Network Working Group
Request for Comments: 3604
Category: Informational

H. Khosravi
Intel
G. Kullgren
S. Shew
Nortel Networks
J. Sadler
Tellabs
A. Watanabe
NTT

October 2003

Requirements for Adding Optical Support to the General Switch Management Protocol version 3 (GSMPv3)

Status of this Memo

Copyright Notice

Abstract

This memo provides requirements for adding optical switching support to the General Switch Management Protocol (GSMP). It also contains clarifications and suggested changes to the GSMPv3 specification.

Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [1].

1. Overview

This document details the changes to GSMP necessary for the support of optical (non-transparent and all optical), SONET/SDH, and spatial switching of IP packets, Layer 2 (L2) frames and TDM data. When implemented, GSMP controllers will then be able to control: photonic cross-connects (optical-optical), transparent optical cross connects (optical-electrical-optical, frame independent), opaque cross connects (optical-electrical-optical, SONET/SDH frames), and traditional TDM switches (all electrical). The resulting systems could form IP based optical routers, optical label switches, wavelength routers, and dynamic optical cross connects.

Several different generic models exist defining how to provide control plane functionality in an optical network [2], [3], [4]. This document takes no position on which model is most appropriate (e.g., single or multiple routing plane instances). The only assumption is that the ability to separate the control mechanisms from the data switching is as useful for the signaling of optical paths (e.g., GMPLS) as it is for the signaling of L2 paths (e.g., MPLS). Therefore, the requirements contained within are focused only on the separation of control functions from data functions in order to provide a more flexible network architecture.

GSMPv3 [5] is well suited for providing the control interface necessary for allowing an IP based controller to direct the activities of an optical switch. In order for GSMP to operate between controllers and optical switches and cross connects, support for optical labels and service and resource abstractions must be added to GSMP.

This document also includes changes recommended by implementers that will facilitate easier development of a GSMP implementation. These changes consist of rearranging PDU formats, clarification of flags, transaction identifiers, and response codes.

2. Requirements for Optical Support

2.1. Label

2.1.1. Label Types

New labels are needed to identify the entities that are to be switched in the optical fabric. These are longer than the labels defined in GSMPv3 as they have physical and structural context. As GMPLS [2], [3] has had very similar requirements for label formats, alignment with GMPLS is proposed. This includes support for:

- Digital Carrier Hierarchy (e.g., DS-1, E1)
- SONET and SDH Hierarchy (e.g., OC-3, STM-1, VT1.5, VC-12)
- Plesiochronous Data Hierarchy (PDH) labels [6]
- OTN G.709 labels
- Lambdas
- Fibers

GSMP MUST include support for all label types list above, as well as for label hierarchies and label lists as defined by GMPLS. Therefore, the ability to perform operations on groups of the above labels MUST also be supported (e.g., 5 OC-3s, contiguous wavebands).

2.1.2. Label Management Issues

An updated label range message MUST be provided. There MUST also be support of multiplexing (e.g., no multiplexing, SONET, Gigabit Ethernet multiplexing etc).

2.2. Statistics messages

Optical switches have a number of different statistics which are not in common with ATM, or Frame Relay switches. Consequently, the statistics messages SHOULD be updated to report Performance Monitoring statistics defined for all new optical transport technologies added to GSMP.

2.3. Configuration Issues

2.3.1. Switch Configuration

2.3.1.1. Layer Switching Identification

Since an Optical Switch may be able to provide connection services at multiple transport layers (i.e., STS-3c, STS-1, VT-1.5, DS3, DS1), and not all switches are expected to support the same transport layers, the switch will need to notify the controller of the specific layers it can support.

Therefore, the Switch Configuration message MUST be extended to provide a list of the transport layers for which an optical switch can perform switching.

2.3.2. Port Configuration

The port configuration message supplies the controller with the configuration information related to a single port. Consequently, extensive additions will need to be made to this command.

2.3.2.1. Port Type extensions

Port types MUST be added to support the mix of optical signals that can operate over a single fiber.

