Network Working Group
Request for Comments: 4860
Category: Standards Track

F. Le Faucheur
B. Davie
Cisco Systems, Inc.
P. Bose
Lockheed Martin
C. Christou
M. Davenport
Booz Allen Hamilton

May 2007

Generic Aggregate Resource ReSerVation Protocol (RSVP) Reservations

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

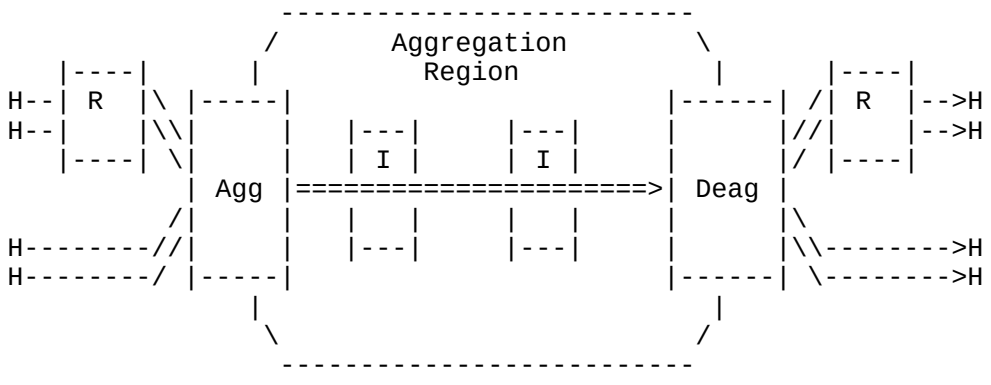
RFC 3175 defines aggregate Resource ReSerVation Protocol (RSVP) reservations allowing resources to be reserved in a Diffserv network for a given Per Hop Behavior (PHB), or given set of PHBs, from a given source to a given destination. RFC 3175 also defines how end-to-end RSVP reservations can be aggregated onto such aggregate reservations when transiting through a Diffserv cloud. There are situations where multiple such aggregate reservations are needed for the same source IP address, destination IP address, and PHB (or set of PHBs). However, this is not supported by the aggregate reservations defined in RFC 3175. In order to support this, the present document defines a more flexible type of aggregate RSVP reservations, referred to as generic aggregate reservation. Multiple such generic aggregate reservations can be established for a given PHB (or set of PHBs) from a given source IP address to a given destination IP address. The generic aggregate reservations may be used to aggregate end-to-end RSVP reservations. This document also defines the procedures for such aggregation. The generic aggregate reservations may also be used end-to-end directly by end-systems attached to a Diffserv network.

Table of Contents

1. Introduction

[RSVP-AGG] defines RSVP aggregate reservations that allow resources to be reserved in a Diffserv network for a flow characterized by its 3-tuple <source IP address, destination IP address, Diffserv Code Point>.

[RSVP-AGG] also defines the procedures for aggregation of end-to-end (E2E) RSVP reservations onto such aggregate reservations when transiting through a Diffserv cloud. Such aggregation is illustrated in Figure 1. This document reuses the terminology defined in [RSVP-AGG].



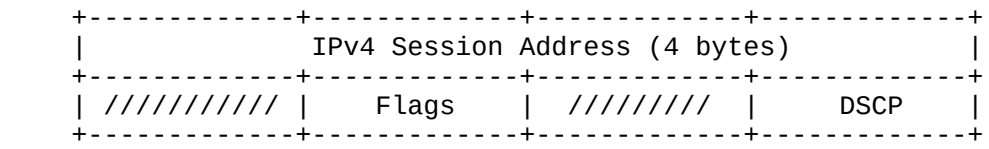
H = Host requesting end-to-end RSVP reservations
R = RSVP router
Agg = Aggregator
Deag = Deaggregator
I = Interior Router

--> = E2E RSVP reservation
==> = Aggregate RSVP reservation

Figure 1 : Aggregation of E2E Reservations over Aggregate RSVP Reservations

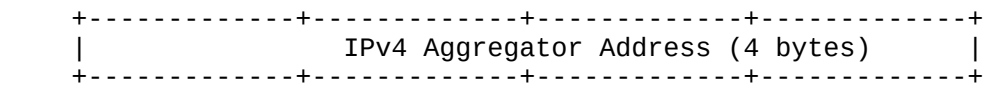
These aggregate reservations use a SESSION type specified in [RSVP-AGG] that contains the receiver (or Deaggregator) IP address and the Diffserv Code Point (DSCP) of the Per Hop Behavior (PHB) from which Diffserv resources are to be reserved. For example, in the case of IPv4, the SESSION object is specified as:

- Class = SESSION,
  C-Type = RSVP-AGGREGATE-IP4



These aggregate reservations use SENDER_TEMPLATE and FILTER_SPEC types, specified in [RSVP-AGG], that contain only the sender (or Aggregator) IP address. For example, in the case of IPv4, the SENDER_TEMPLATE object is specified as:

- Class = SENDER_TEMPLATE,
  C-Type = RSVP-AGGREGATE-IP4



Thus, it is possible to establish, from a given source IP address to a given destination IP address, separate such aggregate reservations for different PHBs (or different sets of PHBs). However, from a given source IP address to a given IP destination address, only a single [RSVP-AGG] aggregate reservation can be established for a given PHB (or given set of PHBs).

Situations have since been identified where multiple such aggregate reservations are needed for the same source IP address, destination IP address, and PHB (or set of PHBs). One example is where E2E reservations using different preemption priorities (as per [RSVP-PREEMP]) need to be aggregated through a Diffserv cloud using the same PHB. Using multiple aggregate reservations for the same PHB allows enforcement of the different preemption priorities within the aggregation region. In turn, this allows more efficient management of the Diffserv resources, and in periods of resource shortage, this allows sustainment of a larger number of E2E reservations with higher preemption priorities.

For example, [SIG-NESTED] discusses in detail how end-to-end RSVP reservations can be established in a nested VPN environment through RSVP aggregation. In particular, [SIG-NESTED] describes how multiple parallel generic aggregate reservations (for the same PHB), each with different preemption priorities, can be used to efficiently support the preemption priorities of end-to-end reservations.

This document addresses this requirement for multiple aggregate reservations for the same PHB (or same set of PHBs), by defining a more flexible type of aggregate RSVP reservations, referred to as generic aggregate reservations. This is achieved primarily by adding the notions of a Virtual Destination Port and of an Extended Virtual Destination Port in the RSVP SESSION object.

The notion of Virtual Destination Port was introduced in [RSVP-IPSEC] to address a similar requirement (albeit in a different context) for identification and demultiplexing of sessions beyond the IP destination address. This document reuses this notion from [RSVP-IPSEC] for identification and demultiplexing of generic aggregate sessions beyond the IP destination address and PHB. This allows multiple generic aggregate reservations to be established for a given PHB (or set of PHBs), from a given source IP address to a given destination IP address.

[RSVP-TE] introduced the concept of an Extended Tunnel ID (in addition to the tunnel egress address and the Tunnel ID) in the SESSION object used to establish MPLS Traffic Engineering tunnels with RSVP. The Extended Tunnel ID provides a very convenient mechanism for the tunnel ingress node to narrow the scope of the session to the ingress-egress pair. The ingress node can achieve this by using one of its own IP addresses as a globally unique identifier and including it in the Extended Tunnel ID and therefore within the SESSION object. This document reuses this notion of Extended Tunnel ID from [RSVP-TE], simply renaming it Extended Virtual Destination Port. This provides a convenient mechanism to narrow the scope of a generic aggregate session to an Aggregator-Deaggregator pair.

