MPLS-TP Linear Protection for ITU-T and IETF

The MPLS-TP linear protection specification has resulted in a single unified solution that has been published in IETF RFCs and in ITU-T Recommendation G.8131.

Jeong-dong Ryoo, Taesik Cheung, Daniel King, Adrian Farrel, and Huub van Helvoort

Abstract

The MPLS Transport Profile (MPLS-TP) is a framework for the construction and operation of reliable packet-switched transport networks based on the architectures for MPLS and Pseudowires. Its development has been shared between the IETF, where the MPLS expertise resides, and the ITU-T, with its historic understanding of transport networks. MPLS-TP adds two significant features to the MPLS toolkit: Operations, Administration, and Maintenance (OAM), and linear protection switching. Unlike OAM, which resulted in two application specific and incompatible standards being approved at the World Telecommunication Standardization Assembly (WTSA) in November 2012, the MPLS-TP linear protection specification has resulted in a single unified solution that has been published in IETF RFCs and in ITU-T Recommendation G.8131. This article outlines the novel concepts and operation principles of the unified MPLS-TP linear protection switching mechanism and discusses how it differs from pre-existing solutions. In addition, the issues of compatibility with pre-existing solutions and the applicability to other network topologies are discussed.

Introduction

The Multiprotocol Label Switching Transport Profile (MPLS-TP) [1] is designed to be used in an environment that operates with or without an IP-based control plane, meaning that MPLS-TP provides functionality for a centrally controlled transport network (such as in Software Defined Networks (SDN)) or may be integrated with an existing IP/MPLS packet network. In order to enable transport network qualities in an MPLS packet network, MPLS-TP enhances the network’s reliability using Operations, Administration, and Maintenance (OAM) to detect and isolate faults, and rapid protection switching (sub-50ms) in the event of failure. These additional functions give the packet network the look and feel of traditional transport networks while building on top of the MPLS architecture.

The roots of MPLS-TP go right back to the original specification of MPLS within the Internet Engineering Task Force (IETF) more than 13 years ago. Since then, MPLS has become an established Internet technology and most packets will traverse an MPLS network somewhere along their end-to-end paths. However, as an Internet technology, MPLS was focused on best-effort routing and connectionless delivery mechanisms. As the possibility arose to utilize the same forwarding hardware in more static environments, it became desirable to add a number of mechanisms to MPLS so that the packet network could perform more like a circuit-switched transport network. The International Telecommunications Union (ITU) used its many years of experience with transport networks to develop requirements and to propose protocol solutions to define what was then called Transport-MPLS (T-MPLS).

As is not uncommon when two worlds collide, the resulting standardization activity resulted in two approaches. Since OAM was the first area worked on, two different, incompatible solutions were developed for MPLS-TP OAM. One is built on pre-existing MPLS diagnostic tools such as Label Switched Path Ping (LSP Ping) and Bidirectional Forwarding Detection (BFD) enhanced through new OAM protocols that can be carried in the MPLS Associated Channel (ACH). The other is developed from the pre-existing ITU-T Ethernet OAM documented in G.8013. Despite very many hours spent debating the merits of the two approaches and the desirability of a single standard for MPLS-TP OAM, agreement could not be reached among ITU-T participants, and the result was the publication of two alternative recommendations: G.8113.1 and G.8113.2.

The next piece of MPLS-TP technology to be worked on was for linear protection switching. As described in [2], linear protection is a protection mechanism that provides rapid protection so that traffic following one path through the network can be switched to a backup path when the working path fails or falls below an acceptable standard, or when an operator command is issued. In a mesh network, linear protection provides a very suitable protection mechanism because it can operate between any pair of points within the network and it can protect against failures in a node, link, transport path segment, or an entire end-to-end transport path. Linear protection relies on a coordination protocol that runs between the end points of the protected path to report errors and to determine what switching actions should be taken. Realizing the problems caused by the existence of two OAM solutions, everyone was particularly concerned to ensure that a single, unified MPLS-TP linear protection protocol and process would be standardized.

