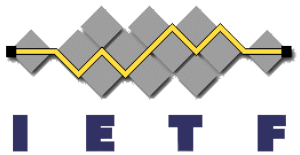
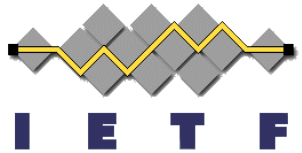
BFD on LAG Interfaces

Manav Bhatia, Alcatel-Lucent
Mach Chen, Huawei
Sami Boutros, Cisco
Marc Binderberger, Cisco
Jeff Haas, Juniper



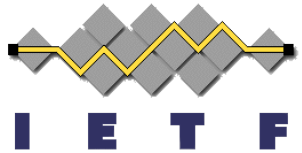
IETF 83, Paris

Current State of BFD on LAGs (1/2)



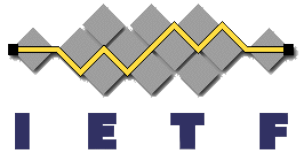
- Single BFD session over the whole LAG
  - Member Link Failures may not get detected – depends upon how the single session is implemented
  - Usually link failures NOT detected till LACP times out
  - All traffic hashed over that link gets dropped
- Can use Connectivity Fault Management (CFM) for link monitoring, but there are environments where BFD is already in use and customers want to extend BFD for this purpose

Current State of BFD on LAGs (2/2)



- Some vendors already have proprietary BFD over LAG implementations where BFD is running on all member links
- Operators want a standardized solution to run BFD over LAGs

Current Proposal (1/3)

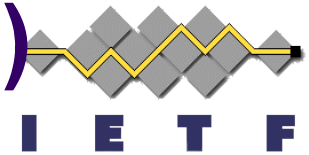


- Run BFD on each member link of the LAG – micro BFD session
- Each micro BFD session is an independent BFD session following the same procedures as defined in 5880/5881
- The BFD control packets are IP/UDP encapsulated with a new UDP destination Port
- The LAG Manager is a client of BFD, i.e., it requests micro sessions to be established per member link

Current Proposal (2/3)

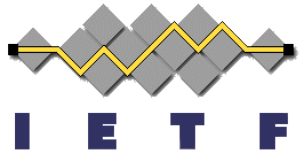
- Micro BFD session for a particular port MUST be requested when a port is attached to an aggregator
  - Session must be deleted when the port is detached from the aggregator
- When the micro BFD session goes down the member link MUST be taken out of the traffic load balance table

Current Proposal – With LACP (3/3)



- Micro BFD session started when link is brought into the LACP Distributing state
  - Although the link is in LACP Distributing state, it is not used for carrying any traffic other than the BFD control packets for the micro session
  - Only when BFD goes up is the link used for carrying general traffic
  - Scheme works – but couples the two protocols (!)
- Link is removed from the active set when the micro session goes down

Next Steps (1/2)



- The current proposal modifies the LACP state machine. We should avoid this. Proposal:
  - LACP and BFD run independently
  - Link can not be in forwarding use when BFD is not in Up state
  - This allows identical rules with/without LACP
- Current draft too specific LACP/LAG?
  - How do we deal with BFD over MPLS component links and other multilink setups?

Next Steps (2/2)

- The other proposal is to get micro BFD sessions to only influence the L3 load balancer
  - The link is added or removed from the L3 traffic load balancer based on the status of the micro BFD sessions