       MPLS Transport Profile (MPLS-TP) Linear Protection to Match the
          Operational Expectations of Synchronous Digital Hierarchy,
       Optical Transport Network, and Ethernet Transport Network Operators

Abstract

   This document describes alternate mechanisms to perform some of the
   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 is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 5741.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc7271.

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
   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 is 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 [RFC6378].

   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 the Manual Switch to Working path (MS-W) command,

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

   5. support of the Exercise (EXER) command.

   The priority modification includes raising the priority of Signal
   Fail on Protection path (SF-P) relative to Forced Switch (FS), and
   raising the priority level of Clear Signal Fail (SFc) above SF-P.

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

   Support of the protection-switching protocol against SD is covered in
   this document. The specifics for the method of identifying SD are
   out of the scope for this document and are treated similarly to
   Signal Fail (SF) in [RFC6378].

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

   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.

   Other modes may be defined as new combinations of the capabilities
   defined in this document or through the definition of additional
   capabilities. In either case, the specification defining a new mode
   will be responsible for documenting the behavior, the priority logic,
   and the state machine of the PSC protocol when the set of
   capabilities in the new mode is enabled.

   This document describes the behavior, the priority logic, and the
   state machine of the PSC protocol when all the capabilities
   associated with the APS mode are enabled. The PSC protocol behavior
   for the PSC mode is as defined in [RFC6378].

   This document updates [RFC6378] by adding a capability advertisement
   mechanism. It is recommended that existing implementations of the
   PSC protocol be updated to support this capability. Backward
   compatibility with existing implementations that do not support this
   mechanism is described in Section 9.2.1.

   Implementations are expected to be configured to support a specific
   set of capabilities (a mode) and to reject messages that indicate the
   use of a different set of capabilities (a different mode). Thus, the
   capability advertisement is not a negotiation but a verification that
   peers are using the same mode.

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 RFC 2119 [RFC2119].

3. Acronyms

   This document uses the following acronyms:

   APS   Automatic Protection Switching
   DNR   Do-not-Revert
   EXER   Exercise
   FS   Forced Switch
   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
   SD-P   Signal Degrade on Protection path
   SD-W   Signal Degrade on Working path
   SDH   Synchronous Digital Hierarchy
   SF   Signal Fail
   SF-P   Signal Fail on Protection path
   SF-W   Signal Fail on Working path
   SFc   Clear Signal Fail
   SFDc   Clear Signal Fail or Degrade
   WTR   Wait-to-Restore

4. Capability 1: Priority Modification

   [RFC6378] defines the priority of FS to be higher than that of SF-P.
   That document also defines the priority of Clear SF (SFc) to be low.
   This document defines the priority modification capability whereby
   the relative priorities of FS and SF-P are swapped, and the priority
   of Clear SF (SFc) is raised. In addition, this capability introduces
   the Freeze command as described in Appendix C. The rationale for
   these changes is detailed in the following subsections from both the
   technical and network operational aspects.

4.1. Motivation 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. When
   the protection path fails, PSC communication may stop as a result.
   In this case, if any input that is supposed to be signaled to the
   other end has a higher priority than SF-P, then this can result in an

   unpredictable protection-switching state. An example scenario that
   may result in an out-of-service situation is presented in Appendix A
   of this document.

   According to Section 2.4 of [RFC5654], it MUST be possible to operate
   an MPLS-TP network without using a control plane. This means that
   the PSC communication channel is very important for the transfer of
   external switching commands (e.g., FS), and these commands should not
   rely on the presence of a control plane. In consequence, the failure
   of the PSC communication channel has higher priority than FS.

   In other transport networks (such as SDH, OTN, and Ethernet transport
   networks), the priority of SF-P has been higher than that of FS. It
   is therefore important to offer network operators the option of
   having the same behavior in their MPLS-TP networks so that they can
   have the same operational protection-switching behavior to which they
   have become accustomed. Typically, an 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 continue to be
   protected, and the operator expects this behavior based on his/her
   experience with traditional transport network operations.

4.2. Motivation for Raising the Priority of SFc

   The priority level of SFc defined in [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.

   A sequence diagram highlighting the problem with the priority level
   of SFc as defined in [RFC6378] is presented in Appendix B.

4.3. Motivation for Introducing the 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 case 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 can be used. Freeze is defined to be a "local"
   command that is not signaled to the remote node. The use of the
   Freeze command is described in Appendix C.

4.4. Procedures in Support of Priority Modification

   When the modified priority order specified in this document is in
   use, the list of local requests in order of priority SHALL be as
   follows (from highest to lowest):

   o Clear Signal Fail

   o Signal Fail on Protection path

   o Forced Switch

   o Signal Fail on Working path

   This requires modification of the PSC Control Logic (including the
   state machine) relative to that described in [RFC6378]. Sections 10
   and 11 present the PSC Control Logic when all capabilities of APS
   mode are enabled.

