MPLS Working Group
Internet-Draft
Updates: 6378 (if approved)
Intended status: Standards Track
Expires: July 23, 2014

J. Ryoo, Ed.
ETRI
E. Gray, Ed.
Ericsson
H. van Helvoort
Huawei Technologies
A. D'Alessandro
Telecom Italia
T. Cheung
ETRI
E. Osborne
Cisco Systems, Inc.
January 19, 2014

MPLS Transport Profile (MPLS-TP) Linear Protection to Match the Operational Expectations of SDH, OTN and Ethernet Transport Network Operators

draft-ietf-mpls-tp-psc-itu-01.txt

Abstract

This document describes alternate mechanisms to perform some of the sub-functions of MPLS Transport Profile (MPLS-TP) linear protection defined in RFC 6378, and also defines additional mechanisms. The purpose of these alternate and additional mechanisms is to provide operator control and experience that more closely models the behavior of linear protection seen in other transport networks.

This document also introduces capabilities and modes for linear protection. A capability is an individual behavior, and a mode is a particular combination of capabilities. Two modes are defined in this document: Protection State Coordination (PSC) mode and Automatic Protection Switching (APS) mode.

This document describes the behavior of the PSC protocol including priority logic and state machine when all the capabilities associated with the APS mode are enabled.

This document updates RFC 6378 in that the capability advertisement method defined here is an addition to that document.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Copyright Notice

Table of Contents

1. Introduction

Linear protection mechanisms for the MPLS Transport Profile (MPLS-TP) are described in RFC 6378 [RFC6378] to meet the requirements described in RFC 5654 [RFC5654].

This document describes alternate mechanisms to perform some of the sub-functions of linear protection, and also defines additional mechanisms. The purpose of these alternate and additional mechanisms is to provide operator control and experience that more closely models the behavior of linear protection seen in other transport

networks, such as Synchronous Digital Hierarchy (SDH), Optical Transport Network (OTN) and Ethernet transport networks. Linear protection for SDH, OTN, and Ethernet transport networks are defined in ITU-T Recommendations G.841 [G841], G.873.1 [G873.1] and G.8031 [G8031], respectively.

The reader of this document is assumed to be familiar with RFC 6378.

The alternative mechanisms described in this document are for the following capabilities:

1. Priority modification,

2. non-revertive behavior modification,

and the following capabilities have been added to define additional mechanisms:

3. support of Manual Switch to Working path (MS-W) command,

4. support of protection against Signal Degrade (SD), and

5. support of Exercise (EXER) command.

Priority modification includes priority swapping between Signal Fail on Protection path (SF-P) and Forced Switch (FS), and raising the priority level of Clear Signal Fail (SFc).

Non-revertive behavior is modified to align with the behavior defined in RFC 4427 [RFC4427] as well as to follow the behavior of linear protection seen in other transport networks.

Support of MS-W command to revert traffic to the working path in non-revertive operation is covered in this document.

Support of protection switching protocol against SD is covered in this document. The specifics for the method of identifying SD is out of the scope of this document similarly to Signal Fail (SF) for RFC 6378.

Support of EXER command to test if the Protection State Coordination (PSC) communication is operating correctly is also covered in this document. More specifically, EXER command tests and validates the linear protection mechanism and PSC protocol including the aliveness of the priority logic, the PSC state machine and the PSC message generation and reception, and the integrity of the protection path, without triggering the actual traffic switching.

This document introduces capabilities and modes. A capability is an individual behavior. The capabilities of a node are advertised using the method given in this document. A mode is a particular combination of capabilities. Two modes are defined in this document: PSC mode and Automatic Protection Switching (APS) mode.

This document describes the behavior of the PSC protocol including the priority logic and the state machine when all the capabilities associated with the APS mode are enabled.

This document updates RFC 6378 in that the capability advertisement method defined here is an addition to that document. For an existing implementation of RFC 6378, it is recommended to be updated with the bug-fixes in [I-D.ietf-mpls-psc-updates] and the capability advertisement in this document.

2. 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 [RFC2119].

3. Acronyms

This document uses the following acronyms:

APS	Automatic Protection Switching
DNR	Do-not-Revert
EXER	Exercise
FS	Forced Switch
LER	Label Edge Router
LO	Lockout of protection
MS	Manual Switch
MS-P	Manual Switch to Protection path
MS-W	Manual Switch to Working path
MPLS-TP	MPLS Transport Profile
NR	No Request
OC	Operator Clear
OTN	Optical Transport Network
PSC	Protection State Coordination
RR	Reverse Request
SD	Signal Degrade
SDH	Synchronous Digital Hierarchy
SD-P	Signal Degrade on Protection path
SD-W	Signal Degrade on Working path
SF	Signal Fail
SFc	Clear Signal Fail
SFDc	Clear Signal Fail or Degrade
SF-P	Signal Fail on Protection path
SF-W	Signal Fail on Working path
WTR	Wait to Restore

4. Capability 1: Priority modification

In this document, the priorities of FS and SF-P are swapped and the priority of Clear SF (SFc) is raised. In addition to the priority modification, this document introduces the use of Freeze command in Appendix C. The reasons for these changes are explained in the following sub-sections from technical and network operational aspects.

4.1. Motivations for swapping priorities of FS and SF-P

Defining the priority of FS higher than that of SF-P can result in a situation where the protected traffic is taken out-of-service. Setting the priority of any input that is supposed to be signaled to the other end to be higher than that of SF-P can result in unpredictable protection switching state, when the protection path has failed and consequently the PSC communication stopped. An example of the out-of-service scenarios is shown in Appendix A.

According to Section 2.4 of RFC 5654 [RFC5654] it MUST be possible to operate an MPLS-TP network without using a control plane. This means that external switch commands, e.g., FS, can be transferred to the remote Label Edge Router (LER) only by using the PSC communication channel and should not rely on the presence of a control plane.

As the priority of SF-P has been higher than FS in other transport networks, such as SDH, OTN and Ethernet transport networks, for network operators it is important that the MPLS-TP protection switching preserves the network operation behavior to which network operators have become accustomed. Typically, FS command is issued before network maintenance jobs, (e.g., replacing optical cables or other network components). When an operator pulls out a cable on the protection path by mistake, the traffic should be protected and the operator expects this behavior based on his/her experience on the traditional transport network operations.

4.2. Motivation for raising the priority of SFc

The priority level of SFc defined in RFC 6378 [RFC6378] can cause traffic disruption when a node that has experienced local signal fails on both the working and the protection paths is recovering from these failures.

An example of sequence diagram showing the problem with the priority level of SFc as defined in RFC 6378 is shown in Appendix B.

4.3. Motivation for introducing Freeze command

With the priority swapping between FS and SF-P, the traffic is always moved back to the working path when SF-P occurs in Protecting administrative state. In the case that network operators need an option to control their networks so that the traffic can remain on the protection path even when the PSC communication channel is broken, the Freeze command, which is a local command (i.e., not signaled to the other end) can be used. The use of the Freeze command is described in Appendix C.

4.4. Modifications to RFC 6378

The list of local requests in order of priority SHALL be modified as follows:

(from higher to lower)

- Clear Signal Fail
- Signal Fail on Protection path

- Forced Switch
- Signal Fail on Working path

The change of the PSC Control logic including the state machine due to this priority modification is incorporated in the PSC Control logic description in Section 10 and Section 11 when all the capabilities are enabled.

5. Capability 2: Modification of non-revertive operation

Non-revertive mode of protection switching is defined in RFC 4427 [RFC4427]. In this mode, the traffic does not return to the working path when switch-over requests are terminated.

