Benchmarking Working Group   Gabor Feher, BUTE
INTERNET-DRAFT   Krisztian Nemeth, BUTE
Expiration Date: August 2004   Andras Korn, BUTE
                                                Istvan Cselenyi, TeliaSonera

February 2004

Benchmarking Terminology for Routers Supporting Resource Reservation

<draft-ietf-bmwg-benchres-term-04.txt>

1. Status of this Memo

2. Table of contents

3. Abstract

The purpose of this document is to define terminology specific to the benchmarking of resource reservation signaling of IntServ IP routers. These terms can be used in additional documents that define benchmarking methodologies for routers that support resource reservation or reporting formats for the benchmarking measurements.

4. Introduction

Signaling based resource reservation (e.g. via RSVP [1]) is an important part of the different QoS provisioning approaches. Therefore network operators who are planning to deploy signaling based resource reservation may want to scrutinize the scalability limitations of reservation capable routers and the impact of signaling on their data forwarding performance.

An objective way of quantifying the scalability constraints of QoS signaling is to perform measurements on routers that are capable of resource reservation. This document defines terminology for a specific set of tests that vendors or network operators can carry out to measure and report the signaling performance characteristics of router devices that support resource reservation protocols. The results of these tests provide comparable data for different products, and thus support the decision process before purchase. Moreover, these measurements provide input characteristics for the dimensioning of a network in which resources are provisioned dynamically by signaling. Finally, the tests are applicable for characterizing the impact of the resource reservation signaling on the forwarding performance of the routers.

This benchmarking terminology document is based on the knowledge gained by examination of (and experimentation with) different resource reservation protocols: the IETF standard RSVP [1] and several experimental ones, such as YESSIR [5], ST2+ [6], SDP [7], Boomerang [8] and Ticket [9]. Some of these protocols are also analyzed in an IETF NSIS working group draft [10]. Although at the moment the authors are only aware of resource reservation capable router products that interpret RSVP, this document defines terms that are valid in general and not restricted to any of the above listed protocols.

In order to avoid any confusion we would like to emphasize that this terminology considers only signaling protocols that provide IntServ resource reservation; the DiffServ world, for example, is out of our scope.

5. Existing definitions

RFC 1242 "Benchmarking Terminology for Network Interconnect Devices" [2] and RFC 2285 "Benchmarking Terminology for LAN Switching Devices" [3] contain discussions and definitions for a number of terms relevant to the benchmarking of signaling performance of reservation capable routers and should be consulted before attempting to make use of this document.

Additionally, this document defines terminology in a way that is consistent with the terms used by Next Steps in Signaling working group laid out in [4].

For the sake of clarity and continuity this document adopts the template for definitions set out in Section 2 of RFC 1242. Definitions are indexed and grouped together into different sections for ease of reference.

6. Definition of Terms

6.1 Traffic Flow Types

This group of definitions describes traffic flow types forwarded by resource reservation capable routers.

6.1.1 Data Flow

Definition:
A data flow is a stream of data packets from one sender to one or more receivers, where each packet has a flow identifier unique to the flow.

Discussion:
The flow identifier can be an arbitrary subset of the packet header fields that uniquely distinguishes the flow from others. For example, the 5-tuple "source address; source port; destination address; destination port; protocol number" is commonly used for this purpose (where port number are applicable). For more comment

on flow identification refer to [4].

The flow identification can be time- and/or resource-consuming, but this can sometimes be avoided as routers do not necessarily have to classify each packet. Instead, packets that should be classified by routers can be marked with special flags that routers understand. One existing marking approach is to use the ToS field of the IP header. Naturally, unmarked packets are not classified by routers and this way valuable resources can be saved.

6.1.2 Distinguished Data Flow

Definition:
Distinguished data flows are flows that resource reservation capable routers intentionally treat better or worse than "ordinary" data flows, according to a QoS agreement defined for the distinguished flow.

Discussion:
Packets of distinguished data flows are marked so that the routers that forward them know they require differentiated treatment. Routers classify these incoming packets and identify the data flow they belong to.

