Audio/Video Transport Working Group
Internet Draft
Document: draft-miyazaki-avt-rtp-selret-01.txt
July 14, 2000
Expires: January 14, 2001

Akihiro Miyazaki
Hideaki Fukushima
Thomas Wiebke
Rolf Hakenberg
Carsten Burmeister

Matsushita

RTP Payload Format to Enable Multiple Selective Retransmissions

Status of this Memo

1. Abstract

This document describes a new RTP payload format and a new RTCP packet format to enable multiple selective retransmissions of generic media data. This payload format is especially applicable in a RTP-RX profile as described in [9].

While RTP is widely used for media streaming applications over the Internet, using it over unreliable links needs some further modifications to achieve a good quality of the media stream at the receiver.

The aforementioned RTP-RX profile describes a generic retransmission mechanism to make RTP more reliable. This payload format description increases the reliability and decreases the bandwidth requirements of such a generic profile by introducing the concept of multiple selective retransmissions.

This paper describes a mechanism to use the lower delay requirements in streaming applications to make RTP more reliable, by the means of multiple selective retransmissions. Therefore a new RTP payload format is defined as well as a new RTCP packet. An example algorithm to use the new payload format fields and the new RTCP packet is introduced and an additional section shows results of a performance evaluation. The evaluation is done by simulating a video streaming application over an mobile communication system.

Table of Contents

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [2].

3. Introduction

The Real-Time Transport Protocol (RTP)[3] was designed as an Internet protocol to transmit real-time (or near real-time) data over multicast or unicast. Even though the multicast capability is a strong feature of RTP, a common use of it is non-interactive unicast transmission of media streams. This application, further referred to as media streaming, has lower delay requirements. A constant offset delay, in which already received packets are stored at the client, is tolerable.

Typically RTP is run over the User Datagram Protocol (UDP)[4]. While the unreliable UDP performs good for reliable channels, the quality of the received media stream is bad, when transmitted over error-prone links. Especially compressed streams, such as MPEG video-data, are highly susceptible to packet loss, which will lead to a bad video quality at the client. E.g. in a wireless environment with a unstable transmission quality (bit error rate (BER) about 10E-3 .. 10E-6) a sufficient quality of video and voice transmitted by RTP cannot be achieved.

The scenario for such a media streaming over an error-prone link might be the one in Figure 1. Media files are stored on a media server which is connected via an internet link to a mobile network. The wireless link between the mobile network and the mobile terminal, which works as the client in this scenario, is generally unreliable and subject to bit errors.

```
      Media                   Mobile                  Mobile
      Server                  Network                 Terminal
                                                      (Client)

                                ***                       |
      +-----+                 **   **                   +--+
      |     |                *       *                  |  |
      |     | -------------- *         *  ~ ~ ~ ~ ~ ~ ~ |  |
      +-----+                *       *                  +--+
                              **   **
                Internet        ***        Wireless
                Link                       Link
```

Figure 1 : Scenario for media streaming

To achieve a better quality of the received media stream, especially when transmitted over error-prone links, retransmissions of lost packets of the stream seems to be a possible solution. However the quite low delay requirements of media streaming applications and the bandwidth limitations of the wireless link have to be considered. In wireless systems bandwidth is a scarce resource and efficient use of this resource is of vital importance. Hence reliability for the complete media stream might not be achievable.

Furthermore in compressed media streams, e.g. MPEG video stream, not every frame is equal important. In this example some frames contain data that other data depends upon, which makes these frames more important. Current protocols applying retransmission for reliable transmission (e.g. TCP) make no use of this characteristic of compressed media streams.

4. Proposed solution

Much work has been done recently to make RTP more reliable. As one result a new profile for unicast sessions is proposed [9]. This profile includes the ability to request retransmissions of lost packets. However the profile is kept very general and leaves space for a payload format to define further details and enhance the functionality.

The solution presented in this document is based on the concept of multiple selective retransmissions of data packets in case of packet loss. This functionality is added to the RTP RX profile by the means of this payload format.

As mentioned before, compressed media streams generally consist of data packets of high importance and data packets of lower importance. In the proposed solution data packets of high importance are retransmitted in case of packet loss and by this the quality of the received media stream is increased significantly.
However single retransmission of lost data packets, which are transmitted over links with high bit error rates are still far from being reliable. Even though no 100% reliability is desired, the delay requirements and bandwidth limitations might allow more than one retransmission of important parts of the media stream. Therefore mechanisms for multiple retransmission attempts are included in our proposal.