After lengthy discussion, Study Group 15 of the ITU-T agreed to develop a single solution for MPLS-TP linear protection that fully meets the ITU-T’s requirements by following the normal procedure for creating an RFC in the IETF. Since then, the IETF has made impressive progress toward RFC 7271 [3]. Subsequently, the ITU-T revised G.8131 [4] with the solution specified in [3]. This solution is called Automatic Protection Coordination (APC) in G.8131 [4], and is called “Protection State Coordination (PSC) in Automatic Protection Switching (APS) mode” in [3].

Jeong-dong Ryoo is with ETRI and the Korea University of Science and Technology.

Taesik Cheung is with ETRI.

Daniel King is with Lancaster University.

Adrian Farrel is with Old Dog Consulting.

Huub van Helvoort is with Hai Gaoming BV.

COMMUNICATIONS STANDARDS
The APS solution was considered in the IETF, but failed to achieve MPLS Working Group consensus. In order to document existing implementations and deployments, this pre-standard solution has been published as an Independent Stream RFC.

The developers of the PSC protocol looked to optimize their solution based on the fact that it would only be applied in a packet network, but still looked to re-use many of the concepts familiar in other protection switching systems. However, this led to some significant differences between the protocol messages, state machines, and principle of operation of this approach and those of the APS protocol be specified by the ITU-T and used in linear protection for traditional transport networks, such as Synchronous Digital Hierarchy (SDH), Optical Transport Network (OTN), and Ethernet transport networks.

As protection switching operation should complete in one message exchange without any acknowledgement from the other side, the protocol complexity in a single-phased protocol is disposed toward the state machine. Some examples of problematic scenarios of the PSC can be found in [3].

**A Unified MPLS-TP Linear Protection Solution**

**MPLS-TP Automatic Protection Coordination (APC)**

The experience gained during the development of the two solutions described above was used to make a new unified solution for MPLS-TP linear protection. Through the convergence of the PSC and the APS, the APC is now able to solve the aforementioned deficiencies and render improvements on those solutions.

**Design Principles of APC**

The APC is designed according to the following principles:

- **Maintain traditional network operational behaviors:** For the network operators who have been accustomed to the linear protection schemes seen in other transport networks, bits on the wire and internal mechanisms may not be so meaningful, but maintaining the same operational methods to manage their transport networks is beneficial; it can reduce training costs and simplify operation across multiple transport networks of different technologies.

- **Define additional mechanisms seen in other transport networks:** The additional mechanisms that are essential, but missing from the PSC, have been identified as: an operator command to manually switch traffic from the protection path to the working path (Manual Switch to
An important feature of an evolving protocol solution is that it should be backward compatible with deployed equipment, facilitating new equipment to be rolled out incrementally without the need for a flag day across the whole network.

Working (MS-W); an operator command to test protection mechanisms (Exercise (EXER)); and protection switching against an SD defect.

Define an efficient way to provide protection against SD: In the APC, traffic duplication is needed only under SD conditions.

Reuse the basic operation principle of the PSC: The basic operation principle of the PSC, which always reflects its local request in the transmitted PSC protocol messages even when the remote request from the other end has a higher priority, is enforced to reduce the issues of the time for round trip protocol message exchanges in out-of-service cases. Even if the round trip time issue is acceptable in the single phase protocol, it is desirable to avoid.

Reuse the PSC PDU structure: Considering the fact that all the necessary information to perform protection switching is defined for both Protocol Data Units (PDUs) of the APS and the PSC, it is natural to reuse the PSC PDU structure since the PSC achieved IETF MPLS WG consensus and was published as a Standard Track RFC.

Strictly decouple priority evaluation from state machine: In order to define a simple and clean description of the state machine inside each end of a protected domain, priority evaluation for various inputs and state transition table lookup are strictly partitioned.

Reduce possible bugs in state transition tables: By categorizing various inputs carefully and defining a comprehensive operation for each grouped input, any potential bugs in state transition tables can be reduced.

PDU Format of APC

Figure 1 depicts the APC PDU format of G.8131, which is framed in the Generic Associated Channel (G-ACh) as described in the IETF RFC 5586 [8]. Like other MPLS-TP OAM PDUs, the APC specific information is preceded by the four-octet G-ACh Label (GAL) and the four-octet Associated Channel Header (ACH).