The most common usage of the distinguished data flow is to get higher priority treatment than that of best-effort data flows (see the next definition). In these cases, a distinguished data flow is sometimes referred to as a "premium data flow". Nevertheless theoretically it is possible to require worse treatment than that of best-effort flows.

6.1.3 Best-Effort Data Flow

Definition:
Best-effort data flows are flows that are not treated in any special manner by resource reservation capable routers; thus, their packets are served (forwarded) in some default way.

Discussion:
"Best-effort" means that the router makes its best effort to forward the data packet quickly and safely, but does not guarantee anything (e.g. delay or loss probability). This type of traffic is the most common in today's Internet.

The packets belonging to the best-effort data flows are not specially marked and thus they are not classified by the routers.

6.2 Resource Reservation Protocol Basics

This group of definitions applies to signaling based resource reservation protocols implemented by IP router devices.

6.2.1 QoS Session

Definition:

A QoS session is an application layer concept, shared between a set of network nodes, that pertains to a specific set of data flows. The information associated with the session includes the data required to identify the set of data flows in addition to a specification of the QoS treatment they require.

Discussion:

A QoS session is an end-to-end relationship. Whenever end-nodes decide to obtain special QoS treatment for their data communication, they set up a QoS session; as part of the process, they or their proxies make a QoS agreement with the network, specifying their data flows and the QoS treatment that the flows require.

It is possible for the same QoS session to span multiple network domains that have different resource provisioning architectures. In this document, however, we only deal with the case where the QoS session is realized over an IntServ architecture. It is assumed that sessions will be established using signaling messages of a resource reservation protocol.

QoS sessions must have unique identifiers; it must be possible to determine which QoS session a given signaling message pertains to. Therefore, each signaling message should include the identifier of its corresponding session. As an example, in the case of RSVP, the "session specification" identifies the QoS session plus refers to the data flow; the "flowspec" specifies the desired QoS treatment and the "filter spec" defines the subset of data packets in the data flow that receive the QoS defined by the flowspec.

QoS sessions can be unicast or multicast depending on the number of participants. In a multicast group there can be several data traffic sources and destinations. Here the QoS agreement does not have to be the same for each branch of the multicast tree forwarding the data flow of the group. Instead, a dedicated network resource in a router can be shared among many traffic sources from the same multicast group (c.f. multicast reservation styles in the case of RSVP).

Issues:

Even though QoS sessions are considered to be unique, resource reservation capable routers might aggregate them and allocate network resources to these aggregated sessions at once. The aggregation can be based on similar data flow attributes (e.g.

similar destination addresses) or it can combine arbitrary sessions as well. While reservation aggregation significantly lightens the signaling processing task of a resource reservation capable router, it also requires the administration of the aggregated QoS sessions and might also lead to the violation of the quality guaranties referring to individual data flows within an aggregation [12].

6.2.2 Resource Reservation Protocol

Definition:
Resource reservation protocols define signaling messages and message processing rules used to control resource allocation in IntServ architectures.

Discussion:
It is the signaling messages of a resource reservation protocol that carry the information related to QoS sessions. This information includes a session identifier, the actual QoS parameters, and possibly flow descriptors.

The message processing rules of the signaling protocols ensure that signaling messages reach all network nodes concerned. Some resource reservation protocols (e.g. RSVP) are only concerned with this, i.e. carrying the QoS-related information to all the appropriate network nodes, without being aware of its content. This latter approach allows changing the way the QoS parameters are described, and different kind of provisioning can be realized without the need to change the protocol itself.

6.2.3 Resource Reservation Capable Router

Definition:
A router is resource reservation capable (it supports resource reservation) if it is able to interpret signaling messages of a resource reservation protocol, and based on these messages is able to adjust the management of its flow classifiers and network resources so as to conform with the content of the messages.

Discussion:
Routers capture signaling messages and manipulate reservation states and/or reserved network resources according to the content of the messages. This ensures that the flows are treated as their specified QoS requirements indicate.

