Network Working Group   Gabor Feher, BUTE
INTERNET-DRAFT   Istvan Cselenyi, TRAB
Expiration Date: May 2002   Andras Korn, BUTE

November 2001

Benchmarking Terminology for Routers Supporting Resource Reservation
<draft-ietf-bmwg-benchres-term-01.txt>

1. Status of this Memo

2. Table of contents

3. Abstract

The purpose of this document is to define terminology specific to the performance benchmarking of the resource reservation signaling of IP routers. These terms are used in additional documents that define benchmarking methodologies for routers supporting resource reservation and define reporting formats for the benchmarking measurements.

4. Introduction

The IntServ over DiffServ framework [1] outlines a heterogeneous Quality of Service (QoS) architecture for multi domain Internet services. Signaling based resource reservation (e.g. via RSVP [2]) is an integral part of that model. While this significantly lightens the load on most of the core routers, the performance of border routers that handle the QoS signaling is still crucial. Therefore network operators, who are planning to deploy this model, shall scrutinize the scalability limitations in reservation capable routers and the impact of signaling on the forwarding performance of the routers.

An objective way for quantifying the scalability constraints of QoS signaling is to perform measurements on routers that are capable of resource reservation. This document defines terminology for specific set of tests that vendors or network operators can use 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 supporting 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 control plane signaling on the forwarding performance of routers.

This benchmarking terminology document is based on the knowledge gained by examination of (and experimentation with) several very different resource reservation protocols: RSVP [2], Boomerang [5], YESSIR [6], ST2+ [7], SDP [8] and Ticket [9]. Nevertheless, this document aspires to compose terms that are valid in general and not restricted to these protocols.

5. Existing definitions

RFC 1242 [3] "Benchmarking Terminology for Network Interconnect Devices" and RFC 2285 [4] "Benchmarking Terminology for LAN Switching Devices" contains 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.

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 in sections for ease of reference.

6. Definition of Terms

6.1 Resource Reservation Protocol Basics

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

6.1.1 Resource Reservation Session

Definition:
A resource reservation session (or shortly reservation) expresses that routers along the data path between two network nodes apply special QoS treatment to a certain traffic flow.

Discussion:
The QoS treatment is specified by giving the amount of networking resources that are dedicated to the traffic flow during the length of the reservation session. Depending on the protocol, there are different approaches to define the network resource requirement of a traffic flow. It can be described by high-level parameters, like the required bandwidth, service class or the maximum traffic delay; or it can be low-level information, like the parameters of a leaky-bucket model of the traffic flow [10].

Issues:
There are resource reservation protocols, where resource dedications in a router are unique for each resource reservation session. However, in this case the number of resource dedications grows along with the number of sessions and working with huge number of resource dedications raise problems (see Reservation Session Maintenance). Therefore, many resource reservation

protocols allow to bunch different reservation sessions into one aggregated session, which takes only one aggregated resource allocation for the whole bunch. The aggregation can be based on the similar attributes of the flows, (e.g. aggregation using DiffServ code-points [11]) or it can combine arbitrary sessions as well.

See Also:

Reservation Session Maintenance

6.1.2 Multicast Resource Reservation Session

Definition:

A multicast resource reservation session (or, in short, multicast reservation) denotes that certain QoS treatment is applied to the packets of every traffic flow related to a multicast group.

Discussion:

Usually, there are several traffic sources and destinations in a multicast group. In order to be able to guarantee the QoS parameters for each packet of the multicast flow, every router that forwards the multicast traffic must dedicate resources to the flow.

Generally, there are two types of multicast resource reservations: many-to-many and one-to-many multicast reservations. Those of the first type allow traffic to be originated from several sources, while those of the second type permit only one traffic source in the whole multicast group and this source should not change during the lifetime of the session. Additionally, in several cases, a many-to-many multicast reservation session does not require the same amount of resources reserved in every involved router. Depending on the resource reservation protocol, the traffic destinations of the multicast group may request different QoS parameters. Furthermore, the protocols may describe more than one reservation styles expressing the resource reservation distribution method among the involved routers. (e.g. RSVP SE/WF/FF [2])

Issues:

Naturally, many-to-many multicast reservation capable protocols are bound to be more complex than one-to-many or non-multicast protocols. Usually, the router has to be aware of the location of the traffic sources and destinations participating in the multicast reservation in the aspect of its network interfaces, plus the resource requirements of the traffic destinations in order to be able to calculate the right amount of resources dedicated to the session.

