OSRM Technology Project

Test Event and Results for January 2004

Introduction

Today's competitive business environment challenges service providers to diversify their portfolios while reducing operational expenses. Increasing network automation supports this objective. Traditional service offerings have been for the most part manually provisioned. This is an expensive, resource intensive task involving circuit planning, design, and physical on-site management of geographically dispersed network switches and routers. New protocols have recently emerged that propose to simplify the service provisioning process for IP and optical networks. This technology is referred to in general throughout this white paper as Optical Signaling, Routing and Management (OSRM). OSRM efforts have the ability to dynamically establish, maintain, and tear down optical connections via Internet Protocol (IP) based routing and signaling over a common control plane. OSRM's benefits are lower operational expenses and new, diverse service offerings. The standardization of OSRM protocols promises to facilitate interoperation between networks with different data plane technologies as well as reduce management plane complexity, which improves overall network scalability.

One such emerging suite of protocols is Generalized Multi-protocol Label Switching (GMPLS). GMPLS simplifies network management and operation by introducing IP-based routing and signaling into the optical domain. GMPLS leverages the efforts of the Internet Engineering Task Force's (IETF) Multi-Protocol Label Switching (MPLS) Working Group. It enables dynamic end-to-end provisioning, maintenance and tear-down of connections across the electrical and optical transport domains. The continuing maturation of GMPLS, as demonstrated in numerous interoperability events, has motivated many service providers to investigate this technology, and in response, numerous vendors are implementing GMPLS in their hardware and software solutions.

Validation of these emerging technologies from a carrier perspective is essential for commercial deployment. Currently service providers are used to the 99.999% uptime of SONET/SDH. To reach this objective it must be demonstrated that OSRM technology is reliable and robust to minimize network down time. Interoperability should be proven to allow multiple equipment suppliers into the same network.

This document describes the testing methodology and results of a GMPLS-focused test event that took place January 12-16 2004 at the University of New Hampshire InterOperability Laboratory (UNH-IOL) in Durham, New Hampshire. The testing was conducted by the UNH-IOL OSRM test group. The OSRM test event involved a variety of test cases that focused on stability testing, control channel fault handling and data channel failure recovery. During its course, the event gave rise to several milestone test scenarios, including:

- Validation of the Constrained Shortest Path First (CSPF) protocol in a GMPLS-enabled network. CSPF is essential to the value proposition of MPLS in three ways: automation, interoperabilty and multilayer capability, which allows devices operating in different domains to communicate with each other.

- Label Switched Path (LSP) behavior in a variety of fault conditions, including nodal fault and temporary loss of the control plane. Failure recovery and network stability is a baseline requirement for carriers intereseted in OSRM technologies.

The test event also uncovered several issues that warrant further discussion, investigation, and validation at future OSRM events. These issues are documented following the conclusion.

Participants

The OSRM test event included service provider Nippon Telegraph and Telephone Corp. (NTT) and the following leading equipment suppliers: Alcatel, Agilent Technologies, Juniper Networks, Navtel Communications, Movaz Networks and Sycamore Networks. Preliminary results show that IP-based optical network technology is maturing for commercial use.

Test Methodology

The UNH-IOL's OSRM test event provided an aggressive carrier-class environment for GMPLS testing including several GMPLS interworking scenarios that had not been achieved in previous multi-vendor test settings. The initial testing consisted of simple topologies of one or two core devices attached to an edge device. Bi-directional LSPs were established during these rotations, and many vendor specific implementation issues were resolved prior to more complex testing scenarios. Some test cases involved control plane only testing, other cases involved both data and control plane testing. While previous test events focused on these basic operations, the UNH-IOL test plan also proved additional aspects of this protocol that are critical to network operations, especially in the area of routing. These cases are detailed in the following sub-sections.

Multi-Path LSP Setup with OSPF-TE

GMPLS traffic engineering, which enables the ability to make intelligent connection selections based on network topology and bandwidth usage information, is critical to dynamic link provisioning, and thus is a critical feature for GMPLS deployment. GMPLS traffic engineering information should be carried by OSPF LSA messages over the control channel. The Unreserved Bandwidth sub-TLV carries the bandwidth and priority of a link transiting a given interface. This allows for all GMPLS neighbors to synchronize the used bandwidth between two interfaces and insures that links are not over subscribed. All nodes on the network exchange and store this OSPF-TE information. The edge device that is setting up the GMPLS LSP uses its database and the CSPF algorithm to automatically construct an ERO. This ERO creates the new LSP over the proper non-provisioned links.