On the other hand by limiting the retransmissions on only the data packets of high importance and transmitting the data packets of lower importance the usual unreliable way the delay requirements of media streaming can still be meet. Additionally the total bandwidth requirements for transmission and retransmission of the media stream data are kept at a low level suitable and advantageous for usage over bandwidth limited, wireless links.

To keep also the bandwidth requirements on the link back from the client to the media server at a minimum level, means for a judgement at the client, whether a retransmission would still arrive in time (retransmission judgement), is proposed. With our solution it is possible to calculate the time when the lost packet has to be received at the client at latest. This time is compared against the round trip time. Only if this calculation predicts that a retransmission of the lost packet will be received in time, a retransmission is generated at the client and send to the server. Therefore no bandwidth is wasted for unnecessary retransmission requests and subsequent retransmissions.

In summary we propose the following enhancements/extensions to RTP:

1) Selective retransmission
   Retransmission of lost packets, but limited to lost packets of high importance.

2) Multiple retransmission attempts
   In case retransmission of a lost data packet fails, further attempts can be requested.

3) Retransmission with time limit based on client judgement
   Means to calculate the time when a lost packet has to be received at the client at latest to perform a efficient client based retransmission judgement.

RTP packet retransmissions in a unicast scenario are an essential part of the proposed RTP-RX profile [9]. The profile is kept very general and could be used as a starting point. However, to enable the items listed above, additions are needed. These are described in the following sections.

5. Concept of a second sequence number

In this section the concept of a second sequence number is described, which enables to do multiple retransmissions very efficiently.

Normally a loss of a transmitted RTP packet can be realized by a gap in the sequence numbers of the received packets. This is a very time efficient way to detect packet loss, since it does not introduce additional traffic or delay.

However, retransmissions of lost RTP packets need to have the same sequence number as the original transmitted packet, because the sequence number is used for other purposes than packet loss detection as well (e.g. synchronization, coping with reordering). Therefore it is not possible to detect a lost retransmission by looking at the sequence numbers.

Other mechanisms to detect packet loss (e.g. acknowledgements, timer-based loss detection) have the disadvantage that they either introduce additional traffic or delay. Both is not desirable, especially not for delay sensitive traffic over a bandwidth limited channel.

To overcome these problems, the following mechanism is used to detect lost retransmissions. A second sequence number (SSN) is introduced. For every packet that is (re)transmitted the source decides if the packet should be retransmitted if it gets lost. If it should be retransmitted it gets a new SSN assigned. If it should not be retransmitted, it carries the SSN of the previous packet. With this mechanism, the client can detect lost packets, which should be retransmitted.

6. Payload header in a new RTP payload format

In order achieve the solution described above, we extend the RTP header in the RTP payload format by the following fields.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M| PT=MulSelRx |       sequence number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           synchronization source (SSRC) identifier            |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|            contributing source (CSRC) identifiers             |
|                             ....                              |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|S|      PT     |      SSN      |D|          Diff Time          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The first part of the header is the usual RTP header with a payload format identifier set to this new payload format "MulSelRx". The payload contains an additional payload header consisting of 4 bytes. The fields of this additional header are explained as follows:

SSN Indicator (S): 1bit

If this bit is set to one, the packet should be retransmitted if it is lost. Therefore it has a new SSN value assigned, which is incremented by one to the previous packet's SSN. If this bit is set to zero, the packet should not be retransmitted if it is lost and therefore carries the SSN value of the previous packet.

Payload Type (PT): 7 bits

This fields identifies the payload format that follows the header additions and is used as described in [3].

Second Sequence Number (SSN): 8 bit

The second sequence number is carried in this field. It is used as describe in section 5.

Diff Time Indicator(D): 1 bit

If DTI field is set to one, the diff time field is included in the RTP header extension. For further information please refer to description of diff time field. If it is set to 0 no diff time is transmitted and the diff time field is not valid.

Diff Time: 15 bit

This field indicates the difference between the time stamp of the last RTP packet that should be retransmitted if lost (i.e. S==1) and that of the current RTP packet. It can be used by the receiver to compute the timestamp of a lost packet. The default unit of this field is [ms] and by default the following table is used to encode the diff time values.

