IEEE 802.1Q – Ethernet-cfm (IEEE 802.1ag)

CFM Enhancements

Karthik Babu H

CFM - Overview and Deployments

- IEEE 802.1ag (also known as connectivity fault management, CFM): A network-level OAM mechanism that provides Ethernet OAM functions, such as continuity check (CC), loopback (LB), and link trace (LT).
- CFM applies to large, end-to-end networks.
- Easy-to-use Ethernet techniques support good bandwidth expansibility on low-cost hardware. With these advantages, Ethernet services and structures have been widely used on

  enterprise networks,

  metropolitan area networks (MANs), and

  wide area networks (WANs)
- CFM is one of the widely deployed protocol in the large/complex scaled service provider networks and with the grow nature of these network there is also opportunity and limitation seen with the protocol.

CFM - Reaching the Limits – *CCM DB*

- Recently SP Customers produced a concern stating that with scaled MEPs(Maintenance end point) on the device and with some MEPs going down, they are concerned ccm-learning-database can grow indefinitely, possibly causing memory issues.

- Since the behaviour is static and the removal happens only after the fixed time even though the MEP is permanently removed from the network, we need to wait for 24 hours to get the cache cleared. This will be causing memory issues in some cases.

- More number of entries in a scaled network will be difficult to troubleshoot in case of any issue.

- AS per IEEE Std 802.1Q-2018 - An entry in the CCM Database is retained for at least 24 h after the last CCM with its triple is received and is removed from the MIP CCM Database after, at most, 48 h. If the resources allocated for the MIP CCM Database are not sufficient to maintain all of the entries for the required time, then the least-recently updated entries should preferentially be removed to make room for new entries.

CFM - Reaching the Limits - *Loop Identification*

- EVPN (RFC 7432) has been widely adopted in all industry platforms. When using an EVPN fabric, layer-2 loops can occur on the access-side of the fabric. Layer-2 loops can cause disruption by consuming available bandwidth, causing traffic drops, disabling users from communicating on the network, and degrading application performance.
- The occurrence of loops can happen due to migration use-cases, incorrect cabling, or mis-configuration on the access side of the fabric. Migration use-cases allow us to support introduction of EVPN VXLAN based networks so that a legacy Ethernet based layer 2 deployment can be gradually and smoothly replaced by a VXLAN overlay.
- In addition to migration use-cases, a network operator can inadvertently create a loop by either adding a new link or by adding an incomplete configuration (mis-configuration). For cases where the customer's access network is simple and does not run an STP, adding a new link can result in a loop.
- CFM is widely used in these deployments and can detect LOOPs but not identify the same

CFM - Reaching the Limits

- Ethernet CFM is a widely used low-level OAM protocol on provider networks. The protocols in use are based on a 1990's view of how to manage networks.

- Ethernet OAM protocols have not changed significantly in 20 years. CFM is complex to configure, requires deep protocol knowledge, and does not include integrated performance monitoring.

- There is an opportunity now to provide defining and delivering the next generation of Ethernet OAM to simplify deployment, include additional diagnostic information to accelerate problem resolution and to leverage telemetry and data analytics to improve the predictive capabilities of the OAM protocol implementation.

- Current complex manual configuration and manual actions used to determine root cause have been partially address by scripting, however the complexity has blocked wider use of OAM in enterprise markets.

Addressing Industry and Customer Needs

- Increasing network automation and shortening the time to issue resolution are very important to the industry itself at the moment.

- Reducing the level of expertise required to resolve issues is required given the massive scale of Ethernet networks. Predicting low-level network issues using embedded data analytics and using telemetry to provide detailed monitoring will help avoid issues that impact the end customer experience.

- Ethernet OAM will reduce customer's cost to deploy and manage large scale ethernet networks. Reduced cost and improved network quality would be the goal of the next generation deployments.

Proposal to enhance CFM

We would like to propose the following enhancements to address the problems that is being faced in the today's ever-growing complex networks,

- CCM Database Enhancement: Method for handling CCM database rapidly using efficient manner (Dynamic Stature).

- Identifying Redundant Link in a Layer-2 Loop using CFM.

- Stateful and efficient CFM Accelerates Ethernet Network Problem Resolution.

Proposal- CCM DB

- The mep-archive-holdtime is user configurable and even if not by default it is set to 100 minutes.

- But the ccm-learning-database entries deletion is not user configurable and it is taken care periodically on its own. When the ccm receiver gets initiated the CLDB (ccm learning database) timer is started by default. The default timer is set to 1 hour, after this gets expired it looks for the following condition to delete or keep the entries.

  Delete any entries that are older than the maximum age
  If it's less than the maximum age then we're keeping the entry around, send an update to the standby RP with the current timestamp value.

- The maximum age value is set to 24 hours, the age that CLDB entries must reach before we choose to delete them.

