A QoS Model for Signaling IntServ Controlled-Load Service with NSIS
draft-kappler-nsis-qosmodel-controlledload-08

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

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

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

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

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on February 22, 2009.

Abstract

This document describes a QoS Model to signal IntServ controlled load service with QoS NSLP. QoS NSLP is QoS Model agnostic. All QoS Model specific information is carried in an opaque object, the QSPEC. This document hence specifies the QSPEC for controlled load service, how the QSPEC must be processed in QoS NSLP nodes, and how QoS NSLP messages must be used.
Table of Contents

1. Introduction ........................................... 3
2. Terminology ............................................ 3
3. Signaling with QoS NSLP ................................. 4
   3.1. QoS NSLP ........................................ 4
   3.2. QSPEC ........................................... 4
   3.3. QoS Model ....................................... 5
4. IntServ Controlled Load Service ......................... 5
5. NSIS QoS Model for IntServ Controlled Load Service .... 6
   5.1. Role of QNEs ..................................... 7
   5.2. QSPEC Definition ................................ 7
       5.2.1. Controlled Load Service Requirements ....... 7
       5.2.2. QSPEC Objects ............................. 8
   5.3. Usage of QoS NSLP Messages -- QSPEC Procedures .... 9
6. Processing Rules in QNEs ................................ 10
   6.1. Admission Control ................................ 10
6.2. Packet Scheduling and Excess Treatment ............... 11
7. Preemption ............................................. 12
8. Interoperation with Controlled Load Service Specified in
   RFC2211 .............................................. 12
9. Security Considerations ................................ 13
10. IANA Considerations .................................. 13
11. Conclusions .......................................... 13
12. References ........................................... 15
   12.1. Normative References ............................ 15
   12.2. Informative References .......................... 15
Appendix A. Bit level Examples of QSPEC objects for
   Controlled Load QOSM ................................ 16
   A.1. Minimal QSPEC objects for Sender-Initiated Reservation .... 16
   A.2. Extended QSPEC objects for Sender-Initiated Reservation .... 17
   A.3. Receiver Initiated Reservation (RSVP Style) ............ 19
   A.4. Resource Queries .................................. 22
Appendix B. Change Tracker ................................ 22
   B.1. Changes in -08 ................................... 22
   B.2. Changes in -07 ................................... 22
   B.3. Changes in -06 ................................... 22
   B.4. Changes in -05 ................................... 23
   B.5. Changes in -04 ................................... 23
   B.6. Changes in -03 ................................... 23
   B.7. Changes in -02 ................................... 23
   B.8. Changes in -01 ................................... 23
Appendix C. Acknowledgements ................................ 23
Authors’ Addresses ........................................ 23
Intellectual Property and Copyright Statements ............. 25
1. Introduction

The QoS NSIS Signaling Layer Protocol, QoS NSLP [I-D.ietf-nsis-qos-nslp] defines how to signal for QoS reservations in the Internet. The protocol is not bound to a specific mechanism for achieving QoS, such as IntServ or DiffServ. Rather, the actual QoS information is carried opaquely in the protocol in a separate object, the QSPEC [I-D.ietf-nsis-qos-nslp]. A method for achieving QoS for a traffic flow is called QoS model. It is expected that a number of QoS models will be developed for QoS NSLP. Examples are [I-D.ietf-nsis-rmd] and [I-D.ietf-nsis-y1541-qosm] and this draft.

The purpose of this document is to describe a QoS model for controlled-load service of IntServ [RFC2211]. In [RFC2210] it is specified how to signal for controlled-load service with RSVP. This document describes how to signal for the same service with QoS NSLP.

The controlled-load service is rather minimal both in terms of information that is signaled - basically bandwidth in the form of a token bucket - and in terms of prescribed realization of the service in the network. It is therefore suited for a wide range of realizations, such as reserving resources per-flow per-network node [RFC1633], achieving QoS in appropriately engineered DiffServ networks with admission control [RFC2998], or across IP tunnels or MPLS Label Switched Paths (LSPs) with reserved bandwidths and admission control [RFC2746] [RFC3031].

The document is structured as follows: It gives a brief overview of QoS NSLP and the QSPEC, and the content and features of a QoS model as described in [I-D.ietf-nsis-qos-nslp] and [I-D.ietf-nsis-qspec]. It then gives a brief overview of the controlled-load service of IntServ. Subsequently, the actual QoS model for controlled-load service is described. A section describing the interoperation of QoS NSLP and RSVP, both for signaling controlled-load service, is also provided.

