Internet Engineering Task Force (IETF)
Request for Comments: 6329
Category: Standards Track
ISSN: 2070-1721

D. Fedyk, Ed.
Alcatel-Lucent
P. Ashwood-Smith, Ed.
Huawei
D. Allan
Ericsson
N. Bragg
Ciena Limited
P. Unbehagen
Avaya

April 2012

IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging

Abstract

802.1aq Shortest Path Bridging (SPB) has been standardized by the IEEE as the next step in the evolution of the various spanning tree and registration protocols. 802.1aq allows for true shortest path forwarding in a mesh Ethernet network context utilizing multiple equal cost paths. This permits it to support much larger Layer 2 topologies, with faster convergence, and vastly improved use of the mesh topology. Combined with this is single point provisioning for logical connectivity membership, which includes point-to-point, point-to-multipoint, and multipoint-to-multipoint variations. This memo documents the IS-IS changes required to support this IEEE protocol and provides some context and examples.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

802.1aq Shortest Path Bridging (SPB) [802.1aq] has been standardized by the IEEE as the next step in the evolution of the various spanning tree and registration protocols. 802.1aq allows for true shortest path forwarding in an Ethernet mesh network context utilizing multiple equal cost paths. This permits SPB to support much larger Layer 2 topologies, with faster convergence, and vastly improved use of the mesh topology. Combined with this is single point provisioning for logical connectivity membership, which includes point-to-point (E-LINE), point-to-multipoint (E-TREE), and multipoint-to-multipoint (E-LAN) variations.

The control protocol for 802.1aq is IS-IS [IS-IS] augmented with a small number of TLVs and sub-TLVs. This supports two Ethernet encapsulating data paths, 802.1ad (Provider Bridges) [PB] and 802.1ah (Provider Backbone Bridges) [PBB]. This memo documents those TLVs while providing some overview.

Note that 802.1aq requires no state machine or other substantive changes to [IS-IS]. 802.1aq simply requires a new Network Layer Protocol Identifier (NLPID) and set of TLVs. In the event of confusion between this document and [IS-IS], [IS-IS] should be taken as authoritative.

2. Terminology

In addition to well-understood IS-IS terms, this memo uses terminology from IEEE 802.1 and introduces a few terms:

```
802.1ad        Provider Bridges (PBs) - Q-in-Q encapsulation
802.1ah        Provider Backbone Bridges (PBBs), MAC-IN-MAC
               encapsulation
802.1aq        Shortest Path Bridging (SPB)
Base VID       VID used to identify a VLAN in management operations
B-DA           Backbone Destination Address 802.1ah PBB
B-MAC          Backbone MAC Address
B-SA           Backbone Source Address in 802.1ah PBB header
B-VID          Backbone VLAN ID in 802.1ah PBB header
B-VLAN         Backbone Virtual LAN
BPDU           Bridge PDU
BridgeID       64-bit quantity = (Bridge Priority:16)<<48 | SYSID:48
BridgePriority 16-bit relative priority of a node for tie-breaking
C-MAC          Customer MAC.  Inner MAC in 802.1ah PBB header
C-VID          Customer VLAN ID
ECT-ALGORITHM  32-bit unique ID of an SPF tie-breaking set of rules
ECT-MASK       64-bit mask XORed with BridgeID during tie-breaking
E-LAN          Bidirectional Logical Connectivity between >2 UNIs
```

E-LINE	Bidirectional Logical Connectivity between two UNIs
E-TREE	Asymmetric Logical Connectivity between UNIs
FDB	Filtering Database: {DA/VID}->{next hops}
I-SID	Ethernet Services Instance Identifier used for Logical Grouping for E-LAN/LINE/TREE UNIs
LAN	Local Area Network
LSDB	Link State Database
LSP	Link State PDU
MAC	Media Access Control
MAC-IN-MAC	Ethernet in Ethernet framing as per 802.1ah [PBB]
MDT	Multicast Distribution Tree
MMRP	Multiple MAC Registration Protocol 802.1ak [MMRP]
MT	Multi-Topology. As used in [MT]
MT ID	Multi-Topology Identifier (12 bits). As used in [MT]
NLPID	Network Layer Protocol Identifier: IEEE 802.1aq= 0xC1
NNI	Network-Network Interface
Q-in-Q	Additional S-VID after a C-VID (802.1ad) [PB]
PBB	Provider Backbone Bridge - forwards using PBB
Ingress Check	Source Forwarding Check - drops misdirected frames
(S,G)	Source & Group - identity of a source-specific tree
(*,G)	Any Source & Group - identity of a shared tree
S-VID	Service VLAN ID
SA	Source Address
SPB	Shortest Path Bridging - generally all of 802.1aq
SPB	Shortest Path Bridge - device implementing 802.1aq
SPB-instance	Logical SPB instance correlated by MT ID
SPBM	Device implementing SPB MAC mode
SPBV	Device implementing SPB VID mode
SPSourceID	20-bit identifier of the source of multicast frames
SPT	Shortest Path Tree computed by one ECT-ALGORITHM
SPT Region	A set of SPBs with identical VID usage on their NNIs
SPVID	Shortest Path VLAN ID: a C-VID or S-VID that identifies the source
STP	Spanning Tree Protocol
UNI	User-Network Interface: customer-to-SPB attach point
VID	VLAN ID: 12-bit logical identifier after MAC header
VLAN	Virtual LAN: a logical network in the control plane

3. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The lowercase forms with an initial capital "Must", "Must Not", "Shall", "Shall Not", "Should", "Should Not", "May", and "Optional" in this document are to be interpreted in the sense defined in

[RFC2119], but are used where the normative behavior is defined in documents published by SDOs other than the IETF.

4. 802.1aq Overview

This section provides an overview of the behavior of [802.1aq] and is not intended to be interpreted as normative text. For the definitive behavior, the reader should consult [802.1aq]. Nonetheless, lowercase forms with initial capitalization of the conventions in RFC 2119 are used in this section to give the reader an indication of the intended normative behaviors as above.

802.1aq utilizes 802.1Q-based Ethernet bridging. The filtering database (FDB) is populated as a consequence of the topology computed from the IS-IS database. For the reader unfamiliar with IEEE terminology, the definition of Ethernet behavior is almost entirely in terms of "filtering" (of broadcast traffic) rather than "forwarding" (the explicit direction of unicast traffic). This document uses the generic term "forwarding", and it has to be understood that these two terms simply represent different ways of expressing the same behaviors.

802.1aq supports multiple modes of operation depending on the type of data plane and the desired behavior. For the initial two modes of 802.1aq (SPBV and SPBM), routes are shortest path, are forward- and reverse-path symmetric with respect to any source/destination pair within the SPB domain, and are congruent with respect to unicast and multicast. Hence, the shortest path tree (SPT) to a given node is congruent with the multicast distribution tree (MDT) from a given node. The MDT for a given VLAN is a pruned subset of the complete MDT for a given node that is identical to its SPT. Symmetry and congruency preserve packet ordering and proper fate sharing of Operations, Administration, and Maintenance (OAM) flows by the forwarding path. Such modes are fully supported by existing [802.1ag] and [Y.1731] OAM mechanisms.

VLANs provide a natural delineation of service instances. 802.1aq supports two modes, SPB VID (SPBV) and SPB MAC (SPBM). In SPBV, multiple VLANS can be used to distribute load on different shortest path trees (each computed by a different tie-breaking rule) on a service basis. In SPBM, service instances are delineated by I-SIDs but VLANs again can be used to distribute load on different shortest path trees.

There are two encapsulation methods supported. SPBM can be used in a PBB network implementing PBB (802.1ah [PBB]) encapsulation. SPBV can be used in PB networks implementing VLANs, PB (802.1aq [PB]), or PBB encapsulation. The two modes can co-exist simultaneously in an SPB network.