5. Capability 2: Non-revertive Behavior Modification

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

   However, the PSC protocol defined in [RFC6378] supports this
   operation only when recovering from a defect condition: it does not
   support the non-revertive function when an operator's switch-over
   command, such as FS or Manual Switch (MS), is cleared. To be aligned
   with the behavior in other transport networks and to be consistent
   with [RFC4427], 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 that requested switch-over is cleared.

   This requires modification to the PSC Control Logic (including the
   state machine) relative to that described in [RFC6378]. Sections 10
   and 11 present the PSC Control Logic when all capabilities of APS
   mode are enabled.

6. Capability 3: Support of the MS-W Command

6.1. Motivation for adding MS-W

   Changing the non-revertive operation as described in Section 5
   introduces the necessity of a new operator command to revert traffic
   to the working path 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 the working

path. According to Section 4.3.3.6 (Do-not-Revert State) in
[RFC6378], "To revert back to the Normal state, the administrator
SHALL issue a Lockout of protection command followed by a Clear
command." However, using the Lockout of protection (LO) command
introduces the potential risk of an unprotected situation while the
LO is in effect.

The "Manual switch-over for recovery LSP/span" command is defined in
[RFC4427]. Requirement 83 in [RFC5654] states that the external
commands defined in [RFC4427] MUST be supported. Since there is no
support for this external command in [RFC6378], this functionality
should be added to PSC. This support is provided by introducing the
MS-W command. The MS-W command, as described here, corresponds to
the "Manual switch-over for recovery LSP/span" command.

6.2. Terminology to Support MS-W

[RFC6378] uses the term "Manual Switch" and its acronym "MS". This
document uses the term "Manual Switch to Protection path" and "MS-P"
to have the same meaning, while avoiding confusion with "Manual
Switch to Working path" and its acronym "MS-W".

Similarly, we modify the name of "Protecting Administrative" state
(as defined in [RFC6378]) to be "Switching Administrative" state to
include the case where traffic is switched to the working path as a
result of the external MS-W command.

6.3. Behavior of MS-P and MS-W

MS-P and MS-W SHALL have the same priority. We consider different
instances of determining the priority of the commands when they are
received either in succession or simultaneously.

o When two commands are received in succession, the command that is
   received after the initial command SHALL be cancelled.

o If two nodes simultaneously receive commands that indicate
   opposite operations (i.e., one node receives MS-P and the other
   node receives MS-W) and transmit the indications to the remote
   node, the MS-W SHALL be considered to have a higher priority, and
   the MS-P SHALL be cancelled and discarded.

Two commands, MS-P and MS-W, are transmitted using the same Request
field value but SHALL indicate in the Fault Path (FPath) value the
path from which the traffic is being diverted. When traffic is
switched to the protection path, the FPath field value SHALL be set
to 1, indicating that traffic is being diverted from the working
path. When traffic is switched to the working path, the FPath field

value SHALL be set to 0, indicating that traffic is being diverted
from the protection path. The Data Path (Path) field SHALL indicate
where user 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).

When an MS command is in effect at a node, any subsequent MS or EXER
command and any other lower-priority requests SHALL be ignored.

6.4. Equal-Priority Resolution for MS

[RFC6378] defines only one rule for the equal-priority condition in
Section 4.3.2 as "The remote message from the far-end 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.

Support of this function requires changes to the PSC Control Logic
(including the state machine) relative to that shown in [RFC6378].
Sections 10 and 11 present the PSC Control Logic when all
capabilities of APS mode are enabled.

7. Capability 4: Support of Protection against SD

7.1. Motivation for Supporting Protection against SD

In the MPLS-TP Survivability Framework [RFC6372], both SF and SD
fault conditions can be used to trigger protection switching.

[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, but
the specifics for the method of identifying SD is out of scope for
the protection protocol in the same way that SF detection and MS or
FS command initiation are out of scope.

7.2. Terminology to Support SD

In this document, the term 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
[RFC6378] 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 [RFC6378] for
SD-P.

The bulleted item on the Protecting Failure state in Section 3.6 of
[RFC6378] includes the degraded condition in the Protecting Failure
state. This document follows the same state grouping as [RFC6378]
for Signal Degrade on Working path (SD-W).

7.3. Behavior of Protection against SD

To better align the behavior of MPLS-TP networks with that of other
transport networks (such as SDH, OTN, and Ethernet transport
networks), we define the following:

o The priorities of SD-P and SD-W SHALL be equal.

o 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.

The SD message indicates that the transmitting node has identified
degradation of the signal or integrity of the packet received on
either the working path or the protection path. The FPath field
SHALL identify the path that is reporting the degraded 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).

