Network Working Group   L. Berger
Request for Comments: 2380   FORE Systems
Category: Standards Track   August 1998

RSVP over ATM Implementation Requirements

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Abstract

This memo presents specific implementation requirements for running RSVP over ATM switched virtual circuits (SVCs). It presents requirements that ensure interoperability between multiple implementations and conformance to the RSVP and Integrated Services specifications. A separate document [5] provides specific guidelines for running over today's ATM networks. The general problem is discussed in [9]. Integrated Services to ATM service mappings are covered in [6]. The full set of documents present the background and information needed to implement Integrated Services and RSVP over ATM.

Table of Contents

1. Introduction

This memo discusses running IP over ATM in an environment where SVCs are used to support QoS flows and RSVP is used as the internet level QoS signaling protocol. It applies when using CLIP/ION, LANE2.0 and MPOA [4] methods for supporting IP over ATM. The general issues related to running RSVP [8] over ATM have been covered in several papers including [9] and other earlier work. This document is intended as a companion to [9,5]. The reader should be familiar with both documents.

This document defines the specific requirements for implementations using ATM UNI3.x and 4.0. These requirements must be adhered to by all RSVP over ATM implementations to ensure interoperability. Further recommendations to guide implementers of RSVP over ATM are provided in [5].

The rest of this section will define terms and assumptions. Section 2 will cover implementation guidelines common to all RSVP session. Section 3 will cover implementation guidelines specific to multicast sessions.

1.1 Terms

The terms "reservation" and "flow" are used in many contexts, often with different meaning. These terms are used in this document with the following meaning:

o Reservation is used in this document to refer to an RSVP initiated request for resources. RSVP initiates requests for resources based on RESV message processing. RESV messages that simply refresh state do not trigger resource requests. Resource requests may be made based on RSVP sessions and RSVP reservation styles. RSVP styles dictate whether the reserved resources are used by one sender or shared by multiple senders. See [8] for details of each. Each new request is referred to in this document as an RSVP reservation, or simply reservation.

- o Flow is used to refer to the data traffic associated with a particular reservation. The specific meaning of flow is RSVP style dependent. For shared style reservations, there is one flow per session. For distinct style reservations, there is one flow per sender (per session).

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [7].

1.2 Assumptions

The following assumptions are made:

- o RSVP

  We assume RSVP as the internet signaling protocol which is described in [8]. The reader is assumed to be familiar with [8].

- o IPv4 and IPv6

  RSVP support has been defined for both IPv4 and IPv6. The guidelines in this document are intended to be used to support RSVP with either IPv4 or IPv6. This document does not require one version over the other.

- o Best effort service model

  The current Internet only supports best effort service. We assume that as additional components of the Integrated Services model are defined, best effort service must continue to be supported.

- o ATM UNI 3.x and 4.0

  We assume ATM service as defined by UNI 3.x and 4.0. ATM provides both point-to-point and point-to-multipoint Virtual Circuits (VCs) with a specified Quality of Service (QoS). ATM provides both Permanent Virtual Circuits (PVCs) and Switched Virtual Circuits (SVCs). In the Permanent Virtual Circuit (PVC) environment, PVCs are typically used as point-to-point link replacements. So the support issues are similar to point-to-point links. This memo assumes that SVCs are used to support RSVP over ATM.

2. General RSVP Session Support

This section provides implementation requirements that are common for all (both unicast and multicast) RSVP sessions. The section covers VC usage, QoS VC initiation, VC teardown, handling requested changes in QoS, and encapsulation.

2.1 RSVP Message VC Usage

There are several RSVP Message VC Usage options available to implementers. Implementers must select which VC to use for RSVP messages and how to aggregate RSVP sessions over QoS VCs. These options have been covered in [9] and some specific implementation guidelines are stated in [5]. In order to ensure interoperability between implementations that follow different options, RSVP over ATM implementations MUST NOT send RSVP (control) messages on the same QoS VC as RSVP associated data packets. RSVP over ATM implementations MAY send RSVP messages on either the best effort data path or on a separate control VC.