However, PSC protocol defined in RFC 6378 [RFC6378] supports this operation only when recovering from a defect condition, but does not operate as non-revertive when an operator's switch-over command such as FS or Manual Switch (MS) is cleared. To be aligned with legacy transport network behavior and RFC 4427, a node should go into the Do-not-Revert (DNR) state not only when a failure condition on the working path is cleared but also when an operator command requesting switch-over is cleared.

The change of the PSC Control logic including the state machine due to the modification of non-revertive operation is incorporated into the PSC Control logic description in Section 10 and Section 11 when all the capabilities are enabled.

6. Capability 3: Support of MS-W command

6.1. Motivation for adding MS-W

Changing the non-revertive operation introduces necessity of a new operator command to revert traffic to the working path when in the DNR state. When the traffic is on the protection path in the DNR state, a Manual Switch to Working (MS-W) command is issued to switch the normal traffic back to working path. According to Section 4.3.3.6 (Do-not-Revert State) in RFC 6378 [RFC6378], "to revert back to Normal state, the administrator SHALL issue a Lockout of protection (LO) command followed by a Clear command." However, using LO command introduces the potential risk of an unprotected situation while the LO is in effect.

Manual Switch-over for recovery LSP/span command, defined in RFC 4427 [RFC4427] and also defined in RFC 5654 [RFC5654], Requirement 83, as one of the mandatory external commands, should be used for this purpose, but is not included in RFC 6378. Note that the "Manual

Switch-over for recovery LSP/span" command is the same as MS-W command.

6.2. Terms modified to support MS-W

The term "Manual Switch" and its acronym "MS" used in RFC 6378 are replaced respectively by "Manual Switch to Protection path" and "MS-P" by this document to avoid confusion with "Manual Switch to Working path" and its acronym "MS-W".

Also, the term "Protecting administrative state" used in RFC 6378 is replaced by "Switching administrative state" by this document to include the case where traffic is switched back to the working path by administrative MS-W command.

6.3. Behavior of MS-P and MS-W

The MS-P and MS-W commands SHALL have the same priority. If one of these commands is already issued and accepted, then the other command that is issued afterwards SHALL be ignored. If two LERs are requesting opposite operations simultaneously, i.e. one LER is sending MS-P while the other LER is sending MS-W, the MS-W SHALL be considered to have a higher priority than MS-P, and MS-P SHALL be ignored and cancelled.

Two commands, MS-P and MS-W are represented by the same Request Field value, but differentiated by the FPath value. When traffic is switched to the protection path, the FPath field SHALL indicate that the working path is being blocked (i.e., FPath set to 1), and the Path field SHALL indicate that user data traffic is being transported on the protection path (i.e., Path set to 1). When traffic is switched to the working path, the FPath field SHALL indicate that the protection path is being blocked (i.e., FPath set to 0), and the Path field SHALL indicate that user data traffic is being transported on the working path (i.e., Path set to 0).

6.4. Equal priority resolution for MS

RFC 6378 defines only one rule for equal priority condition in Section 4.3.2 as "The remote message from the remote LER is assigned a priority just below the similar local input." In order to support the manual switch behavior described in Section 6.3, additional rules for equal priority resolution are required. Since the support of protection against signal degrade also requires a similar equal priority resolution, the rules are described in Section 7.4.

The change of the PSC Control logic including the state machine due to the support of MS-W command is incorporated into the PSC Control

logic description in Section 10 and Section 11 when all the capabilities are enabled

7. Capability 4: Support of protection against SD

7.1. Motivation for supporting protection against SD

In MPLS-TP survivability framework [RFC6372], fault conditions include both SF and SD that can be used to trigger protection switching.

RFC 6378 [RFC6378], which defines the protection switching protocol for MPLS-TP does not specify how the SF and SD are detected, and specifies the protection switching protocol associated with SF only.

The PSC protocol associated with SD is covered in this document, and the specifics for the method of identifying SD is out of the scope of the protection protocol similar to the facts that how SF is detect and how MS and FS commands are initiated in a management system and signaled to protection switching are out of its scope.

7.2. Terms modified to support SD

Instead of SFc, Clear Signal Fail or Degrade (SFDc) is used to indicate the clearance of either a degraded condition or a failure condition.

The second paragraph of Section 4.3.3.2 Unavailable state in RFC 6378 shows the intention of including Signal Degrade on Protection path (SD-P) in the Unavailable state. Even though the protection path can be partially available under the condition of SD-P, this document follows the same state grouping as RFC 6378 for SD-P.

The bullet item "Protecting failure state" in Section 3.6 in RFC 6378 includes the degraded condition in Protecting failure state. This document follows the same state grouping as RFC 6378 for Signal Degrade on Working path (SD-W).

7.3. Behavior of protection against SD

In order to maintain the network operation behavior to which transport network operators have become accustomed, the priorities of SD-P and SD-W are defined to be equal as in other transport networks, such as SDH, OTN and Ethernet transport networks. Once a switch has been completed due to SD on one path, it will not be overridden by SD on the other path (first come, first served behavior), to avoid protection switching that cannot improve signal quality.

SD indicates that the transmitting end point has identified a degradation of the signal, or integrity of the packet transmission on either the working path or the protection path. The FPath field SHALL identify the path that is reporting the degrade condition (i.e., if the protection path, then FPath is set to 0; if the working path, then FPath is set to 1), and the Path field SHALL indicate where the data traffic is being transported (i.e., if the working path is selected, then Path is set to 0; if the protection path is selected, then Path is set to 1).

The Wait to Restore (WTR) timer is used when the protected domain is configured for revertive behavior and started at the node that recovers from a local degraded condition on the working path.

Protection switching against SD is always provided by a selector bridge duplicating user data traffic and feeding it to both the working path and the protection path under SD condition. When a local or remote SD occurs on either the working path or the protection path, the LER SHALL duplicate user data traffic and SHALL feed to both the working path and the protection path. The packet duplication SHALL continue as long as any SD condition exists in the protected domain, and SHALL stop when there is no SD condition. Additionally, the packet duplication SHALL continue in the WTR state in revertive mode. In non-revertive mode, the packet duplication SHALL stop when there is no SD condition.

The selector bridge with the packet duplication under SD condition, which is a non-permanent bridge, is considered to be a 1:1 protection architecture.

7.4. Equal priority resolution

In order to support the manual switch behavior described in Section 6.3 and the protection against Signal Degrade described in Section 7.3, the rules to resolve the equal priority requests are required.

For the equal priority local inputs, such as MS and SD, first-come, first-served rule is applied. Once a local input is determined as the highest priority local input, then a subsequent equal priority local input requesting a different action, i.e., the action results in the same PSC Request Field but different FPath value, will not be presented to the PSC Control logic as the highest local request. Furthermore, in the case of MS command, the subsequent local MS command requesting a different action will be cancelled.

If the LER is in a remote state due to a remote SD (or MS) message, a subsequent local input having the same priority but requesting

different action to the PSC Control logic, will be considered as having lower priority than the remote message, and will be ignored. If the LER is in remote Switching administrative state due to a remote MS-P, then subsequent local MS-W SHALL be ignored and automatically cancelled. If the LER is in remote Unavailable state due to a remote SD-P, then subsequent local SD-W input will be ignored. However, the local SD-W SHALL appear in the Local Request logic as long as the SD condition exists, but SHALL NOT be the top priority global request, which determines the state transition at the PSC Control logic.

There is a case where one LER receives a local input and the other LER receives, simultaneously, a local input with the same priority but requesting different action. In this case, each of the two LERs receives a subsequent remote message having the same priority but requesting different action, while the LER is in a local state due to the local input. When this case happens, a priority must be set for the inputs with the same priority regardless of its origin (local input or remote message).

When MS-W and MS-P occur simultaneously at both LERs, MS-W SHALL be considered as having higher priority than MS-P at both LERs.

