        Generic Aggregation of Resource ReSerVation Protocol (RSVP)
             for IPv4 And IPv6 Reservations over PCN domains
                      draft-ietf-tsvwg-rsvp-pcn-07

Abstract

   This document specifies extensions to Generic Aggregated RSVP
   [RFC4860] for support of the PCN Controlled Load (CL) and Single
   Marking (SM) edge behaviors over a Diffserv cloud using Pre-
   Congestion Notification.

Status of this Memo

Copyright Notice

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 [RFC2119].

Table of Contents

1. Introduction

1.1 Objective

Pre-Congestion Notification (PCN) can support the quality of service
(QoS) of inelastic flows within a Diffserv domain in a simple,
scalable, and robust fashion. Two mechanisms are used: admission
control and flow termination. Admission control is used to decide
whether to admit or block a new flow request, while flow termination
is used in abnormal circumstances to decide whether to terminate some
of the existing flows. To support these two features, the overall
rate of PCN-traffic is metered on every link in the domain, and PCN-
packets are appropriately marked when certain configured rates are
exceeded. These configured rates are below the rate of the link,
thus providing notification to boundary nodes about overloads before
any congestion occurs (hence "pre-congestion" notification). The
PCN-egress-nodes measure the rates of differently marked PCN traffic
in periodic intervals and report these rates to the Decision Points
for admission control and flow termination; the Decision Points use
these rates to make decisions. The Decision Points may be collocated
with the PCN-ingress-nodes, or their function may be implemented in a
centralized node. For more details see [RFC5559], [RFC6661], and
[RFC6662].

The main objective of this document is to specify the signaling
protocol that can be used within a Pre-Congestion Notification (PCN)
domain to carry reports from a PCN-egress-node to a PCN Decision
point, considering that the PCN decision Point and PCN-ingress-node
are collocated.
If the PCN decision point is not collocated with the PCN-ingress-node
then additional signaling procedures are required that are out of
the scope of this document. Moreover, as mentioned above this
architecture conforms with PBAC (Policy-Based Admission Control),
when decision point is located in a centralized node [RFC2753].

Several signaling protocols can be used to carry reports from a PCN-
egress-node to a PCN-ingress-nodes. However, since both PCN-egress-
node and PCN-ingress-nodes are located on the data path, a signaling
protocol that follows the same path as the data path, like RSVP
(Resource Reservation Protocol), is more suited for this purpose. In
particular, this document specifies extensions to Generic
Aggregated RSVP [RFC4860] for support of the PCN Controlled Load (CL)
and Single Marking (SM) edge behaviors over a Diffserv cloud using
Pre-Congestion Notification.

1.2 Overview and Motivation

Two main Quality of Service (QoS) architectures have been specified

by the IETF. These are the Integrated Services (Intserv) [RFC1633]
architecture and the Differentiated Services (DiffServ) architecture
([RFC2475]).

Intserv provides methods for the delivery of end-to-end Quality of
Service (QoS) to applications over heterogeneous networks. One of the
QoS signaling protocols used by the Intserv architecture is the
Resource reServation Protocol (RSVP) [RFC2205], which can be used by
applications to request per-flow resources from the network. These
RSVP requests can be admitted or rejected by the network.
Applications can express their quantifiable resource requirements
using Intserv parameters as defined in [RFC2211] and [RFC2212]. The
Controlled Load (CL) service [RFC2211] is a quality of service (QoS)
closely approximating the QoS that the same flow would receive from a
lightly loaded network element. The CL service is useful for
inelastic flows such as those used for real-time media.

The DiffServ architecture can support the differentiated treatment of
packets in very large scale environments. While Intserv and RSVP
classify packets per-flow, Diffserv networks classify packets into
one of a small number of aggregated flows or "classes", based on the
Diffserv codepoint (DSCP) in the packet IP header. At each Diffserv
router, packets are subjected to a "per-hop behavior" (PHB), which is
invoked by the DSCP. The primary benefit of Diffserv is its
scalability, since the need for per-flow state and per-flow
processing, is eliminated.

However, DiffServ does not include any mechanism for communication
between applications and the network. Several solutions have been
specified to solve this issue. One of these solutions is Intserv over
Diffserv [RFC2998] including resource-based admission control (RBAC),
PBAC, assistance in traffic identification/classification, and
traffic conditioning. Intserv over Diffserv can operate over a
statically provisioned Diffserv region or RSVP aware. When it is RSVP
aware, several mechanisms may be used to support dynamic provisioning
and topology-aware admission control, including aggregate RSVP
reservations, per-flow RSVP, or a bandwidth broker. [RFC3175]
specifies aggregation of Resource ReSerVation Protocol (RSVP) end-to-
end reservations over aggregate RSVP reservations. In [RFC3175] the
RSVP aggregated reservation is characterized by a RSVP SESSION object
using the 3-tuple <source IP address, destination IP address,
Diffserv Code Point>.

Several scenarios require the use of multiple generic aggregate
reservations that are established for a given PHB from a given source
IP address to a given destination IP address, see [SIG-NESTED],
[RFC4860]. For example, multiple generic aggregate reservations
can be applied in the situation that multiple e2e reservations using
different preemption priorities need to be aggregated through a PCN-
domain using the same PHB. By using multiple aggregate reservations
for the same PHB allows enforcement of the different preemption
priorities within the aggregation region. 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. In particular,
[SIG-NESTED] discusses in detail how end-to-end RSVP reservations can
be established in a nested VPN environment through RSVP aggregation.

[RFC4860] provides generic aggregate reservations by extending
[RFC3175] to support multiple aggregate reservations for the same
source IP address, destination IP address, and PHB (or set of PHBs).
In particular, multiple such generic aggregate reservations can be
established for a given PHB from a given source IP address to a given
destination IP address. This is achieved by adding the concept of a
Virtual Destination Port and of an Extended Virtual Destination Port
in the RSVP SESSION object. In addition to this, the RSVP SESSION
object for generic aggregate reservations uses the PHB Identification
Code (PHB-ID) defined in [RFC3140], instead of using the Diffserv
Code Point (DSCP) used in [RFC3175]. The PHB-ID is used to identify
the PHB, or set of PHBs, from which the Diffserv resources are to be
reserved.

The RSVP like signaling protocol required to carry reports from a
PCN-egress-node to a PCN-ingress-node needs to follow the PCN
signaling requirements defined in [RFC6663]. In addition to
that the signaling protocol functionality supported by the PCN-
ingress-nodes and PCN-egress-nodes needs to maintain logical
aggregate constructs (i.e. ingress-egress-aggregate state) and be
able to map e2e reservations to these aggregate constructs. Moreover,
no actual reservation state is needed to be maintained inside the PCN
domain, i.e., the PCN-interior-nodes are not maintaining any
reservation state.

This can be accomplished by two possible approaches:

Approach (1):

  o) adapting the RFC 4860 aggregation procedures to fit the PCN
     requirements with as little change as possible over the RFC 4860
     functionality

  o) hence performing aggregate RSVP signaling (even if it is to be
     ignored by PCN interior nodes)

  o) using this aggregate RSVP signaling procedures to carry PCN
     information from PCN-egress-node to the PCN-ingress-node.

Approach (2):

  o) adapting the RFC 4860 aggregation procedures to fit the PCN
     requirements with more significant changes over RFC4860 (i.e.
     the aspect of the procedures that have to do with maintaining
     aggregate states and to do with mapping the e2e reservations to
     aggregate constructs are kept, but the procedures that have to
     do with the aggregate RSVP signaling and aggregate reservation
     establishment/maintenance are dropped).

  o) hence not performing aggregate RSVP signaling

o) piggy-backing of the PCN information inside the e2e RSVP
      signaling.