When the SD condition is cleared and the protected domain is
recovering from the situation, the Wait-to-Restore (WTR) timer SHALL
be used if the protected domain is configured for revertive behavior.
The WTR timer SHALL be 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 node SHALL duplicate user data traffic and SHALL
feed it 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. When the SD condition is cleared, in revertive
operation, the packet duplication SHALL continue in the WTR state and
SHALL stop when the node leaves the WTR state; while in non-revertive
operation, the packet duplication SHALL stop immediately.

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.

Protection switching against SD does not introduce any modification
to the operation of the selector at the sink node described in
[RFC6378]. The selector chooses either the working or protection
path from which to receive the normal traffic in both 1:1 and 1+1
architectures. The position of the selector, i.e., which path to
receive the traffic, is determined by the PSC protocol in
bidirectional switching or by the local input in unidirectional
switching.

7.4. Equal-Priority Resolution

In order to support the MS behavior described in Section 6.3 and the
protection against SD described in Section 7.3, it is necessary to
expand rules for treating equal-priority inputs.

For equal-priority local inputs, such as MS and SD, apply a simple
first-come, first-served rule. 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 an MS command, the subsequent local MS
command requesting a different action will be cancelled.

If a node is in a remote state due to a remote SD (or MS) message, a
subsequent local input having the same priority but requesting a
different action to the PSC Control Logic will be considered as
having lower priority than the remote message and will be ignored.
For example, if a node is in remote Switching Administrative state
due to a remote MS-P, then any subsequent local MS-W SHALL be ignored
and automatically cancelled. If a node is in remote Unavailable
state due to a remote SD-P, then any subsequent local SD-W input will
be ignored. However, the local SD-W SHALL continue to appear in the
Local Request Logic as long as the SD condition exists, but it SHALL
NOT be the top-priority global request, which determines the state
transition at the PSC Control Logic.

Cases where two end-points of the protected domain simultaneously
receive local triggers of the same priority that request different
actions may occur (for example, one node receives SD-P and the other
receives SD-W). Subsequently, each node will receive a remote
message with the opposing action indication. To address these cases,
we define the following priority resolution rules:

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

o When SD-W and SD-P occur simultaneously at both nodes, 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, "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 an SD message receives an SD (or MS) message that
indicates a different value of Path field from the value of Path
field in the transmitted SD (or MS) message, both the local and
remote SD requests are considered to occur simultaneously.

The addition of support for protection against SD requires
modification to the PSC Control Logic (including the state machine)
relative to that described in [RFC6378]. Sections 10 and 11 present
the PSC Control Logic when all capabilities of APS mode are enabled.

8. Capability 5: Support of EXER Command

The EXER command is used to verify the correct operation of the PSC
communication, such as the aliveness of the Local Request Logic, the
integrity of the PSC Control Logic, the PSC message generation and
reception mechanism, and the integrity of the protection path. EXER
does not trigger any actual traffic switching.

The command is only relevant for bidirectional protection switching,
since it is dependent upon receiving a response from the remote node.
The EXER command is assigned lower priority than any switching
message. It may be used regardless of the traffic usage of the
working path.

When a node receives a remote EXER message, it SHOULD respond with a
Reverse Request (RR) message with the FPath and Path fields set
according to the current condition of the node. The RR message SHALL
be generated only in response to a remote EXER message.

This command is documented in R84 of [RFC5654].

If EXER commands are input at both ends, then a race condition may
arise. This is resolved as follows:

o If a node has issued EXER and receives EXER before receiving RR,
   it MUST treat the received EXER as it would an RR, and it SHOULD
   NOT respond with RR.

The following PSC Requests are added to the PSC Request field to
support the Exercise command (see also Section 14.1):

   (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
   message whose transmission is stopped by EXER.

   (2) Reverse Request - indicates that the transmitting end-point is
   responding to an EXER command from the remote node. FPath and
   Path are set to the same value of the NR or DNR message whose
   transmission is stopped by RR.

The relative priorities of EXER and RR are defined in Section 10.2.

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 [RFC6378]. The format of the
Capabilities TLV is:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type = Capabilities         |   Length                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Value = Flags                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
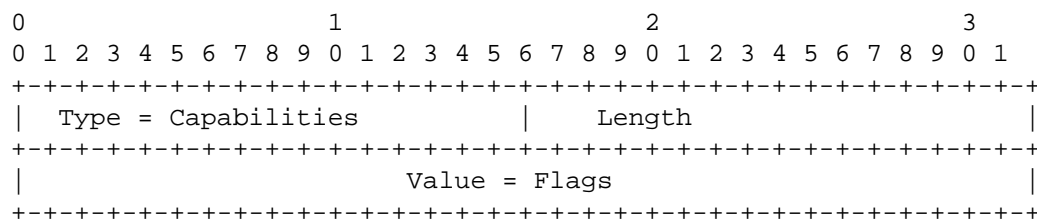
Figure 1: Format of Capabilities TLV

The value of the Type field is 1.