The practical design goals for SPBV and SPBM in the current 802.1aq specification are networks of size 100 nodes and 1000 nodes respectively. However, since SPBV can be sparsely used in an SPB region it can simply span a large SPB region with a small number of SPVIDs.

In SPBM and SPBV each bridge has at least one unique "known" MAC address which is advertised by IS-IS in the SYSID.

In the forwarding plane, SPBM uses the combination of one or more B-VIDs and "known" Backbone-MAC (B-MAC) addresses that have been advertised in IS-IS. The term Backbone simply implies an encapsulation that is often used in the backbone networks, but the encapsulation is useful in other types of networks where hiding C-MACs is useful.

The SPBM filtering database (FDB) is computed and installed for unicast and multicast MAC addresses, while the SPBV filtering database is computed and installed for unidirectional VIDs (referred to as SPVIDs), after which MAC reachability is learned (exactly as in bridged Ethernet) for unicast MACs.

Both SPBV and SPBM use source-specific multicast trees. If they share the same ECT-ALGORITHM (32-bit worldwide unique definition of the computation), the tree is the same SPT. For SPBV, (S,G) is encoded by a source-specific VID (the SPVID) and a standard Group MAC address. For SPBM, (S,G) is encoded in the destination B-MAC address as the concatenation of a 20-bit SPB wide unique nodal nickname (referred to as the SPSourceID) and the 24-bit I-SID together with the B-VID that corresponds to the ECT-ALGORITHM network wide.

802.1aq supports membership attributes that are advertised with the I-SID (SPBM) or Group Address (SPBV) that defines the group. Individual members can be transmitters (T) and/or receivers (R) within the group, and the multicast state is appropriately sized to these requests. Multicast group membership is possible even without transmit membership by performing head-end replication to the receivers thereby eliminating transit multicast state entirely.

Some highly connected mesh networks provide for path diversity by offering multiple equal cost alternatives between nodes. Since congruency and symmetry Must be honored, a single tree may leave some links under-utilized. By using different deterministic tie-breakers, up to 16 shortest paths of arbitrary diversity are possible between any pair of nodes. This distributes the traffic on a VLAN basis.

SPBV and SPBM May share a single SPT with a single ECT-ALGORITHM or use any combination of the 16 ECT-ALGORITHMs. An extensible framework permits additional or alternative algorithms with other properties and parameters (e.g., ECMP, (*,G)) to also be supported without any changes in this or the IEEE documents.

4.1. Multi-Topology Support

SPB incorporates the multi topology features of [MT] thereby allowing multiple logical SPB instances within a single IS-IS instance.

To accomplish this, all SPB-related information is either explicitly or implicitly associated with a Multi-Topology Identifier (MT ID). SPB information related to a given MT ID thus forms a single logical SPB instance.

Since SPB has its own adjacency metrics and those metrics are also associated with an MT ID, it is possible to have different adjacency metrics (or infinite metrics) for SPB adjacencies that are not only distinct from IP or other NLPIDs riding in this IS-IS instance, but also distinct from those used by other SPB instances in the same IS-IS instance.

Data plane traffic for a given MT ID is intrinsically isolated by the VLANs assigned to the SPB instance in question. Therefore, VLANs (represented by VIDs in TLVs and in the data plane) Must Not overlap between SPB instances (regardless of how the control planes are isolated).

The [MT] mechanism when applied to SPB allows different routing metrics and topology subsets for different classes of services.

The use of [MT] other than the default MT ID #0 is completely OPTIONAL.

The use of [MT] to separate SPB from other NLPIDs is also OPTIONAL.

4.2. Data Path SPBM - Unicast

Unicast frames in SPBM are encapsulated as per 802.1ah [PBB]. A Backbone Source Address (B-SA), Backbone Destination Address (B-DA), Backbone VLAN ID (B-VID), and an I-Component Service Instance ID (I-TAG) are used to encapsulate the Ethernet frame. The B-SA is a B-MAC associated with the ingress 802.1aq bridge, usually the "known" B-MAC of that entire bridge. The B-DA is one of the "known" B-MACs associated with the egress 802.1aq bridge. The B-VID and I-TAG are mapped based on the physical or logical UNI port (untagged, or tagged either by S-TAG or C-TAG) being bridged. Normal learning and broadcast to unknown C-MACs is applied as per [PBB] at the ingress/egress SPBs only.

Unlike [PBB] on a (*,G) tree, the B-DA forwarding on tandem nodes (NNI to NNI) is performed without learning. Instead, the output of 802.1aq computations, based on the TLVs specified in this document, is used to populate the filtering databases (FDBs). The FDB entries map {B-DA, B-VID} to an outgoing interface and are only populated from the IS-IS database and computations.

The B-SA/B-VID is checked on tandem nodes against the ingress port. If the B-SA/B-VID (as a destination) entry in the FDB does not point to the port on which the packet arrived, the packet is discarded. This is referred to as an ingress check and serves as a very powerful loop mitigation mechanism.

4.3. Data Path SPBM - Multicast (Head-End Replication)

Head-end replication is supported for instances where there is a sparse community of interest or a low likelihood of multicast traffic. Head-end replication requires no multicast state in the core. A UNI port wishing to use head-end replication Must Not advertise its I-SID membership with the Transmit (T) bit set but instead Must locally and dynamically construct the appropriate unicast serial replication to all the other receivers (Receive (R) bit set) of the same I-SID.

When an unknown customer unicast or a multicast frame arrives at an SPBM User-Network Interface (UNI) port that has been configured to replicate only at the head end, the packet is replicated once for each receiver, encapsulated, and sent as a unicast frame. The set of receivers is determined by inspecting the IS-IS database for other SPBs that have registered interest in the same I-SID with the R bit set. This R bit / I-SID pair is found in the SPBM Service Identifier and Unicast Address (SPBM-SI) sub-TLV. The packets are encapsulated as per the SPBM unicast forwarding above.

4.4. Data Path SPBM - Multicast (Tandem Replication)

Tandem replication uses the shortest path tree to replicate frames only where the tree forks and there is at least one receiver on each branch. Tandem replication is bandwidth efficient but uses multicast FDB entries (state) in core bridges, which might be unnecessary if there is little multicast traffic demand. The head-end replication mode is best suited for the case where there is little or no true multicast traffic for an I-SID. Tandem replication is triggered on transit nodes when the I-SID is advertised with the T bit set.

Broadcast, unknown unicast, or multicast frames arriving at an SPBM UNI port are encapsulated with a B-DA multicast address that uniquely identifies the encapsulating node (the root of the Multicast Distribution Tree) and the I-SID scoping this multicast.