- It means after the last CCM received it takes a day for the entry to get cleared from the ccm-learning-database.

What - CCM DB

- All the basic technique's in caching is understood they are not required or applied for CFM CCM Database..

- This resolves key issues reported by the multiple SP customers who has a very complex networks and need help in optimising memory constraints.

- Avoiding stale entries and control overrun with protection is a must have feature for all protocols but this key element is currently missing in the CFM standard which we are addressing with this proposal. This needs to be published to the outside world for inter-operability.

- This enhances the IEEE 802.1ag/IEEE 802.1Q standard.

- In addition to using TTL, we also use a set of protocol specific events to manage entry lifetime (fast ageing). This significantly reduces the size of the database optimising the memory use. This fast-age explained in draft is being proposed to solve the issue mentioned in efficient and quick way.

How – CCM DB

- Making the CCM learning database entry deletion rapid in efficient manner with user configurable option. The maximum age value is set to 24 hours which is the main disadvantage and no user configurable option is provided.

- So now the proposed solution would be configuring the CCM learning database using CLI/models/controllers, the minimum value would be set to 1 hour(to maintain and keep track of the flapping CCM entries without re learning and wasting the resources) and the maximum configurable age would be 48 hours based on the user need with the default value being the 24 hours.

- Enable fast aging based upon events such as,

  when the Local MEP is removed/changed
  when the Link is permanently removed
  when the service instance for that port is deleted
  LC Failure in which this MEP is hosted
  the port in which the MEP is hosted is shut for some other reason.
  BDI is down and will not be made up.

Proposal– Loop Identification

- The main component of this proposal are

  (a) having a framework for automatic loop detection (based on triggered event) with provision for automatic remediation and generation of edge graph and

  (b) determining/identifying remote (node ,interface) which is causing layer-2 loop in network followed by subsequent mitigation.

- If a system already has an implementation for loop-detection, then our solution can be added as an input trigger to start the loop detection mechanism and upon loop detection, our proposed solution can help the existing implementation in terms of doing the correct mitigation.

- Wherever the CFM is deployed, we want to provide the customer with an extra service forward loop detection/identification without adding any new overhead to the networks.

Triggering Loop Detection

- Our proposed solution is to automate this behaviour and make them as an LTM periodic trigger to the remote MEP and store the last triggered sequence of path traversal in a LTM (Link Trace Message) database as a graph, which is not available right now.

- The Link trace message is currently triggered manually after an error is detected, this will take time to isolate the fault and carry out the remediation. We propose to have an automatic trigger of LTM based on the following events:

  CCM Loop trap
  Routing change
  Interface down/up
  Periodic-probes: User configured interval limit is met

- In any network where the SLA is more crucial, our proposal to identify and localize the redundant link within quickly will be a great value add for our customers.

Executive Summary

- Proposing a CFM protocol-based solution which dwells well with both legacy layer 2 and VXLAN overlay networks.
- Proposing loop identification feature to the existing CFM armoury
- Proposing automatic trigger of LTM to the existing CFM armoury
- Proposing a solution to maintain the new topology change/topology update to the CFM
- Proposing a strong convergence time of loop detection/identification in the range of 3.3ms
- Proposing continuous monitoring of links/performance of the links in the VXLAN overlay networks with added TLV to the existing CCM to measure the delay/jitter/latency in the links.

Deficiencies to Solve - CFM

- Heavy Configuration; Leads to tougher troubleshooting
- In-depth CFM Knowledge needed to debug the protocol
- Performance monitoring is a separate protocol driven by the IPSLA
- No telemetry; No Data and Analytics to predict the link bandwidth and fault
- CCM timeout is the way to tell the peer node that local node is dead which is a periodic update and not a triggered one, Fault verification and fault isolation are administrative actions, typically performed after fault detection.
- The receiving MP responds by transforming the LBM into a unicast Loopback Reply sent back to the originating MEP. That MEP records the responses for examination by the administrator.
- ALL THESE ACTIONS ARE PERFORMED BY THE ADMINISTARTORS MANUALLY OR THROUGH NMS SCRIPTS ONCE THE FAULT IS DETECTED

Deficiencies to Solve – Y.1731

- For a given one-way Ethernet frame delay measurement, frame delay and frame delay variation values are available only on the router that contains the receiver MEP.

- Single-ended ETH-LM is used for on-demand and proactive OAM. In this case, a MEP sends frames with ETH-LM request information to its peer MEP and receives frames with ETH-LM reply information from its peer MEP to carry out loss measurements.

- The main disadvantage is that the end points are not monitored for packet delay due to various congestion reason on continuous basis, whereas the packet loss/delay/delay variation are performed only when the administrator triggers on demand.