The RSVP SESSION object for generic aggregate reservations uses the PHB Identification Code (PHB-ID) defined in [PHB-ID] to identify the PHB, or set of PHBs, from which the Diffserv resources are to be reserved. This is instead of using the Diffserv Code Point (DSCP) as per [RSVP-AGG]. Using the PHB-ID instead of the DSCP allows explicit indication of whether the Diffserv resources belong to a single PHB or to a set of PHBs. It also facilitates handling of situations where a generic aggregate reservation spans two (or more) Diffserv domains that use different DSCP values for the same Diffserv PHB (or set of PHBs) from which resources are reserved. This is because the PHB-ID allows conveying of the PHB (or set of PHBs) independently of what DSCP value(s) are used locally for that PHB (or set of PHBs).

The generic aggregate reservations may be used to aggregate end-to-end RSVP reservations. This document also defines the procedures for such aggregation. These procedures are based on those of [RSVP-AGG], and this document only specifies the differences from those.

The generic aggregate reservations may also be used end-to-end directly by end-systems attached to a Diffserv network.

1.1. Related IETF Documents

This document is heavily based on [RSVP-AGG]. It reuses [RSVP-AGG] wherever applicable and only specifies the necessary extensions beyond [RSVP-AGG].

The mechanisms defined in [BW-REDUC] allow an existing reservation to be reduced in allocated bandwidth by RSVP routers in lieu of tearing that reservation down. These mechanisms are applicable to the generic aggregate reservations defined in the present document.

[RSVP-TUNNEL] describes a general approach to running RSVP over various types of tunnels. One of these types of tunnel, referred to as a "type 2 tunnel", has some similarity with the generic aggregate reservations described in this document. The similarity stems from the fact that a single, aggregate reservation is made for the tunnel while many individual flows are carried over that tunnel. However, [RSVP-TUNNEL] does not address the use of Diffserv-based classification and scheduling in the core of a network (between tunnel endpoints), but rather relies on a UDP/IP tunnel header for classification. This is why [RSVP-AGG] required additional objects and procedures beyond those of [RSVP-TUNNEL]. Like [RSVP-AGG], this document also assumes the use of Diffserv-based classification and scheduling in the aggregation region and, thus, requires additional objects and procedures beyond those of [RSVP-TUNNEL].

As explained earlier, this document reuses the notion of Virtual Destination Port from [RSVP-IPSEC] and the notion of Extended Tunnel ID from [RSVP-TE].

1.2. Organization Of This Document

Section 2 defines the new RSVP objects related to generic aggregate reservations and to aggregation of E2E reservations onto those. Section 3 describes the processing rules for handling of generic aggregate reservations. Section 4 specifies the procedures for aggregation of end-to-end RSVP reservations over generic aggregate RSVP reservations. Section 5 provides example usage of how the generic aggregate reservations may be used.

The Security Considerations and the IANA Considerations are discussed in Sections 6 and 7, respectively.

Finally, Appendix A provides an example signaling flow that illustrates aggregation of E2E RSVP reservations onto generic aggregate RSVP reservations.

1.3. Requirements Language

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 [KEYWORDS].

2. Object Definition

This document reuses the RSVP-AGGREGATE-IP4 FILTER_SPEC, RSVP-AGGREGATE-IP6 FILTER_SPEC, RSVP-AGGREGATE-IP4 SENDER_TEMPLATE, and RSVP-AGGREGATE-IP6 SENDER_TEMPLATE objects defined in [RSVP-AGG].

This document defines:

- two new objects (GENERIC-AGGREGATE-IP4 SESSION and GENERIC-AGGREGATE-IP6 SESSION) under the existing SESSION Class, and

- two new objects (GENERIC-AGG-IP4-SOI and GENERIC-AGG-IP6-SOI) under a new SESSION-OF-INTEREST Class.

Detailed description of these objects is provided below in this section.

The GENERIC-AGGREGATE-IP4 SESSION and GENERIC-AGGREGATE-IP6 SESSION objects are applicable to all types of RSVP messages.

This specification defines the use of the GENERIC-AGG-IP4-SOI and GENERIC-AGG-IP6-SOI objects in two circumstances:

- inside an E2E PathErr message that contains an error code of NEW-AGGREGATE-NEEDED in order to convey the session of a new generic aggregate reservation that needs to be established.

- inside an E2E Resv message in order to convey the session of the generic aggregate reservation onto which this E2E reservation needs to be mapped.

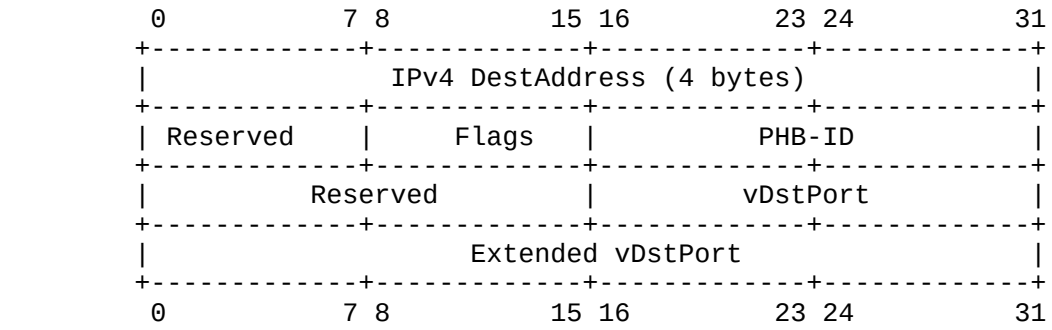
Details of the corresponding procedures can be found in Section 4.

However, it is envisioned that the ability to signal, inside RSVP messages, the Session of another reservation (which has some relationship with the current RSVP reservation) might have some other applicability in the future. Thus, those objects have been specified in a more generic manner under a flexible SESSION-OF-INTEREST class.

All the new objects defined in this document are optional with respect to RSVP so that general RSVP implementations that are not concerned with generic aggregate reservations do not have to support these objects. RSVP routers supporting generic aggregate IPv4 or IPv6 reservations MUST support the GENERIC-AGGREGATE-IP4 SESSION object or the GENERIC-AGGREGATE-IP6 SESSION object, respectively. RSVP routers supporting RSVP aggregation over generic aggregate IPv4 or IPv6 reservations MUST support the GENERIC-AGG-IP4-SOI object or GENERIC-AGG-IP6-SOI object, respectively.

2.1. SESSION Class

o GENERIC-AGGREGATE-IP4 SESSION object:
Class = 1 (SESSION)
C-Type = 17



IPv4 DestAddress (IPv4 Destination Address)

IPv4 address of the receiver (or Deaggregator).

Reserved

An 8-bit field. All bits MUST be set to 0 on transmit. This field MUST be ignored on receipt.

Flags

An 8-bit field. The content and processing of this field are the same as for the Flags field of the IPv4/UDP SESSION object (see [RSVP]).

PHB-ID (Per Hop Behavior Identification Code)

A 16-bit field containing the Per Hop Behavior Identification Code of the PHB, or of the set of PHBs, from which Diffserv resources are to be reserved. This field MUST be encoded as specified in Section 2 of [PHB-ID].

