

QoS Signaling Across Heterogeneous Wired/Wireless Networks: Resource Management in Diffserv Using the NSIS Protocol Suite

Attila Báder¹, Georgios Karagiannis², Lars Westberg³, Cornelia Kappler⁴,
Tom Phelan⁵, Hannes Tschofenig⁶, Geert Heijenk²

¹ Ericsson Research,

P.O. Box 107, Budapest, Hungary, H-1300

attila.bader@ericsson.com

² University of Twente,

P.O. Box 217, 7500 AE Enschede, the Netherlands

{g.karagiannis, geert.heijenk}@utwente.nl

³ Ericsson Research,

Torshamnsgatan 23, SE-16480, Stockholm, Sweden

lars.westberg@ericsson.com

⁴ Siemens AG,

Siemensdamm 62, Berlin 13627, Germany

cornelia.kappler@siemens.com

⁵ Sonus Networks,

250 Apollo Dr., Chelmsford, MA 01824, USA

tphelan@sonusnet.com

⁶ Siemens AG

Otto-Hahn-Ring 6, Munich 81739, Germany

hannes.tschofenig@siemens.com

Abstract

Reservation-based Quality of Service (QoS) in a mixed wireless and wireline environment requires an end-to-end signaling protocol that is capable of adapting to the idiosyncrasies of the different networks. The QoS NSIS Signaling Protocol (QoS-NSLP) has been created by the Next Steps In Signaling working group at the IETF to fulfill this need for an adaptive reservation protocol. It allows reservation requests to be interpreted by equipment implementing different QoS models along the path between a data sender and a data receiver. This paper describes the QoS-NSLP, and an example of a particular QoS model that is based on Resource Management in Diffserv (RMD). RMD provides a scalable dynamic resource management method for Diffserv networks. RMD has two basic functions to control the traffic load in a Diffserv domain: it provides admission control for flows entering the network and it has an algorithm that terminates the required amount of flows in case of congestion caused by failures (e.g. link or router)

within a Diffserv domain. The admission control within the domain can be either measurement- or reservation-based. The basic signaling mechanism is described for different signaling scenarios and the expected performance of the protocol is discussed.

1. Introduction

The disparate technologies of wireless and wireline networking make resource reservations across heterogeneous networks difficult. Some resource description parameters, such as bit-rate, are relatively easy to understand in almost all environments. Other parameters tend to be special to one environment and do not make sense in the others. For example, parameters used in wireless networks, such as bit error rate (BER), link-level retransmission, and error correction do not translate easily into the wireline environment.

Furthermore, in a mixed environment, wireless networks tend to be access networks with fairly limited

bandwidth and require per-flow reservations. On the other hand, the wireline networks tend to form the backbones and have relatively abundant bandwidth and carry a large number of flows, where aggregation is necessary since per-flow reservations suffer from scalability constraints.

The Integrated Services (IntServ) model [1], and its signaling protocol, Resource ReSerVation Protocol (RSVP) [2], created by the Internet Engineering Task Force (IETF), is a one-size-fits-all answer for end-to-end reservations. End systems request an end-to-end reservation for each flow, if desired. Intermediary nodes interpret the reservation requests and implement the resource guarantees according to the IntServ algorithms.

This per-flow orientation of Intserv/RSVP has led to problems in nodes that support a large number of flows, and to several attempts to improve the scaling characteristics through flow aggregation. One attempt is the Differentiated Services (Diffserv) model [3]. Diffserv allows the aggregation of flows with similar QoS requirements into Per Hop Behavior (PHB) groups. Nodes at the edges of Diffserv domains perform this aggregation by marking incoming packets with appropriate Diffserv Code Points (DSCPs). Interior nodes only need to recognize the relatively small set of DCSPs and implement the corresponding PHBs.