2. Terminology

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 [RFC2119].

The terminology defined in [I-D.ietf-nsis-qos-nslp] and [I-D.ietf-nsis-qspec] applies to this document.
3. Signaling with QoS NSLP

3.1. QoS NSLP

QoS NSLP [I-D.ietf-nsis-qos-nslp] is an NSIS signaling layer protocol for signaling QoS reservations in the Internet. Together with GIST [I-D.ietf-nsis-ntlp], it provides functionality similar to RSVP and extends it, e.g. by supporting both sender-initiated and receiver-initiated reservations. QoS NSLP however does not support multicast. QoS NSLP establishes and maintains reservation state in QoS NSLP aware nodes, called QNEs, along the path of a data flow. The number or frequency of QNEs is not prescribed. The node initiating a reservation request is called QNI, the node terminating the request is called QNR. QNI and QNR are also QNEs, and are not necessarily the actual sender and receiver of the data flow they are signaling for as they may also be proxying for them.

QoS NSLP defines four message types, RESERVE, QUERY, RESPONSE and NOTIFY. The message type identifies whether a message manipulates state (e.g. RESERVE) or not (e.g. QUERY, RESPONSE). The RESERVE message is used to create, refresh, modify or remove reservation state in QNEs. The QUERY message is used to request information about the data path without making a reservation. This functionality may be used to ‘probe’ the path for certain characteristics. The RESPONSE message is used to provide information about the results of a previous RESERVE or QUERY message, e.g. confirmation of a successful reservation, error, or for transferring results of a QUERY back towards the querying node. A NOTIFY message is sent asynchronously and need not refer to any previously received message. The information conveyed by a NOTIFY message is typically related to error conditions.

3.2. QSPEC

QoS NSLP carries QoS Model specific information encapsulated in an opaque object, the QSPEC [I-D.ietf-nsis-qspec]. The QSPEC thus fulfills a similar purpose as TSpec, RSpec and AdSpec in RSVP [RFC2205]. The QSPEC is not interpreted by the QoS NSLP Processing unit on a QNE, but passed as-is to the Resource Management Function RMF, usually located on the same node, where it is interpreted.

The QSPEC is composed of QSPEC objects, namely <QoS Desired>, <QoS Available>, <QoS Reserved> and <Minimum QoS>. A QSPEC typically only contains a subset of these objects. QSPEC objects contain a set of QSPEC parameters that govern the processing of the resource request in the RMF, e.g. information on excess treatment.
<QoS Desired> contains parameters describing the QoS desired by a QNI.

<QoS Available> contains parameters describing the available resources. In the controlled load service QOSM, this QSPEC object is used to collect information on the available bandwidth along a path.

<QoS Reserved> describes the actual QoS reserved.

Minimum QoS can be included by a QNI together with QoS Desired to signal a range of QoS (between QoS Desired and Minimum QoS) is acceptable.

The QSPEC template [I-D.ietf-nsis-qspec] defines a number of QSPEC parameters. <TMOD1> provides a description of the traffic for which resources are reserved. This parameter MUST be interpreted by each QNE along the path. All other QSPEC parameters MAY be signaled by the QNI if they are applicable to the underlying QoS desired. The QNI sets the M-Flag if they must be interpreted by downstream QNEs. If the parameter cannot be interpreted by a QNE the reservation fails. A QSPEC parameter without set M-Flag should be interpreted by the QNE but may be ignored if it cannot be interpreted. In a given QoS Model, new optional parameters may be defined.

3.3. QoS Model

A QoS-enabled domain supports a particular QoS model (QOSM), which is a method to achieve QoS for a traffic flow, such as IntServ Controlled Load or DiffServ [RFC2475]. QoS NSLP is independent of the QOSM, just as RSVP [RFC2205] is independent of IntServ. A QOSM hence incorporates QoS provisioning methods and a QoS architecture. It however also defines how to use QoS NSLP. It therefore defines the behavior of the resource management function (RMF), including inputs and outputs, and how QSPEC information on traffic description, resources required, resources available, and control information required by the RMF is interpreted. A QOSM also specifies the QSPEC parameters that describe the QoS and how resources will be managed by the RMF.