Reserved

A 16-bit field. All bits MUST be set to 0 on transmit. This field MUST be ignored on receipt.

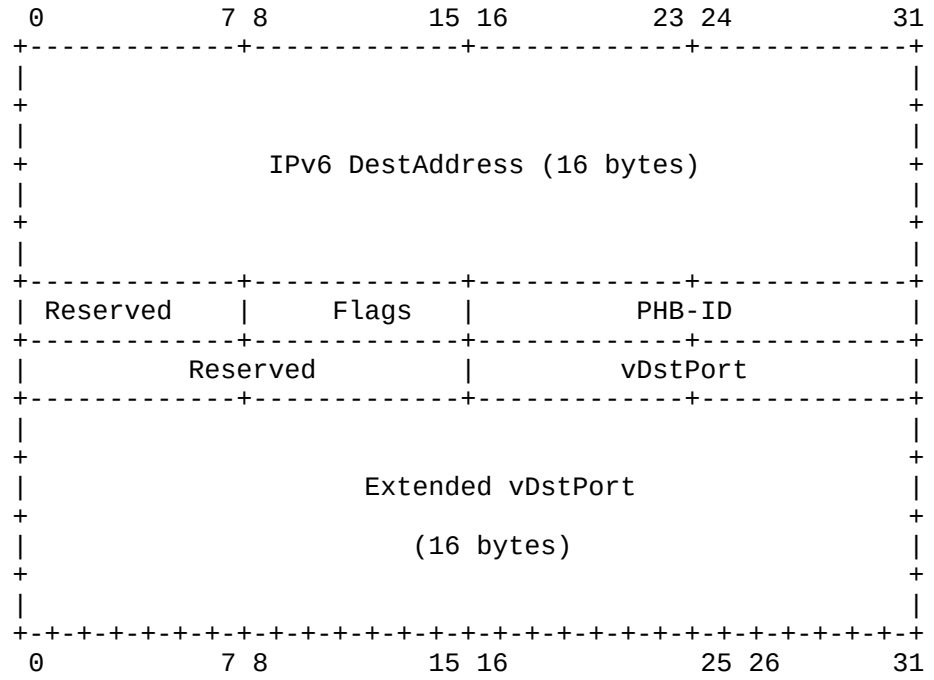
VDstPort (Virtual Destination Port)

A 16-bit identifier used in the SESSION that remains constant over the life of the generic aggregate reservation.

Extended vDstPort (Extended Virtual Destination Port)

A 32-bit identifier used in the SESSION that remains constant over the life of the generic aggregate reservation. A sender (or Aggregator) that wishes to narrow the scope of a SESSION to the sender-receiver pair (or Aggregator-Deaggregator pair) SHOULD place its IPv4 address here as a network unique identifier. A sender (or Aggregator) that wishes to use a common session with other senders (or Aggregators) in order to use a shared reservation across senders (or Aggregators) MUST set this field to all zeros.

o GENERIC-AGGREGATE-IP6 SESSION object:
Class = 1 (SESSION)
C-Type = 18



IPv6 DestAddress (IPv6 Destination Address)

IPv6 address of the receiver (or Deaggregator).

Reserved

An 8-bit field. All bits MUST be set to 0 on transmit. This field MUST be ignored on receipt.

Flags

An 8-bit field. The content and processing of this field are the same as for the Flags field of the IPv6/UDP SESSION object (see [RSVP]).

PHB-ID (Per Hop Behavior Identification Code)

A 16-bit field containing the Per Hop Behavior Identification Code of the PHB, or of the set of PHBs, from which Diffserv resources are to be reserved. This field MUST be encoded as specified in Section 2 of [PHB-ID].

Reserved

A 16-bit field. All bits MUST be set to 0 on transmit. This field MUST be ignored on receipt.

VDstPort (Virtual Destination Port)

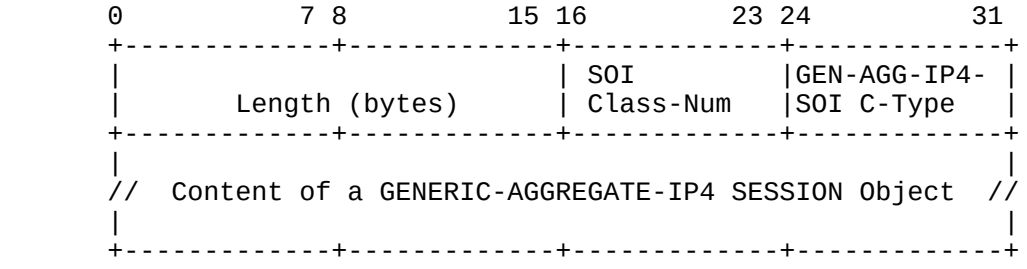
A 16-bit identifier used in the SESSION that remains constant over the life of the generic aggregate reservation.

Extended vDstPort (Extended Virtual Destination Port)

A 128-bit identifier used in the SESSION that remains constant over the life of the generic aggregate reservation. A sender (or Aggregator) that wishes to narrow the scope of a SESSION to the sender-receiver pair (or Aggregator-Deaggregator pair) SHOULD place its IPv6 address here as a network unique identifier. A sender (or Aggregator) that wishes to use a common session with other senders (or Aggregators) in order to use a shared reservation across senders (or Aggregators) MUST set this field to all zeros.

2.2. SESSION-OF-INTEREST (SOI) Class

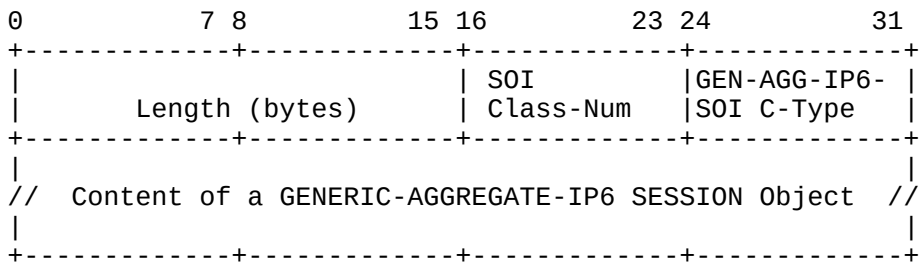
o GENERIC-AGG-IP4-SOI object:
Class = 132
C-Type = 1



Content of a GENERIC-AGGREGATE-IP4 SESSION Object:

This field contains a copy of the SESSION object of the session that is of interest for the reservation. In the case of a GENERIC-AGG-IP4-SOI, the session of interest conveyed in this field is a GENERIC-AGGREGATE-IP4 SESSION.