The designers of Diffserv have not defined a signaling protocol for it. QoS reservations are made by out-of-band Service Level Agreements (SLAs). In addition to aggregating flows, Diffserv Edge Nodes police incoming packets for conformance to the appropriate SLA.

Both, IntServ/RSVP and Diffserv, have problems when they are applied to a mixed wireless and wireline situation. IntServ's network-layer QoS mechanisms are sometimes unsuitable for, or need extensions in wireless networks. Intserv's per-flow reservation semantics make the scale of wired core networks difficult. Diffserv's mechanisms are mostly suited to core networks, and do not fit as well for access networks. Plus, Diffserv's reliance on static SLAs can create inefficiencies in resource-scarce access networks.

To advance the state of reservation protocols, the IETF created the Next Steps In Signaling (NSIS) working group. NSIS provides a model for the network entities taking part in control signaling and for the relationship between signaling and the rest of the network operation. The NSIS protocol suite is decomposed into a generic (lower) layer named NSIS Transport Layer Protocol (NTLP), and a separate upper layer, which is different for each signaling application, known as NSIS Signaling Layer Protocol (NSLP).

The NTLP primarily consists of a messaging layer which is used to transport the signaling application layer messages upstream and downstream along the data path between NSIS peer-to-peer basis. The NTLP runs over standard transport and security protocols such as UDP, TCP, SCTP, and DCCP.

The NSIS working group is currently working on a QoS signaling application (QoS-NSLP) [5] and an application for configuring Network Address Translators and Firewalls (NATFW-NSLP) [6]. Other NSLPs have been proposed, for example an NSLP for configuration of metering entities [7]. The QoS-NSLP protocol establishes and maintains reservation states at nodes along the path of a data flow for providing forwarding resources for that flow. The design of the QoS NSLP is conceptually similar to RSVP [2], and uses soft-state peer-to-peer refresh messages as the primary state management mechanism. However, the QoS NSLP supports sender- and receiver-initiated reservations, bi-directional reservations and more flexible deployment scenarios (for example, signaling between edge-to-edge or end-to-edge). NSIS does not aim to provide a signaling solution for multicast traffic. Providing a solution for IP multicast has often been seen as one of the reasons for the complexity of RSVP.

The QoS-NSLP separates the signaling protocol from the Resource Management Function (RMF). Therefore, the actual parameters describing the QoS reservation are encapsulated in an object, the QSPEC [8], which is only interpreted by the RMF. Each QoS-NSLP node along the path is expected to implement a QoS Model (QOSM) appropriate to its underlying technology. The QSPEC parameters are interpreted and implemented according to this QOSM. There might be a number of different QOSMs implemented along the path of the reservation. The system initiating a reservation describes its needs with a set of generic parameters that must be understood by all QOSMs. The initiator may also include QOSM-specific parameters that provide additional data for certain QOSMs along the path.

With QoS-NSLP, an end system can make generic end-to-end resource reservations across a mixed wireless and wireline network, and have those reservations sensibly implemented at each point in the network. Furthermore, if the application requires some special treatment by some network technology along the path (say the wireless access network in use), it can include technology-specific requests.

One of the QOSMs that can be used is the Resource Management in Diffserv (RMD) QoS Model (RMD-QOSM). The RMD concept was developed to provide dynamic resource reservation within a Diffserv domain in a scalable way [9, 10]. Admission control at the edges of the Diffserv domain is based on the feedback

of a load control mechanism, which checks the resource availability in all nodes in the communication path between the edges of the Diffserv domain. The routers between the edge nodes (referred as interior routers) apply a simple resource management function based on per-PHB traffic aggregation.

In this paper, the QoS-NSLP protocol is described through the example of RMD-QOSM. In particular, Section 2 describes the protocol model used by the NSIS protocol suite and the basic signaling principles in an RMD-QOSM domain. Section 3 introduces the resource management methods used in the interior routers, Section 4 describes the basic signaling operation examples and Section 5 presents the performance evaluation of an important feature of the protocol, i.e., severe congestion handling. Finally, Section 6 concludes.