Both approaches are probably viable, however, since the RFC 4860
operations have been thoroughly studied and implemented, it can be
considered that the RFC 4860 solution can better deal with the more
challenging situations (rerouting in the PCN domain, failure of an
PCN-ingress-node, failure of an PCN-egress-node, rerouting towards a
different edge, etc.). This is the reason for choosing Approach (1)
for the specification of the signaling protocol used to carry
PCN information from the PCN-egress-node to the PCN-ingress-node.

In particular, this document specifies extensions to Generic
Aggregated RSVP [RFC4860] for support of the PCN Controlled Load (CL)
and Single Marking (SM) edge behaviors over a Diffserv cloud using
Pre-Congestion Notification.

This document follows the PCN signaling requirements defined in
[RFC6663] and specifies extensions to Generic Aggregated RSVP
[RFC4860] for support of PCN edge behaviors as specified in
[RFC6661] and [RFC6662]. Moreover, this document specifies how RSVP
aggregation can be used to setup and maintain: (1) Ingress Egress
Aggregate (IEA) states at Ingress and Egress nodes and (2) generic
aggregation of RSVP end-to-end RSVP reservations over PCN (Congestion
and Pre-Congestion Notification) domains.

To comply with this specification, PCN-nodes MUST be able to
support the functionality specified in [RFC5670], [RFC5559],
[RFC6660], [RFC6661], [RFC6662]. Furthermore, the PCN-boundary-nodes
MUST support the RSVP generic aggregated reservation procedures
specified in [RFC4860] which are augmented with procedures specified
in this document.

1.3. Terminology

This document uses terms defined in [RFC4860], [RFC3175], [RFC5559],
[RFC5670], [RFC6661], [RFC6662].

For readability, a number of definitions from [RFC3175] as well as
definitions for terms used in [RFC5559], [RFC6661], and [RFC6662] are
provided here, where some of them are augmented with new meanings:

Aggregator   This is the process in (or associated with) the
                  router at the ingress edge of the aggregation region
                  (with respect to the end-to-end RSVP reservation)
                  and behaving in accordance with [RFC4860]. In this
                  document, it is also the PCN-ingress-node and the
                  decision point. It is important to notice that in
                  the context of this document the Aggregator MUST be
                  able to determine the Deaggregator using the
                  procedures specified in Section 4 of [RFC4860] and

in Section 1.4.2 of [RFC3175].

Deaggregator   This is the process in (or associated with) the
                  router at the egress edge of the aggregation region
                  (with respect to the end-to-end RSVP reservation)
                  and behaving in accordance with [RFC4860]. In this
                  document, it is also the PCN-egress-node.

E2E (or e2e)   end to end

E2E Reservation   This is an RSVP reservation such that:

                  (i)   corresponding RSVP Path messages are initiated
                         upstream of the Aggregator and terminated
                         downstream of the Deaggregator, and

                  (ii)   corresponding RSVP Resv messages are initiated
                         downstream of the Deaggregator and terminated
                         upstream of the Aggregator, and

                  (iii) this RSVP reservation is aggregated over an
                         Ingress Egress Aggregate (IEA) between the
                         Aggregator and Deaggregator.
                  An E2E RSVP reservation may be a per-flow
                  reservation, which in this document is only
                  maintained at the PCN-ingress-node and PCN-egress-
                  node. Alternatively, the E2E reservation may itself
                  be an aggregate reservation of various types (e.g.,
                  Aggregate IP reservation, Aggregate IPsec
                  reservation, see [RFC4860]). As per regular RSVP
                  operations, E2E RSVP reservations are
                  unidirectional.

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 [RFC3140].

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)

                     An identifier used in the SESSION that remains
                     constant over the life of the generic aggregate
                     reservation. The length of this idenitifier is 32-
                     bits when IPv4 addresses are used and 128 bits when

IPv6 addresses are used.

                          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
                          or 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. In this document, the Extended
                          vDstPort SHOULD contain the IPv4 or IPv6 address of
                          the Aggregator.

      PCN-domain: a PCN-capable domain; a contiguous set of
                          PCN-enabled nodes that perform Diffserv scheduling
                          [RFC2474]; the complete set of PCN-nodes that in
                          principle can, through PCN-marking packets,
                          influence decisions about flow admission and
                          termination for the PCN-domain; includes
                          the PCN-egress-nodes, which measure these
                          PCN-marks, and the PCN-ingress-nodes.

      PCN-boundary-node: a PCN-node that connects one PCN-domain to a node
                          either in another PCN-domain or in a non-PCN-domain.

      PCN-interior-node: a node in a PCN-domain that is not a PCN-
                          boundary-node.

      PCN-node: a PCN-boundary-node or a PCN-interior-node.

      PCN-egress-node: a PCN-boundary-node in its role in handling
                          traffic as it leaves a PCN-domain. In this
                          document the PCN-ingress-node operates also as a
                          deaggregator.

      PCN-ingress-node: a PCN-boundary-node in its role in handling
                          traffic as it enters a PCN-domain. In this
                          document the PCN-ingress-node operates also as a
                          Decision Point and aggregator.

      PCN-traffic,
      PCN-packets,
      PCN-BA: a PCN-domain carries traffic of different Diffserv
                          behavior aggregates (BAs) [RFC2474]. The PCN-BA
                          uses the PCN mechanisms to carry PCN-traffic, and
                          the corresponding packets are PCN-packets.
                          The same network will carry traffic of other
                          Diffserv BAs. The PCN-BA is

distinguished by a combination of the Diffserv
codepoint (DSCP) and ECN fields.

   Microflow: a single instance of an application-to-application
  (from [RFC2474])   flow of packets which is identified by source
                       address, destination address, protocol id, and
                       source port, destination port (where applicable).

   e2e microflow   a microflow where its associated packets are being
                       forwarded on an E2E path.

   PCN-flow: the unit of PCN-traffic that the PCN-boundary-node
                       admits (or terminates); the unit could be a single
                       e2e microflow (as defined in [RFC2474]) or some
                       identifiable collection of microflows.

RSVP generic aggregated reservation: an RSVP reservation that is
                       identified by using the RSVP SESSION object
                       for generic RSVP aggregate reservation. This RSVP
                       SESSION object is based on the RSVP SESSION object
                       specified in [RFC4860] augmented with the following
                       information:

                       o) the IPv4 DestAddress, IPv6 DestAddress SHOULD be
                          set to the IPv4 or IPv6 destination addresses,
                          respectively, of the Deaggregator (PCN-egress-
                          node)

                       o) PHB-ID (Per Hop Behavior Identification Code)
                          SHOULD be set equal to PCN-compatible Diffserv
                          codepoint(s).

                       o) Extended vDstPort SHOULD be set to the IPv4 or
                          IPv6 destination addresses, of the Aggregator
                          (PCN-ingress-node)

   Ingress-egress-aggregate (IEA):
                       The collection of PCN-packets from all PCN-flows
                       that travel in one direction between a specific pair
                       of PCN-boundary-nodes. An ingress-egress-aggregate
                       is identified by the combination of (1) PCN-BA
                       (i.e., combination of the DSCP and ECN fields),(2)
                       IP addresses of the specific pair of
                       PCN-boundary-nodes used by the
                       ingress-egress-aggregate. In this document one RSVP
                       generic aggregated reservation is mapped to only
                       one ingress-egress-aggregate, while one
                       ingress-egress-aggregate is mapped to either
                       one or to more than one RSVP generic aggregated
                       reservations.

   PCN-admission-state:

                  The state ("admit" or "block") derived by the
                  Decision Point for a given ingress-egress-aggregate
                  based on statistics about PCN-packet marking. The
                  Decision Point decides to admit or block new flows
                  offered to the aggregate based on the current value
                  of the PCN-admission-state.

