
ForCES Intra-NE High Availability draft 01

IETF – 80 Prague

Kentaro Ogawa (ogawa.kentaro@lab.ntt.co.jp)

Weiming Wang (wmwang@mail.zjgsu.edu.cn)

Evangelos Haleplidis (ehalep@ece.upatras.gr)

Jamal Hadi Salim (hadi@mojatatu.com)

Changes from 00 – FEPO model

- Required some changes in the FEPO LFB.
 - Addition of a new data type: Status (unsigned char)
 - 0 – Disconnected
 - 1 – Connected
 - 2 – Associated
 - 3 – Lost Connection
 - 4 – Unreachable
 - Change BackupCEs to AllCEs.
 - Now an Array of {CEIDs – Status}
 - Add two special values for CEFailover Policy
 - 2 – High Availability w/o Graceful Restart
 - 3 – High Availability with Graceful Restart.
-

Changes from 00 – CEHA init

- Added CEHA initialization steps
 - In pre-association
 - FEM configures AllCEs. AllCE IDs are in FCFS order.
 - FE will try to associate with the first CEID and set it as master.
 - If it fails it goes to the next.
 - In post-association
 - The master CE MAY change the list of AllCEs.
 - The master CE MAY instruct the FE to use a different master CE.
 - How the CEs communicate with each other and decide the order of BackupCEs. – out of scope.
-

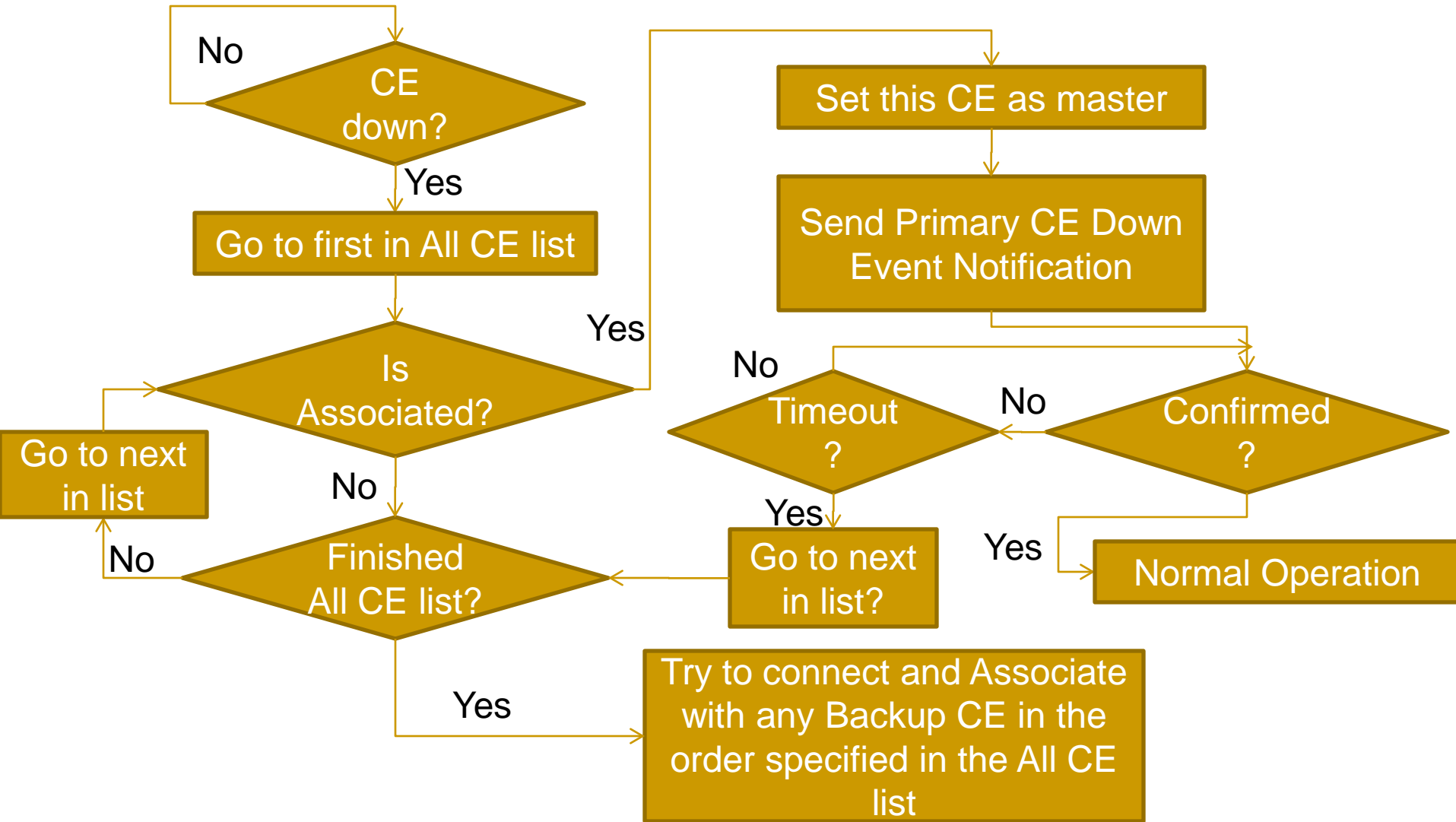
Changes from 00 – CEHA setup

- CEHA setup (CEFailover Policy set to 2 or 3)
 - FE will attempt to connect and associate with all CEs in the AllCE array in a FCFS fashion.
 - If it cannot connect to a CE the FE MAY flag it as unreachable.
-

Changes from 00 - CEHA

- CEHA (Master CE is considered down)
 - FE will search for an associated CE from the AllCE array.
 - It will select as master CE the FIRST CE ID it finds as associated.
 - Once a CE has been found, the FE sets it as master CE and sends an Primary CE down event to ALL associated CEs.
 - The temporary master CE MUST configure the CEID of the FE with a value. If it configures with its own CEID (same value) then the CE has accepted that it is the master CE and in the FE it becomes the master CE.

Changes from 00 – CEHA flowchart



Open issues

- What should happen to unreachable CEs in the AllCEs list when attempting to reconnect?
 - What happens if a CE doesn't respond within a FEPO Failover Timeout?
 - Split-brain - out of scope. Since this issue is mostly CE-CE issue and FE accepts messages only from master CE.
-

Interoperability test

- Presented approach tested within the Interoperability test.
 - FE & Backup CE from Greece – CEs from Japan & China.
 - Test was successful.
 - Interesting issue found
-

Interoperability test – Issue

- Bug in FE (added 1” delay in message handling).
 - HBs was set to 0.5”.
 - The master CE first processed messages from master CE and then from Backup CEs.
 - The master CE couldn't process messages from any Backup CE at all – so it couldn't successfully associate with a Backup CE.
-

Interoperability test – Issue

- Implementation issue.
 - Proposed solution:
 - While the FE is attempting to associate with Backup CEs, it checks per priority channel.
 - On all other occasions the FE handles messages first from master CE.
-