2. Protocol model of NSIS and basic signaling principles in an RMD domain

RMD can be applied within a local domain of an NSIS aware network, as shown in Figure 1. The NSIS nodes outside the RMD domain implement QoS models suitable to their technologies, e.g., a wireless specific QOSM. The two signaling end points are the QoS-NSLP Initiator (QNI) and the QoS NSLP Receiver (QNR). RMD operates between the QoS NSIS Forwarder (QNF) Ingress node and the QNF Egress node.

One of the most important characteristics of RMD signaling and resource management is that QNF Interior nodes do not store and use per-flow states. RMD requires per-flow NTLP and QoS-NSLP states in QNF edge nodes only. QNF Interior nodes are either QoS-NSLP stateless (in the case of measurement-based operation), or they are reduced-state nodes storing per PHB aggregated QoS-NSLP states (in the case of reservation-based operation). NTLP message association and routing states do not need to be stored in QNF Interior nodes.

As a consequence of the stateless/reduced state operation, only sender-initiated reservation can be performed in the RMD domain. QNF Interior routers send signaling messages always in the direction of the data receiver only. QNF edge nodes aggregate received reservations for QNF Interior nodes, sending RESPONSE and NOTIFY messages back towards the QNF Initiator.

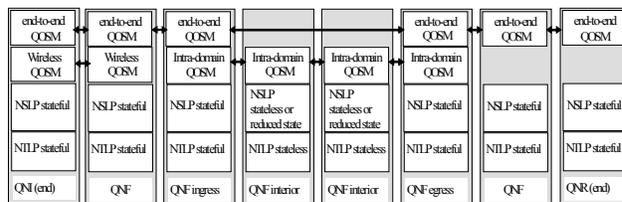


Figure 1. Protocol model of RMD operation in an NSIS network.

The RMD domain supports two signaling sessions for each flow, namely, one end-to-end session and one intra-domain session that are bound together at the QNF edges. Each session is identified by a globally unique session identifier. The binding between the two sessions indicates a (possibly asymmetric) dependency relation between these sessions by using a bound session identifier. The QNF Ingress can use this information e.g. to tear down an intra-domain session when the end-to-end session is terminated.

3. Resource Management in the QNF Interior nodes

The RMD is based on resource units (e.g. transmission rate or bandwidth) and service differentiation. The number of resource units that is assigned to a particular flow and the resource unit used within an RMD domain are determined by a local policy configured in the QNF Ingress. The resource units can reflect e.g. a peak rate, an average rate or an effective rate supporting statistical multiplexing at packet level. The resource management with the RMD aware domain can be:

Reservation-based: In this method the resource reservation is accomplished by adding and removing resource units. For each flow or aggregate a number of resource units is reserved by the RMF. For the established reservation state within the local RMD domain the soft state principle is applied (as it is done with all QoS-NSLP states). State must periodically be refreshed. The refresh messages include the number of resource units that should be refreshed for given flows. If a refresh message is not sent for a flow, the corresponding number of resource units will be removed after the time-out period from the reservation state in the QNF Interior nodes along the data path. Resource units can be removed by explicit release message as well. In this case the resources are removed immediately and it is not necessary to wait until the time-out period.

Measurement-based method: This method does not require the use of refresh messages since no

reservation states are installed and maintained. The RMF is responsible for measuring and periodically updating the available resources for each PHB. The frequency of the update is based on a local policy. The admission decision is based on the comparison of the available and requested resources. It is beneficial to keep very limited state information per PHB, regarding the amount of resources for which requests have been granted, but which are not yet accounted for in the measurements, since the actual flow has not started yet.