4. IntServ Controlled Load Service

As specified in [RFC2211], the controlled-load service defined for IntServ supports applications which are highly sensitive to overload conditions, e.g. real-time applications. The controlled-load service provides to an application approximately the end-to-end service of an unloaded best-effort network. "Unloaded" thereby is used in the sense of "not heavily loaded or congested" rather than in the sense of "no other network traffic whatsoever".
The definition of controlled-load service is intentionally imprecise. It implies a very high percentage of transmitted packets will be successfully delivered to the end nodes. Furthermore, the transit delay experienced by a very high percentage of the delivered packets will not greatly exceed the minimum transmit delay experienced by any successfully delivered packet. In other words, a short disruption of the service is viewed as a statistical effect which may occur in normal operation. Events of longer duration are indicative of failure to allocate sufficient resources to the controlled-load flow.

In order to ensure that the conditions on controlled-load service are met, clients requesting the service provide network elements on the data path with an estimation of the data traffic they are going to generate. When signaling with RSVP, the object carrying this estimation is called TSpec. In return, the service ensures that in each network element on the data path, resources adequate to process traffic falling within this descriptive envelope will be available to the client. This must be accomplished by admission control.

The controlled-load service is implemented per-flow in each network element on the data path. Thereby, a network element may be an individual node such as a router. However, a network element can also be a subnet, e.g. a DiffServ cloud within a larger IntServ network [RFC2998]. In this case, the per-flow traffic description (e.g. carried in the RSVP TSpec) together with the DiffServ Code Point (carried e.g. in the DCLASS object [RFC2996] of RSVP) is used for admission control into the DiffServ cloud. The DiffServ cloud MUST ensure it provides controlled-load service. It is also possible to operate controlled-load service over logical links such as IP tunnels [RFC2746] or MPLS LSPs [RFC3031]. The per-flow traffic descriptor is in this case used for admission control into the tunnel/LSP.

5. NSIS QoS Model for IntServ Controlled Load Service

According to [I-D.ietf-nsis-qspec], a QOSM MUST include the following information:
- role of QNEs, e.g., location, frequency, statefulness, etc.
- QSPEC definition including QSPEC parameters
- QSPEC procedures applicable to this QOSM
- QNE processing rules describing how QSPEC information is treated and interpreted in the RMF, e.g., admission control, scheduling, policy control, QoS parameter accumulation (e.g., delay).
- at least one bit-level QSPEC example
- QSPEC parameter behavior for new QSPEC parameters the QOSM specification defines
A QOSM specification MAY include the following information:
- define additional QOSM-specific error codes
- can state which QoS NSLP options a QOSM wants to use, when several options are available for a QOSM (e.g., local QSPEC to either a) hide initiator QSPEC within a local domain message, or b) encapsulate initiator QSPEC).

Subsequent sections treat these points one-by-one. An example bit-level QSPEC format is given in Appendix A.

5.1. Role of QNEs

Controlled-load service network elements can be individual routers or subnets. I.e. it is not necessary for each network node on the data path to interpret the signaling for the service. Rather, dedicated nodes MAY interpret signaling information and take on responsibility that the subnet they represent delivers adequate service. In fact, this setting maps nicely onto QoS NSLP – and the NSIS protocol suite in general. In NSIS, QNEs are just required to be located on the data path. However there are no prescriptions regarding their number or frequency. Hence, in the controlled-load QoS model, there MUST be (at least) one QNE acting on behalf of every network element. For example all ingress routers to a DiffServ cloud could be QNEs, performing admission control. If there is more than one network element per QNE, they MUST be coordinated among to ensure they delivers controlled-load service. Controlled Load QNEs are always stateful.

5.2. QSPEC Definition

5.2.1. Controlled Load Service Requirements

The controlled-load service QOSM uses TMOD1 parameters[I-D.ietf-nsis-qspec], which consist of a token bucket specification (i.e. bucket rate r and a bucket depth b) plus a peak rate (p) and a minimum policed unit (m). The minimum policed unit m is an integer measured in bytes. All IP datagrams of size less than m are counted against the token bucket as being of size m. For more details, including value ranges of the parameters see [RFC2210].

