802.1Qbp CFM

Ali Sajassi

July 20, 2011
IEEE Plenary Meeting

Agenda

- Network-level ECMP CFM
- Flow-level & Service-level ECMP CFM
- MEPs and MIPs Location

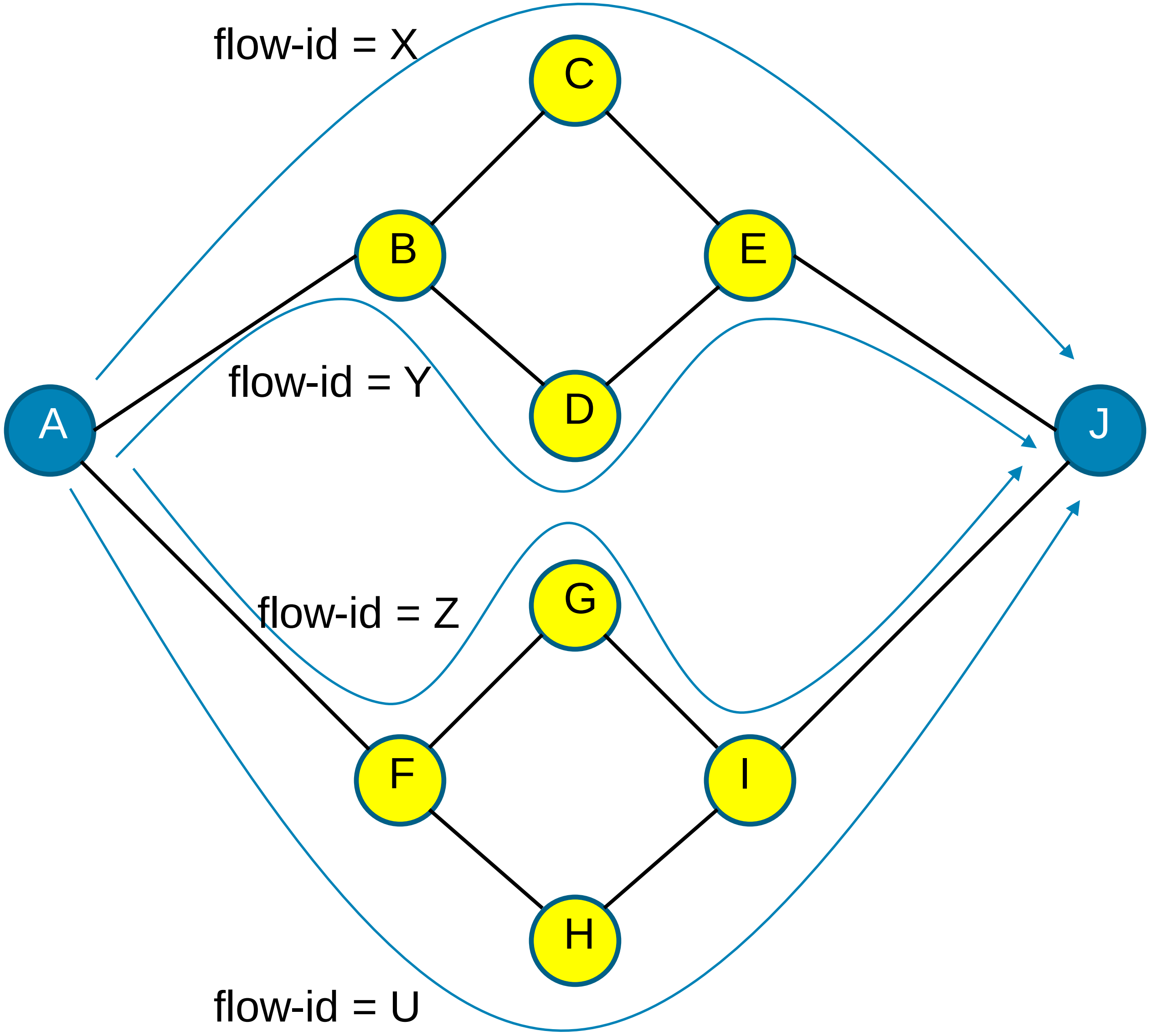
Terminology

- Flow-Level CFM: CFM functions performed on the user flows.
- Network-Level CFM: CFM functions performed on a test "VLAN" that covers the network. Test flows are chosen to exercise all ECMPs for the Test VLAN.
- Service-Level CFM: CFM functions performed on a service VLAN. Test flows are choses to exercise all ECMPs for that VLAN. It can be considered as a special case of Network-Level CFM.