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
in 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, a node SHALL set the
Flags field to 0xF8000000.

9.1.1. Sending and Receiving the Capabilities TLV

A node MUST include its Capabilities TLV in every PSC message that it
transmits. The transmission and acceptance of the PSC message is
described in Section 4.1 of [RFC6378].

When a node receives a Capabilities TLV, it MUST compare the Flags
value to its most recent Flags value transmitted by the node. 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 TLV mismatch. When this happens, the node MUST alert
the operator and MUST NOT perform any protection switching until the
operator resolves the mismatch between the two end-points.

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. Capabilities TLVs with
other combinations than the one specified by a mode are not supported
in this specification.

9.2.1. PSC Mode

PSC mode is defined as the lack of support for any of the additional
capabilities defined in this document -- that is, a Capabilities set
of 0x0. It is the behavior specified in [RFC6378].

There are two ways to declare PSC mode. A node can send no
Capabilities TLV at all since there are no TLV units defined in
[RFC6378], or it can send a Capabilities TLV with Flags value set to
0x0. In order to allow backward compatibility between two end-points
-- one which supports sending the Capabilities TLV, and one which
does not, the node that has the ability to send and process 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 options.

9.2.2. APS Mode

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

10. PSC Protocol in APS Mode

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

10.1. Request Field in PSC Protocol Message

   This document defines two new values for the "Request" field in the
   PSC protocol message that is shown in Figure 2 of [RFC6378] as
   follows:

      (2) Reverse Request

      (3) Exercise

   See also Section 14.1 of this document.

10.2. Priorities of Local Inputs and Remote Requests

   Based on the description in Sections 3 and 4.3.2 in [RFC6378], 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 that is
   present. The top-priority global request is used to determine the
   state transition, which is described in Section 11. In this
   document, in order to simplify the description on the PSC Control
   Logic, we strictly decouple the priority evaluation from the state
   transition table lookup.

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

   o Operator Clear (Local only)

   o Lockout of protection (Local and Remote)

   o Clear Signal Fail or Degrade (Local only)

   o Signal Fail on Protection path (Local and Remote)

   o Forced Switch (Local and Remote)

o Signal Fail on Working path (Local and Remote)

o Signal Degrade on either Protection path or Working path (Local
   and Remote)

o Manual Switch to either Protection path or Working path (Local and
   Remote)

o WTR Timer Expiry (Local only)

o WTR (Remote only)

o Exercise (Local and Remote)

o Reverse Request (Remote only)

o Do-Not-Revert (Remote only)

o No Request (Remote and Local)

Note that the "Local only" requests are not transmitted to the remote
node. 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 node. The remote node that is running the WTR timer in the
WTR state has no local request.

The remote SF and SD on either the working path or the protection
path and the remote MS to either the working path or the protection
path are indicated by the values of the Request and FPath fields in
the PSC message.

The remote request from the remote node is assigned a priority just
below the same local request except for NR and equal-priority
requests, such as SD and MS. Since a received NR message needs to be
used in the state transition table lookup when there is no
outstanding local request, the remote NR request SHALL have a higher
priority than the local NR. For the equal-priority requests, see
Section 10.2.1.

10.2.1. Equal-Priority Requests

As stated in Section 10.2, the remote request from the remote node is
assigned a priority just below the same local request. However, for
equal-priority requests, such as SD and MS, the priority SHALL be
evaluated as described in this section.

For equal-priority local requests, the first-come, first-served rule
SHALL be applied. Once a local request appears in the Local Request
Logic, a subsequent equal-priority local request requesting a
different action, i.e., the action results in the same Request value
but a different FPath value, SHALL be considered to have a lower
priority. Furthermore, in the case of an MS command, the subsequent
local MS command requesting a different action SHALL be rejected and
cleared.

When the priority is evaluated in the PSC Control Logic between the
highest local request and a remote request, the following equal-
priority resolution rules SHALL be applied:

o If two requests request the same action, i.e., the same Request
   and FPath values, then the local request SHALL be considered to
   have a higher priority than the remote request.

o When the highest local request comes to the PSC Control Logic, if
   the remote request that requests a different action exists, then
   the highest local request SHALL be ignored and the remote request
   SHALL remain to be the top-priority global request. In the case
   of an MS command, the local MS command requesting a different
   action SHALL be cancelled.

o When the remote request comes to the PSC Control Logic, if the
   highest local request that requests a different action exists,
   then the top-priority global request SHALL be determined by the
   following rules:

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

   * For SD requests, the SD on the standby path (the path from
      which the selector does not select the user data traffic) SHALL
      be considered to have a 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 node 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 node 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 selected for the user data
        traffic when each node detected its local SD.