This B-DA address is a well-formed multicast group address (as per 802.1Q and 802.1ah) that concatenates the SPSourceID A' with the I-SID M (written as DA=<A',M> and uniquely identifying the (S,G) tree). This exact format is given in Figure 1 below:

```
 M L TYP
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|1|0|0|SPSrcMS|  SPSrc [8:15] |  SPSrc [0:7]  | I-SID [16:23] |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| I-SID [8:15]  |  I-SID [0:7]  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: SPBM Multicast Address Format
(SPSrcMS represents SPSrc [16:19])

Note: In Figure 1, the index numbering from less significant bit to more significant bit within a byte or field within a byte gives the wire order of the bits in the address consistent with the IETF format in the rest of this document. (The IEEE convention for number representation reverses the bits within an octet compared with IETF practice.)

- o M is the multicast bit, always set to 1 for a multicast DA. (It is the lowest bit in the most significant byte.)

- o L is the local bit, always set to 1 for an SPBM-constructed multicast DA.

- o TYP is the SPSourceID type. 00 is the only type supported at this time.

- o SPSrc (SPSourceID) is a 20-bit quantity that uniquely identifies a SPBM node for all B-VIDs allocated to SPBM operation. This is just the SPSourceID advertised in the SPB Instance (SPB-Inst) sub-TLV. The value SPSourceID = 0 has special significance; it is advertised by an SPBM node that has been configured to assign its SPSourceID dynamically, which requires LSDB synchronization, but where the SPSourceID assignment has not yet completed.

- o I-SID is the 24-bit I-Component Service ID advertised in the SPBM Service Identifier TLV. It occupies the lower 24 bits of the SPBM multicast DA. The I-SID value 0xfff is reserved for SPBM control traffic (refer to the default I-SID in [802.1aq]).

This multicast address format is used as the DA on frames when they are first encapsulated at ingress to the SPBM network. The DA is also installed into the FDBs on all SPBM nodes that are on the corresponding SPT between the source and other nodes that have registered receiver interest in the same I-SID.

Just as with unicast forwarding, the B-SA/B-VID May be used to perform an ingress check, but the SPSourceID encoded in the DA and the "drop-on-unknown" functionality of the FDB in [PBB] achieve the same effect.

The I-Component at the egress SPBM device has completely standard [PBB] behavior and therefore will:

1) learn the remote C-SA to B-SA relationship and

2) bridge the original customer frame to the set of local UNI ports that are associated with the I-SID.

4.5. Data Path SPBV Broadcast

When a packet for an unknown DA arrives at an SPBV UNI port, VID translation (or VID encapsulation for untagged Frames) with the corresponding SPVID for this VLAN and ingress SPB is performed.

SPVID forwarding is simply an SPT that follows normal VLAN forwarding behavior, with the exception that the SPVID is unidirectional. As a result, shared VLAN learning (SVL) is used between the forward- and reverse-path SPVIDs associated with the same Base VID to allow SPBV unicast forwarding to operate in the normal reverse learning fashion.

Ingress check is done by simply verifying that the bridge to which the SPVID has been assigned is indeed "shortest path" reachable over the link over which the packet tagged with that SPVID arrived. This is computed from the IS-IS database and is implied when the SPVID is associated with a specific incoming port.

4.6. Data Path SPBV Unicast

When a packet for a known DA arrives at an SPBV UNI port, VID translation (or VID encapsulation for untagged Frames) with the corresponding SPVID for this VLAN and ingress bridge is performed.

Since the SPVID will have been configured to follow a source-specific SPT and the DA is known, the packet will follow the source-specific path towards the destination C-MAC.

Ingress check is as per the previous SPBV section.

4.7. Data Path SPBV Multicast

C-DA multicast addresses May be advertised from SPBV UNI ports. These may be configured or learned through the Multiple MAC Registration Protocol (MMRP). MMRP is terminated at the edge of the SPBV network and IS-IS carries the multicast addresses. Tandem SPBV devices will check to see if they are on the SPF tree between SPBV UNI ports advertising the same C-DA multicast address, and if so will install multicast state to follow the SPBV SPF trees.

Ingress check is as per the previous two SPBV sections.

5. SPBM Example

Consider the small example network shown in Figure 2. Nodes are drawn in boxes with the last nibble of their B-MAC address :1..:7. The rest of the B-MAC address nibbles are 4455-6677-00xx. Links are drawn as "--" and "/", while the interface indexes are drawn as numbers next to the links. UNI ports are shown as "<==>" with the desired I-SID shown at the end of the UNI ports as "i1".

```
                   +----+           +----+
                   | :4 | 2 ------1 | :5 | <==> i1
                   +----+           +----+
                  1      3         3      2
                 /        \       /        \
                1          4     3          2
             +----+        +----+         +----+
   i1 <==> | :1 | 2----1 | :2 | 2------1 | :3 | <==> i1
             +----+        +----+         +----+
                3          6     5          3
                 \        /       \        /
                  3      2         1      2
                   +----+           +----+
                   | :6 | 1-------3 | :7 | <==> i1
                   +----+           +----+
```

Figure 2: SPBM Example 7-Node Network

Using the default ECT-ALGORITHM (00-80-C2-01), which picks the equal cost path with the lowest BridgeID, this ECT-ALGORITHM is assigned to B-VID 100. When all links have the same cost, then the 1-hop shortest paths are all direct and the 2-hop shortest paths (which are of course symmetric) are as follows:

```
{ 1-2-3,  1-2-5, 1-2-7, 6-2-5,
  4-2-7,  4-1-6, 5-2-7, 6-2-3, 4-2-3 }
```

Node :1’s unicast forwarding table therefore routes toward B-MACs :7, :3, and :5 via interface/2, while its single-hop paths are all direct as can be seen from its FDB given in Figure 3.

Node :1 originates multicast since it is at the head of the MDT to nodes :3, :5, and :7 and is a transmitter of I-SID 1, which nodes :3, :5, and :7 all wish to receive. Node :1 therefore produces a multicast forwarding entry whose DA contains its SPSourceID (which is the last 20 bits of the B-MAC in the example) and the I-SID 1. Node :1 thereafter sends packets matching this entry to interface if/2 with B-VID=100. Node :1’s full unicast (U) and multicast (M) table is shown in Figure 3. Note that the IN/IF (incoming interface) field is not specified for unicast traffic, and for multicast traffic has to point back to the root of the tree, unless it is the head of the tree -- in which case, we use the convention if/00. Since node :1 is not transit for any multicast, it only has a single entry for the root of its tree for I-SID=1.

	IN/IF	DESTINATION ADDR	BVID	OUT/IF(s)
U	if/**	4455-6677-0002	0100	{if/2 }
U	if/**	4455-6677-0003	0100	{if/2 }
U	if/**	4455-6677-0004	0100	{if/1 }
U	if/**	4455-6677-0005	0100	{if/2 }
U	if/**	4455-6677-0006	0100	{if/3 }
U	if/**	4455-6677-0007	0100	{if/2 }
M	if/00	7300-0100-0001	0100	{if/2 }

Figure 3: SPBM Node :1 FDB - Unicast (U) and Multicast (M)

Node :2, being at the center of the network, has direct 1-hop paths to all other nodes; therefore, its unicast FDB simply sends packets with the given B-MAC/B-VID=100 to the interface directly to the addressed node. This can be seen by looking at the unicast entries (the first 6) shown in Figure 4.

	IN/IF	DESTINATION ADDR	BVID	OUT/IF(s)
U	if/**	4455-6677-0001	0100	{if/1 }
U	if/**	4455-6677-0003	0100	{if/2 }
U	if/**	4455-6677-0004	0100	{if/4 }
U	if/**	4455-6677-0005	0100	{if/3 }
U	if/**	4455-6677-0006	0100	{if/6 }
U	if/**	4455-6677-0007	0100	{if/5 }
M	if/01	7300-0100-0001	0100	{if/2,if/3,if/5 }
M	if/02	7300-0300-0001	0100	{if/1 }
M	if/03	7300-0500-0001	0100	{if/1,if/5 }
M	if/05	7300-0700-0001	0100	{if/1,if/3 }

Figure 4: SPBM Node :2 FDB Unicast (U) and Multicast (M)

Node :2's multicast is more complicated since it is a transit node for the 4 members of I-SID=1; therefore, it requires 4 multicast FDB entries depending on which member it is forwarding/replicating on behalf of. For example, node :2 is on the shortest path between each of nodes {:3, :5, :7} and :1. So it must replicate from node :1 I-SID 1 out on interfaces { if/2, if/3 and if/5 } (to reach nodes :3, :5, and :7). It therefore creates a multicast DA with the SPSourceID of node :1 together with I-SID=1, which it expects to receive over interface/1 and will replicate out interfaces { if/2, if/3 and if/5 }. This can be seen in the first multicast entry in Figure 4.

Note that node :2 is not on the shortest path between nodes :3 and :5 nor between nodes :3 and :7; however, it still has to forward packets to node :1 from node :3 for this I-SID, which results in the second multicast forwarding entry in Figure 4. Likewise, for packets originating at nodes :5 or :7, node :2 only has to replicate twice, which results in the last two multicast forwarding entries in Figure 4.

6. SPBV Example

Using the same example network as Figure 2, we will look at the FDBs produced for SPBV mode forwarding. Nodes :1, :5, :3, and :7 wish to transmit and receive the same multicast MAC traffic using multicast address 0300-0000-000f and at the same time require congruent and symmetric unicast forwarding. In SPBV mode, the only encapsulation is the C-TAG or S-TAG, and the MAC addresses SA and DA are reverse-path learned, as in traditional bridging.