Sixteen octets are allocated for the APC specific information. In the APC specific information, all the values except “Request,” “Fault Path,” and “Data Path” remain the same as configured by the operator. The first two bits are for the Version (V) of the protocol. The Protection Type (PT) field is to indicate the switching type, which can be unidirectional or bidirectional, and the bridge type, which can be a permanent bridge or a selector bridge. The Revertive (R) field is to indicate either revertive or non-revertive operation.

The values of “Request,” “Fault Path” and “Data Path” can be changed to provide protection switching against defects and operator commands. They are shown in Table 1.

The remaining fields in the APC specific information are to indicate the protocol capabilities encoded in Capabilities TLV. The description on the Capabilities TLV can be found in [3]. For the protocol operation in G.8131, the fields of the Capabilities TLV should be set as shown in Fig. 1.

APC Process Operation

Figure 2 shows the APC process algorithm, which is performed at both ends of the protected domain. The APC process algorithm is initiated immediately every time one of the input signals changes, i.e. when the status of any local defect (signal fail and signal degrade) changes, when an operator command (lockout of protection, forced switch, manual switch, exercise, etc.) is issued, or when a different APC-specific information is received from the remote end. When there is a need to coordinate timing of protection switches at multiple layers or across cascaded protected domains, a defect may be delayed in the “hold-off timer logic” before being processed.

As multiple local inputs may be active at one time, the “local request logic” determines which of these inputs is of highest priority. The highest priority local input (highest local request) is passed to the “global request logic,” that will determine the higher priority request (top priority global request) between the highest local request and the last received remote request.

When a remote APC protocol message arrives, its APC-specific information is subject to the “validity check.” By comparing the received APC-specific information with the transmitted, the “validity check” declares a protocol failure if the bridge type or the Capabilities TLV mismatches or the protection switching is not completed within 50 ms. When the remote request specified in the APC-specific information comes to the “global request logic,” the top priority global request is determined between this remote request and the highest local request that is present. If the remote request becomes the top priority global request and the highest local request is an operator command, the local command is cancelled.

The top priority global request is then presented to the “state transition logic” to determine the state transition. The consequent actions
of the state transition are to set the local bridge and selector positions and to determine the values of the variable fields for new APC-specific information. If revertive operation is configured, then the Wait-to-Restore (WTR) timer is started to prevent frequent operation of the protection switch due to an intermittent defect. For detailed descriptions of the APC process algorithm, refer to [3] and [4]. Some operation examples of the APC protocol can also be found in [3].

### Backward Compatibility

An important feature of an evolving protocol solution is that it should be backward compatible with deployed equipment, facilitating new equipment to be rolled out incrementally without the need for a flag day across the whole network. This section examines how this has been achieved with the unified MPLS-TP linear protection solution, and calls up the evolution of APS as an example of how problems can arise.

### APC and PSC

Fundamental to how MPLS-TP linear protection manages to be a unified solution is the way that an implementer can upgrade an existing implementation so that it can support both APC and PSC. Similarly, an operator can introduce APC into a PSC network without breaking anything and continuing to use PSC functionality across the whole network or on the LSPs where one of the end points does not support APC. The main issue to be resolved is that early implementations and deployments of MPLS-TP linear protection are limited to PSC as defined in [6] and need to be able to interoperate with full implementations as defined in [3] and [4].

APC and PSC are defined as operational modes in MPLS-TP linear protection. A mode is a set of capabilities to perform specific functions and to operate in particular ways as indicated in the Flags field of the Capabilities TLV as shown in Fig. 1. For the PSC mode, the Flags value is set to 0x0, and for the APS mode, the Flags value is set to 0xF800000 as shown in Fig. 1. When two implementations use PSC messages to communicate they include the Capabilities TLV that announces the capabilities that they support. Capabilities TLV with other Flags values than 0xF800000 or 0x0 are treated as an error.

MPLS-TP linear protection can only operate if both ends of an LSP announce support for the same mode. Nodes can be configured to support one mode or the other, and this configuration may be per node, per interface, or even per LSP.

