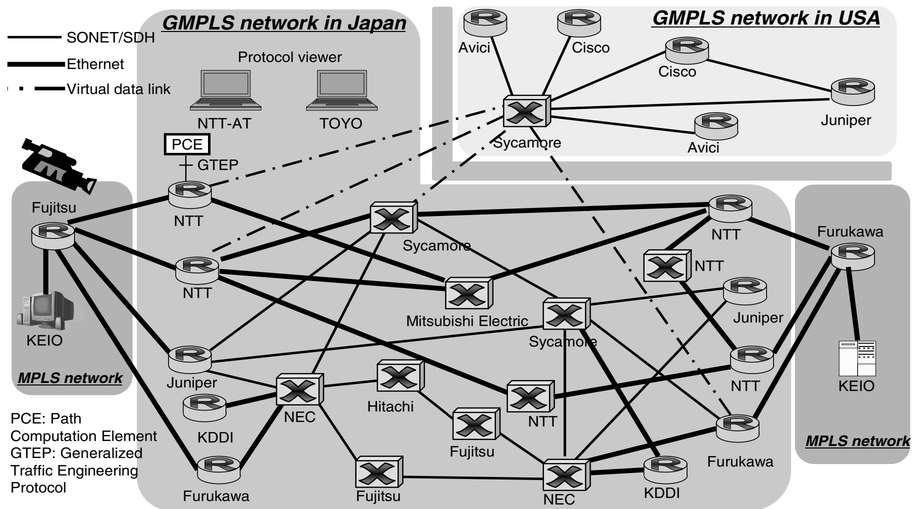

Fig. 5. iPOP2005 inter-operability test network

Another LSP was established by OUNI from FURUKAWA to AVICI via SYCAMORE. That LSP was established between Japan and USA across the virtual Ethernet.

Subset of the system of route selection according to the requested content was demonstrated. The client requested a large movie file from a content server via a proxy server. The proxy server forwarded the request message to the content server via 100 BASE-TX Ethernet, i.e. a narrow band connection. Upon receiving the request, the content server judged the content being requested and sent the movie file to the client via a wideband LSC LSP that differed from the route of the request message.

V. Conclusion

In this trial, we demonstrated four items. First, Demandled GMPLS LSP that was triggered by MPLS signaling message was established. Second, we confirmed collaboration between OIF UNI and GMPLS signaling. Third, we succeeded GMPLS-based recovery. We never found alarm when GMPLS LSP switched to backup route. Finally, we provided an example of GMPLS use by demonstrating some applications that demanded wide bandwidth GMPLS LSPs. Application-triggered GMPLS LSP whose switching type depended on the file type was successfully established.

This demonstration shows that GMPLS basic function is now available and Time that GMPLS is introduced into commercial network is near.

VI. Acknowledgement

The authors wish to thank Sycamore Networks, Juniper Networks, KDDI R&D Laboratories, TOYO Corporation, Avici Systems, and NTT Advance Technology for their participation in the GMPLS interoperability test and their technical cooperation. This experiment used the JGN II network to implement the part of the control plane network of our GMPLS test network. We are grateful to NTT Communications for providing a private line to connect the iPOP PIL site to JGN II access point and Internet access point. It must be remembered that ISOCORE contributed to the realization of the GMPLS test network at ISOCORE Inter-working Lab. in the United States. The authors are indebted to PIL members for their technical discussions and helpful support. PIL activities are supported by the SCOPE (Strategic Information and Communications R&D Promotion Program) of the MIC (Ministry of Internal Affairs and Communications) of Japan.

References

[1] "Ministry of Internal Affairs and Communications Information and Communications Statistics Database," http://www.soumu.go.jp/joho_tsusin/eng/main_data.html
[2] FCC, "Broadband Internet Access in OECD Countries: A Comparative Analysis," Oct. 2003.
[3] http://www.soumu.go.jp/joho_tsusin/eng/Releases/NewsLetter/Vol15/Vol21/index.html
[4] E. Mannie et al., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," IETF RFC, RFC 3945, Oct 2004.
[5] ITU-T Rec. G.8080/Y.1304, "Architecture for the Automatically Switched Optical Network (ASON)," Nov. 2001 (rev. Jan. 2003).
[6] http://www.pilab.org/
[7] N. Yamanaka, "Photnic Internet Lab.: Breakthrough for leading edge pohotonic-GMPLS" IEICE Trans. Commun. vol. E87-B, no. 3, pp. 573-578, Mar. 2004.
[8] http://www.pilab.org/news/e_030520.html
[9] http://www.pilab.org/news/e_040123.html
[10] http://www.isocore.com/
[11] http://www.pilab.org/ipop2005/
[12] E. Oki, K. Shiomoto, D. Shimazaki, N. Yamanaka, W. Imajuku, and Y. Takigawa, "Dynamic Multilayer Routing Schemes in GMPLS-based IP+Optical Networks," IEEE Commun. Mag., pp. 108-114, vol. 43, no. 1, Jan. 2005.
[13] D. Brungard, et al., "IP/MPLS-GMPLS interworking in support of IP/MPLS to GMPLS migration," IETF Internet Draft draft-oki-ccamp-gmpls-ip-interworking-05.txt, Feb. 2005, work in progress.
[14] Adrian Farrel, Jean-Philippe Vasseur, "Path Computation Element (PCE) Architecture," IETF Internet Draft draft-ietf-pce-architecture-00.txt, Mar. 2005, work in progress.
[15] Jerry Ash, J.L. Le Roux, "PCE Communication Protocol Generic Requirements," IETF Internet Draft draft-ietf-pce-comm-protocol-gen-reqs-00.txt, May. 2005, work in progress.
[16] E. Oki, D. Shimazaki, K. Shiomoto, "Generalized Traffic Engineering Protocol," IETF Internet Draft draft-oki-ccamp-gtep-02.txt, Feb. 2005, work in progress.
[17] http://www.oiforum.com/public/documents/UNIWP.pdf
[18] http://www.yamanaka.ics.keio.ac.jp/
[19] I. Nishioka, S. Kano, K. Kusama, N. Chaki,D. Muto, E. Horiuchi, E. Oki, "Interoperability Demonstration of GMPLS-based Path Restoration," Proc. ECOC2004, Mo3.6.1 Sep. 2004.
[20] R. MacDonald et al., "Requirements of Optical Layer Network Restoration," OFC2000 ThE3, Mar. 2000.
[21] J. P. Lang, Y. Rekhter and D. Papadimitriou, "RSVP-TE Extensions in support of End-to-End GMPLS-based Recovery," IETF Internet Draft draft-lang-ccamp-gmpls-recovery-e2e-signaling-03.txt, Feb. 2004.
[22] N. Wauters, et al., "Multi-Layer Resilience Tests in an IP Over Optical Switched Test-Bed in IST Project LION," ECOC2003, WI3, Sep. 2003.
[23] I. Nishioka et al. "Monolithic Control of Multi-Layer Optical Networks," Proc. ECOC2002, S204, Sep. 2002.
[24] http://www.jgn.nict.go.jp/e/index.html