When SD-W and SD-P occur simultaneously at both LERs, the SD on the standby path (the path from which the selector does not select the user data traffic) is considered as having higher priority than the SD on the active path (the path from which the selector selects the user data traffic) regardless of its origin (local or remote message). Therefore, no unnecessary protection switching is performed and the user data traffic continues to be selected from the active path.

In the preceding paragraphs, the "simultaneously" refers to the case a sent SD (or MS) request has not been confirmed by the remote end in bidirectional protection switching. When a local node that has transmitted a SD message receives a SD (or MS) message that indicates a different value of data path (Path) field than the value of the Path field in the transmitted SD (or MS) message, both the local and the remote SD requests are considered to occur simultaneously.

The change of the PSC Control logic including the state machine due to the support of protection against SD is incorporated into the PSC Control logic description in Section 10 and Section 11 when all the capabilities are enabled.

8. Capability 5: Support of EXER command

EXER is a command to test if the PSC communication is operating correctly. More specifically, EXER is to test and validate the linear protection mechanism and PSC protocol including the aliveness of the Local Request logic, the PSC state machine and the PSC message generation and reception, and the integrity of the protection path, without triggering the actual traffic switching. It is used while the working path is either carrying the traffic or not. It has lower priority than any "real" switch request. It is only valid in bidirectional switching, since this is the only place where one can get a meaningful test by looking for a response.

This command is documented in R84 of RFC 5654 [RFC5654].

A received EXER message indicates that the remote end point is operating under an operator command to validate the protection mechanism and PSC protocol including the aliveness of the Local Request logic, the PSC state machine and the PSC message generation and reception, and the integrity of the protection path, without triggering the actual traffic switching. The valid response to EXER message is an Reverse Request (RR) with the corresponding FPath and Path numbers. The local LER SHALL signal a RR only in response to an EXER command from the remote LER.

When Exercise commands are input at both ends, an EXER, instead of RR, SHALL be transmitted from both ends.

The following PSC Requests SHALL be added to PSC Request field to support Exercise:

(3) Exercise - indicates that the transmitting end point is exercising the protection channel and mechanism. FPath and Path are set to the same value of the No Request (NR), RR or DNR request that EXER replaces.

(2) Reverse Request - indicates that the transmitting end point is responding to an EXER command from the remote LER. FPath and Path are set to the same value of the NR or DNR request that RR replaces.

The priority of Exercise SHALL be inserted between the priorities of WTR Expires and No Request.

9. Capabilities and modes

9.1. Capabilities

A Capability is an individual behavior whose use is signaled in a Capabilities TLV, which is placed in Optional TLVs field inside the PSC message shown in Figure 2 of RFC 6378 [RFC6378]. The format of the Capabilities TLV is:

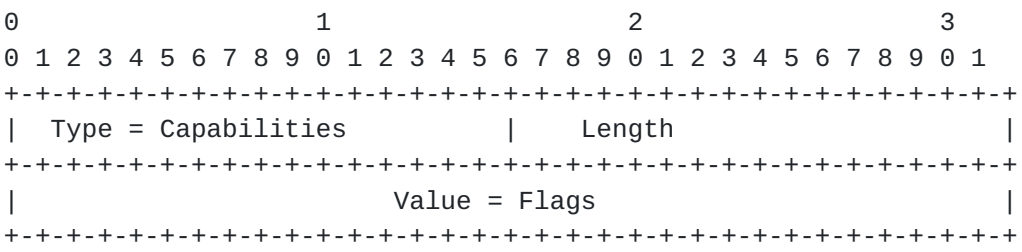


Figure 1: Format of Capabilities TLV

The value of the Type field is TBD pending IANA allocation.

The value of the Length field is the length of the Flags field in octets. The length of the Flags field MUST be a multiple of 4 octets and MUST be the minimum required to signal all the required capabilities.

Section 4 to Section 8 discuss five capabilities that are signaled using the five most significant bits; if a node wishes to signal these five capabilities, it MUST send a Flags field of 4 octets. A node would send a Flags field greater than 4 octets only if it had more than 32 Capabilities to indicate. All unused bits MUST be set to zero.

If the bit assigned for an individual capability is set to 1, it indicates the sending node's intent to use that capability in the protected domain. If a bit is set to 0, the sending node does not intend to use the indicated capability in the protected domain. Note that it is not possible to distinguish between the intent not to use a capability and a node's complete non-support (i.e., lack of implementation) of a given capability.

This document defines five specific capabilities that are described from Section 4 to Section 8. Each capability is assigned bit as follows:

0x80000000: priority modification

0x40000000: non-revertive behavior modification

0x20000000: support of MS-W command

0x10000000: support of protection against SD

0x08000000: support of EXER command

If all the five capabilities should be used, an LER SHALL set 0xF8000000 in the Flags field.

9.1.1. Sending the Capabilities TLV

PSC sends messages in response to external events and in periodic retransmission of current status. It may be expensive to send and to parse an Capabilities TLV attached to a packet intended to trigger a protection switch or other real-time behavior. However, if a node does not periodically send its Capabilities TLV, the receiving node cannot discriminate a deliberate omission of the Capabilities TLV for performance reasons from an accidental omission due to an implementation issue. To guard against this, a node MUST include its Capabilities TLV in every PSC message that it sends.

9.1.2. Receiving the Capabilities TLV

A node MUST establish a receive timer for the Capabilities TLV. By default this MUST be 3.5 times the periodic retransmission timer of five seconds - i.e., 17.5 seconds. Both the periodic retransmission time and the timeout SHOULD be configurable by the operator. When a node receives a Capabilities TLV it resets the timer to 17.5 seconds. If the timer expires, the node behaves as in Section 9.1.3.

When a node receives a Capabilities TLV it MUST compare it to its most recent transmitted Capabilities TLV. If the two are equal, the protected domain is said to be running in the mode indicated by that set of capabilities (see Section 9.2). If the sent and received Capabilities TLVs are not equal, this indicates a capabilities mismatch. When this happens, the node MUST alert the operator and MUST behave as in Section 9.1.3.

9.1.3. Handling Capabilities TLV errors

This section covers the two possible errors - a TLV timeout and a TLV mismatch - and the error handling procedures in both cases.

9.1.3.1. Capabilities TLV Timeout

If the Capabilities TLV receive timer expires and there is no defect on the protection path, the node MUST alert the operator and MUST behave as in Section 9.1.3.3.

9.1.3.2. Capabilities TLV Mismatch

If the sent and received Capabilities TLVs are not equal, this indicates a capabilities mismatch. When this happens, the node MUST alert the operator and MUST behave as in Section 9.1.3.3. A node MAY retain the received TLV for logging, alert or debug purposes.

9.1.3.3. Handling Capabilities TLV error conditions

When a node enters in Capabilities protocol error conditions, the following actions MUST be taken:

1. Indicate the error condition (e.g., either mismatch or timeout) to the operator by the usual alert mechanisms (e.g., syslog).

2. Not make any state transitions based on the contents of any PSC messages

To expand on point 2 - assume node A is receiving NR(0,0) from its PSC peer node Z and is also receiving a mismatched set of capabilities (e.g., received 0x20000000, transmitted 0xA0000000). If node Z detects a local SF-W and wants to initiate a protection switch (that is, by sending SF(1,1)), node A MUST NOT react to this input by changing its state. A node MAY increase the severity or urgency of its alarms to the operator, but until the operator resolves the mismatch in the Capabilities TLV the protected domain will likely operate in an inconsistent state.

9.2. Modes

A Mode is a given set of Capabilities. Modes are shorthand; referring to a set of capabilities by their individual values or by the name of their mode does not change the protocol behavior. This document defines two modes - PSC and APS.

9.2.1. PSC Mode

PSC Mode is defined as the lack of any Capabilities - that is, a Capabilities set of 0x0. It is the behavior specified in RFC 6378. There are two ways to declare PSC Mode. A node can send a Capabilities TLV of 0x0, or it can send no Capabilities TLV at all. This is further explored in Section 9.3.