  Congestion level estimate (CLE):

                  The ratio of PCN-marked to total PCN-traffic
                  (measured in octets) received for a given ingress-
                  egress-aggregate during a given measurement period.
                  The CLE is used to derive the PCN-admission-state
                  and is also used by the report suppression procedure
                  if report suppression is activated.

  t-meas:

                  A configurable time interval that defines the
                  measurement period over which the PCN-egress-node
                  collects statistics relating to PCN-traffic marking.

  t-fail:

                  An interval after which the Decision Point (i.e.,
                  PCN-ingress-node in this document) concludes that
                  communication from a given PCN-egress-node has failed
                  if it has received no reports from the
                  PCN-egress-node during that interval.

  t-recvFail:

                  A timer per ingress-egress-aggregate that the
                  Decision point (i.e., PCN-ingress-node) sets every
                  time it receives an RSVP Aggregated RESV message for
                  that ingress-egress-aggregate. When its value
                  reaches t-fail it is assumed that the PCN-ingress-
                  node has lost contact with the PCN-egress-node.
                  Therefore the PCN-ingress-node blocks admission of
                  new PCN-flows into that aggregate and raises a
                  management alarm.

1.4. Organization of This Document

This document is organized as follows. Section 2 gives an overview of
RSVP extensions and operations. The elements of the used procedures
are specified in Section 3. Section 4 describes the protocol
elements. The security considerations are given in section 5 and the
IANA considerations are provided in Section 6.

2. Overview of RSVP extensions and Operations

2.1 Overview of RSVP Aggregation Procedures in PCN domains

The PCN-boundary-nodes, see Figure 1, can support RSVP SESSIONS for
generic aggregated reservations {RFC4860], which are depending on
ingress-egress-aggregates. In particular, one RSVP generic aggregated
reservation matches to only one ingress-egress-aggregate.

However, one ingress-egress-aggregate matches to either
one or more than one RSVP generic aggregated reservations.
In addition, to comply with this specification it is considered that
the PCN-boundary nodes are able to distinguish by using the addresses
that the RSVP messages are addressed to, and process (1) RSVP
SESSIONS for generic aggregated sessions and their messages according
to [RFC4860], (2) e2e RSVP sessions and messages according to
[RFC2205]. Furthermore, it is considered that by configuration the
PCN-interior-nodes are not able to distinguish neither RSVP generic
aggregated sessions and their associated messages [RFC4860], nor e2e
RSVP sessions and their associated messages [RFC2205].

```
                        ---------------------------
                      /         PCN-domain         \
     |----|        |                          |       |----|
  H--| R  |\ |-----|                          |------| /| R   |-->H
  H--|    |\\|      |    |---|       |---|     |      |//|     |-->H
    |----| \|       |    | I |       | I |     |      |/ |----|
             | Agg |=====================>| Deag |
           /|      |   |   |     |   |   |      |\
  H--------//|      |   |---|     |---|     |      |\\-------->H
  H--------/ |-----|                          |------| \-------->H
              |                              |
              \                             /
                ---------------------------
```
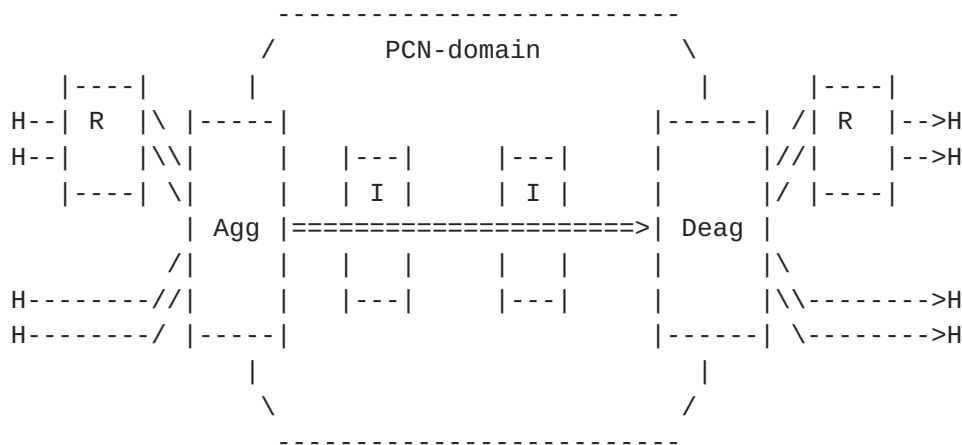
H   = Host requesting end-to-end RSVP reservations
R   = RSVP router
Agg   = Aggregator (PCN-ingress-node)
Deag   = Deaggregator (PCN-egress-node)
I   = Interior Router (PCN-interior-node)
-->   = E2E RSVP reservation
==>   = Aggregate RSVP reservation

          Figure 1 : Aggregation of E2E Reservations
           over Generic Aggregate RSVP Reservations
               in PCN domains, based on [RFC4860]

Moreover, each Aggregator and Deaggregator (i.e., PCN-boundary-nodes)
MUST support policies to initiate and maintain for each pair of
PCN-boundary-nodes of the same PCN-domain one ingress-egress-
aggregate. Both the Aggregator and Deaggregator can maintain one or
more RSVP generic aggregated Reservations, but the Deaggregator is
the entity that initiates these RSVP generic aggregated reservations.
Note that one RSVP generic aggregated reservation matches to only
one ingress-egress-aggregate, while one ingress-egress-aggregate
matches to either one or to more than one RSVP generic aggregated
reservations. This can be accomplished by using for the different
RSVP generic aggregated reservations the same combinations of

ingress and egress identifiers, but with a different PHB-ID value
(see [RFC4860]). The procedures for aggregation of E2E reservations
over generic aggregate RSVP reservations are the same as the
procedures specified in Section 4 of [RFC4860], augmented with the
following ones, see also Section 2.5:

o) Aggregator (PCN-ingress-node) and Deaggregator (PCN-egress-node)
   MUST be able to determine, for each received e2e Path message, in
   which ingress-egress-aggregate it can be mapped to.

o) Depending on policies the Aggregator and Deaggregator MUST be able
   to decide whether a RSVP generic aggregate reservations can be
   mapped into an ingress-egress-aggregate, see Section 2.5 for more
   details.

2.2 PCN Marking and encoding and transport of pre-congestion information

The method of PCN marking within the PCN domain is based on
[RFC5670]. In addition, the method of encoding and transport of pre-
congestion information is based [RFC6660]. The PHB-ID (Per Hop
Behavior Identification Code) used SHOULD be set equal
to PCN-compatible Diffserv codepoint(s).

2.3. Traffic Classification Within The Aggregation Region

The PCN-ingress marks a PCN-BA using PCN-marking (i.e., combination
of the DSCP and ECN fields), which interior nodes use to
classify PCN-traffic. The PCN-traffic (e.g., e2e microflows)
belonging to an ingress-egress-aggregate can be classified only at
the PCN-boundary-nodes using the combination of (1) PCN-BA (i.e.,
combination of the DSCP and ECN fields), (2) IP addresses of the
specific pair of PCN-boundary-nodes used by a ingress-egress-
aggregate. The method of classification and traffic conditioning of
PCN-traffic and non-PCN traffic and PHB configuration is described in
[RFC6661] and [RFC6662]. Moreover, the PCN-traffic (e.g., e2e
microflows) belonging to a RSVP generic aggregated reservation can be
classified only at the PCN-boundary-nodes (i.e., Aggregator and
Deaggregator) by using the RSVP SESSION object for RSVP generic
aggregated reservations, see Section 2.1 of [RFC4860]. It is
considered that tunnels need to be used between Aggregators and
Deaggregators, using the same procedures as specified in Section 4 of
   [RFC4860].