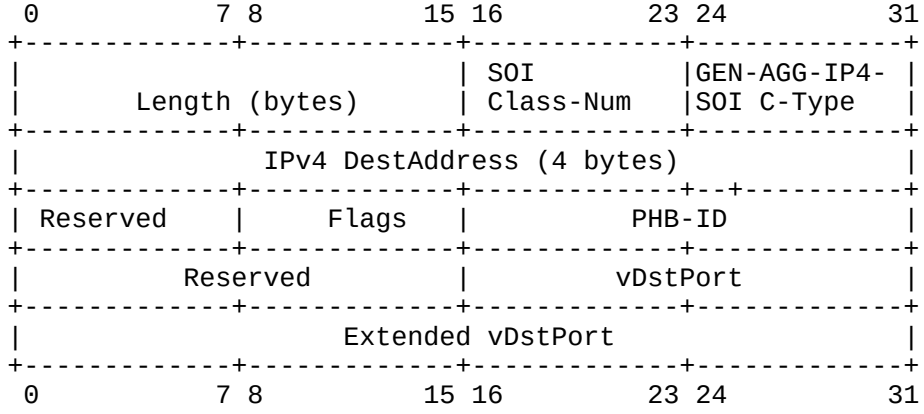
o GENERIC-AGG-IP6-SOI object:
Class = 132
C-Type = 2



Content of a GENERIC-AGGREGATE-IP6 SESSION Object:

This field contains a copy of the SESSION object of the session that is of interest for the reservation. In the case of a GENERIC-AGG-IP6-SOI, the session of interest conveyed in this field is a GENERIC-AGGREGATE-IP6 SESSION.

For example, if a SESSION-OF-INTEREST object is used inside an E2E Resv message (as per the procedures defined in Section 4) to indicate which generic aggregate IPv4 session the E2E reservation is to be mapped onto, then the GENERIC-AGG-IP4-SOI object will be used, and it will be encoded like this:



Note that a SESSION-OF-INTEREST object is not a SESSION object in itself. It does not replace the SESSION object in RSVP messages. It does not modify the usage of the SESSION object in RSVP messages. It simply allows conveying the Session of another RSVP reservation inside RSVP signaling messages, for some particular purposes. In the context of this document, it is used to convey, inside an E2E RSVP

message pertaining to an end-to-end reservation, the Session of a generic aggregate reservation associated with the E2E reservation. Details for the corresponding procedures are specified in Section 4.

3. Processing Rules for Handling Generic Aggregate RSVP Reservations

This section presents extensions to the processing of RSVP messages required by [RSVP] and presented in [RSVP-PROCESS]. These extensions are required in order to properly process the GENERIC-AGGREGATE-IP4 or GENERIC-AGGREGATE-IP6 SESSION object and the RSVP-AGGREGATE-IP4 or RSVP-AGGREGATE-IP6 FILTER_SPEC object. Values for referenced error codes can be found in [RSVP]. As with the other RSVP documents, values for internally reported (API) errors are not defined.

When referring to the new GENERIC-AGGREGATE-IP4 and GENERIC-AGGREGATE-IP6 SESSION objects, IP version will not be included, and they will be referred to simply as GENERIC-AGGREGATE SESSION, unless a specific distinction between IPv4 and IPv6 is being made.

When referring to the [RSVP-AGG] RSVP-AGGREGATE-IP4 and RSVP-AGGREGATE-IP6 SESSION, FILTER_SPEC, and SENDER_TEMPLATE objects, IP version will not be included, and they will be referred to simply as RSVP-AGGREGATE, unless a specific distinction between IPv4 and IPv6 is being made.

3.1. Extensions to Path and Resv Processing

The following PATH message processing changes are defined:

- When a session is defined using the GENERIC-AGGREGATE SESSION object, only the [RSVP-AGG] RSVP-AGGREGATE SENDER_TEMPLATE may be used. When this condition is violated in a PATH message received by an RSVP end-station, the RSVP end-station SHOULD report a "Conflicting C-Type" API error to the application. When this condition is violated in a PATH message received by an RSVP router, the RSVP router MUST consider this as a message formatting error.

- For PATH messages that contain the GENERIC-AGGREGATE SESSION object, the VDstPort value, the Extended VDstPort value, and the PHB-ID value should be recorded (in addition to the destination/Deaggregator address and source/Aggregator address). These values form part of the recorded state of the session. The PHB-ID may need to be passed to traffic control; however the vDstPort and Extended VDstPort are not passed to traffic control since they do not appear inside the data packets of the corresponding reservation.

The following changes to RESV message processing are defined:

- When a RESV message contains a [RSVP-AGG] RSVP-AGGREGATE FILTER_SPEC, the session MUST be defined using either the RSVP-AGGREGATE SESSION object (as per [RSVP-AGG]) or the GENERIC-AGGREGATE SESSION object (as per this document). If this condition is not met, an RSVP router or end-station MUST consider that there is a message formatting error.

- When the RSVP-AGGREGATE FILTER_SPEC is used and the SESSION type is GENERIC-AGGREGATE, each node uses data classifiers as per the following:

  * to perform Diffserv classification the node MUST rely on the Diffserv data classifier based on the DSCP only. The relevant DSCP value(s) are those that are associated with the PHB-ID of the generic aggregate reservation.

  * If the node also needs to perform fine-grain classification (for example, to perform fine-grain input policing at a trust boundary) then the node MUST create a data classifier described by the 3-tuple <DestAddress, SrcAddress, DSCP>.

    The relevant DSCP value(s) are those that are associated with the PHB-ID of the generic aggregate reservation.

    Note that if multiple generic aggregate reservations are established with different Virtual Destination Ports (and/or different Extended Virtual Destination Ports) but with the same <DestAddress, SrcAddress, PHB-ID>, then those cannot be distinguished by the classifier. If the router is using the classifier for policing purposes, the router will therefore police those together and MUST program the policing rate to the sum of the reserved rate across all the corresponding reservations.

4. Procedures for Aggregation over Generic Aggregate RSVP Reservations

The procedures for aggregation of E2E reservations over generic aggregate RSVP reservations are the same as the procedures specified in [RSVP-AGG] with the exceptions of the procedure changes listed in this section.

As specified in [RSVP-AGG], the Deaggregator is responsible for mapping a given E2E reservation on a given aggregate reservation. The Deaggregator requests establishment of a new aggregate reservation by sending to the Aggregator an E2E PathErr message with an error code of NEW-AGGREGATE-NEEDED. In [RSVP-AGG], the

Deaggregator conveys the DSCP of the new requested aggregate reservation by including a DCLASS Object in the E2E PathErr and encoding the corresponding DSCP inside. This document modifies and extends this procedure. The Deaggregator MUST include in the E2E PathErr message a SESSION-OF-INTEREST object that contains the GENERIC-AGGREGATE SESSION to be used for establishment of the requested generic aggregate reservation. Since this GENERIC-AGGREGATE SESSION contains the PHB-ID, the DCLASS object need not be included in the PathErr message.

Note that the Deaggregator can easily ensure that different Aggregators use different sessions for their Aggregate Path towards a given Deaggregator. This is because the Deaggregator can easily select VDstPort and/or Extended VDstPort numbers which are different for each Aggregator (for example, by using the Aggregator address as the Extended VDstPort) and can communicate those inside the GENERIC-AGGREGATE SESSION included in the SESSION-OF-INTEREST object. This provides an easy solution to establish separate reservations from every Aggregator to a given Deaggregator. Conversely, if reservation sharing were needed across multiple Aggregators, the Deaggregator could facilitate this by allocating the same VDstPort and Extended VDstPort to the multiple Aggregators, and thus including the same GENERIC-AGGREGATE SESSION inside the SESSION-OF-INTEREST object in the E2E PathErr messages sent to these Aggregators. The Aggregators could then all establish an Aggregate Path with the same GENERIC-AGGREGATE SESSION.

Therefore, various sharing scenarios can easily be supported. Policies followed by the Deaggregator to determine which Aggregators need shared or separate reservations are beyond the scope of this document.

