Network Working Group   K. Ogawa
Internet-Draft   NTT Corporation
Intended status: Standards Track   W. M. Wang
Expires: January 17, 2013   Zhejiang Gongshang University
                                                      E. Haleplidis
                                               University of Patras
                                                      J. Hadi Salim
                                                  Mojatatu Networks
                                                      July 16, 2012

ForCES Intra-NE High Availability
draft-ietf-forces-ceha-04

Abstract

   This document discusses CE High Availability within a ForCES NE.

Status of this Memo

Copyright Notice

   described in the Simplified BSD License.

Table of Contents

[1](#). Definitions

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](#).

The following definitions are taken from [[RFC3654](#)]and [[RFC3746](#)]:

Logical Functional Block (LFB) -- A template that represents a fine-grained, logically separate aspects of FE processing.

ForCES Protocol -- The protocol used at the Fp reference point in the ForCES Framework in [[RFC3746](#)].

ForCES Protocol Layer (ForCES PL) -- A layer in the ForCES architecture that embodies the ForCES protocol and the state transfer mechanisms as defined in [[RFC5810](#)].

ForCES Protocol Transport Mapping Layer (ForCES TML) -- A layer in ForCES protocol architecture that specifically addresses the protocol message transportation issues, such as how the protocol messages are mapped to different transport media (like SCTP, IP, TCP, UDP, ATM, Ethernet, etc), and how to achieve and implement reliability, security, etc.

2. Introduction

   Figure 1 illustrates a ForCES NE controlled by a set of redundant CEs
   with CE1 being active and CE2 and CEn-1 being a backup.

```
                       ------------------------------------------
                       | ForCES Network Element                 |
                       |                      +-----------+ |
                       |                      |  CEn-1    | |
                       |                      | (Backup)  | |
     --------------  Fc | +------------+    +------------+ | |
    | CE Manager |--------+-|    CE1    |------|    CE2    |-+ |
     --------------      | | (Active) | Fr | (Backup) |   |
       |                | +-------+--+-+    +---+---+----+   |
       | Fl             |        |  |  Fp      /  |          |
       |                |        |  | +---------+ /   |      |
       |                |        Fp|            |/   |Fp     |
       |                |        |  |            |   |       |
       |                |        |  |   Fp    /+--+  |       |
       |                |        |  | +-------+  | |          |
       |                |        |  | |  |     | |          |
     --------------  Ff | --------+--+--   ----+---+----+   |
    | FE Manager |--------+-|    FE1    | Fi |    FE2    |   |
     --------------      | |           |------|          |   |
                       | --------------   --------------   |
                       | | | | | |       | | | |       |
                       ----+--+--+--+----------+--+--+--+-------
                         | | | |         | | | |
                         | | | |         | | | |
                          Fi/f              Fi/f
```
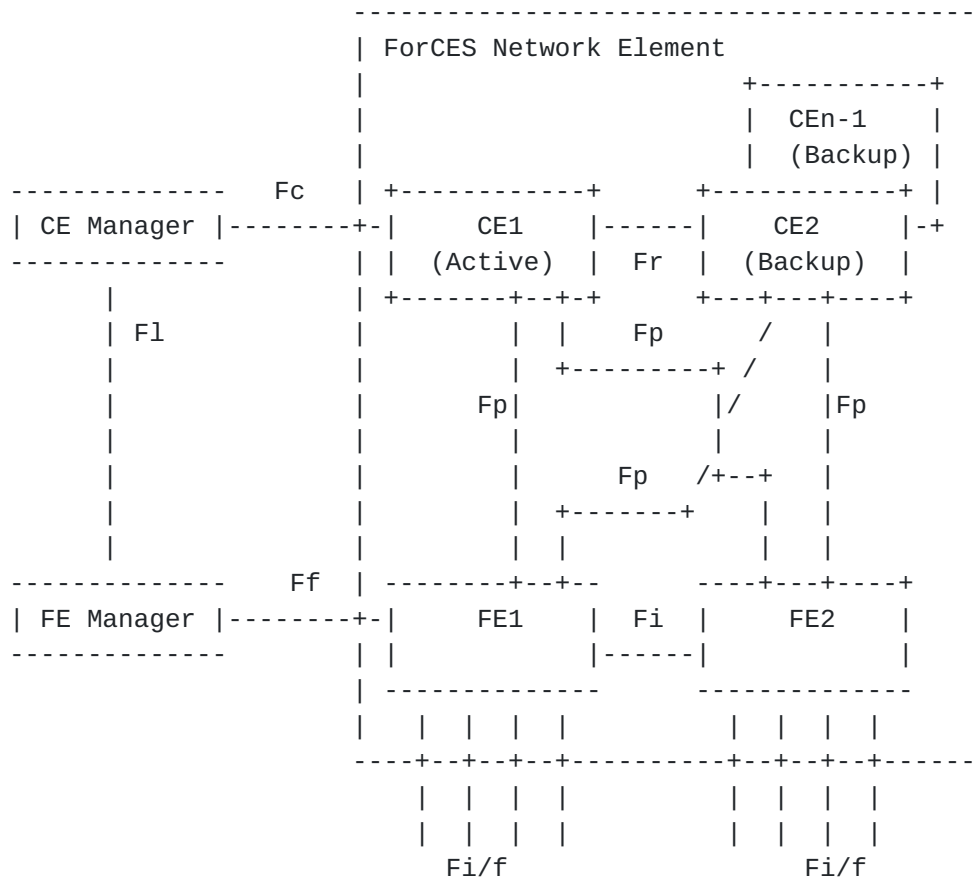
          Fp: CE-FE interface
          Fi: FE-FE interface
          Fr: CE-CE interface
          Fc: Interface between the CE Manager and a CE
          Ff: Interface between the FE Manager and an FE
          Fl: Interface between the CE Manager and the FE Manager
          Fi/f: FE external interface

                        Figure 1: ForCES Architecture

   The ForCES architecture allows FEs to be aware of multiple CEs but
   enforces that only one CE be the master controller. This is known in
   the industry as 1+N redundancy. The master CE controls the FEs via
   the ForCES protocol operating in the Fp interface. If the master CE
   becomes faulty, a backup CE takes over and NE operation continues.
   By definition, the current documented setup is known as cold-standby.
   The CE set is static and is passed to the FE by the FE Manager (FEM)

   via the Ff interface and to each CE by the CE Manager (CEM) in the Fc
   interface during the pre-association phase.

   From an FE perspective, the knobs of control for a CE set are defined
   by the FEPO LFB in [RFC5810], Appendix B. Section 3.1 of this
   document details these knobs further.