6.2.4 Reservation State

Definition:
A reservation state is the set of entries in the routerÆs memory that contain all relevant information about a given QoS session

registered with the router.

Discussion:

States are needed because IntServ related resource reservation protocols require the routers to keep track of QoS session and data-flow-related metadata. The reservation state includes the parameters of the QoS treatment; the description of how and where to forward the incoming signaling messages; refresh timing information; etc.

Based on how reservation states are stored in a reservation capable router, the routers can be categorized into two classes:

Hard-state resource reservation protocols (e.g. ST2 [6]) require routers to store the reservation states permanently, established by a set-up signaling primitive, until the router is explicitly informed that the QoS session is canceled.

There are also soft-state resource reservation capable routers, where there are no permanent reservation states, and each state has to be regularly refreshed by appropriate refresh signaling messages. If no refresh signaling message arrives during a certain period then the router stops the maintenance of the QoS session assuming that the end-points do not intend to keep the session up any longer or the communication lines are broken somewhere along the data path. This feature makes soft-state resource reservation capable routers more robust than hard-state routers, since no failures can cause resources to stay permanently stuck in the routers. (Note, it is still possible to have an explicit teardown message in soft-state protocols for quicker resource release.)

Issues:

Based on the initiating point of the refresh messages, soft-state resource reservation protocols can be divided into two groups. First, there are protocols where it is the responsibility of the end-points or their proxies to initiate refresh messages. These messages are forwarded along the path of the data flow refreshing the corresponding reservation states in each router affected by the flow. Secondly, there are other protocols, where routers and end-points have their own schedule for the reservation state refreshes and they signal these refreshes to the neighboring routers.

6.2.5 Resource Reservation Protocol Orientation

Definition:

The orientation of a resource reservation protocol tells which end of the protocol communication initiates the allocation of the network resources. Thus, the protocol can be sender or receiver initiated, depending on the location of the data flow source (sender) and destination (receiver) compared to the reservation initiator.

Discussion:
In the case of sender-initiated protocols the resource reservation propagates the same directions as of the data flow. Consequently, in the case of receiver-initiated protocols the signaling messages reserving resources are forwarded backward on the path of the data flow. Due to the asymmetric routing nature of the Internet, in this latter case, the path of the desired data flow should be known before the reservation initiator would be able to send the resource allocation messages. For example in the case of RSVP, the RSVP PATH message, traveling from the data flow sources towards the destinations, first marks the path of the data flow on which the resource allocation messages will travel backward.

This definition considers only protocols that reserve resources for just one data flow between the end-nodes. The reservation orientation of protocols reserving more than one data flow is not defined here.

Issues:
The location of the reservation initiator affects the basics of the resource reservation protocols and therefore it is an important design decision. In the case of multicast QoS sessions, the sender-oriented protocols require the traffic sources to maintain a list of receivers and send their allocation messages considering the different requirements of the receivers. Using multicast QoS sessions, the receiver-oriented protocols give the chance to the receivers to manage their own resource allocation requests and thus ease the task of the sources.

6.3 Router Load Factors

The router load expressing the utilization of the device naturally affects the performance of the router. During the benchmarking process several load conditions have to be examined.

This group of definitions describes different load components that impact only a specific part of the resource reservation capable router.

6.3.1 Best-Effort Traffic Load Factor

Definition:
The best-effort traffic load factor is defined as the volume of the best-effort data traffic that traverses the router in a second.

Discussion:
Forwarding the best-effort data packets, which requires obtaining the routing information and transferring the data packet between

network interfaces, requires processing power, which is related to this load factor.

Issues:
The same amount of data segmented into differently sized packets causes different amounts of load on the router, which has to be considered during the benchmarking measurements.

Measurement unit:
bits per second (bps)

6.3.2 Distinguished Traffic Load Factor

Definition:
The distinguished traffic load factor is defined as the volume of the distinguished data traffic that traverses the router in a second.