The Deaggregator MAY also include in the E2E PathErr message (with an error code of NEW-AGGREGATE-NEEDED) additional RSVP objects which are to be used for establishment of the newly needed generic aggregate reservation. For example, the Deaggregator MAY include in the E2E PathErr an RSVP Signaled Preemption Priority Policy Element (as specified in [RSVP-PREEMP]).

The [RSVP-AGG] procedures for processing of an E2E PathErr message received with an error code of NEW-AGGREGATE-NEEDED by the Aggregator are extended correspondingly. On receipt of such a message containing a SESSION-OF-INTEREST object, the Aggregator MUST trigger establishment of a generic aggregate reservation. In particular, it MUST start sending aggregate Path messages with the GENERIC-AGGREGATE SESSION found in the received SESSION-OF-INTEREST object. When an RSVP Signaled Preemption Priority Policy Element is contained in the received E2E PathErr message, the Aggregator MUST include this object

in the Aggregate Path for the corresponding generic aggregate reservation. When other additional objects are contained in the received E2E PathErr message and those can be unambiguously interpreted as related to the new needed generic aggregate reservation (as opposed to related to the E2E reservation), the Aggregator SHOULD include those in the Aggregate Path for the corresponding generic aggregate reservation. The Aggregator MUST use as the Source Address (i.e., as the Aggregator Address in the Sender-Template) for the generic aggregate reservation, the address it uses to identify itself as the PHOP (RSVP previous hop) when forwarding the E2E Path messages corresponding to the E2E PathErr message.

The Deaggregator follows the same procedures as described in [RSVP-AGG] for establishing, maintaining and clearing the aggregate Resv state. However, a Deaggregator behaving according to the present specification MUST use the generic aggregate reservations and hence use the GENERIC-AGGREGATE SESSION specified earlier in this document.

This document also modifies the procedures of [RSVP-AGG] related to exchange of E2E Resv messages between Deaggregator and Aggregator. The Deaggregator MUST include the new SESSION-OF-INTEREST object in the E2E Resv message, in order to indicate to the Aggregator the generic aggregate session to map a given E2E reservation onto. Again, since the GENERIC-AGGREGATE SESSION (included in the SESSION-OF-INTEREST object) contains the PHB-ID, the DCLASS object need not be included in the E2E Resv message. The Aggregator MUST interpret the SESSION-OF-INTEREST object in the E2E Resv as indicating which generic aggregate reservation session the corresponding E2E reservation is mapped onto. The Aggregator MUST not include the SESSION-OF-INTEREST object when sending an E2E Resv upstream towards the sender.

Based on relevant policy, the Deaggregator may decide at some point that an aggregate reservation is no longer needed and should be torn down. In that case, the Deaggregator MUST send an aggregate ResvTear. On receipt of the aggregate ResvTear, the Aggregator SHOULD send an aggregate PathTear (unless the relevant policy instructs the Aggregator to do otherwise or to wait for some time before doing so, for example in order to speed up potential re-establishment of the aggregate reservation in the future).

[RSVP-AGG] describes how the Aggregator and Deaggregator can communicate their respective identities to each other. For example, the Aggregator includes one of its IP addresses in the RSVP HOP object in the E2E Path that is transmitted downstream and received by the Deaggregator once it traversed the aggregation region. Similarly, the Deaggregator identifies itself to the Aggregator by

including one of its IP addresses in various fields, including the ERROR SPECIFICATION of the E2E PathErr message (containing the NEW-AGGREGATE-NEEDED Error Code) and in the RSVP HOP object of the E2E Resv message. However, [RSVP-AGG] does not discuss which IP addresses are to be selected by the Aggregator and Deaggregator for such purposes. Because these addresses are intended to identify the Aggregator and Deaggregator and not to identify any specific interface on these devices, this document RECOMMENDS that the Aggregator and Deaggregator SHOULD use interface-independent addresses (for example, a loopback address) whenever they communicate their respective identities to each other. This ensures that respective identification of the Aggregator and Deaggregator is not impacted by any interface state change on these devices. In turn, this results in more stable operations and considerably reduced RSVP signaling in the aggregation region. For example, if interface-independent addresses are used by the Aggregator and the Deaggregator, then a failure of an interface on these devices may simply result in the rerouting of a given generic aggregate reservation, but will not result in the generic aggregate reservation having to be torn down and another one established. Moreover, it will not result in a change of mapping of E2E reservations on generic aggregate reservations (assuming the Aggregator and Deaggregator still have reachability after the failure, and the Aggregator and Deaggregator are still on the shortest path to the destination).

However, when identifying themselves to real RSVP neighbors (i.e., neighbors that are not on the other side of the aggregation region), the Aggregator and Deaggregator SHOULD continue using interface-dependent addresses as per regular [RSVP] procedures. This applies for example when the Aggregator identifies itself downstream as a PHOP for the generic aggregate reservation or identifies itself upstream as a NHOP (RSVP next hop) for an E2E reservation. This also applies when the Deaggregator identifies itself downstream as a PHOP for the E2E reservation or identifies itself upstream as a NHOP for the generic aggregate reservation. As part of the processing of generic aggregate reservations, interior routers (i.e., routers within the aggregation region) SHOULD continue using interface-dependent addresses as per regular [RSVP] procedures.

More generally, within the aggregation region (i.e., between Aggregator and Deaggregator) the operation of RSVP should be modeled with the notion that E2E reservations are mapped to aggregate reservations and are no longer tied to physical interfaces (as was the case with regular RSVP). However, generic aggregate reservations (within the aggregation region) as well as E2E reservations (outside the aggregation region) retain the model of regular RVSP and remain tied to physical interfaces.

As discussed above, generic aggregate reservations may be established edge-to-edge as a result of the establishment of E2E reservations (from outside the aggregation region) that are to be aggregated over the aggregation region. However, generic aggregate reservations may also be used end-to-end by end-systems directly attached to a Diffserv domain, such as Public Switched Telephone Network (PSTN) gateways. In that case, the generic aggregate reservations may be established by the end-systems in response to application-level triggers such as voice call signaling. Alternatively, generic aggregate reservations may also be used edge-to-edge to manage bandwidth in a Diffserv cloud even if RSVP is not used end-to-end. A simple example of such a usage would be the static configuration of a generic aggregate reservation for a certain bandwidth for traffic from an ingress (Aggregator) router to an egress (Deaggregator) router.

In this case, the establishment of the generic aggregate reservations is controlled by configuration on the Aggregator and on the Deaggregator. Configuration on the Aggregator triggers generation of the aggregate Path message and provides sufficient information to the Aggregator to derive the content of the GENERIC-AGGREGATE SESSION object. This would typically include Deaggregator IP address, PHB-ID and possibly VDstPort. Configuration on the Deaggregator would instruct the Deaggregator to respond to a received generic aggregate Path message and would provide sufficient information to the Deaggregator to control the reservation. This may include bandwidth to be reserved by the Deaggregator (for a given <Deaggregator, PHB-ID, VDstPort> tuple).

In the absence of E2E microflow reservations, the Aggregator can use a variety of policies to set the DSCP of packets passing into the aggregation region and how they are mapped onto generic aggregate reservations, thus determining whether they gain access to the resources reserved by the aggregate reservation. These policies are a matter of local configuration, as is typical for a device at the edge of a Diffserv cloud.

5. Example Usage Of Multiple Generic Aggregate Reservations per PHB from a Given Aggregator to a Given Deaggregator

Let us consider the environment depicted in Figure 2 below. RSVP aggregation is used to support E2E reservations between Cloud-1, Cloud-2, and Cloud-3.