Diff Time [ms]	bits
more than 16383	011 1111 1111 1111
16382	011 1111 1111 1110
16381	011 1111 1111 1101
:	:
2	000 0000 0000 0010
1	000 0000 0000 0001
0	000 0000 0000 0000
-1	111 1111 1111 1111
-2	111 1111 1111 1110
:	:
-16382	100 0000 0000 0010
-16383	100 0000 0000 0001
less than –16384	100 0000 0000 0000

It is possible to define other units or coding tables, however, how to specify and negotiate this is outside the scope of this document.

7. RTCP packet type

The retransmission request, which is sent from the sender to the receiver to request the retransmission of a lost packet has to contain the SSN value that is requested. The RTCP packet that is used in the RTP-RX profile to request retransmissions can be used for this purposes. It contains a 16 bit packet identification (PID) and 16 payload specific bits. These fields are used as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|   RC    |PT=RTCP_NACK   |           length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             SSRC                              |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|                             SSRC_1                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          PID = SSN            |            SSN BLP            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

version(V), padding(P), report count(RC), packet type(PT), length, SSRC and SSRC_1:
   as described in [9], Section 4.2.

packet identification(PID):
   The SSN value of the requested packet is inserted in this field. Since the SSN is only 8 bit long, the upper 8 bits are not used in this field.
   If several packets are requested (i.e. SSN_BLP != 0) this indicates the lowest SSN value of all requested packets.

SSN bitmask of following lost packets(SSN_BLP):
   By the means of this bitmask it is possible to request more than one lost packet. The use is similar as described in [9] section 4.3.

8. Indicating usage in SDP

Packets of this format contain RTP packets with dynamic payload type values. These values have to be indicated in SDP as well as the value for this payload format. An example indication within SDP would be:

```
m= video 0 RTP/AVP-RX xx yy
```

The first payload type value indicate the multiple selective retransmission payload type and the second the used media payload type.

9. An example use of the payload format

In sections 4 and 5 we proposed a solution to enhance the quality of media stream, transmitted over an unreliable link. How this solution can be achieved by using the header fields as described in section 6 and 7 is shown in form of an example in this section.

9.1 Selective retransmission

As described before a media stream contains more or less important parts. In section 6 we introduced a SSN indicator bit (S) which is used to mark packets, that should be retransmitted if they are lost. Hence these packets should contain the important parts of the media stream. With this it is possible to distinguish between two priority levels, and to allow retransmissions of only the important packets.

9.2 Multiple retransmission attempts

If a retransmission fails, further attempts to retransmit the lost packet MAY be requested. While the detection of a lost packet is normally straight forward (e.g. by checking the sequence number), detecting a lost retransmission requires further mechanisms. This is done by the means of the SSN field as described in section 5.

The SSN value MUST be incremented by one for every packet that should be retransmitted, regardless whether it is a retransmission or sent for the first time. By comparing the actual received with the most recent SSN value, the client will detect the loss of a high priority packet immediately (i.e. if the SSN value was increased, but no packet with S=1 was received, it must have been lost). This enables the client to send another retransmission request and the server to do multiple retransmissions.

9.3 Retransmission with time limit based on client judgement

As described in section 3, bandwidth might be a very scarce resource. Therefore requesting the retransmission or retransmitting packets that will not be received in time anyway, SHOULD be avoided. Hence in this example the client estimates the arrival time of the requested lost packet and the time after which it will become useless. Therefore the RTP timestamp of the lost packet is needed, which is achieved by using the diff-timestamp field.

In these fields the difference between the actual packet’s time stamp and the most recent packet’s timestamp that should be retransmitted is calculated. At detecting a lost packet with S=1, it is possible to calculate the timestamp of the lost packet by the means of this field, because the received packet contains its own timestamp and the difference to the lost packet's timestamp.

The knowledge of the lost packet's time stamp enables the client to judge if a retransmission request would still make sense or if the bandwidth should be saved.

10. Performance evaluation of the proposed solution

The previous sections introduced a solution, how to enhance the capabilities of media streaming application over unreliable links. The solution implies a new RTP payload format and a new RTCP packet as described in sections 6 and 7. How these fields can be used in general is described in an example in section 8.

This section will show the increased performance that is possible with our solution. This is done by simulating a video streaming application over a unreliable wireless link.

