Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 15, 2014

E. Osborne
Cisco Systems

September 11, 2013

Updates to PSC
draft-osborne-mpls-psc-updates-02

Abstract

This document contains four updates to RFC6378, "MPLS Transport Profile (MPLS-TP) Linear Protection". Two of them correct existing behavior. The third clears up a behavior which was not explained in the RFC, and the fourth adds rules around handling capabilities mismatches.

Requirements Language

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].

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 15, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 2
2. Incorrect local status after failure 2
3. Reversion deadlock due to a race condition 3
4. Clarifying PSC's behavior in the face of multiple inputs . . 4
5. Handling a capabilities mismatch 6
5.1. PT mismatch 6
5.2. R mismatch 7
5.3. Unsupported modes 7
6. Security Considerations 7
7. IANA Considerations 7
8. Acknowledgements 7
9. References 7
9.1. Normative References 8
9.2. Informative References 8
Author's Address 8

1. Introduction

This document contains four updates to PSC [RFC6378]. Three of them fix issues identified in the ITU's liaison statement "Recommendation ITU-T G.8131/Y.1382 revision - Linear protection switching for MPLS-TP networks" [LIAISON]. The fourth clears up a behavior which was not well explained in RFC6378. These updates are not changes to the protocol's packet format or to PSC's design, but are corrections and clarifications to specific aspects of the protocol's procedures.

2. Incorrect local status after failure

Issue #2 in the liaison identifies a case where a strict reading of RFC6378 leaves a node reporting an inaccurate status

. A node can end up sending incorrect status - NR(0,1) - despite the failure of the protection LSP (P-LSP). This is clearly not correct, as a node should not be sending NR if it has a local failure. To address this issue, the fourth bullet in section 4.3.3.3 is replaced with the following three bullets:

o If the current state is due to a local or remote Manual Switch, a local Signal Fail indication on the protection path SHALL cause

the LER to enter local Unavailable state and begin transmission of an SF(0,0) message.

- If the LER is in local Protecting Administrative state due to a local Forced Switch, a local Signal Fail indication on the protection path SHALL be ignored.

- If the LER is in remote Protecting Administrative state due to a remote Forced Switch, a local Signal Fail indication on the protection path SHALL cause the LER to remain in remote Protecting administrative state and transmit an SF(0,1) message.

3. Reversion deadlock due to a race condition

Issue #8 in the liaison identifies a deadlock case where each node can end up sending NR(0,1) when it should instead be in the process of recovering from the failure (i.e. entering into WTR or DNR, as appropriate for the protection domain). The root of the issue is that a pair of nodes can simultaneously enter WTR state, receive an out of date SF-W indication and transition into a remotely triggered WTR, and remain in remotely triggered WTR waiting for the other end to trigger a change in status.

In the case identified in issue #8, each node can end up sending NR(0,1), which is an indication that the transmitting node has no local failure, but is instead reacting to the remote SF-W. If a node which receives NR(0,1) is in fact not indicating a local error, the receive node can take the received NR(0,1) as an indication that there is no error in the protection domain, and recovery procedures (WTR or DNR) should begin.

This is addressed by adding the following text as the penultimate bullet in section 4.3.3.4:

- If a node is in Protecting Failure state due to a remote SF-W and receives NR(0,1), this SHALL cause the node to begin recovery procedures. If the LER is configured for revertive behavior, it enters into Wait-to-Restore state, starts the WTR timer, and begins transmitting WTR(0,1). If the LER is configured for non-revertive behavior, it enters into Do-Not-Revert state and begins transmitting a DNR(0,1) message.

Additionally, the final bullet in section 4.3.3.3 is changed from

- A remote NR(0,0) message SHALL be ignored if in local Protecting administrative state.

to

- A remote No Request message SHALL be ignored if in local Protecting administrative state.

This indicates that a remote NR triggers the same behavior regardless of the value of FPath and Path. This change does not directly address issue #8, but fixes a similar issue - if a node receives NR while in Remote administrative state, the value of FPath and Path have no bearing on the node's reaction to this NR.

4. Clarifying PSC's behavior in the face of multiple inputs

RFC6378 describes the PSC state machine. Figure 1 in section 3 shows two inputs into the PSC Control logic - Local Request logic and Remote PSC Request. When there is only one input into the PSC Control logic - a local request or a remote request but not both - the PSC Control logic decides what that input signifies and then takes one or more actions, as necessary. This is what the PSC State Machine in section 4.3 describes.

RFC6378 does not sufficiently describe the behavior in the face of multiple inputs into the PSC Control Logic (one Local Request and one Remote Request). This section clarifies the desired behavior.

There are two cases to think about when considering dual inputs into the PSC Control logic. The first is when the same request is presented from both local and remote sources. One example of this case is a failure of the Working LSP. A bidirectional fiber cut will result in the PSC Control logic receiving both a local SF-W (due to loss of light on the underlying fiber) and a remote SF-W (due to the peer node's loss of light). Incidentally, a unidirectional fiber cut will very likely result in a bidirectional failure scenario as it is expected that most MPLS-TP deployments will be running CC/CV [RFC6428]. For convenience, this scenario is written as [L(SF-W), R(SF-W)]

The second case, which is handled in exactly the same way as the first, is when the two inputs into the PSC Control logic describe different events. There are a number of variations on this case. One example is when there is a Lockout of Protection from the Local request logic and a Forced Switch (that is, a forced switch to the protection LSP) from the Remote PSC Request. This is shortened to [L(LO), R(FS)].

