Proposed text of a response to liaison-itut-sg15-ols-288-0310.doc – rev 5

After reading and discussing the referenced liaison letter from ITU-T SG15 to IEEE 802.1, the members of 802.1 believe that some information about previous IEEE 802.1 work would be helpful to SG15. Specifically, the Multiple VLAN Registration Protocol (MVRP), the Multiple I-SID Registration Protocol (MIRP), and the Multiple MAC Registration Protocol (MMRP) should be of interest.

MVRP/MIRP/MMRP introduction

These three protocol s are all based on a single underlying protocol with state machines, called the Multiple Registration Protocol (MRP); they are three "applications" of MRP. They work together in IEEE 802.1 networks to:

1. Propagate VLAN ID (MVRP) or MAC address (MMRP) registration information in order to determine to what parts of a network a given service or MAC address need be transmitted; and

2. Propagate "New" messages for VLAN IDs (MVRP) or I-SIDs (MIRP) in order to indicate that certain learned (not configured) MAC address location information should be discarded.

MVRP and MMRP are defined in IEEE Std. 802.1Q-2005. MIRP is defined, and MVRP is modified, in IEEE Std. 802.1Qbe-2010. What may not be clear to the reader of these documents is that:

a) MMRP can register unicast, as well as multicast addresses; and

b) These three protocols are ***not*** tied to the spanning tree protocols, STP, RSTP, and MSTP.

c) The "New" messages can be used to discard learned information without using the service registration (pruning) state machines of MVRP and MIRP.

It is certainly true that, for MAC address learning to work at all, a service (e.g., a VLAN) must be fully connected (spanned) and there must be a unique path from bridge to bridge between any two connected points in the service (tree connectivity). In this sense, a "spanning tree" is necessary for MAC address learning. However, M{V,I,M}RP impose no requirement that that spanning tree be established via one of the IEEE 802.1Q protocols. They operate independently of the means used to establish the tree. The tree could be created via IS-IS, configuration, or by some other means.

In operation, these protocols are efficient in terms of exchanges. There are no explicit acknowledgements, only exchanges of state information, but the net result is a protocol that corrects very quickly for the loss of a single protocol data unit (PDU), and corrects on a slower timescale for any possible connectivity changes (e.g., reroute, protection, or additional ports) or multiple PDU losses. The PDUs are peer-bridge-to-peer-bridge, and each PDU carries information about all services shared by those bridges. The information carried in the PDUs is propagated through a bridge from one link to another link along the same path as the data in the various services is propagated. Thus, the PDUs do not follow the spanning tree(s), but the per-service information carried by the PDUs does follow the spanning tree(s). The initiator of an action, i.e., the bridge or station that first discovers (or is

configured) that a service or MAC address is or is not required at a certain point in the network, need not know the topology of the network; it only needs to transmit a registration to its neighbor(s), and that registration information then propagates as required through the network.

Note that the definition of "neighbor" varies with the use case. A neighbor is always the nearest peer bridge that participates in the protocol. Depending on the usage scenario, a neighbor can be a physically adjacent bridge, or it can be a peer edge bridge (e.g., an I-component) across a backbone network. Data plane parameters determine neighbors. At a given level of encapsulation, a bridge that participates in one of these protocols terminates the multicast MAC address of the protocol PDUs and does not propagate frames with that destination MAC address; a bridge that does not participate in a protocol passes frames that protocol's multicast destination address as ordinary data.

MVRP/MIRP/MMRP capabilities

In the usage scenario described in your liaison letter, device A would, after losing one of its links to the network, transmit two MVRP or MIRP PDUs over the remaining link. Each PDU would contain a "New" command for each of the services that formerly used the failed link and are now using the remaining link. If necessary, that network bridge could propagate those "New" commands further through the network. The "New" commands cause the receiving bridges to discard learned MAC address locations for the indicated service on all links carrying that service *except* for the link on which the "New" command was received. The "New" command also serves to indicate that broadcasts and unknown multicasts or unicasts for its service need to be propagated over the link to bridge A. Similarly, bridge A can use MMRP to register its need to receive a given unicast or multicast MAC address. (This request can be tied to a particular service, or can apply to all services.)