2.1. Document Scope

   It is assumed that the reader is aware of the ForCES architecture to
   make sense of the changes made here. This document provides minimal
   background to set the context of the discussion in Section 4.

   By current definition, the Fr interface is out of scope for the
   ForCES architecture. However, it is expected that organizations
   implementing a set of CEs will need to have the CEs communicate to
   each other via the Fr interface in order to achieve the
   synchronization necessary for controlling the FEs.

   The problem scope addressed by this document falls into 2 areas:

   1. To describe with more clarity (than [RFC5810]) how current cold-
       standby approach operates within the NE cluster.

   2. To describe how to evolve the cold-standby setup to a hot-standby
       redundancy setup so as to improve the failover time and NE
       availability.

2.2. Quantifying Problem Scope

   The NE recovery and availability is dependent on several time-
   sensitive metrics:

   1. How fast the CE plane failure is detected the FE.

   2. How fast a backup CE becomes operational.

   3. How fast the FEs associate with the new master CE.

   4. How fast the FEs recover their state and become operational.

   The design goals of the current [RFC5810] choices to meet the above
   goals are driven by desire for simplicity.

   To quantify the above criteria with the current prescribed ForCES CE
   setup in [RFC5810]:

1. How fast the CE side detects a CE failure is left undefined. To
    illustrate an extreme scenario, we could have a human operator
    acting as the monitoring entity to detect faulty CEs. How fast
    such detection happens could be in the range of seconds to days.
    A more active monitor on the Fr interface could improve this
    detection.

2. How fast the backup CE becomes operational is also currently out
    of scope. In the current setup, a backup CE need not be
    operational at all (for example, to save power) and therefore it
    is feasible for a monitoring entity to boot up a backup CE after
    it detects the failure of the master CE. In this document
    Section 4 we suggest that at least one backup CE be online so as
    to improve this metric.

3. How fast an FE associates with new master CE is also currently
    undefined. The cost of an FE connecting and associating adds to
    the recovery overhead. As mentioned above we suggest having at
    least one backup CE online. In Section 4 we propose to zero out
    the connection and association cost on failover by having each FE
    associate with all online backup CEs after associating to the
    active CE. Note that if an FE pre-associates with backup CEs,
    then the system will be technically operating in hot-standby
    mode.

4. And last: How fast an FE recovers its state depends on how much
    NE state exists. By ForCES current definition, the new master CE
    assumes zero state on the FE and starts from scratch to update
    the FE. So the larger the state, the longer the recovery.

3. RFC5810 CE HA Framework

To achieve CE High Availabilty, FEs and CEs MUST inter-operate per
[RFC5810] definition which is repeated for contextual reasons in
Section 3.1. It should be noted that in this default setup, which
MUST be implemented by CEs and FEs needing HA, the Fr plane is out of
scope (and if available is proprietary to an implementation).

3.1. Current CE High Availability Support

As mentioned earlier, although there can be multiple redundant CEs,
only one CE actively controls FEs in a ForCES NE. In practice there
may be only one backup CE. At any moment in time only one master CE
can control the FEs. In addition, the FE connects and associates to
only the master CE. The FE and the CE PL are aware of the primary
and one or more secondary CEs. This information (primary, secondary
CEs) is configured on the FE and the CE PLs during pre-association by

the FEM and the CEM respectively.

Figure 2 below illustrates the Forces message sequences that the FE
uses to recover the connection in current defined cold-standby
scheme.


```
      FE                       CE Primary         CE Secondary
      |                         |                  |
      |  Asso Estb,Caps exchg   |                  |
    1 |<--------------------->|                    |
      |                         |                  |
      |         state update    |                  |
    2 |<--------------------->|                    |
      |                         |                  |
      |                         |                  |
      |              FAILURE                        |
      |                                             |
      |          Asso Estb,Caps exchange            |
    3 |<------------------------------------------->|
      |                                             |
      |            Event Report (pri CE down)       |
    4 |-------------------------------------------->|
      |                                             |
      |                  state update               |
    5 |<------------------------------------------->|
```
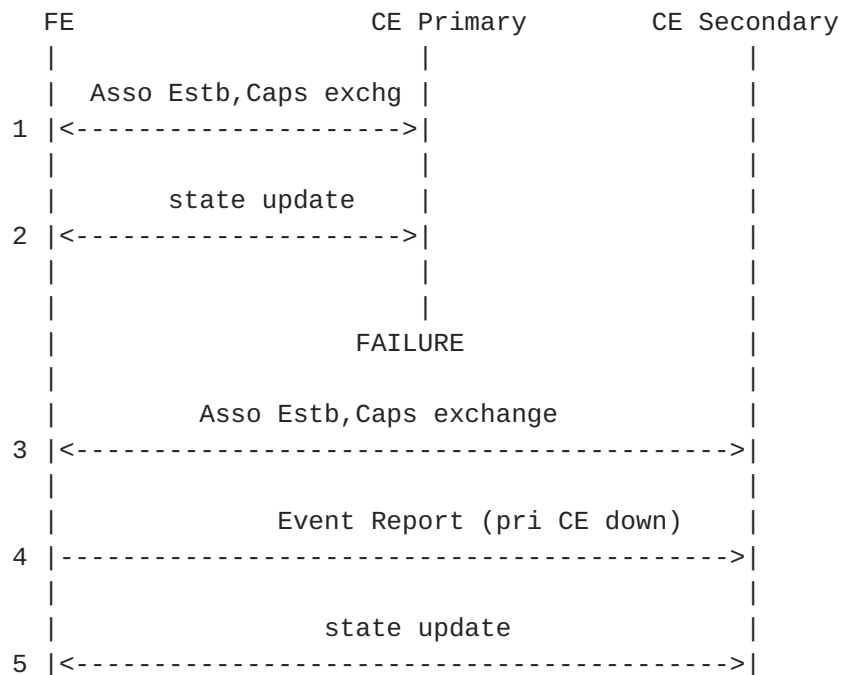

                 Figure 2: CE Failover for Cold Standby

3.1.1. Cold Standby Interaction with ForCES Protocol

   High Availability parameterization in an FE is driven by configuring
   the FE Protocol Object (FEPO) LFB.

   The FEPO CEID component identifies the current master CE and the
   component table BackupCEs identifies the backup CEs. The FEPO FE
   Heartbeat Interval, CE Heartbeat Dead Interval, and CE Heartbeat
   policy help in detecting connectivity problems between an FE and CE.
   The CE Failover policy defines how the FE should react on a detected
   failure.

