Detecting SRP domain boundary ports

Tony Jeffree

Summary

There is a problem with the domain boundary port detection method as described in P802.1Qat/D5.0. Two approaches were proposed during the January 2010 802.1 Interim meeting in http://www.ieee802.org/1/files/public/docs2010/at-cgunther-srp-lldp.pdf that would deal with the problem; one based on changes to the existing method and the other based on use of LLDP (IEEE Std 802.1AB). This paper discusses some of the issues and proposes a third solution based on extensions to SRP.

1. Introduction

The current method of detection of SRP domains and their boundaries leverages the operation of MSRP itself; essentially, if a port can see that there is an entity connected to its LAN that is issuing MSRP protocol relevant to a particular SR class, then it assumes that it is not a boundary Port. The current method has a characteristic that the domain boundary expands as the set of stations/Bridges that are involved in streaming for a given SR class expands, and similarly contracts as the set of stations/Bridges that are involved in streaming for a given SR class contracts. For non-streaming devices that are outside the boundary, but communicate using the streaming priorities, if the domain boundary expands to encompass part of the communication path that those devices are using, the priority they are using will be mapped to 0 at the boundary. If the boundary then contracts again, the mapping to 0 ceases. So the effective priority used by such devices is either affected, or not affected, according to the presence or absence of active streams in their neighbourhood. It is undesirable for the network behaviour to “flap” in this way, as one of the possible side-effects is that the non-stream traffic could experience misordering as the priority remapping is removed; it is therefore preferable to establish a domain boundary that is static, and which only changes based on changes in topology rather than on changes in the streaming behaviour of its components.

2. Potential solution 1

The solution proposed in slide 10 of at-cgunther-srp-lldp.pdf changes the behaviour of the existing boundary detection so that the boundary expands as the set of stations/Bridges that are involved in streaming for a given SR class expands, but does not contract as the set of devices involved contracts. So, the only event that can cause a Bridge port to change from believing itself to be inside the SRP domain to believing itself to be at the boundary is if the Port leaves the active topology. While this solution fixes the problem of the priority mapping behaviour changing as stream reservations come and go, it does not appear to fix the problem of what happens when a device attached to a Bridge Port is reconfigured to no longer support SRP without any attendant change that causes that port to be removed from the active topology. This solution therefore doesn’t address the comment in bullet 3 of slide 9, and so would seem to be inappropriate.

3. Potential solution 2

The solution proposed in slide 11 of at-cgunther-srp-lldp.pdf involves the inclusion of LLDP (IEEE Std 802.1AB) in the set of mandatory features required for support of 802.1Qat. Essentially, what is proposed is that LLDP be used to exchange tuples of information; for each SR class supported on a given port, the SRP capable device transmits the priority on which it expects to transmit and/or receive streaming information. By comparing, for a given SR class, the priority value supported on the port with the priority supported by the neighbouring device(s), the port can determine whether it is within the boundary or at the boundary.

1. If the priority values match, then it is within a domain and not at the boundary;
2. If the priority values don't match, or if multiple values have been received from the neighbouring device[1] that give conflicting information, or if there is no neighbouring device, then the port is at the boundary of a domain. Note that MSRP doesn't support half duplex Ethernet, so if the Port is attached to a half duplex Ethernet LAN, then it is a boundary Port by definition.
3. If the neighbouring device(s) disappear, and therefore LLDPDUs are no longer received, then LLDP will in time purge its remote database of any stored information received from the old neighbour, and the port will (as a consequence of the second bullet) become a boundary port.

This solution is certainly a workable one; it has the characteristic that it will create stable domain boundaries that change only when the active topology changes, or when devices are reconfigured either to support or not support MSRP. It also, as observed in at-cgunther-srp-lldp.pdf, is a step in the right direction if SRPv2 goes down the route of automatic configuration for SR classes and their associated priorities.

However, the fact that it relies on the availability of LLDP means that it raises the bar in terms of the support requirements for low-cost consumer AV Bridges; it would no longer be possible to contemplate building un-managed consumer AV Bridges, as LLDP conformance requires the device to be capable of storage and retrieval of the LLDP MIB data[2]. So this has complexity implications with regard to the design of consumer devices that may be considered to be undesirable.

4. Potential solution 3

Solution 2 will work very well, except for the fact that it requires the addition of LLDP, plus SNMP (or an equivalent retrieval mechanism), to the set of protocols that are needed for support of AV. So, the obvious alternative is to exchange the desired information, but do it either via a simpler protocol, or via an extension to the existing MSRP protocol machinery. The appeal of the latter approach is that the MSRP state machines already exist, and already cope with the problem of keeping declarations and registrations of such data current, both on point-to-point links and shared media LANs, both of which we need to be able to handle in an AVB environment.