A legacy node implemented according to [6] would send no Capabilities TLV since it would be unaware of the new Capabilities TLV; this behavior is taken to mean PSC mode. To facilitate backward compatibility between a legacy and a new end-point, a new node that has the ability to send and process the Capabilities TLV must be able to both send the PSC mode Capabilities TLV and send no Capabilities TLV at all.

### APS-2007 and Its Upgrade Strategy

Both APC and APS enhance the behavior of the previous version of ITU-T G.8131 (2007), APS-2007, by adding support for the MS-W, SD, and EXER functions, but the APS described in [7] is based on the same protocol and operation principles as the APS-2007. It is common perception that new vision of a protocol would be interoperable with the previous implementation under the limited use of old features. For the network operators who have already deployed the APS-2007, the APS might seem to be beneficial as network upgrade can be done gradually. One good example is ITU-T G.8032 Ethernet Ring Protection [9], where both old and new version nodes can reside on the same ring and provide protection switching with limited functionalities. However, as for the revisions of the APS based linear protection recommendation,1 the gradual upgrade strategy has not been considered seriously. It is well-recognized among the ITU-T linear protection experts that both end nodes in the protected domain should have the same implementation.

One of the examples that two different versions of the APS protocol lead to a problematic issue to be resolved is that early implementations and deployments of MPLS-TP linear protection are limited to PSC as defined in [6] and need to be able to interoperate with full implementations as defined in [3] and [4].

1 ITU-T G.8131 (2007) is consistent with the behaviour of G.8031 (2006) - Ethernet linear protection switching, and the APS mentioned in this article is also consistent with the behaviour defined in G.8031 (2011).

### Table 1: The values of the fields that vary according to protection switching operation.

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Request</td>
<td>4</td>
<td>Wait-to-Restore (WTR)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Exercise (EXER)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Reverse Request (RR)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Do-not-Revert (DNR)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>No Request (NR)</td>
</tr>
<tr>
<td>Fault Path</td>
<td>1</td>
<td>Indicates that the working path is identified to be in a fault condition or is affected by an administrative command</td>
</tr>
<tr>
<td></td>
<td>2–255</td>
<td>For future extensions and ignored upon receipt</td>
</tr>
<tr>
<td>Data Path</td>
<td>0</td>
<td>Indicates that the protection transport entity is not transporting user data traffic (in 1:1 architecture) or transporting redundant user data traffic (in 1:1 under SD on Protection condition or in 1+1 architecture)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Indicates that the protection path is transmitting user traffic replacing the use of the working path</td>
</tr>
<tr>
<td></td>
<td>2–255</td>
<td>For future extensions and ignored upon receipt</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>Indicates that the protection path is identified to be in a fault condition or is affected by an administrative command</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>For future use and ignored upon receipt</td>
</tr>
</tbody>
</table>

---

IEEE Communications Magazine • December 2014
situation is illustrated in Fig. 3. After the signal fail on the working path is recovered, both nodes go to the No Request (NR) state holding traffic on the protection path, which is indicated in NR(1,1) messages. When the NR(1,1) message is received, node A running APS-2007 switches traffic to the working path and indicates the No Request state with traffic on the working path in NR(0,0) messages. In the meantime, node Z running the APS-2007 starts the WTR timer to see if the clearance of the defect is persistent before reverting traffic to the working path. This results in the traffic discontinuity between two ends during the WTR period which is configured by an operator between five and 12 minutes. As it is obvious that the interoperation between APS-2007 and APS is not possible even for this basic signal fail and recovery scenario, from the perspective of APS-2007 deployment, its upgrade to the APS does not have any real benefits but deviates from the international standard.

**Applicability to Other Topologies**

As long as what to protect is end-to-end traffic, which is in fact the majority in transport networks, a linear protection mechanism can be used regardless of network topologies. In this section we consider the applicability of the APC to other topologies: ring and shared mesh.

