MSRP V2 Extended Scheme for Multiple Path Selection and Redundancy

Philippe Klein, PhD
Broadcom Corporation

V.01	Nov 2011	Initial Version
V.02	Jan 2012	
V.03	Feb 2012	

The idea presented here is to extend the usage of MSRP as defined in IEEE Std 802.1Q-2011 to provide support for redundancy and loop-free path selection beyond the limitation of spanning tree networks.

- Propagate the MSRP TA through all the possible paths offered by the topology (beyond the spanning tree as it is the case in MSRP rev1).

- For a given stream ID, TAs sent through a different port (could either be the talker or the first bridge "splitting" the TAs are marked with a subStream ID to differentiate these TAs.

- TAs propagate link metrics accumulated along the path:

    - lowest link max bandwidth,
    - accumulated latency,
    - accumulated hop count : incremented by each bridge,
    - highest ref count : indicates how many TAs with the same StreamID but different subStreams ID crosses this particular bridge. This reference count allows the listener to detect a path without any single point of failure (ref count = 0).
    - link ref count: indicates to the downstream bridge how many TAs with the same StreamID but different subStrream IDs were propagated on this particular link.

- Loop protection: TA includes a sequence ID incremented by the Talker for each TA. Bridge rejects duplicated TAs received on a different port that the port registered for a given TA tuple {StreamID, StreamsubID, SEq ID} .

- Downstream bridges eliminates duplicated TA with the same {StreamID, subStreamID} on configurable criteria:
    - Path Metrics (multi criteria: bw, latency, link type,…)

  - Bridge Reference count (single point of failure)
  - Hop Count

- Listeners receive multiple TAs for the same stream ID but different substreamIDs and select the path(s) they will register for this given stream by sending a Listener Ready along this/these selected path(s). The path selection is application specific based on the metrics and references provided by the TAs.

Figure 1 tries to illustrate the proposed scheme for a given stream:

1) Talker: advertises Stream ID =1

2) B1: "splitting" bridge i.e. it propagates the TA through 2 different paths. The two TAs are differentiated by their Stream sub ID.

   Note: "Splitting" could be either done by a talker with multiple ports idea or a downstream bridge from the downstream topological point from where multiple paths are available.

3) B2: propagate the TA{1,2} received on port [1] on every egress port [2,3] and the TA {1,3} received on port [3] on egress port [1,2] (regardless of the RSTP port state of the egress ports). Since the bridge forwards TAs with the same Stream ID but different Stream subID, the bridge reference count is incremented to indicate a single point of failure.

4) B3: propagate the TA{1,3} received on port [1] on every egress port [2,3] and the TA {1,2} received on port [3] on egress port [1,2] (regardless of the RSTP port state of the egress ports). Since the bridge forwards TAs with the same Stream ID but different Stream subID, the bridge reference count is incremented to indicate a single point of failure.

5) B4: receives multiple TAs with the same stream ID but with different stream subIDs and propagates both on every egress port (regardless of the RSTP port state).

6) B5: receives multiple TAs with the same stream ID and the same stream sub IDs . The bridge prunes these TAs by selecting which single TA will be propagated, based on configurable preference /selection between TA attributes : (link metrics, hop count,bridge ref count,..)

   Note: The precedence should be consistent across the network but this is a management issue. There is several ways to handle this:

   - *A default configuration*

- *Multiple preference sets selected by an additional stream parameter to keep flexibility (i.e some streams could select bandwhith preference, while other could select latency preference,…)*

In this example, the bridge B5 selects *(no logic here, just an illustration of the selection mentioned above)*

- for TA{1,2 }, the path with the best link metric over the bridge count and bridge reference count
- for TA{1,3}, the path with the lowest hop count

7) Listener: receives multiple TAs and could select one or multiple TAs based on the application requirements and the attributes provided by the TAs .

