Class of Service in 802.1

Norman Finn, Cisco Systems, Inc.

1.0 End-to-end Class Of Service Indication

An important facility provided by 802.1p/D3 [1] is an end-to-end layer 2 Class of Service (CoS) indication, where "end-to-end" refers to the users of the MAC layer. In Figure 1, we see an example which, though somewhat farfetched, serves to illustrate the value of an end-to-end CoS indication.

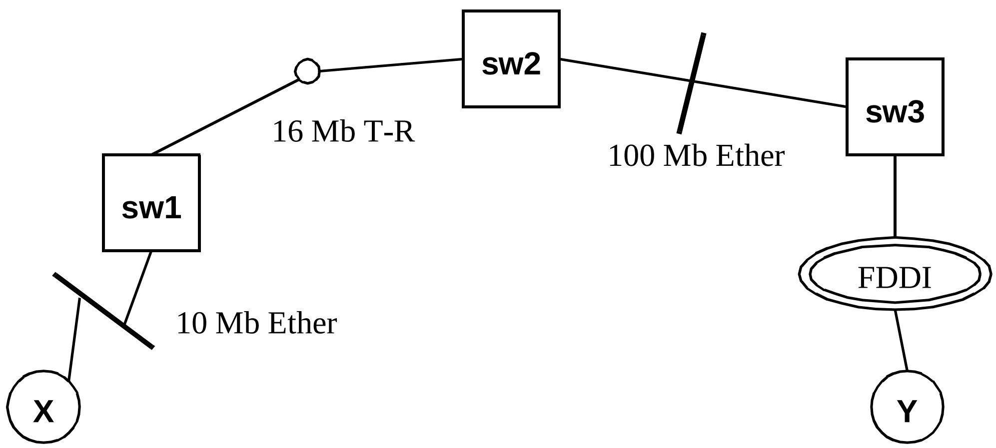

FIGURE 1. CoS Bridging Example

If forced to use only the access priorities available in Token Ring and FDDI LANs, but not in 802.3/Ethernet LANs, a packet sent from Y to X would lose its priority indication when it traversed the 100 Mb Ether link, and would have to use a default value across the T-R link. A packet from X to Y would start off with no access priority indication, and would get no benefit from the access priorities available on FDDI and T-R.

However, a packet can carry its CoS indication in such a manner that it is carried from LAN segment to LAN segment as the packet is switched over different media. This indication can provide both an access priority and an output queue selection at each switch output port, regardless of the packet's history of transmission over different media. Furthermore, if this end-to-end indication is transparent to non-CoS-aware switches, then those switches do not affect the CoS-aware switches' ability to provide the CoS facility.

2.0 Methods Available for End-to-end CoS Indication

At present, there are four methods on the table for an end-to-end CoS indication:

1. Use the high-order 3 bits of the VLAN-ID field. (Discussed at the Boston interim meeting and in [2].)
2. Using the GARP protocol, assign each multicast MAC address a 3-bit CoS value on a per-LAN segment basis. (The obvious way to use 802.1p/D3.)
3. Redefine the Universal/Local bit in the destination MAC address as a 1-bit CoS indication. (As discussed in [3].)
4. Create a protocol to assign blocks of locally-administered unicast addresses to end stations in such a manner that the CoS indication is in a fixed location in the MAC address.

3.0 Making a Choice

Should we (802.1) standardize one, two, three, or all four methods for end-to-end Class of Service indication? As it stands, three methods (all except the VLAN-ID method) are sanctioned by 802.1p/D3, but none are explicitly selected. Table 1 summarizes the advantages and disadvantages of each method.

TABLE 1. Comparison of End-to-End CoS Solutions

	multi/ unicast	**CoS levels**	**Frame length**	**bridge has CoS in:**	**CoS-unaware bridges**	**CoS-unaware end stations**
high bits of VLAN-ID	multicast/ unicast	8	Adds 4 bytes	frame	transparent[a]	conflict avoided
assignment via GARP	multicast only	8	Same size	fwd table	transparent	conflict avoided
redefine U/L bit	multicast/ unicast[b]	2	Same size	frame or fwd table	transparent[c]	possible conflict[d]
blocks of local addrs	unicast only	many	Same size	frame or fwd table	transparent[c]	conflict[d]

a. If bridge does not parse above layer 2, and if frame length is no problem.

b. Subject to incompatibilities with IGMP-aware bridges.

c. Subject to issues with bridges learning all of a station’s MAC addresses.

d. Conflict with existing protocols for allocating locally-administered MAC addresses.

3.1 Applicability to Multicast and Unicast Frames

Assignment via GARP is not useful for unicast traffic, both because of the inordinate overhead of registration, and because existing Layer 3 protocols are unable to associate multiple MAC addresses with a single Layer 3 address. This seems a significant drawback for this method, compared to the VLAN-ID and U/L bit methods, which are applicable to one degree or another with both unicast and multicast traffic.

Redefining the U/L bit as a CoS indication has certain problems when CoS-unaware bridges are present. GARP does not eliminate the need for the IETF's IGMP protocol for IP multicasts, because IGMP includes essential layer 3 information. GARP could accommodate moving the CoS registration to an explicit U/L bit. However, there currently exist a number of "smart" bridges, which utilize the IETF's IGMP protocol to perform the GARP multicast pruning function. These bridges would be confused if the U/L were used as a CoS indication, independent of the rest of the MAC address.