Discussion:
Similarly to the best-effort data, forwarding the distinguished data packets requires obtaining the routing information and transferring the data packet between network interfaces. However, in this case packets have to be classified as well, which requires additional processing capacity.

Issues:
Just as in the best-effort case, the same amount of data segmented into differently sized packets causes different amounts of load on the router, which has to be considered during the benchmarking measurements.

Measurement unit:
bits per second (bps)

6.3.3 Session Load Factor

Definition:
The session load factor is the number of QoS sessions the router is keeping track of.

Discussion:
Resource reservation capable routers maintain reservation states keeping track of the QoS sessions. Obviously, the more reservation states are registered with the router, the more complex the traffic classification becomes, and the longer time it takes to look up the corresponding resource reservation sate. Moreover, not only the traffic flows, but also the signaling messages that control the reservation states have to be identified first, before taking any other action, and this kind of classification also means extra work for the router.

In the case of soft-state resource reservation protocols, the session load also affects reservation state maintenance. For example, the supervision of timers that watchdog the reservation state refreshes may cause further load on the router.

Measurement unit:
This factor is measured by the number of QoS sessions impacting the router, thus it has no unit.

6.3.4 Signaling Intensity Load Factor

Definition:
The signaling intensity load factor is defined as the number of signaling messages that hit the router during one second.

Discussion:
The processing of signaling messages requires processor power that raises the load on the control plane of the router.

In routers where the control plane and the data plane are not totally independent (e.g. certain parts of the tasks are served by the same processor; or the architecture has common memory buffers, transfer buses or any other resources) the signaling load can have an impact on the router's packet forwarding performance as well.

Naturally, just as everywhere else in this document, the term "signaling messages" refer only to the resource reservation protocol related primitives.

Issues:
Most of the resource reservation protocols have several protocol primitives realized by different signaling message types. Each of these message types may require a different amount of processing power from the router. This fact has to be considered during the benchmarking measurements.

Measurement unit:
The unit of this factor is 1/second.

6.3.5 Signaling Burst Load Factor

Definition:
The signaling burst load factor is defined as the number of signaling messages that arrive to one input port of the router back-to-back ([2]), causing persistent load on the signaling message handler.

Discussion:
The definition focuses on one input port only and does not consider the traffic arriving at the other input ports.
As a consequence, a set of messages arriving at different ports, but with such a timing that would be a burst if the messages arrived at the same port, is not considered to be a burst. The reason for this is that it is not guaranteed at a black-box test that this would have the same effect on the router as a burst

(incoming at the same interface) has.

This definition conforms to the burst definition given in [3].

Issues:
Most of the resource reservation protocols have several protocol primitives realized by different signaling message types. Bursts built up of different messages may have a different effect on the router. Consequently, during measurements the content of the burst has to be considered as well.

Measurement unit:
This load factor is measured by the number of messages in the burst, thus it has no unit.

6.4 Performance Metrics

This group of definitions is a collection of measurable quantities that describe the impact the different load components have on the router.

6.4.1 Signaling Message Handling Time

Definition:
The signaling message handling time (or, in short, signal handling time) is the latency ([2]) of a signaling message passing through the router.

Discussion:
The router interprets the signaling messages, acts based on their content and usually forwards them in an unmodified or modified form. Thus the message handling time is usually longer than the forwarding time of data packets of the same size.

There might be signaling message primitives, however, that are drained or generated by the router, like certain refresh messages. In this case the signal handling time is immeasurable, therefore it is not defined for such messages.

In the case of signaling messages that carry information pertaining to multicast flows, the router might issue multiple signaling messages after processing them. In this case, by definition, the signal handling time is the latency between the incoming signaling message and the last outgoing signaling message related to the received one.

The signal handling time is an important characteristic as it directly affects the setup time of a QoS session.

Issues:
The signal handling time may be dependent on the type of the signaling message. For example, it usually takes a shorter time for the router to remove a reservation state than to set it up.

This fact has to be considered during the benchmarking process.