Requirements for Network-Level CFM

- To verify ECMP coverage between two edge nodes
  - ➢ If N ECMPs exist between two nodes, then N test flows can be found – one test flow for each ECMP
- To verify each ECMP path
  - ➢ To verify that each test flow takes the path that originator node expects – to verify that hashing function works properly at each node
- To verify network-level connectivity for all ECMPs between a pair of nodes (e.g., edge nodes)
  - ➢ to verify connectivity for each ECMP

ECMP CFM: An Example



Fault-Monitoring (network-level)

1. Test Flow discovery: Node A needs to find four test flows (X, Y, Z, U) to cover the four possible ECMPs (one test flow for each ECMP)
2. ECMP Verification: Node A needs to verify that each test flow takes the path that it thinks it does
3. ECMP Fault Monitoring: Once Node A verify all the test flows, then it starts fault monitoring of ECMPs between the two edge nodes A and J
4. Continuous ECMP Verification: Node A can continue verification of ECMPs while performing ECMP fault monitoring

1. Test Flow Discovery

- Each edge node knows the topology of the entire network – e.g., node A knows how many ECMPs exist between self and B
- Since a uniform Hash(2) function is used in the entire network, An edge node (e.g., A) can predict the exact path of an ECMP for a given flow-id
- An edge node can identify N flow-ids for N ECMPs (e.g., using brute force) between self and the destination edge node

2. ECMP Verification

- Now that an edge node discovers N flow-ids to cover all N ECMPs between self and the target destination node, it needs to verify each ECMP to ensure that the actual path is the same as it thinks it is

- FlowTrace is used for such verification
  - ➢ in the above example, Node A wants to verify path [A,B,C,E,J] using flow-id X
  - ➢ Node A sends a FlowTrace: target node (J) and flow-id X
  - ➢ FlowTrace uses TTL=1 for stopping at each hop
  - ➢ FlowTrace carried a flag indicating that the response should only be received from target node
  - ➢ FlowTrace traverses hop-by-hop through nodes B, C, E, and J
  - ➢ At each node path information is recorded in the payload (e.g., node-id, ingress i/f, egress i/f)
  - ➢ Target node J sends a response to node A with path info [A,B, C, E, J]. Node A compares and verifies the path info.

3. ECMP Fault Monitoring

- The receiving node (J), upon receiving FlowTrace in the verification step, initializes its state accordingly to receive CCMs from node A (and sends a response back to A)
- Once the edge node (A) verifies a given ECMP (e.g., the actual path for that test flow is the same as what A thinks it is), then A can start fault monitoring for that ECMP by sending periodic CCM
- For ECMPs that node A doesn’t receive a response, it can try again and/or raise an alarm and/or initiate fault isolation procedures