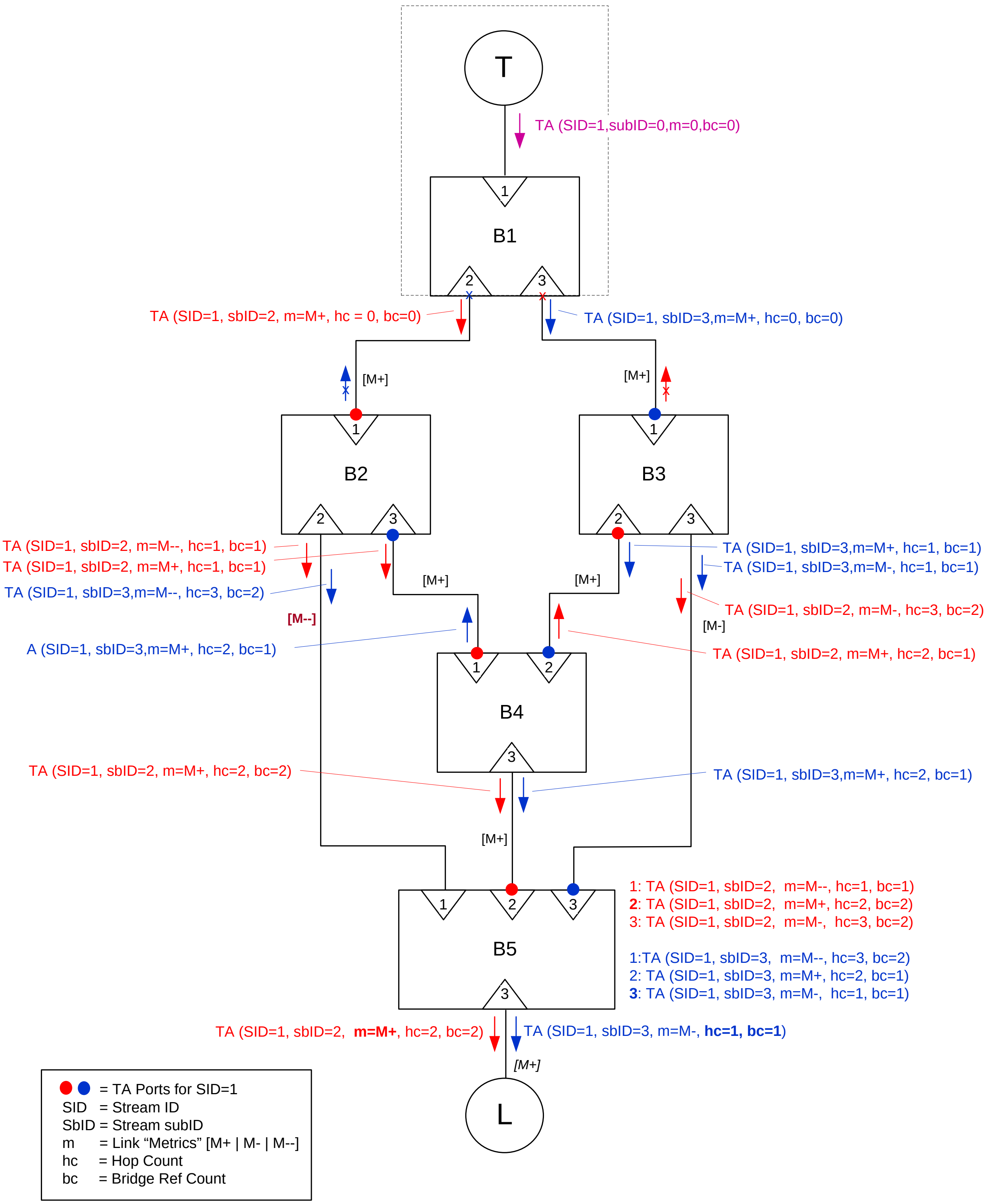

Figure 1: Example of TA propagation

Figure 2 & Figure 3 illustrate the TA propagation for a given flow when the option of avoiding single points of failure i.e. paths sharing the same bridge.