9.2.2. APS Mode

APS Mode is defined as the use of all the five specific capabilities, which are described from Section 4 to Section 8 in this document. APS Mode is indicated with the Flags value of 0xF8000000.

9.3. Backward compatibility

As defined in Section 9.2.1, PSC Mode is indicated either with a Capabilities TLV of 0x0 or the lack of Capabilities TLV. This is to allow backward compatibility between two nodes - one which can send the Capabilities TLV, and one which cannot.

RFC 6378 does not define how to handle an unrecognized TLV. There may be some implementations that silently discard an unrecognized TLV, and some that take more drastic steps like refusing to allow PSC to operate. Thus, a node which has the ability to send and receive the PSC Mode Capabilities TLV MUST be able to both send the PSC Mode Capabilities TLV and send no Capabilities TLV at all. An implementation MUST be configurable between these two choices.

One question that arises from this dual definition of PSC Mode is, what happens if a node which was sending a non-null Capabilities TLV (e.g., APS Mode) sends PSC packets without any Capabilities TLV? This case is handled as follows:

If a node has never, during the life of a PSC session, received a Capabilities TLV from its peer, the lack of a Capabilities TLV is treated as receipt of a PSC Capabilities TLV. This allows for interoperability between nodes which support the PSC Mode TLV and nodes which do not, and are thus implicitly operating in PSC Mode.

If a node has received a non-null Capabilities TLV (e.g., APS Mode) during the life of a PSC session and then receives a PSC packet with no Capabilities TLV, the receiving node MUST treat the lack of Capabilities TLV as simply a lack of refresh. That is, the receipt of a PSC packet with no Capabilities TLV simply does not reset the receive timer defined in Section 9.1.2.

10. PSC Protocol in APS Mode

This section and Section 11 define the behavior of PSC protocol when all of the aforementioned capabilities are enabled, i.e., APS mode.

10.1. Request field in PSC protocol message

The values of "Request" field in PSC protocol message, which is shown in Figure 2 of RFC 6378 [RFC6378], are redefined as follows:

(14) Lockout of protection

(12) Forced Switch

(10) Signal Fail

(7) Signal Degrade

(5) Manual Switch

(4) Wait-to-Restore

(3) Exercise

(2) Reverse Request

(1) Do-not-Revert

(0) No Request

10.2. Priorities of local inputs and remote requests

Based on the description in Section 3 and Section 4.3.2 in RFC 6378, the priorities of multiple outstanding local inputs are evaluated in the Local Request logic, where the highest priority local input (highest local request) is determined. This highest local request is passed to the PSC Control logic, that will determine the higher priority input (top priority global request) between the highest local request and the last received remote message. When a remote message comes to the PSC Control logic, the top priority global request is determined between this remote message and the highest local request which is present. The top priority global request is used to determine the state transition, which is described in Section 11.

The priorities for both local and remote requests are defined as follows from highest to lowest:

- Operator Clear (Local only)
- Lockout of protection (Local and Remote)
- Clear Signal Fail or Degrade (Local only)
- Signal Fail on Protection path (Local and Remote)
- Forced Switch (Local and Remote)
- Signal Fail on Working path (Local and Remote)
- Signal Degrade on either Protection path or Working path (Local and Remote)

- Manual Switch to either Protection path or Working path (Local and Remote)
- WTR Expires (Local only)
- WTR (Remote only)
- Exercise (Local and Remote)
- Reverse Request (Remote only)
- Do-Not-Revert (Remote only)
- No Request (Remote and Local)

Note that the "Local only" requests are not signaled to the remote LER. Likewise, the "Remote only" requests do not exist in the Local Request logic as local inputs. For example, the priority of WTR only applies to the received WTR message, which is generated from the remote LER. The remote LER that is running the WTR timer in the WTR state has no local request.

The remote request from the remote LER is assigned a priority just below the same local request. However, for the equal priority requests, such as SD and MS, the following equal priority resolution rules are defined:

- If two local inputs having the same priority but requesting different action come to the Local Request logic, then the input coming first SHALL be considered to have a higher priority than the other coming later (first-come, first-served).
- If the PSC Control logic has both the highest local request and a remote message with the same priority and requesting the same action, i.e., the same PSC Request Field and the same FPath value, then the local input SHALL be considered to have a higher priority than the remote message.
- If the PSC Control logic has both the highest local request and a remote message with the same priority but requesting different action and the remote message exists when the highest local request comes to the PSC Control logic, the highest local request is ignored and the remote Request SHALL be the top priority global request.
- If the PSC Control logic has both the highest local request and a remote message with the same priority but requesting different action and the highest local request exists when the remote

message comes to the PSC Control logic, the top priority global request SHALL be determined by the following rules for each simultaneous condition:

- For simultaneous MS requests, the MS-W request SHALL be considered to have a higher priority than the MS-P request. The LER that has local MS-W request SHALL maintain the local MS-W request as the top priority global request, but the other LER that has local MS-P request SHALL clear the MS-P command and internally generate "Operator Clear" request.

- For simultaneous SD requests, the SD on the standby path (the path from which the selector does not select the user data traffic) SHALL be considered as having higher priority than the SD on the active path (the path from which the selector selects the user data traffic) regardless of its origin (local or remote message). The LER that has the SD on the standby path SHALL maintain the local SD on the standby path request as the top priority global request. The other LER that has local SD on the active path SHALL use the remote SD on the standby path as the top priority global request to lookup the state transition table. The differentiation of the active and standby paths is based upon which path had been used for the user data traffic at the time just before an LER selected its local SD as the top priority global request.

No Request is another exception to the rule of assigning a remote request a priority just below the same local request. Since a received NR message needs to be used in the state transition table lookup when there is no outstanding local request, the received remote NR request SHALL be the top priority global request when there is no request in the local LER.

10.3. Acceptance and retention of local inputs

A local input indicating a defect, such as SF-P, SF-W, SD-P and SD-W, SHALL be accepted and retained persistently in the Local Request logic as long as the defect condition exists. If there is any higher priority local input than the local defect input, the higher priority local input is passed to the PSC Control logic as the highest local request, but the local defect input cannot be removed but remains in the Local Request logic. When the higher priority local input disappears, the local defect will become the highest local request if the defect condition still exists.

Operator Clear command, SFDc and WTR Expires are not persistent. Once they appear to the Local Request logic and complete the operation, they SHALL be disappeared.

Operator LO, FS, MS, and EXER commands SHALL be rejected if there is any higher priority local input in the Local Request logic. If a new operator command is accepted, any previous lower-priority local operator command SHALL be cancelled. When any higher priority remote request is received, a lower-priority local operator command SHALL be cancelled. The cancelled operator command is forgotten and will never return, unless the operator reissues the command.

11. State Transition Tables in APS Mode

When there is a change in the highest local request or in remote PSC messages, the top priority global request SHALL be evaluated and the state transition tables SHALL be looked up in the PSC Control logic. The following rules are applied to the operation related to the state transition table lookup.

- If the top priority global request, which determines the state transition, is the highest local request, the local state transition table in Section 11.1 SHALL be used to decide the next state of the LER. Otherwise, remote messages state transition table in Section 11.2 SHALL be used.

- If in remote state, the highest local defect condition (SF-P, SF-W, SD-P or SD-W) SHALL always be reflected in the Request Field and Fpath.

- For the LER currently in the local state, if the top priority global request is changed to OC or SFDc causing the next state to be Normal, WTR or DNR, then all the local and remote requests should be re-evaluated as if the LER is in the state specified in the footnotes to the state transition tables, before deciding the final state. This re-evaluation is an internal operation confined within the local LER, and PSC messages are generated according to the final state.