**Ring Protection Switching**

The applicability of MPLS-TP linear protection mechanisms to ring topologies is described in [10], and the APC can also be used to provide protection of the traffic that traverses an MPLS-TP ring without any new additional mechanisms or protocol. Considering the processing speeds of the current implementations of linear protection processes and OAM sessions, multiple APC processes required to provide protection over a ring would not be a concern. In particular, for the steering architecture, where an ingress ring node determines the forwarding direction between two ring ports, reusing the existing linear protection would be a reasonable choice compared to ring-specific protocols available in other technologies, such as SDH and OTN.

For the other ring protection architecture, wrapping, which has been used in other technologies and which needs to be supported according to [5], the application of linear protection could be quite troublesome and a ring-specific mechanism can be beneficial. It can also be noted that the application of linear protection is limited to a single ring. Interconnected rings cannot be covered efficiently without additional mechanisms, which are not yet available. In the meantime, the benefits of a ring-specific solution also need to be justified against the costs of developing and deploying the ring-optimized solution.

**Shared Mesh Protection Switching**

In shared mesh protection, the network resources are shared to provide protection for multiple working paths that may not have the same end points. Each working path is protected by a dedicated protection path as in linear protection, but the network resources in a protection path might not be sufficient to simultaneously protect all of the paths for which it offers protection.

One approach can be to define a hop-by-hop restoration mechanism along the protection path. When an end node receives a protection trigger, the end node communicates with each intermediate node along the protection path in a hop-by-hop manner and performs protection switching only after the availability of the shared resources is confirmed by the other end node.

The other approach to achieve shared mesh protection can be to reuse an existing linear protection mechanism for the protection switching action to switch the traffic and define a coordination mechanism to control the use of the shared resources. The additional coordination mechanism focuses on notifying the end points of other working paths not to make any protection switching if the protection resources are insufficient. By separating the traffic switching action from the coordination protocol, which can be done rather more slowly, the protection switching time will be as fast as that in the linear protection. For MPLS-TP networks, the APC protocol can be used as is for the protection switching action for the shared mesh protection.

**Conclusions**

Linear protection switching is an important component of a circuit-switched transport network enabling the delivery of reliable services even in the event of network faults. Its inclusion in the MPLS-TP portfolio is, therefore, critical to the development of packet switching transport networks based on MPLS technology.

The development of widely agreed and open standards for MPLS-TP linear protection allows all equipment vendors to participate in the market on a level playing field. It also means that network operators can purchase equipment that conforms to a well-known international standard
so that it has a high likelihood of interoperating “out of the box.” Of course, the fact that there is a single standard, published in their respective ways by both the IETF (the home of MPLS) and the ITU-T (the home of transport networks) makes this situation considerably simpler compared with, for example, the case of OAM in MPLS-TP networks. Furthermore, the manner in which designers have unified the two desired behaviors within MPLS-TP linear protection to be options within a single protocol specification means that operators are free to choose how to run their networks without having to make detailed technical decisions at the time of purchase.

Finally we observe that the evolutionary approach adopted in the latest MPLS-TP linear protection specifications means that the roll-out of the APC function can proceed incrementally in networks that already deploy PSC, and that operators can make their own deployment choices to support PSC, APC, or both according to their own preferences.

ACKNOWLEDGEMENT

This work was supported by the ICT R&D program of MSIP/IITP (Project 10041414, Terabit optical-circuit-packet converged switching system technology development for the next-generation optical transport network).

REFERENCES


BIOGRAPHIES

Jeong-Don Rho (rhyo@etri.re.kr) is a Principal Researcher at the Electronics and Telecommunications Research Institute (ETRI), South Korea. He holds B.S., M.S. and Ph. D. degrees in electrical engineering from Yong-In National University, South Korea. After completing his Ph. D. study in the area of telecommunication networks and optimization, he started working for Bell Labs, Lucent Technologies, New Jersey, in 1999. While he worked for Bell Labs, he was mainly involved in performance analysis/evaluation/enhancement study for various wireless and wired networks systems. Since he joined ETRI in 2004 his work has been focused on next generation network, carrier class Ethernet and packet transport network technology research, especially participating in OAM and standardization activities in ITU-T. He is the co-editor of the G.8131 (MPLS-TP linear protection) and G.8132 (MPLS-TP ring protection) recommendations and a vice-chairman of ITU-T Study Group 15. He co-authored TCP/IP Essentials: A Lab-Based Approach (Cambridge University Press, 2004). He is a member of Eta Kappa Nu.