6.1.3 Resource Reservation Capable Router

Definition:

By definition, a router is resource reservation capable - supports resource reservation - if it understands a resource reservation protocol that signals the set-up, tear-down and modification of resource reservation sessions.

Discussion:

Resource reservation protocols define signaling messages that are interpreted by resource reservation capable routers. The router captures the signaling message and manipulates resource reservation sessions according to the content of the message. In addition, the protocol might declare to forward the same or a modified signaling message to other routers as well.

Issues:

There are resource reservation protocols where routers are required to initiate signaling messages besides the signaling message forwarding. The benefits of such protocols are that changes in resource reservation sessions can be signaled to other routers immediately, even if the change is not caused by signaling messages directly. In contrast, the message initiation takes time that slows down the signaling message processing, so there are protocols declaring instant responses, where all the signaling messages are forwarded to other routers immediately, avoiding the signaling message initiation.

6.1.4 Signaling End-point

Definition:

A signaling end-point is a network node capable of initiating and terminating resource reservation sessions.

Discussion:

Typically, signaling end-points have a separate protocol stack that is capable of generating and understanding the signaling messages. However, in some special cases, the resource reservation initiation is carried out without the notice of the signaling terminating node. For example, the Boomerang resource reservation protocol encapsulates the reservation requests in an ICMP Echo message. This message is bounced back from the signaling terminating network node and as a result the node becomes a signaling end-point without understanding the reservation protocol.

There are reservation gateways that translate the signaling messages of one resource reservation protocol into messages of another resource reservation protocol. Thus the reservation gateway represents two signaling end-points in one, as it is both a signaling terminator and a signaling initiator.

6.1.5 Reservation Initiator

Definition:

The reservation initiator is the signaling end-point that initiates the resource reservation session setup.

Discussion:

Resource reservation protocols can be classified depending on the relationship between the reservation initiators and their role in the traffic flow.

In the case of receiver-oriented protocols, the traffic destinations, which are the receivers of the data traffic, initiate the reservation session setup, unlike the sender-oriented protocols where this is done by traffic sources. Moreover, there are protocols where both the traffic source and destination can act as the reservation initiator.

The importance of the reservation initiator orientation is only dominant in case of multicast reservation sessions. Generally, in multicast groups the number of traffic destinations changes more frequently than the number of traffic sources. In this case, the receiver-oriented protocols do not require the traffic sources to change their states generating signaling messages when a new traffic destination joins or an existing one leaves the group, since the reservation session is managed by the traffic destinations.

Issues:

Receiver-oriented resource reservation protocols often require two pass to setup the reservation session. Since the resource dedication should take place on all the routers that are on the path from the sources to the destinations, thus the reservation initiator must have a method to reach every such router. In the case of the sender-oriented protocols this method is assured by sending the signaling traffic same way as the data traffic, which guarantees that signaling messages goes through the same routers as the data traffic. However, in the case of the receiver-oriented protocols the reservation initiation requests go from the destinations to the sources, which is the opposite direction compared to the data traffic. Thus, receiver-oriented protocols must provide a mechanism that at first discovers all the routers along the path of the data traffic (RSVP PATH message [2]), before the second pass, where the traffic destinations are ready to initiate resource reservation requests.

6.1.6 Reservation Session Maintenance

Definition:

Soft-state resource reservation protocols require the routers to maintain the resource reservation sessions from the initiation until the teardown of the session. This maintenance often involves the regular checking and refreshing of the session.

Discussion:

Based on the approach of reservation session maintenance, resource reservation protocols can be divided into two categories: soft-state protocols and hard-state protocols.