10.1 Application

In this performance evaluation we simulate a MPEG4 video streaming application. MPEG4 video data is streamed from a server to a client (i.e. unicast). In the client the received frames are buffered before they are displayed at their scheduled playing time. Therefore the delay requirements are not exactly real-time. The scheduled playing time includes an additional buffering delay.

We consider a transmission of a real MPEG4 video stream, with a total play time of about 120s and an average bit rate of 33100 bps. The frame rate is 10 fps. The frame length is variable between 100byte and 2000byte. The additional allowed buffering delay is set in our simulations to 2 seconds.

The MPEG4 video stream consists of Intra coded (I-)frames and Predictive coded (P-)frames. While the I-frames are completely independent and can be displayed all by themselves, the P-frames contain only the difference to the previous frame and hence need the previous frame to be received correctly to be displayed. This leads to I-frames having a higher importance than P-frames. The different frames are mapped into different RTP packets with packets containing an I-frame having S=1 and packets containing P-frames having S=0 assigned.

10.2 Environment

A mobile environment is considered as a typical representation for unreliable, bandwidth limited links. Figure 1 illustrates the general scenario that is evaluated. The used protocols for the evaluation are shown in Figure 2.

```
Media Server            Mobile Network          Mobile Client

+-----------+                                   +-----------+
|   Video-  |                                   |  Display- |
|Application|                                   |Application|
+-----------+                                   +-----------+
|    RTP    |                                   |    RTP    |
+-----------+                                   +-----------+
|    UDP    |                                   |    UDP    |
+-----------+           +-----------+           +-----------+
|    IP     |<--------->|    IP     |           |    IP     |
+-----------+           +-----------+           +-----------+
                        |  W-CDMA   |<--------->|  W-CDMA   |
                        +-----------+           +-----------+
              Internet                 Wireless
              Link                     Link
```

Figure 2

The RTP layer implies the newly defined RTP payload format and RTCP packet as described in sections 5 and 6 and the use of the additional fields as given in the example in section 7. Additional to the client’s retransmission judgement, a similar mechanism is used on the server side, to avoid unnecessary retransmissions.

The Internet link is modeled as error-free, loss-less and with a constant delay of 50 ms.

For the wireless link a W-CDMA channel is considered. W-CDMA, as a third generation mobile communication system with high Quality-of-Service (QoS) capabilities, is a well suited network for this kind of applications in a mobile use.

The W-CDMA system is standardized at the 3rd Generation Partnership Project (3GPP). The standardization is still under progress and no final specifications are finished. For the simulations the specifications of December 1999 are considered. The wireless link and its link layer protocols are simulated according to this specifications (see [7] for details).

A single user is considered for the mobile channel, hence the bandwidth is not shared between different users here. The channel is simulated with the fading model, as described in [8]. The fading process has a mean signal-to-noise ratio per received symbol of Es/N0. Different mean Es/N0 values are simulated, which reflect different channel conditions.

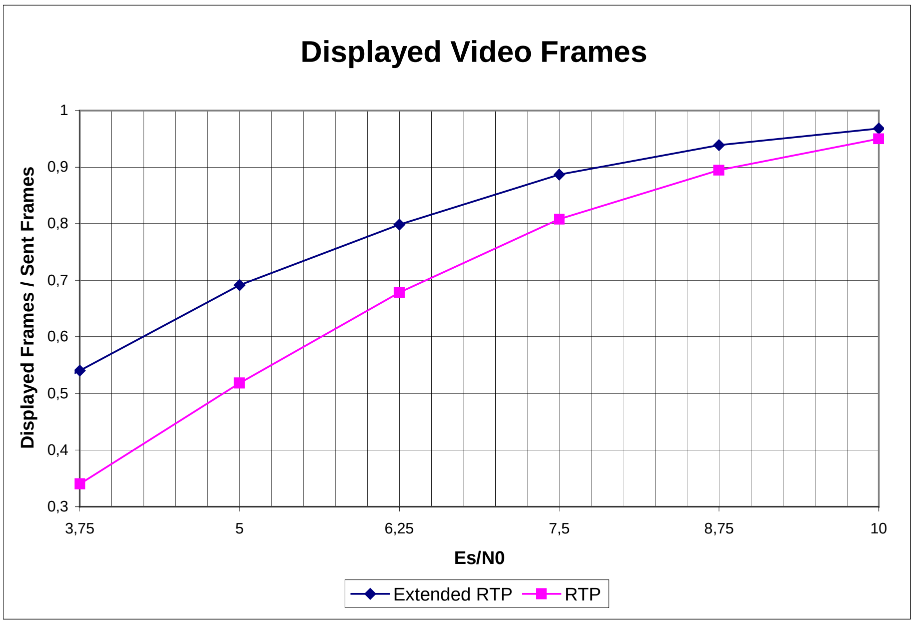
The back channel in our simulations is modeled as error-free and therefore loss-less.