   Figure 3 illustrates the defined state machine that facilitates
   connection recovery.

   The FE connects to the CE specified on FEPO CEID component. If it
   fails to connect to the defined CE, it moves it to the bottom of

table BackupCEs and sets its CEID component to be the first CE
retrieved from table BackupCEs. The FE then attempts to associate
with the CE designated as the new primary CE. The FE continues
through this procedure until it successfully connects to one of the
CEs.

```
                              FE tries to associate
                                +-->-----+
                                |        |
     (CE issues Teardown ||   +---+--------v----+
       Lost association) &&   | Pre-Association |
     CE failover policy = 0   | (Association    |
        +------------->-->-->|   in             +<----+
        |                     | progress)        |    |
        |        CE Issues    +--------+--------+     |
        |        Association           |             | CEFTI
        |          Response            V             | timer
        |        +-----------------+                 | expires
        |        |                                   ^
        |        V                                   |
      +-+-+----------+                    +------+-----+
      |              |                    |  Not       |
      |              | (CE issues Teardown || | Associated |
      |              |   Lost association) &&  |            +->---+
      | Associated   | CE Failover Policy = 1 |(May        | FE  |
      |              |                        | Continue   |try  v
      |              |-------->------->------>| Forwarding)|assn |
      |              |    Start CEFTI timer   |            |-<---+
      |              |                        |            |
      +-------------+                         +-------+-----+
           ^                                          |
           |           CE Issues                      V
           |           Association                    |
           |           Setup                          |
           |           (Cancel CEFTI Timer)           |
           +_____+
```

                Figure 3: FE State Machine considering HA

When communication fails between the FE and CE (which can be caused
by either the CE or link failure but not FE related), either the TML
on the FE will trigger the FE PL regarding this failure or it will be
detected using the HB messages between FEs and CEs. The
communication failure, regardless of how it is detected, MUST be
considered as a loss of association between the CE and corresponding
FE.

If the FE's FEPO CE Failover Policy is configured to mode 0 (the
default), it will immediately transition to the pre-association
phase. This means that if association is again established, all FE
state will need to be re-established.

If the FE's FEPO CE Failover Policy is configured to mode 1, it
indicates that the FE is capable of HA restart recovery. In such a
case, the FE transitions to the Not Associated state and the CEFTI
timer[RFC 5810] is started. The FE MAY continue to forward packets
during this state. It MAY also recycle through any configured backup
CEs in a round-robin fashion. It first adds its primary CE to the
bottom of table BackupCEs and sets its CEID component to be the first
secondary retrieved from table BackupCEs. The FE then attempts to
associate with the CE designated as the new primary CE. If it fails
to re-associate with any CE and the CEFTI expires, the FE then
transitions to the pre-association state.

If the FE, while in the not associated state, manages to reconnect to
a new primary CE before CEFTI expires it transitions to the
Associated state. Once re-associated, the CE tries to synchronize
any state that the FE may have lost during the not associated state.
How the CE re-synchronizes such state is out of scope for the current
ForCES architecture but would typically constitute the issuing of new
configs and queries.

An explicit message (a Config message setting Primary CE component in
ForCES Protocol object) from the primary CE, can also be used to
change the Primary CE for an FE during normal protocol operation. In
this case, the FE transitions to the Not Associated State and
attempts to Associate with the new CE.

3.1.2. Responsibilities for HA

TML Level:

1. The TML controls logical connection availability and failover.

2. The TML also controls peer HA management.

At this level, control of all lower layers, for example transport
level (such as IP addresses, MAC addresses etc) and associated links
going down are the role of the TML.

PL Level:
All other functionality, including configuring the HA behavior during
setup, the CE IDs used to identify primary and secondary CEs,
protocol messages used to report CE failure (Event Report), Heartbeat
messages used to detect association failure, messages to change the

primary CE (Config), and other HA related operations described in
Section 3.1, are the PL's responsibility.

To put the two together, if a path to a primary CE is down, the TML
would take care of failing over to a backup path, if one is
available. If the CE is totally unreachable then the PL would be
informed and it would take the appropriate actions described before.

4. CE HA Hot Standby

In this section we describe small extensions to the existing scheme
to enable hot standby HA. To achieve hot standby HA, we target
specific goals defined in Section 2.2, namely:

o How fast a backup CE becomes operational.

o How fast the FEs associate with the new master CE.

As described in Section 3.1, in the pre-association phase the FEM
configures the FE to make it aware of all the CEs in the NE. The FEM
MUST configure the FE to make it aware which CE is the master and MAY
specify any backup CE(s).

4.1. Changes to the FEPO model

In order for the above to be achievable there is a need to make a few
changes in the FEPO model. Section 1 contains the xml definition of
the new version 2 of the FEPO LFB.

Changes from the version 1 of FEPO are:

1. Added four new datatypes:

    1. CEStatusType an unsigned char to specify status of a
        connection with a CE. Special values are 0 (Disconnected), 1
        (Connected), 2 (Associated), 3 (Lost_Connection) and 4
        (Unreachable)

    2. HAModeValues an unsigned char to specify selected HA mode.
        Special values are 0 (No HA Mode), 1 (HA Mode - Cold Standby)
        and 2 (HA Mode - Hot Standby)

    3. FEHACapab an unsigned char to specify HA capabilities of the
        FE. Special values are 0 (Graceful Restart), 1 (Cold
        Standby) and 2 (Hot Standby)

   4. AllCEType a struct of CE ID and CEStatusType to contain
       connection information for one CE. Used in the AllCEs array.

2. Appended three new components:

   1. AllCEs to hold status for all CEs. AllCEs is an Array of the
       AllCEType.

   2. HAMode to specify current High Availability mode selected.
       An unsigned char with three special values 0 (No HA), 1
       (Running Cold-Standby) and 2 (Running Hot-Standby)

   3. AcceptBackupGets to provide the master CE to control whether
       the FE will accept incoming queries from backup CEs.

3. Added two new capabilities. The HACapabilities, a table that
    defines which HA capabilities the FE supports and the
    MaximumMultipleCEAssocations which defines the maximum
    associations with CEs this FE can have.

4. Added one additional Event, the HAPrimaryCEDown event which
    reports last known CEID and tentative new master CEID.

Since no component has been altered there is no compatibility issue
between this version of the FEPO with the old one. This means that a
CE which understands only version 1.0 of the FEPO can use the newer
version without making use of the High Availability options that it
provides.