What

- Lightweight protocol with lesser config targeting enterprise market.
- Easier troubleshooting and debuggability with state machine illustrated.
- CFM does fault detection/verification/notification, now it adds fault device/link identification to its armory.
- Adding new fields to get the remote peer host name/port state/l3 Mgmt address will be easier for the administrator to find the issue.
- Monitoring the CFM MEPs doesn’t need any manual/script intervention and can be taken to the next level automation to perform the EEM actions based on the situation.
- Performance and the bandwidth can be on the control CCM packet itself and no need of separate protocol and probe for the same.
- Telemetry is used in the SRC MEP to collect the t1-t4 and do the performance monitoring on live wire.

How

Performance Monitoring using CCM

- Measure quality-of-service attributes such as frame delay/delay variation on a continuous basis and not “on demand", to achieve this we can use the existing CCM mechanism to determine the delay in the end point.

- Introducing a new TLV in CCM packet which carries the Timestamp of the transmitting node, receiving MEP process the CCM TLV with its Timestamp of its receiver and calculate the delay and delay variance based on the threshold configured.

- Identify problems before customers are impacted by network defects on a timely manner and the respective action is triggered for alternate option can be done without any impact.

How

CCM Packet Enhancement

We are introducing 2 new fields and 2 TLV in the existing CCM packet format.

- This TLV contains the CCM loopback port mac address, this will be swapped in the destination side and will be sent again to the source side.
- This TLV contains the Timestamp of the source(t1) to destination(t2) as well destination(t3) to source(t4).

Data Analytics and telemetry

- SRC MEP receives the looped CCM frames with all the t value (t1+t2+t3+t4).
- This data will be egressed out of the SRC MEP Using NETFLOW and will convert this data's into data frames.
- The data model prediction is based out on the data pattern from which we could find the delay/jitter/loss.