Taek Cheong (ctacket@etri.re.kr) is a principal researcher at the Electronics and Telecommunications Research Institute (ETRI), South Korea. He holds B.S., M.S. and Ph. D. degrees in engineering from Yonsei University, South Korea. After completing his Ph. D. study in the area of high-speed circuit design in 2000, he started working for ETRI, where he was engaged in system hardware design such as flow-based QoS switches, Carrier Ethernet switches and packet-optical integrated transport systems. Since 2005 he has participated in ITU-T Q0/15 and contributed to standardization of protection mechanisms for packet transport networks. Since 2010 he has participated in the IETF MPLS WG and contributed to MPLS-TP standardization especially in the area of survivability. His current work focuses on the standardization of MPLS-TP shared mesh protection in the IETF MPLS WG and new protection functionalities such as multi-point Ethernet connection protection and multi-domain segment network protection in ITU-T Q0/15.

Daniel King (dking@lancaster.ac.uk) is a senior consultant at Old Dog Consulting and is currently studying for his Ph.D. at Lancaster University, where he is researching Network Functions Virtualisation (NFV). He has 15 years of experience working within market leading technology companies. He co-founded Aria Networks with Adrian Farrell and held key roles at Marconi, Movaz Networks, Redback Networks, Cisco Systems, and Bell Labs. Daniel is an active contributor within the IETF, specifically within the PCE, MPLS, LISP/VPN, and CCAMP working groups, and is an editor or author of numerous IETF Internet-Drafts and RFCs related to path computation, MPLS and network optimization. Daniel is the secretary of two IETF working groups, namely PCE and LISP/VPN.

Adrian Farrell (adriangoldod.co.uk) currently works as one of two routing area directors in the Internet Engineering Task Force (IETF). His responsibilities include the MPLS, CCAMP, LISP/VNP, and PCE working groups. He is currently funded in this role by Juniper Networks. Adrian has been heavily involved with the IETF for a number of years and is the author of over 50 RFCs. He was among the leaders in the development of some key Internet technologies including GMPLS and PCE. He also runs a successful consultancy company, Old Dog Consulting, providing advice on implementation, deployment, and standardization of Internet Protocol-based solutions, especially in the arena of routing, MPLS, and GMPLS. Adrian is the author or editor of five books on Internet protocols including The Internet and Its Protocols: A Comparative Approach (Morgan-Kaufmann, 2004), GMPLS Architecture and Applications (Morgan-Kaufmann, 2005), and MPLS: Next Steps (Morgan Kaufmann, 2008).

Husei Van Helvoort (hvbpan-vanhelvoort.eu) is senior networking consultant for Hai Gaoming BV. He received his MSEE at Eindhoven University of Technology in 1979. He has been a senior member IEEE since 2003. He worked as a designer and architect of SDH, OTN, and PTN equipment at Philips Telecom, AT&T, Lucent, TradSwitch, and since 2005 for Huawei. He is an expert in functional modeling, fault management, performance monitoring, diagnostics and network management. Since 1998 he has been an active participant in transport network standardization: ITU-T, ETSI, ANSI, IETF. He is co-editor of several SDH, OTN, and PTN ITU-T Recommendations for equipment specification, protection,
performance monitoring and OAM, and co-editor of MPLS-TP related RFCs. Since 2007 he has been ITU-T Study Group15 rapporteur of expert group Q18: “Interfaces, Interworking, OAM and Equipment specifications for Packet based Transport Networks.” He is the author of The ComSoc Guide to Next Generation Optical Transport: SDH/SONET/OTN (Wiley/IEEE-Press 2009), Modeling the Optical Transport Network (Wiley 2005), and Next Generation SDH/SONET: Evolution or Revolution? (Wiley 2005), and he is co-author of Optical Transport Networks from TDM to Packet (ITU-T 2011) and Optical Networking Standards (Springer Verlag 2006). See also www.van-helvoort.eu.