Internet Draft
draft-hamid-issll-rsvp-cap-dsmark-00.txt

Syed, Hamid
Nortel Networks

February, 2001

**The DS marking Capability Negotiation:**
**A Usage Case for the RSVP CAP Object**

Status of this Memo



## 1. Abstract

The DCLASS object is proposed in [DCLASS] to represent and carry Differentiated Services Code Points (DSCPs) within RSVP messages. The principle use of the DCLASS object is to carry DSCP information between a DS network and upstream nodes that may wish to mark packets with DSCP values. A network element in the DS network determines the value for DSCP which is further carried as a DCLASS object in RSVP RESV message to the sender host. The RSVP capability negotiation CAP Object [RSVP_CAP] is proposed to convey end host/upstream node Capabilities to the downstream network.

This draft proposes a usage case for the capability object (CAP object) in an Intserv/Diffserv network and defines one bit in the CAP field of

the CAP object to convey the host/upstream node's capability or willingness to mark the downstream packets.

## 2. Introduction

The mechanics of using RSVP [RSVP] signaling and the DCLASS object for requesting and applying the QoS in a differentiated services [DS] network are described fully in [INTDIFF]. It assumes architecture with RSVP senders and receivers and a differentiated services network somewhere between the sender and the receiver. At least one RSVP aware network element resides in the DiffServ network. This network element interacts with RSVP messages arriving from outside the DS network.

The principle use of the DCLASS object is to carry DSCP information between a DS network and upstream nodes that may wish to mark packets with DSCP values. A network element in the DS network determines the appropriate DSCP value which is further carried as a DCLASS object in the RSVP RESV message to the sender host. If the network element determines that the request represented by the PATH and RESV messages is admissible to the DiffServ network, a decision is made to mark the arriving data packets for this traffic using MF classification, or to request upstream marking of packets with the appropriate DSCPs. If the network element decides that packets are to be marked at the sender host for the data traffic, it adds a DCLASS object in the RSVP RESV message to the host. The use and format of DCLASS object is fully specified in [DCLASS]. Technically the downstream network edge device only needs to install the packet forwarding rules assuming the classification and marking will be performed by the upstream device when it provides a DCLASS value to the upstream node. There may be situations where the upstream node/network does not understand DCLASS so it will not be able to perform a packet marking or the upstream node may decide to leave the packet marking to the downstream device. In such scenarios the downstream network device need to install all classification, marking and forwarding rules for the bearer traffic. The decision at the downstream device on what configuration rules are needed for a flow request must be made on the RSVP RESV message. This requires that the downstream node be able to know whether the upstream node/network will perform packet marking or it will outsource it to the downstream network. The current definition of the DCLASS object does not address such a scenario.

This draft attempts to solve the problem by proposing a usage case for the RSVP CAP object [RSVP_CAP] in an Intserv-Diffserv network. The intelligent decisions of where the data packets should be marked and what configuration rules are required to be installed can be made at the downstream network nodes assuming that the network edge devices receives a prior indication of the marking capability of the upstream nodes. The draft also defines one bit in the CAP field of the CAP object to convey the host/upstream node's marking capability or willingness to the downstream nodes.

## 3. Marking Capability Negotiation

The processing of the bearer traffic at the DiffServ edge device could be different for the case where an upstream node performs the packet marking and the case where the downstream edge device has

to perform the packet classification and marking. In the former case, the DCLASS object is sent to the device and the device may perform device configuration necessary for packet forwarding based on the DSCP received in the packet header. While in the later situation, the network device needs to install filters to carry out packet classification and marking/forwarding of the bearer packets. A priori knowledge of the upstream node's capabilities would enable the edge device to figure out whether a DCLASS should be provided to the upstream node and prepare the device for packet forwarding only or install necessary packet classification, marking and forwarding rules for the incoming traffic. In the current definition of the DCLASS object, the network edge device inserts the DCLASS object in the RSVP RESV message without having any priori knowledge of whether or not the host can make use of this object. Moreover, the definition of DCLASS object allows any DS domain to supply the object on a flow to the upstream DS domains. There may be situations where the sender host or an upstream node is not capable or is not willing to mark the packets. The provision of DCLASS object to such nodes would be meaningless as the edge device has to install the multifield classification and marking rules to treat the packets. Advance knowledge of whether or not the upstream node is capable and willing to perform the packet marking can enable the edge device to make intelligent decisions on what filters need to be installed and whether or not to insert a DCLASS object in the RESV message.

