Y.Bernet, Microsoft

Internet Draft
Document: draft-bernet-dclass-01.txt June, 1999

Usage and Format of the DCLASS Object With RSVP Signaling

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet- Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Distribution of this memo is unlimited.

1. Abstract

RSVP signaling may be used to enhance the manageability of application traffic's QoS in a differentiated service (diff-serv) network [intdiff]. In this model, certain network elements within or at the edges of the diff-serv network may use RSVP messages to effect admission control or to apply QoS policy. One mechanism by which network elements may apply QoS policy is by causing a DCLASS object to be returned to a sending host in an RSVP RESV message. The DCLASS object indicates one or more diff-serv codepoints (DSCPs) that the sender should include when submitting packets on the admitted flow, to the diff-serv network. This draft describes the usage and format of the DCLASS object.

3. Signaling Protocol

This section describes the mechanics of using RSVP signaling and the

DCLASS object for effecting admission control and applying QoS policy within a diff-serv network. It assumes a standard RSVP sender

and a diff-serv network somewhere in the path between sender and receiver. At least one RSVP aware network element resides in the diff-serv network. This network element may be a policy enforcement point (PEP) associated with a PDP, or may simply act as an admission control agent, admitting or denying resource requests based exclusively on the availability of resources. The network element is typically a router and will be considered to be so for the purpose of this draft.

The sender composes a standard RSVP PATH message and sends it towards the receiver on the remote end of the diff-serv network. The PATH message traverses one or more network elements that are PEPs and/or admission control agents for the diff-serv network. These elements install appropriate state and forward the PATH message towards the receiver. If admission control is successful downstream of the diff-serv network, then a RESV message will arrive from the direction of the receiver. As this message arrives at the PEPs and/or admission control agents that are RSVP enabled, each of these network elements must make a decision regarding the admissibility of the signaled flow to the diff-serv network.

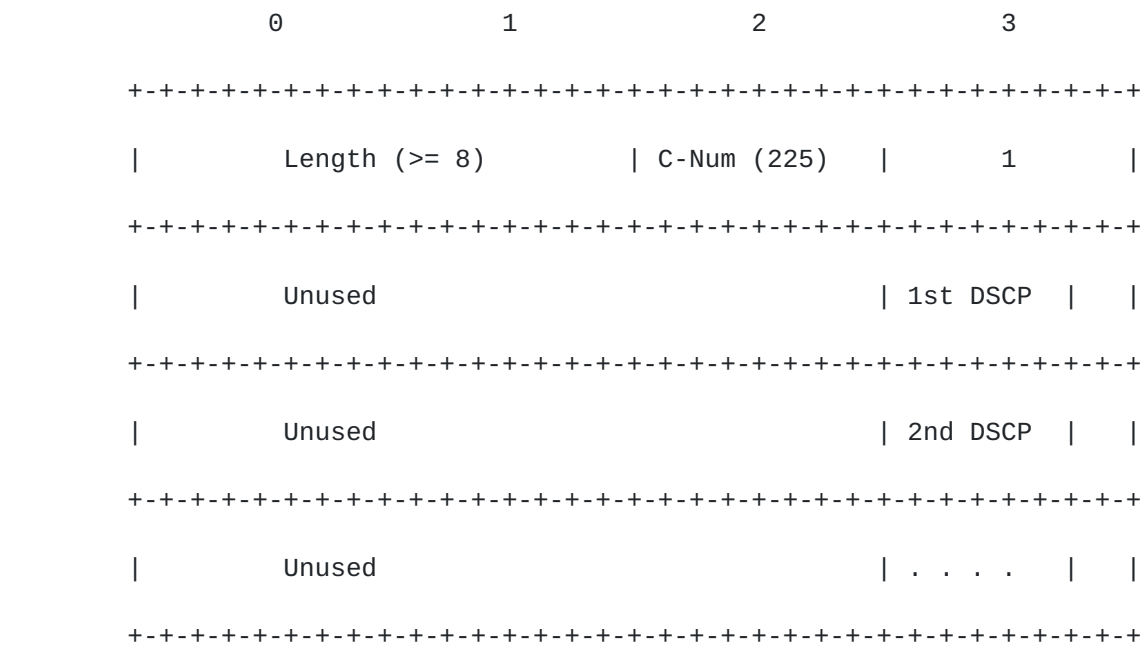
If the network element determines that the request represented by the PATH and RESV messages is admissible to the diff-serv network, it must decide which diff-serv service level (or behaviour aggregate) is appropriate for the traffic represented in the RSVP request. It then adds a DCLASS object containing one or more DSCPs corresponding to the behaviour aggregate, to the RESV message. The RESV message is then sent upstream towards the RSVP sender.

If the network element determines that the RSVP request is not admissible to the diff-serv network, it sends a RESV error message towards the receiver. No DCLASS is required.

Note that a network element may terminate RSVP signaling, in which case it effectively provides admission control to all regions of the network downstream (including the receiver). In this case, no actual RESV message will arrive from the receiver. Instead, the network element may act as a proxy, composing the RESV message on behalf of the downstream nodes.