Extending MSRP to make this possible would appear to consist of the following changes:

1. Define a 4th attribute type for registering attribute values that are tuples of an SR class and its associated priority.
2. Define the "first value" to be a single octet where the upper 4 bit "nibble" defines the SR class and the lower "nibble" defines the associated priority[3]. Define the encoding of SR Class such that the highest class (A) is encoded as 6 and the lowest (G) as 0, and the priority as 7 through 0 as normal. Define the incrementing rule for the case where the number of values is >1 such that both the SR Class and the Priority fields are incremented by 1. Using the tuple of SR class and priority as the value registered means that a distinct MRP state machine will be instantiated for each value registered,

[1] It is possible that a transitory condition could exist where the remote device's priority for a given SR class is being reconfigured, and as a result, two different priorities are registered locally until the old value ages out. While this conflict remains, the Port has to be considered to be a boundary Port.

[2] It doesn't require SNMP support per se, but in the absence of SNMP support it requires that "...the system shall provide storage and retrieval capability equivalent to the functionality specified in 10.1 for the operating mode being implemented." So if it isn't SNMP, then it needs to be something equivalent.

[3] Alternatively, use 2 octets, one for SR Class, the other for Priority.

which will allow values for different SR classes to be registered independently of each other and hence, one could be changed without affecting the state of the other registration. If different priorities have been registered for the same SR class, then this is a boundary even if one of the remote priorities matches the local priority for that SR class (see footnote 1).

3. Define how these attributes are propagated. The desired behaviour here is that they are NOT propagated by MAP at all – we are only interested in information local to the LAN attached to a port, not stuff propagated over more than one hop. So the rule is that on any port that is part of the active topology, SRP makes declarations of the SR classes and associated priorities supported by the port, regardless of what may have been received on other Ports.

The end result will be that each port will know (a) what its own supported SR classes are, and what priorities are associated with them (from existing information defined in Qav) and (b) what set of SR classes and associated priorities exists in the other devices attached to the LAN (this could be the empty set if no-one on that LAN supports MSRP). If the match is perfect for a given SR class, i.e., there is one registration on the Port, and it exactly matches the value being declared by that Port, then the Port is a core port; if the remote data does not match the local data (more than one priority registered for a given SR class, or no registration at all, or just one registration but the priority differs from the declared value), then the Port is a boundary port. Also, if the Port is attached to a half duplex Ethernet LAN, then it is a boundary port.

The choice of encoding and incrementing rule means that if one class (B) is supported, then the first value will be {5, 2} (class B, priority 2, assuming the defaults are in play) and the number of values field will be set to 1. If 2 classes are supported, then the first value is still {5, 2} but the number of values field is 2. When incremented according to the rule, {5, 2} becomes {6, 3}, i.e., class A priority 3, which is what we need for the default case.

The approach described does not require any change to the existing PDU structure for MSRP other than the addition of the 4th attribute type; however, if the priorities used for multiple SR classes are discontinuous (e.g., A=5, B=2), then the packing is less optimal, as separate vcetors have to be used for each tuple. As this is unlikely to be the case, and as we are in any case anticipating a small number of SR classes to be the norm, this doesn't seem to be an issue.

In operation, this approach will result in a stable configuration of a set of SRP domains within a contiguous LAN, with boundaries being created as soon as the devices attached to the LAN are powered on. Boundaries will be established wherever there is a configuration mismatch; it is therefore appropriate for MSRP to convert "Talker Advertise" to "Talker Fail" at such boundaries, in order to prevent a stream from being established across a set of incompatible SRP domains[4].

[4] It is also trivial to extend the set of rules that define a boundary to include 802.1AS detection of buffered repeaters etc..

5. Conclusion

It is feasible to extend MSRP to fix this problem. The additional complexity in MSRP seems to be very minor compared with the additional complexity involved in supporting LLDP and all that it brings with it (SNMP MIBs, SNMP protocol stack, etc., or equivalent functionality), so I consider this approach to be preferable to the LLDP approach, although in terms of the mechanics of boundary detection, the two approaches would be exactly equivalent. The MSRP-based approach may also have applications elsewhere, such as in DCB[5].

---

[5] As the problem being addressed here is rather similar to the problem being addressed in DCB, where they also have a need to discover the configuration of adjacent systems, it may be that a similar approach would be applicable there as well, particularly if DCB needs to use one or more MRP-based protocol for other reasons, and has no other need to make use of LLDP.