The port information that MAY need to be conveyed includes [7]:

- wavelengths available per interface
- bit rate per wavelength
- type of fiber

2.3.2.2. Supported Signal Type extensions

Since a port on an optical switch may support signals at multiple transport layers, it is necessary to understand the signals supported, as well as the possible ways that one signal can be transported within another.

For OTN, SONET/SDH and PDH optical switches, the Port configuration message MUST be extended to detail the different transport layer signals that are supported by a port. Furthermore, this extension MUST detail which signals may be transported by another signal.

This mechanism MUST also provide information about optional capabilities (such as virtual concatenation and arbitrary concatenation for SONET/SDH) available on a port.

2.3.2.3. Trace Mechanism support Identification

A number of transport layer signals include overhead channels that can be used to identify the source of a signal. Since they are embedded in the signal, only the network element has access to the signals. However, not all network elements have the capability to set or read the messages in these channels on every port. Consequently, this port attribute needs to be reported to the controller.

The Port Configuration message MUST be extended to report which trace mechanisms are supported by a port.

2.3.2.4. Port Location Identification

Since contemporary Optical switches have the ability to support tens of thousands of ports in hundreds of shelves located in as potentially as many bays, the current "Slot/Port" location identifier is inadequate.

The Slot/Port Location Identifier MUST be extended to encode Bay/Shelf/Slot/Port.

2.3.2.5. Port-related Partitioning Extensions

Partitioning can be done for any resource that exists in the network element. The GSMP partitioning draft currently defines ports and switching resources as partitionable resources. Since optical switches may support multiple transport network layers, an additional resource type is introduced: the transport layer signal.

The point where a transport layer signal is inserted into a lower layer signal (called an "access point" by the ITU [8]), is very similar to a port. Therefore, when partitioning is done on a transport layer signal basis, the partition that is the user of the access point MUST have a port that associated with the access point. Labels will then be used in the to describe the subordinate signals.

2.3.3. Service Configuration

While new capability sets MUST be added to support quality parameters in optical switches, no changes are foreseen to the service configuration message as its role to carry the service information as defined in the applicable service model.

2.4. Service Model Issues

While one assumption of using optical media is that bandwidth is plentiful, it should be expected that traffic engineering will be necessary in any case [5]. GSMP provides the means for each connection to be created with specific attributes. Therefore, service parameters will need to be defined for each of the Different Optical technologies.

2.4.1. Transparent Optical

Capability to control re-timing and re-shaping on a per port level MUST be added.

2.4.2. SONET/SDH and OTN

The capability to control the adaptation parameters used when a transport signal is inserted into another transport signal MUST be added. These parameters SHOULD be modifiable at times other than adding a branch so that functions such as Tandem Connection Monitoring can be configured. Currently, the default set of service models in GSMP are all based on the services models defined elsewhere, e.g., the Intserv model [9], [10], the Diffserv [11] model, ATM QoS models and the Frame relay forum QoS models. A determination needs to be made of the applicable service models for optical channel trails. These models MUST then be mapped to the GSMP capability set mechanism.

2.5. Encapsulation issues

The working group needs to decide whether a new encapsulation is required. In other words, will all optical switches used in either the MPLS over Optics and the IP over optics applications require that IP be implemented on the control channel connecting the GSMP controller and Optical switch (the GSMP target).

A new encapsulation SHOULD be defined allowing the use of a non-IP raw wavelength control connection.

Likewise, a new encapsulation SHOULD be defined allowing GSMP to be used in legacy Data Communication Network (DCN) environments that use OSI CLNP.

The security risks of additional non-IP encapsulations MUST be described, since the mandatory to implement mechanism of IPsec is not available for these control channels, as in the RFC 3293 Ethernet and ATM cases. It is in scope to perform risk analysis and describe if mechanisms for link-level security mitigate the risk.

2.6. MIB Issues

If a new encapsulation is defined, then the encapsulation group SHOULD be updated. No other changes should be required.

2.7. OXC Transaction Model

2.7.1. Serial Transactions

Many existing OXCs use a command interface which assumes a serial transaction model. That is, a new command cannot be issued or processed until the existing command is completed. Under provisioning control via a network management application, and with non-dynamic path setup, this model has been adequate.