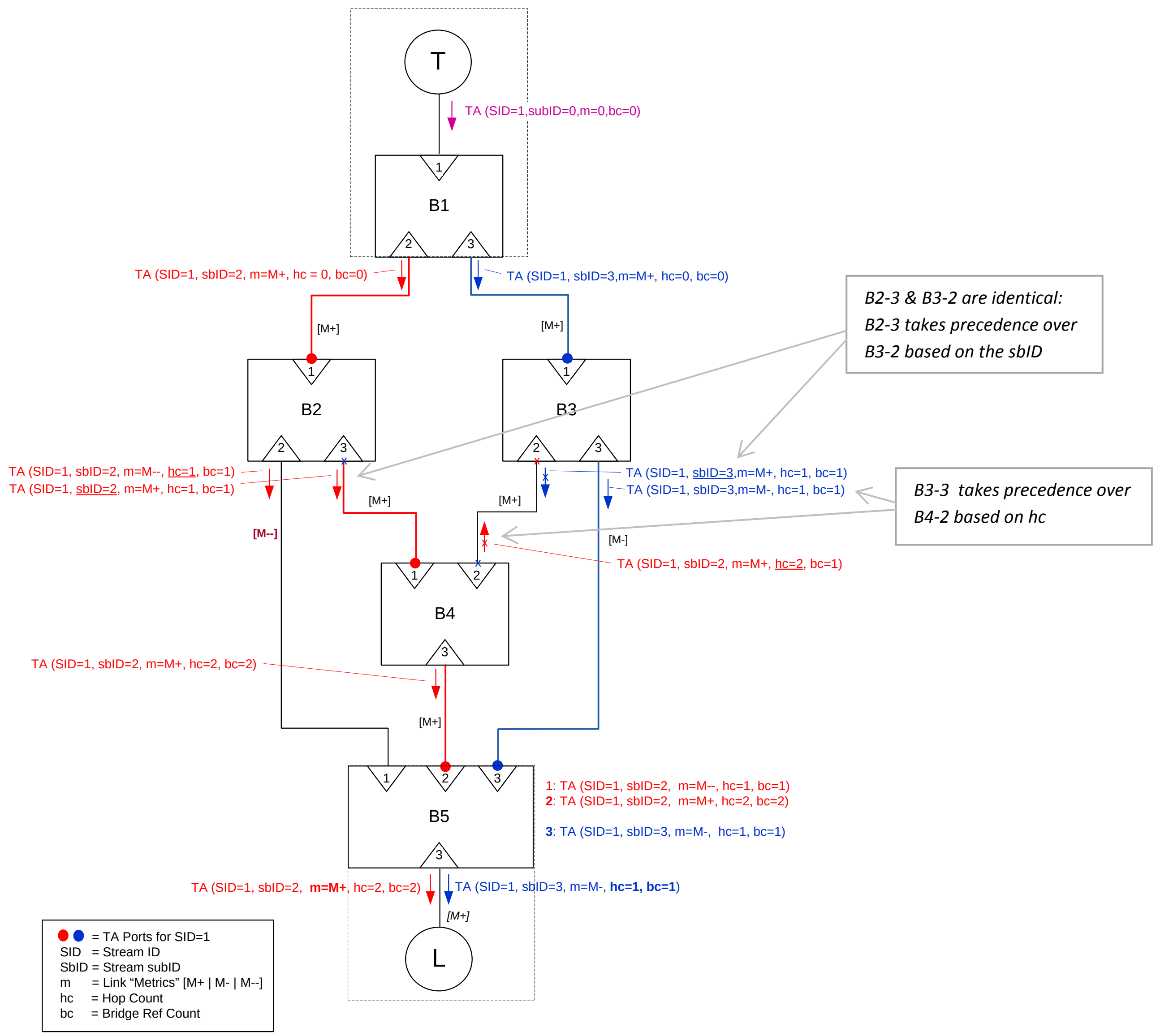

Figure 2: Example of TA propagation without single point of failure - 1