The capability object has been defined as a mechanism for conveying a node's capabilities or willingness in RSVP messages. As an example, we will focus on the marking capability of nodes throughout this document by defining a single bit for host marking information to be carried in the CAP field inside the CAP object of RSVP PATH message. To explain this usage case of CAP object, we will describe two scenarios

- Host/Edge router interaction
- Border Router/Border Router interaction

It should be noted that how and when the packets will be marked and what configuration at the device is required for the flow request is a decision governed by the network policies. The network policy domain may or may not trust a end host marking. Hence, even though the network may have supplied the DCLASS object to the end host on request (via CAP) it may overwrite the marking based on the domain policy.

## 3.1 Host/Edge Router Capability Negotiation

The advance knowledge of the end host's capabilities may help the network edge devices to make policy decisions on end host's requests. These capabilities can be indicated in the RSVP PATH message to the downstream edge devices.

The end hosts can be classified in two categories. The first category groups those end hosts capable of marking downstream packets and decide to do so. The second category of hosts either do not have the capability to mark packets or they decide not to mark packets. In either case, the

network element needs to know the host's packet marking capability or willingness. This information can help the network element to decide whether or not a DCLASS object must be added in a RSVP message for the flow and what kind of packet filtering rules be installed for the bearer traffic. One way to convey the host capability/willingness to the network is to use the CAP object in the RSVP PATH message. We give examples here to explain the scenarios.

If the sender host is ready to mark the downstream traffic (based on the DCLASS provided by the network element), it sets the marking bit of the CAP field inside the CAP object of the RSVP PATH message. On receiving the RSVP message, the network element at the DS edge records the host marking capability with the PATH state. It then resets the marking bit and sends the RSVP message to the downstream nodes. The treatment of the CAP object at the downstream nodes will be explained in the next section. For now, consider the RESV message comes back to the edge device, which performs the necessary admission control. If the network element determines that the request represented by the PATH and RESV messages is admissible to the DiffServ network, it adds a DCLASS object after consulting the recorded state. It may decide to overwrite any DCLASS object inserted by the downstream node/domain based on its own domain policies. The edge device may now be able to decide what kind of filtering rules could be installed for the bearer traffic. Assuming the network policy allows the edge device to trust the packet marking from the end host, it would only configure the device for packet forwarding based on the received DS code points in the packet header.

Another example could be the end host that is not capable of downstream packet marking. This either will not include a CAP object or the host will reset the marking bit of the CAP object as an indication of his unwillingness to mark packets. The network edge router will then know that the upstream node/end host does not require a DCLASS object. The edge router, in this case, would be responsible for enforcing the packet classification and marking rules in addition to the packet forwarding rules.

### 3.2 Boundry router/Boundry Router Interaction

The CAP object could be carried in the PATH message end-to-end. The RSVP PATH message is generated by the end host. The network edge router 'A' of the DS domain processes the message, resets the marking bit of the CAP object (if it comes as set from the host) and passes the PATH message to the next RSVP Hop. For a DS domain, the boundary router 'B' of the access/stub network receives the RSVP PATH message as next RSVP enabled node (Figure 1). It may set the marking bit again to advertise the marking capability of its own domain. The decision must be governed by the domain policy. The ingress boundary router 'C' of the downstream domain receives the CAP object with the marking bit set providing an indication of the

marking capability of the upstream node/domain. It again stores this information as the PATH state, resets the marking bit and passes it to the downstream RSVP enabled network element. The boundary router 'D' of this domain may decide to set the marking bit again based on the domain policy. The PATH message may pass through more domains like this until

it is received by the host. The RSVP RESV message is then generated and passed through the same route. The RSVP message arrives at the router 'C' and it may contain a DCLASS object provided by an downstream node/domain. The PATH state of router 'C' indicates that the upstream node/domain is capable of packet marking and a DCLASS object is to be passed back. The domain policy/admission control decisions of router 'C'may not allow the router to use the same DCLASS value as it received from the downstream. So it may decide to overwrite the DCLASS value. The edge router 'A' may also decide to remark the DCLASS value in the RESV message following its admission control outcome and knowing the end host's willingness for packet marking. Finally, the end host receives the DCLASS value in RESV message and it may start marking the downstream packets with the appropriate DSCP.

In the above scenario, the routers 'A' and 'C' would install the device configuration rules based on the knowledge of the upstream node/network capabilities and the DS code point provided by the domain policy.