It is tempting to pair GARP assignment, which is multicast only, and local address block allocation, which is unicast only. These methods would be compatible together in a bridge that based CoS on its forwarding table. It is also desirable, however, for unicast and multicast frames to use a single mechanism.

3.2 Number of Levels of CoS Supported

Clearly, it is desirable that any end-to-end CoS indication be able to express at least as many CoS levels as the number of access priority levels currently supported by existing LANs. Several media, including 802.5 and FDDI, support 8 priority levels. Several recent contributions have argued that two levels are insufficient for significant classes of applications.

3.3 Frame Length Issues

The VLAN tag either adds 4 bytes to every frame or reduces the number of bytes available for layer 3 by the same amount. If VLAN tagging is used, this penalty is already accepted. If not, and if an installation has CoS- and VLAN-unaware bridges that cannot handle slightly-too-large frames, and if reduction of MTU by 4 bytes is unacceptable, then that installation is perhaps overconstrained for the VLAN-ID solution.

3.4 Bridge Internal Model for CoS

This is a fundamental issue. The GARP registration, U/L bit, and local address block assignment models are all compatible with a bridge architecture that obtains CoS information from the bridge forwarding table (as described in 802.1p/D3). The VLAN-ID, U/L bit, and local address block assignment models are all compatible with a bridge architecture that obtains CoS information directly from the frame. Therefore, the combination of "GARP registration for multicast, but VLAN-ID for unicast and/or multicast" would make an exceptionally poor choice, as it would require bridges to implement both architectural models.

3.5 Compatibility with CoS-unaware bridges

Three of the four models (excepting the GARP registration model) exhibit some incompatibilities with CoS-unaware bridges. Potential problems with the VLAN-ID model and CoS-unaware bridges are 1) "smart" bridges that parse above layer 2 will be unable to handle the CoS tag; and 2) the user must choose between extra-large packets and reduced MTU, both of which could conceivably be unacceptable.

The two methods that carry the unicast CoS indication in the destination MAC address must solve the problem of ensuring that Cos-unaware bridges can learn about all of an end station's MAC addresses. That is, for every bridge that forwards frames to a given MAC address with imbedded CoS information, that bridge must receive frames with that same MAC address and imbedded CoS indication in the source MAC address field.

There are a number of ways to improve this situation. An endstation receiving traffic for different CoS levels can keep track of its various correspondents and use the appropriate MAC source address(es) with each. This greatly increases the complexity of the end station's MAC layer. It might send periodic broadcast messages using all its MAC addresses and CoS indications in the source MAC fields, or it might rotate through all of its MAC addresses and CoS indications in order or at random as it sends frames. Both of these techniques result in expanding the working size of bridges' forwarding databases which have no interest in many of the MAC addresses received on broadcasts. If nothing is done, CoS-unaware bridges will unnecessarily flood traffic for MAC address/CoS indications they have not seen in the source MAC address field.

3.6 Compatibility with CoS-unaware end stations

None of these methods allow an end station designed for current standards to receive unicast traffic at multiple CoS levels without modification. The two methods that use locally administered MAC addresses may conflict with usage of the D/L bit by existing protocols. At worst, a locally administered MAC address assigned for CoS purposes may duplicate another end station's MAC address that is locally assigned by an existing protocol. The VLAN-ID and GARP assignment methods do not interfere with existing end stations, unless giant packets are used on the same LAN segment with an end station that complains about their existence on that LAN segment.

4.0 Summary

The "blocks of local addresses" should be ruled out because no allocation algorithm exists, and because running such an algorithm involves numerous problems:

1. It would be difficult to add yet another protocol to a diskless end station's (or bridge's or router's) boot PROMs.
2. It would be difficult to change an end station's MAC address if it used its L3 address with its own, universally allocated MAC address, before acquiring its locally allocated address.
3. It is difficult to ensure that the same end station gets the same block of MAC addresses each time it boots, and difficult for some L3 protocols (IP, for example) to adjust to changed MAC addresses.

Of the remaining possibilities, no solution is perfect. In the author's opinion, the balance of pros and cons favors the VLAN-ID method. This one method can be used for both unicast and multicast traffic, it is independent of addressing, is adequately transparent to CoS-unaware end stations and bridges. By separating the addressing and CoS functions, it minimizes undesirable interactions between these functions, at the cost of 4 bytes per packet.

5.0 References

[1] 802.1p/D3 "Draft Standard for Traffic Class and Dynamic Multicast Filtering Services in Bridged Local Area Networks (Draft Supplement to 802.1D)"

[2] Norman Finn, Gideon Prat, Anil Rijsinghani, John Wakerly, “802.1 Class of Service Alternatives”, contribution presented to Boston IEEE 802.1 interim meeting, June, 1996.

[3] Norman Finn, Gideon Prat, John Wakerly, “Applying 802.1p/D3”, contribution presented to Boston IEEE 802.1 interim meeting, June, 1996.