4. Format of the DCLASS Object

The DCLASS object has the following format:



The first word contains the standard RSVP object header (the Class Num for the DCLASS object is 225). The length field indicates the total object length in bytes. The object header is followed by one or more 32-bit words, each containing a DSCP in the six high-order bits of the least significant byte. The length field in the object header indicates the number of DSCPs included in the object.

The network may return multiple DSCPs in the DCLASS object in order to enable the host to discriminate sub-flows within a behaviour aggregate. For example, in the case of the AF PHB group [AF], the network may return the DSCPs 001010, 001100, and 001110 corresponding to increasing levels of drop precedence within Class 1 of the AF PHB group. Note that the DSCPs must be ordered within the object in increasing order of drop probability or non-conformance.

The first DSCP always corresponds to the most favorable treatment by the network. Further interpretation of DSCP sets is dependent on the specific service requested by the host and is beyond the scope of this draft.

Hosts may simply mark all packets on a signaled flow with the first (most favored treatment) DSCP in the list. Alternatively, hosts may 'pre-demote' certain packets by marking them with one of the less favorable DSCPs returned. In the first case, the network controls which packets are treated less favorably in the case of congestion or non-conformance to a signaled traffic profile. In the latter case, the host may exercise control over which packets are treated less favorably by the network.

Note that the Class-Num for the DCLASS object is chosen from the space of unknown class objects that should be ignored and forwarded by nodes that do not recognize it. This is to assure maximal backward compatibility.

5. Admission Control Functionality

From a black-box perspective, admission control and policy functionality amounts to the decision whether to accept or reject a request and the determination of the DSCPs that should be used for

the corresponding traffic. The specific details of admission control are beyond the scope of this document. In general the admission control decision is based both on resource availability and on policies regarding the use of resources in the diff-serv network. The admission control decision made by RSVP aware network elements represents both considerations.

In order to decide whether the RSVP request is admissible in terms of resource availability, one or more network elements within or at the boundary of the diff-serv network must understand the impact that admission would have on specific diff-serv resources, as well as the availability of these resources along the relevant data path in the diff-serv network.

In order to decide whether the RSVP request is admissible in terms of policy, the network element may use identity objects describing users and/or applications that may be included in the request. The router may act as a PEP/PDP and use data from a policy database or directory to aid in this decision.

See Appendix A for a simple mechanism for configurable resource based admission control.

8. Security Considerations

There are no security considerations beyond those of standard RSVP.

9. References

[INTDIFF], Bernet, Y., Yavatkar, R., Ford, P., Baker, F., Zhang, L., Speer, M., Braden, R., Davie, B., "Integrated Services Operation over Diffserv Networks", Internet Draft, June 1999

[AF], Heinanen, J., Baker, F., Weiss, W., Wroclawski, J., "Assured Forwarding PHB Group", RFC 2597, June 1999

10. Acknowledgments

Thanks to Fred Baker and Carol Iturralde for reviewing this draft. Thanks to Ramesh Pabbati, Tim Moore, Bruce Davie and Kam Lee for input.

11. Author's Addresses

Bernet, Yoram
Microsoft
One Microsoft Way,
Redmond, WA 98052
Phone: (425) 936-9568
Email: yoramb@microsoft.com

Appendix A - Simple Configurable Resource Based Admission Control

Routers may use quite sophisticated mechanisms in making the admission control decision, including policy considerations, various intra-domain signaling protocols, results of traffic monitoring and so on. It is recommended that the following basic functionality be provided to enable simple resource based admission control in the absence of more sophisticated mechanisms. This functionality can be used with configurable, standalone routers. It applies to standard RSVP/Intserv requests. This minimal functionality assumes only a single DSCP is included in the DCLASS object, but may readily be extended to support multiple DSCPs.

It must be possible to configure two tables in the router. These are described below.

A.1 Service Type to DSCP Mapping

One table provides a mapping from the intserv service-type specified in the RSVP request to a DSCP that can be used to obtain a corresponding service in the diff-serv network. This table contains a row for each intserv service type for which a mapping is available. Each row has the following format:

Intserv service type : DSCP

The table would typically contain at least three rows; one for Guaranteed service, one for Controlled Load service and one for Best-Effort service. (The best-effort service will typically map to DSCP 000000, but may be overridden). It should be possible to add rows for as-yet-undefined service types.

This table allows the network administrator to statically configure a DSCP that the router will return in the DCLASS object for an admitted RSVP request. In general, more sophisticated and likely more dynamic mechanisms may be used to determine the DSCP to be returned in the DCLASS object. In this case, these mechanisms may override the static table based mapping.

A.2 Quantitative Resource Availability

Standard intserv requests are quantitative in nature. They include token bucket parameters describing the resources required by the traffic for which admission is requested. The second table enables the network administrator to statically configure quantitative parameters to be used by the router when making an admission control decision for quantitative service requests. Each row in this table has the following form:

DSCP : Token bucket profile

The first column specifies those DSCPs for which quantitative admission control is applied. The second column specifies the token bucket parameters which represent the total resources available in

the diff-serv network to accommodate traffic in the service class specified by the DSCP.