CSPF Validation

Test Case #1. Basic CSPF Topology

OSPF-TE was enabled on the platforms and the databases were synchronized. The blue LSP was formed across the network as illustrated in Figure 1 with an ERO automatically generated by CSPF. The OSPF databases refreshed their values and data was properly passed through the transport network.

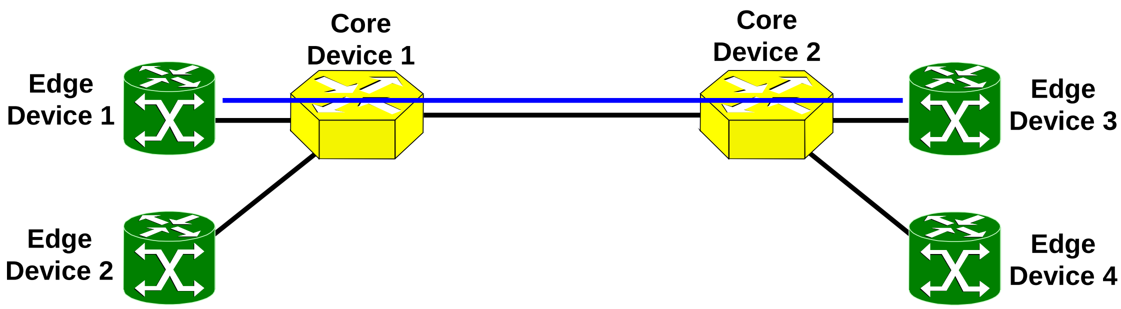

Figure 1: CSPF with Data Plane and Control Plane Functionality

Test Case #2. Advanced CSPF Topology

OSPF-TE was enabled on the platforms and the databases were synchronized. Displayed in Figure 2, each link has a metric of 1. The blue LSP was established and torn down. It was verified that the OSPF databases would properly update their values. The blue LSP was re-established. Edge device 2 then established the pink LSP across the topology, properly routed by CSPF. Attempted LSP setups from Edge Device 3 and Edge Device 6 properly failed due to no available bandwidth. The green LSP was properly setup.

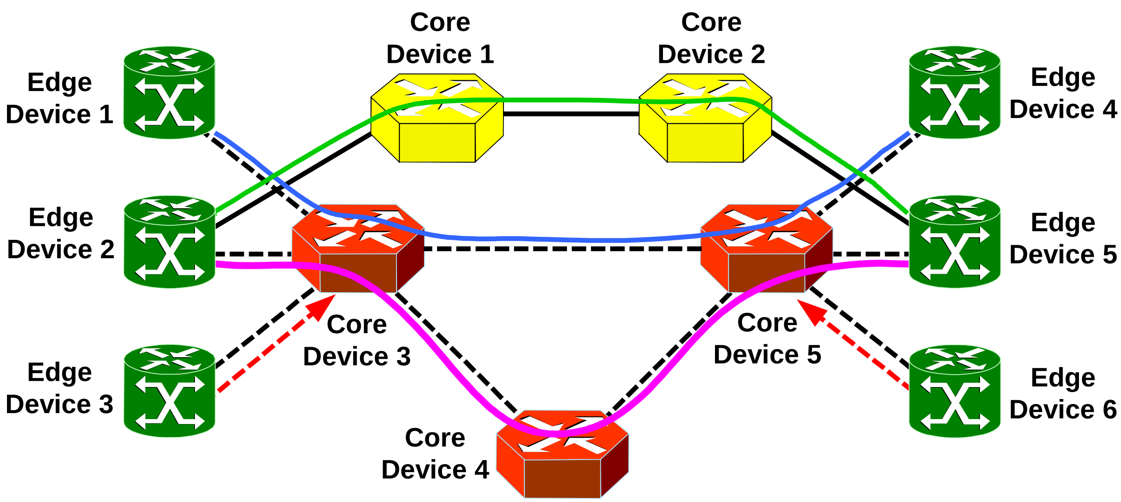

Figure 2: Advanced CSPF Topology

The LSPs were then all taken down. Edge Device 2 was configured to setup a LSC LSP to edge device 5. The edge device properly checked the configuration with the routing protocol and the expected (green) LSP was setup over the proper interface, demonstrating that the CSPF process can determine the difference in switching capabilities.

