January 08, 2019

PROXY MONITORING OF SESSION STATE FOR SHARED FATE ENTITIES

Stefan Olofsson

Khalid Raza

PROXY MONITORING OF SESSION STATE FOR SHARED FATE ENTITIES

AUTHORS:
Stefan Olofsson
Khalid Raza

ABSTRACT

By using a general protocol that allows for sharing of bidirectional forwarding detection (BFD) state across a shared path between neighboring routers, scale, simplicity and flexibility in designs may be achieved with at least a 2:1 ratio compared to a traditional implementation. Each participating node hosts BFD sessions that monitor a locally attached link and then shares the state of the BFD sessions to other neighboring nodes, effectively proxying the BFD state for consumption by participating devices in an efficient manner.

DETAILED DESCRIPTION

BFD is used as a vehicle for failure detection in different network environments, providing either single-hop or multi-hop path monitoring. In the SD-WAN (Viptela) platform, the multi-hop capability is used on every WAN-facing connection to ensure liveness of remote endpoints across each of the connected underlay networks. This approach works well, but is challenged from a scalability perspective.

The techniques presented herein enable the use of a single BFD session to monitor an arbitrary number of BFD sessions hosted by a neighboring device, and hence, aids scalability by an initially assessed factor of 2:1, from the current situation. Any architecture deploying a similar architecture, whether based on BFD or other mechanisms will be able to enjoy a similar benefit, given that a pair of devices share certain common attributes and sources of information.

An example of an existing implementation and use of BFD is provided to build context. Furthermore, the challenges associated with the existing approach are discussed as well as how the described techniques benefit the sample application as well as other general use cases where the described techniques may be used.

Published by Technical Disclosure Commons, 2019

In an existing application, an SD-WAN site equipped with a redundant pair of devices can be configured to use a feature known as a Transport Locator (TLOC) extension. In such a site, a pair of devices are interconnected via a logical or physical link. Each device is also connected to a minimum of one underlay network, typically with different underlay networks attached to each node. Each underlay connection is terminated on what is referred to as a TLOC-interface, which is a combination of a physical interface and a logical SD-WAN interface behavior profile. Assume that node A is connected to MPLS and node B is connected to the Internet (see, FIG. 1). The MPLS and Internet TLOCs respectively are then logically extended across the interconnection link, such that each device is actively connected to both underlays.

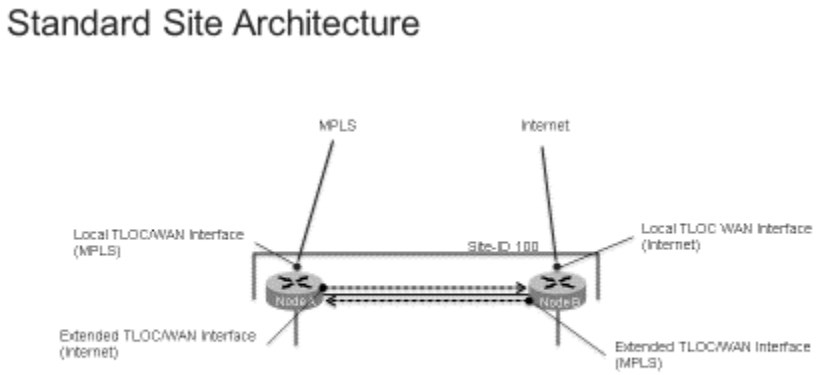

FIG. 1

The TLOC interface function involves establishing BFD sessions to other known TLOCs in the network for overlay data plane establishment (see, FIG. 2). This behavior is identical on both devices for both TLOC interfaces. In this configuration, both node A and node B have active MPLS and Internet connections and are independently establishing BFD sessions on both interfaces. The BFD sessions to the remote devices reachable across the extended TLOC through the peer device, are simply forwarded by that device without any processing or other intelligence being applied. This is a significant scalability impairment. The remote TLOCs are learnt via the Overlay Management Protocol (OMP), a path vector protocol designed to share information about

other devices in the overlay, their WAN connections, VRF configurations and routing for said VRFs. Hence, each of the nodes has an identical OMP TLOC table being kept up to date via communication with the routing controller.