```
                        +----+           +----+
                        | :4 | 2 ------1 | :5 | <==> MMAC ..:f
                        +----+           +----+
                       1      3         3      2
                      /        \       /        \
                     1          4     3          2
                  +----+        +----+          +----+
         MMAC<==> | :1 | 2----1 | :2 | 2------1 | :3 | <==> MMAC ..:f
          ..:f    +----+        +----+          +----+
                     3          6     5          3
                      \        /       \        /
                       3      2         1      2
                        +----+           +----+
                        | :6 | 1-------3 | :7 | <==> MMAC ..:f
                        +----+           +----+
```

Figure 5: SPBV Example 7-Node Network

Assuming the same ECT-ALGORITHM (00-80-C2-01), which picks the equal cost path with the lowest BridgeID, this ECT-ALGORITHM is assigned to Base VID 100, and for each node the SPVID = Base VID + Node ID (i.e., 101, 102..107). When all links have the same cost, then the 1-hop shortest paths are all direct, and the 2-hop shortest paths (which are of course symmetric) are as previously given for Figure 2.

Node :1's SPT for this ECT-ALGORITHM is therefore (described as a sequence of unidirectional paths):

{ 1->4, 1->6, 1->2->3, 1->2->5, 1->2->7 }

The FDBs therefore must have entries for the SPVID reserved for packets originating from node :1, which in this case is VID=101.

Node :2 therefore has an FDB that looks like Figure 6. In particular, it takes packets from VID 101 on interface/01 and sends to nodes :3, :5, and :7 via if/2, if/3, and if/5. It does not replicate anywhere else because the other nodes (:4 and :6) are reached by the SPT directly from node :1. The rest of the FDB unicast entries follow a similar pattern; recall that the shortest path between :4 and :6 is via node :1, which explains replication onto only two interfaces from if/4 and if/6. Note that the destination addresses are wild cards, and SVL exists between these SPVIDs because they are all associated with Base VID = 100, which defines the VLAN being bridged.

	IN/IF	DESTINATION ADDR	VID	OUT/IF(s)
U	if/01	**************	0101	{if/2,if/3,if/5 }
U	if/02	**************	0103	{if/1,if/4,if/6 }
U	if/04	**************	0104	{if/2,if/5 }
U	if/03	**************	0105	{if/1,if/5,if/6 }
U	if/06	**************	0106	{if/2,if/3 }
U	if/05	**************	0107	{if/1,if/3,if/4 }

Figure 6: SPBV Node :2 FDB Unicast

Now, since nodes :5, :3, :7 and :1 are advertising membership in the same multicast group address :f, Node 2 requires additional entries to replicate just to these specific nodes for the given multicast group address. These additional multicast entries are given below in Figure 7.

	IN/IF	DESTINATION ADDR	VID	OUT/IF(s)
M	if/01	0300-0000-000f	0101	{if/2,if/3,if/5 }
M	if/02	0300-0000-000f	0103	{if/1 }
M	if/03	0300-0000-000f	0105	{if/1,if/5 }
M	if/05	0300-0000-000f	0107	{if/1,if/3 }

Figure 7: SPBV Node :2 FDB Multicast (M)

7. SPB Supported Adjacency types

IS-IS for SPB currently only supports peer-to-peer adjacencies. Other link types are for future study. As a result, pseudonodes and links to/from pseudonodes are not considered as part of the IS-IS SPF computations and will be avoided if present in the physical topology. Other NLPIDs MAY of course use them as per normal.

IS-IS for SPB Must use the IS-IS three-way handshake for IS-IS point-to-point adjacencies described in RFC 5303.

8. SPB IS-IS Adjacency Addressing

The default behavior of 802.1aq is to use the normal IS-IS Ethernet multicast addresses for IS-IS.

There are however additional Ethernet multicast addresses that have been assigned for 802.1aq for special use cases. These do not in any way change the state machinery or packet formats of IS-IS but simply recommend and reserve different multicast addresses. Refer to [802.1aq] for additional details.

9. IS-IS Area Address and SYSID

A stand-alone implementation (supporting ONLY the single NLPID=0xC1) of SPB Must use an IS-IS area address value of 0, and the SYSID Must be the well-known MAC address of the SPB device.

Non-stand-alone implementations (supporting other NLPIDs) MUST use the normal IS-IS rules for the establishment of a level 1 domain (i.e., multiple area addresses are allowed only where immediate adjacencies share a common area address). Level 2 operations of course place no such restriction on adjacent area addresses.

10. Level 1/2 Adjacency

SPBV and SPBM will operate within either an IS-IS level 1 or level 2. As a result, the TLVs specified here MAY propagate in either level 1 or level 2 LSPs. IS-IS SPB implementations Must support level 1 and May support level 2 operations. Hierarchical SPB is for further study; therefore, these TLVs Should Not be leaked between level 1 and level 2.

11. Shortest Path Default Tie-Breaking

The default algorithm is ECT-Algorithm = 00-80-c2-01.

Two mechanisms are used to ensure symmetry and determinism in the shortest path calculations.

The first mechanism addresses the problem when different ends (nodes) of an adjacency advertise different values for the SPB-LINK-METRIC. To solve this, SPB shortest path calculations Must use the maximum value of the two nodes' advertised SPB-LINK-METRICs when accumulating and minimizing the (sub)path costs.

The second mechanism addresses the problem when two equal sums of link metrics (sub)paths are found. To solve this, the (sub)path with the fewest hops between the fork/join points Must win the tie. However, if both (sub)paths have the same number of hops between the fork and join points, then the default tie-breaking Must pick the path traversing the intermediate node with the lower BridgeID. The BridgeID is an 8-byte quantity whose upper 2 bytes are the node's BridgePriority and lower 6 bytes are the node's SYSID.

For example, consider the network in Figure 2 when a shortest path computation is being done from node :1. Upon reaching node :7, two competing sub-paths fork at node :1 and join at node :7, the first via :2 and the second via :6. Assuming that all the nodes advertise a Bridge Priority of 0, the default tie-breaking rule causes the path traversing node :2 to be selected since it has a lower BridgeID {0...:2} than node :6 {0...:6}. Note that the operator may cause the tie-breaking logic to pick the alternate path by raising the Bridge Priority of node :2 above that of node :6.

The above algorithm guarantees symmetric and deterministic results in addition to having the critical property of transitivity (shortest path is made up of sub-shortest paths).

12. Shortest Path ECT

Standard ECT Algorithms initially have been proposed ranging from 00-80-c2-01 to 00-80-c2-10.

To create diversity in routing, SPB defines 16 variations on the above default tie-breaking algorithm; these have worldwide unique designations 00-80-C2-01 through 00-80-C2-10. These designations consist of the IEEE 802.1 OUI (Organizationally Unique Identifier) value 00-80-C2 concatenated with indexes 0X01..0X10. These individual algorithms are implemented by selecting the (sub)path with the lowest value of:

```
     XOR BYTE BY BYTE(ECT-MASK{ECT-ALGORITHM.index},BridgeID)
```

Where:

```
     ECT-MASK{17} = { 0x00, 0x00, 0xFF, 0x88,
                      0x77, 0x44, 0x33, 0xCC,
                      0xBB, 0x22, 0x11, 0x66,
                      0x55, 0xAA, 0x99, 0xDD,
                      0xEE };