Since RSVP (control) messages and RSVP associated data packets are not sent on the same VCs, it is possible for a VC supporting one type of traffic to fail while the other remains in place. When the VC associated with data packets fails and cannot be reestablished, RSVP SHOULD treat this as an allocation failure. When the VC used to forward RSVP control messages is abnormally released and cannot be reestablished, the RSVP associated QoS VCs MUST also be released. The release of the associated data VCs is required to maintain the synchronization between forwarding and reservation states for the associated data flows.

2.2 VC Initiation

There is an apparent mismatch between RSVP and ATM. Specifically, RSVP control is receiver oriented and ATM control is sender oriented. This initially may seem like a major issue but really is not. While RSVP reservation (RESV) requests are generated at the receiver, actual allocation of resources takes place at the subnet sender.

For data flows, this means that subnet senders MUST establish all QoS VCs and the RSVP enabled subnet receiver MUST be able to accept incoming QoS VCs. These restrictions are consistent with RSVP version 1 processing rules and allow senders to use different flow to VC mappings and even different QoS renegotiation techniques without interoperability problems. All RSVP over ATM approaches that have VCs initiated and controlled by the subnet senders will interoperate. Figure 1 shows this model of data flow VC initiation.

```
                    Data Flow ==========>

                +-----+
                |     |      -------------->  +----+
                | Src |    -------------->    | R1 |
                |    *|  -------------->      +----+
                +-----+       QoS VCs
                    /\
                    ||
                 VC ||
                 Initiator
```

Figure 1: Data Flow VC Initiation

RSVP over ATM implementations MAY send data in the backwards direction on an RSVP initiated QoS point-to-point VC. When sending in the backwards data path, the sender MUST ensure that the data conforms to the backwards direction traffic parameters. Since the traffic parameters are set by the VC initiator, it is quite likely that no resources will be requested for traffic originating at the called party. It should be noted that the backwards data path is not available with point-to-multipoint VCs.

2.3 VC Teardown

VCs supporting IP over ATM data are typically torndown based on inactivity timers. This mechanism is used since IP is connectionless and there is therefore no way to know when a VC is no longer needed. Since RSVP provides explicit mechanisms (messages and timeouts) to determine when an associated data VC is no longer needed, the traditional VC timeout mechanisms are not needed. Additionally, under normal operations RSVP implementations expect to be able to allocate resources and have those resources remain allocated until released at the direction of RSVP. Therefore, data VCs set up to support RSVP controlled flows should only be released at the direction of RSVP. Such VCs must not be timed out due to inactivity by either the VC initiator or the VC receiver. This conflicts with VCs timing out as described in RFC 1755 [11], section 3.4 on VC Teardown. RFC 1755 recommends tearing down a VC that is inactive for a certain length of time. Twenty minutes is recommended. This timeout is typically implemented at both the VC initiator and the VC receiver. Although, section 3.1 of the update to RFC 1755 [12] states that inactivity timers must not be used at the VC receiver.

In RSVP over ATM implementations, the configurable inactivity timer mentioned in [11] MUST be set to "infinite" for VCs initiated at the request of RSVP. Setting the inactivity timer value at the VC initiator should not be problematic since the proper value can be relayed internally at the originator. Setting the inactivity timer at the VC receiver is more difficult, and would require some mechanism to signal that an incoming VC was RSVP initiated. To avoid this complexity and to conform to [12], RSVP over ATM implementations MUST not use an inactivity timer to clear any received connections.

2.4 Dynamic QoS

As stated in [9], there is a mismatch in the service provided by RSVP and that provided by ATM UNI3.x and 4.0. RSVP allows modifications to QoS parameters at any time while ATM does not support any modifications to QoS parameters post VC setup. See [9] for more detail.

