Network Working Group   D. Li
Internet Draft   H. Xu
Category: Standards Track   Huawei
                                                  S. Bardalai
                                                     Fujitsu
                                                   J. Meuric
                                              France Telecom
                                                 D. Caviglia
                                                    Ericsson

Expires: February 2010   August 14, 2009

                Data Channel Status Confirmation Extensions
                    for the Link Management Protocol

              draft-ietf-ccamp-confirm-data-channel-status-06.txt

Status of this Memo

Abstract

   This document defines simple additions to the Link Management
   Protocol (LMP) to provide a control plane tool that can assist in
   the location of stranded resources by allowing adjacent LSRs to
   confirm data channel statuses, and provides triggers for notifying
   the management plane if any discrepancies are found. As LMP is
   already used to verify data plane connectivity, it is considered to
   be an appropriate candidate to support this feature.

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].

Table of Contents

1. Introduction

   Generalized Multiprotocol Label Switching (GMPLS) networks are
   constructed from Traffic Engineering (TE) links connecting Label
   Switching Routers (LSRs). The TE links are constructed from a set of
   data channels. In this context, a data channel corresponds to a
   resource label in a non-packet technology (such as a timeslot or a
   lambda).

A data channel status mismatch exists if the LSR at one end of a TE
link believes that the data channel is assigned to carry data, but
the LSR at the other end does not. The term "ready to carry data"
means cross-connected or bound to an end-point for the receipt or
delivery of data.

Data channel mismatches cannot be detected from the TE information
advertised by the routing protocols [RFC4203], [RFC4205]. The
existence of some data channel mismatch problems may be detected by
a mismatch in the advertised bandwidths where bidirectional TE links
and bidirectional services are in use, but where unidirectional
services exist, or where multiple data channel mismatches occur, it
is not possible to detect such errors through the routing protocol-
advertised TE information. In any case, there is no mechanism to
isolate the mismatches by determining which data channels are at
fault.

If a data channel mismatch exists, any attempt to use the data
channel for a new LSP will fail. One end of the TE link may attempt
to assign the TE link for use, but the other end will report the
data channel as unavailable when the control plane or management
plane attempts to assign it to an LSP.

Although such a situation can be resolved through the use of the
Acceptable Label Set object in GMPLS signaling [RFC3473], such a
procedure is inefficient since it may require an additional
signaling exchange for each LSP that is set up. When many LSPs are
to be set up, and when there are many data channel mismatches, such
inefficiencies become significant. It is desirable to avoid the
additional signaling overhead, and to report the problems to the
management plane so that they can be resolved to improve the
efficiency of LSP setup.

Correspondingly, such a mismatch situation may give rise to
misconnections in the data plane especially when LSPs are set up
using management plane operations.

Resources (data channels) that are in a mismatched state are often
described as "stranded resources". They are not in use for any LSP,
but they cannot be assigned for use by a new LSP because they appear
to be in use. Although it is theoretically possible for management
plane applications to audit all network resources to locate stranded
resources and to release them, this process is rarely performed
because of the difficulty of coordinating different Element
Management Systems (EMSs), and the associated risks of accidentally
releasing in-use resources. It is desirable to have a control plane
mechanism that detects and reports stranded resources.

This document defines simple additions to the Link Management
Protocol (LMP) [RFC4204] to provide a control plane tool that can
assist in the location of stranded resources by allowing adjacent
LSRs to confirm data channel statuses, and provides triggers for
notifying the management plane if any discrepancies are found. As LMP
is already used to verify data plane connectivity, it is considered
to be an appropriate candidate to support this feature.

2. Problem Explanation

Examples of data channel mismatches are described in the following
three scenarios.

In all of the scenarios, the specific channel resource of a data link
will be unavailable because of the data channel status mismatch, and
this channel resource will be wasted. Furthermore, a data channel
status mismatch may reduce the possibility of successful LSP
establishment, because a data channel status mismatch may result in
failure when establishing an LSP.

So it is desirable to confirm the data channel statuses as early as
possible.

2.1. Mismatch Caused by Manual Configuration

The operator may have configured a cross-connect at only one end of
a TE link using an EMS. The resource at one end of the data channel
is allocated, but the corresponding resource is still available at
the other end of the same data channel. In this case, the data
channel may appear to be available for use by the control plane when
viewed from one end of the TE link, but will be considered to be
unavailable by the other end of the TE link. Alternatively, the
available end of the data channel may be cross-connected by the
management plane and a misconnection may result from the fact that
the other end of the data channel is already cross-connected.