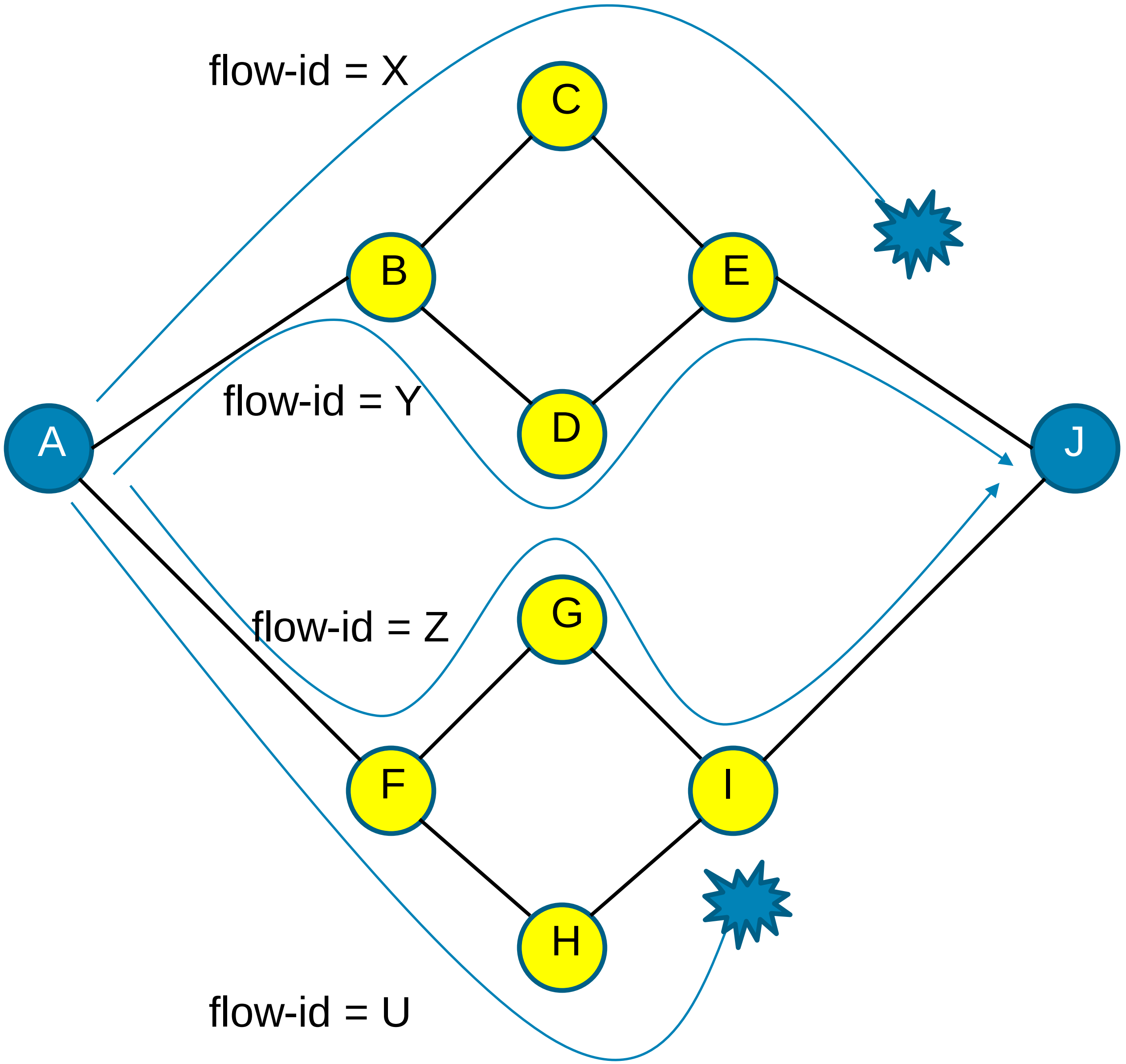
4. Continuous ECMP Verification

- Once the originating edge node (A) starts period CCMs for all its ECMPs toward the target node (J), then it can continuously verify each ECMP by repeating the procedure in step (2)
- The timer for CCMs transmission is different than the timer for transmitting period FlowTrace, thus the frequency of each can be adjusted independently