When purging MAC addresses, MVRP/MIRP do not purge all addresses; they do not purge addresses learned on the interface from which the "New" command is received. This often purges more addresses than are absolutely necessary for a given fault in a given network. However, in the absence of sure knowledge of the overall topology of the network, MVRP and MIRP purge the fewest possible number of MAC addresses.

We may also note that MVRP/MIRP transmit the "New" message for a service, causing a purge, on the *new* link, not the old link, and that it is transmitted when that service is *brought up* on the new link, not when the service is *lost* on the old link. Users and vendors of bridges have found, over nearly 30 years of experience, that there is no point in flooding traffic until the new path is available, and that it is best to "black-hole" the traffic destined for the failed link until then.

MVRP/MIRP/MMRP limitations

Some of the requirements expressed by the ITU-T liaison are met by MVRP, MIRP, and MMRP, and some are not.

1. We believe that MIRP and MVRP can purge dynamic addresses quite satisfactorily, in the case where arbitrary network topologies are supported.

2. MIRP and MVRP cannot take advantage of configured knowledge of the network topology in order to purge fewer addresses.

3. MMRP can populate MAC address database entries with unicast or multicast MAC addresses. This information supersedes learned MAC address location information. Subsequently learned information cannot override the MMRP-propagated information. This may or may not be the behavior perceived as necessary by SG15. (However, see the next point.)

4. Neither MVRP, MIRP, nor MMRP can purge single learned MAC addresses (or, as mentioned, register them). This has found to be an unnecessary feature in enterprise bridged networks, which can be very similar in nature to provider networks, because individual stations, when moved, immediately transmit one or two ARP broadcasts that result in updating the bridges' learned MAC address information.

5. If propagating the extent of a service (also called, "pruning") is not a capability that is required, MVRP/MIRP carry more information than is necessary (2.8 bits, rather than 1 bit, per service) to indicate the need to forget learned MAC addresses.

Using 802.1ag/Y.1731 PDUs for MAC address database operations

While it may be true that Y.1731 OAM/802.1ag CFM is implemented by all of the devices in a given network, that fact does not make OAM/CFM the right protocol to use for all purposes. OAM/CFM is defined in terms of MEPs and MIPs, which reside in ports. MAC learning is a matter of filtering databases and forwarding decisions. The awkwardness of the LinkTrace mechanism (as opposed, for example, to IP traceroute) demonstrates the difficulty inherent in connecting OAM/CFM to the filtering database.

The MIPs or MEPs of an MA are often placed in locations, e.g., an 802.1aj Two-Port MAC Relay, where no notice need be taken of MAC address operations. Furthermore, the need to signal MAC flushing may extend beyond the range of a Maintenance Association, and indeed, may involve potentially circular relationships that cannot be covered by Maintenance Associations. We may illustrate this using the example in the SG15 liaison letter. If a connection existed between bridges A and B that is outside the green cloud, some means (perhaps 802.1 MSTP, perhaps some other means) of breaking the loop is required. Should a loss of a link results in that break being healed to restore connectivity, it is not clear that any configuration of MEPs and MIPs will provide the necessary points of control over which the proper MAC flushing signals can be sent. MVRP/MIRP/MMRP can accommodate any topology. MVRP/MIRP would presumably be used in the portion of that network exterior to the service cloud.

Other Layer 2 technologies

IEEE 802.1 is not the only author of Ethernet technology standards. Other technologies, e.g., TRILL, G.8032 rings, or VPLS exist. Each has some means for signaling the need to flush learned MAC address information. MVRP/MIRP/MMRP has the advantage over at least some of these technologies in that MVRP/MIRP/MMRP work in-band.

Summary

To summarize, it is the opinion of 802.1 that:

a) MMRP, MIRP, and MVRP together come close to meeting the needs expressed by SG15;

b) Further exchanges between SG15 and 802.1 can result in either a modification of SG15's expressed requirements to match those of these protocols, a project in 802.1 to enhance these protocols to meet SG15's needs, or both; and that

c) Because 802.1ag/Y.1731 PDUs are tied to MEPs and MIPs that reside in ports, rather than filtering and forwarding functions, an approach to signaling MAC learning via those PDUs is awkward from an architectural point of view, and we would not recommend using that approach.

We hope that further exchanges on this subject will lead to a better understanding of our groups' capabilities and needs, and to a satisfactory solution to this problem.