Figure 1 shows a data channel between nodes A and B. The resource at
A's end of the TE link is allocated through manual configuration,
while the resource at B's end of the TE link available, so the data
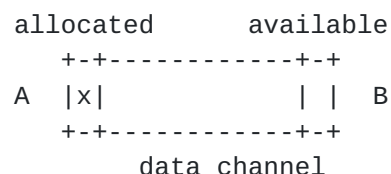channel status is mismatched.

```
            allocated      available
             +-+------------+-+
        A   |x|            | |  B
             +-+------------+-+
                  data channel
```

        Figure 1. Mismatch caused by manual configuration

2.2. Mismatch Caused by LSP Deletion

   The channel status of a data link may become mismatched during the
   LSP deletion process. If the LSP deletion process is aborted in the
   middle of the process (perhaps because of a temporary control plane
   failure), the cross-connect at the upstream node may be removed while
   the downstream node still keeps its cross-connect, if the LSP
   deletion was initiated by the source node.

   For example, in Figure 2 an LSP traverses nodes A, B, and C. Node B
   resets abnormally when the LSP is being deleted. This results in the
   cross-connects of node A and C being removed, but the cross-connect
   of node B still being in use. So the data channel statuses between
   nodes A and B, and between nodes B and C are both mismatched.

```
             <---------LSP--------->
             +-+-------+-+-------+-+
             | |       |X|       | |
             +-+-------+-+-------+-+
              A         B         C
```
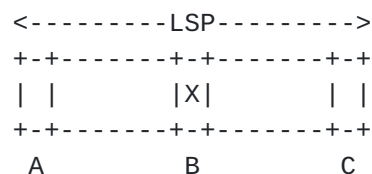   Figure 2. Mismatch caused by LSP deletion

   In [RFC2205] and [RFC3209], a soft state mechanism was defined to
   prevent state discrepancies between LSRs. RSVP-TE restart processes
   ([RFC3473], [RFC5063]) have been defined: adjacent LSRs may
   resynchronize their control plane state to reinstate information
   about LSPs that have persisted in the data plane. Both mechanisms aim
   at keeping state consistency among nodes and allow LSRs to detect
   mismatched data plane states. The data plane handling of such
   mismatched state can be treated as a local policy decision. Some
   deployments may decide to automatically clean up the data plane state
   so it matches the control plane state, but others may choose to raise
   an alert to the management plane and leave the data plane untouched
   just in case it is in use.

   In such cases, data channel mismatches may arise after restart and
   might not be cleared up by the restart procedures.

2.3. Failed Resources

   Even if the situation is not common, it might happen that a
   termination point of a TE-link is seen as failed by one end, while
   on the other end it is seen as OK. This problem may arise due to

some failure either in the hardware or in the status detection of
the termination point.

This mismatch in the termination point status can lead to failure
in case of bidirectional LSP set-up.

```
                Good            Failed
                +-+------------+-+
            A   | |            |X|  B
                +-+------------+-+
                    data channel
            Path Message with Upstream Label---->
```
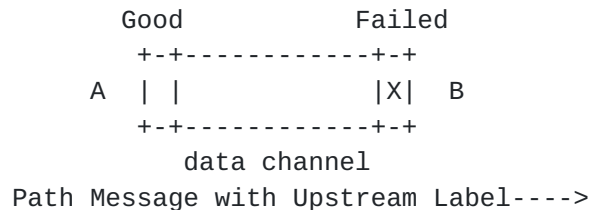
                Figure 3. Mismatch caused by resource failure

In this case upstream node chooses to use termination point A in
order to receive traffic from downstream node. From the upstream
node's point of view, the resource is available thus usable; however,
in the downstream node, the corresponding termination point (resource
B) is broken. This leads to a set-up failure.

3. Motivation

The requirement does not come from a lack in GMPLS specifications
themselves but rather from operational concerns because, in most
cases, GMPLS-controlled networks will co-exist with legacy networks
and legacy procedures.

The protocol extensions defined in this document are intended to
detect data plane problems resulting from mis-use or mis-
configurations triggered by user error, or resulting from failure to
clean up the data plane after control plane disconnection. It is
anticipated that human mistake is probably the major source of errors
to deal with. It is not the intention to provide a protocol mechanism
to deal with broken implementations.