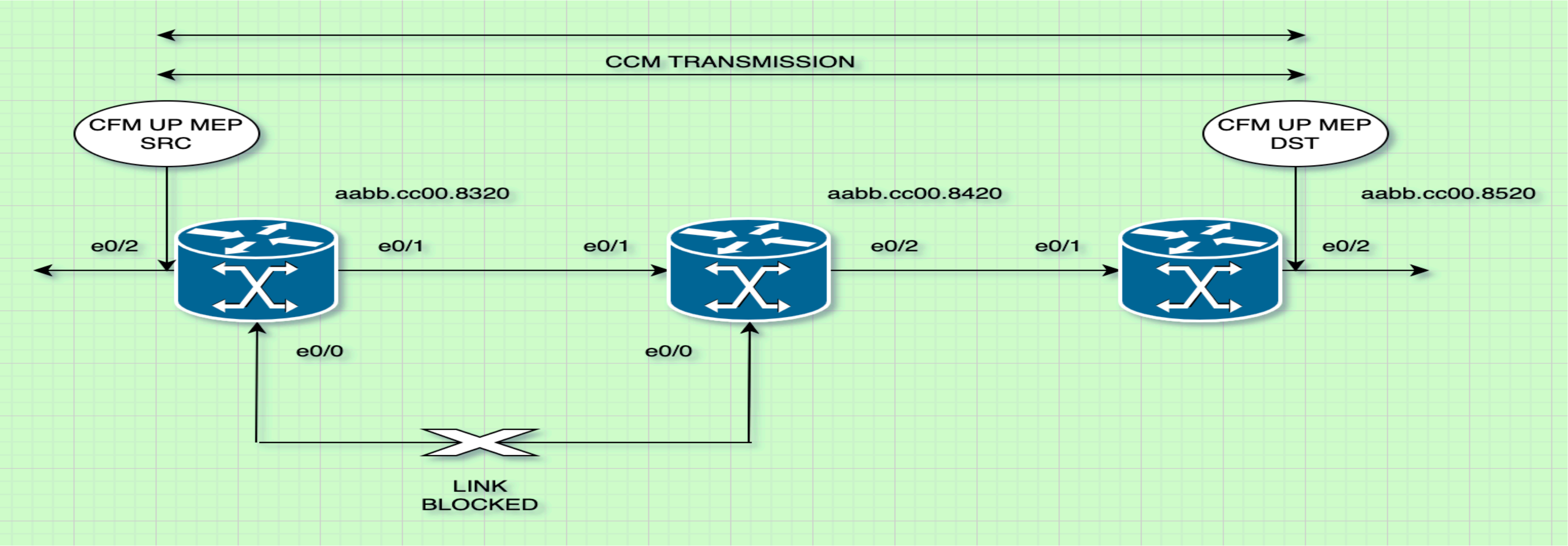
cisco

Back Up SLIDES

Karthik Babu H

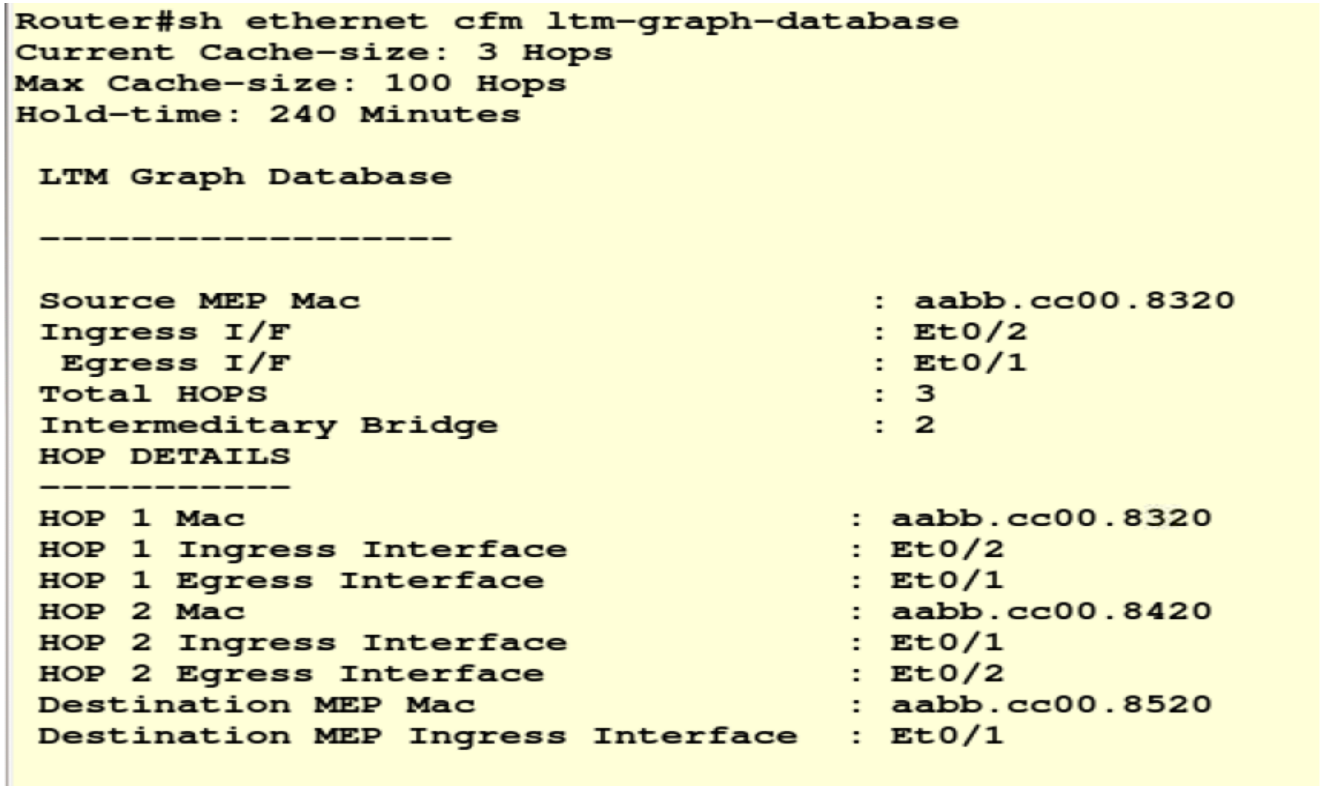
Prototype – Normal Initial Topology

Last update path traversal:
Device 1 -> Device 2-> Device 3



Prototype – LTM Graph Database

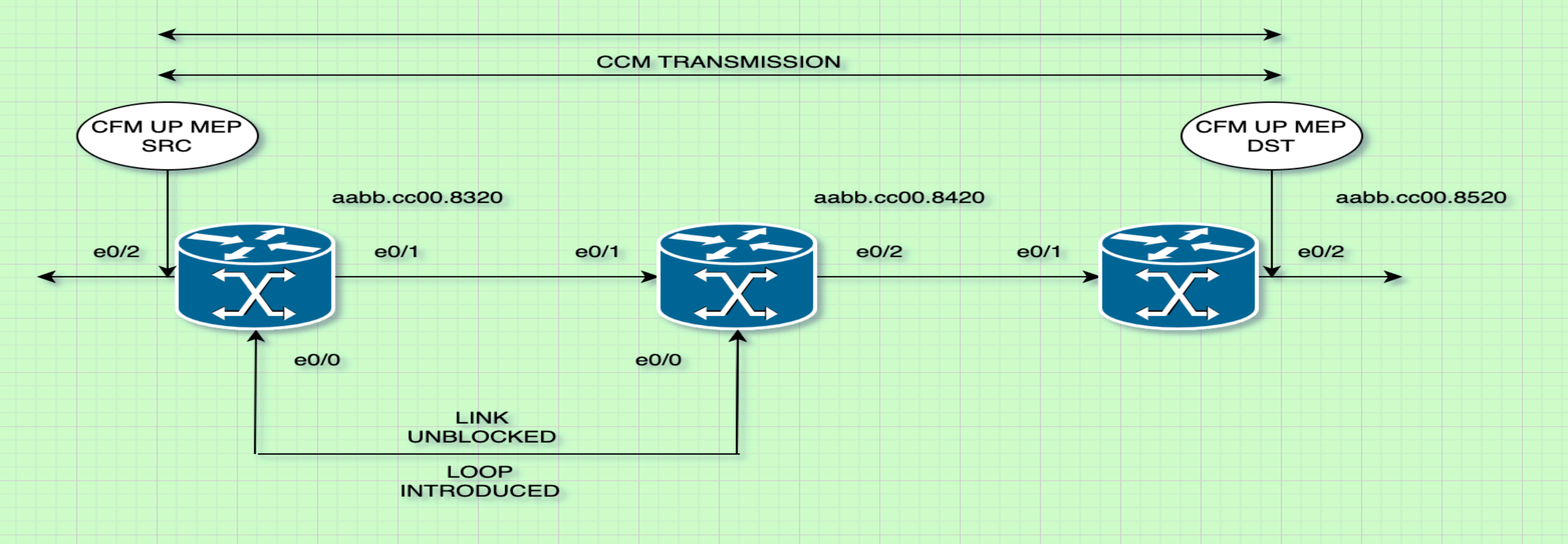
It took 3 hops(since it is an UP-CFM MEP) to successfully reach the destination MEP, the LTM database would be storing the graph of the path traversal in the order,



```
Router#sh ethernet cfm ltm-graph-database
Current Cache-size: 3 Hops
Max Cache-size: 100 Hops
Hold-time: 240 Minutes

 LTM Graph Database

 ------------------

 Source MEP Mac                               : aabb.cc00.8320
 Ingress I/F                                  : Et0/2
  Egress I/F                                  : Et0/1
 Total HOPS                                   : 3