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 is cleared, the local defect will become the highest
   local request if the defect condition still exists.

   The Operator Clear (OC) command, SFDc, and WTR Timer Expiry are not
   persistent. Once they appear to the Local Request Logic and complete
   all the operations in the protection-switching control, they SHALL
   disappear.

   The 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
   higher-priority local request (including an 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 cleared. If the operators wish to
   renew the cancelled command, then they should reissue 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.

   o 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 node. Otherwise, the remote state transition table
      in Section 11.2 SHALL be used.

   o 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 and
      FPath fields.

o For the node 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
   SHALL be re-evaluated as if the node is in the state specified in
   the footnotes to the state transition tables, before deciding the
   final state. If there are no active requests, the node enters the
   state specified in the footnotes to the state transition tables.
   This re-evaluation is an internal operation confined within the
   local node, and the PSC messages are generated according to the
   final state.

o The WTR timer is started only when the node that has recovered
   from a local failure or degradation enters the WTR state. A node
   that is entering into the WTR state due to a remote WTR message
   does not start the WTR timer. The WTR timer SHALL be 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
WTR       Wait-to-Restore state
DNR       Do-not-Revert state
E::L      Exercise state due to local EXER command
E::R      Exercise state due to remote EXER message
```

Each state corresponds to the transmission of a particular set of
Request, FPath, and Path fields. 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 of the state transition tables.

```
State     Request(FPath,Path)
-------   ----------------------------------
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 APS mode are shown in Appendix D.

In the state transition tables below, the letter 'i' stands for
"ignore" and is an indication to remain in the current state and
continue transmitting the current PSC message

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  |
```

(Continued)

```
          | 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 the final state as if the node is in
     the Normal state. If there are no active requests, the node
     enters the Normal State.

(2) In the case that both local input after SFDc and the last
     received remote message are NR, the node enters into the WTR
     state when the domain is configured for revertive behavior, or
     the node enters into the DNR state when the domain is configured
     for non-revertive behavior. In all the other cases, where one
     or more active requests exist, re-evaluate to determine the
     final state as if the node is in the Normal state.

(3) Re-evaluate to determine final state as if the node is in the
     Normal state when the domain is configured for revertive
     behavior, or as if the node is in the DNR state when the domain
     is configured for non-revertive behavior. If there are no
     active requests, the node enters either the Normal state when
     the domain is configured for revertive behavior or the DNR state
     when the domain is configured for non-revertive behavior.

(4) Remain in the WTR state and send an NR(0,1) message. Stop the
     WTR timer if it is running. In APS mode, OC can cancel the WTR
     timer and hasten the state transition to the Normal state as in
     other transport networks.

(5) If Path value is 0, re-evaluate to determine final state as if
     the node is in the Normal state. If Path value is 1,
     re-evaluate to determine final state as if the node is in the
     DNR state. If there are no active requests, the node enters the
     Normal state when Path value is 0, or the DNR state when Path
     value is 1.

(6) Remain in the WTR state and send an NR(0,1) message.

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 |
```

(Continued)

```
          | 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 | (13)| E::R | i  | i    | i
  E::L    | SA:MW:R | SA:MP:R | i   | 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 the PF:DW:R state
     and transmit an SD(0,1) message.

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

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

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

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

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

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

   o The value of Request field in the received remote message is
      ignored and always assumed to be no request.

   o Replace footnote (4) with "Stop the WTR timer and transit to the
      Normal state."

   o Replace footnote (6) with "Transit to the Normal state."

   o Exercise command is not relevant.

12. Provisioning Mismatch and Protocol Failure in APS Mode

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

   o 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
      until the operator resolves this path configuration mismatch.

   o In the case that the mismatch happens in the two-bit "Protection
      Type (PT)" field, which indicates permanent/selector bridge type
      and uni/bidirectional switching type:

* If the value of the PT field of one side is 2 (i.e., selector
  bridge) and that of the other side is 1 or 3 (i.e., permanent
  bridge), then this event MUST be notified to the operator and
  each node MUST NOT perform any protection switching until the
  operator resolves this bridge type mismatch.

* If the bridge type matches but the switching type mismatches,
  i.e., one side has PT=1 (unidirectional switching) while the
  other side has PT=2 or 3 (bidirectional switching), then the
  node provisioned for bidirectional switching SHOULD fall back
  to unidirectional switching to allow interworking. The node
  SHOULD notify the operator of this event.

o If the "Revertive (R)" bit mismatches, two sides will interwork
  and traffic is protected according to the state transition
  definition given in Section 11. The node SHOULD notify the
  operator of this event.

o If the Capabilities TLV mismatches, the node MUST alert the
  operator and MUST NOT perform any protection switching until the
  operator resolves the mismatch in the Capabilities TLV.