4.2. FEPO processing

The FE's FEPO LFB version 2 AllCEs table contains all the CEIDs that
the FE may connect and associate with. The ordering of the CE IDs in
this table defines the priority order in which an FE will connect to
the CEs. In the pre-association phase, the first CE ID (lowest table
index) in the AllCEs table MUST be the first CE ID that the FE will
attempt to connect and associate with. If the FE fails to connect
and associate with the first CE ID, it will attempt to connect to the
second CE ID and so forth, and cycles back to the beggining of the
list until there is a connection and an association. The FE MUST
associate with at least one CE. Upon a successful association, the
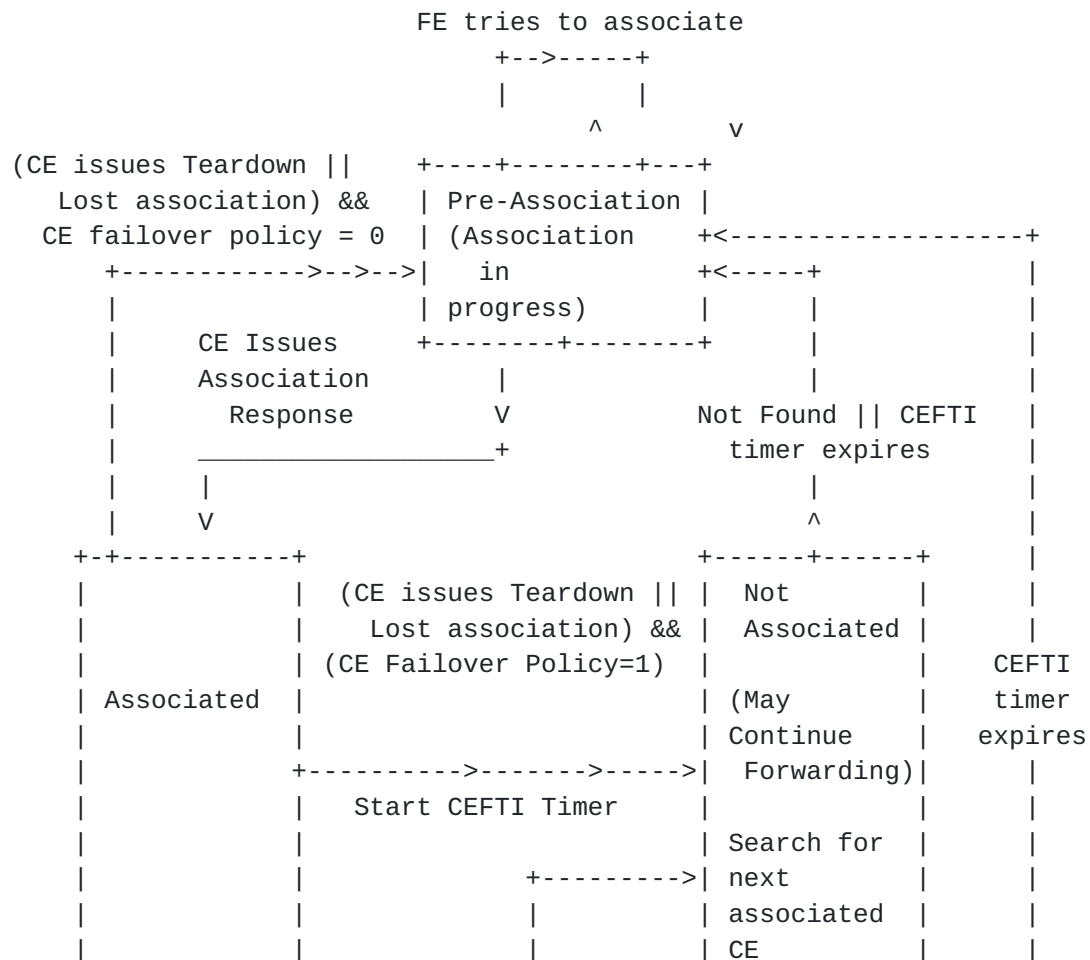FEPO's CEID component identifies the current associated master CE.

While it would be much simpler to have the FE not respond to any
messages from CE other than the master, it may be useful for the
backup CEs to be able to query the FE. Query commands are sent
always on the high priority channel. In order to avoid missing
critical configuration or query commands from the master CE, all

query commands from backup CEs MUST be sent on the high priority
channel but with the least priority, the value of which is 4.
However since queries are high priority from heartbeats, if the
master CE waits for heartbeat responses and the backup CEs flood the
FE, the master CE may think that the FE is down. Therefore it is
prudent to add a control mechanism that will be able to control
whether the FE can respond to query messages from backup CEs. The
AcceptBackupGets component, a boolean, is designed for this occasion.
If the master CE sets it to true, the FE MUST accept and process
query commands from backup CEs. If the AcceptBackupGets is false,
the FE MUST drop query commands from backup CEs.

Asynchronous events that the master CE has subscribed to, as well as
heartbeats are sent to all associated-to CEs. Packet redirects
continue to be sent only to the master CE. The Heartbeat Interval,
the CEHB Policy and the FEHB Policy MUST be the same for all CEs.

Figure 4 illustrates the state machine that facilitates connection
recovery with High Availability enabled.

```
                            FE tries to associate
                               +-->-----+
                               |        |
                                   ^        v
 (CE issues Teardown ||     +----+--------+---+
   Lost association) &&     | Pre-Association |
   CE failover policy = 0   | (Association     +<-------------------+
       +------------->-->-->|   in              +<-----+           |
       |                    | progress)        |      |           |
       |       CE Issues    +--------+--------+      |           |
       |       Association          |                |           |
       |         Response           V          Not Found || CEFTI  |
       |     _____+             timer expires      |
       |     |                                   |             |
       |     V                                   ^             |
     +-+-----------+                    +------+------+        |
     |            | (CE issues Teardown || | Not       |      |
     |            |   Lost association) && | Associated |     |
     |            | (CE Failover Policy=1)  |            |     CEFTI
     | Associated |                         | (May       |     timer
     |            |                         | Continue   |     expires
     |            +---------->------->----->| Forwarding)|      |
     |            |   Start CEFTI Timer     |            |      |
     |            |                         | Search for |      |
     |            |              +--------->| next       |      |
     |            |              |          | associated |      |
     |            |              |          | CE         |      |
```

```
     |               |                  |  | (HAMode 2)  |        |
     +-------------+                   |      +-------------+       |
        ^                              |            V               |
        |                             |            |               |
        |                             |         Found CE            |
        |                     CEHDI Expires   Send Event of         |
        |                             |           New CE ID.        |
        |                             |         Start CEHDI Timer   |
        |                             |            |               |
        |                             |            V               |
        |                             |      +------+------+        |
        |                     ^---------+ Confirm     +-------^
        |                             | State       |
        |           Received      +---->|            |
        |           different     |     | Wait for CE |
        |           CE ID.        ^     | to confirm  |
        |           Resend Event |      | new CE ID   |
        |     Restart CEHDI Timer +----<|            |
        |                             +-----+-------+
        |           Received same CE ID        |
        |        (Cancel CEFTI & CEHDI Timer)    |
        +_____+
```