 Intermeditary Bridge                         : 2
 HOP DETAILS
 -----------
 HOP 1 Mac                                   : aabb.cc00.8320
 HOP 1 Ingress Interface                     : Et0/2
 HOP 1 Egress Interface                      : Et0/1
 HOP 2 Mac                                   : aabb.cc00.8420
 HOP 2 Ingress Interface                     : Et0/1
 HOP 2 Egress Interface                      : Et0/2
 Destination MEP Mac                         : aabb.cc00.8520
 Destination MEP Ingress Interface           : Et0/1
```

Prototype – Introduce Loop

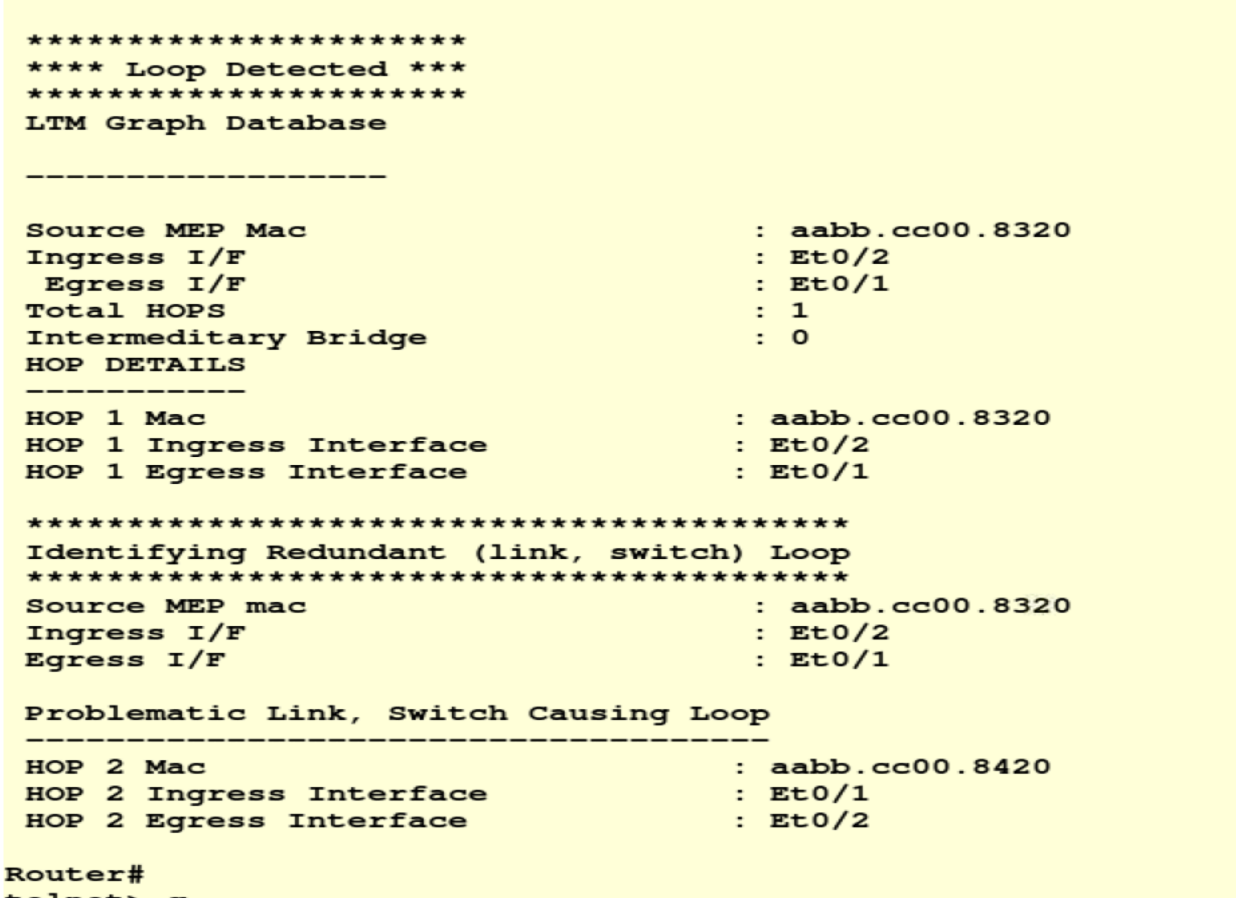
Device 1 -> <- Gets the same CCM which it has sent due to the redundant link and the loop caused during that.
The remote defect is marked with loop trap error from which our implementation takes control and act to resolve the issue automatically without any manual intervention.