In the case of hard-state protocols (e.g. ST2 [7]), the resource reservation session established by a set-up signaling primitive is permanent and is cancelled only when the corresponding tear-down signaling primitive arrives to the router. Contrary, in the case of the soft-state protocols there are no permanent resource reservations. The resource reservation session have to be regularly refreshed by appropriate signaling messages. No refresh signaling message arrived during a certain period is assumed as the indication that the resource reservation session is not maintained by the signaling end-points any longer. In such case, the router tears the session down waiting for no explicit request. For this reason, soft-state protocols exhibit more robust behavior than hard-state protocols, since no failures can cause permanent resource stuck in the routers.

Issues:

Although soft-state protocols are more robust than hard-state protocols, the frequent processing of refresh signaling messages might cause serious increase in the router load. Moreover, session maintenance mechanisms often use timers to watch the refresh period expirations of the sessions. The maintenance of such timers and the actions due to the expiration of such timers also contributes to the router load.

In order to reduce the large number of refresh signaling message processing overhead, the resource reservation protocol may support various mechanisms to pack several refresh signaling messages into one signaling message.

6.1.7 Signaling Path

Definition:

A signaling path is a sequence of network nodes and links along which signaling messages travel from one signaling end-point to the other.

Discussion:

The resource reservation protocol must provide that each router, which is responsible to handle a signaling message, truly receives the signaling message. Usually it is assured by passing through the signaling messages along the signaling path, which involves every router affected by the resource reservation session.

Resource reservation protocol must be also prepared that there are routers forwarding the data traffic of a resource reservation session that do not support the actual resource reservation protocol. In this case the signaling traffic must be tunneled through the zones of the routers that can not interpret the signaling messages in order to keep the continuity of the signaling path.

Issues:

It is not unusual for routers to change their routing from time to time. One reason for the change can be a failure of a neighboring router or link. In case of route changes the data traffic will be forwarded along a different path than the signaling messages used in establishing the resource dedications for the reservation session. In order to properly handle this situation, hard-state protocols have to be much more sophisticated in order to detect the route change and to re-reserve the resources on the new path. However, soft-state protocols do not have to worry about such situation, since the refresh messages can be used to set up the reservation on the new path and the dedicated resources will eventually disappear from routers of the obsoleted path.

6.2 Traffic Types

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

6.2.1 Premium Traffic

Definition:
Premium traffic is a traffic type that the router distinguishes from best-effort traffic (to be defined later) and forwards its packets according to a QoS agreement.

Discussion:
Traffic that corresponds to a resource reservation session in the router is premium traffic. The QoS treatment is defined in the associated flow descriptor that is established by the signaling messages during the reservation session setup.

The router may distinguish several types of premium traffic (e.g. delay sensitive traffic, loss sensitive traffic, etc.). Different types of premium traffic may receive different QoS treatment.

Issues:
The router has to identify every packet whether it has dedicated resource or not. This can be done by either multi-field classification [11] using the IP 5-tuple or behavior-aggregate classification using the DSCP field. However, if a packet claims that it has an associated resource reservation session in the router, the router has to find the flow descriptor, which might be time consuming in routers with vast amounts of resource reservation sessions.

6.2.2 Best-Effort Traffic

Definition:
Best-effort traffic is a traffic type that has no reservation entry in the router.

Discussion:

Traffic flows that do not require QoS guarantees along their path are considered to be best-effort traffic. "Best-effort" means that the router makes its best effort to forward every data packet, but does not guarantee anything. This is the most common type of traffic on today's Internet. (There may be scenarios where resource reservation is done for BE traffic too, but those are outside of the focus of this memo.)

6.3 Router Load Types

This group of definitions describes different load component types that impact only a specific part of the resource reservation capable router. Categorizing the router load is crucial, since the conventional router load metric expressing the processing power utilization of the router does not characterize perfectly the resource reservation capable router. In the case of routers supporting resource reservations it is also important to know the source of the processing power utilization.

6.3.1 Traffic Load

Definition:
Traffic load is the load that is raised by forwarding data traffic on the router.

Discussion:
It is obvious that forwarding the data packets, which requires obtaining the routing information and transferring the data packet between network interfaces, requires processing power. Speaking of general router measurements only this type of load is considered as the source of the processing power utilization expressed by the router load metric. Although the traffic load is the dominant load component, benchmarking routers supporting resource reservations must consider other load components also in line with the resource reservation handling.