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

o 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 this event.

o 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 with a default message interval of 5 seconds), and there
  is no defect on the protection path: The node MUST alert the
  operator and MUST NOT perform any protection switching until the
  operator resolves this defect.

13. Security Considerations

This document introduces no new security risks. [RFC6378] points out
that MPLS relies on assumptions about the difficulty of traffic
injection and assumes that the control plane does not have end-to-end
security. [RFC5920] describes MPLS security issues and generic
methods for securing traffic privacy and integrity. MPLS use should
conform to such advice.

14. IANA Considerations

14.1. MPLS PSC Request Registry

   In the "Generic Associated Channel (G-ACh) Parameters" registry, IANA
   maintains the "MPLS PSC Request Registry".

   IANA has assigned the following two new code points from this
   registry.

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

14.2. MPLS PSC TLV Registry

   In the "Generic Associated Channel (G-ACh) Parameters" registry, IANA
   maintains the "MPLS PSC TLV Registry".

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

      Value Description   Reference
      ------ -------------------- --------------
        1   Capabilities   (this document)

14.3. MPLS PSC Capability Flag Registry

   IANA has created and now maintains a new registry within the "Generic
   Associated Channel (G-ACh) 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 each flag 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. The 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

The authors would like to thank Yaacov Weingarten, Yuji Tochio, Malcolm Betts, Ross Callon, Qin Wu, and Xian Zhang for their valuable comments and suggestions on this document.

We would also like to acknowledge explicit text provided by Loa Andersson and Adrian Farrel.

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.

16.2. Informative References

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

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

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

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

[RFC5920]   Fang, L., "Security Framework for MPLS and GMPLS
            Networks", RFC 5920, July 2010.

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

Appendix A. An Example of an Out-of-Service Scenario

   The sequence diagram shown is an example of the out-of-service
   scenarios based on the priority level defined in [RFC6378]. The
   first PSC message that differs from the previous PSC message is
   shown.

```
                       A                 Z
                       |                 |
                   (1) |-- NR(0,0) ------>| (1)
                       |<----- NR(0,0) ---|
                       |                 |
                       |                 |
                       | (FS issued at Z) | (2)
                   (3) |<------ FS(1,1) --|
                       |-- NR(0,1) ------>|
                       |                 |
                       |                 |
                   (4) | (SF on P(A<-Z))  |
                       |                 |
                       |                 |
                       | (Clear FS at Z)  | (5)
                   (6) |    X <- NR(0,0) --|
                       |                 |
                       |                 |
```
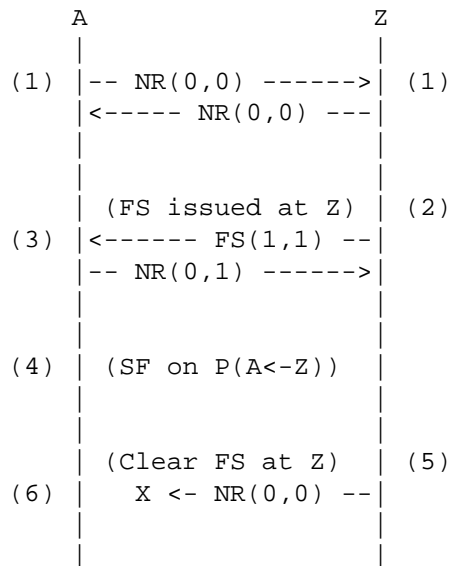
   (1)   Each end is in the Normal state and transmits NR(0,0) messages.

   (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) message.

   (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
         an 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 selector and bridge
   positions of node A (transmitting an NR(0,1) message) and node Z
   (transmitting an NR(0,0) message). It results in an out-of-service
   situation even when there is neither SF-W nor FS.

Appendix B. An Example of a Sequence Diagram Showing the Problem with
             the Priority Level of SFc

   An example of a sequence diagram showing the problem with the
   priority level of SFc defined in [RFC6378] is given below. The
   following sequence diagram depicts the case when the bidirectional
   signal fails. However, other cases with unidirectional signal fails
   can result in the same problem. The first PSC message that differs
   from the previous PSC message is shown.

```
                    A                Z
                    |                |
           (1)  |-- NR(0,0) ------>| (1)
                |<----- NR(0,0) ---|
                    |                |
                    |                |
           (2)  | (SF on P(A<->Z)) | (2)
                |-- SF(0,0) ------>|
                |<------ SF(0,0) --|
                    |                |
                    |                |
           (3)  | (SF on W(A<->Z)) | (3)
                    |                |
                    |                |
           (4)  |    (Clear SF-P)   | (4)
                    |                |
                    |                |
           (5)  |    (Clear SF-W)   | (5)
                    |                |
                    |                |
```
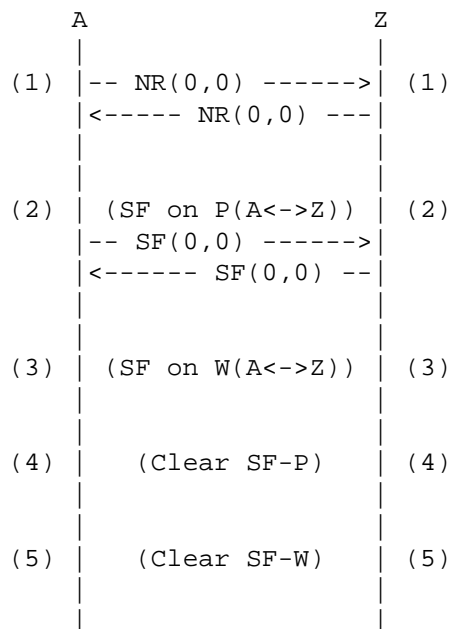
   (1) Each end is in the Normal state and transmits NR(0,0) messages.