Measurement unit:
The unit of the signaling message handling time is the second.

6.4.2 Distinguished Traffic Delay

Definition:
Distinguished traffic delay is the latency ([2]) of a distinguished data packet passing through the tested router device.

Discussion:
Distinguished traffic packets must be classified first in order to assign the network resources dedicated to the flow. The time of the classification is added to the usual forwarding time (including the queuing) that a router would spend on the packet without any resource reservation capability. This classification procedure might be quite time consuming in routers with vast amounts of reservation states.

There are routers where the processing power is shared between the control plane and the data plane. This means that the processing of signaling messages may have an impact on the data forwarding performance of the router. In this case the distinguished traffic delay metric also indicates the influence the two planes have on each other.

Issues:
Queuing of the incoming data packets in routers can bias this metric, so the measurement procedures have to consider this effect.

Measurement unit:
The unit of the distinguished traffic delay is the second.

6.4.3 Best-effort Traffic Delay

Definition:
Best-effort traffic delay is the latency of a best-effort data packet traversing the tested router device.

Discussion:
If the processing power of the router is shared between the control and data plane, then the processing of signaling messages may have an impact on the data forwarding performance of the router. In this case the best-effort traffic delay metric is an indicator of the influence the two planes have on each other.

Issues:
Queuing of the incoming data packets in routers can bias this metric as well, so measurement procedures have to consider this

effect.

Measurement unit:
The unit of the best-effort traffic delay is the second.

6.4.4 Signaling Message Loss

Definition:
Signaling message loss is the ratio of the actual and the expected number of signaling messages leaving a resource reservation capable router, subtracted from one.

Discussion:
This definition gives the same value as the ratio of the lost and the expected messages. The reason for choosing the given definition is that the number of lost messages cannot be measured directly.

There are certain types of signaling messages that are required to be forwarded by reservation capable routers as soon as their processing is finished. However, due to the high router load or for other reasons, the forwarding or even the processing of these signaling messages might be canceled. There are other kinds of signaling messages, that should have been generated by the router, without any corresponding incoming message. In case of high router load, it is possible that such a message never leaves the router. To characterize these situations we introduce the signaling message loss metric expressing the ratio of the signaling messages that actually have left the router and those ones that were expected to leave the router.

Since the most frequent reason for signaling message loss is high router load, this metric is suitable for sounding out the scalability limits of resource reservation capable routers.

During the measurements one must be able to determine whether a signaling message is still in the queues of the router or if it has already been dropped. For this reason we define a signaling message as lost if no forwarded signaling message is emitted within a reasonably long time period. This period is defined along with the benchmarking methodology.

Measurement unit:
This measure has no unit; it is expressed as a real number, which is between zero and one, including the limits.

6.4.5 Session Maintenance Capacity

Definition:
The session maintenance capacity metric is used in the case of soft-state resource reservation protocols only. It is defined as

the ratio of the number of QoS sessions actually maintained and the number of QoS sessions that should have been maintained during one refresh period.

Discussion:
For soft-state protocols maintaining a QoS session means refreshing the reservation states associated with it.

When a soft-state resource reservation capable router is overloaded, it may happen that the router is not able to refresh all the registered reservation states, because it does not have the time to run the state refresh task. In this case sooner or later some QoS sessions will be lost even if the endpoints still require their maintenance.

The session maintenance capacity sounds out the maximal number of QoS sessions that the router is capable of maintaining.

Issues:
The actual process of session maintenance is protocol and implementation dependent, so is the method to examine that a session is maintained or not.

In the case of soft-state resource reservation protocols a router that fails to maintain a QoS session may not emit refresh signaling messages either. This has direct consequences on the signaling message loss metric.

Measurement unit:
This measure has no unit; it is expressed as a real number, which is between zero and one (including the limits).

6.5 Scalability Limit

Definition:
The scalability limit of the router is the critical load condition, when the router is still in the steady state but the smallest amount of additional load would drive it to the overloaded state.