The procedures defined in this document are designed to be operated
on a periodic or on-demand basis. It is NOT RECOMMENDED that the
procedures be used to provide a continuous and on-line monitoring
process.

As LMP is already used to verify data plane connectivity, it is
considered to be an appropriate candidate to support this feature.

4. Extensions to LMP

A control plane tool to detect and isolate data channel mismatches is provided in this document by simple additions to the Link Management Protocol (LMP) [RFC4204]. It can assist in the location of stranded resources by allowing adjacent LSRs to confirm data channel statuses.

Outline procedures are described in this section. More detailed procedures are found in Section 5.

4.1. Confirm Data Channel Status Messages

Extensions to LMP to confirm a data channel status are described below. In order to confirm a data channel status, the new LMP messages are sent between adjacent nodes periodically or driven by some event (such as an operator command, a configurable timer, or the rejection of an LSP setup message because of an unavailable resource). The new LMP messages run over the control channel, encapsulated in UDP with an LMP port number and IP addressing as defined in Link Management Protocol (LMP) [RFC4204].

Three new messages are defined to check data channel status. Message Type numbers are found in Section 7.1.

If the message is a Confirm Data Channel Status message, and the Message_Id value is less than the largest Message_Id value previously received from the sender for the specified TE link, then the message SHOULD be treated as being out-of-order.

4.1.1. ConfirmDataChannelStatus Messages

The ConfirmDataChannelStatus message is used to tell the remote end of the data channel what the status of the local end of the data channel is, and to ask the remote end to report its data channel. The message may report on (and request information about) more than one data channel.

```
<ConfirmDataChannelStatus Message> ::= <Common Header>
                                       <LOCAL_LINK_ID>
                                       <MESSAGE_ID>
                                       <DATA_LINK>[<DATA_LINK>...]
```

When a node receives the ConfirmDataChannelStatus message, and the data channel status confirmation procedure is supported at the node, the node compares its own data channel statuses with all of the data channel statuses sent by the remote end in the ConfirmDataChannelStatus message. If a data channel status mismatch

   is found, this mismatch result is expected to be reported to the
   management plane for further action. Management plane reporting
   procedures and actions are outside the scope of this document.

4.1.2. **ConfirmDataChannelStatusAck Messages**

   The ConfirmDataChannelStatusAck message is sent back to the node
   which originated the ConfirmDataChannelStatus message to return the
   requested data channel statuses.

   When the ConfirmDataChannelStatusAck message is received, the node
   compares the received data channel statuses at the remote end with
   those at the local end (the same operation as performed by the
   receiver of the ConfirmDataChannelStatus message). If a data channel
   status mismatch is found, the mismatch result is expected to be
   reported to the management plane for further action.

   <ConfirmDataChannelStatusAck Message> ::= <Common Header>
                                             <MESSAGE_ID_ACK>
                                             <DATA_LINK>[<DATA_LINK>...]

   The contents of the MESSAGE_ID_ACK objects MUST be obtained from the
   ConfirmDataChannelStatus message being acknowledged.

   Note that the ConfirmDataChannelStatusAck message is used both when
   the data channel statuses match and when they do not match.

4.1.3. **ConfirmDataChannelStatusNack Messages**

   When a node receives the ConfirmDataChannelStatus message, if the
   data channel status confirmation procedure is not supported but the
   message is recognized, a ConfirmDataChannelStatusNack message
   containing an ERROR_CODE indicating "Channel Status Confirmation
   Procedure not supported" MUST be sent.

   If the data channel status confirmation procedure is supported, but
   the node is unable to begin the procedure, a
   ConfirmDataChannelStatusNack message containing an ERROR_CODE
   indicating "Unwilling to Confirm" MUST be sent. If a
   ConfirmDataChannelStatusNack message is received with such an
   ERROR_CODE, the node which originated the ConfirmDataChannelStatus
   message MAY schedule the ConfirmDataChannelStatus message
   retransmission after a configured time. A default value of 10 minutes
   is suggested for this timer.

```
<ConfirmDataChannelStatusNack Message> ::= <Common Header>
                                           [<LOCAL_LINK_ID>]
                                           <MESSAGE_ID_ACK>
                                           <ERROR_CODE>
```

The contents of the MESSAGE_ID_ACK objects MUST be obtained from the
ConfirmDataChannelStatus message being rejected.

4.2. Data Channel Status Subobject

A new Data Channel Status subobject type is introduced to the DATA
LINK object to hold the data channel status and Data Channel
Identification.