     XOR BYTE BY BYTE  - XORs BridgeID bytes with ECT-MASK
```

ECT-MASK{1}, since it XORs with all zeros, yields the default algorithm described above (00-80-C2-01); while ECT-MASK{2}, since it XORs with a mask of all ones, will invert the BridgeID, essentially picking the path traversing the largest Bridge ID. The other ECT-MASKs produce diverse alternatives. In all cases, the BridgePriority, since it is the most significant part of the BridgeID, permits overriding the SYSID as the selection criteria and gives the operator a degree of control on the chosen ECT paths.

To support many other tie-breaking mechanisms in the future, two opaque ECT TLVs are defined, which may be used to provide parameters to ECT-ALGORITHMs outside of the currently defined space.

ECT-ALGORITHMs are mapped to VIDs, and then services can be assigned to those VIDs. This permits a degree of traffic engineering since service assignment to VID is consistent end to end through the network.

13. Hello (IIH) Protocol Extensions

IEEE 802.1aq can run in parallel with other network layer protocols such as IPv4 and IPv6; therefore, failure of two SPB nodes to establish an adjacency MUST NOT cause rejection of an adjacency for the purposes of other network layer protocols.

IEEE 802.1aq has been assigned the NLPID value 0xC1 [RFC6328], which MUST be used by Shortest Path Bridges (SPBs) to indicate their ability to run 802.1aq. This is done by including this NLPID value in the IS-IS IIH PDU Protocols Supported TLV (type 129). 802.1aq frames MUST only flow on adjacencies that advertise this NLPID in both directions of the IIH PDUs. 802.1aq computations MUST consider an adjacency that has not advertised 0xC1 NLPID in both directions as non-existent (infinite link metric) and MUST ignore any IIH SPB TLVs they receive over such adjacencies.

IEEE 802.1aq augments the normal IIH PDU with three new TLVs, which like all other SPB TLVs, travel within Multi-Topology [MT] TLVs, therefore allowing multiple logical instances of SPB within a single IS-IS protocol instance.

Since SPB can use many VIDs and Must agree on which VIDs are used for which purposes, the IIH PDUs carry a digest of all the used VIDs (on the NNIs) referred to as the SPB-MCID TLV, which uses a common and compact encoding reused from 802.1Q.

SPB neighbors May support a mechanism to verify that the contents of their topology databases are synchronized (for the purposes of loop prevention). This is done by exchanging a digest of SPB topology information (computed over all MT IDs) and taking specific actions on forwarding entries when the digests indicate a mismatch in topology. This digest is carried in the Optional SPB-Digest sub-TLV.

Finally, SPB needs to know which SPT Sets (defined by ECT-ALGORITHMs) are being used by which VIDs, and this is carried in the Base VLAN Identifiers (SPB-B-VID) sub-TLV.

13.1. SPB-MCID Sub-TLV

This sub-TLV is added to an IIH PDU to indicate the digest for the multiple spanning tree configuration a.k.a. MCID. This TLV is a digest of local configuration of which VIDs are running which protocols. (The information is not to the level of a specific algorithm in the case of SPB.) This information Must be the same on all bridges in the SPT Region controlled by an IS-IS instance. The data used to generate the MCID is populated by configuration and is a digest of the VIDs allocated to various protocols. Two MCIDs are carried to allow non-disruptive transitions between configurations when the changes are non-critical.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type=SPB-MCID  | = 4
+-+-+-+-+-+-+-+-+
|   Length      |    (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            MCID (51 Bytes)                    |
|                            ...............                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Aux   MCID (51 Bytes)                    |
|                            ...............                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- o Type: sub-TLV type 4.

- o Length: The size of the value defined below (102).

- o MCID (51 bytes): The complete MCID defined in IEEE 802.1Q, which identifies an SPT Region on the basis of matching assignments of VIDs to control regimes (xSTP, SPBV, SPBM, etc.). Briefly, the MCID consists of a 1-byte format selector, a 32-byte configuration name, a 2-byte revision level, and finally a 16-byte signature of type HMAC-MD5 over an array of 4096 elements that contain identifiers of the use of the corresponding VID. Refer to Section 13.8 of [802.1aq] for the exact format and procedure. Note that the use of the VID does not include specification of a specific SPB ECT-ALGORITHM; rather, it is coarser grain.

- o Aux MCID (51 bytes): The complete MCID defined in IEEE 802.1Q, which identifies an SPT Region. The aux MCID allows SPT Regions to be migrated by the allocation of new VLAN to FDB Mappings without interruption to existing traffic.

The SPB-MCID sub-TLV is carried within the MT-Port-Cap TLV [RFC6165] with the MT ID value of 0, which in turn is carried in an IIH PDU.

13.2. SPB-Digest Sub-TLV

This sub-TLV is Optionally added to an IIH PDU to indicate the current SPB topology digest value. It is always carried in an MT-Port-Cap TLV [RFC6165] with an MT ID value of 0. This information should settle to be the same on all bridges in an unchanging topology. Matching digests indicate (with extremely high probability) that the topology view between two SPBs is synchronized; this match (or lack of match) is used to control the updating of forwarding information. The SPB Agreement Digest is computed based on contributions derived from the current topologies of all SPB MT instances and is designed to change when significant topology changes occur within any SPB instance.

During the propagation of LSPs, the Agreement Digest may vary between neighbors until the key topology information in the LSPs is common. The digest is therefore a summarized means of determining agreement between nodes on database commonality, and hence of inferring agreement on the distance to all multicast roots. When present, it is used for loop prevention as follows: for each shortest path tree where it has been determined the distance to the root has changed, "unsafe" multicast forwarding is blocked until the exchanged Agreement Digests match, while "safe" multicast forwarding is allowed to continue despite the disagreement in digests and hence topology views. Section 28.2 of [802.1aq] defines in detail what constitutes "safe" vs. "unsafe".

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type=SPB-Digest| = 5
+-+-+-+-+-+-+-+-+
|   Length      | (1 byte)
+-----+-+---+---+
| Res |V| A | D | (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Agreement Digest (Length - 1)                    |
|                             ...............                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- o Type: sub-TLV type 5.

- o Length: The size of the value.

- o V bit: Agreed digest valid bit. See Section 28.2 of [802.1aq].

- A (2 bits): The Agreement Number 0-3, which aligns with the BPDU's Agreement Number concept [802.1aq]. When the Agreement Digest for this node changes, this number is incremented. The node then checks for Agreement Digest match (as below). The new local Agreement Number and the updated local Discarded Agreement Number are then transmitted with the new Agreement Digest to the node's neighbors in the Hello PDU. Once an Agreement Number has been sent, it is considered outstanding until a matching or more recent Discarded Agreement Number is received from the neighbor.

- D (2 bits): The Discarded Agreement Number 0-3, which aligns with BPDU's Agreement Number concept. When an Agreement Digest is received from a neighbor, this number is set to the received Agreement Number to signify that this node has received this new agreement and discarded any previous ones. The node then checks whether the local and received Agreement Digests match. If they do, this node then sets:

  the local Discarded Agreement Number = received Agreement Number + 1

  If the Agreement Digests match, AND received Discarded Agreement Number ==
  local Agreement Number + N (N = 0 || 1)

  then the node has a topology matched to its neighbor.

  Whenever the local Discarded Agreement Number relating to a neighbor changes, the local Agreement Digest, Agreement Number, and Discarded Agreement Number are transmitted.

- Agreement Digest. This digest is used to determine when SPB is synchronized between neighbors for all SPB instances. The Agreement Digest is a hash computed over the set of all SPB adjacencies in all SPB instances. In other words, the digest includes all VIDs and all adjacencies for all MT instances of SPB (but not other network layer protocols). This reflects the fact that all SPB nodes in a region Must have identical VID allocations (see Section 13.1), and so all SPB instances will contain the same set of nodes. The exact procedure for computing the Agreement Digest and its size are defined in Section 28.2 of [802.1aq].

The SPB-Digest sub-TLV is carried within the MT-Port-Cap TLV [RFC6165] (with the MT ID value 0), which in turn is carried in an IIH PDU.

When supported, this sub-TLV MUST be carried on every IIH between SPB neighbors, not just when a Digest changes.

When one peer supports this TLV and the other does not, loop prevention by Agreement Digest Must Not be done by either side.

13.3. SPB Base VLAN Identifiers (SPB-B-VID) Sub-TLV

This sub-TLV is added to an IIH PDU to indicate the mappings between ECT algorithms and Base VIDs (and by implication the VID(s) used on the forwarding path for each SPT Set for a VLAN identified by a Base VID) that are in use. Under stable operational conditions, this information should be the same on all bridges in the topology identified by the MT-Port-Cap TLV [RFC6165] it is being carried within.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+
|Type= SPB-B-VID| = 68
+-+-+-+-+-+-+-+-+-+
|   Length      |    (1 byte)
+-+-+-+-+-+-+-+-+-+-----------------------------------------------+
|        ECT-VID Tuple (1)  (6 bytes)                             |
+---------------------------------+-------------------------------+
|      ...                        | ECT-VID Tuple(2) (6 bytes)    |
+---------------------------------+-------------------------------+
|                          .....                                  |
+-----------------------------------------------------------------+
|                          .....                                  |
|                          .....                                  |
+-----------------------------------------------------------------+
```

o Type: sub-TLV type 6.