Moving to a dynamic path setup capability with a distributed control plane, a parallel transaction model is likely required for performance. This is particularly helpful when the performance of setting up a TDM style connection is much slower than setting up an L2 connection table. If the OXC is not able to support a parallel transaction model, a GSMP controller MUST be informed of this and adopt serial transaction behavior.

2.7.2. Bulk Transactions

Again due to the time it may take some OXCs to setup TDM connections relative to L2 fabrics (e.g., VC-4/STS-1 SPE fabric in an HOVC/STS switch), support for sending multiple transactions in the same message is a useful optimization. When an OXC receives a bulk message, the individual transactions are acted upon and a single reply is sent. If parallel transactions are not supported, bulk messages can improve performance by reducing transaction overhead. Bulk transactions SHOULD be supported.

2.8. OXC Protection Capabilities

To achieve good link protection performance (e.g., 50 ms after failure detection), SONET/SDH and some OXC systems use hardware based protection schemes (e.g., ring protection). Achieving this level of performance solely using a data control plane such as GMPLS is a serious challenge. An alternate approach is to utilize protection capabilities of an OXC with a dynamic control plane. An implication of this hybrid approach is that extensions are needed to GSMP to provision the behavior of an OXC in anticipation of a link failure.

This differs from the strict master-slave relationship in GSMP for Layer 2 switches in that here the OXC is capable of taking an action independent of the GSMP controller and then informing the controller afterwards. Consequently, the GSMP port configuration command MUST be extended to allow autonomous protection behaviors to be provisioned into the Network Element.

Furthermore, the controller MUST be able to provide the parameters for when reversion from a backup link to the original link is allowed. This may take the form of hold-off timers, BER parameters, or the requirement for controller directed reversion.

2.8.1. Non-Reserved Protection Links

An example of protection OXC behavior is that when a link fails, a backup link may be used to protect traffic on. This backup link could be selected from a set of links, none of which are pre-reserved. A backup link could be shared with one or more "working" links which is a form of 1:n shared protection. Specifying the set of possible backup links SHOULD be done as an option to the Add-Branch message.

When a backup link is used or the OXC reverts back to the original link, the control plane (i.e., signaling) may need to know about the new path state in order to notify the operator, or take some other OAM action (e.g., billing, SLA monitoring). An additional GSMP message to inform the controller SHOULD be added to do this.

2.8.2. Dedicated Protection Links

A more specialized form of restoration called "1+1" defines a (usually node disjoint) protection path in a transport/optical network for a given working path. At the ingress node to the path, the traffic signal is sent simultaneously along both working and protection paths. Under non-failure conditions at the egress node, only the last link of the working path is connected to the client. When any link in the working path fails, traffic on the working path ceases to be received at end of the path. The egress OXC detects this condition and then switches to use the last link of the protection path without the controller having to issue a Move-Input-Branch message. At no time is the ingress node aware which link the egress node is using. Selection of the protection path and all of its links is outside the scope of GSMP.

Specification of the two output branches at the ingress node can be done with the usual Add-Branch semantics. The ingress node protection link is not shared with any other working link.

Specification of the two input branches at the egress node should be done when the Add-Branch message is sent. This SHOULD be an option to that message. The egress node protection link is not shared with any other working link.

When a protection link is used or the OXC reverts back to the working link, the control plane (i.e., signaling) may need to know about the new path state in order to notify the operator, or take some other OAM action (e.g., billing, SLA monitoring). An additional GSMP message to inform the controller SHOULD be added to do this.

If an alternate input port is not specified with an original Add-Branch message, it MAY be specified in a subsequent Add-Branch message. In this case, it is useful to include information about existing users of the output port in that Add-Branch message. This helps the OXC immediately learn of the association between the new input port and an existing one. The association is used to enable OXC protection procedures. This capability MUST be added to the add-branch message.

Similar contextual information is needed for a Delete-Branch message so that the OXC can determine if a path becomes unprotected. This capability MUST be added to the Delete-branch message.