Note the TMOD1 parameter does not contain a maximum transmission unit (MTU), as the original token bucket does. When using RSVP to signal for controlled-load services, the PATH message collects information on MTU and available bandwidth which is used by the receiver to adapt the reservation parameters in the RESV message [RFC2210][RFC2215]. It is hence related to the signaling for Controlled Load rather than to the Controlled Load Service itself. Indeed, while collecting path MTU can be useful for achieving QoS, it is not considered to be part
of QoS signaling or QOSMs [I-D.ietf-nsis-qspec] in NSIS; rather, an independent path MTU discovery mechanism (e.g., [pmtud-charter]) during the flow setup phase is assumed to provide means to learn about the path MTU.

Available bandwidth may be collected if desired and used for controlled load service QOSM. The controlled-load service has no required characterization parameters the QNI needs to be informed about, i.e. current measurement and monitoring information need not be exported by QNEs, although individual implementations may do so if they wish.

5.2.2. QSPEC Objects

The QSPEC can contain some or all of the following objects:

<QoS Desired> = <TMOD1>

<QoS Available> = <TMOD1>

<Minimum QoS> = <TMOD1>

<QoS Reserved> = <TMOD1>

Among them, <QoS Desired>, <QoS Available> and <QoS Reserved> MUST be supported by all QOSM implementations, as defined in [I-D.ietf-nsis-qspec].

<QoS Available> is required for receiver-initiated reservations case 2 and 3, and case 2 and 3 in sender-initiated reservations. It is used for gathering available bandwidth information along the path. This information can be used by the QNI (or QNR, for receiver-initiated reservations) to make an appropriate reservation thereafter, particularly to re-issue a failed reservation. Since only bandwidth is needed, set the <TMOD1> parameters \( r = \text{peak rate} = p, b = \text{large}, m = \text{large} \) and for TCP traffic, \( r = \text{average rate}, b = \text{large}, p = \text{large} \).

<QoS Available> MAY be supported by QNEs and is optional to implement and use. If supported it always travels together with <QoS Desired>. It signifies that the QNI can accept a downgrade of resources for particular parameters in the reservation, down to the value of the respective parameter in <Minimum QoS>. For parameters not appearing in <Minimum QoS>, it cannot accept a downgrade. For controlled load service this means if <Minimum QoS> is included, a downgrade of all TMOD1 parameters is possible.

Furthermore, the Excess Treatment parameter MAY be included as
parameter. Currently supported values are "reshape" or "drop". The
default value for the Controlled Load QoSM if not included is
"reshape". This parameter is used for a controlled load service
implementation to handle the received data traffic belonging to a
controlled load flow which is "non-conformant" to the TMOD1
specification reserved. Traffic is considered "non-conformant" when:
over time period T, the amount of data received exceeds rT+b; or
data rate of the traffic exceeds the peak rate p; or
data packet size is larger than M or the QNE’s outgoing link MTU

In all QSPEC objects additional parameters MAY be included, as
described in [I-D.ietf-nsis-qspec].

5.3. Usage of QoS NSLP Messages -- QSPEC Procedures

QoS NSLP allows a variety of message sequences for reserving
resources ("QSPEC Procedures"). Particularly, sender-initiated,
receiver-initiated and bi-directional messages are possible. E.g.,
in sender-initiated reservations, a RESERVE is issued by the QNI. If
the reservation is successful, the QNR replies with a RESPONSE. If
the reservation fails, the QNE at which it failed sends an INFO_SPEC
object indicating this failure towards the QNI.

The QSPEC template defines what QSPEC objects are carried in which of
these messages, and how they are translated from message-to-message.
For each of the message patterns defined in QoS NSLP, a variety of
QSPEC object usages, the so-called QSPEC Procedures, are possible.
- In the simplest message sequence, sender-initiated reservations,
  the RESERVE may carry just <QoS Desired> to indicate the exact QoS
  it wants, and the corresponding RESPONSE carries solely <QoS
  Reserved>. This implies either the exact resources described in
  <QoS Desired> are reserved, or the reservation fails.
- A more advanced QNI would include, in addition to <QoS Desired>, a
  <QoS Available> QSPEC object, or even a <Minimum QoS>. <QoS
  Available> allows collecting path properties, e.g. currently
  available TMOD1, and <Minimum QoS> signals that (and how much)
  less resources than <QoS Desired> are acceptable. The RESPONSE
  message carries <QoS Reserved>, and additionally copies the <QoS
  Available> QSPEC Object from RESERVE. This information may be of
  particular interest if a reservation failed. Note however, that
  since the QNE failing the reservation sends the RESPONSE, no
  complete end-to-end information on e.g. bandwidth can be collected
  and delivered to the QNI.
- In an "RSVP-style" receiver-initiated reservation, the sender
  (QNR) issues a QUERY with <QoS Desired> specifying the desired
  resources and <QoS Available> collecting information on available
  TMOD1 parameters. The receiver (QNI) reacts with a RESERVE
  message containing <QoS Desired> with a TMOD1. <QoS Available> is
copied from the QUERY message. The signaling exchange is concluded with a RESPONSE by the QNR including a <QoS Reserved> echoing the TMOD1 that was reserved.