See Section 7.2 for the Subobject Type value.

```
 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       |    Length     |      Data Channel Status      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                      Data Channel ID                        //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Data Channel Status:

This is a series of bit flags to indicate the status of the data
channel. The following values are defined.

0x0000 : The channel is available/free.
0x0001 : The channel is unavailable/in-use.

Data Channel ID

This identifies the data channel. The length of this field can be
deduced from the Length field in the subobject. Note that all
subobjects must be padded to a four byte boundary with trailing zeros.
If such padding is required, the Length field MUST indicate the
length of the subobject up to, but not including, the first byte of
padding. Thus, the amount of padding is deduced and not represented
in the Length field.

Note that the Data Channel ID is given in the context of the sender
of the ConfirmChannelStatus message.

   The data-channel ID must be encoded as a label value. Based on the
   type of signal e.g. SONET/SDH, Lambda etc. the encoding methodology
   used will be different. For SONET/SDH the label value is encoded as
   per RFC4606.

4.3. Message Construction

   Data_Link Class is included in ConfirmDataChannelStatus and
   ConfirmDataChannelStatusAck messages, which is defined in section
   13.12 in [RFC4204].

   The status of the TE link end MUST be carried by the Data Channel
   Status subobject which is defined in section 4.2 of this document.
   The new subobject MUST be part of Data_Link Class.

   In the case of SDH/SONET, DATA Channel ID in the new subobject SHOULD
   be used to identify each timeslot of the data link.

4.4. Backward Compatibility

   Some nodes running in the network may only support the LMP Message
   Type from 1 to 20, which are already defined in [RFC4204]. The three
   new types of LMP message (Message Type from 21 to 23) defined in this
   document can not be recognized by these nodes. The unknown message
   behavior is not being specified in [RFC4204], it's suggested to
   discard the unknown message silently. This document's defined
   mechanisms presume a certain non-standard behavior of existing/non-
   document supporting nodes. To use this mechanisms all nodes MUST have
   the extensions described in this document for compatibility.

5. Procedures

   The data channel status confirmation related LMP messages MAY be sent
   between adjacent nodes which are triggered by timer periodically or
   driven by some events to confirm the channel status for the data
   links. It's a local police decision to start the data channel status
   confirmation process. The procedure is described below:

   . The SENDER constructs a ConfirmDataChannelStatus message which
      MUST contain one or more DATA_LINK objects. DATA_LINK object is
      defined in [RFC4204]. Each DATA_LINK object MUST contain one or
      more Data Channel Status subobjects. The Data Channel ID field in
      the Data Channel Status subobject MUST indicate which data channel
      needs to be confirmed, and MUST report the data channel status at
      the SENDER. The ConfirmDataChannelStatus message is sent to the
      RECEIVER.

. The RECEIVER MUST extract the data channel statuses from the
   ConfirmDataChannelStatus message, and SHOULD compare these with
   its data channel statuses for the reported data channels. If a
   data channel status mismatch is found, the mismatch result SHOULD
   be reported to the management plane for further action. The
   RECEIVER also SHOULD send the ConfirmDataChannelStatusAck message
   which MUST carry all the local end statuses of the requested data
   channels to the SENDER.

. If the RECEIVER is not able to support or to begin the
   confirmation procedure, the ConfirmDataChannelStatusNack message
   MUST be responded with the ERROR_CODE which indicates the reason
   of rejection.

. When the SENDER receives the response ConfirmDataChannelStatusAck
   message, and MUST compare the received data channel statuses at
   the remote end with the data channel statuses at the local end. If
   a data channel status mismatch is found, the mismatch result
   SHOULD be reported to the management plane for further action.

The data channel status mismatch issue identified by LMP may be
automatically resolved by RSVP restart. For example, the restarting
node may also have damaged its data plane. This leaves the data
channels mismatched. But RSVP restart will re-install the data plane
state in the restarting node. The issue may also be resolved via RSVP
soft state timeout.

If the ConfirmDataChannelStatus message is not recognized by the
RECEIVER, the RECEIVER ignores this message, and will not send out an
acknowledgment message to the SENDER.

Due to message loss problem, the SENDER may not be able to receive
the acknowledgment message.

ConfirmDataChannelStatus SHOULD be sent using LMP [RFC4204] reliable