- The WTR timer is started only when the LER which has recovered from a local failure/degradation enters the WTR state. An LER which is entering into the WTR state due to a remote WTR message does not start the WTR timer. The WTR timer is stopped when any local or remote request triggers the state change out of the WTR state.

The extended states, as they appear in the table, are as follows:

N	Normal state
UA:LO:L	Unavailable state due to local LO command
UA:P:L	Unavailable state due to local SF-P
UA:DP:L	Unavailable state due to local SD-P
UA:LO:R	Unavailable state due to remote LO message
UA:P:R	Unavailable state due to remote SF-P message
UA:DP:R	Unavailable state due to remote SD-P message
PF:W:L	Protecting failure state due to local SF-W
PF:DW:L	Protecting failure state due to local SD-W
PF:W:R	Protecting failure state due to remote SF-W message
PF:DW:R	Protecting failure state due to remote SD-W message
SA:F:L	Switching administrative state due to local FS command
SA:MW:L	Switching administrative state due to local MS-W command
SA:MP:L	Switching administrative state due to local MS-P command
SA:F:R	Switching administrative state due to remote FS message
SA:MW:R	Switching administrative state due to remote MS-W message
SA:MP:R	Switching administrative state due to remote MS-P message
E::L	Exercise state due to local EXER command
E::R	Exercise state due to remote EXER message
WTR	Wait-to-Restore state
DNR	Do-not-Revert state

Each state corresponds to the transmission of a particular set of Request, FPath and Path bits. The table below lists the message that is generally sent in each particular state. If the message to be sent in a particular state deviates from the table below, it is noted in the footnotes to the state transition tables.

State	REQ(FP,P)
N	NR(0,0)
UA:LO:L	LO(0,0)
UA:P:L	SF(0,0)
UA:DP:L	SD(0,0)
UA:LO:R	highest local request(local FPath,0)
UA:P:R	highest local request(local FPath,0)
UA:DP:R	highest local request(local FPath,0)
PF:W:L	SF(1,1)
PF:DW:L	SD(1,1)
PF:W:R	highest local request(local FPath,1)
PF:DW:R	highest local request(local FPath,1)
SA:F:L	FS(1,1)
SA:MW:L	MS(0,0)
SA:MP:L	MS(1,1)
SA:F:R	highest local request(local FPath,1)
SA:MW:R	NR(0,0)
SA:MP:R	NR(0,1)
WTR	WTR(0,1)
DNR	DNR(0,1)
E::L	EXER(0,x), where x is the existing Path value when Exercise command is issued.
E::R	RR(0,x), where x is the existing Path value when RR message is generated.

Some operation examples of the APS mode are shown in Appendix D.

11.1. State transition by local inputs

	OC	LO	SFDc	SF-P	FS	SF-W
N	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
UA:LO:L	(1)	i	i	i	i	i
UA:P:L	i	UA:LO:L	(1)	i	i	i
UA:DP:L	i	UA:LO:L	(1)	UA:P:L	SA:F:L	PF:W:L
UA:LO:R	i	UA:LO:L	i	UA:P:L	i	PF:W:L
UA:P:R	i	UA:LO:L	i	UA:P:L	i	PF:W:L
UA:DP:R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
PF:W:L	i	UA:LO:L	(2)	UA:P:L	SA:F:L	i
PF:DW:L	i	UA:LO:L	(2)	UA:P:L	SA:F:L	PF:W:L
PF:W:R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
PF:DW:R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
SA:F:L	(3)	UA:LO:L	i	UA:P:L	i	i
SA:MW:L	(1)	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
SA:MP:L	(3)	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
SA:F:R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
SA:MW:R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
SA:MP:R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
WTR	(4)	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
DNR	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
E::L	(5)	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L
E::R	i	UA:LO:L	i	UA:P:L	SA:F:L	PF:W:L

	SD-P	SD-W	MS-W	MS-P	WTRExp	EXER
N	UA:DP:L	PF:DW:L	SA:MW:L	SA:MP:L	i	E::L
UA:LO:L	i	i	i	i	i	i
UA:P:L	i	i	i	i	i	i
UA:DP:L	i	i	i	i	i	i
UA:LO:R	UA:DP:L	PF:DW:L	i	i	i	i
UA:P:R	UA:DP:L	PF:DW:L	i	i	i	i
UA:DP:R	UA:DP:L	PF:DW:L	i	i	i	i
PF:W:L	i	i	i	i	i	i
PF:DW:L	i	i	i	i	i	i
PF:W:R	UA:DP:L	PF:DW:L	i	i	i	i
PF:DW:R	UA:DP:L	PF:DW:L	i	i	i	i
SA:F:L	i	i	i	i	i	i
SA:MW:L	UA:DP:L	PF:DW:L	i	i	i	i
SA:MP:L	UA:DP:L	PF:DW:L	i	i	i	i
SA:F:R	UA:DP:L	PF:DW:L	i	i	i	i
SA:MW:R	UA:DP:L	PF:DW:L	SA:MW:L	i	i	i
SA:MP:R	UA:DP:L	PF:DW:L	i	SA:MP:L	i	i
WTR	UA:DP:L	PF:DW:L	SA:MW:L	SA:MP:L	(6)	i
DNR	UA:DP:L	PF:DW:L	SA:MW:L	SA:MP:L	i	E::L
E::L	UA:DP:L	PF:DW:L	SA:MW:L	SA:MP:L	i	i
E::R	UA:DP:L	PF:DW:L	SA:MW:L	SA:MP:L	i	E::L

NOTES:

(1) Re-evaluate to determine final state as if the LER is in the Normal state.

(2) In the case that both local input after SFDc and the last received remote message are no requests, the LER enters into the WTR state when the domain is configured for revertive behavior, or the LER enters into the DNR state when the domain is configured for non-revertive behavior. In all the other cases, re-evaluate to determine the final state as if the LER is in the Normal state.

(3) Re-evaluate to determine final state as if the LER is in the Normal state when the domain is configured for revertive behavior, or as if the LER is in the DNR state when the domain is configured for non-revertive behavior,

(4) Remain in WTR and send NR(0,1). Stop the WTR timer if it is running.

(5) If Path value is 0, re-evaluate to determine final state as if the LER is in the Normal state. If Path value is 1, re-evaluate to determine final state as if the LER is in the DNR state.

(6) Remain in WTR and send NR(0,1).

11.2. State transition by remote messages

	LO	SF-P	FS	SF-W	SD-P	SD-W
N	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
UA:LO:L	i	i	i	i	i	i
UA:P:L	UA:LO:R	i	i	i	i	i
UA:DP:L	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	i	(7)
UA:LO:R	i	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
UA:P:R	UA:LO:R	i	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
UA:DP:R	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	i	PF:DW:R
PF:W:L	UA:LO:R	UA:P:R	SA:F:R	i	i	i
PF:DW:L	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	(8)	i
PF:W:R	UA:LO:R	UA:P:R	SA:F:R	i	UA:DP:R	PF:DW:R
PF:DW:R	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	i
SA:F:L	UA:LO:R	UA:P:R	i	i	i	i
SA:MW:L	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
SA:MP:L	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
SA:F:R	UA:LO:R	UA:P:R	i	PF:W:R	UA:DP:R	PF:DW:R
SA:MW:R	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
SA:MP:R	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
WTR	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
DNR	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
E::L	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R
E::R	UA:LO:R	UA:P:R	SA:F:R	PF:W:R	UA:DP:R	PF:DW:R