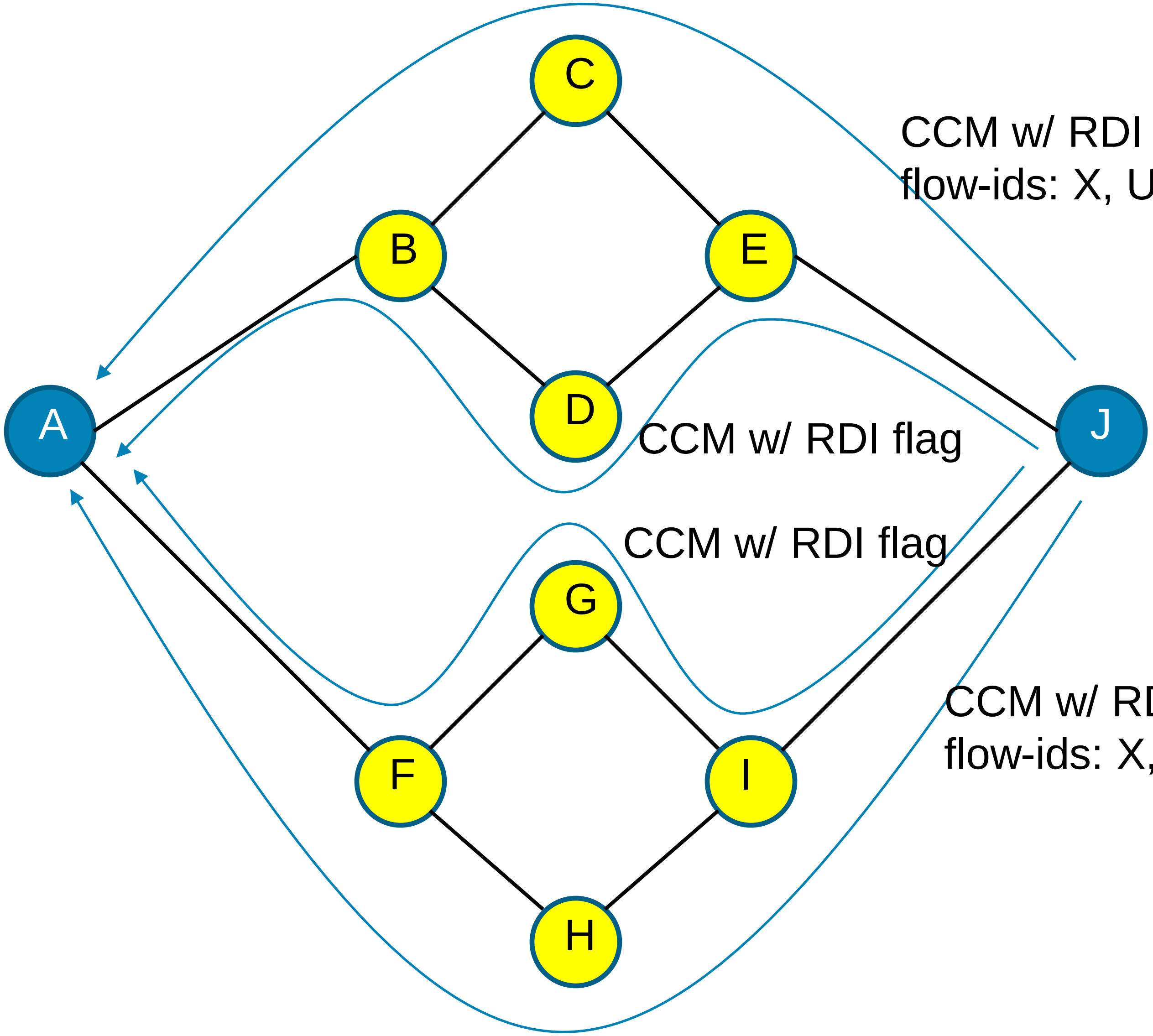
Fault Isolation

- During periodic transmission of flow CCMs from A to J, if J doesn’t receive some of the test flows, it sends an RDI to A, listing flow-ids that are not received by J
  - J can send this list on every ECMP RDI or a single EMCP RDI (going out next) or on some of them
- A can subsequently start fault isolation procedures for these test flows by
  - sending FlowTrace hop-by-hop and expecting to receive a response from each hop
  - If A doesn’t receive a response from a given hop, then it know the failure is in the vicinity of that hop

Fault Isolation



Fault Isolation



CCM w/ RDI flag:
flow-ids: X, U

CCM w/ RDI flag

CCM w/ RDI flag

CCM w/ RDI flag:
flow-ids: X, U

Agenda

- Network-level ECMP CFM
- Flow-level & Service-level ECMP CFM
- MEPs and MIPs Location

Requirements – Cont.