   (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, the local "Clear SF-P" request cannot be
     presented to the PSC Control Logic, which takes the highest
     local request and runs the 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 SF 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, the local "Clear SF-W" request can be
     passed to the PSC Control Logic, as there is no higher-priority
     local input, but it will be ignored in the PSC Control Logic
     according to the state transition definition in [RFC6378].
     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 in the UA:P:L state and transmitting SF(0,0) messages,
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 node of the protection
group and is not signaled to the remote node. This command freezes
the state of the protection group. Until the Freeze is cleared,
additional local commands are rejected, and condition changes and
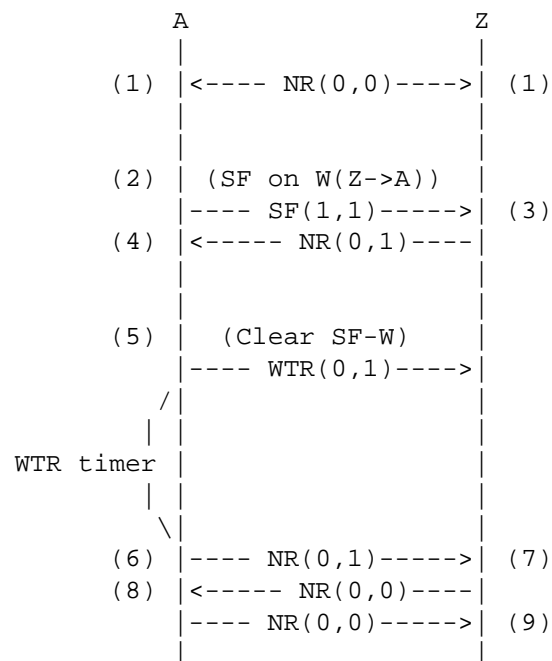received PSC information are ignored.

The "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 that
   differs from the previous message is shown. The operation of the
   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 a PSC message with the following field
   values: Request=SF, FPath=1, and Path=0. 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
   operation) - Unidirectional SF case

```
                       A                 Z
                       |                 |
             (1)  |<---- NR(0,0)---->|  (1)
                       |                 |
                       |                 |
             (2)  |  (SF on W(Z->A))  |
                       |---- SF(1,1)----->|  (3)
             (4)  |<----- NR(0,1)----|
                       |                 |
                       |                 |
             (5)  |   (Clear SF-W)   |
                       |---- WTR(0,1)---->|
                      /|                 |
                       | |                 |
             WTR timer |                 |
                       | |                 |
                      \|                 |
             (6)  |---- NR(0,1)----->|  (7)
             (8)  |<----- NR(0,0)----|
                       |---- NR(0,0)----->|  (9)
                       |                 |
```

   (1) The protected 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 an SF(1,1) message. Both the
        selector and bridge of node A are pointing at the protection
        path.