Fault Handling

Fault handling and failure protection are requirements for carrier class network deployments. Control channel fault occurs when control communication is lost between two nodes. Nodal fault occurs when a control module in a device fails or resets and a node losses its control state. GMPLS contains an extension that allows LSP data channel forwarding state to remain up in the event of the above failures. This feature can be negotiated among network devices for the forwarding state to remain either infinitely or for a set period of time. In order to insure proper notification of control channel faults and nodal faults, the Restart_Cap object is added to RSVP Hello Messages. The Restart_Cap object contains two time values. A device should wait Restart Time before initializing lost communication procedures in the event that a fault occurs with one of its neighbors. Recovery Time is how long a device will wait for a signaling refresh before determining that the forwarding state was lost. Restart Time starts at the point RSVP Hello state transitions to down with the neighbor.

The following test initialization was executed for all fault handling tests. The network was set up with two edge devices and two core devices connected between, similar to Figure 1. Once the LSP was set up and traffic was passing on the data channel, the below tests were executed. Restart Time was set to a valid fixed value for the below tests.

Test Case #3. Control Channel Fault, Restart Time set to infinity
The LSP was established and an infinite Restart Time was signaled into the network via RSVP Hello messages. The control channel link was disconnected and traffic was properly passed on the LSP for six minutes. When the link was reconnected the LSP was refreshed and the traffic flow on the data channel was uninterrupted.

Test Case #4. Control Channel Fault, Restart Time set to two minutes
The LSP was established and a Restart Time of two minutes was signaled into the network via RSVP Hello messages. The control channel link was disconnected and traffic was properly passed on the LSP for two minutes. At that point the devices properly ceased forwarding traffic. When the link was reconnected the LSP was re-established and traffic was again passed on the data channel.

Test Case #5. Nodal Fault, Restart Time set to infinity
The LSP was established and an infinite Restart Time was signaled into the network via RSVP Hello messages, the management module of one of the core devices was removed and traffic was properly passed on the LSP for six minutes. When the management module was replaced the LSP was refreshed and the traffic flow on the data channel was uninterrupted.

Test Case #6. Nodal Fault, Restart Time set to two minutes
The LSP was established and a Restart Time of two minutes was signaled into the network via RSVP Hello messages. The management module of one of the core devices was removed and and traffic was properly passed on the LSP for two minutes. At that point the devices properly ceased forwarding traffic. When the management module was replaced the LSP was re-established and traffic was again passed on the data channel.

Test Cases #7 and #8. Simultaneous Control Plane and Nodal Fault
The above scenarios were repeated, but with testing the simultaneous interruption of nodal fault and control channel fault. This was successfully demonstrated for a restart time of both infinity and two minutes.

Overall the testing for fault handling scenarios was demonstrated as described in the specification. One issue that was found was that some devices did not contain a full-featured user interface to configure different time values for LSP up time, as only one time value was supportted. There was also an implementation problem observed where the Hello Adjacency was dropped.

Test Case #9. Restoration
OSPF-TE was enabled on the platforms and the databases were synchronized. The red devices advertised FSC. The yellow devices advertised LSC. As displayed in Figure 2, each link has a metric of 1. The blue LSP was established (dashed blue). This is the best route according to the OSPF-TE database. The TE link between Core Device 3 and Core Device 5 was then torn down.