- To perform 802.1ag comparable CFM on a per flow basis
- To perform 802.1ag comparable CFM on a per service level (I-SID level) basis

Flow-Level CFM

- User supplies flow information, including one or more of:
  - MAC SA and/or DA
  - IP Src and/or Dst
  - Src and/or Dst Port (TCP or UDP)
- Flow parameters are converted to a flow ID (e.g., NMS can query platform using flow parameters and get back flow ID)
- MEP monitors the flow by sending periodic CCMs for that flow.
  - Monitoring of unicast flows uses unicast CCMs
  - Monitoring of multicast flows uses multicast CCMs

Flow-Level Fault Detection

- Per flow CCM is used to detect a fault
- MEPs are configured at the end points but MIPs don’t need to be explicitly configured
- Fault detection is always in the direction of the received traffic – e.g., not receiving a CCM from remote bridge doesn’t mean there is a problem with the reception of your CCM
- When a loss of CCM is detected, RDI flag is set to notify the remote bridge that its CCM was not received and there is a problem on the path from remote bridge to self (flow-id is conveyed to the originating side along with RDI flag)
- Remote bridge can initiate fault isolation procedure if needed

NOTE: Fault Isolation procedure needs to start from the remote bridge because of the non-congruent nature of forward/reverse traffic

Flow-Level Fault Isolation

- “Flow Trace” is used for fault isolation
- It works similar to “Link Trace”
  - A singel FTM is generated from MEP and traverses the network hop by hop. Each hop generates a response to the originating MEP
  - FTM is generated with the proper flow-id and TTL=1
  - Because of TTL=1, the message is stopped in the next hop and a respond is generated (FTR) with similar info as current LTR
  - Next hop sets the TTL=1 and forwards the message to its next hop with the same B-MAC SA and B-MAC DA.

Servic-Level CFM

- A MEP, knowing the topology and how to exercise the ECMPs, first calculates the necessary Test Flows for full coverage of all paths in a given service instance.
  - At this point brute force method is used to generate test flows until someone comes with better scheme
- On a per service instance basis, MEPs perform monitoring of all unicast and multicast paths using the Test Flows.
- MEPs follow a ‘round-robin of Test Flows’ scheme to verify connectivity over all ECMP paths (unicast) and shared trees (multicast).

  –Round-robin scheduling reduces processing burden on nodes, and modulates the volume of OAM messaging over the network.

  – Comes at the expense of relatively longer fault detection time

  – For critical flows, it is possible to schedule their connectivity check continuously.

  – MEP CCDB will track every flow independently (timer per flow per remote MEP rather than per remote MEP in CFM)

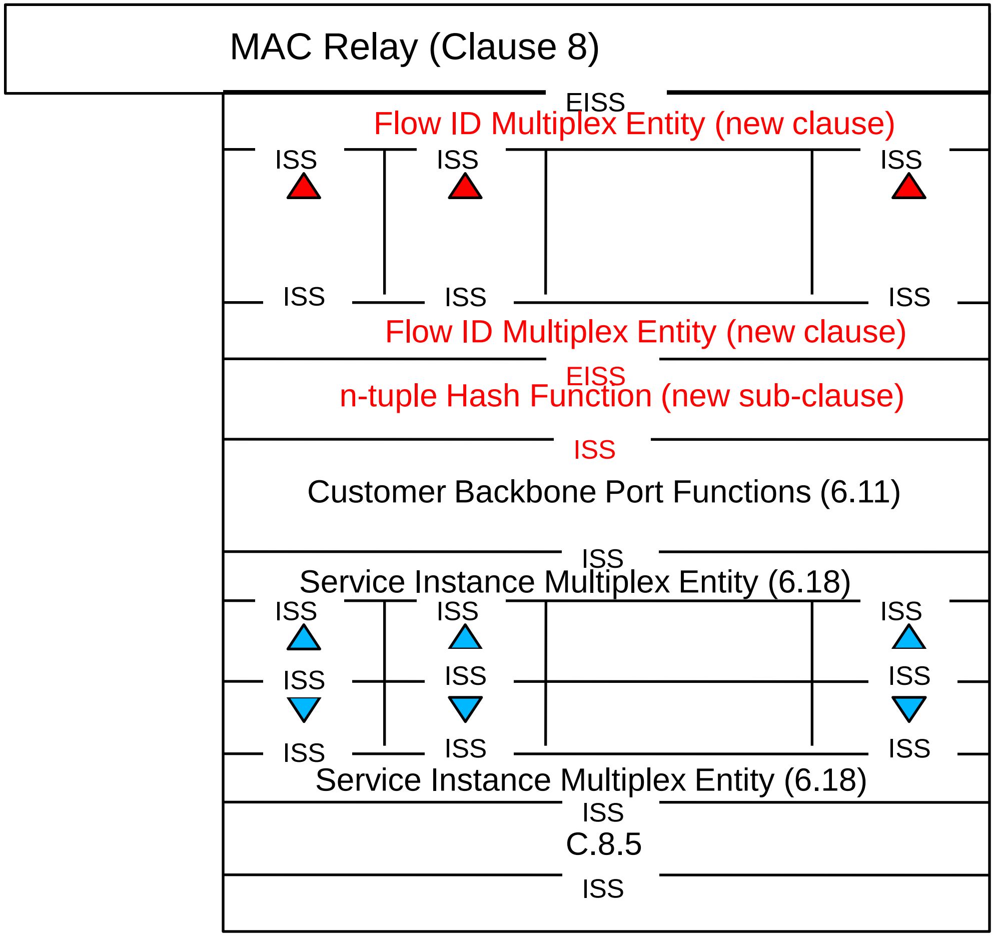
Service-Level CFM – Cont.