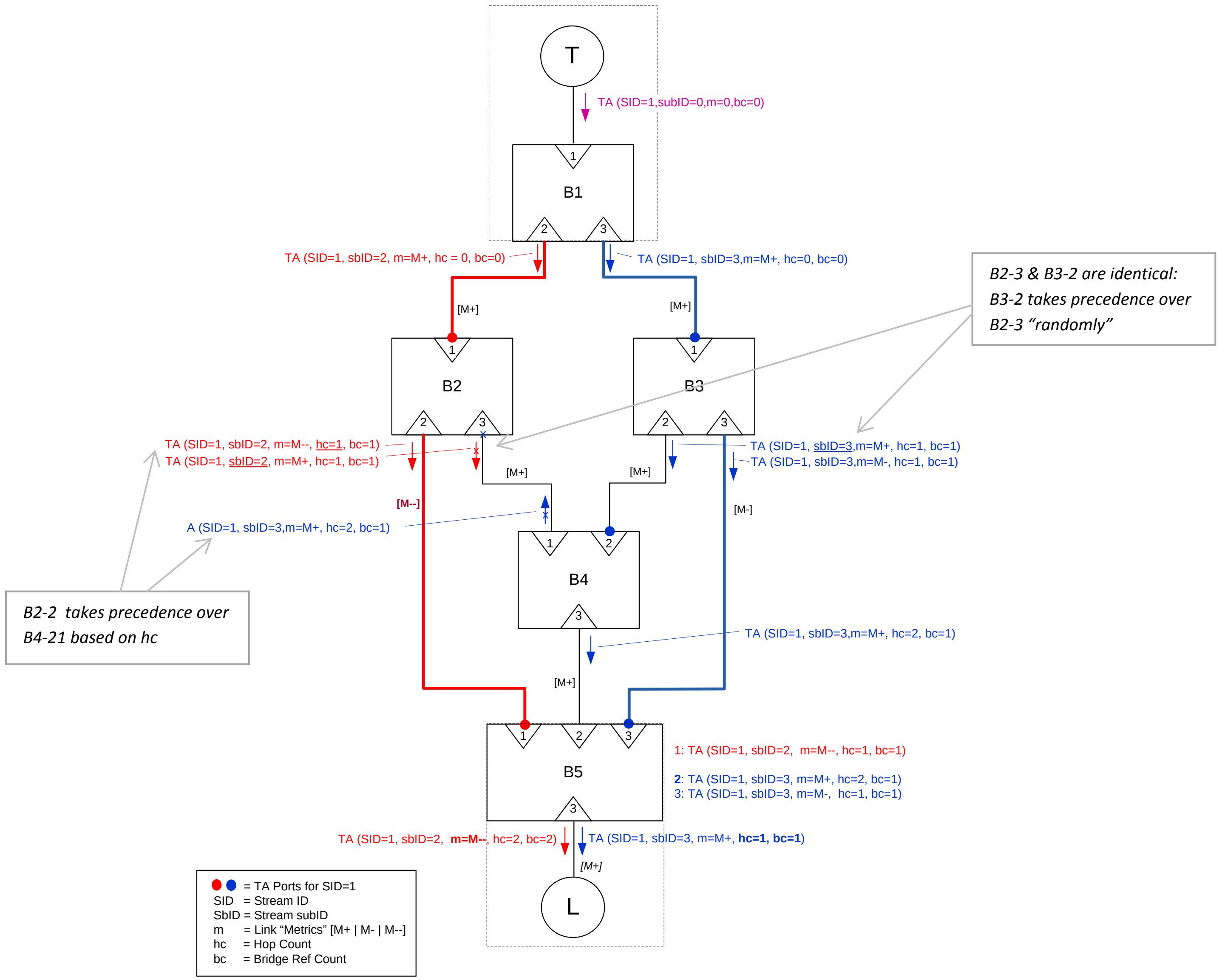

Figure 3: Example of TA propagation without single point of failure - 2