                  Figure 4: FE State Machine considering HA
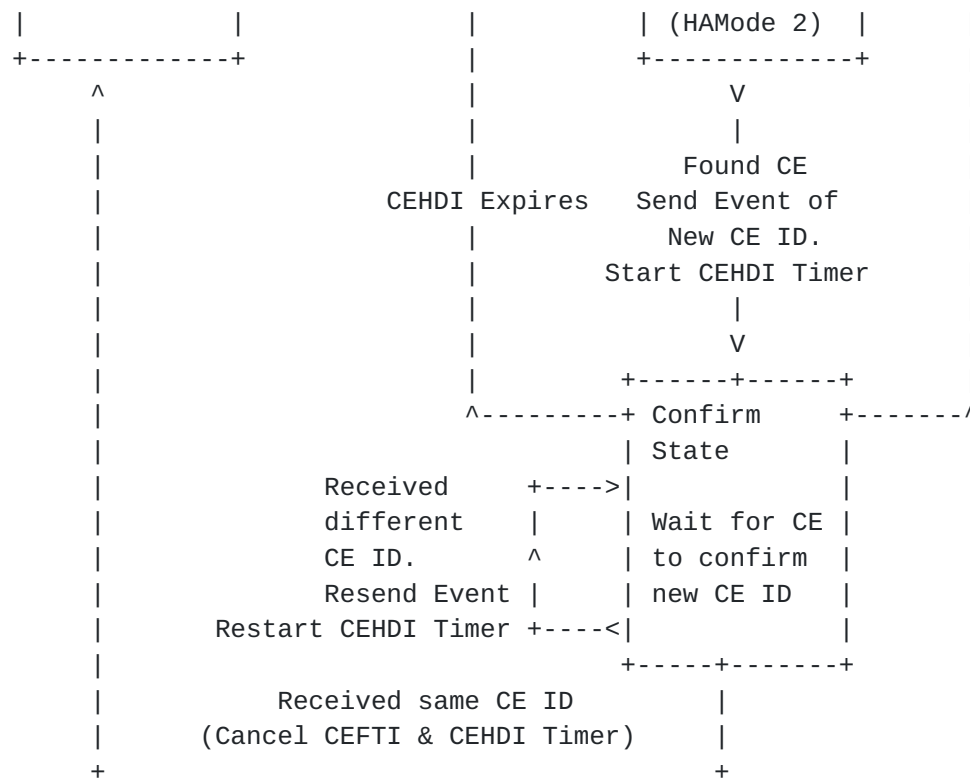
   Once the FE has associated with a master CE it moves to the post-
   association phase (Associated state). It MAY also instruct the FE to
   use a different master CE. It is assumed that the master CE will
   communicate with other CEs within the NE for the purpose of
   synchronization via the CE-CE interface. The CE-CE interface is out
   of scope for this document.

```
        FE                    CE#1           CE#2 ... CE#N
         |                     |              |       |
         | Asso Estb,Caps exchg |              |       |
       1 |<-------------------->|              |       |
         |                     |              |       |
         |      state update    |              |       |
       2 |<-------------------->|              |       |
         |                     |              |       |
         |       Asso Estb,Caps exchg          |       |
      3I |<----------------------------------->|       |
        ...              ...              ...       ...
         |              Asso Estb,Caps exchg           |
      3N |<------------------------------------------->|
         |                     |              |       |
       4 |<-------------------->|              |       |
         .                     .              .       .
      4x |<-------------------->|              |       |
         |                  FAILURE            |       |
         |                     |              |       |
         | Event Report (CE#2 is new master) |       |
       5 |---------------------------------->|------->|
         |                     |              |       |
         | Config (Set CEID to CEID of CE#2) |       |
       6 |<----------------------------------|       |
       7 |<---------------------------------->|       |
         .                     .              .       .
      7x |<---------------------------------->|       |
         .                     .              .       .
```
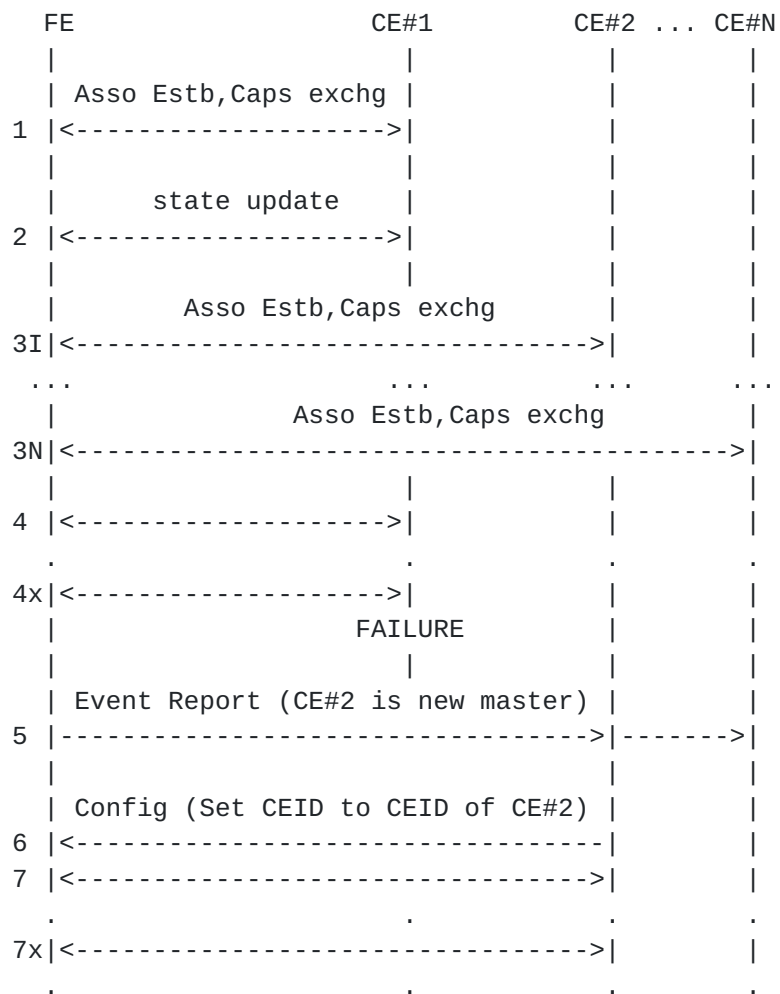
Figure 5: CE Failover for Hot Standby