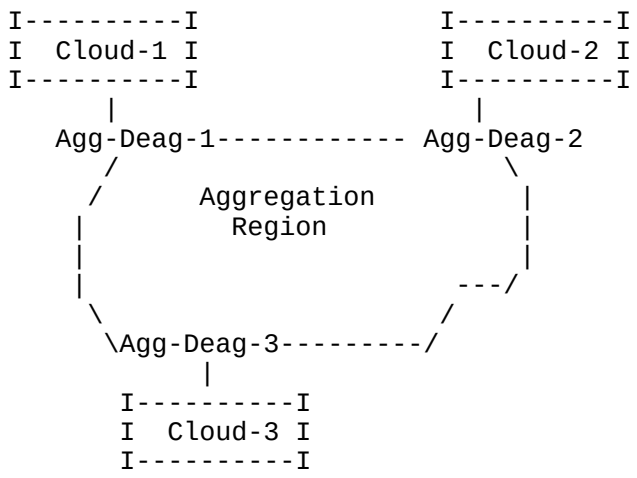


Figure 2 : Example Usage of Generic Aggregate IP Reservations

Let us assume that:

- The E2E reservations from Cloud-1 to Cloud-3 have a preemption of either P1 or P2.
- The E2E reservations from Cloud-2 to Cloud-3 have a preemption of either P1 or P2.
- The E2E reservations are only for Voice (which needs to be treated in the aggregation region using the EF -Expedited Forwarding- PHB).
- Traffic from the E2E reservations is encapsulated in aggregate IP reservations from Aggregator to Deaggregator using Generic Routing Encapsulation [GRE] tunneling.

Then, the following generic aggregate RSVP reservations may be established from Agg-Deag-1 to Agg-Deag-3 for aggregation of the end-to-end RSVP reservations:

(1) A first generic aggregate reservation for aggregation of Voice reservations from Cloud-1 to Cloud-3 requiring use of P1:

* GENERIC-AGGREGATE-IP4 SESSION:
  IPv4 DestAddress = Agg-Deag-3
  vDstPort = V1
  PHB-ID = EF
  Extended VDstPort = Agg-Deag-1

* STYLE = FF or SE

* IPv4/GPI FILTER_SPEC:
  IPv4 SrcAddress = Agg-Deag-1

* POLICY_DATA (PREEMPTION_PRI) = P1

(2) A second generic aggregate reservation for aggregation of Voice reservations from Cloud-1 to Cloud-3 requiring use of P2:

* GENERIC-AGGREGATE-IP4 SESSION:
  IPv4 DestAddress = Agg-Deag-3
  vDstPort = V2
  PHB-ID = EF
  Extended VDstPort = Agg-Deag-1

* STYLE = FF or SE

* IPv4/GPI FILTER_SPEC:
  IPv4 SrcAddress = Agg-Deag-1

* POLICY_DATA (PREEMPTION_PRI) = P2

where V1 and V2 are arbitrary VDstPort values picked by Agg-Deag-3.

The following generic aggregate RSVP reservations may be established from Agg-Deag-2 to Agg-Deag-3 for aggregation of the end-to-end RSVP reservations:

(3) A third generic aggregate reservation for aggregation of Voice reservations from Cloud-2 to Cloud-3 requiring use of P1:

* GENERIC-AGGREGATE-IP4 SESSION:
  IPv4 DestAddress = Agg-Deag-3
  vDstPort = V3
  PHB-ID = EF
  Extended VDstPort = Agg-Deag-2

* STYLE = FF or SE

* IPv4/GPI FILTER_SPEC:
  IPv4 SrcAddress = Agg-Deag-2

* POLICY_DATA (PREEMPTION_PRI) = P1

(4) A fourth generic aggregate reservation for aggregation of Voice reservations from Cloud-2 to Cloud-3 requiring use of P2:

* GENERIC-AGGREGATE-IP4 SESSION:
  IPv4 DestAddress = Agg-Deag-3
  vDstPort = V4
  PHB-ID = EF
  Extended VDstPort = Agg-Deag-2

* STYLE = FF or SE

* IPv4/GPI FILTER_SPEC:
  IPv4 SrcAddress = Agg-Deag-2

* POLICY_DATA (PREEMPTION_PRI) = P2

where V3 and V4 are arbitrary VDstPort values picked by Agg-Deag-3.

Note that V3 and V4 could be equal to V1 and V2 (respectively) since, in this example, the Extended VDstPort of the GENERIC-AGGREGATE Session contains the address of the Aggregator and, thus, ensures that different sessions are used from each Aggregator.

6. Security Considerations

In the environments addressed by this document, RSVP messages are used to control resource reservations for generic aggregate reservations and may be used to control resource reservations for E2E reservations being aggregated over the generic aggregate reservations. To ensure the integrity of the associated reservation and admission control mechanisms, the RSVP Authentication mechanisms defined in [RSVP-CRYPTO1] and [RSVP-CRYPTO2] may be used. These protect RSVP message integrity hop-by-hop and provide node authentication as well as replay protection, thereby protecting against corruption and spoofing of RSVP messages. These hop-by-hop integrity mechanisms can be naturally used to protect the RSVP messages used for generic aggregate reservations and to protect RSVP messages used for E2E reservations outside the aggregation region. These hop-by-hop RSVP integrity mechanisms can also be used to protect RSVP messages used for E2E reservations when those transit through the aggregation region. This is because the Aggregator and

Deaggregator behave as RSVP neighbors from the viewpoint of the E2E flows (even if they are not necessarily IP neighbors).

[RSVP-CRYPTO1] discusses several approaches for key distribution. First, the RSVP Authentication shared keys can be distributed manually. This is the base option and its support is mandated for any implementation. However, in some environments, this approach may become a burden if keys frequently change over time. Alternatively, a standard key management protocol for secure key distribution can be used. However, existing key distribution protocols may not be appropriate in all environments because of the complexity or operational burden they involve.

The use of RSVP Authentication in parts of the network where there may be one or more IP hops in between two RSVP neighbors raises an additional challenge. This is because, with some RSVP messages such as a Path message, an RSVP router does not know the RSVP next hop for that message at the time of forwarding it. In fact, part of the role of a Path message is precisely to discover the RSVP next hop (and to dynamically re-discover it when it changes, say because of a routing change). Hence, the RSVP router may not know which security association to use when forwarding such a message. This applies in particular to the case where RSVP Authentication mechanisms are to be used for protection of RSVP E2E messages (e.g., E2E Path) while they transit through an aggregation region and where the dynamic Deaggregator determination procedure defined in [RSVP-AGG] is used. This is because the Aggregator and the Deaggregator behave as RSVP neighbors for the E2E reservation, while there may be one or more IP hops in between them, and the Aggregator does not know ahead of time which router is going to act as the Deaggregator.

In that situation, one approach is to share the same RSVP Authentication shared key across all the RSVP routers of a part of the network where there may be RSVP neighbors with IP hops in between. For example, all the Aggregators or Deaggregators of an aggregation region could share the same RSVP Authentication key, while different per-neighbor keys could be used between any RSVP router pair straddling the boundary between two administrative domains that have agreed to use RSVP signaling.