	MS-W	MS-P	WTR	EXER	RR	DNR	NR
N	SA:MW:R	SA:MP:R	i	E::R	i	i	i
UA:LO:L	i	i	i	i	i	i	i
UA:P:L	i	i	i	i	i	i	i
UA:DP:L	i	i	i	i	i	i	i
UA:LO:R	SA:MW:R	SA:MP:R	i	E::R	i	i	N
UA:P:R	SA:MW:R	SA:MP:R	i	E::R	i	i	N
UA:DP:R	SA:MW:R	SA:MP:R	i	E::R	i	i	N
PF:W:L	i	i	i	i	i	i	i
PF:DW:L	i	i	i	i	i	i	i
PF:W:R	SA:MW:R	SA:MP:R	(9)	E::R	i	(10)	(11)
PF:DW:R	SA:MW:R	SA:MP:R	(9)	E::R	i	(10)	(11)
SA:F:L	i	i	i	i	i	i	i
SA:MW:L	i	i	i	i	i	i	i
SA:MP:L	i	i	i	i	i	i	i
SA:F:R	SA:MW:R	SA:MP:R	i	E::R	i	DNR	N
SA:MW:R	i	SA:MP:R	i	E::R	i	i	N
SA:MP:R	SA:MW:R	i	i	E::R	i	DNR	N
WTR	SA:MW:R	SA:MP:R	i	i	i	i	(12)
DNR	SA:MW:R	SA:MP:R	i	E::R	i	i	i
E::L	SA:MW:R	SA:MP:R	(13)	i	i	i	i
E::R	SA:MW:R	SA:MP:R	i	i	i	DNR	N

NOTES:

(7) If the received SD-W message has Path=0, ignore the message. If the received SD-W message has Path=1, go to PF:DW:R state and transmit SD(0,1)

(8) If the received SD-P message has Path=1, ignore the message. If the received SD-P message has Path=0, go to UA:DP:R state and transmit SD(1,0).

(9) Transition to WTR state and continue to send the current message.

(10) Transition to DNR state and continue to send the current message.

(11) If the received NR message has Path=1, transition to WTR if domain configured for revertive behavior, else transition to DNR. If the received NR message has Path=0, transition to N.

(12) If the receiving LER's WTR timer is running, maintain current state and message. If the WTR timer is not running, transition to N.

(13) Transit to WTR state and send NR(0,1) message. The WTR timer is not initiated.

11.3. State transition for 1+1 unidirectional protection

The state transition tables given in Section 11.1 and Section 11.2 are for bidirectional protection switching, where remote PSC protocol messages are used to determine the protection switching actions. The 1+1 unidirectional protection switching does not require the remote information in PSC protocol message and acts upon local inputs only. The state transition by local inputs in Section 11.1 SHALL be reused for the 1+1 unidirectional protection under the following conditions:

- The value of Request field in the received remote message is ignored and always assumed to be no request.
- Replace footnote (4) with "Stop the WTR timer and transit to Normal state."
- Replace footnote (6) with "Transit to Normal state."
- Exercise is not applicable.

12. Provisioning mismatch and protocol failure in the APS mode

The remote PSC message that is received from the remote LER is subject to the detection of provisioning mismatch and protocol failure conditions. In the APS mode, provisioning mismatches are handled as follows:

- If the PSC message is received from the working path due to working/protection path configuration mismatch, the node MUST alert the operator and MUST NOT perform any protection switching.

- If the "Protection Type (PT)" field mismatches and two sides are unable to converge as described in Section 5.1 in [I-D.ietf-mpls-psc-updates], the node MUST alert the operator and MUST NOT perform any protection switching.

- If the "Revertive (R)" bit mismatches, two sides will interwork and traffic is protected in the APS mode. The node MAY notify the operator of this event.

- If the Capabilities TLV mismatches, the node MUST alert the operator and MUST NOT perform any protection switching.

The followings are the protocol failure situations and the actions to be taken:

- No match in sent "Data Path (Path)" and received "Data Path (Path)" for more than 50 ms: The node MAY continue to perform protection switching and SHOULD notify the operator of these events:

- No PSC message is received on the protection path during at least 3.5 times the long PSC message interval (e.g. at least 17.5 seconds) and there is no defect on the protection path (The Capabilities TLV Timeout error specifies in Section 9.1.3 is included in this situation.): The node MUST alert the operator and MUST NOT perform any protection switching.

13. Security considerations

No specific security issue is raised in addition to those ones already documented in RFC 6378 [RFC6378]

14. IANA considerations

14.1. MPLS PSC Request Registry

In the "Multiprotocol Label Switching (MPLS) Operations, Administration, and Management (OAM) Parameters" registry, IANA maintains the "MPLS PSC Request Registry".

IANA is requested to assign two new code points from this registry. The values shall be allocated as follows:

Value	Description	Reference
2	Reverse Request	(this document)
3	Exercise	(this document)

14.2. MPLS PSC TLV Registry

In the "Multiprotocol Label Switching (MPLS) Operations, Administration, and Management (OAM) Parameters" registry, IANA maintains the "MPLS PSC TLV Registry".

This document defines a new value for the Capabilities TLV type in the "MPLS PSC TLV Registry".

Value	Description	Reference
TBD	Capabilities	(this document)

14.3. MPLS PSC Capability Flag Registry

IANA is requested to create and maintain a new registry within the "Multiprotocol Label Switching (MPLS) Operations, Administration, and Management (OAM) Parameters" registry called "MPLS PSC Capability Flag Registry". All flags within this registry SHALL be allocated according to the "Standards Action" procedures as specified in RFC 5226 [RFC5226].

The length of the flags MUST be a multiple of 4 octets. This document defines 4 octet flags. Flags greater than 4 octets SHALL be used only if more than 32 Capabilities need to be defined. Flags defined in this document are:

Bit	Hex Value	Capability	Reference
0	0x80000000	priority modification	(this document)
1	0x40000000	non-revertive behavior modification	(this document)
2	0x20000000	support of MS-W command	(this document)
3	0x10000000	support of protection against SD	(this document)
4	0x08000000	support of EXER command	(this document)
5-31		Unassigned	(this document)

15. Acknowledgements

16. References

16.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC5654] Niven-Jenkins, B., Brungard, D., Betts, M., Sprecher, N., and S. Ueno, "Requirements of an MPLS Transport Profile", RFC 5654, September 2009.

[RFC6378] Weingarten, Y., Bryant, S., Osborne, E., Sprecher, N., and A. Fulignoli, "MPLS Transport Profile (MPLS-TP) Linear Protection", RFC 6378, October 2011.

[I-D.ietf-mpls-psc-updates] Osborne, E., "Updates to PSC", draft-ietf-mpls-psc-updates-00 (work in progress), October 2013.

16.2. Informative References

[RFC4427] Mannie, E. and D. Papadimitriou, "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4427, March 2006.

[RFC6372] Sprecher, N. and A. Farrel, "MPLS Transport Profile (MPLS-TP) Survivability Framework", RFC 6372, September 2011.

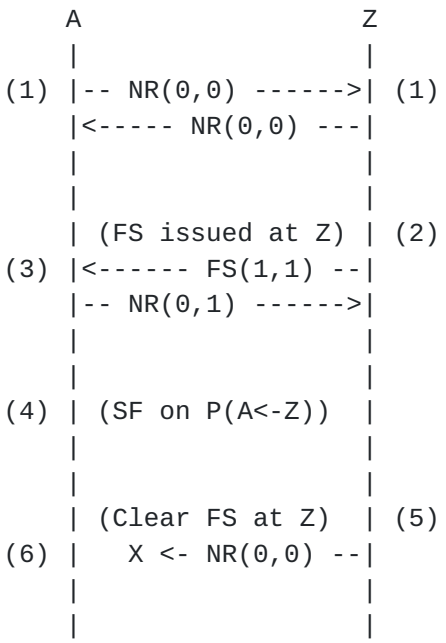
[G841] International Telecommunications Union, "Types and characteristics of SDH network protection architectures", ITU-T Recommendation G.841, October 1998.