o Length: The size of the value is ECT-VID Tuples*6 bytes. Each 6-byte part of the ECT-VID tuple is formatted as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       ECT-ALGORITHM (32 bits)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Base VID (12 bits)    |U|M|RES|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

o ECT-ALGORITHM (4 bytes): The ECT-ALGORITHM is advertised when the bridge supports a given ECT-ALGORITHM (by OUI/Index) on a given Base VID. There are 17 predefined IEEE algorithms for SPB with index values 0X00..0X10 occupying the low 8 bits and the IEEE OUI=00-80-C2 occupying the top 24 bits of the ECT-ALGORITHM.

o Base VID (12 bits): The Base VID that is associated with the SPT Set.

o Use-Flag (1 bit): The Use-Flag is set if this bridge, or any bridge in the LSDB, is currently using this ECT-ALGORITHM and Base VID. Remote usage is discovered by inspection of the U bit in the SPB-Inst sub-TLV of other SPB bridges (see Section 14.1)

o M bit (1 bit): The M bit indicates if this Base VID operates in SPBM (M = 1) or SPBV (M = 0) mode.

The SPB-B-VID sub-TLV is carried within the MT-Port-Cap TLV [RFC6165], which in turn is carried in an IIH PDU.

14. Node Information Extensions

All SPB nodal information extensions travel within a new multi-topology capability TLV MT-Capability (type 144).