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

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

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

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

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

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

    Example 2. 1:1 bidirectional protection switching (revertive
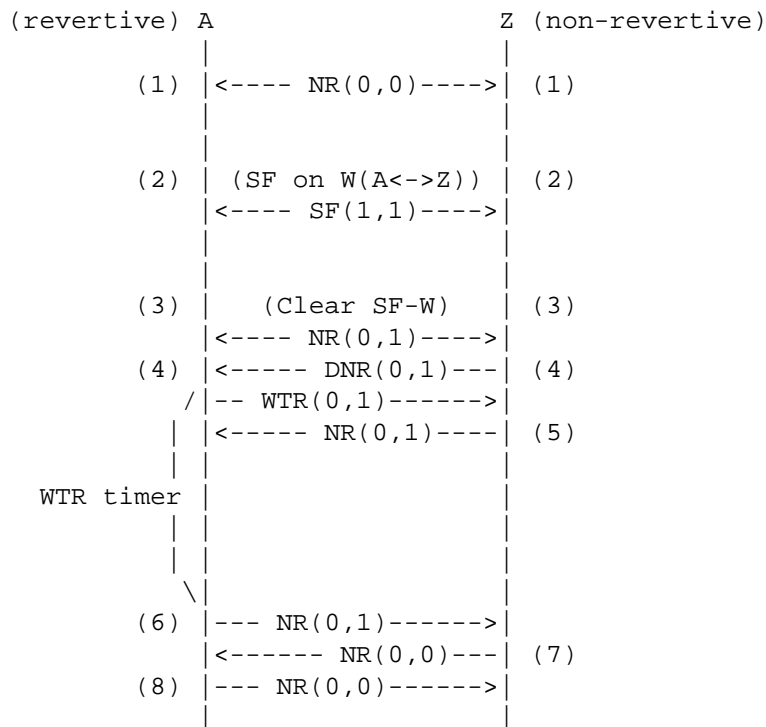   operation) - Bidirectional SF case - Inconsistent WTR timers
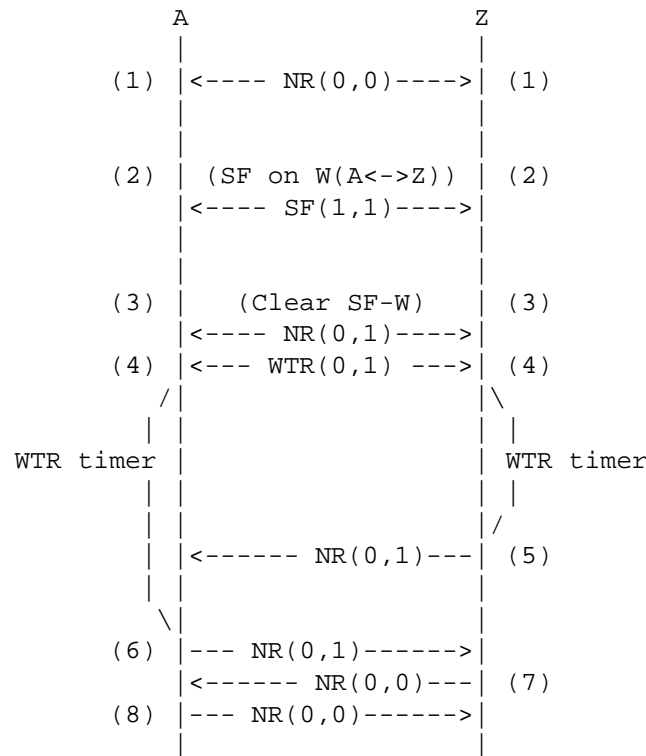
```
                     A                Z
                     |                |
            (1)  |<---- NR(0,0)---->| (1)
                     |                |
                     |                |
            (2)  | (SF on W(A<->Z)) | (2)
                     |<---- SF(1,1)---->|
                     |                |
                     |                |
            (3)  |   (Clear SF-W)   | (3)
                     |<---- NR(0,1)---->|
            (4)  |<--- WTR(0,1) --->| (4)
                   /|                |\
                   | |                | |
        WTR timer  |                | WTR timer
                   | |                | |
                   | |                |/
                   | |<------ NR(0,1)---| (5)
                   | |                |
                   \|                |
            (6)  |--- NR(0,1)------>|
                     |<------ NR(0,0)---| (7)
            (8)  |--- NR(0,0)------>|
                     |                |
```

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

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

   (7) When the NR(0,1) message arrives at node Z, node Z moves to the
        Normal state, sets both the selector and bridge to the working
        path, and sends an 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 operation
   and the other (node Z) is configured as non-revertive operation. The
   interworking is covered in the state transition tables.

```
          (revertive) A              Z (non-revertive)
                       |              |
              (1)  |<---- NR(0,0)---->| (1)
                       |              |
                       |              |
              (2)  | (SF on W(A<->Z)) | (2)
                     |<---- SF(1,1)---->|
                       |              |
                       |              |
              (3)  |    (Clear SF-W)  | (3)
                     |<---- NR(0,1)---->|
              (4)  |<----- DNR(0,1)---| (4)
                   /|-- WTR(0,1)------>|
                    | |<----- NR(0,1)----| (5)
                    | |              |
         WTR timer  | |              |
                    | |              |
                    | |              |
                   \|              |
              (6)  |--- NR(0,1)------>|
                     |<------ NR(0,0)---| (7)
              (8)  |--- NR(0,0)------>|
                       |              |
```

   (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 an SF(1,1) message, each node
        confirms that the remote node 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 Z transits to the DNR state and sends a DNR(0,1)
     message.

(5) When the WTR message arrives at node Z, node Z transits to the
     WTR state and sends an 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
     operation, is operating as if in revertive operation.

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

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

(8) The received NR(0,0) message causes node A to transit 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
   EMail: eric.osborne@notcom.com