Once again, It should be noted that how and when the packets will be marked is a decision governed by the network policies. The network policy domain may or may not trust the upstream node marking (specially in the case of end host marking). Hence, even though the network may have supplied the DCLASS object to the end host on request (via CAP) it may overwrite the marking based on the domain policy.

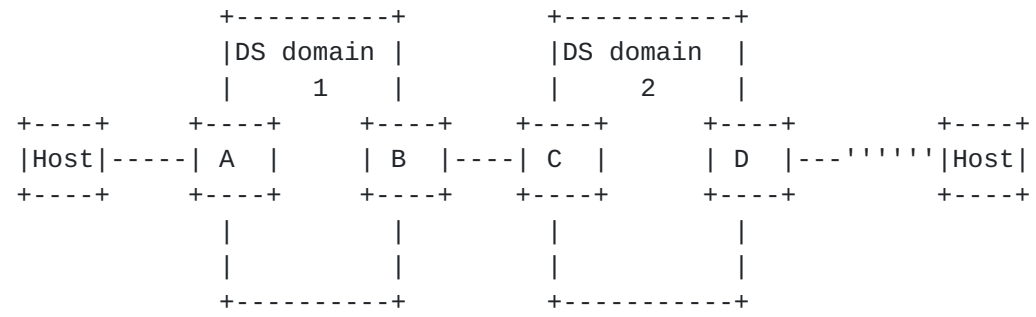


Figure 1

## 4. The D_Mark Bit

The first bit in the CAP field can be used to indicate the marking capability/willingness of the downstream nodes as follows

0x01: D_MARK
The host marking capability/willingness identifier.
If D_MARK bit is reset, the sender host/upstream node is not able to mark packets
If D_MARK bit is set, the sender host/upstream node is able/willing to mark packets

Note: The processing of the D_MARK bit should follow the rules specified by the Capability Object definition [RSVP_CAP].

## 5. Deployment Scenarios

There are a number of hosts today that do have the marking capability and they even do not depend on a DCLASS object from the network. The marking is based on a default mapping from requested service type to the DSCP. In this section, we will briefly address the deployment scenarios for such hosts which do mark without signaling the network about their marking capability.

If a host does not provide a CAP object, then the network edge must be provisioned (or be given policies) as to how it should react. This may be one of:

- send a DCLASS object.
- install a filter to mark the appropriate flow at the edge.
- do both.

The problem here is ensuring that the mapping configured in the host matches the allowed mappings configured in the edge router. If there is a mismatch, the edge router will, at best, remark the packets to match its policies (possibly resulting in a treatment different from that expected by the host) or, at worst, mark packets as non-conforming and discard them. The policy may be for a specific host address, for a specific interface, for a specific edge router or for the entire domain. The bottom line is that manual provisioning would be required in the interim until hosts support the CAP option. Once hosts support the CAP option, manual provisioning would no longer be required.

In a multi-domain scenario, the boundary router 'B' could be the first and the only router in the first DS domain who is dealing with the CAP/DCLASS objects (maintaining the state information and deciding for a DSCP for the upstream end host). This will allow only one router in a domain with the knowledge of the host's capability and will be the one responsible for deciding/providing a DCLASS object in a RSVP RESV message. In this scenario, the boundary router 'B' becomes the DS edge for the end host.

## 6. References

[RSVP_CAP] Syed, H., "Capability Negotiation: The RSVP CAP Object.", IETF<draft-ietf-issll-rsvp-cap-02.txt>, Februray 2001.

[INTDIFF] Bernet, Y., Yavatkar, R., Ford, P., Baker, F., Zhang, L., Speer, M., Braden, R., Davie, B., Wroclawski, J., "Integrated Services Operation over Diffserv Networks", RFC 2998, November 2000

[DS] An Architecture for Differentiated Services. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, RFC 2475, December 1998.

[RSVP] Braden, R. ed., "Resource ReSerVation Protocol (RSVP) - Functional Specification.", IETF RFC 2205, Sep. 1997.

[DCLASS] Bernet, Y., "Format of the RSVP DCLASS Object",
RFC , Oct., 1999.

## 7. Acknowledgments

Thanks to Bill Gage, Yoram Bernet, Goran Janevski, Gary Kenward, kwok Ho chan, Muhammad Jaseemuddin and Louis-Nicolas Hamer for reviewing this draft and providing useful input.

## 8. Author's Address

Syed, Hamid
Nortel Networks
100 - Constellation Crescent,
Nepean, ON K2G 6J8
Phone: (613) 763-6553
Email: hmsyed@nortelnetworks.com

## 9. Full Copyright Statement