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+
|Type = MT-CAP   | = 144
+-+-+-+-+-+-+-+-+-+
|   Length       |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|O R R R|       MT ID            | (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  (sub-TLVs ... )
```

The format of this TLV is identical in its first 2 bytes to all current MT TLVs and carries the MT ID as defined in [MT].

The O (overload) bit carried in bit 16 has the same semantics as specified in [MT], but in the context of SPB adjacencies only.

There can be multiple MT-Capability TLVs present, depending on the amount of information that needs to be carried.

14.1. SPB Instance (SPB-Inst) Sub-TLV

The SPB-Inst sub-TLV gives the SPSourceID for this node/topology instance. This is the 20-bit value that is used in the formation of multicast DAs for frames originating from this node/instance. The SPSourceID occupies the upper 20 bits of the multicast DA together with 4 other bits (see the SPBM 802.1ah multicast DA address format section). This sub-TLV MUST be carried within the MT-Capability TLV in the fragment ZERO LSP. If there is an additional SPB instance, it

MUST be declared under a separate MT-Capability TLV and also carried in the fragment ZERO LSP.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type = SPB-Inst| = 1
+-+-+-+-+-+-+-+-+
|   Length      |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                CIST Root Identifier  (4 bytes)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                CIST Root Identifier (cont)  (4 bytes)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            CIST External Root Path Cost     (4 bytes)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Bridge Priority         |          (2 bytes)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R R R R R R R R R R R|V|              SPSourceID               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Num of Trees  |        (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  VLAN-ID (1) Tuples    (8 bytes)              |
|                  ...........................                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                   ...........................
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  VLAN-ID (N) Tuples    (8 bytes)              |
|                  ...........................                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where VLAN-ID tuples have the format as:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|U|M|A|  Res    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        ECT-ALGORITHM (32 bits)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Base VID (12 bits)    |   SPVID (12 bits)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

o Type: sub-TLV type 1.

o Length: Total number of bytes contained in the value field.

- CIST Root Identifier (64 bits): The CIST Root Identifier is for SPB interworking with Rapid STP (RSTP) and Multiple STP (MSTP) at SPT Region boundaries. This is an imported value from a spanning tree.

- CIST External Root Path Cost (32 bits): The CIST External Root Path Cost is the cost to root, derived from the spanning tree algorithm.

- Bridge Priority (16 bits): Bridge priority is the 16 bits that together with the 6 bytes of the System ID form the Bridge Identifier. This allows SPB to build a compatible spanning tree using link state by combining the Bridge Priority and the System ID to form the 8-byte Bridge Identifier. The 8-byte Bridge Identifier is also the input to the 16 predefined ECT tie-breaker algorithms.

- V bit (1 bit): The V bit (SPBM) indicates this SPSourceID is auto-allocated (Section 27.11 of [802.1aq]). If the V bit is clear, the SPSourceID has been configured and Must be unique. Allocation of SPSourceID is defined in IEEE [802.1aq]. Bridges running SPBM will allocate an SPSourceID if they are not configured with an explicit SPSourceID. The V bit allows neighbor bridges to determine if the auto-allocation was enabled. In the rare chance of a collision of SPsourceID allocation, the bridge with the highest priority Bridge Identifier will win conflicts. The lower priority bridge will be re-allocated; or, if the lower priority bridge is configured, it will not be allowed to join the SPT Region.

- SPSourceID: a 20-bit value used to construct multicast DAs as described below for multicast frames originating from the origin (SPB node) of the Link State Packet (LSP) that contains this TLV. More details are in IEEE [802.1aq].

- Number of Trees (8 bits): The Number of Trees is set to the number of {ECT-ALGORITHM, Base VID plus flags} tuples that follow. Each ECT-ALGORITHM has a Base VID, an SPVID, and flags described below. This Must contain at least the one ECT-ALGORITHM (00-80-C2-01).

Each VID Tuple consists of:

- U bit (1 bit): The U bit is set if this bridge is currently using this ECT-ALGORITHM for I-SIDs it sources or sinks. This is a strictly local indication; the semantics differ from the Use-Flag found in the Hello, which will set the Use-Flag if it sees other nodal U bits are set OR it sources or sinks itself.

o M bit (1 bit): The M bit indicates if this is SPBM or SPBV mode. When cleared, the mode is SPBV; when set, the mode is SPBM.

o A bit (1 bit): The A bit (SPB), when set, declares this is an SPVID with auto-allocation. The VID allocation logic details are in IEEE [802.1aq]. Since SPVIDs are allocated from a small pool of 12-bit resources, the chances of collision are high. To minimize collisions during auto-allocation, LSPs are initially advertised with the originating bridge setting the SPVID to 0. Only after learning the other bridges' SPVID allocations does this bridge re-advertise this sub-TLV with a non-zero SPVID. This will minimize but not eliminate the chance of a clash. In the event of a clash, the highest Bridge Identifier is used to select the winner, while the loser(s) with lower Bridge Identifier(s) Must withdraw their SPVID allocation(s) and select an alternative candidate for another trial. SPVID May also be configured. When the A bit is set to not specify auto-allocation and the SPVID is set to 0, this SPBV bridge is used for transit only within the SPB region. If a port is configured with the Base VID as a neighbor using RSTP or MSTP, the bridge will act as an ingress filter for that VID.

o ECT-ALGORITHM (4 bytes): ECT-ALGORITHM is advertised when the bridge supports a given ECT-ALGORITHM (by OUI/Index) on a given VID. This declaration Must match the declaration in the Hello PDU originating from the same bridge. The ECT-ALGORITHM and Base VID Must match what is generated in the IIHs of the same node. The ECT-ALGORITHM, Base VID tuples can come in any order, however. There are currently 17 worldwide unique 802.1aq defined ECT-ALGORITHMs given by values 00-80-C2-00 through 00-80-C2-10.

o Base VID (12 bits): The Base VID that associated the SPT Set via the ECT-ALGORITHM.

o SPVID (12 bits): The SPVID is the Shortest Path VID assigned for the Base VID to this node when using SPBV mode. It is not defined for SPBM mode and Must be 0 for SPBM mode B-VIDs.

14.1.1. SPB Instance Opaque ECT-ALGORITHM (SPB-I-OALG) Sub-TLV

There are multiple ECT algorithms defined for SPB; however, for the future, additional algorithms may be defined including but not limited to ECMP- or hash-based behaviors and (*,G) multicast trees. These algorithms will use this Optional TLV to define new algorithm parametric data. For tie-breaking parameters, there are two broad classes of algorithm, one that uses nodal data to break ties and one that uses link data to break ties. This sub-TLV is used to associate opaque tie-breaking data with a node. This sub-TLV, when present, MUST be carried within the MT-Capability TLV (along with a valid SPB-Inst sub-TLV). Multiple copies of this sub-TLV MAY be carried for different ECT-ALGORITHMs relating to this node.

There are of course many other uses of this opaque data that have yet to be defined.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type=SPB-I-OALG| = 2
+-+-+-+-+-+-+-+-+
|   Length      |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Opaque ECT-ALGORITHM    (4 bytes)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Opaque ECT Information (variable)            |
|                   .......................                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- o Type: sub-TLV type 2.
- o Length: Total number of bytes contained in the value field.
- o ECT-ALGORITHM: ECT-ALGORITHM is advertised when the bridge supports a given ECT-ALGORITHM (by OUI/Index) on a given VID.
- o ECT Information: ECT-ALGORITHM Information of variable length which SHOULD be in sub-TLV format with an IANA numbering space where appropriate.

15. Adjacency Information Extensions

15.1. SPB Link Metric (SPB-Metric) Sub-TLV

The SPB-Metric sub-TLV (type 29) occurs within the Multi-Topology Intermediate System Neighbor (MT-ISN) TLV (type 222) or within the Extended IS Reachability TLV (type 22). If this sub-TLV is not present for an IS-IS adjacency, then that adjacency Must not carry SPB traffic for the given topology instance.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type=SPB-Metric| = 29
+-+-+-+-+-+-+-+-+
|   Length      |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       SPB-LINK-METRIC                         |   (3 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Num of Ports    |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Port Identifier          |    ( 2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- o Type: sub-TLV type 29.

- o Length: Total number of bytes contained in the value field.

- o SPB-LINK-METRIC: the administrative cost or weight of using this link as a 24-bit unsigned number. This metric applies to the use of this link for SPB traffic only. Smaller numbers indicate lower weights and are more likely to carry SPB traffic. Only one metric is allowed per SPB instance per link. If multiple metrics are required, multiple SPB instances Must be used, either within IS-IS or within several independent IS-IS instances. If this metric is different at each end of a link, the maximum of the two values Must be used in all SPB calculations for the weight of this link. The maximum SPB-LINK-METRIC value 2^24 - 1 has a special significance; this value indicates that although the IS-IS adjacency has formed, incompatible values have been detected in parameters configured within SPB itself (for example, different regions), and the link Must Not be used for carrying SPB traffic. Full details are found in [802.1aq].

- o Num of Ports: the number of ports associated with this link.

- o Port Identifier: the standard IEEE port identifier used to build a spanning tree associated with this link.

15.1.1. SPB Adjacency Opaque ECT-ALGORITHM (SPB-A-OALG) Sub-TLV

There are multiple ECT algorithms defined for SPB; however, for the future, additional algorithms may be defined. The SPB-A-OALG sub-TLV occurs within the Multi-Topology Intermediate System TLV (type 222) or the Extended IS Reachability TLV (type 22). Multiple copies of this sub-TLV MAY be carried for different ECT-ALGORITHMs related to this adjacency.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type=SPB-A-OALG| = 30
+-+-+-+-+-+-+-+-+
|   Length      |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Opaque ECT Algorithm    (4 bytes)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Opaque ECT Information (variable)            |
|                  .........................                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- o Type: sub-TLV type 30.
- o Length: Total number of bytes contained in the value field.
- o ECT-ALGORITHM: ECT-ALGORITHM is advertised when the bridge supports a given ECT-ALGORITHM (by OUI/Index) on a given VID.
- o ECT Information: ECT-ALGORITHM Information of variable length in sub-TLV format using new IANA type values as appropriate.

16. Service Information Extensions

16.1. SPBM Service Identifier and Unicast Address (SPBM-SI) Sub-TLV

The SPBM-SI sub-TLV (type 3) is used to introduce service group membership on the originating node and/or to advertise an additional B-MAC unicast address present on, or reachable by the node.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|Type = SPBM-SI | = 3
+-+-+-+-+-+-+-+-+
|   Length      |     (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        B-MAC ADDRESS                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    B-MAC ADDRESS  (6 bytes)    |  Res. |   Base VID (12 bits)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T|R| Reserved  |                 I-SID  #1                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T|R| Reserved  |                 I-SID  #2                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                         .................
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T|R| Reserved  |                 I-SID  #n                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Type: sub-TLV type 3.

- Length: Total number of bytes contained in the value field.

- B-MAC ADDRESS: a unicast address of this node. It may be the single nodal address, or it may address a port or any other level of granularity relative to the node. In the case where the node only has one B-MAC address, this Should be the same as the SYSID of the node. To add multiple B-MACs this TLV MUST be repeated per additional B-MAC.

- Base VID (12 bits): The Base VID associated with the B-BMAC allows the linkage to the ECT-ALGORITHM and SPT Set defined in the SPB-Inst sub-TLV.

- I-SID #1 .. #n: 24-bit service group membership identifiers. If two nodes have an I-SID in common, intermediate nodes on the unique shortest path between them will create forwarding state for the related B-MAC addresses and will also construct multicast forwarding state using the I-SID and the node’s SPSourceID to construct a multicast DA as described in IEEE 802.1aq LSB. Each I-SID has a Transmit (T) and Receive (R) bit that indicates if the membership is as a transmitter, a receiver, or both (with both bits set). In the case where the Transmit (T) and Receive (R) bits are both zero, the I-SID instance is ignored for the purposes of distributed multicast computation, but the unicast B-MAC address Must be processed and installed at nodes providing transit to that address. If more I-SIDs are associated with a particular

B-MAC than can fit in a single sub-TLV, this sub-TLV can be repeated with the same B-MAC but with different I-SID values.

o Note: When the T bit is not set, an SPB May still multicast to all the other receiving members of this I-SID (those advertising with their R bits set), by configuring edge replication and serial unicast to each member locally.

The SPBM-SI sub-TLV, when present, MUST be carried within the MT-Capability TLV and can occur multiple times in any LSP fragment.

16.2. SPBV MAC Address (SPBV-ADDR) Sub-TLV

The SPBV-ADDR sub-TLV is IS-IS sub-TLV type 4. It Should be used for advertisement of Group MAC addresses in SPBV mode. Unicast MAC addresses will normally be distributed by reverse-path learning, but carrying them in this TLV is not precluded. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+
| Type=SPBV-ADDR|   = 4            (1 byte)
+-+-+-+-+-+-+-+-+-+
|   Length      |                  (1 byte)
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R| SR|       SPVID           |  (2 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+
|T|R| Reserved  |       MAC 1 Address              |  (1+6 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+
                         ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+
|T|R| Reserved  |       MAC N Address              |  (1+6 bytes)
+-+-+-+-+-+-+-+-+-+-+-+-+-+...+-+-+-+-+-+-+-+-+-+-+-+-+
```

o Type: sub-TLV type 4.