Measurement-based admission control (MBAC) algorithms have been studied extensively in literature. These range from fairly simple algorithms, taking just the average measured load into account (e.g., the Measured Sum algorithm [11]), to more complicated algorithms, taking also higher order statistics into account (e.g., Time Scale Decomposition (TSD) [12]). The latter periodically estimates the mean and variance of the carried traffic using measurements. The admission decision is positive if the currently carried traffic, as characterized by the measured statistics, plus the requested resources for the new flow do exceed the system capacity with a probability smaller than some α . Otherwise, the admission decision is negative. This algorithm can be extended using the refresh and explicit release signaling procedures and by keeping very limited state information per PHB. This allows for separation of the measured variance into variance due to flow arrival/departure and variance due to traffic variability, and can result in higher utilization. Also, information regarding the number of flows generating the measured traffic can help interior nodes to increase their utilization.

4. Basic Signaling Operation Examples

When integrating RMD into the NSIS protocol suite four functional scenarios are supported [13]:

- Unidirectional normal operation within the RMD aware NSIS domain describing successful and unsuccessful resource reservation (in the reservation-based method) or resource query (in the measurement-based method).
- Unidirectional severe congestion handling within the RMD aware domain.
- Bi-directional normal operation within the RMD aware domain that describes the successful and unsuccessful resource reservation or resource query procedure.
- Bi-directional severe congestion handling within the RMD aware domain.

A unidirectional resource reservation/resource query refers to a procedure that is initiated end-to-end by the QNI and is propagated only in one direction

towards the QNR. This procedure can be either be sender- or receiver-initiated. A bi-directional reservation/query is composed by binding two unidirectional reservations/queries using bound session identifier.

The basic RMD-QOSM signaling is shown in Figure 2. A RESERVE message is created by a QNI with an Initiator QSPEC describing the reservation and forwarded along the path towards the QNR. When the original RESERVE message arrives at the ingress node, an RMD-QSPEC is constructed based on the top-most QSPEC in the message (usually the Initiator QSPEC). The RMD-QSPEC is sent in a local, independent RESERVE message through the interior nodes towards the QNR. This local RESERVE message uses the NTLP hop-by-hop datagram signaling mechanism. Meanwhile, the original RESERVE message is sent to the egress node on the path to the QNR using the reliable transport mode of NTLP.

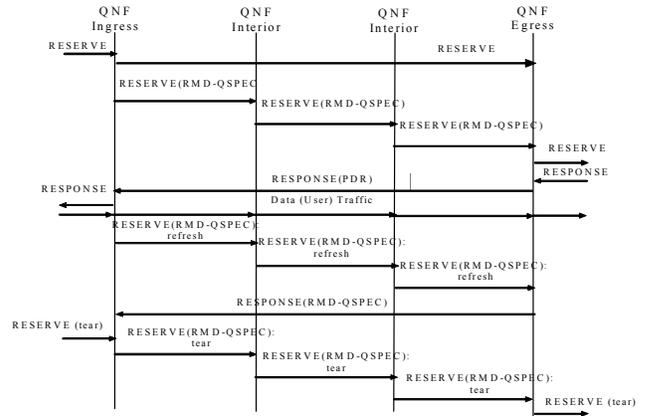


Figure 2. RMD functional operation for successful reservation.

Each QoS NSLP node on the data path processes the local RESERVE message and checks the availability of resources with either the reservation-based or the measurement-based method (see Section 3). When the message reaches the egress node, and the reservation is successful in each interior node, the original end-to-end RESERVE message is forwarded to the next domain. When the egress node receives a RESPONSE message from the downstream end, it is forwarded directly to the ingress node.

If an intermediate node cannot accommodate the new request, it indicates this by marking a single bit in the message, and continues forwarding the message until the egress node is reached. From the egress node a RESPONSE message is sent directly the ingress node.

5. Performance evaluation of severe congestion handling

When a failure along the end-to-end communication path occurs, for example a router failure or a link failure, the routing algorithms will adapt to failures by changing the routing decisions to reflect changes in the topology and traffic volume. Since the traffic admission decisions were based on the old routes, the re-routed traffic may create overloads along the new paths. This may cause severe congestion along some paths.