Measurement unit:
The amount of the traffic load is represented by the volume of the data traffic. The volume is measured with the transferred bits during a specified time unit, which is typically bit per seconds (bps).

6.3.2 Session Load

Definition:
Session load is the load that manifests itself as the excess processing power required to keep track of reservation sessions.

Discussion:
All signaling based resource reservation protocol implementation

employ a packet classifier algorithm that distinguishes the flows having reservations in the router from the others that do not. Therefore each implementation maintains a list of reservation session descriptors that is instrumental in keeping track of the resource reservation dedication. Obviously, the more reservation sessions are set up on the router, the more complex traffic classification becomes, and the more time it takes for the classification algorithm to identify the session descriptor list.

Moreover, in most protocols, not only the traffic flows, but also signaling messages that manipulate resource reservations on the router have to identify themselves first, before taking any other actions. This kind of classification gives extra work for the router.

Session load also involves the duties related to reservation session maintenance. The maintenance of timers that watchdog the reservation session refreshes and the signaling of the reservation session refresh may cause severe load on the router. Based on the initiating point of the refresh messages, resource reservation protocols can be divided into two groups. First, there are protocols where it is the responsibility of the signaling end-points to initiate refresh messages, which messages are forwarded by the routers along the signaling path refreshing the corresponding session. Second, there are other protocols, where the session refresh happens between the two peering network nodes from the signaling path only. In this latter case, the routers and signaling end-points have their own schedule for the refresh message initiation. The first approach lightens the load of the session maintenance task; however, the second approach bears the ability to adjust the signaling message traffic intensity along the signaling path.

Measurement unit:

The session load is represented by the number of reservation sessions in the router.

6.3.3 Signaling Load

Definition:

Signaling load is the load that manifests itself as the time required to process the incoming signaling messages.

Discussion:

The processing of signaling messages requires processing power that raises load on the control plane of the router. In the case of 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 or transfer buses) the signaling load can have an impact on the router's packet forwarding performance as well.

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.

Measurement unit:
The signaling load is characterized by the signaling intensity, which expresses how many signaling messages arrive to the router within a time unit. The typical unit of the signaling intensity is [1/s], which is the number of signaling messages that arrive within one second.

6.3.4 Signaling Burst

Definition:
The signaling burst denotes a certain number of signaling messages that arrive to the input port(s) of the router without interruption, causing persistent load on the signaling message handler.

Discussion:
Back-to-back signaling messages on one port of the router form a typical signaling burst. However, other cases are imaginable, for example when signaling messages arrive on different ports simultaneously or with an overlap in time (i.e. when the tail of one signaling message is behind the head of another one arriving on another port).

Measurement unit:
The signaling burst is characterized by its length, which is the number of messages that have arrived during the burst.

6.4 Performance Metrics

This group of definitions is the collection of the measurable impacts that a resource reservation protocol has over the tested router device it is running on.

6.4.1 Signaling Message Handling Time

Definition:
The signaling message handling time (or, in short, signal handling time) is the time that a signaling message spends inside the router before it is forwarded to the next node on the signaling path.

Discussion:
Depending on the type of the signaling message, the router also interprets the signaling messages, acts on them and forwards an extended signaling message, which might contain information about the result of the message processing. Thus the message handling time is usually longer than forwarding time of data packets of the same size. In addition, there might be also signaling message primitives that are drained or generated by the router. Thus, the

signal handling time is defined as the time difference between the time when a signaling message is received and the time the corresponding processed signaling message is transmitted. If a message is not forwarded on the router, 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. In this case, by definition, the signal handling time is the time interval elapsed between the arrival of the incoming signaling message and the departure of the last signaling message related to the received one.

This metric depends on the load on the router, as other tasks may limit the processing power available to signaling message handling. In addition to the router load, the signal handling time may also be dependent on the type of the signaling message. For example, it usually takes a shorter time for the router to tear down a resource reservation session than to set it up.

Issues:

In the case of soft-state protocols, where refresh messages are exchanged between peering network nodes only (see Reservation Session Maintenance) the incoming refresh messages are drained by the router making impossible to measure the signaling message handling time on them.

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

Measurement unit:

The typical unit of the signaling message handling time is microsecond.