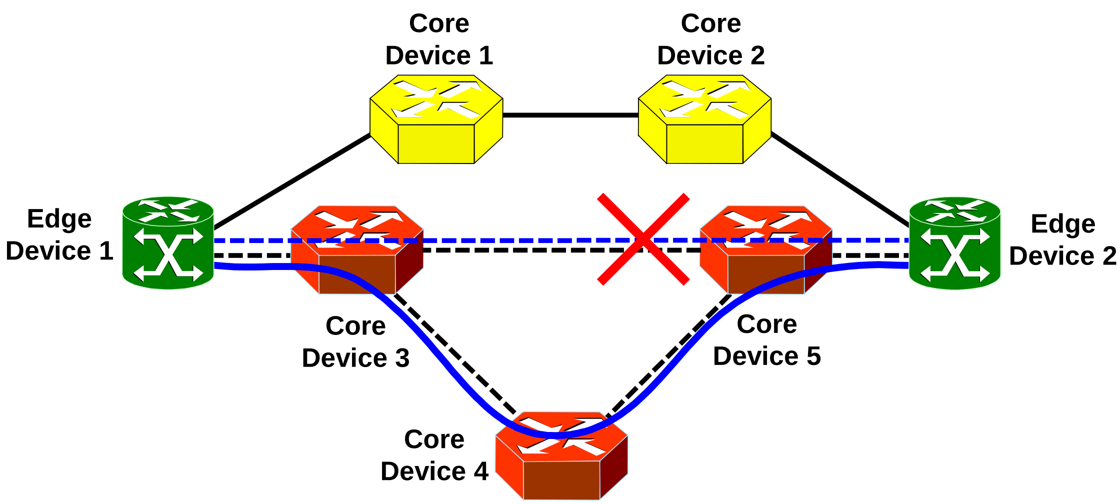

Figure 3: Restoration

The LSP was then reestablished across the higher cost route between Core Device 3, Core Device 4 and Core Device 5. Because Core Device 1 and Core Device 2 only support LSC, the route was correctly not chosen for the LSP restoration.

Discoveries for Further Investigation

Different Expectations for ERO Sub-objects

The tested devices expected different interface address types in the ERO sub-objects. There is no clear requirement for the ERO to contain sub-objects with either the incoming interface or outgoing interface of the next hop device. What was discovered was that Edge Device 1 sends a Path message with a strict ERO, containing sub-object address of the outgoing interface and

Device 2 sent a PathErr message, as it expected the ERO sub-object to contain the IP address of the incoming interface.

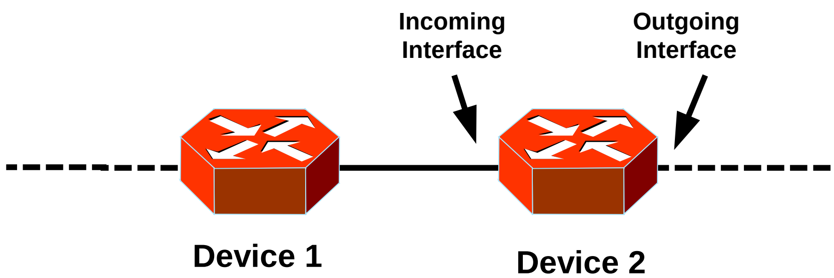

Figure 4: Incoming and Outgoing Interface, Considering LSP setup originates from Device 1 or a network behind Device 1

In the case of loose ERO, the route considering the address to be incoming interface address differs from the route considering the address to be outgoing interface address. The overall recommendation to maximize interoperability is implementations should be flexible in accepting any possible address in the ERO. RFC 3209 leaves this option open to allow for flexibility with the address sub-objects.

Tunneling for the Control Plane

There are several configurations for the GMPLS control plane. Each neighbor pair must select the same configuration for its control plane link. There are several ways of establishing proper control channels, none of which is easily described as "best." The issue is that some devices supported a subset of the possible features, creating some instances where the GMPLS control plane could not be realized. It is recommended that an implementation support a large enough subset of the tunneling options so that there will be overlap with other vendors.

1. Numbered, GRE tunnel
2. Unnumbered, GRE tunnel
3. Numbered, IPinIP tunnel
4. Unnumbered, IPinIP tunnel
5. Unicast point-to-point

In the future, an auto-configuration mechanism may ease the complicated configuration issues.

Advertising the Control Plane Link Information with OSPF

Control plane link information may be advertised in OSPF. Non-router devices can easily divide data plane from control plane because IP packets are used only for the control plane and lower layer transport technologies are used for the data plane. IP router devices route IP packets transiting both the data plane and the control plane. An optical interface and an IP routed interface behave differently with regards to advertising the control plane, and the test event highlighted the difference.

During the test, this was resolved by the creation of a static route in the node that neighbors the router. Considering that the topology of TE-link is different from that of the control link, the control plane interfaces should not be advertised in the routing protocols. The control channel is a management interface over which TE-link information is exchanged. The control channel is also a link, but it should not be advertised into OSPF. This will prevent data traffic from traversing the control channel.