The QoS-NSLP and RMD are able to cope with congested situations using the refresh procedure. If the refresh is not successful in a QNF Interior node, QNF edge nodes are notified by “S” marking the refresh messages and the flows that cannot be supported are terminated.

In general, relying the soft state refresh mechanism solves the congestion within the time frame of the refresh period. To speed up the response to this situation, the RMD QOSM provides a “proportional marking” method, where the number of the remarked packets is proportional to the detected overload. Note that this severe congestion method can be used for the situations when either the RMD resource reservation procedure or the RMD resource query is supported within the RMD aware domain.

In this method the QNF Interior node experiencing the severe congestion marks data packets passing through the node in proportion to the degree of congestion (using at least one additional Diffsev codepoint). The QNF Egress node applies a predefined policy based on the number of the marked bytes to solve the severe congestion, by selecting a number of end-to-end flows that should be terminated. For these flows, the QNF Egress node generates and sends a NOTIFY (PDR) message upstream towards the QNF Ingress node (its upstream stateful QoS-NSLP peer) to indicate the severe congestion in the communication path.

Upon receiving this message, the QNF Ingress node resolves the severe congestion by a predefined policy, e.g., refusing new incoming reservation requests, terminating affected reservations or shifting them to an alternative RMD traffic class PHB.

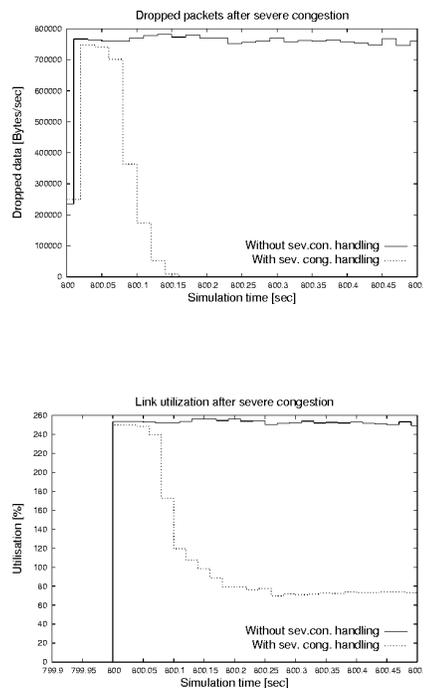


Figure 3 Severe congestion handling using data packet marking.

The severe congestion handling algorithm has been evaluated using the NS2 simulation environment. The evaluation results showed that the normal operating conditions can be restored within a few hundred ms, depending on the network size and traffic conditions. Figure 3 shows the NS2 simulation results where the amount of dropped packets and the link utilization in a congested link was measured in the function of the simulation time. The bottom figure shows a theoretical curve comparing the actual load in a large link to a preconfigured limit that corresponds to a real link capacity or a preconfigured limit. The simulations were done for a typical 3G access network, see [14], in which voice traffic is dominant. The link capacities were between 2-4 Mbps. Flows were simulated with ON-OFF traffic sources with 60% activity, 50 byte packet sizes, and 20 ms inter-packet time periods. Flows were generated by a Poisson call arrival process. After a link goes down in the access network, the flows are re-routed to a new path at 800 seconds simulation time. The new path, however, has not enough free capacity, therefore it gets congested and experiences packet drops. The results show that the normal operation is restored within 200 ms by using the RMD severe congestion function based on data marking.

A more detailed analysis of the transient behavior and description of the severe congestion handling algorithm can be found in [14].