   While in the post-association phase, if the CE Failover Policy is set
   to 1 and HAMode set to 2 (HotStandby) then the FE, after succesfully
   associating with the master CE, MUST attempt to connect and associate
   with all the CEs that is aware of. Figure 5 steps #1 and #2
   illustrates the FE associating with CE#1 as the master and then
   proceeding to steps #3I to #3N the association with backup CE's CE#2
   to CE#N. If the FE fails to connect or associate with some CEs, the
   FE MAY flag them as unreachable to avoid continuous attempts to
   connect. The FE MAY retry to reassociate with unreachable CEs when
   possible.

   When the master CE for any reason is considered to be down, then the
   FE will try to find the first associated CE from the list of all CEs
   in a round-robin fashion.

   If the FE is unable to find an associated FE in its list of CEs, then
   it will attempt to connect and associate with the first from the list

of all CEs and continue in a round-robin fashion until it connects
and associates with a CE.

Once the FE selects the associated CE to use as the new master, the
FE then sends a High Availability Primary CE Changed Event
Notification to all associated CEs to notifying them that the primary
CE is down as well as which CE the reporting FE considers to be the
new master.

The new master CE MUST configure the CEID component of the FE within
the time limit defined in the CEHDI Failover Timeout as a
confirmation that the FE made the right choice.

```
        FE                      CE#1          CE#2 ... CE#N
        |                        |             |        |
        | Asso Estb,Caps exchg   |             |        |
     1  |<--------------------->|             |        |
        |                        |             |        |
        |      state update      |             |        |
     2  |<--------------------->|             |        |
        |                        |             |        |
        |       Asso Estb,Caps exchg           |        |
     3I |<------------------------------------>|        |
        |                        |             |        |
       ...                      ...           ...      ...
        |             Asso Estb,Caps exchg              |
     3N |<--------------------------------------------->|
        |                        |             |        |
     4  |<--------------------->|             |        |
        .                        .             .        .
     4x |<--------------------->|             |        |
        |                    FAILURE           |        |
        |                        |             |        |
        | Event Report (CE#2 is new master) |        |
     5  |---------------------------------->|------->|
        |                        |             |        |
        |        CEHDI Failover Timeout        |        |
        |                        |             |        |
        | Event Report (CE#N is new master) |        |
     6  |---------------------------------->|------->|
        |                        |             |        |
        |     Config (Set CEID to CEID of CE#N)      |
     7  |<------------------------------------------|
     8a |<-------------------------------------------->|
        .                        .             .        .
     8x |<-------------------------------------------->|
```
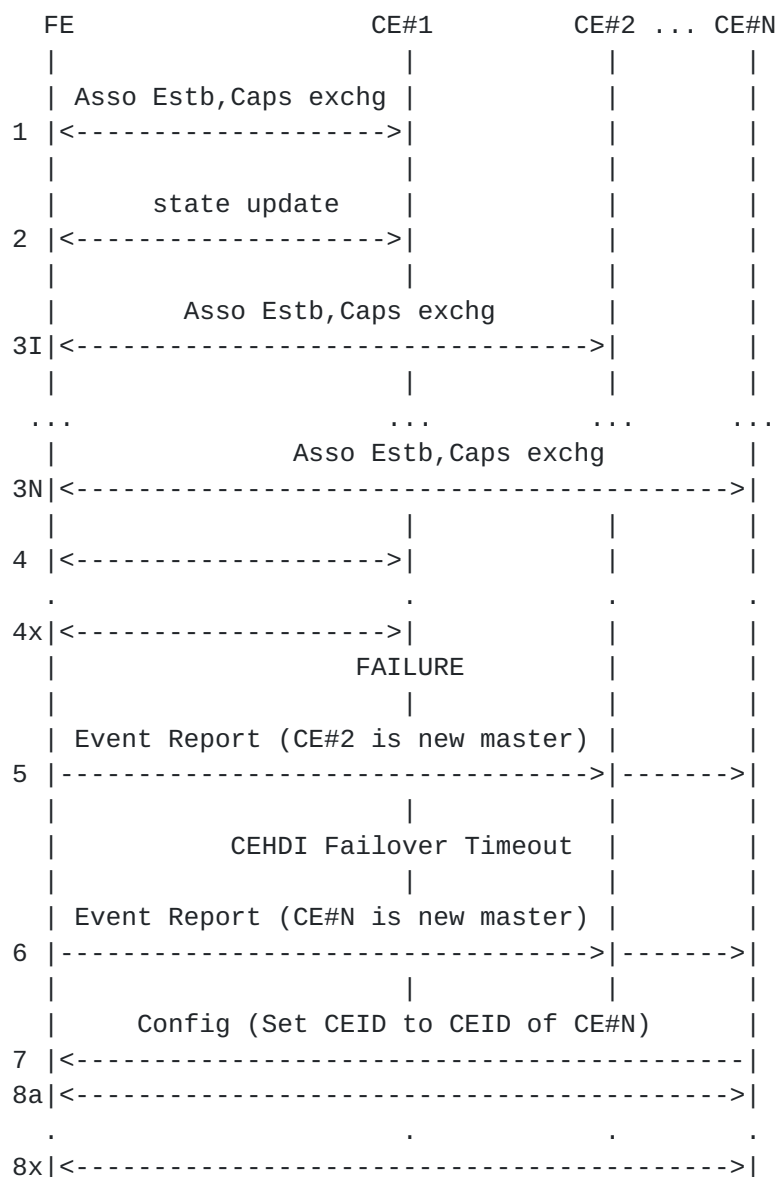
Figure 6: CE Failover for Hot Standby

If the FE does not get confirmation within the CEHDI Failover
Timeout, it picks the next CE on its list and advertises it as the
new master. Figure 6 illustrates in step #5 selecting CE#2 as its
new master. In step #6, the timeout occurs and it picks CE#N as its
new master. The FE receives confirmation that CE#N is the new master
in step #7.

If the CE the FE assumed to be the master discovers that it should
not be the new master CE, then it will configure the CEID with the ID
of the proper master CE. How the CE decides who the new master CE
is, is also out of scope of this document and is assumed to be done
via a CE-CE communication protocol. The FE must then associate with
then new CE.