- In order to do proper fault suppression during topology change, once a network topology change is detected, a flag is set on the CCM message to notify the far end not to generate faults as the result of missing CCMs till the network settles down
- Once the network settles down and new test flows are generated, then the flag is cleared

Agenda

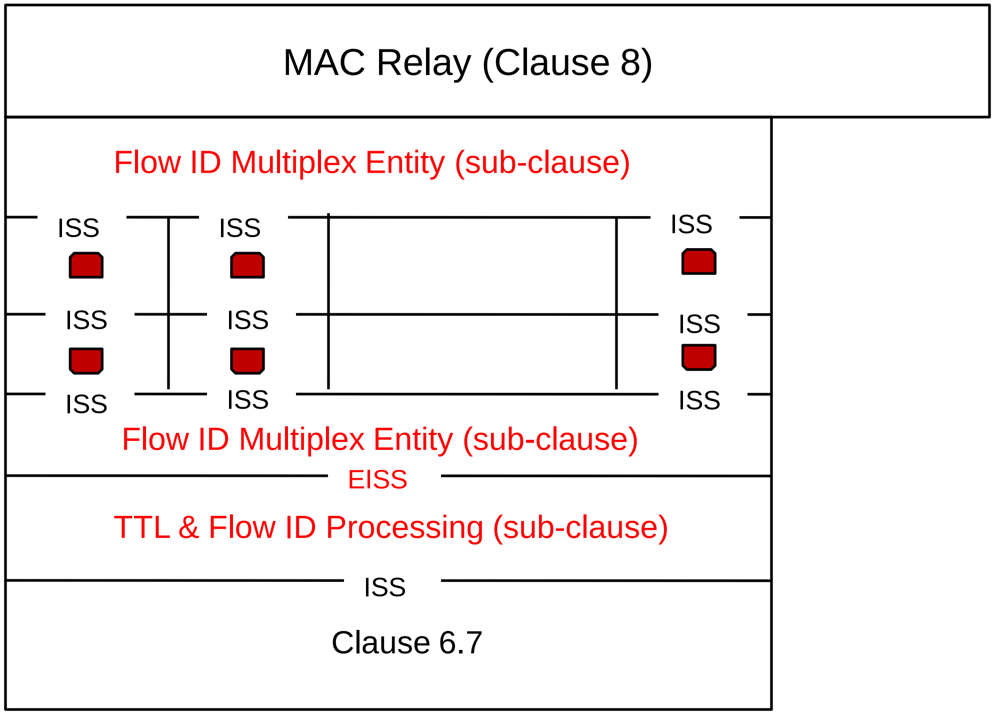
- Network-level ECMP CFM
- Flow-level & Service-level ECMP CFM
- MEPs and MIPs Location