Note that the initial message triggering the signaling exchange fully determines the sequencing of subsequent messages, and also determines what QSPEC objects will be carried in them. That is, only the QNI for sender initiated reservation and the QNR for receiver initiated reservation have freedom in choosing a particular QSPEC procedure. Other QNEs can only react to this.

The controlled load service parameters can be signaled with any QSPEC procedure. Note, in contrast, in RSVP only one type of message exchange is defined (receiver-initiated reservations, and the equivalent of <Minimum QoS> = 0). However, this is a characteristic of RSVP rather than of the controlled load service.

6. Processing Rules in QNEs

6.1. Admission Control

For controlled-load service, QNEs are required to perform admission control. All resources important to the operation of the network element MUST be considered when admitting a request. Common examples of such resources include link bandwidth, router or switch port buffer space, and computational capacity of the packet forwarding engine. It is not prescribed how a QNE determines adequate resources are available. It is however required that they make bandwidth greater than the token rate available to the flow in certain situations in order to account for fluctuations. E.g. statistical methods may be used to determine how much bandwidth is necessary.

During the admission control, the controlled-load service TMOD1 parameters MUST be met according to the following rule: a TMOD1 A from the available resources for a flow MUST be "as good or better than" or "greater than or equal to" TMOD1 B (which is carried in the received QoS Description, e.g., <QoS Desired>, or <Minimum QoS> if available), i.e.,:

- the TMOD1 rate r for TMOD1 A is greater than or equal to that of TMOD1 B, and
- the TMOD1 depth b for TMOD1 A is greater than or equal to that of TMOD1 B, and
- the peak rate p for TMOD1 A is greater than or equal to that of TMOD1 B, and
- the minimum policed unit m for TMOD1 A is less than or equal to that of TMOD1 B, and
Remark: these rules come originally from rules for ordering TMOD1s in [RFC2211].

There are no target values for other parameters, e.g. delay or loss, other than providing a service closely equivalent to that provided to best-effort traffic under lightly loaded conditions.

Resource requests for new flows are accepted if capacity is available. Reservation modifications are accepted if the new <TMOD1> is strictly smaller than the old one. Otherwise they are treated like new reservations from an admission control perspective.

6.2. Packet Scheduling and Excess Treatment

A QNE MUST ensure the TMOD1 requirements for any individual flow given at setup time are met locally. That is, traffic MUST obey the rule that over all time periods, the amount of data sent does not exceed rT+b. Packets smaller than m are counted as of size m. A basic requirement for packet scheduling is that the QNE MUST ensure the QoS requirements are met for traffic belonging to flows whose traffic are all conformant.

In presence of arriving non-conformant traffic, the QNE MUST behave as follows:
- the QNE MUST continue to provide the contracted QoS for traffic belonging to flows which are all conformant.
- the QNE SHOULD prevent excess control load traffic from unfairly impacting the handling of arriving best-effort traffic.
- While fulfilling the above two requirements, the QNE MUST attempt to forward the excess traffic on a best-effort basis if sufficient resources are available, unless indicated differently by <Excess Treatment>.

Several basic approaches for excess treatment are suggested in [RFC2211] and reused here, although other alternatives are possible, if available. A simple approach is the priority mechanism, namely, to let the QNE process excess controlled-load traffic at a lower priority than the elastic best-effort traffic, especially when the most controlled-load traffic arises from non-rate-adaptive real-time applications.

The second approach is that a QNE can maintain separate flow classes (e.g., one for each non-conformant controlled-load traffic, one for inelastic best-effort flows, and another from elastic best-effort flows), where packet scheduling mechanisms like Fair Queueing or Weight Fair Queueing can be used. One implementation, for instance, could allocate each controlled-load flow a 1/N "fair share" percentage of the available best effort bandwidth for its excess...
traffic.