In both cases the question is not how the PSC Control logic decides which of these is the one it acts upon. Section 4.3.2 of RFC6378 lists the priority order, and prioritizes the local input over the remote input in case both inputs are of the same priority. So in the first example it is the local SF that drives the PSC Control logic,

and in the second example it is the local Lockout which drives the PSC Control logic.

The point that this section clears up is around what happens when the highest priority input goes away. Consider the first case. Initially, the PSC Control logic has [L(SF-W), R(SF-W)] and L(SF-W) is driving PSC's behavior. When L(SF-W) is removed but R(SF-W) remains, what does PSC do? A strict reading of the FSM would suggest that PSC transition from PA:F:L into N, and at some future time (perhaps after the remote request refreshes) PSC would transition from N to PA:F:R. This is an unreasonable behavior, as there is no sensible justification for a node behaving as if things were normal (i.e. N state) when it is clear that they are not.

The second case is similar. If a node starts with [L(LO), R(FS)] and the local lockout is removed, a strict reading of the state machine would suggest that the node transition from UA:LO:L to N, and then at some future time presumably notice the R(FS) and transition from N to PA:F:R. As with the first case, this is clearly not a useful behavior.

In both cases, the request which was driving PSC's behavior was removed. What should happen is that the PSC Control logic should, upon removal of an input, immediately reevaluate all other inputs to decide on the next course of action. This requires an implementation to store the most recent local and remote inputs regardless of their eventual use as triggers for the PSC Control Logic.

There is a third case. Consider a node with [L(FS), R(LO)]. At some point in time the remote node replaces its Lockout request with a Signal Fail on Working, so that the inputs into the PSC Control logic on the receiving node go to [L(FS), R(SF-W)]. Similar to the first two cases, the node should immediately reevaluate both its local and remote inputs to determine the highest priority among them, and act on that input accordingly. That is in fact what happens, as defined in Section 4.3.3:

"When a LER is in a remote state, i.e., state transition in reaction to a PSC message received from the far-end LER, and receives a new PSC message from the far-end LER that indicates a contradictory state, e.g., in remote Unavailable state receiving a remote FS(1,1) message, then the PSC Control logic SHALL reevaluate all inputs (both the local input and the remote message) as if the LER is in the Normal state."

This section extends that paragraph to handle the first two cases. The essence of the quoted paragraph is that when faced with multiple inputs, PSC must reevaluate any changes as if it was in Normal state. So the quoted paragraph is replaced with the following text:

"The PSC Control logic may simultaneously have Local and Remote requests, and the highest priority of these requests ultimately drives the behavior of the PSC Control logic. When this highest priority request is removed or is replaced with another input, then the PSC Control logic SHALL immediately reevaluate all inputs (both the local input and the remote message), transitioning into a new state only upon reevaluation of all inputs".

5. Handling a capabilities mismatch

PSC has no explicit facility to negotiate any properties of the protection domain. It does, however, have the ability to signal two properties of that domain, via the Protection Type and Revertive bits. RFC6378 specifies that if these bits do not match an operator "SHALL [be notified]" (PT, section 4.2.3) or "SHOULD be notified" (R, section 4.2.4). However, there is no text which specifies the behavior of the end nodes of a protection domain in case of a mismatch

5.1. PT mismatch

The behavior of the protection domain depends on the exact PT mismatch. Section 4.2.3 of RFC6378 specifies three protection types - bidirectional switching using a permanent bridge, bidirectional switching using a selector bridge, and unidirectional switching using a permanent bridge. They are abbreviated here as BP, BS and UP.

There are three possible mismatches: [BP, UP], [BP, BS], and [UP, BS]. The priority is:

UP > BS > BP

In other words:

- If the PT mismatch is [BP, UP], the node transmitting BP MUST switch to UP mode if it is supported.

- If the PT mismatch is [BP, BS] the node transmitting BP MUST switch to BS mode if it is supported.

- If the PT mismatch is [UP, BS] the node transmitting BS MUST switch to UP mode if it is supported.

5.2. R mismatch

The R bit indicates whether the protection domain is in Revertive or Non-Revertive behavior. If the R bits do not match, the node indicating Non-Revertive MUST switch to Revertive if it is supported.

5.3. Unsupported modes

An implementation may not support all three PT modes and/or both R modes, and thus a pair of nodes may be unable to converge on a common mode. This creates a permanent mismatch, resolvable only by operator intervention. An implementation SHOULD alert the operator to an irreconcilable mismatch.

It is desirable to allow the protection domain to function in a non-failuremode even if there is a mismatch, as the mismatches of PT or R have to do with how nodes recover from a failure. An implementation SHOULD allow traffic to be sent on the Working LSP as long as there is no failure (e.g. NR state) regardless of any PT or R mismatch.

If there is a trigger which would cause the protection LSP to be used, such as SF or MS, a node MUST NOT use the protection LSP to carry traffic.

6. Security Considerations

These changes and clarifications raise no new security concerns.

7. IANA Considerations

None.

Note to RFC Editor: this section may be removed on publication as an RFC.

8. Acknowledgements

The author of this document thanks Taesik Cheung, Alessandro D'Alessandro, Annamaria Fulignoli, Sagar Soni, George Swallow and Yaacov Weingarten for their contributions and review.

9. References

9.1. Normative References

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

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

[RFC6428] Allan, D., Swallow Ed. , G., and J. Drake Ed. , "Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile", RFC 6428, November 2011.

9.2. Informative References

[LIAISON] ITU-T SG15, "Liaison Statement: Recommendation ITU-T G.8131/Y.1382 revision - Linear protection switching for MPLS-TP networks", , <https://datatracker.ietf.org/liaison/1205/>.

Author's Address

Eric Osborne
Cisco Systems

Email: eosborne@cisco.com