2.8.3. Protection Triggers

Aside from link or equipment failures, there are a variety of maintenance conditions that could cause the backup/protection link(s) to be used. These may include:

- Scheduled maintenance of the working link. Here the network operator deliberately takes a link out of service to perform maintenance.
- Reconfiguration of fiber/node/network which causes temporary need to use backup links.

It may be useful to specify these triggers when the backup/protection links are defined with the Add-Branch message. This depends on how the OXC is implemented to be aware of such triggers. This is for further study.

2.8.4. Protection Link Capabilities

When an OXC has the capability to perform protection switching independently from the Optical Call Controller (OCC), it may be useful for the OCC to be informed of these capabilities at switch and/or port configuration. Applications in the GSMP controller could use this information. For example, signaling clients could define a path protection scheme over multiple GSMP enabled OXCs. This is for further study.

2.9. Controller directed restoration

Bi-directional Connection Replacement

Connections in the transport network are inherently point-to-point bi-directional. Unfortunately, GSMPv3 currently does not allow for the B and R flags to be set on an add branch message. This means that it is not possible to do an atomic replacement of a bi-directional connection -- an action that is desirable for controller directed restoration. Consequently, the protocol MUST be changed to allow these flags to be used at the same time.

2.10. Support for optical burst switching

GSMP for Optical Switching should also support optical burst switching. As described in [12], [13], and [14], part of burst switching connection setup includes reserving time on the transport medium for the client.

This time is characterized by two parameters: a start time and the duration. These values MAY define a one-time reservation or a repeating reservation. Upon a request for setup of a burst connection, the GSMP controller MUST perform appropriate Connection Admission Control for the time and duration specified and, if the connection is allowed, MUST signal these parameters to the burst switching device to reserve the exact bandwidth required [12], [14]. The burst switch MUST perform the switching operation autonomously, using the synchronization methods prescribed for the burst network it is operating in.

3. Requirements from Implementers

This section describes requirements to GSMP v3 based on some implementation experience. They address areas of ambiguity, missing semantics, and configuration recommendations.

3.1. GSMP Packet Format

The Basic GSMP Message Format in chapter 3.1.1 in [5] describes the common fields present in all GSMP messages except for the Adjacency protocol.

3.1.1. Message segmentation

If a message exceeds the MTU of the link layer it has to be segmented. This was originally done with the "More" value in the Result field. The addition of the I flag and the SubMessage Number to the header has made the "More" value obsolete.

The I flag and SubMessage numbers should be used in all messages that can be segmented.

3.1.1.1. SubMessage Number and I flag

It should be specified if the SubMessage Number starts on 0 or 1 in a segmented message and what value the I flag should have in an message that is not segmented.

3.1.1.2. Message Length

Clarification of what value should be used in the Length field for segmented messages. Specifically, does the Length field contain the total length of the message or the length of the current segment.

3.1.1.3. Message Segmentation example

To avoid all ambiguity an example of message segmentation should be provided.

3.1.2. Transaction Identifier

The Transaction Identifier in [5] does not distinguish between replies from a request with "AckAll" and "NoSuccessAck". It also does not provide any information about how to handle replies where the Transaction ID doesn’t match a Transaction ID from a previously sent request.

If multiple controllers are connected to a single switch and the switch sends an event message with "ReturnReceipt" set to all of them, there is no way for the switch to identify which controller the receipt is coming from.

The "ReturnReceipt" value should not be permitted for Events.

3.2. Window Size

The Switch Configuration Message defined in chapter 8.1 in [5] defines a Window size to be used by the controller when sending messages to the switch. It is not stated if this window should apply to all messages or only to messages that will always generate a reply.

If messages that may not generate a reply should be counted against the window a time-out period when they are to be removed from the window should be defined.

It is not defined if the window should be cleared when the adjacency is lost and later recovered.

3.3. Retransmission

A retransmission policy with a well-designed exponential backoff should be used if no reply is received for a message with "AckAll" set.

3.4. Delete Branches Message

The "Delete Branch Element" has a 4 bit Error field that should be redefined to match the size of the "Failure Response Codes".