Finally, Random Early Detection (RED) queueing mechanism may be used.

7. Preemption

The controlled-load service QOSM makes use of the <Preemption Priority> and <Defending Priority> parameter to preempt flows. A new flow with higher <Preemption Priority> can preempt an already admitted flow with lower <Defending Priority>. One single higher priority flow can replace more than one lower priority flow until the requested amount of <TMOD1> is met. Once the flow is admitted <Preemption Priority> becomes irrelevant. More details about Preemption and Defending Priority are provided in [RFC3181].

8. Interoperation with Controlled Load Service Specified in RFC2211

The controlled-load service QOSM is intended to be consistent with the RSVP/Controlled Load Service specified in [RFC2211], although the signaling protocols used are QoS NSLP and RSVP, respectively. This section discusses how a router implementing both RSVP and QoS NSLP could translate from one to the other.

The following is a table that contains a mapping of messages, objects and parameters between QoS NSLP and RSVP for the specific case of controlled-load signaling using the "RSVP-style" receiver-initiated signaling described in Appendix A.3.

<table>
<thead>
<tr>
<th>Message</th>
<th>Object(s)</th>
<th>Parameter(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSVP</td>
<td>Path</td>
<td>Sender TSpec</td>
</tr>
<tr>
<td></td>
<td>ADSpec</td>
<td>avail. bw and MTU</td>
</tr>
<tr>
<td>QoS NSLP</td>
<td>QUERY</td>
<td>&lt;QoS Desired&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;QoS Available&gt;</td>
</tr>
<tr>
<td>RSVP</td>
<td>Resv</td>
<td>FlowSpec</td>
</tr>
<tr>
<td>QoS NSLP</td>
<td>RESERVE</td>
<td>&lt;QoS Desired&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;QoS Available&gt;</td>
</tr>
<tr>
<td>RSVP</td>
<td>ResvConfirm</td>
<td></td>
</tr>
<tr>
<td>QoS NSLP</td>
<td>RESPONSE</td>
<td>&lt;QoS Reserved&gt;</td>
</tr>
</tbody>
</table>
Please note that TMod1 in <QoS Available> specifies bandwidth only. See Section Section 5.2.2 how to set bandwidth in TMod.

A RSVP Path Message includes a SenderTSPEC specifying the traffic an application is going to send. The RSVP AdSpec is optionally included. It probes for the available bandwidth on the data path. This reservation model is referred to as One Pass with Advertising (OPWA). When the AdSpec is omitted, the Receiver cannot determine the available resources for the resulting end-to-end QoS. This reservation model is referred to as One Pass. On arrival at the Receiver the FlowSpec, consisting of the TSPEC, is created. The TSPEC is usually derived from the SenderTSPEC and if available from the AdSpec. It contains the desired QoS. The Resv message is sent upstream to the Sender. At each hop the desired resource reservation is reserved. The last node sends a ResvConf back to the Receiver to indicate that end-to-end reservation has been installed.

In QoS NSLP, the sender populates the QoS which it desires by including a <QoS Desired> and optionally queries the network for the QoS that is available. In this case it carries the <QoS Available> parameter which is updated by all QNEs to reflect the QoS that is actually available. With the <QoS Desired> object, the Receiver (QNI) is informed about the requested resources. See Section 5.3 for a detailed description of QSPEC procedure for controlled-load service.

Note that under controlled-load QOSM, there is no MTU discovery as in RSVP/CLS, where path MTU is a mandatory parameter. This relieves the QNE from being overloaded with the orthogonal task of determining path MTU.

9. Security Considerations

This document does not raise additional security issues beyond those described in [I-D.ietf-nsis-qos-nslp].

10. IANA Considerations

A new QOSM ID ("Controlled-Load Service QOSM") needs to be assigned by IANA.

11. Conclusions

This document describes a QoS Model to signal IntServ controlled load service with QoS NSLP. Up to now, it was only described how to
signal for IntServ controlled load service with RSVP. Since no independent document exists that describes IntServ controlled load by its own, i.e. without RSVP, it is sometimes difficult to determine what features are specific to IntServ controlled load, and which features are specific to RSVP.