Prototype – LTM Graph Database

Syslogs Logged:

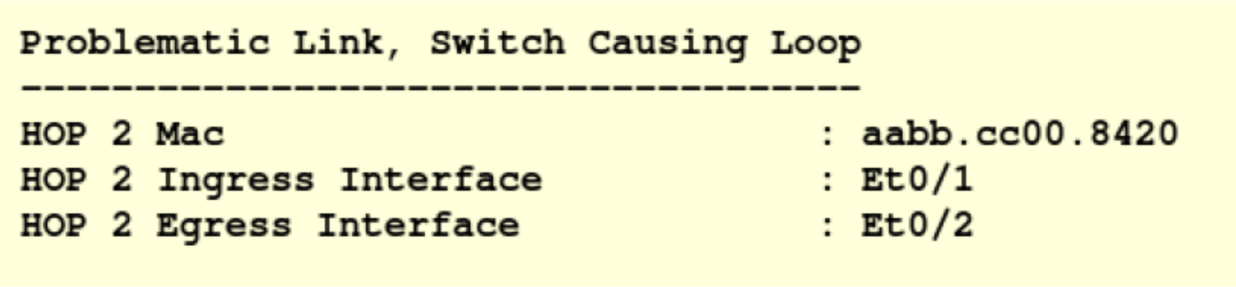
*Apr 12 15:47:20.119: %E_CFM-3-FORWARDING_LOOP: Continuity Check message is received with same source MAC aabb.cc00.8320 and mpid 8000 evc green vlan 2000 of its own in the MA name red.



```
 **********************
 **** Loop Detected ***
 **********************
 LTM Graph Database

 ------------------

 Source MEP Mac                                : aabb.cc00.8320
 Ingress I/F                                   : Et0/2
  Egress I/F                                   : Et0/1
 Total HOPS                                    : 1
 Intermeditary Bridge                          : 0
 HOP DETAILS
 -----------
 HOP 1 Mac                                    : aabb.cc00.8320
 HOP 1 Ingress Interface                      : Et0/2
 HOP 1 Egress Interface                       : Et0/1

 *****************************************
 Identifying Redundant (link, switch) Loop
 *****************************************
 Source MEP mac                                : aabb.cc00.8320
 Ingress I/F                                   : Et0/2
 Egress I/F                                    : Et0/1

 Problematic Link, Switch Causing Loop
 -------------------------------------
 HOP 2 Mac                                    : aabb.cc00.8420
 HOP 2 Ingress Interface                      : Et0/1
 HOP 2 Egress Interface                       : Et0/2

Router#
```

Prototype – Loop detection & pointing node, interface:

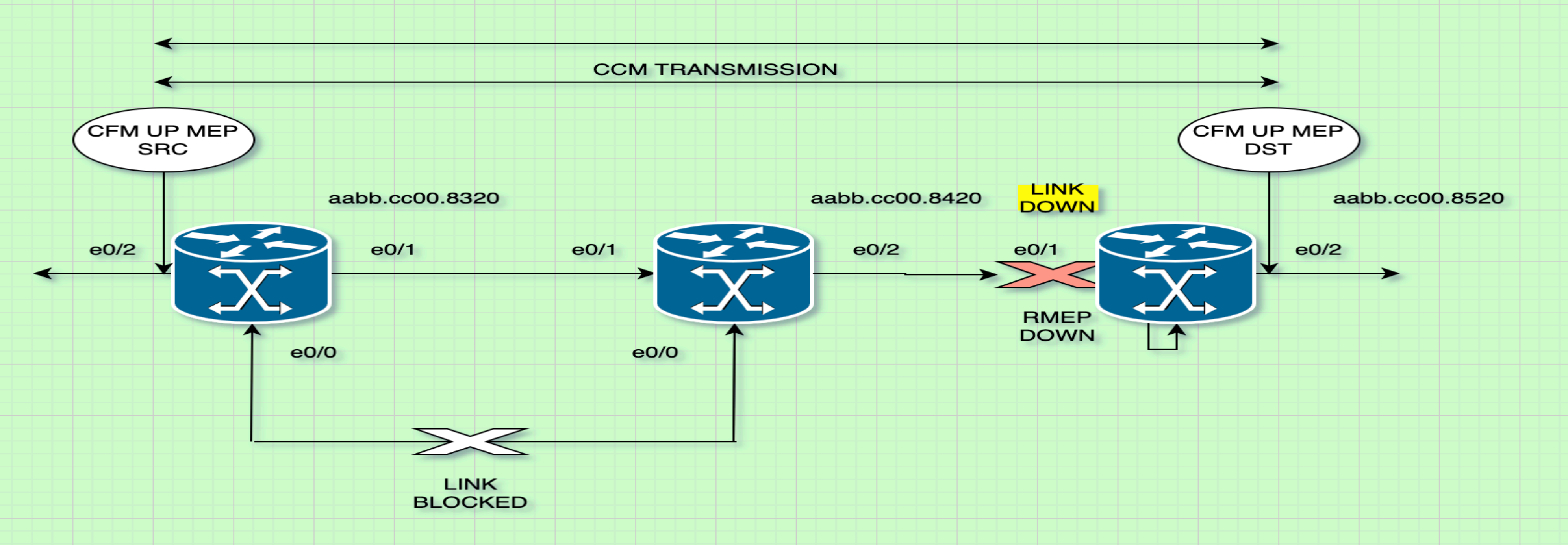
This LTM graph database is compared with the updated existing graph and the difference is then reported to network management to enable rapid remediation. So in the above case the fault is localised using the automatic LTM trigger and the following information is passed as the faulty device/link in the network.



```
Problematic Link, Switch Causing Loop
-------------------------------------
HOP 2 Mac                           : aabb.cc00.8420
HOP 2 Ingress Interface             : Et0/1
HOP 2 Egress Interface              : Et0/2
```

Prototype – Introduce Connectivity Issue (RMEP/HOP/Topology Change/Path Change)