o Length: Total number of bytes contained in the value field. The number of MAC address associated with the SPVID is computed by (Length - 2)/7.

o SR bits (2 bits): The SR bits are the service requirement parameter from MMRP. The service requirement parameters have the value 0 (Forward all Groups) and 1 (Forward All Unregistered Groups) defined. However, this attribute May also be missing. So the SR bits are defined as 0 not declared, 1 Forward all Groups, and 2 Forward All Unregistered Groups. The two 'R' reserved bits immediately preceding these SR bits Shall be set to zero when originating this sub-TLV and Shall be ignored on receipt.

o SPVID (12 bits): The SPVID and by association Base VID and the ECT-ALGORITHM and SPT Set that the MAC addresses defined below will use. If the SPVID is not allocated the SPVID Value is 0. Note that if the ECT-ALGORITHM in use is spanning tree algorithm this value Must be populated with the Base VID and the MAC Must be populated.

o T bit (1 bit): This is the Transmit allowed bit for a following group MAC address. This is an indication that the Group MAC address in the context of the SPVID of the bridge advertising this Group MAC Must be installed in the FDB of transit bridges, when the bridge computing the trees is on the corresponding ECT-ALGORITHM shortest path between the bridge advertising this MAC with the T bit set and any receiver of this Group MAC address. A bridge that does not advertise this bit set for a MAC address Must Not cause multicast forwarding state to be installed on other transit bridges in the network for traffic originating from that bridge.

o R bit (1 bit): This is the Receive allowed bit for the following MAC address. This is an indication that MAC addresses as the receiver Must be populated and installed when the bridge computing the trees lies on the corresponding shortest path for this ECT-ALGORITHM between this receiver and any transmitter to this MAC address. An entry that does not have this bit set for a Group MAC address is prevented from receiving on this Group MAC address because transit bridges Must Not install multicast forwarding state towards it in their FDBs.

o MAC Address (48 bits): The MAC address declares this bridge as part of the multicast interest for this destination MAC address. Multicast trees can be efficiently constructed for destination by populating FDB entries for the subset of the shortest path tree that connects the bridges supporting the MAC address. This replaces the function of MMRP for SPTs. The T and R bits above have meaning as specified above.

The SPBV-ADDR sub-TLV, when present, MUST be carried within the MT-Capability TLV and can occur multiple times in any LSP fragment.

17. Security Considerations

This document adds no additional security risks to IS-IS, nor does it provide any additional security for IS-IS when used in a configured environment or a single-operator domain such as a data center.

However, this protocol may be used in a zero-configuration environment. Zero configuration may apply to the automatic detection and formation of an IS-IS adjacency (forming an NNI port). Likewise, zero configuration may apply to the automatic detection of VLAN-tagged traffic and the formation of a UNI port, with resultant I-SID advertisements.

If zero configuration methods are used to autoconfigure NNIs or UNIs, there are intrinsic security concerns that should be mitigated with authentication procedures for the above cases. Such procedures are beyond the scope of this document and are yet to be defined.

In addition, this protocol can create significant amounts of multicast state when an I-SID is advertised with the T bit set. Extra care should be taken to ensure that this cannot be used in a denial-of-service attack [RFC4732] in a zero-configuration environment.

18. IANA Considerations

Note that the NLPID value 0xC1 [RFC6328] used in the IIH PDUs has already been assigned by IANA for the purpose of 802.1aq; therefore, no further action is required for this code point.

Since 802.1aq operates within the IS-IS Multi-Topology framework, every sub-TLV MUST occur in the context of the proper MT TLV (with the exception of the SPB-Metric sub-TLV, which MAY travel in TLV 22 where its MT ID is unspecified but implied to be 0). IANA has allocated sub-TLVs for three Multi-Topology TLVs per 802.1aq. These are the MT-Port-Cap TLV [RFC6165] used in the IIH, the MT-Capability TLV (new) used within the LSP, and finally the MT-ISN TLV [MT] used to contain adjacency information within the LSP.

This document creates the following TLVs and sub-TLVs within the IIH and LSP PDUs MT TLVs as described below. The '*' indicates new IANA assignments (per this document). Other entries are shown to provide context only.

The MT-Capability TLV is the only TLV that required a new sub-registry. Type value 144 has been assigned, with a starting sub-TLV value of 1, and managed by Expert Review.

```
   +-----+----+-----------------+--------+------+-------------+
   | PDU |TLV | SUB-TLV         | TYPE   | TYPE | #OCCURRENCE |
   +-----+----+-----------------+--------+------+-------------+
     IIH
          MT-Port-Cap              143
*               SPB-MCID                    4       1
*               SPB-Digest                  5       >=0
*               SPB-B-VID                   6       1

     LSP
*         MT-Capability            144              >=1
*               SPB-Inst                    1       1
*               SPB-I-OALG                  2       >=0
*               SPBM-SI                     3       >=0
*               SPBV-ADDR                   4       >=0

          MT-ISN                   222
       or Extended IS Reachability  22
*               SPB-Metric                 29       1
*               SPB-A-OALG                 30       >=0
```

19. References

19.1. Normative References

[802.1aq] "Standard for Local and Metropolitan Area Networks: Virtual Bridges and Virtual Bridged Local Area Networks - Amendment 9: Shortest Path Bridging", IEEE P802.1aq, Draft 4.6, 2012.

[IS-IS] ISO/IEC 10589:2002, Second Edition, "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", 2002.

[MT] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC6165] Banerjee, A. and D. Ward, "Extensions to IS-IS for Layer-2 Systems", RFC 6165, April 2011.

[RFC6328] Eastlake 3rd, D., "IANA Considerations for Network Layer Protocol Identifiers", BCP 164, RFC 6328, July 2011.

19.2. Informative References

[802.1ag] "Standard for Local and Metropolitan Area Networks / Virtual Bridged Local Area Networks / Amendment 5: Connectivity Fault Management", IEEE STD 802.1ag, 2007.

[MMRP] "Standard for Local and Metropolitan Area Networks Virtual Bridged Local Area Networks - Amendment 07: Multiple Registration Protocol", IEEE STD 802.1ak, 2007.

[PB] "Standard for Local and Metropolitan Area Networks / Virtual Bridged Local Area Networks / Amendment 4: Provider Bridges", IEEE STD 802.1ad, 2005.

[PBB] "Standard for Local and Metropolitan Area Networks / Virtual Bridged Local Area Networks / Amendment 7: Provider Backbone Bridges", IEEE STD 802.1ah, 2008.

[RFC4732] Handley, M., Ed., Rescorla, E., Ed., and IAB, "Internet Denial-of-Service Considerations", RFC 4732, December 2006.

[Y.1731] ITU-T, "OAM Functions and Mechanisms for Ethernet based networks", ITU-T Y.1731, 2006.

20. Acknowledgments

The authors would like to thank Ayan Banerjee, Mick Seaman, Janos Farkas, Les Ginsberg, Stewart Bryant , Donald Eastlake, Matthew Bocci and Mike Shand for contributions and/or detailed review.

Authors' Addresses

Don Fedyk (editor)
Alcatel-Lucent
Groton, MA 01450
USA
EMail: Donald.Fedyk@alcatel-lucent.com

Peter Ashwood-Smith (editor)
Huawei Technologies Canada Ltd.
303 Terry Fox Drive, Suite 400
Kanata, Ontario, K2K 3J1
CANADA
EMail: Peter.AshwoodSmith@huawei.com

Dave Allan
Ericsson
300 Holger Way
San Jose, CA 95134
USA
EMail: david.i.allan@ericsson.com

Nigel Bragg
Ciena Limited
Ciena House
43-51 Worship Street
London EC2A 2DX
UK
EMail: nbragg@ciena.com

Paul Unbehagen
Avaya
8742 Lucent Boulevard
Highlands Ranch, CO 80129
USA
EMail: unbehagen@avaya.com