The QoS NSLP QOSM for controlled load service allows a variety of message exchanges all eventually resulting in a reservation, e.g. sender-initiated, receiver-initiated and bidirectional signaling. The controlled load service when signaled with RSVP was bound to receiver-initiated reservations.

When signaling with RSVP, it is not possible to define a range of acceptable QoS. Also this seems to be a characteristic of RSVP rather than a feature of the controlled load service.

RSVP allows discovery of path MTU. Since independent mechanisms area exist to this end, this feature has not been reproduced by the Controlled Load QOSM (and QoS NSLP in general)

An issue of general interest discovered here concerns feedback of information in sender-initiated scenarios (In receiver-initiated scenarios it does not occur because path information is collected before the RESERVE is issued). A QNI may include in <QoS Available> several parameters, e.g. bandwidth, which it would like to measure along the data path. If the reservation fails, e.g. because the desired bandwidth was too large, the QNE failing the reservation returns a RESPONSE, including the <QoS Available> QSPEC object with accumulated information up to this point. The QNI can learn from this why the reservation failed at this particular QNE. However it cannot be sure a subsequent downgraded RESERVE will be more successful. This is because there may be even more difficult conditions (e.g. even less bandwidth) down the path. That is, in sender-initiated scenarios it is not straightforward to receive feedback from a failed reservation that allows to make a good guess at what size of reservation would be more successful. Of course it would be possible for the QNI to issue a QUERY first to find out about a suitable value for, e.g. maximum packet size. However this adds another round-trip time to the reservation, thereby obsoleting one of the main benefits of sender-initiated reservations compared to receiver-initiated ones.

In this draft, the feedback problem is solved by including a <Minimum QoS> QSPEC object in sender-initiated reservations. This gives some flexibility as it implicitly says the QNI would also accept a downgraded reservation, up to the value specified. Note however as currently specified in [I-D.ietf-nsis-qspec], the <Minimum QoS> QSPEC object is not necessarily supported by all QNEs.
12. References

12.1. Normative References


12.2. Informative References


Appendix A.  Bit level Examples of QSPEC objects for Controlled Load QOSM

A.1.  Minimal QSPEC objects for Sender-Initiated Reservation

The first example shows a "minimal" QSPEC for Controlled Load containing the least number of objects and parameters. It signals for sender initiated reservations, containing the TMOD1 for <QoS Desired> and for <QoS Reserved>. The difference between the QSPEC in the RESERVE and the RESPONSE message is only slight.
A.2.  Extended QSPEC objects for Sender-Initiated Reservation

The following QSPEC offers a range of acceptable bandwidth in case the request of <QoS Desired> cannot be fulfilled. When <QoS Available> becomes lower than <Minimum QoS> the reservation fails. The requesting node gets informed by <QoS Available> about the remaining resources. See [I-D.ietf-nsis-qspec] for details. The optional <Excess Treatment> parameter defines the behavior of the traffic conditioner how to handle out of profile traffic.
<table>
<thead>
<tr>
<th>Vers. QType=I</th>
<th>QSPEC Proc.=0/3</th>
<th>I</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>R</th>
<th>Length = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>Type = 0 (QoS Des.)</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
<td>E</td>
<td>0</td>
<td>r</td>
<td>ID = 1 &lt;TMOD-1&gt;</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>TMOD Rate-1 [r] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>TMOD Size-1 [b] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Peak Data Rate-1 [p] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Minimum Policed Unit-1 [m] (32-bit unsigned integer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>M</td>
<td>E</td>
<td>0</td>
<td>r</td>
<td>ID = 11 &lt;Excess Tr.&gt;</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Excess Trtmnt</td>
<td>Remark Value</td>
<td>Reserved</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>E</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>Type = 1 (QoS Avail.)</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
<td>E</td>
<td>0</td>
<td>r</td>
<td>ID = 1 &lt;TMOD-1&gt;</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>TMOD Rate-1 [r] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>TMOD Size-1 [b] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Peak Data Rate-1 [p] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Minimum Policed Unit-1 [m] (32-bit unsigned integer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>E</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>Type = 3 (Min. QoS)</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
<td>E</td>
<td>0</td>
<td>r</td>
<td>ID = 1 &lt;TMOD-1&gt;</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>TMOD Rate-1 [r] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>TMOD Size-1 [b] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Peak Data Rate-1 [p] (32-bit IEEE floating point number)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Minimum Policed Unit-1 [m] (32-bit unsigned integer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
</tr>
</tbody>
</table>