Discussion:
All existing routers have finite buffer memory and finite processing power. In the steady state of the router, the buffer memories are not fully utilized and the processing power is enough to cope with all tasks running on the router. As the router load increases the buffers are starting to fill up and/or the router has to postpone more and more tasks. However, there is a certain point where no more buffer memory is available, or a task cannot be postponed any longer; thus the router is forced to drop a packet or an activity. This is the overloaded state of the resource reservation capable router, which can be recognized by the fact that some kind of data (signaling or packet) or task (e.g. QoS session maintenance) loss occurs.

7. Security Considerations

As this document only provides terminology and describes neither a protocol nor an implementation or a procedure, there are no security considerations associated with it.

8. Acknowledgements

We would like to thank the following individuals for their help in the research and development work, which contributed to form this document: Joakim Bergkvist and Norbert Vegh from Telia Research AB, Sweden; Balazs Szabo, Gabor Kovacs and Peter Vary from the High Speed Networks Laboratory, Budapest University of Technology and Economics, Hungary.

9. References

[1] B. Braden, Ed., et. al., "Resource Reservation Protocol (RSVP) - Version 1 Functional Specification", RFC 2205, September 1997.

[2] S. Bradner, "Benchmarking Terminology for Network Interconnection Devices", RFC 1242, July 1991

[3] R. Mandeville, "Benchmarking Terminology for LAN Switching Devices", RFC 2285, February 1998

[4] R. Hancock (ed.), et. al., "Next Steps in Signaling: Framework", IETF draft: draft-ietf-nsis-fw-05.txt (work in progress), October 2003

[5] P. Pan, H. Schulzrinne, "YESSIR: A Simple Reservation Mechanism for the Internet", Computer Communication Review, on-line version, volume 29, number 2, April 1999

[6] L. Delgrossi, L. Berger, "Internet Stream Protocol Version 2 (ST2) Protocol Specification - Version ST2+", RFC 1819, August 1995

[7] P. White, J. Crowcroft, "A Case for Dynamic Sender-Initiated Reservation in the Internet", Journal on High Speed Networks, Special Issue on QoS Routing and Signaling, Vol. 7 No. 2, 1998

[8] J. Bergkvist, D. Ahlard, T. Engborg, K. Nemeth, G. Feher, I. Csel nyi, M. Maliosz, "Boomerang : A Simple Protocol for Resource Reservation in IP Networks", Vancouver, IEEE Real-Time Technology and Applications Symposium, June 1999

[9] A. Eriksson, C. Gehrmann, "Robust and Secure Light-weight Resource Reservation for Unicast IP Traffic", International WS

on QoS'98, IWQoS'98, May 18-20, 1998

[10] J. Manner (ed.), X. Fu, Ping Pan, "Analysis of Existing Quality of Service Signaling Protocols", IETF draft: draft-ietf-nsis-signalling-analysis-03.txt (work in progress), October 2003

[11] J. Wroclawski, "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997

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

Authors' Addresses

Gabor Feher
Budapest University of Technology and Economics
Department of Telecommunications and Mediainformatics
Magyar Tudosok krt. 2, H-1117, Budapest, Hungary
Phone: +36 1 463-1538
Email: Gabor.Feher@tmit.bme.hu

Krisztian Nemeth
Budapest University of Technology and Economics
Department of Telecommunications and Mediainformatics
Magyar Tudosok krt. 2, H-1117, Budapest, Hungary
Phone: +36 1 463-1565
Email: Krisztian.Nemeth@tmit.bme.hu

Andras Korn
Budapest University of Technology and Economics
Institute of Mathematics, Department of Analysis
Egry Jozsef u. 2, H-1111 Budapest, Hungary
Phone: +36 1 463-2475
Email: Korn@math.bme.hu

Istvan Cselenyi
TeliaSonera International Carrier
Vaci ut 22-24, H-1132 Budapest, Hungary
Phone: +36 1 412-2705
Email: Istvan.Cselenyi@teliasonera.com