When the same RSVP Authentication shared key is to be shared among multiple RSVP neighbors, manual key distribution may be used. For situations where RSVP is being used for multicast flows, it might also be possible, in the future, to adapt a multicast key management method (e.g. from IETF Multicast Security Working Group) for key distribution with such multicast RSVP usage. For situations where RSVP is being used for unicast flows across domain boundaries, it is not currently clear how one might provide automated key management.

Specification of a specific automated key management technique is outside the scope of this document. Operators should consider these key management issues when contemplating deployment of this specification.

The RSVP Authentication mechanisms do not provide confidentiality. If confidentiality is required, IPsec ESP [IPSEC-ESP] may be used, although it imposes the burden of key distribution. It also faces the additional issue discussed for key management above in the case where there can be IP hops in between RSVP hops. In the future, confidentiality solutions may be developed for the case where there can be IP hops in between RSVP hops, perhaps by adapting confidentiality solutions developed by the IETF MSEC Working Group. Such confidentiality solutions for RSVP are outside the scope of this document.

Protection against traffic analysis is also not provided by RSVP Authentication. Since generic aggregate reservations are intended to reserve resources collectively for a whole set of users or hosts, malicious snooping of the corresponding RSVP messages could provide more traffic analysis information than snooping of an E2E reservation. When RSVP neighbors are directly attached, mechanisms such as bulk link encryption might be used when protection against traffic analysis is required. This approach could be used inside the aggregation region for protection of the generic aggregate reservations. It may also be used outside the aggregation region for protection of the E2E reservation. However, it is not applicable to the protection of E2E reservations while the corresponding E2E RSVP messages transit through the aggregation region.

When generic aggregate reservations are used for aggregation of E2E reservations, the security considerations discussed in [RSVP-AGG] apply and are revisited here.

First, the loss of an aggregate reservation to an aggressor causes E2E flows to operate unreserved, and the reservation of a great excess of bandwidth may result in a denial of service. These issues are not confined to the extensions defined in the present document: RSVP itself has them. However, they may be exacerbated here by the fact that each aggregate reservation typically facilitates communication for many sessions. Hence, compromising one such aggregate reservation can result in more damage than compromising a typical E2E reservation. Use of the RSVP Authentication mechanisms to protect against such attacks has been discussed above.

An additional security consideration specific to RSVP aggregation involves the modification of the IP protocol number in RSVP Path messages that traverse an aggregation region. Malicious modification

of the IP protocol number in a Path message would cause the message to be ignored by all subsequent RSVP devices on its path, preventing reservations from being made. It could even be possible to correct the value before it reached the receiver, making it difficult to detect the attack. Note that, in theory, it might also be possible for a node to modify the IP protocol number for non-RSVP messages as well, thus interfering with the operation of other protocols. It is RECOMMENDED that implementations of this specification only support modification of the IP protocol number for RSVP Path, PathTear, and ResvConf messages. That is, a general facility for modification of the IP protocol number SHOULD NOT be made available.

Network operators deploying routers with RSVP aggregation capability should be aware of the risks of inappropriate modification of the IP protocol number and should take appropriate steps (physical security, password protection, etc.) to reduce the risk that a router could be configured by an attacker to perform malicious modification of the protocol number.

7. IANA Considerations

IANA modified the RSVP parameters registry, 'Class Names, Class Numbers, and Class Types' subregistry, and assigned two new C-Types under the existing SESSION Class (Class number 1), as described below:

Class Number	Class Name	Reference
1	SESSION	[RFC2205]
	Class Types or C-Types:	
	17 GENERIC-AGGREGATE-IP4	[RFC4860]
	18 GENERIC-AGGREGATE-IP6	[RFC4860]

IANA also modified the RSVP parameters registry, 'Class Names, Class Numbers, and Class Types' subregistry, and assigned one new Class Number for the SESSION-OF-INTEREST class and two new C-Types for that class, according to the table below:

Class Number	Class Name	Reference
132	SESSION-OF-INTEREST	[RFC4860]
	Class Types or C-Types:	
	1 GENERIC-AGG-IP4-SOI	[RFC4860]
	2 GENERIC-AGG-IP6-SOI	[RFC4860]

These allocations are in accordance with [RSVP-MOD].

8. Acknowledgments

This document borrows heavily from [RSVP-AGG]. It also borrows the concepts of Virtual Destination Port and Extended Virtual Destination Port from [RSVP-IPSEC] and [RSVP-TE], respectively.

Also, we thank Fred Baker, Roger Levesque, Carol Iturralde, Daniel Voce, Anil Agarwal, Alexander Sayenko, and Anca Zamfir for their input into the content of this document. Thanks to Steve Kent for insightful comments on usage of RSVP reservations in IPsec environments.

Ran Atkinson, Fred Baker, Luc Billot, Pascal Delprat, and Eric Vyncke provided guidance and suggestions for the security considerations section.

9. Normative References

[IPSEC-ESP] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.

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

[PHB-ID] Black, D., Brim, S., Carpenter, B., and F. Le Faucheur, "Per Hop Behavior Identification Codes", RFC 3140, June 2001.

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

[RSVP-AGG] Baker, F., Iturralde, C., Le Faucheur, F., and B. Davie, "Aggregation of RSVP for IPv4 and IPv6 Reservations", RFC 3175, September 2001.

[RSVP-CRYPTO1] Baker, F., Lindell, B., and M. Talwar, "RSVP Cryptographic Authentication", RFC 2747, January 2000.

[RSVP-CRYPTO2] Braden, R. and L. Zhang, "RSVP Cryptographic Authentication -- Updated Message Type Value", RFC 3097, April 2001.

[RSVP-IPSEC] Berger, L. and T. O'Malley, "RSVP Extensions for IPSEC Data Flows", RFC 2207, September 1997.

[RSVP-MOD] Kompella, K. and J. Lang, "Procedures for Modifying the Resource reSerVation Protocol (RSVP)", BCP 96, RFC 3936, October 2004.

10. Informative References

[BW-REDUC] Polk, J. and S. Dhesikan, "A Resource Reservation Protocol (RSVP) Extension for the Reduction of Bandwidth of a Reservation Flow", RFC 4495, May 2006.

[GRE] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RSVP-PREEMP] Herzog, S., "Signaled Preemption Priority Policy Element", RFC 3181, October 2001.

[RSVP-PROCESS] Braden, R. and L. Zhang, "Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules", RFC 2209, September 1997.