The method for supporting changes in RSVP reservations is to attempt to replace an existing VC with a new appropriately sized VC. During setup of the replacement VC, the old VC MUST be left in place unmodified. The old VC is left unmodified to minimize interruption of QoS data delivery. Once the replacement VC is established, data transmission is shifted to the new VC, and only then is the old VC closed.

If setup of the replacement VC fails, then the old QoS VC MUST continue to be used. When the new reservation is greater than the old reservation, the reservation request MUST be answered with an error. When the new reservation is less than the old reservation, the request MUST be treated as if the modification was successful. While leaving the larger allocation in place is suboptimal, it maximizes delivery of service to the user. The behavior is also required in order to conform to RSVP error handling as defined in sections 2.5, 3.1.8 and 3.11.2 of [8]. Implementations SHOULD retry replacing a too large VC after some appropriate elapsed time.

One additional issue is that only one QoS change can be processed at one time per reservation. If the (RSVP) requested QoS is changed while the first replacement VC is still being setup, then the replacement VC SHOULD BE released and the whole VC replacement process is restarted. Implementations MAY also limit number of changes processed in a time period per [9].

2.5 Encapsulation

There are multiple encapsulation options for data sent over RSVP triggered QoS VCs. All RSVP over ATM implementations MUST be able to support LLC encapsulation per RFC 1483 [10] on such QoS VCs. Implementations MAY negotiate alternative encapsulations using the B-LLI negotiation procedures defined in ATM Signalling, see [11] for details. When a QoS VC is only being used to carry IP packets, implementations SHOULD negotiate VC based multiplexing to avoid incurring the overhead of the LLC header.

3. Multicast RSVP Session Support

There are several aspects to running RSVP over ATM that are unique to multicast sessions. This section addresses multicast end-point identification, multicast data distribution, multicast receiver transitions and next-hops requesting different QoS values (heterogeneity) which includes the handling of multicast best effort receivers. Handling of best effort receivers is not strictly an RSVP issue, but needs to be addressed by any RSVP over ATM implementation in order to maintain expected best effort internet service.

3.1 Data VC Management for Heterogeneous Sessions

The issues relating to data VC management of heterogeneous sessions are covered in detail in [9]. In summary, heterogeneity occurs when receivers request different levels of QoS within a single session, and also when some receivers do not request any QoS. Both types of heterogeneity are shown in figure 2.

```
                                 +-----+
                        +------> | R1 |
                        |        +-----+
                        |
                        |        +-----+
           +-----+ -----+   +--> | R2 |
           |     | ---------+    +-----+       Receiver Request Types:
           | Src |                             ---->  QoS 1 and QoS 2
           |     | .........+    +-----+       ....>  Best-Effort
           +-----+ .....+   +..> | R3 |
                        :        +-----+
                    /\  :
                    ||  :        +-----+
                    ||  +......> | R4 |
                    ||           +-----+
                  Single
               IP Mulicast
                  Group
```

Figure 2: Types of Multicast Receivers

[9] provides four models for dealing with heterogeneity: full heterogeneity, limited heterogeneity, homogeneous, and modified homogeneous models. No matter which model or combination of models is used by an implementation, implementations MUST NOT normally send

more than one copy of a particular data packet to a particular next-hop (ATM end-point). Some transient duplicate transmission is acceptable, but only during VC setup and transition.

Implementations MUST also ensure that data traffic is sent to best effort receivers. Data traffic MAY be sent to best effort receivers via best effort or QoS VCs as is appropriate for the implemented model. In all cases, implementations MUST NOT create VCs in such a way that data cannot be sent to best effort receivers. This includes the case of not being able to add a best effort receiver to a QoS VC, but does not include the case where best effort VCs cannot be setup. The failure to establish best effort VCs is considered to be a general IP over ATM failure and is therefore beyond the scope of this document.