If the CEFTI timer expires at either the not-associated or confirm
states without a new master CE confirmed, then the FE MUST revert to
the pre-association stage.

In most High Availability architectures there exists the possibility
of split-brain. However, since in our setup the FE will never accept
any configuration messages from any other than the master CE, we
consider the FE as fenced against data corruption from the other CEs
that consider themselves as the master. The split-brain issue
becomes mostly a CE-CE communication problem which is considered to
be out of scope.

By virtue of having multiple CE connections, the FE switchover to a
new master CE will be relatively much faster. The overall effect is
improving the NE recovery time in case of communication failure or
faults of the master CE. This satisfies the requirement we set to
achieve.

5. IANA Considerations

TBA

6. Security Considerations

TBA

7. References

7.1. Normative References

[RFC5810]   Doria, A., Hadi Salim, J., Haas, R., Khosravi, H., Wang,
            W., Dong, L., Gopal, R., and J. Halpern, "Forwarding and
            Control Element Separation (ForCES) Protocol
            Specification", RFC 5810, March 2010.

7.2. Informative References

[RFC3654]   Khosravi, H. and T. Anderson, "Requirements for Separation
            of IP Control and Forwarding", RFC 3654, November 2003.

[RFC3746]   Yang, L., Dantu, R., Anderson, T., and R. Gopal,
            "Forwarding and Control Element Separation (ForCES)
            Framework", RFC 3746, April 2004.

[RFC5812]   Halpern, J. and J. Hadi Salim, "Forwarding and Control
            Element Separation (ForCES) Forwarding Element Model",
            RFC 5812, March 2010.