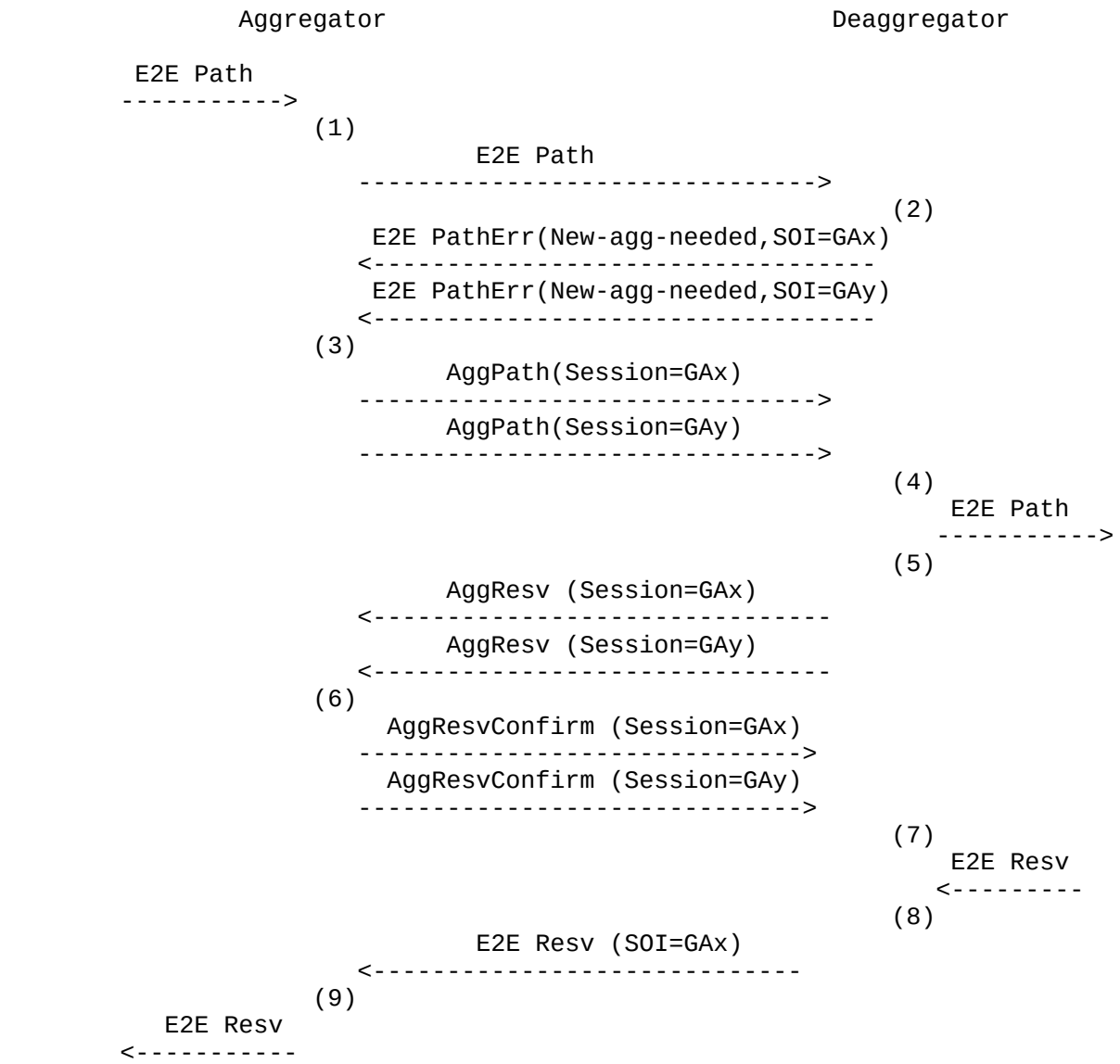
[RSVP-TE] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RSVP-TUNNEL] Terzis, A., Krawczyk, J., Wroclawski, J., and L. Zhang, "RSVP Operation Over IP Tunnels", RFC 2746, January 2000.

[SIG-NESTED] Baker, F. and P. Bose, "QoS Signaling in a Nested Virtual Private Network", Work in Progress, February 2007.

Appendix A. Example Signaling Flow

This appendix does not provide additional specification. It only illustrates the specification detailed in Section 4 through a possible flow of RSVP signaling messages. This flow assumes an environment where E2E reservations are aggregated over generic aggregate RSVP reservations. It illustrates a possible RSVP message flow that could take place in the successful establishment of a unicast E2E reservation that is the first between a given pair of Aggregator/Deaggregator.



(1) The Aggregator forwards E2E Path into the aggregation region after modifying its IP protocol number to RSVP-E2E-IGNORE

(2) Let's assume no Aggregate Path exists. To be able to accurately update the ADSPEC of the E2E Path, the Deaggregator needs the ADSPEC of Aggregate Path. In this example, the Deaggregator elects to instruct the Aggregator to set up Aggregate Path states for the two supported PHB-IDs. To do that, the Deaggregator

sends two E2E PathErr messages with a New-Agg-Needed PathErr code. Both PathErr messages also contain a SESSION-OF-INTEREST (SOI) object. In the first E2E PathErr, the SOI contains a GENERIC-AGGREGATE SESSION (GAx) whose PHB-ID is set to x. In the second E2E PathErr, the SOI contains a GENERIC-AGGREGATE SESSION (GAy) whose PHB-ID is set to y. In both messages the GENERIC-AGGREGATE SESSION contains an interface-independent Deaggregator address inside the DestAddress and appropriate values inside the vDstPort and Extended vDstPort fields.

(3) The Aggregator follows the request from the Deaggregator and signals an Aggregate Path for both GENERIC-AGGREGATE Sessions (GAx and GAy).

(4) The Deaggregator takes into account the information contained in the ADSPEC from both Aggregate Paths and updates the E2E Path ADSPEC accordingly. The Deaggregator also modifies the E2E Path IP protocol number to RSVP before forwarding it.

(5) In this example, the Deaggregator elects to immediately proceed with establishment of generic aggregate reservations for both PHB-IDs. In effect, the Deaggregator can be seen as anticipating the actual demand of E2E reservations so that resources are available on the generic aggregate reservations when the E2E Resv requests arrive, in order to speed up establishment of E2E reservations. Assume also that the Deaggregator includes the optional Resv Confirm Request in these Aggregate Resv.

(6) The Aggregator merely complies with the received ResvConfirm Request and returns the corresponding Aggregate ResvConfirm.

(7) The Deaggregator has explicit confirmation that both Aggregate Resvs are established.

(8) On receipt of the E2E Resv, the Deaggregator applies the mapping policy defined by the network administrator to map the E2E Resv onto a generic aggregate reservation. Let’s assume that this policy is such that the E2E reservation is to be mapped onto the generic aggregate reservation with PHB-ID=x. The Deaggregator knows that a generic aggregate reservation (GAx) is in place for the corresponding PHB-ID since (7). The Deaggregator performs admission control of the E2E Resv onto the generic aggregate reservation for PHB-ID=x (GAx). Assuming that the generic aggregate reservation for PHB-ID=x (GAx) had been established with sufficient bandwidth to support the E2E Resv, the Deaggregator adjusts its counter, tracking the unused bandwidth on the generic aggregate reservation. Then it forwards the E2E Resv to the Aggregator including a SESSION-OF-INTEREST object

conveying the selected mapping onto GAx (and hence onto PHB-ID=x).

(9) The Aggregator records the mapping of the E2E Resv onto GAx (and onto PHB-ID=x). The Aggregator removes the SOI object and forwards the E2E Resv towards the sender.

Authors' Addresses

Francois Le Faucheur
Cisco Systems, Inc.
Village d'Entreprise Green Side - Batiment T3
400, Avenue de Roumanille
06410 Biot Sophia-Antipolis
France
EMail: flefauch@cisco.com

Bruce Davie
Cisco Systems, Inc.
1414 Massachusetts Ave.
Boxborough, MA 01719
USA
EMail: bds@cisco.com

Pratik Bose
Lockheed Martin
700 North Frederick Ave.
Gaithersburg, MD 20879
USA
EMail: pratik.bose@lmco.com

Chris Christou
Booz Allen Hamilton
13200 Woodland Park Road
Herndon, VA 20171
USA
EMail: christou_chris@bah.com

Michael Davenport
Booz Allen Hamilton
Suite 390
5220 Pacific Concourse Drive
Los Angeles, CA 90045
USA
EMail: davenport_michael@bah.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.