There is an interesting interaction between dynamic QoS and heterogeneous requests when using the limited heterogeneity, homogeneous, or modified homogeneous models. In the case where a RESV message is received from a new next-hop and the requested resources are larger than any existing reservation, both dynamic QoS and heterogeneity need to be addressed. A key issue is whether to first add the new next-hop or to change to the new QoS. This is a fairly straight forward special case. Since the older, smaller reservation does not support the new next-hop, the dynamic QoS process SHOULD be initiated first. Since the new QoS is only needed by the new next-hop, it SHOULD be the first end-point of the new VC. This way signaling is minimized when the setup to the new next-hop fails.

3.2 Multicast End-Point Identification

Implementations must be able to identify ATM end-points participating in an IP multicast group. The ATM end-points will be IP multicast receivers and/or next-hops. Both QoS and best effort end-points must be identified. RSVP next-hop information will usually provide QoS end-points, but not best effort end-points.

There is a special case where RSVP next-hop information will not provide the appropriate end-points. This occurs when a next-hop is not RSVP capable and RSVP is being automatically tunneled. In this case a PATH message travels through a non-RSVP egress router on the way to the next-hop RSVP node. When the next-hop RSVP node sends a RESV message it may arrive at the source via a different route than used by the PATH message. The source will get the RESV message, but will not know which ATM end-point should be associated with the reservation. For unicast sessions, there is no problem since the ATM end-point will be the IP next-hop router. There is a problem with multicast, since multicast routing may not be able to uniquely identify the IP next-hop router. It is therefore possible for a multicast end-point to not be properly identified.

In certain cases it is also possible to identify the list of all best effort end-points. Some multicast over ATM control mechanisms, such as MARS in mesh mode, can be used to identify all end-points of a multicast group. Also, some multicast routing protocols can provide all next-hops for a particular multicast group. In both cases, RSVP over ATM implementations can obtain a full list of end-points, both QoS and non-QoS, using the appropriate mechanisms. The full list can then be compared against the RSVP identified end-points to determine the list of best effort receivers.

While there are cases where QoS and best effort end-points can be identified, there is no straightforward solution to uniquely identifying end-points of multicast traffic handled by non-RSVP next-hops. The preferred solution is to use multicast control mechanisms and routing protocols that support unique end-point identification. In cases where such mechanisms and routing protocols are unavailable, all IP routers that will be used to support RSVP over ATM should support RSVP. To ensure proper behavior, baseline RSVP over ATM implementations MUST only establish RSVP-initiated VCs to RSVP capable end-points. It is permissible to allow a user to override this behavior.

3.3 Multicast Data Distribution

Two basic models exist for IP multicast data distribution over ATM. In one model, senders establish point-to-multipoint VCs to all ATM attached destinations, and data is then sent over these VCs. This model is often called "multicast mesh" or "VC mesh" mode distribution. In the second model, senders send data over point-to-point VCs to a central point and the central point relays the data onto point-to-multipoint VCs that have been established to all receivers of the IP multicast group. This model is often referred to as "multicast server" mode distribution. Figure 3 shows data flow for both modes of IP multicast data distribution.

```
                            _________
                           /         \
                          / Multicast \
                          \   Server  /
                           \_________/
                             ^  |  |
                             |  |  +--------+
              +-----+        |  |           |
              |     | -------+  |           |         Data Flow:
              | Src | ...+......|....+      V         ---->  Server
              |     |    :           :    +-----+     ....>  Mesh
              +-----+    :      |    +...>| R1 |
                         :      |         +-----+
                         :      V
                         :    +-----+
                         +..> | R2 |
                              +-----+
```

Figure 3: IP Multicast Data Distribution Over ATM

The goal of RSVP over ATM solutions is to ensure that IP multicast data is distributed with appropriate QoS. Current multicast servers [1,2] do not support any mechanisms for communicating QoS requirements to a multicast server. For this reason, RSVP over ATM implementations SHOULD support "mesh-mode" distribution for RSVP controlled multicast flows. When using multicast servers that do not support QoS requests, a sender MUST set the service, not global, break bit(s). Use of the service-specific break bit tells the receiver(s) that RSVP and Integrated Services are supported by the router but that the service cannot be delivered over the ATM network for the specific request.