transmission mechanisms. If after the retry limit is reached, a
ConfirmDataChannelStatusAck message or a ConfirmDataChannelStatusNack
message is not received by the SENDER, the SENDER SHOULD terminate
the data channel confirmation procedure.

6. Security Considerations

[RFC4204] describes how LMP messages between peers can be secured,
and these measures are equally applicable to the new messages defined
in this document.

   The operation of the procedures described in this document does not
   of themselves constitute a security risk since they do not cause any
   change in network state. It would be possible, if the messages were
   intercepted or spoofed to cause bogus alerts in the management plane
   and so the use of the LMP security measures are RECOMMENDED.

   Note that operating the procedures described in this document may
   provide a useful additional security measure to verify that data
   channels have not been illicitly modified.

7. IANA Considerations

7.1. LMP Message Types

   IANA maintains the "Link Management Protocol (LMP)" registry which
   has a subregistry called "LMP Message Type". IANA is requested to
   make three new allocations from this registry as follows. The message
   type values are suggested and to be confirmed by IANA.

   Value   Description
   ------   --------------------------------
    32   ConfirmDataChannelStatus
    33   ConfirmDataChannelStatusAck
    34   ConfirmDataChannelStatusNack

7.2. LMP Data Link Object Subobject

   IANA maintains the "Link Management Protocol (LMP)" registry which
   has a subregistry called "LMP Object Class name space and Class type
   (C-Type)". This subregistry has an entry for the DATA_LINK object,
   and there is a further embedded registry called "DATA_LINK Sub-object
   Class name space". IANA is requested to make the following allocation
   from this embedded registry. The value shown is suggested and to be
   confirmed by IANA.

   Value   Description
   ------   --------------------------------
     9   Data Channel Status

8. Acknowledgments

   We would like to thank Adrian Farrel, Dimitri Papadimitriou, Lou
   Berger for their useful comments.

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.

[RFC4204]   J. Lang, Ed., "Link Management Protocol (LMP)", RFC 4204,
             October 2005.

9.2. Informative References

[RFC2205]   R. Braden, Ed., "Resource ReSerVation Protocol (RSVP) --
             Version 1 Functional Specification", RFC 2205, September
             1997

[RFC3209]   D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G.
             Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels",
             RFC 3209, December 2001

[RFC3473]   L. Berger, Ed., "Generalized Multi-Protocol Label
             Switching (GMPLS) Signaling Resource ReserVation
             Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC
             3473, January 2003

[RFC5063]   A. Satyanarayana, R. Rahman, "Extensions to GMPLS RSVP
             Graceful Restart", RFC 5063, September 2007

[RFC4203]   K. Kompella, Ed., "OSPF Extensions in Support of
             Generalized Multi-Protocol Label Switching (GMPLS) ", RFC
             4203, October 2005

[RFC4205]   K. Kompella, Ed., "Intermediate System to Intermediate
             System (IS-IS) Extensions in Support of Generalized
             Multi-Protocol Label Switching (GMPLS) ", RFC 4205,
             October 2005

10. Authors' Addresses

Dan Li
Huawei Technologies
F3-5-B R&D Center, Huawei Base,
Shenzhen 518129 China

Phone: +86 755-289-70230
Email: danli@huawei.com

Huiying Xu
Huawei Technologies
F3-5-B R&D Center, Huawei Base,
Shenzhen 518129 China

Phone: +86 755-289-72910
Email: xuhuiying@huawei.com

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base,
Shenzhen 518129 China

Phone: +86 755-289-72912
Email: zhangfatai@huawei.com

Snigdho C. Bardalai
Fujitsu Network Communications
2801 Telecom Parkway,
Richardson, Texas 75082, USA

Phone: +1 972 479 2951
Email: snigdho.bardalai@us.fujitsu.com

Julien Meuric
France Telecom Orange Labs
2, avenue Pierre Marzin
22307 Lannion Cedex, France

Phone: +33 2 96 05 28 28
Email: julien.meuric@orange-ftgroup.com

Diego Caviglia
Ericsson
Via A. Negrone 1/A 16153
Genoa Italy

Phone: +39 010 600 3736
Email: diego.caviglia@ericsson.com

13. Disclaimer of Validity

   All IETF Documents and the information contained therein are provided
   on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE
   REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE
   IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL
   WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY
   WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE
   ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS
   FOR A PARTICULAR PURPOSE. Provisions Relating to IETF Documents in
   effect on the date of publication of this document
   (http://trustee.ietf.org/license-info). Please review these documents
   carefully, as they describe your rights and restrictions with respect
   to this document.