2.4. Deaggregator (PCN-egress-node) Determination

To comply with this specification it is considered that in order to
determine the Deaggregator, the same methods can be used as the ones
described in Section 4 of [RFC4860] and in Section 1.4.2 of
   [RFC3175]. In the context of this document this can be determined
very easily, since from the point of RSVP, the next RSVP hop for the
Aggregator in the downstream direction is the Deaggregator and the
next RSVP hop for the Deaggregator in the upstream direction is the
Aggregator.

2.5. Mapping E2E Reservations Onto Aggregate Reservations

   To comply with this specification it is considered that for the
   mapping of e2e reservations onto aggregate reservations, the same
   methods can be used as the ones described in Section 4 of [RFC4860],
   augmented by the following rules:

   o) An Aggregator (PCN-ingress-node) MUST be able to determine for
      each e2e Path message that arrives at its external interface in
      which ingress-egress-aggregate it can be mapped to. This is
      possible, see also Section 2.4, since from the point of RSVP, the
      Deaggregator (PCN-egress-node) is one RSVP hop away from the
      Aggregator (PCN-ingress-node). The Aggregator (PCN-ingress-node)
      uses PCN related information sent by the Deaggregator to
      map RSVP generic aggregated states to ingress-egress-aggregates.

   o) A PCN-ingress-node (Aggregator) or PCN-egress-node (Deaggregator)
      MUST use one or more policies to determine whether a RSVP generic
      aggregated reservation can be mapped into an ingress-egress-
      aggregate. Note that one RSVP generic aggregated reservation
      matches to only one ingress-egress-aggregate, while one ingress-
      egress-aggregate matches to either one or to more than one RSVP
      generic aggregated reservations. The Aggregator or the
      Deaggregator MUST be able to map RSVP generic aggregated
      reservations into ingress-egress-aggregates. This can be
      accomplished by using for the different RSVP generic aggregated
      reservations the same combinations of ingress and egress
      identifiers, but with a different PHB-ID value (see [RFC4860]). In
      particular, each RSVP generic aggregated reservation is identified
      by using the RSVP SESSION object [RFC4860]. The RSVP SESSION
      object for generic aggregate reservations is based on the RSVP
      SESSION object specified in [RFC4860] augmented with the following
      information:

      o) the IPv4 DestAddress, IPv6 DestAddress MUST be set to the IPv4
         or IPv6 destination addresses, respectively, of the
         Deaggregator (PCN-egress-node), see [RFC4860]. Note that the
         PCN-domain is considered as being only one RSVP hop (for
         Generic aggregated RSVP or e2e RSVP). This means that the next
         RSVP hop for the Aggregator in the downstream direction is the
         Deaggregator and the next RSVP hop for the Deaggregator in the
         upstream direction is the Aggregator.

      o) PHB-ID (Per Hop Behavior Identification Code) SHOULD be set
         equal to PCN-compatible Diffserv codepoint(s).

      o) Extended vDstPort SHOULD be set to the IPv4 or IPv6 destination
         addresses, of the Aggregator (PCN-ingress-node), see [RFC4860].

2.6. Size of Aggregate Reservations

To comply with this specification it is considered that for the determination of the size of the RSVP generic aggregate reservations, the same methods can be used as the ones described in [RFC4860] and Section 1.4.4. of [RFC3175].

2.7. E2E Path ADSPEC update

To comply with this specification it is considered that for the update of the e2e Path ADSPEC, the same methods can be used as the ones described in [RFC4860].

2.8. Intra-domain Routes

The PCN-interior-nodes are neither maintaining e2e RSVP nor RSVP generic aggregation states and reservations. Therefore, intra-domain route changes will not affect intra-domain reservations since such reservations are not maintained by the PCN-interior-nodes. Furthermore, it is considered that by configuration, the PCN-interior-nodes are not able to distinguish neither RSVP generic aggregated sessions and their associated messages [RFC4860], nor e2e RSVP sessions and their associated messages [RFC2205].

2.9. Inter-domain Routes

The PCN-charter scope precludes inter-domain considerations. However, for solving inter-domain routes changes associated with the operation of the RSVP messages, the same methods SHOULD be used as the ones described in [RFC4860] and in Section 1.4.7 of [RFC3175].

2.10. Reservations for Multicast Sessions

PCN does not consider reservations for multicast sessions.

2.11. Multi-level Aggregation

PCN does not consider multi-level aggregations within the PCN domain. Therefore, the PCN-interior-nodes are not supporting multi-level aggregation procedures. However, the Aggregator and Deaggregator SHOULD support the multi-level aggregation procedures specified in [RFC4860] and in Section 1.4.9 of [RFC3175].

2.12. Reliability Issues

To comply with this specification it is considered that for solving

possible reliability issues, the same methods can be used as the ones
described in Section 4 of [RFC4860].

2.13. Message Integrity and Node Authentication

   To comply with this specification it is considered that for message
   integrity and node authentication, the same methods can be used as
   the ones described in Section 4 of [RFC4860] and [RFC5559].

3. Elements of Procedure

   This section describes the procedures used to implement the
   aggregated RSVP procedure over PCN. It is considered that the
   procedures for aggregation of e2e reservations over generic aggregate
   RSVP reservations are same as the procedures specified in Section
   4 of [RFC4860]. Please refer to [RFC4860] for all the below error
   cases:
       *) Incomplete message
       *) Unexpected objects

3.1. Receipt of E2E Path Message By PCN-ingress-node (aggregating
         router)

   When the e2e Path message arrives at the exterior interface of the
   Aggregator, i.e., PCN-ingress-node, then standard RSVP generic
   aggregation [RFC4860] procedures are used, augmented with the
   following rules:

  o) The e2e RSVP reservation session associated with an e2e Path
     message that arrives at the external interface of the PCN-
     ingress-node is mapped/matched onto an PCN ingress-egress-
     aggregate.

  o) If the timer t-recvFail does NOT expire for a given PCN-egress-
     node, then:

         o) If (1) the PCN-admission state for the ingress-egress-
            aggregate associated with the received e2e Path is "admit",
            the Decision Point (i.e., PCN-ingress-node) SHOULD allow the
            new flow to be admitted to that PCN ingress-egress-
            aggregate, see [RFC6661] and [RFC6662]. The e2e Path message
            is then forwarded towards destination.

         o) If for the same PCN ingress-egress-aggregate
            the PCN-admission-state is "block", the request SHOULD NOT
            be admitted by the PCN-ingress-node (Aggregator) and an e2e
            PathErr message SHOULD be generated, using standard e2e RSVP
            procedures [RFC4495]. This e2e PathErr message is sent to
            the originating sender of the e2e Path message, using
            standard e2e RSVP procedures [RFC2205], [RFC4495]. This e2e
            PathErr message is sent to the originating sender of the e2e
            Path message. The e2e RSVP error code "01: Admission Control

failure" and the "Sub-code = 2: Requested bandwidth
unavailable " specified in Appendix B of [RFC2205] SHOULD be
used for this purpose.

When the originating sender receives this e2e PathErr
message it SHOULD apply a PCN specific policy to generate an
e2e PathTear message to release all the possible Path states
initiated on the e2e RSVP aware nodes on the path towards
the PCN-ingress-node (Aggregator).

o) If the timer t-recvFail expires for a given PCN-egress-node, the
   Decision Point (i.e., PCN-ingress-node) SHOULD NOT
   allow the e2e microflow (i.e., PCN-flow) to be admitted to that
   PCN ingress-egress-aggregate, see [RFC6661], [RFC6662]. The
   admission or rejection procedure of a PCN-flow into the PCN-
   domain is defined in detail in: [RFC6661] and [RFC6662].
   If the Aggregator is not able to admit the e2e microflow it
   SHOULD then generate an e2e PathErr message using standard e2e
   RSVP procedures [RFC4495]. This e2e PathErr message is sent to
   the originating sender of the e2e Path message. The e2e RSVP
   error code "01: Admission Control failure" and the "Sub-code =
   2: Requested bandwidth unavailable " specified in Appendix B of
   [RFC2205] SHOULD be used for this purpose. When the originating
   sender receives this e2e PathErr message it SHOULD apply a PCN
   specific policy to generate an e2e PathTear message to release all
   the possible Path states initiated on the e2e RSVP aware nodes on
   the path towards the PCN-ingress-node (Aggregator).