In the case of MARS [1], the selection of distribution modes is administratively controlled. Therefore network administrators that desire proper RSVP over ATM operation MUST appropriately configure their network to support mesh mode distribution for multicast groups that will be used in RSVP sessions. For LANE1.0 networks the only multicast distribution option is over the LANE Broadcast and Unknown Server which means that the break bit MUST always be set. For LANE2.0 [3] there are provisions that allow for non-server solutions with which it may be possible to ensure proper QoS delivery.

3.4 Receiver Transitions

When setting up a point-to-multipoint VCs there will be a time when some receivers have been added to a QoS VC and some have not.

During such transition times it is possible to start sending data on the newly established VC. If data is sent both on the new VC and the old VC, then data will be delivered with proper QoS to some receivers and with the old QoS to all receivers. Additionally, the QoS receivers would get duplicate data. If data is sent just on the new QoS VC, the receivers that have not yet been added will miss data. So, the issue comes down to whether to send to both the old and new VCs, or to just send to one of the VCs. In one case duplicate data will be received, in the other some data may not be received. This issue needs to be considered for three cases: when establishing the first QoS VC, when establishing a VC to support a QoS change, and when adding a new end-point to an already established QoS VC.

The first two cases are essentially the same. In both, it is possible to send data on the partially completed new VC. In both, there is the option of duplicate or lost data. In order to ensure predictable behavior and to conform to the requirement to deliver data to all receivers, data MUST NOT be sent on new VCs until all parties have been added. This will ensure that all data is only delivered once to all receivers.

The last case differs from the others and occurs when an end-point must be added to an existing QoS VC. In this case the end-point must be both added to the QoS VC and dropped from a best effort VC. The issue is which to do first. If the add is first requested, then the end-point may get duplicate data. If the drop is requested first, then the end-point may miss data. In order to avoid loss of data, the add MUST be completed first and then followed by the drop. This behavior requires receivers to be prepared to receive some duplicate packets at times of QoS setup.

4. Security Considerations

The same considerations stated in [8] and [11] apply to this document. There are no additional security issues raised in this document.

5. Acknowledgments

This work is based on earlier drafts and comments from the ISSLL working group. The author would like to acknowledge their contribution, most notably Steve Berson who coauthored one of the drafts.

6. Author's Address

Lou Berger
FORE Systems
1595 Spring Hill Road
5th Floor
Vienna, VA 22182

Phone: +1 703-245-4527
EMail: lberger@fore.com

REFERENCES

[1] Armitage, G., "Support for Multicast over UNI 3.0/3.1 based ATM Networks," RFC 2022, November 1996.

[2] The ATM Forum, "LAN Emulation Over ATM Specification", Version 1.0.

[3] The ATM Forum, "LAN Emulation over ATM Version 2 - LUNI Specification", April 1997.

[4] The ATM Forum, "MPOA Baseline Version 1", May 1997.

[5] Berger, L., "RSVP over ATM Implementation Guidelines", BCP 24, RFC 2379, August 1998.

[6] Borden, M., and M. Garrett, "Interoperation of Controlled-Load and Guaranteed-Service with ATM", RFC 2381, August 1998.

[7] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[8] Braden, R., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[9] Crawley, E., Berger, L., Berson, S., Baker, F., Borden, M., and J. Krawczyk, "A Framework for Integrated Services and RSVP over ATM", RFC 2382, August 1998.

[10] Heinanen, J., "Multiprotocol Encapsulation over ATM Adaptation Layer 5", RFC 1483, July 1993.

[11] Perez, M., Liaw, F., Grossman, D., Mankin, A., Hoffman, E., and A. Malis, "ATM Signalling Support for IP over ATM", RFC 1755, February 1995.

[12] Maher, M., "ATM Signalling Support for IP over ATM - UNI 4.0 Update", RFC 2331, April 1998.

Full Copyright Statement