1. Appendix I - New FEPO version

```
<LFBLibrary xmlns="urn:ietf:params:xml:ns:forces:lfbmodel:1.0"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" provides="FEPO">
   <dataTypeDefs>
      <dataTypeDef>
         <name>FEHBPolicyValues</name>
         <synopsis>
                        The possible values of FE heartbeat policy
               </synopsis>
         <atomic>
            <baseType>uchar</baseType>
            <specialValues>
               <specialValue value="0">
                  <name>FEHBPolicy0</name>
                  <synopsis>
            The FE heartbeat policy 0
         </synopsis>
               </specialValue>
               <specialValue value="1">
                  <name>FEHBPolicy1</name>
                  <synopsis>
            The FE heartbeat policy 1
          </synopsis>
               </specialValue>
            </specialValues>
         </atomic>
```

```
          </dataTypeDef>
          <dataTypeDef>
            <name>HAModeValues</name>
              <synopsis>
                 The possible values of HA modes
              </synopsis>
              <atomic>
              <baseType>uchar</baseType>
              <specialValues>
                 <specialValue value="0">
                   <name>NoHA</name>
                   <synopsis>
                        The FE is not running in HA mode
                   </synopsis>
                  </specialValue>
                 <specialValue value="1">
                   <name>ColdStandby</name>
                   <synopsis>
                        The FE is running in HA mode cold Standby
                   </synopsis>
                  </specialValue>
                 <specialValue value="2">
                   <name>HotStandby</name>
                   <synopsis>
                        The FE is running in HA mode hot Standby
                   </synopsis>
                  </specialValue>
               </specialValues>
               </atomic>
          </dataTypeDef>
          <dataTypeDef>
             <name>FERestartPolicyValues</name>
             <synopsis>
              The possible values of FE restart policy
             </synopsis>
             <atomic>
                <baseType>uchar</baseType>
                <specialValues>
                   <specialValue value="0">
                      <name>FERestartPolicy0</name>
                      <synopsis>
             The FE restart policy 0
          </synopsis>
                   </specialValue>
                </specialValues>
             </atomic>
          </dataTypeDef>
          <dataTypeDef>
```

```
          <name>CEHBPolicyValues</name>
          <synopsis>The possible values of CE heartbeat policy</synopsis>
          <atomic>
            <baseType>uchar</baseType>
            <specialValues>
              <specialValue value="0">
                <name>CEHBPolicy0</name>
                <synopsis>The CE heartbeat policy 0</synopsis>
              </specialValue>
              <specialValue value="1">
                <name>CEHBPolicy1</name>
                <synopsis>The CE heartbeat policy 1</synopsis>
              </specialValue>
            </specialValues>
          </atomic>
        </dataTypeDef>
        <dataTypeDef>
         <name>FEHACapab</name>
          <synopsis>
              The supported HA features
          </synopsis>
          <atomic>
            <baseType>uchar</baseType>
            <specialValues>
               <specialValue value="0">
                 <name>GracefullRestart</name>
                 <synopsis>
                       The FE supports Graceful Restart
                 </synopsis>
                 </specialValue>
               <specialValue value="1">
                  <name>HA</name>
                  <synopsis>
                        The FE supports cold-standby mode
                  </synopsis>
               </specialValue>
               <specialValue value="2">
                  <name>HOtStandBy</name>
                  <synopsis>
                        The FE supports hot-standby mode
                  </synopsis>
               </specialValue>
              </specialValues>
          </atomic>
        </dataTypeDef>
        <dataTypeDef>
          <name>CEStatusType</name>
           <synopsis>
```

```
                   Status values. Status for each CE.
             </synopsis>
            <atomic>
               <baseType>uchar</baseType>
               <specialValues>
                  <specialValue value="0">
                     <name>Disconnected</name>
                     <synopsis>
              No connection attempt with the CE yet.
             </synopsis>
                  </specialValue>
                  <specialValue value="1">
                     <name>Connected</name>
                     <synopsis>
                The FE has connected with the CE.
              </synopsis>
                  </specialValue>
                  <specialValue value="2">
                     <name>Associated</name>
                     <synopsis>
                The FE has associated with the CE.
             </synopsis>
                  </specialValue>
                  <specialValue value="3">
                     <name>Lost_Connection</name>
                     <synopsis>
                The FE was associated with the CE
                but lost the connection.
              </synopsis>
                  </specialValue>
                  <specialValue value="4">
                     <name>Unreachable</name>
                     <synopsis>
                The CE is deemed as unreachable by the FE.
              </synopsis>
                  </specialValue>
               </specialValues>
            </atomic>
         </dataTypeDef>
         <dataTypeDef>
            <name>AllCEType</name>
            <synopsis>
                Table Type for AllCE component.
             </synopsis>
            <struct>
               <component componentID="1">
                  <name>CEID</name>
                  <synopsis>ID of the CE</synopsis>
```

```
                <typeRef>uint32</typeRef>
            </component>
            <component componentID="2">
                <name>CEStatus</name>
                <synopsis>Status of the CE</synopsis>
                <typeRef>CEStatusType</typeRef>
            </component>
        </struct>
    </dataTypeDef>
</dataTypeDefs>
<LFBClassDefs>
    <LFBClassDef LFBClassID="2">
        <name>FEPO</name>
        <synopsis>
    The FE Protocol Object
    </synopsis>
        <version>2.0</version>
        <components>
            <component componentID="1" access="read-only">
                <name>CurrentRunningVersion</name>
                <synopsis>Currently running ForCES version</synopsis>
                <typeRef>u8</typeRef>
            </component>
            <component componentID="2" access="read-only">
                <name>FEID</name>
                <synopsis>Unicast FEID</synopsis>
                <typeRef>uint32</typeRef>
            </component>
            <component componentID="3" access="read-write">
                <name>MulticastFEIDs</name>
                <synopsis>
            the table of all multicast IDs
      </synopsis>
                <array type="variable-size">
                    <typeRef>uint32</typeRef>
                </array>
            </component>
            <component componentID="4" access="read-write">
                <name>CEHBPolicy</name>
                <synopsis>
        The CE Heartbeat Policy
      </synopsis>
                <typeRef>CEHBPolicyValues</typeRef>
            </component>
            <component componentID="5" access="read-write">
                <name>CEHDI</name>
                <synopsis>
        The CE Heartbeat Dead Interval in millisecs
```

```
            </synopsis>
                    <typeRef>uint32</typeRef>
            </component>
            <component componentID="6" access="read-write">
                <name>FEHBPolicy</name>
                <synopsis>
      The FE Heartbeat Policy
</synopsis>
                    <typeRef>FEHBPolicyValues</typeRef>
            </component>
            <component componentID="7" access="read-write">
                <name>FEHI</name>
                <synopsis>
      The FE Heartbeat Interval in millisecs
</synopsis>
                    <typeRef>uint32</typeRef>
            </component>
            <component componentID="8" access="read-write">
                <name>CEID</name>
                <synopsis>
       The Primary CE this FE is associated with
</synopsis>
                    <typeRef>uint32</typeRef>
            </component>
          <component componentID="9" access="read-write">
                <name>BackupCEs</name>
                 <synopsis>
                   The table of all backup CEs other than the primary
                 </synopsis>
                 <array type="variable-size">
                   <typeRef>uint32</typeRef>
                 </array>
          </component>
            <component componentID="10" access="read-write">
                <name>CEFailoverPolicy</name>
                <synopsis>
      The CE Failover Policy
</synopsis>
                    <typeRef>CEFailoverPolicyValues</typeRef>
            </component>
            <component componentID="11" access="read-write">
                <name>CEFTI</name>
                <synopsis>
      The CE Failover Timeout Interval in millisecs
</synopsis>
                    <typeRef>uint32</typeRef>
            </component>
            <component componentID="12" access="read-write">
```

```
                  <name>FERestartPolicy</name>
                  <synopsis>
            The FE Restart Policy
      </synopsis>
                  <typeRef>FERestartPolicyValues</typeRef>
            </component>
            <component componentID="13" access="read-write">
                  <name>LastCEID</name>
                  <synopsis>
          The Primary CE this FE was last associated with
      </synopsis>
                  <typeRef>uint32</typeRef>
            </component>
        <component componentID="14" access="read-only">
                  <name>AllCEs</name>
                  <synopsis>
            The table of all CEs.
       </synopsis>
                  <array type="variable-size">
                     <typeRef>AllCEType</typeRef>
                  </array>
              </component>
   <component componentID="15" access="read-write">
                        <name>HAMode</name>
                        <synopsis>
         Mode selection for action in HA after loss of master CE
       </synopsis>
                        <typeRef>HAModeValues</typeRef>
                </component>
   <component componentID="16" access="read-write">
       <name>AcceptBackupGets</name>
       <synopsis>If true, the FE will accept and respond to Queries
       from BackupCEs.</synopsis>
       <typeRef>Boolean</typeRef>
      </component>
     </components>
         <capabilities>
           <capability componentID="30">
              <name>SupportableVersions</name>
              <synopsis>
           the table of ForCES versions that FE supports
       </synopsis>
                  <array type="variable-size">
                     <typeRef>uchar</typeRef>
                  </array>
               </capability>
               <capability componentID="31">
                  <name>HACapabilities</name>
```

```
              <synopsis>
            the table of HA capabilities the FE supports
         </synopsis>
              <array type="variable-size">
                 <typeRef>FEHACapab</typeRef>
              </array>
           </capability>
        <capability componentID="32">
           <name>MaximumMultipleCEAssocations</name>
           <synopsis>
             The number of CEs this FE can associate with at the same
             time
           </synopsis>
           <atomic>
              <baseType>uint32</baseType>
           </atomic>
         </capability>
         </capabilities>
         <events baseID="61">
            <event eventID="1">
               <name>PrimaryCEDown</name>
               <synopsis>
            The pimary CE has changed
         </synopsis>
               <eventTarget>
                  <eventField>LastCEID</eventField>
               </eventTarget>
               <eventChanged/>
               <eventReports>
                  <eventReport>
                     <eventField>LastCEID</eventField>
                  </eventReport>
               </eventReports>
            </event>
            <event eventID="2">
               <name>HAPrimaryCEDown</name>
               <synopsis>The primary CE has changed</synopsis>
               <eventTarget>
                  <eventField>LastCEID</eventField>
               </eventTarget>
               <eventChanged/>
               <eventReports>
                  <eventReport>
                     <eventField>CEID</eventField>
                     <eventField>LastCEID</eventField>
                  </eventReport>
               </eventReports>
            </event>
```

```
              </events>
            </LFBClassDef>
         </LFBClassDefs>
      </LFBLibrary>
```

Authors' Addresses

    Kentaro Ogawa
    NTT Corporation
    3-9-11 Midori-cho
    Musashino-shi, Tokyo 180-8585
    Japan

    Email: ogawa.kentaro@lab.ntt.co.jp

    Weiming Wang
    Zhejiang Gongshang University
    149 Jiaogong Road
    Hangzhou 310035
    P.R.China

    Phone: +86-571-88057712
    Email: wmwang@mail.zjgsu.edu.cn

    Evangelos Haleplidis
    University of Patras
    Patras
    Greece

    Email: ehalep@ece.upatras.gr

    Jamal Hadi Salim
    Mojatatu Networks
    Ottawa, Ontario
    Canada

    Email: hadi@mojatatu.com