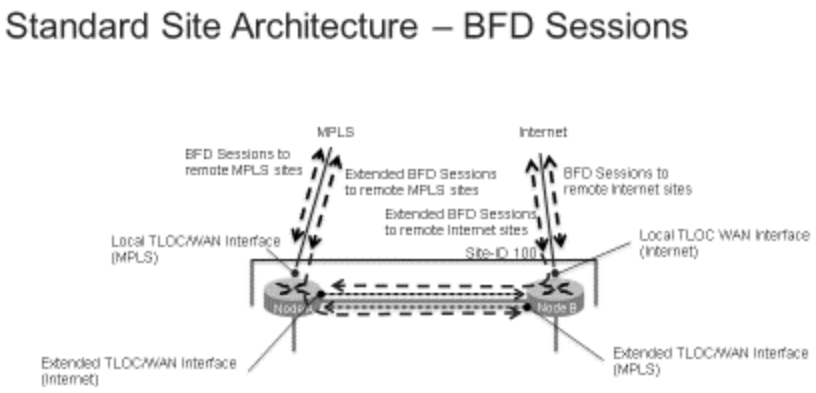

FIG. 2

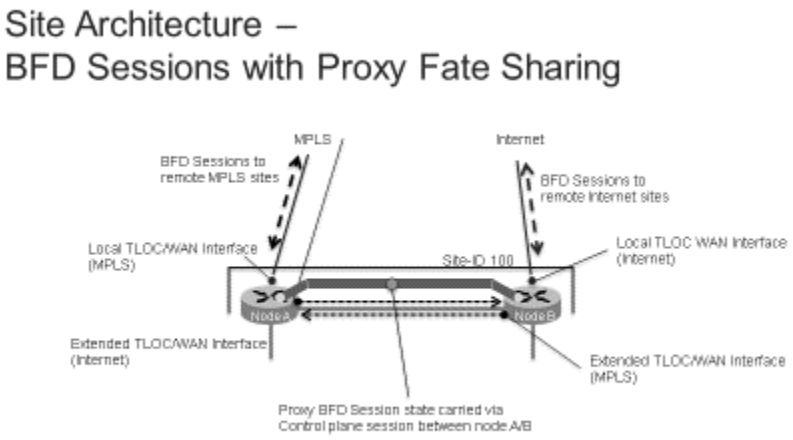

FIG. 3

Applying the present techniques to this environment, the following differences would apply:

1) A control plane session is established between node A and node B, this could be BFD, OMP or another suitable protocol (see, FIG. 3).

2) The extended TLOC connection on each node facing the other node does not establish any BFD sessions specific to the extended TLOC.

3) The protocol session between the nodes is used to relay a table of TLOCs that are down from the perspective of each node. In this case, if a remote TLOC is down on the node owning the underlay connection, it would always be down on the node with the extended TLOC connection as well. Upon any change in the state of a TLOC, an update message is generated from the node owning the BFD session towards the neighboring node, allowing the neighbor to alter its local state for the path, reflecting the actual state as determined via BFD by the originating node.

This approach has the following immediate benefits:

1) Each node can now easily scale to support twice the amount of monitored TLOCs. For example, one set of TLOCs may be directly monitored via locally driven BFD-sessions, while the other set may be monitored via the directed protocol session with its neighbor.

2) A given node could now support additional TLOC connections via partner nodes without any scaling implications. This would potentially yield scalability far beyond the 2:1 discussed above as a set of extended TLOCs would be monitored via the control plane session(s) to one or more neighboring nodes.

Any applications reliant on multi-hop BFD would benefit from these techniques as well.

In summary, by using a general protocol that allows for sharing of bidirectional forwarding detection (BFD) state across a shared path between neighboring routers, scale, simplicity and flexibility in designs may be achieved with at least a 2:1 ratio compared to a traditional implementation. Each participating node only hosts BFD sessions that monitor a locally attached link and then shares the state of the BFD sessions to other neighboring nodes, effectively proxying the BFD state for consumption by participating devices in an efficient manner.