The way of how the PCN-admission-state is maintained is specified in
[RFC6661] and [RFC6662]. The way of how the RSVP generic aggregated
reservation state is maintained is specified in [RFC4860].

3.2. Handling Of E2E Path Message By Interior Routers

The e2e Path messages traverse zero or more PCN-interior-nodes.
The PCN-interior-nodes receive the e2e Path message on an interior
interface and forward it on another interior interface. It is
considered that by configuration the PCN-interior-nodes are not able
to distinguish neither e2e RSVP sessions and their associated
messages [RFC2205]. Therefore, the e2e Path messages are simply
forwarded as normal IP datagrams.

3.3. Receipt of E2E Path Message By PCN-egress-node (Deaggregating router)

When receiving the e2e Path message the PCN-egress-node
(Deaggregator) performs main regular [RFC4860] procedures, augmented
with the following rules:

  o) The PCN-egress-node MUST NOT perform the RSVP-TTL vs IP TTL-
     check and MUST NOT update the ADspec Break bit. This is because

the whole PCN-domain is effectively handled by e2e RSVP as a
virtual link on which integrated service is indeed supported
(and admission control performed) so that the Break bit MUST
NOT be set, see also [draft-lefaucheur-rsvp-ecn-01].

o) If the Deaggregator does not maintain any RSVP generic
   aggregated reservation states, then one or more of such states
   are created during this step. Moreover, also at this step
   the Deaggregator maps the new generated RSVP generic
   aggregated reservations onto one ingress-egress-aggregate
   maintained by the Deaggregator (PCN-egress-node), see Section
   2.5.

The PCN-egress-nodes forwards the e2e Path message towards the
receiver.

3.4. Initiation of new Aggregate Path Message By PCN-ingress-node
     (Aggregating Router)

To comply with this specification it is considered that for the
initiation of the new RSVP aggregated Path message by the PCN-
ingress-node (Aggregator), the same methods can be used as the ones
described in [RFC4860].

3.5. Handling Of new Aggregate Path Message By Interior Routers

The Aggregate Path messages traverse zero or more PCN-interior-nodes.
The PCN-interior-nodes receive the e2e Path message on an interior
interface and forward it on another interior interface.
It is considered that by configuration, the PCN-interior-nodes are
not able to distinguish neither RSVP generic aggregated sessions and
their associated messages [RFC4860]. Therefore, the Aggregated Path
messages are simply forwarded as normal IP datagrams.

3.6. Handling of E2E Resv Message by Deaggregating Router

When the e2e Resv message arrives at the exterior interface of the
Deaggregator, i.e., PCN-egress-node, then standard RSVP
aggregation [RFC4860] procedures are used. It is important to be
noticed that according to [RFC4860] the Deaggregator is responsible
of performing admission control of the E2E RESV onto the generic
aggregate reservation.

3.7. Handling Of E2E Resv Message By Interior Routers

The e2e Resv messages traverse zero or more PCN-interior-nodes. The
PCN-interior-nodes receive the e2e Resv message on an interior
interface and forward it on another interior interface. It is
considered that by configuration the PCN-interior-nodes are not able
to distinguish neither e2e RSVP sessions and their associated
messages [RFC2205]. Therefore, the e2e Resv messages are simply
forwarded as normal IP datagrams.

3.8. **Initiation of New Aggregate Resv Message By Deaggregating Router**

   To comply with this specification it is considered that for the
   initiation of the new RSVP aggregated Resv message by the PCN-
   egress-node (Deaggregator), the same methods can be used as the ones
   described in Section 4 of [RFC4860] augmented with the following
   rules:

     o) At the end of each t-meas measurement interval, or less
        frequently if "optional report suppression" is activated, see
        [RFC6661], and [RFC6662], the PCN-egress-node MUST include the
        new PCN object that will be sent to the associated Decision
        Point (i.e., PCN-ingress-node). The PCN-egress-node reports the
        data it measures for a particular ingress-egress-aggregate in a
        PCN object, as specified in Section 4 of this document (see
        [RFC6661], and [RFC6662]). The address of the PCN-ingress-
        node, i.e., Aggregator, is the one specified in the same
        ingress-egress-aggregate. It is considered that the ingress-
        egress-aggregate state stores both IP addresses of the PCN-
        ingress-node, i.e., Aggregator, and of the IP-egress-node, i.e.,
        Deaggregator.

3.9. **Handling of Aggregate Resv Message by Interior Routers**

   The Aggregated Resv messages traverse zero or more PCN-interior-
   nodes. The PCN-interior-nodes receive the Aggregated Resv message on
   an interior interface and forward it on another interior interface.
   It is considered that by configuration, the PCN-interior-nodes are
   not able to distinguish neither RSVP generic aggregated sessions and
   their associated messages [RFC4860]. Therefore, the Aggregated Resv
   messages are simply forwarded as normal IP datagrams.

3.10. **Handling of E2E Resv Message by Aggregating Router**

   When the e2e Resv message arrives at the interior interface of the
   Aggregating router, i.e., PCN-ingress-node, then standard RSVP
   aggregation [RFC4860] procedures are used.

3.11. **Handling of Aggregated Resv Message by Aggregating Router**

   When the Aggregated Resv message arrives at the interior interface of
   the Aggregating router, i.e., PCN-ingress-node, then standard RSVP
   aggregation [RFC4860] procedures are used, augmented with the
   following rules:

     o) If the Decision Point is not collocated with the PCN-ingress-
         node, then other procedures need to be specified of handling the
         Aggregated Resv Message by the Aggregating router, i.e., PCN-
         ingress-node. Even though these procedures are out of the scope
         of this document, the PCN-ingress-node can refer to a central

decision point which can respond to the PCN ingress as per
[RFC2753]

   o) If the Decision point is collocated with the PCN-ingress-node,
       then the PCN-ingress-node (i.e. Aggregator) SHOULD use the
       information carried by the PCN objects, see Section 4, to map
       the RSVP generic aggregated state onto the maintained ingress-
       egress-aggregate state at the Aggregator (PCN-ingress-node).
       Furthermore, the Aggregator follows the steps specified in
       [RFC6661], [RFC6662]. Using the information contained in the PCN
       object the Aggregator (i.e., PCN-ingress-node) can decide
       whether the PCN-admission state for the ingress-egress-aggregate
       is "admit" or "reject". Moreover, when the Aggregator (i.e.,
       PCN-ingress-node) needs to terminate an amount of traffic
       associated with one ingress-egress-aggregate (see bullet 2 in
       Section 3.3.2 of [RFC6661] and [RFC6662]), then several
       procedures of terminating e2e microflows can be deployed. The
       default procedure of terminating e2e microflows (i.e., PCN-
       flows) is as follows, see e.g., [RFC6661].
       For the same ingress-egress-aggregate, select a number of e2e
       microflows to be terminated in order to decrease the total
       incoming amount of bandwidth associated with one ingress-egress-
       aggregate by the amount of traffic to be terminated, see above.
       In this situation the same mechanisms for terminating an e2e
       microflow can be followed as specified in [RFC2205]. However,
       based on a local policy, the Aggregator could use other ways of
       selecting which microflows should be terminated.
       For example, for the same ingress-egress-aggregate, select a
       number of e2e microflows to be terminated or to reduce their
       reserved bandwidth in order to decrease the total incoming
       amount of bandwidth associated with one ingress-egress-aggregate
       by the amount of traffic to be terminated. In this situation the
       same mechanisms for terminating an e2e microflow or reducing
       bandwidth associated with an e2e microflow can be followed as
       specified in [RFC4495].