[G873.1] International Telecommunications Union, "Optical Transport Network (OTN): Linear protection", ITU-T Recommendation G.873.1, July 2011.

[G8031] International Telecommunications Union, "Ethernet Linear Protection Switching", ITU-T Recommendation G.8031/Y.1342, June 2011.

Appendix A. An example of out-of-service scenarios

The sequence diagram shown is an example of the out-of-service scenarios based on the priority level defined in RFC 6378. The first PSC message which differs from the previous PSC message is shown.



(1) Each end is in the Normal state, and transmits NR(0,0) messages.

(2) When a FS command is issued at node Z, node Z goes into local Protecting administrative state (PA:F:L) and begins transmission of an FS(1,1) messages.

(3) A remote FS message causes node A to go into remote Protecting administrative state (PA:F:R), and node A begins transmitting NR(0,1) messages.

(4) When node A detects a unidirectional SF-P, node A keeps sending NR(0,1) message because SF-P is ignored under the PA:F:R state.

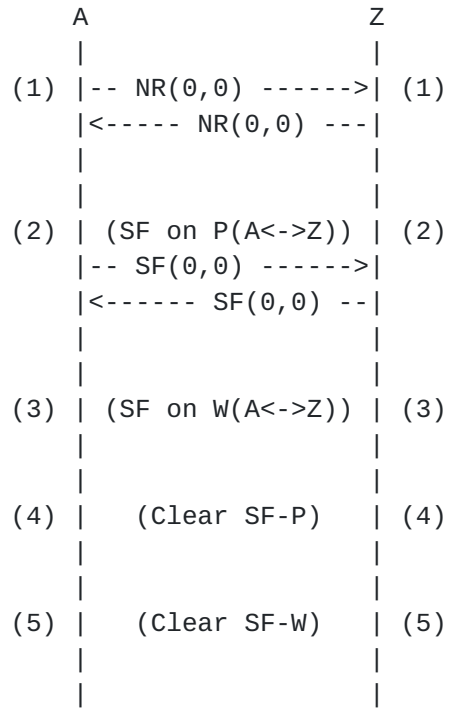
(5) When a Clear command is issued at node Z, node Z goes into the Normal state and begins transmission of NR(0,0) messages.

(6) But, node A cannot receive PSC message because of local unidirectional SF-P. Because no valid PSC message is received, over a period of several successive message intervals, the last valid received message remains applicable and the node A continue to transmit an NR(0,1) message in the PA:F:R state.

Now, there exists a mismatch between the bridge/selector positions of node A (transmitting an NR(0,1)) and node Z (transmitting an NR(0,0)). It results in out-of-service even when there is neither SF-W nor FS.

Appendix B. An example of sequence diagram showing the problem with the priority level of SFc

An example of sequence diagram showing the problem with the priority level of SFc defined in RFC 6378 is given below. The following sequence diagram is depicted for the case of bidirectional signal fails. However, other cases with unidirectional signal fails can result in the same problem. The first PSC message which differs from the previous PSC message is shown.



(1) Each end is in the Normal state, and transmits NR(0,0) messages.

(2) When SF-P occurs, each node enters into the UA:P:L state and transmits SF(0,0) messages. Traffic remains on the working path.

(3) When SF-W occurs, each node remains in the UA:P:L state as SF-W has a lower priority than SF-P. Traffic is still on the working path. Traffic cannot be delivered as both the working path and the protection path are experiencing signal fails.

(4) When SF-P is cleared, local "Clear SF-P" request cannot be presented to the PSC Control logic, which takes the highest local request and runs PSC state machine, since the priority of "Clear SF-P" is lower than that of SF-W. Consequently, there is no change in state, and the selector and/or bridge keep pointing at the working path, which has signal fail condition.

Now, traffic cannot be delivered while the protection path is recovered and available. It should be noted that the same problem will occur in the case that the sequence of SF-P and SF-W events is changed.

If we further continue with this sequence to see what will happen after SF-W is cleared,

(5) When SF-W is cleared, local "Clear SF-W" request can be passed to the PSC Control logic as there is no higher priority local input, but this will be ignored in the PSC Control logic according to the state transition definition in RFC 6378. There will be no change in state or protocol message transmitted.

As SF-W is now cleared and the selector and/or bridge are still pointing at the working path, traffic delivery is resumed. However, each node is the in UA:P:L state and transmitting SF(0,0) message, while there exists no outstanding request for protection switching. Moreover, any future legitimate protection switching requests, such as SF-W, will be rejected as each node thinks the protection path is unavailable.

Appendix C. Freeze Command

The "Freeze" command applies only to the local LER of the protection group and is not signaled to the remote LER. This command freezes the state of the protection group. Until the Freeze is cleared, additional local commands are rejected and condition changes and received PSC information are ignored.

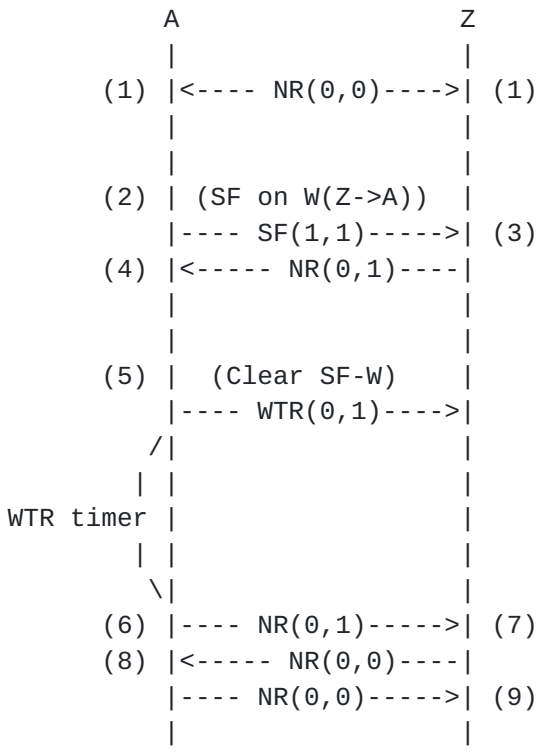
"Clear Freeze" command clears the local freeze. When the Freeze command is cleared, the state of the protection group is recomputed based on the persistent condition of the local triggers.

Because the freeze is local, if the freeze is issued at one end only, a failure of protocol can occur as the other end is open to accept any operator command or a fault condition.

Appendix D. Operation examples of the APS mode

The sequence diagrams shown in this section are only a few examples of the APS mode operations. The first PSC protocol message which differs from the previous message is shown. The operation of hold-off timer is omitted. The Request, FPath and Path fields, whose values are changed during PSC message exchange are shown. For an example, SF(1, 0) represents an PSC message with the following field values: Request = SF, FPath = 1, and Path = 1. The values of the other fields remain unchanged from the initial configuration. W(A->Z) and P(A->Z) indicate the working path and the protection path in the direction of A to Z, respectively.

Example 1. 1:1 bidirectional protection switching (revertive mode) - Unidirectional SF case



(1) The protection domain is operating without any defect, and the working path is used for delivering the traffic in the Normal state.

(2) SF-W occurs in the Z to A direction. Node A enters into the PF:W:L state and generates SF(1, 1) message. Selector and bridge of node A are pointing at the protection path.

(3) Upon receiving SF(1, 1), node Z sets selector and bridge to the protection path. As there is no local request in node Z, node Z generates NR(0, 1) message in the PF:W:R state.

(4) Node A confirms that the remote LER is also selecting protection path.

(5) Node A detects clearing of SF condition, starts the WTR timer, and sends WTR(0, 1) message in the WTR state.