3.5. Adjacency

The chapter about how to handle a new adjacency and re-established adjacencies should be clarified.

3.5.1. Loss of Synchronization

The switch must not reset the connection states if another adjacency has already been established since this would destroy an already valid state.

4. Security Considerations

The security of GSMP's TCP/IP control channel has been addressed in [15]. Any potential remaining security considerations are not addressed in this requirements document.

5. Acknowledgements

The list of authors provided with this document is a reduction of the original list. Currently listed authors wish to acknowledge that a substantial amount was also contributed to this work by: Avri Doria and Kenneth Sundell

The authors would like to acknowledge the careful review and comments of Dimitri Papadimitriou, Torbjorn Hedqvist, Satoru Okamoto, and Kohei Shiomoto.

6. References

6.1. Normative References

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

6.2. Informative References

[2] Berger, L., Ed., "Generalized MPLS - Signaling Functional Description", RFC 3471, January 2003.

[3] Mannie, E., et al., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", Work in Progress, May 2003.

[4] ITU-T Recommendation, "Architecture for the Automatically Switched Optical Network (ASON)", G.8080/Y.1304, January 2003

[5] Doria, A., Sundell, K., Hellstrand, F. and T. Worster, "General Switch Management Protocol V3", RFC 3292, June 2002.

[6] Sadler, J., Mack-Crane, B., "TDM Labels for GSMP", Work in Progress, February 2001.

[7] Rajagopalan, B., et al., "IP over Optical Networks: A Framework", Work in Progress, September 2003.

[8] ITU-T Recommendation, "Generic functional architecture of transport networks", G.805, March 2000.

[9] Braden, R., Clark, D. and S. Shenker, "Integrated Services in the Internet Architecture: An Overview", RFC 1633, June 1994.

[10] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, September 1997.

[11] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, _"An Architecture for Differentiated Services", RFC 2475, December 1998.

[12] C. Qiao, M. Yoo, "Choice, and Feature and Issues in Optical Burst Switching", Optical Net. Mag., vol.1, No.2, Apr.2000, pp.36-44.

[13] Ilia Baldine, George N. Rouskas, Harry G. Perros, Dan Stevension, "JumpStart: A Just-in-time Signaling Architecture for WDM Burst-Switching Networks", IEEE Comm. Mag., Fab. 2002.

[14] Sanjeev Verma, et al. "Optical burst switching: a viable solution for terabit IP backbone", IEEE network, pp. 48-53, Nov/Dec 2000.

[15] Worster, T., Doria, A. and J. Buerkle, "GSMP Packet Encapsulations for ATM, Ethernet and TCP", RFC 3293, June 2002.

7. Authors' Addresses

Hormuzd Khosravi
Intel
2111 NE 25th Avenue
Hillsboro, OR 97124 USA

Phone: +1 503 264 0334
EMail: hormuzd.m.khosravi@intel.com

Georg Kullgren
Nortel Networks AB
S:t Eriksgatan 115 A
P.O. Box 6701
SE-113 85 Stockholm Sweden

EMail: geku@nortelnetworks.com

Jonathan Sadler
Tellabs Operations, Inc.
1415 West Diehl Road
Naperville, IL 60563

Phone: +1 630-798-6182
EMail: Jonathan.Sadler@tellabs.com

Stephen Shew
Nortel Networks
PO Box 3511 Station C
Ottawa, ON
K1Y 4H7

EMail: sdshew@nortelnetworks.com

Kohei Shiomoto

EMail: Shiomoto.Kohei@lab.ntt.co.jp

Atsushi Watanabe
Nippon Telegraph and Telephone Corporation
807A 1-1 Hikari-no-oka, Yokosuka-shi
Kanagawa 239-0847, Japan

EMail: atsushi@exa.onlab.ntt.co.jp

Satoru Okamoto
Nippon Telegraph and Telephone Corporation
9-11 Midori-cho 3-chome, Musashino-shi
Tokyo 180-8585, Japan

EMail: okamoto@exa.onlab.ntt.co.jp

8. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.