3.12. Removal of E2E Reservation

   To comply with this specification it is considered that for the
   removal of e2e reservations, the same methods can be used as the ones
   described in Section 4 of [RFC4860] and [RFC4495], augmented by the
   methods described in Section 3.11.

3.13. Removal of Aggregate Reservation

   To comply with this specification it is considered that for the
   removal of RSVP generic aggregated reservations, the same methods can
   be used as the ones described in Section 4 of [RFC4860] and Section
   2.10 of [RFC3175]. In particular, should an aggregate reservation go
   away (presumably due to a configuration change, route change, or
   policy event), the e2e reservations it supports are no longer active.
   They must be treated accordingly.

3.14. Handling of Data On Reserved E2E Flow by Aggregating Router

The handling of data on the reserved e2e Flow by Aggregating Router
is using the procedures described in [RFC4860] augmented with:

o) Regarding, PCN marking and traffic classification the procedures
    defined in Section 2.2 and 2.4 of this document are used.

3.15. Procedures for Multicast Sessions

In this document no multicast sessions are considered.

3.16. Misconfiguration of PCN-node

In an event where a PCN-node is misconfigured within a PCN-domain,
the desired behavior is same as described in Section 3.9. Therefore,
the Aggregated Resv messages are simply forwarded as normal IP
datagrams.

4. Protocol Elements

The protocol elements in this document are using the protocol
elements defined in Section 4 [RFC4860] and Section 3 of [RFC3175]
augmented with the following rules:

o) A PCN-egress-node (i.e., Deaggregator) SHOULD send periodically
   and at the end of each t-meas measurement interval, or less
   frequently if "optional report suppression" is activated, an
   (refresh) aggregated RSVP message to the PCN-ingress-node (i.e.
   aggregator), see [RFC6661] and [RFC6662].

o) the DSCP value included in the SESSION object, SHOULD be set equal
   to a PCN-compatible Diffserv codepoint.

o) An aggregated Resv message MUST carry one or more PCN
   objects, see Section 4.1, to report the data measured by an
   PCN-egress-node (i.e., Deaggregator).

o) As described in [RFC6661], [RFC6663], PCN reports
   from the PCN-egress-node (Deaggregator) to the decision point may
   contain flow identifiers for individual flows within an
   ingress-egress-aggregate that have recently experienced
   excess-marking. Hence, the PCN report messages used by the PCN CL
   edge behavior MUST be capable of carrying sequences of octet
   strings constituting such identifiers. When the PCN CL edge
   behavior is used, the individual flow identifiers need to be
   included in specific PCN objects, see Section 4.1
   (RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs,
    RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs)

4.1 PCN objects

The PCN object reports data measured by a PCN-egress-node and
carried by the generic aggregated RESV messages specified in
[RFC4860]. PCN objects are defined for different PCN edge behavior

drafts. This document defines six types of PCN objects that belong
into the SESSION Class and need to be carried by
Aggregate RESV messages. These objects are:

   RSVP-AGGREGATE-IPv4-PCN-SM, RSVP-AGGREGATE-IPv6-PCN-SM,
   RSVP-AGGREGATE-IPv4-PCN-CL, RSVP-AGGREGATE-IPv6-PCN-CL,
   RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs, RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs.

   o) RSVP-AGGREGATE-IPv4-PCN-SM: Single Marking (SM) PCN object, when
      IPv4 addresses are used:
      Class = 1 (SESSION)
      C-Type = RSVP-AGGREGATE-IPv4-PCN-SM (to be replaced by IANA)

```
       +------------+------------+------------+------------+
       |       IPv4 PCN-ingress-node Address (4 bytes)    |
       +------------+------------+------------+------------+
       |       IPv4 PCN-egress-node Address (4 bytes)     |
       +------------+------------+------------+------------+
       |        rate of not marked PCN-traffic (NM-rate)  |
       +------------+------------+------------+------------+
       |         rate of PCN-marked PCN-traffic (PM-rate) |
       +------------+------------+------------+------------|
```

   o) RSVP-AGGREGATE-IPv6-PCN-SM: Single Marking (SM) PCN object, when
      IPv6 addresses are used:

      Class = 1 (SESSION)
      C-Type = RSVP-AGGREGATE-IPv6-PCN-SM (to be replaced by IANA)

```
       +------------+------------+------------+------------+
       |                                                  |
       +                                                  +
       |                                                  |
       +      IPv6 PCN-ingress-node Address (16 bytes)    +
       |                                                  |
       +                                                  +
       |                                                  |
       +------------+------------+------------+------------+
       |                                                  |
       +                                                  +
       |                                                  |
       +      IPv6 PCN-egress-node Address (16 bytes)     +
       |                                                  |
       +                                                  +
       |                                                  |
       +------------+------------+------------+------------+
       |        rate of not marked PCN-traffic (NM-rate)  |
       +------------+------------+------------+------------+
       |         rate of PCN-marked PCN-traffic (PM-rate) |
       +------------+------------+------------+------------+
```

o) RSVP-AGGREGATE-IPv4-PCN-CL: Controlled (CL) PCN object, IPv4
      addresses are used:
      Class = 1 (SESSION)
      C-Type = RSVP-AGGREGATE-IPv4-PCN-CL (To be replaced by IANA)

```
   +------------+------------+------------+------------+
   |        IPv4 PCN-ingress-node Address (4 bytes)    |
   +------------+------------+------------+------------+
   |        IPv4 PCN-egress-node Address (4 bytes)     |
   +------------+------------+------------+------------+
   |        rate of not marked PCN-traffic (NM-rate)   |
   +------------+------------+------------+------------+
   |  rate of threshold-marked PCN-traffic (ThM-rate)  |
   +------------+------------+------------+------------+
   |  rate of excess-traffic-marked PCN-traffic (ETM-rate) |
   +------------+------------+------------+------------+
```

  o) RSVP-AGGREGATE-IPv6-PCN-CL: Controlled (CL) PCN object, IPv6
     addresses are used:
     Class = 1 (SESSION)
     C-Type = RSVP-AGGREGATE-IPv6-PCN-CL (to be replaced by IANA)

```
   +------------+------------+------------+------------+
   |                                                   |
   +                                                   +
   |                                                   |
   +      IPv6 PCN-ingress-node Address (16 bytes)     +
   |                                                   |
   +                                                   +
   |                                                   |
   +------------+------------+------------+------------+
   |                                                   |
   +                                                   +
   |                                                   |
   +      IPv6 PCN-egress-node Address (16 bytes)      +
   |                                                   |
   +                                                   +
   |                                                   |
   +------------+------------+------------+------------+
   |        rate of not marked PCN-traffic (NM-rate)   |
   +------------+------------+------------+------------+
   |  rate of threshold-marked PCN-traffic (ThM-rate)  |
   +------------+------------+------------+------------+
   |  rate of excess-traffic-marked PCN-traffic (ETM-rate) |
   +------------+------------+------------+------------+
```