Baggy Pants Model for OAM operation at BEB

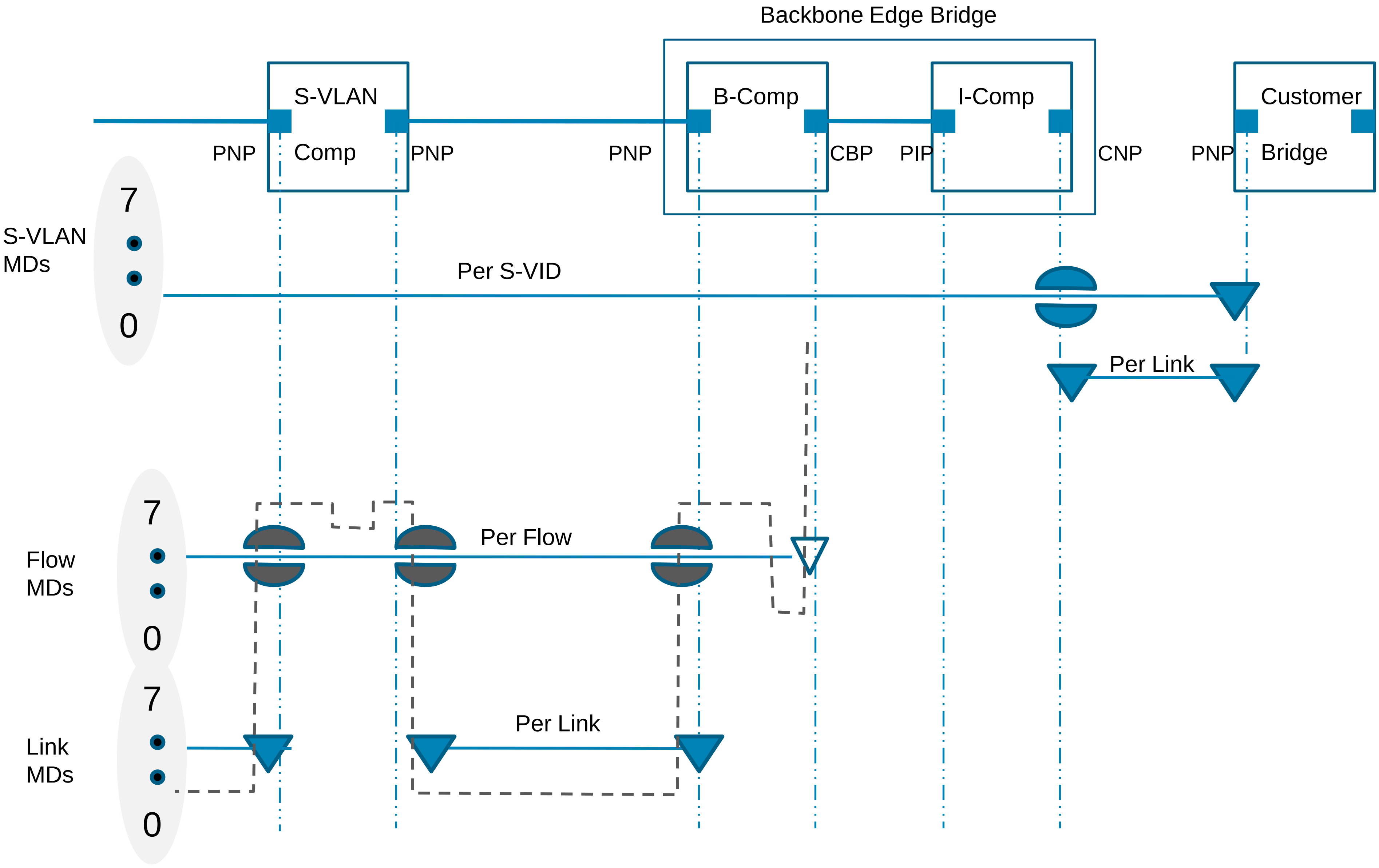

MAC Relay (Clause 8)
EISS
Flow ID Multiplex Entity (new clause)
ISS
ISS
ISS
ISS
ISS
ISS
Flow ID Multiplex Entity (new clause)
EISS
n-tuple Hash Function (new sub-clause)
ISS
Customer Backbone Port Functions (6.11)
ISS
Service Instance Multiplex Entity (6.18)
ISS
ISS
ISS
ISS
ISS
ISS
ISS
ISS
ISS
Service Instance Multiplex Entity (6.18)
ISS
C.8.5
ISS

NOTE: Clause 6.11 needs to be modified to indicate that all ECMP I-SIDs are mapped to a single default B-VID