The propagation delay of the mobile channel, is strongly dependent on the length of the packet that is to be sent. The minimum delay of 50ms applies for all transmitted frames, due to the coding, interleaving etc. in the physical layer.

10.3 Performance results

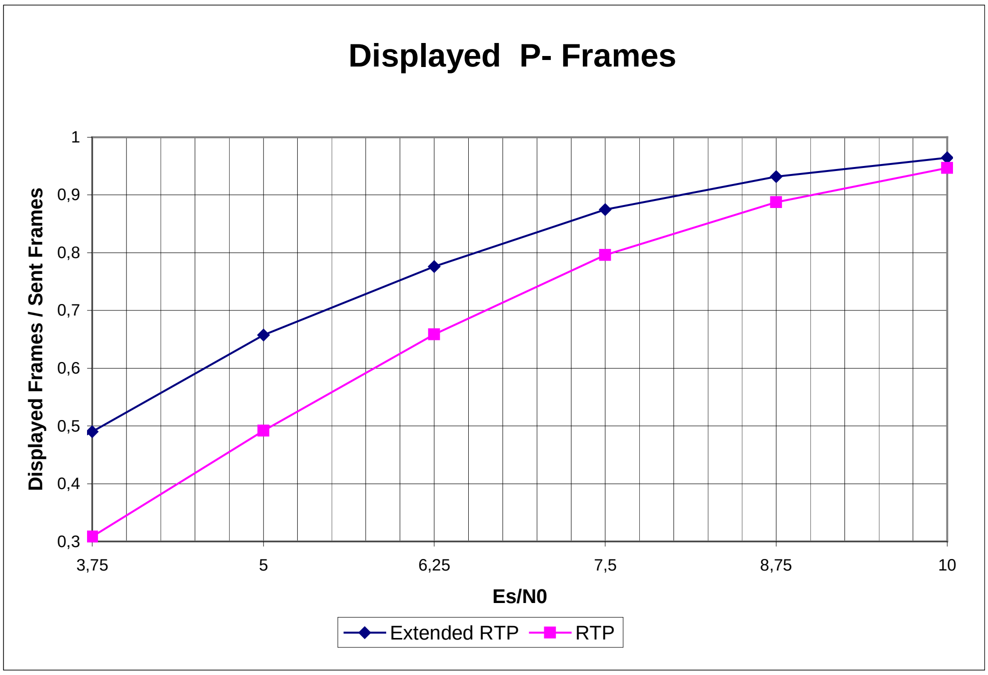
This section shows some results of the simulations as described above. To measure the quality of the received video stream, the number of displayed frames, relative to the number of sent frames, is considered.

The next figure shows the amount of received frames for RTP and the extended RTP for different channel conditions.



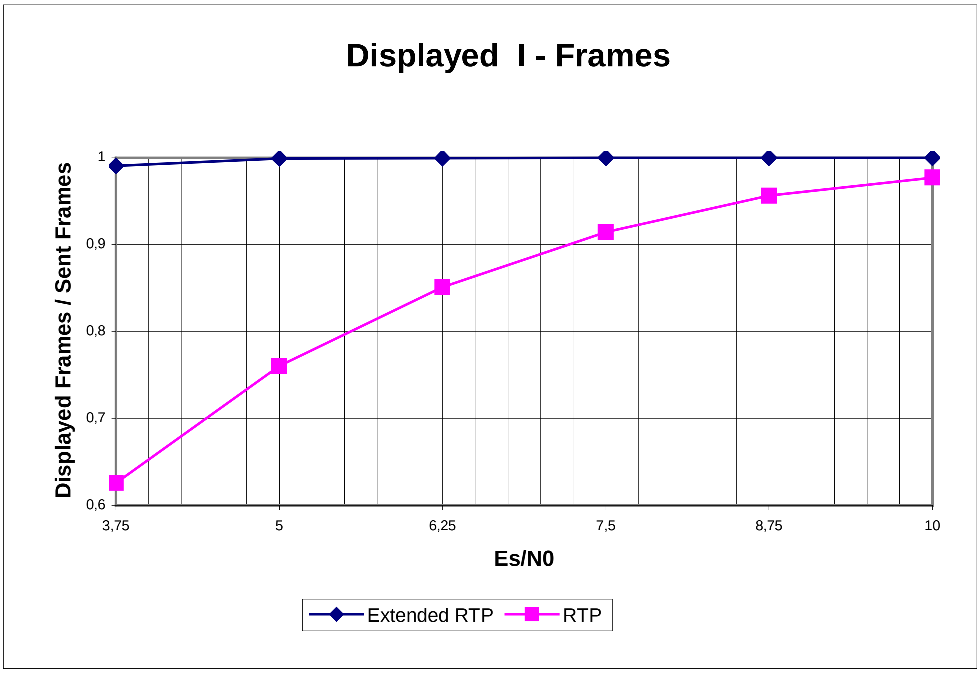
It can be seen from this figures, that both applications suffer from bad channel conditions.