   The fields carried by the PCN object are specified in
   [RFC6663], [RFC6661] and [RFC6662]:

     o the IPv4 or IPv6 address of the PCN-ingress-node and the IPv4
       or IPv6 address of the PCN-egress-node; together they specify the
       ingress-egress-aggregate to which the report refers. According to
       [RFC6663] the report should carry the identifier of the PCN-

ingress-node and the identifier of the PCN-egress-node (typically
their IP addresses);

o rate of not-marked PCN-traffic (NM-rate) in octets/second; its
format is a 32-bit IEEE floating point number;

        o rate of PCN-marked traffic (PM-rate) in octets/second; its format
          is a 32-bit IEEE floating point number;

        o rate of threshold-marked PCN traffic (ThM-rate) in
          octets/second; its format is a 32-bit IEEE floating point number;

        o rate of excess-traffic-marked traffic (ETM-rate) in
          octets/second; its format is a 32-bit IEEE floating point number;

     o) RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs: Controlled (CL) PCN CL Flow IDs
        object, IPv4 addresses are used:
         Class = 1 (SESSION)
         C-Type = RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs (to be replaced by IANA)


```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
                                            +-+-+-+-+-+-+-+-+-+
                                            | Length          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                      Source Address                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Destination Address                         |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |          Source Port          |       Destination Port        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Protocol     |    Reserved                                   |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    //                                                             //
    +                                                               +
    |                      Source Address                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Destination Address                         |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |          Source Port          |       Destination Port        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Protocol     |    Reserved                                   |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

        o) Length (1 byte): the length of the
            RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs object in units of 16 bytes.
            This field is used to specify the number of IPv4 flow IDs
            carried by this object. Each flow ID is represented by the
            combination of each subsequent 5 tuple:
            Source address, Destination address, Source Port,
            Destination Port and Protocol number.
            If Length is 0 then the RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs is
            empty.

o) Source address (4 bytes): The IPv4 source address.

o) Destination address (4 bytes): The IPv4 destination address.

        o) Protocol (1 byte): The IP protocol number. It refers to the
            true upper layer protocol carried by the packets.

        o) Source Port (2 bytes): contains the source port number.

        o) Destination Port (2 bytes): contains the destination port
            number.

    o) RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs: Controlled (CL) PCN CL Flow IDs
        object, IPv6 addresses are used:
        Class = 1 (SESSION)
        C-Type = RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs (To be replaced by IANA)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
                                            +-+-+-+-+-+-+-+-+
                                            | Length        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 |                    Source Address                             |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 |                 Destination Address                           |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          Source Port          |       Destination Port        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  Protocol    |    Reserved                                    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 //                                                             //
 +                                                               +
 |                                                               |
 |                    Source Address                             |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 |                 Destination Address                           |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          Source Port          |       Destination Port        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  Protocol    |    Reserved                                    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

o) Length (1 byte): the length of the
   RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs object in units of 40 bytes.
   This field is used to specify the number of flow IDs carried by
   this object. Each flow ID is represented by the combination of
   each subsequent 5 tuple fields:

      Source address, Destination address, Source Port,
      Destination Port and Protocol number.
      If Length is 0 then the RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs object
      is empty.

   o) Source address (16 bytes): The IPv6 source address.

   o) Destination address (16 bytes): The IPv6 destination address.

   o) Protocol (1 byte): The IP protocol number. It refers to the
       true upper layer protocol carried by the packets.

   o) Source Port (2 bytes): contains the source port number.

   o) Destination Port (2 bytes): contains the destination port
      number.

5. Security Considerations

   The same security considerations specified in [RFC2205], [RFC4230],
   [RFC4860], [RFC5559] and [RFC6411].

6. IANA Considerations

   This document makes the following requests to the IANA.
   IANA needs to modify the RSVP parameters registry, 'Class Names,
   Class Numbers, and Class Types' subregistry, and assigned 6 new C-
   Types under the existing SESSION Class (Class number 1), as described
   Below, see Section 4.1:

   Class
   Number Class Name   Reference
   ------ -----------------------   ---------

      1 SESSION   [RFC2205]

         Class Types or C-Types:

          19?   RSVP-AGGREGATE-IPv4-PCN-SM   this document
          20?   RSVP-AGGREGATE-IPv6-PCN-SM   this document
          21?   RSVP-AGGREGATE-IPv4-PCN-CL   this document
          22?   RSVP-AGGREGATE-IPv6-PCN-CL   this document
          23?   RSVP-AGGREGATE-IPv4-PCN-CL-FLIDs this document
          24?   RSVP-AGGREGATE-IPv6-PCN-CL-FLIDs this document

7. Acknowledgments

   We would like to thank the authors of [draft-lefaucheur-rsvp-ecn-
   01.txt], since some ideas used in this document are based on the work

initiated in [draft-lefaucheur-rsvp-ecn-01.txt]. Moreover, we would
like to thank Bob Briscoe, David Black, Ken Carlberg, Tom Taylor,
Philip Eardley, Michael Menth, Toby Moncaster, Francois Le Faucheur,
James Polk and Lixia Zhang for the provided comments.

8. Normative References

   [RFC6661] T. Taylor, A, Charny, F. Huang,
   G. Karagiannis, M. Menth, "PCN Boundary Node Behaviour for the
   Controlled Load (CL) Mode of Operation", July
   2012.

   [RFC6662] A. Charny, J. Zhang,
   G. Karagiannis, M. Menth, T. Taylor, "PCN Boundary Node Behaviour
   for the Single Marking (SM) Mode of Operation",
   July 2012.

   [RFC6663] G. Karagiannis, T. Taylor,
   K. Chan, M. Menth, P. Eardley, " Requirements for Signaling of (Pre-)
   Congestion Information in a DiffServ Domain",
   July 2012.

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
    Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC2205] Braden, R., ed., et al., "Resource ReSerVation Protocol
   (RSVP)- Functional Specification", RFC 2205, September 1997.

   [RFC3140] Black, D., Brim, S., Carpenter, B., and F. Le
   Faucheur, "Per Hop Behavior Identification Codes",
   RFC 3140, June 2001.

   [RFC3175] Baker, F., Iturralde, C., Le Faucheur, F., and B. Davie,
   "Aggregation of RSVP for IPv4 and IPv6 Reservations", RFC 3175,
   September 2001.

   [RFC4495] Polk, J. and S. Dhesikan, "A Resource Reservation
   Protocol (RSVP) Extension for the Reduction of
   Bandwidth of a Reservation Flow", RFC 4495, May 2006.

   [RFC4860] F. Le Faucheur, B. Davie, P. Bose, C. Christou, M.
   Davenport, "Generic Aggregate Resource ReSerVation Protocol (RSVP)
   Reservations", RFC4860, May 2007.

   [RFC5670] Eardley, P., "Metering and Marking Behaviour of PCN-Nodes",
    RFC 5670, November 2009.

   [RFC6660] Moncaster, T., Briscoe, B., and M. Menth, "Baseline
    Encoding and Transport of Pre-Congestion Information", RFC 6660,
    July 2012.

9. Informative References

   [draft-lefaucheur-rsvp-ecn-01.txt] Le Faucheur, F., Charny, A.,

Briscoe, B., Eardley, P., Chan, K., and J. Babiarz, "RSVP Extensions
for Admission Control over Diffserv using Pre-congestion
Notification (PCN) (Work in progress)", June 2006.

[RFC1633]   Braden, R., Clark, D., and S. Shenker, "Integrated
Services in the Internet Architecture: an Overview", RFC 1633, June
1994.

[RFC2211] J. Wroclawski, Specification of the Controlled-Load Network
Element Service, September 1997

[RFC2212] S. Shenker et al., Specification of Guaranteed Quality of
Service, September 1997

[RFC2474]   Nichols, K., Blake, S., Baker, F., and D. Black,
"Definition of the Differentiated Services Field (DS Field) in the
IPv4 and IPv6 Headers", RFC 2474, December 1998.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and
W. Weiss, "A framework for Differentiated Services", RFC 2475,
December 1998.

[RFC2998] Bernet, Y., Yavatkar, R., Ford, P., Baker, F., Zhang, L.,
Speer, M., Braden, R., Davie, B., Wroclawski, J. and E. Felstaine, "A
Framework for Integrated Services Operation Over DiffServ Networks",
RFC 2998, November 2000.

[RFC4230] H. Tschofenig, R. Graveman, "RSVP Security Properties",
RFC 4230, December 2005.

[RFC5559]   Eardley, P., "Pre-Congestion Notification (PCN)
Architecture", RFC 5559, June 2009.

[RFC6411] M. Behringer, F. Le Faucheur, B. Weis, "Applicability of
Keying Methods for RSVP Security", RFC 6411, October 2011.

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

[RFC2753] Yavatkar, R., D. Pendarakis and R. Guerin, "A Framework for
Policy-based Admission Control", January 2000.

10. Appendix A: Example Signaling Flow

This appendix is based on the appendix provided in [RFC4860]. In
particular, it provides an example signaling flow of the
specification detailed in Section 3 and 4. This signaling flow
assumes an environment where E2E reservations are aggregated over
generic aggregate RSVP reservations and applied over a PCN domain. In
particular the Aggregator (PCN-ingress-node) and Deaggregator (PCN-
egress-node) are located at the boundaries of the PCN domain. The
PCN-interior-nodes are located within the PCN-domain, between the
PCN-boundary nodes, but are not shown in this Figure. 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.