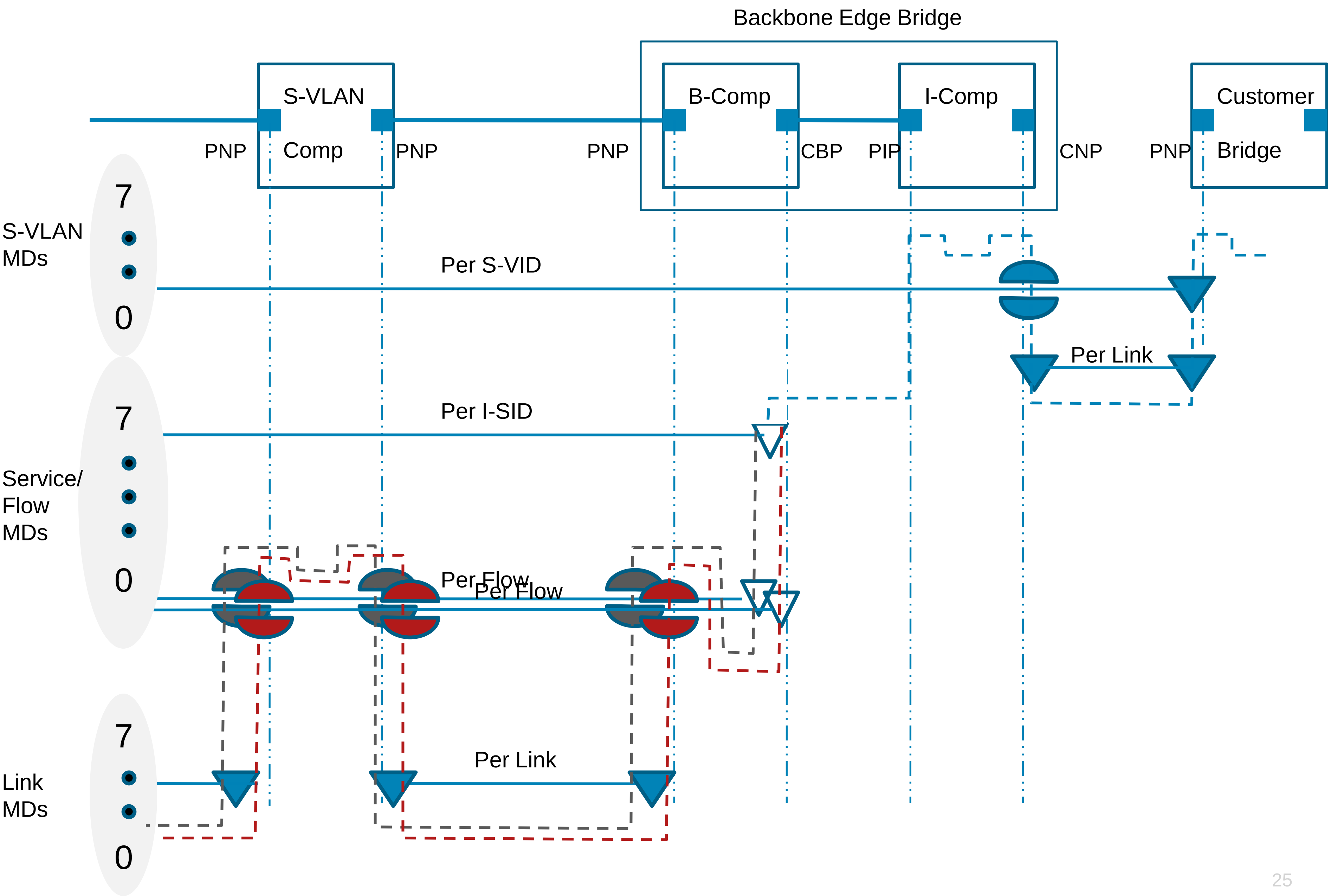
Baggy Pants Diagram for OAM operation at BCB


MAC Relay (Clause 8)
Flow ID Multiplex Entity (sub-clause)
ISS
ISS
ISS
ISS
ISS
ISS
ISS
ISS
ISS
Flow ID Multiplex Entity (sub-clause)
EISS
TTL & Flow ID Processing (sub-clause)
ISS
Clause 6.7

Flow-Level CFM



Service-Level CFM



CISCO