6. Conclusions and future work

End-to-end QoS signaling protocols, such as those developed within the IETF, can significantly enhance the quality perceived by the end users and at the same time increase the network utilization. In a mixed wireless and wireline environment, reservation-based Quality of Service (QoS) requires a signaling protocol that is capable of adapting to the idiosyncrasies of the different networks. The QoS NSIS Signaling Layer Protocol (QoS-NSLP) has been defined by the Next Steps In Signaling (NSIS) working group at the IETF to fulfill this need. The flexibility of NSIS makes it possible to use different QoS models in different parts of the network.

In this paper the QoS-NSLP protocol is described through the example of a particular QoS model, based on the Resource Management in Diffserv. In particular, the basic signaling mechanisms are described and the expected performance of the protocol for severe congestion handling is discussed.

The RMD QoS model can be applied for resource reservation in large-scale networks, transporting large numbers of real-time flows at the same time, or in mobile access networks where the setup time is critical. In particular, the expected processing time of QoS-NSLP messages in stateful QNF edge nodes is in the order of a few milliseconds in an unloaded system. The simple RMF in QNF Interior nodes make an implementation possible for which the message processing time is in the order of a few microseconds.

Though RMD is an aggregation method, the proposed severe congestion algorithm is able to give a prompt response to a link or node failure. In particular, the performance results show that the normal operation can be restored within 200 ms by using the RMD severe congestion function based on data marking.

Implementations of the QoS-NSLP and the RMD-QOSM protocol schemes are currently under development.

7. References

- [1] R. Braden, D. Clark, S. Shenker: Integrated services in the internet architecture: an overview. RFC 1633, IETF Informational RFC, 1994
- [2] R. Braden., L. Zhang, S. Berson, S. Herzog, S. Jamin: Resource reservation protocol (RSVP) functional specification”, RFC2205, IETF proposed standard RFC, 1997.
- [3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss: An Architecture for Differentiated Services. RFC2475, IETF Informational RFC, 1998.
- [5] S. Van den Bosch, G. Karagiannis, and A. McDonald: NSLP for Quality-of-Service signaling. Internet draft, work in progress, Feb. 2005
- [6] M. Stiernerling, H. Tschofenig, and C. Aoun: NAT/Firewall NSIS Signaling Layer Protocol (NSLP). Internet draft, work in progress, May 2005
- [7] F. Dressler, G. Carle, C. Fan, C. Kappler, and H. Tschofenig: NSLP for Metering Configuration Signaling. Internet draft, work in progress, Oct. 2004
- [8] J. Ash, A. Báder, and C. Kappler: QoS-NSLP QSPEC Template, Internet draft, work in progress, Feb. 2005
- [9] L. Westberg, A. Császár, G. Karagiannis, Á. Marquetant, D. Partain, O. Pop, V. Rexhepi, R. Szabó, A. Takács.: Resource Management in Diffserv (RMD): A Functionality and Performance Behavior Overview. IFIP PfHSN’02, 2002, Berlin
- [10] A. Császár and A. Takács: Comparative Performance Analysis of RSVP and RMD, Proceedings of the Fourth COST 263 International Workshop on Quality of Future Internet Services (QoFIS2003), Stockholm, Sweden, pp. 41--51, 2003. Oct., LNCS 2811, Springer Verlag
- [11] S. Jamin, S. Shenker, P. Danzig: Comparison of Measurement-based Admission Control Algorithms for Controlled-Load Service. Proceedings IEEE Infocom '97, Kobe, Japan, April 1997
- [12] M. Grossglauser, D.N.C. Tse: A Time-Scale Decomposition Approach to Measurement-Based Admission Control. IEEE/ACM Transactions on Networking, Vol. 11, No. 4, August 2003
- [13] A. Báder, L. Westberg, G. Karagiannis, C. Kappler, T. Phelan, H. Tschofenig: RMD-QOSM – The Resource Management in Diffserv QoS Model, Internet draft, work in progress, Feb. 2005
- [14] A. Császár, A. Takács, A. Báder: A Practical Method for Efficient Resolution of Congestion in an On-Path Reduced State Signaling Environment, accepted for IWQOS 2005