```
      Aggregator (PCN-ingress-node)     Deaggregator (PCN-egress-node)

      E2E Path
    ---------->
                (1)
                            E2E Path
                ------------------------------->
                                                (2)
                  E2E PathErr(New-agg-needed,SOI=GAx)
                <---------------------------------
                  E2E PathErr(New-agg-needed,SOI=GAy)
                <---------------------------------
                (3)
                        AggPath(Session=GAx)
                ------------------------------->
                        AggPath(Session=GAy)
                ------------------------------->
                                                (4)
                                                  E2E Path
                                                ---------->
                                                (5)
                        AggResv (Session=GAx) (PCN object)
                <------------------------------
                        AggResv (Session=GAy) (PCN object)
                <------------------------------
                (6)
                    AggResvConfirm (Session=GAx)
                ------------------------------->
                    AggResvConfirm (Session=GAy)
                ------------------------------->
                                                (7)
                                                  E2E Resv
                                                <---------
                                                (8)
                            E2E Resv (SOI=GAx)
                <-----------------------------
                (9)
        E2E Resv
    <----------
```
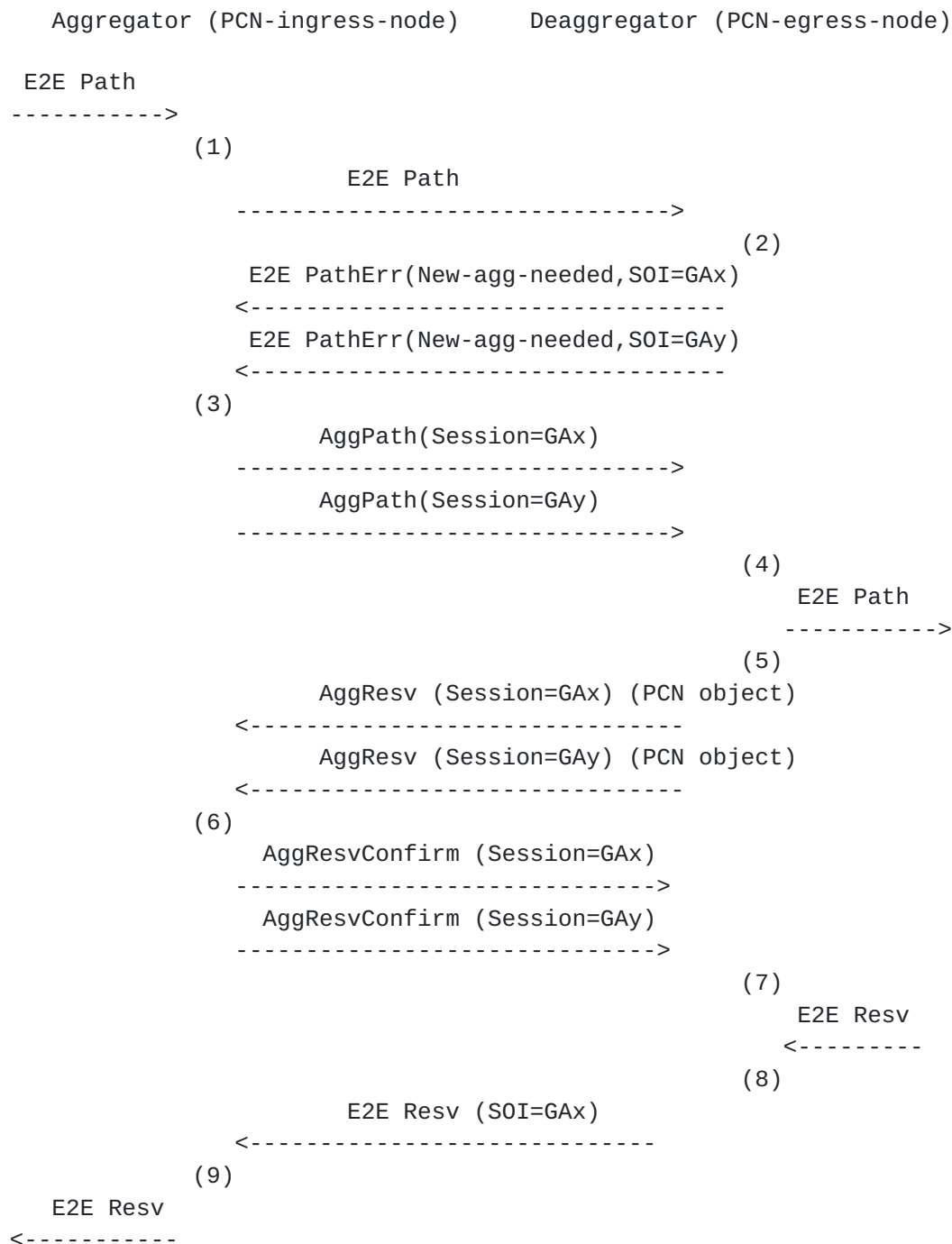
(1) The Aggregator (PCN-ingress-node) maps the e2e RSVP reservation
    session associated with the e2e Path message onto an PCN ingress-
    egress-aggregate. The Aggregator forwards e2e Path into the
    aggregation region after modifying its IP protocol number to
      RSVP-E2E-IGNORE. Note that in this case it is considered that the
      PCN-admission-state is "admit", see [Section 3.1](#). Otherwise, the
      e2e Path will not be forwarded into the aggregation region and
      the steps associated with the PCN-admission-state is "block"

situation specified in [Section 3.1](#) will be followed.

(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.
    At this step the Deaggregator maps the new generated RSVP generic
    aggregated reservations onto one ingress-egress-aggregate
    maintained by the Deaggregator (PCN-egress-node), see Section
    3.3. Moreover, the Deaggregator, depending on the used
    PCN edge behaviour and IP version, it includes one of the
    following PCN objects specified in Section 4.1:
    RSVP-AGGREGATE-IPv4-PCN-SM, RSVP-AGGREGATE-IPv6-PCN-SM,
    RSVP-AGGREGATE-IPv4-PCN-CL or RSVP-AGGREGATE-IPv6-PCN-CL.
    Here it is also Assumed 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.
    Moreover, the PCN-ingress-node functionality uses the PCN object
    to map the RSVP generic aggregated reservation state onto the
    maintained PCN ingress-egress-aggregate state. Moreover, the
    Aggregator performs the steps specified in Section 3.11.

(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.

11. Authors' Address

Georgios Karagiannis
University of Twente
P.O. Box 217
7500 AE Enschede,
The Netherlands
EMail: g.karagiannis@utwente.nl

Anurag Bhargava
Cisco Systems, Inc.
7100-9 Kit Creek Road
PO Box 14987
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709-4987
USA
Email: anuragb@cisco.com