Fig.3 Example QSPEC for Sender-Initiated Reservation(RESERVE)
A.3. Receiver Initiated Reservation (RSVP Style)

This is an example for an ‘RSVP-style’ reservation using a 3-way handshake. The QNR as the sender issues a QUERY and informs the QNI at the receiver about the desired bandwidth. The requested resources are contained in <QoS Desired>. Resource information about the path is collected in <QoS Available>. The receiver copies the content of <QoS Available> into <QoS Desired>. The QNI is updated about available resources before sending the RESERVE. A RESPONSE is sent back to confirm the reservation.
Fig. 5 Example QSPEC for Receiver-Initiated Reservation (QUERY)
Fig. 6 Example QSPEC for Receiver-Initiated Reservation (RESERVE)

Fig. 7 Example QSPEC for Receiver-Initiated Reservation (RESPONSE)
A.4. Resource Queries

The QUERY message is used to collect information about available bandwidth along the path. It does not manipulate any state. In response to the <QoS Desired> a <QoS Available> object describing the resources is returned.

```
+-----------------+-----------------+-----------------+-----------------+-----------------+-
|     Version     |   QSPEC Type    |      2         |      1         |       I        |
|-----------------+-----------------+-----------------+-----------------+-----------------+-
|E                |                 |                 |                 | Reserved       |
+-----------------+-----------------+-----------------+-----------------+-----------------+-
| Type = 1 (QoS Avail.) | r | r | r | r | Length = 5 |
+-----------------+-----------------+-----------------+-----------------+-----------------+-
| ID = 1 <TMOD-1> | r | r | r | r | Length = 4 |
+-----------------+-----------------+-----------------+-----------------+-----------------+-
| TMOD Rate-1 [r] (32-bit IEEE floating point number) |
+-----------------+-----------------+-----------------+-----------------+-----------------+-
| TMOD Size-1 [b] (32-bit IEEE floating point number) |
+-----------------+-----------------+-----------------+-----------------+-----------------+-
| Peak Data Rate-1 [p] (32-bit IEEE floating point number) |
+-----------------+-----------------+-----------------+-----------------+-----------------+-
| Minimum Policed Unit-1 [m] (32-bit unsigned integer) |
+-----------------+-----------------+-----------------+-----------------+-----------------+
```

Fig. 8 Example QSPEC for Resource Queries (QUERY and RESPONSE)

Other scenarios can be easily derived by adapting to the QoS NSLP signaling procedure and used QoS specifications.

Appendix B. Change Tracker

B.1. Changes in -08

1. Removed discussion about number of hops not implementing the CLS QoSM in section Conclusions

B.2. Changes in -07

1. Editorial changes made

B.3. Changes in -06

1. Updated requirements for QOSM specification
2. Clarified the use of <Minimum QoS>
3. Added section about preemption
B.4. Changes in -05

1. Included additional bit-level examples.

2. Updated section about interoperation with RSVP-CLS.

B.5. Changes in -04

1. Adapted terminology and content to latest version of QSPEC (v17).
   E.g. removed QOSM ID, removed MTU,...

B.6. Changes in -03

1. Adapted terminology and updated references.

B.7. Changes in -02

1. Added "RSVP-style reservation" as running example

2. Updated interoperability section

3. Aligned QSPEC example in Appendix A with update of QSPEC draft and added more details

B.8. Changes in -01

1. Clarifications about path MTU, scheduling/excess treatment and QOSM Hops.

2. Added a section "Interoperation with RFC2211" and QSPEC format as Appendix.

Appendix C. Acknowledgements

The authors would like to thank Andrew McDonald for fruitful discussions. John Loughney, Bob Braden and Hannes Tschofenig provided helpful comments.
Authors’ Addresses

Cornelia Kappler
deZem GmbH
Knesebeckstr. 86/87
Berlin 10623
Germany

Email: cornelia.kappler@googlemail.com

Xiaoming Fu
University of Goettingen
Institute for Informatics
Lotzestr. 16-18
Goettingen 37083
Germany

Email: fu@cs.uni-goettingen.de

Bernd Schloer
University of Goettingen
Institute for Informatics
Lotzestr. 16-18
Goettingen 37083
Germany

Email: bschloer@cs.uni-goettingen.de
Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.