Some of the operational solutions currently envisioned are:

- Set 2 separate OSPF instances for the packet switching node. One would be the IP network instance, where the device was participating as a full OSPF node.
- Only install control channel information into TE-LSAs, not into router LSAs.
- Define different OSPF areas.
- Define different Autonomous Systems (ASes) .

Conclusion

Carriers seeking a competitive edge in today's communications market must control the cost of network operation and provisioning while providing high-quality end-user services that are applicable to both existing and new market opportunities.

The requirements for meeting this multi-service future include the ability to:

- offer products and services at a lower cost. Optical routing and signaling reduce management expense by replacing costly, centralized designs and automating elements of the provisioning cycle;
- innovate and diversify service operations. By obviating the need for manual provisioning, OSRM enables dynamic service creation—bandwidth on demand for unique applications;
- increase speed of reaction time to customer demands. OSRM simplifies provisioning and management so new services can be deployed faster and carriers can be more responsive to service requests.

As GMPLS continues to progress in areas that drive new service creation and cost reduction, a key determinant of emerging protocol standardization and commercial adoption is validation in operative networks. To further these efforts, the OSRM test event at the UNH-IOL provided an aggressive test scenario built around service providers' requirements.

The UNH-IOL is committed to contributing to this validation process in whatever ways are needed. In addition to a wide variety of MPLS and GMPLS testing already completed, the UNH-IOL is particularly pleased with the outcome of the milestone testing achievements that emerged from the January OSRM event, particularly in areas of CSPF testing and failure recovery. CSPF is an essential technical requirement in order to realize the benefits of GMPLS as a whole. Validation of LSP failure recovery – demonstrating that even when the control plane is taken down, existing service will not be disrupted – is an essential requirement for commercial adoption.

OSRM technologies will become increasingly important as carriers consider new ways to build out their networks and diversify their service offerings. They are likely to be especially relevant to high-bandwidth and bandwidth-on-demand services such as those associated with grid computing and e-science applications.

Recommendations for Further Investigation

As a result of the OSRM test methodologies and findings described above, several facets of OSRM technology emerged as compelling candidates for further testing. Among these are the following:

- LSP formation with multiple switching capabilities: The testing and theory of LSP formation with multiple nodes that have different switching capabilities. There is still significant debate as to how this should work.
- scalability: Testing with a realistic number of Nodes, LSPs and setup/teardown patterns.
- extended OSPF-TE and CSPF topologies
- ERO and RRO label sub-objects
- additional protection and restoration scenarios
- Link Management Protocol (LMP)
- UNI and NNI signaling; NNI routing
- Hierarchical routing

The UNH-IOL looks forward to working with carriers and all participants in the first OSRM test event to further investigate these and other aspects that are equally important to realizing validation in complex operative networks.

References

Request for Comments 2205 – Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification

Request for Comments 3031 – Multiprotocol Label Switching Architecture

Request for Comments 3032 – MPLS Label Stack Encoding

Request for Comments 3209 – RSVP-TE: Extensions to RSVP for LSP Tunnels

Request for Comments 3471 – Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description

Request for Comments 3473 – Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions

Internet Draft draft-ietf-ccamp-gmpls-architecture-07.txt- Generalized Multi-Protocol ns Label Switching (GMPLS) Architecture

Internet Draft draft-ietf-ccamp-gmpls-routing-09.txt - Routing Extensions in Support of Generalized MPLS

Internet Draft draft-ietf-ccamp-ospf-gmpls-extensions-12.txt - OSPF Extensions in Support of Generalized Multi-Protocol Label Switching

Internet Draft draft-ietf-ccamp-gmpls-recovery-terminology-03.txt - Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)

Internet Draft draft-ietf-mpls-lsp-hierarchy-08.txt – LSP Hierarchy with Generalized MPLS TE

Contributors

The whitepaper content is an accumulation of agreements, input, and comments from all participants. Ichiro Inoue, Kaori Shimizu and Yumiko Kawashima from NTT Network Service Systems Laboratories and Ben Schultz from UNH-IOL wrote the initial content. Special thanks to Tom Dimicelli from Sycamore Networks and Chris Volpe from UNH-IOL for creating the introduction and conclusion sections. Much appreciation to Vijay Pandian from Sycamore Networks for his work and documentation of LSP Establishment Across Multiple Switching Capabilities.