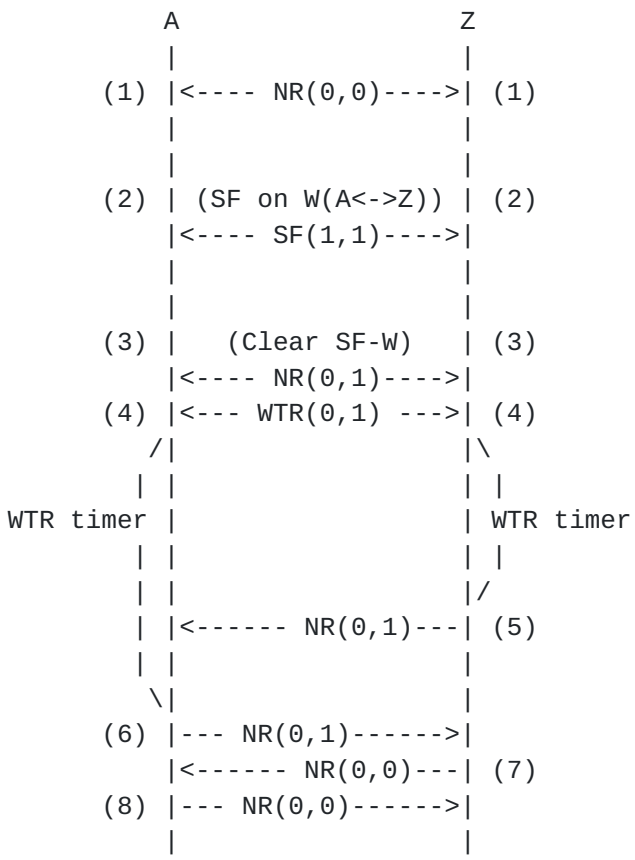
(6) At expiration of the WTR timer, node A sets selector and bridge to the working path and sends NR(0, 1) message.

(7) Node Z is notified that the remote request has been cleared. Node Z transits to the Normal state and sends NR(0,0) message.

(8) Upon receiving NR(0,0) message, node A transits to the Normal state and sends NR(0,0) message.

(9) It is confirmed that the remote LER is also selecting the working path.

Example 2. 1:1 bidirectional protection switching (revertive mode) - Bidirectional SF case - Inconsistent WTR timers



(1) Each end is in the Normal state, and transmits NR(0,0) messages.

(2) When SF-W occurs, each node enters into the PF:W:L state and transmits SF(1,1) messages. Traffic is switched to the protection path. Upon receiving SF(1,1), each node confirms that the remote LER is also sending and receiving the traffic from the protection path.

(3) When SF-W is cleared, each node transits to the PF:W:R state and transmits NR(0,1) messages as the last received message is SF-W.

(4) Upon receiving NR(0,1) messages, each node goes into the WTR state, starts the WTR timer, and sends the WTR(0,1) messages.

(5) At expiration of the WTR timer in node Z, node Z sends NR(0,1) as the last received APS message was WTR. When NR(0,1) arrives at node A, node A maintains the WTR state and keeps sending current WTR messages as described in the state transition table.

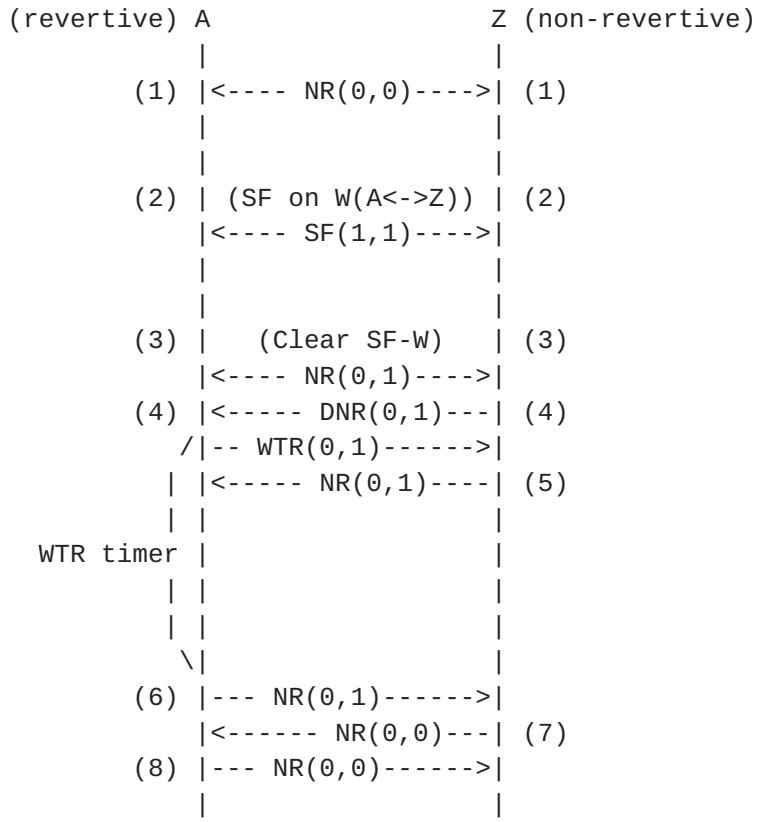
(6) At expiration of the WTR timer in node A, node A sends NR(0,1).

(7) When the NR(0,1) message arrives at node Z, node Z moves to the Normal state, sets selector and bridge to the working path, and sends NR(0, 0) message.

(8) The received NR(0,0) message causes node A to go to the Normal state. Now, the traffic is switched back to the working path.

Example 3. 1:1 bidirectional protection switching - R bit mismatch

This example shows that both sides will interwork and the traffic is protected when one side (node A) is configured as revertive mode and the other (node Z) is configured as non-revertive mode. The interworking is covered in the state transition tables.



(1) Each end is in the Normal state, and transmits NR(0,0) messages.

(2) When SF-W occurs, each node enters into the PF:W:L state and transmits SF(l,l) messages. Traffic is switched to the protection path. Upon receiving SF(1,1), each node confirms that the remote LER is also sending and receiving the traffic on the protection path.

(3) When SF-W is cleared, each node transits to the PF:W:R state and transmits NR(0,1) messages as the last received message is SF-W.

(4) Upon receiving NR(0,1) messages, node A goes into the WTR state, starts the WTR timer, and sends WTR(0,1) messages. At the same time, node B transits to the DNR state and sends DNR(0,1) message.

(5) When the WTR message arrives at node Z, node Z transits to the WTR state and send NR(0,1) message according to the state transition table. At the same time, the DNR message arrived at node Z is ignored according to the state transition table. Therefore, node Z, which is configured as non-revertive mode, is operating as if in revertive mode.

(6) At expiration of the WTR timer in node A, node A sends NR(0,1).

(7) When the NR(0,1) message arrives at node Z, node Z moves to the Normal state, sets selector and bridge to the working path, and sends NR(0, 0) message.

(8) The received NR(0,0) message causes node A to transits to the Normal state. Now, the traffic is switched back to the working path.

Authors' Addresses

Jeong-dong Ryoo (editor)
ETRI
218 Gajeongno
Yuseong-gu, Daejeon 305-700
South Korea

Phone: +82-42-860-5384
Email: ryoo@etri.re.kr

Eric Gray (editor)
Ericsson

Email: eric.gray@ericsson.com

Huub van Helvoort
Huawei Technologies
Karspeldreef 4,
Amsterdam 1101 CJ
the Netherlands

Phone: +31 20 4300936
Email: huub.van.helvoort@huawei.com

Alessandro D'Alessandro
Telecom Italia
via Reiss Romoli, 274
Torino 10148
Italy

Phone: +39 011 2285887
Email: alessandro.dalessandro@telecomitalia.it

Taesik Cheung
ETRI
218 Gajeongno
Yuseong-gu, Daejeon 305-700
South Korea

Phone: +82-42-860-5646
Email: cts@etri.re.kr

Eric Osborne
Cisco Systems, Inc.

Email: eosborne@cisco.com