6.4.2 Premium Traffic Delay

Definition:

Premium traffic delay is the forwarding time of a packet that belongs to a premium traffic flow passing through a resource reservation capable router.

Discussion:

Premium traffic packets must be classified first in order to find the resources dedicated to the flow. The time of the classification is added to the usual forwarding time that a router would spend on the packet without any resource reservation capability.

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 premium traffic delay metric reflects the influence the two planes have on each other.

Measurement unit:

The typical unit of the premium traffic delay is the microsecond.

6.4.3 Best-effort Traffic Delay

Definition:

Best-effort traffic delay is the forwarding time of a packet that does not belong to any premium traffic flow passing through a resource reservation capable router.

Discussion:

It looks trivial that the classification algorithms do not have any influence on the best-effort traffic. However, the processing power sharing between the control and data plane may cause delays in the forwarding procedure of each packet.

Measurement unit:

The typical unit of the best-effort traffic delay is microsecond.

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:

There are certain types of signaling messages, which are required to be forwarded by the router immediately when their processing is finished. However, due to the high router load or for other reasons, the forwarding or even the processing of the signaling message might be canceled. To characterize such situations we introduce the signaling message loss metric, which expresses the ratio of the signaling messages that actually have left the router and those ones that were expected to leave the router as a result of the incoming sequence of signaling messages.

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

Issues:

In the case of routers where network packets are queued in several places, we have to be aware of that a signaling message may be delayed seriously. Therefore, it may be hard or impossible to determine whether the signaling message is still in the queues or whether it was already dropped. By definition we say that a signaling message is lost if there is no appearing forwarded signaling message within a reasonable long time period. This time period should be adjusted to the actual resource reservation protocol (e.g. soft-state protocols may wait as much as the refresh period to determine the loss of a signaling message).

Measurement unit:

Usually, we measure the signaling message loss over a longer period of time and then we express it as a percentage value. Besides, in many cases it is enough to know that there was signaling loss.

6.4.5 Session Refreshing Capacity

Definition:

The session refreshing capacity is the ratio of the truly refreshed sessions and the number of session that have to be refreshed during one refresh period. This metric is applied for soft-state routers only.

Discussion:

The session refreshing capacity informs about condition of the session maintenance. When the router is overloaded it may happen that the router is not capable to refresh all the allocated reservation sessions due to other tasks with higher priorities. In this case sooner or later the resource reservation sessions over the session refresh capacity are dropped even if the resource reservation end-points are still refreshing them.

The session refreshing capacity sounds out the limit of resource reservation session number that the router is capable to maintain.

Measurement unit:

The session refreshing capacity is expressed as a percentage value.

6.4.6 Scalability Limit

Definition:

The scalability limit is the threshold between the steady state and the overloaded state of the device under test.

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 router has to postpone more and more tasks. These tasks (e.g. forwarding certain packets) are queued in the buffers, and processed later. However, there is a certain point where no more buffer memory is available; thus, the router becomes overloaded and it is unable to store any more tasks for future processing, so it is forced to drop them. Therefore the overloaded state of the router can be recognized by the fact that some kind of data (signaling or packet) loss occurs. A resource reservation capable router may drop signaling messages, data packets or entire resource reservation sessions.

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

7. Acknowledgement

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

8. References

[1] Y. Bernet, et. al., "A Framework for Integrated Services Operation over Diffserv Networks", RFC 2998, November 2000

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

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

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

[5] G. Feher, K. Nemeth, M. Maliosz, "Boomerang : A Simple Protocol for Resource Reservation in IP Networks," Vancouver, IEEE Real-Time Technology and Applications Symposium, June 1999

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

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

[8] 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

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

[11] K. Nichols et al., " Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers ", RFC 2474

9. Authors' Addresses:

Gabor Feher
Budapest University of Technology and Economics (BUTE)
Department of Telecommunications and Telematics
Pazmany Peter Setany 1/D, H-1117, Budapest, Hungary
Phone: +36 1 463-3110
Email: feher@ttt-atm.ttt.bme.hu

Istvan Cselenyi
Telia Research AB
Vitsandsgatan 9B
SE 12386, Farsta
SWEDEN,
Phone: +46 8 713-8173
Email: istvan.i.cselenyi@telia.se

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