In states, where Es/N0 is smaller than 5dB, not even half of the video frames can be displayed, using normal RTP. In those conditions many errors occur, which leads to a loss of many frames.
Additionally, if an I-frame gets lost, all following P-frames, even if they would be received correctly, are discarded.
The application using the extended RTP as described in section 7, performs much better. Even though it suffers from bad channel conditions as well, the repetition of lost I-frames, leads to an increased performance.
The next figure shows the amount of displayed P-frames for different channel conditions.



Even though the lost P-frames are not retransmitted, more frames are displayed, using the extended RTP. This is due to the fact, that fewer correctly received frames are discarded, because of a lost I-frame.

The next figure shows the amount of received I-frames for different channel conditions.



From this figure the effect of the extended RTP is visible best. Nearly all I-frames are displayed at the receiver in this case, while only half of the frames can be displayed with normal RTP in bad channel conditions.

10.4 Conclusions

It was shown in the previous sections that repetition of important media data, leads to a better media quality at the receiver. These repetitions are possible, if the lower delay requirements in streaming applications are used wisely. To do so, a newly defined RTP payload format and RTCP packet are introduced, which enable a media server to do multiple, selective retransmissions.

11. Security Considerations

Security is not considered in this draft.

12. Intellectual property considerations

Matsushita has filed patent applications that might possibly have technical relation to this contribution.

13. References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

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

3 Schulzrinne, H., Casner, S., Frederick, R. and V. Jacobson, "RTP: A Transport Protocol for real-time applications", RFC 1889, January 1996.

4 Postel, J.,"User Datagram Protocol", STD 6, RFC 768, August 1980.

5 Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

6 Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

7 Homepage of 3GPP: http://www.3gpp.org

8 "Procedure for Evaluation of Transmission Technologies for FPLMTS", ITU-R TG8-1, 8-1/TEMP/233-E, September 1995.

9 Yano, K., Podolsky, M., McCanne, S., "RTP Profile for RTCP-based Retransmission Request for Unicast session", IETF Internet Draft, work in progress, July 2000.

14. Author’s Addresses

Akihiro Miyazaki
Matsushita Electric Industrial Co., Ltd
1006, Kadoma, Kadoma City, Osaka, Japan
Tel. +81-6-6900-9192
Fax. +81-6-6900-9193
Mail akihiro@isl.mei.co.jp

Hideaki Fukushima
Matsushita Electric Industrial Co., Ltd
1006, Kadoma, Kadoma City, Osaka, Japan
Tel. +81-6-6900-9192
Fax. +81-6-6900-9193
Mail fukusima@isl.mei.co.jp

Thomas Wiebke
Panasonic European Laboratories GmbH
Monzastr. 4c, 63225 Langen, Germany
Tel. +49-(0)6103-766-161
Fax. +49-(0)6103-766-166
Mail wiebke@panasonic.de

Rolf Hakenberg
Panasonic European Laboratories GmbH
Monzastr. 4c, 63225 Langen, Germany
Tel. +49-(0)6103-766-162
Fax. +49-(0)6103-766-166
Mail hakenberg@panasonic.de

Carsten Burmeister
Panasonic European Laboratories GmbH
Monzastr. 4c, 63225 Langen, Germany
Tel. +49-(0)6103-766-263
Fax. +49-(0)6103-766-166
Mail burmeister@panasonic.de

Full Copyright Statement