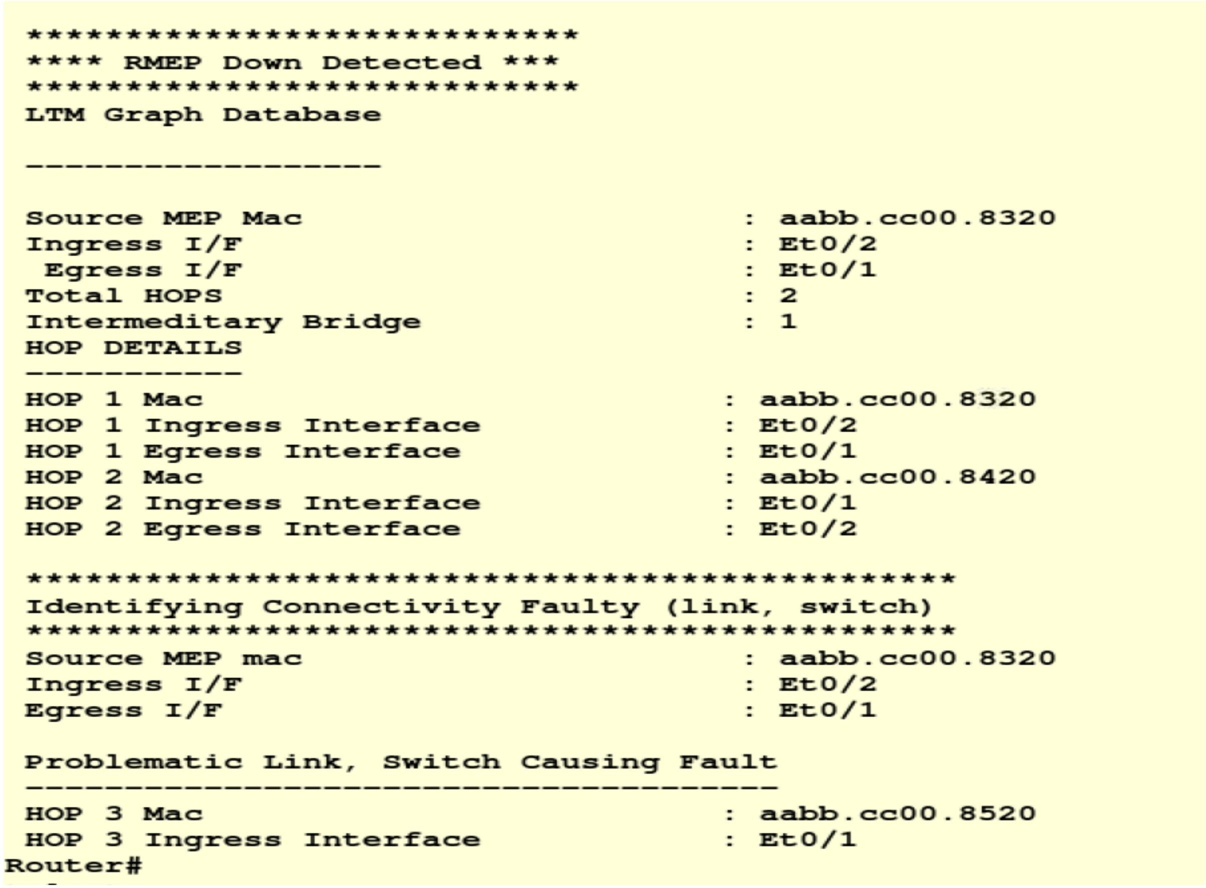
Device 3-> <- Shut the interface to introduce the connectivity fault by bringing the RMEP Down.
The remote defect is marked with DefRemoteCCM error from which our implementation takes control and act to resolve the issue automatically without any manual intervention.



Prototype – LTM Graph Database

Syslogs Logged:

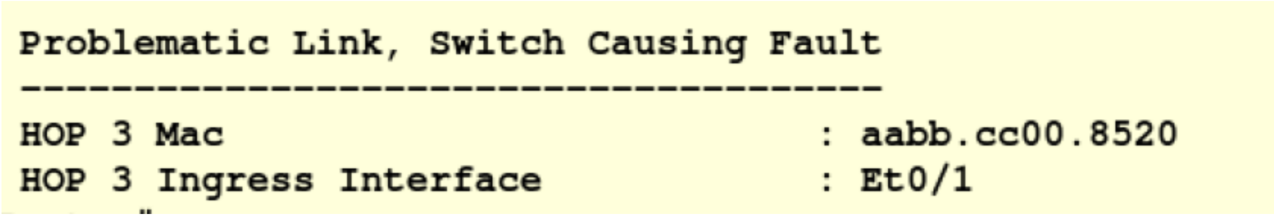
*Apr 12 15:35:28.132: %E_CFM-3-REMOTE_MEP_DOWN: Remote MEP mpid 8001 evc green vlan 2000 MA name red in domain kar changed state to down with event code TimeOut.



```
 ****************************
 **** RMEP Down Detected ***
 ****************************
 LTM Graph Database

 ------------------

 Source MEP Mac                       : aabb.cc00.8320
 Ingress I/F                          : Et0/2
  Egress I/F                          : Et0/1
 Total HOPS                           : 2
 Intermeditary Bridge                 : 1
 HOP DETAILS
 -----------
 HOP 1 Mac                           : aabb.cc00.8320
 HOP 1 Ingress Interface             : Et0/2
 HOP 1 Egress Interface              : Et0/1
 HOP 2 Mac                           : aabb.cc00.8420
 HOP 2 Ingress Interface             : Et0/1
 HOP 2 Egress Interface              : Et0/2

 ***********************************************
 Identifying Connectivity Faulty (link, switch)
 ***********************************************
 Source MEP mac                       : aabb.cc00.8320
 Ingress I/F                          : Et0/2
 Egress I/F                           : Et0/1

 Problematic Link, Switch Causing Fault
 --------------------------------------
 HOP 3 Mac                           : aabb.cc00.8520
 HOP 3 Ingress Interface             : Et0/1
Router#
```

Prototype – Connectivity detection & pointing node, interface:

This LTM graph database is compared with the updated existing graph and the difference is then reported to network management to enable rapid remediation. So in the above case the fault is localised using the automatic LTM trigger and the following information is passed as the faulty device/link in the network.



```
Problematic Link, Switch Causing Fault
--------------------------------------
HOP 3 Mac                            : aabb.cc00.8520
HOP 3 Ingress Interface              : Et0/1
```