

Audio/Video Transport Working
Group
Internet-Draft
Expires: August 23, 2008

A. Clark
Telchemy Incorporated
G. Hunt
BT
A. Pendleton
Nortel
R. Kumar
K. Connor
Cisco Systems
February 25, 2008

**RTCP HR - High Resolution VoIP Metrics Report Blocks
draft-ietf-avt-rtcphr-03.txt**

Status of this Memo

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

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

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

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

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

This Internet-Draft will expire on August 23, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2008).

Abstract

This document defines extensions to the RTCP XR extended report packet type blocks to support Voice over IP (VoIP) monitoring for services that require higher resolution or more detailed metrics than those supported by [RFC3611](#).

Table of Contents

1.	Introduction	5
2.	Definitions	6
2.1.	Cumulative and Interval Metrics	6
2.2.	Bursts, Gaps, and Concealed Seconds	6
2.3.	Numeric formats	6
3.	High Resolution VoIP Metrics Report Block	8
3.1.	Block Description	8
3.2.	Header	10
3.2.1.	Block type	10
3.2.2.	Map field	10
3.2.3.	Block Length	11
3.2.4.	SSRC	11
3.2.5.	Duration	11
3.3.	Basic Loss/ Discard Metrics	11
3.3.1.	Loss Proportion	12
3.3.2.	Discard Proportion	12
3.3.3.	Number of frames expected	12
3.4.	Burst/Gap metrics sub-block	12
3.4.1.	Threshold	13
3.4.2.	Burst Duration (ms)	13
3.4.3.	Gap Duration (ms)	13
3.4.4.	Burst Loss/Discard Proportion	13
3.4.5.	Gap Loss/Discard Proportion	14
3.5.	Playout Metrics sub-block	14
3.5.1.	On-time Playout Duration	14
3.5.2.	On-time Active Speech Playout Duration	15
3.5.3.	Loss Concealment Duration	15
3.5.4.	Buffer Adjustment Concealment Duration (optional)	15
3.6.	Concealed Seconds metrics sub-block	16
3.6.1.	Unimpaired Seconds	17
3.6.2.	Concealed Seconds	17
3.6.3.	Severely Concealed Seconds	17
3.6.4.	SCS Threshold	18
3.7.	Delay and Packet Delay Variation (PDV) metrics sub-block	18
3.7.1.	Network Round Trip Delay (ms)	18
3.7.2.	End System Delay (ms)	18
3.7.3.	External Delay (ms)	18

3.7.4.	PDV/Jitter Metrics	19
3.7.5.	PDV Type	21
3.7.6.	Jitter Buffer / PLC Configuration	22
3.7.7.	Jitter Buffer Size parameters	22
3.8.	Call Quality Metrics sub-block	22
3.8.1.	Listening and Conversation Quality R Factors - R-LQ, R-CQ	22
3.8.2.	Listening and Conversation Quality MOS - MOS-LQ, MOS-CQ	22
3.8.3.	R-LQ Ext In and Out	23
3.8.4.	RFC3550 RTP Payload Type	24
3.8.5.	Media Type	24
3.8.6.	Received Signal and Noise Levels - IP side	24
3.8.7.	Received Signal and Noise Levels - External	24
3.8.8.	Local and Remote Residual Echo Return Loss	25
3.8.9.	Metric Status	25
4.	RTCP HR Configuration Block	28
4.1.	Header	28
4.1.1.	Block type	28
4.1.2.	Map field	29
4.1.3.	Block Length	29
4.1.4.	SSRC	29
4.2.	Correlation Tag	29
4.3.	Algorithm descriptor	30
5.	SDP Signalling	32
5.1.	The SDP Attributes	32
5.2.	Usage in Offer/Answer	34
5.3.	Usage Outside of Offer/Answer	35
6.	Practical Applications	36
6.1.	Overview	36
6.2.	Supplementary Services: Call Hold and Transfer	36
6.2.1.	General	36
6.2.2.	Supplementary Service: Call Transfer	36
6.2.3.	Supplementary Service: Call Hold	37
6.3.	Bitrate efficiency improvements: VAD/Silence Suppression based on Voice Activity Detection (VAD) Elimination	37
6.4.	Endpoint configuration changes mid-call	37
6.4.1.	Changes due to mid-call transitions between different voice codec types	37
6.4.2.	Changes due to mid-call transitions from VoIP to RTP-based VBDoIP	37
6.4.3.	Changes due to mid-call transitions from VoIP to non-RTP -based VBDoIP	38
6.5.	SSRC changes mid-call	38
7.	IANA Considerations	39
8.	Security Considerations	40
9.	Contributors	41

10 . Informative References	42
Authors' Addresses	44
Intellectual Property and Copyright Statements	45

1. Introduction

This draft defines several new block types to augment those defined in [[RFC3611](#)] for use in Quality of Service reporting for Voice over IP. The new block types support the reporting of metrics to a higher resolution to support certain applications, for example carrier backbone networks.

For certain types of VoIP service it is desirable to report VoIP performance metrics to a higher resolution than provided in the [[RFC3611](#)] VoIP Metrics block or [[RFC3550](#)] Receiver Reports. The report blocks described in this section provide both interval based and cumulative metrics with a higher resolution than that provided in the [[RFC3611](#)] VoIP metrics report block.

The new block types defined in this draft are the High Resolution VoIP Metrics Report Block, and the High Resolution VoIP Metrics Configuration Block.

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. Definitions

2.1. Cumulative and Interval Metrics

Cumulative metrics relate to the entire duration of the call to the point at which metrics are determined and reported, and are typically used to report call quality. Cumulative metrics generally result in a lower volume of data that may need to be stored, as each report supersedes earlier reports.

Interval metrics relate to the period since the last Interval report. Interval data may be easier to correlate with specific network events for which timing is known, and may also be used as a basis for threshold crossing alerts.

Note that interval metrics for the start and end of calls may be unreliable due to factors such as irregular start and end interval length and the difficulty in knowing when packet transmission started and ended.

2.2. Bursts, Gaps, and Concealed Seconds

The terms Burst and Gap are used in a manner consistent with that of RTCP XR [[RFC3611](#)]. RTCP XR views a call as being divided into bursts, which are periods during which the combined packet loss and discard rate is high enough to cause noticeable call quality degradation (generally over 5 percent loss/discard rate), and gaps, which are periods during which lost or discarded packets are infrequent and hence call quality is generally acceptable.

The recommended value for Gmin in [[RFC3611](#)] results in a Burst being a period of time during which the call quality is degraded to a similar extent to a typical PCM Severely Errored Second.

The term Concealed Seconds defines a count of seconds during which some proportion of the media stream was lost through packet loss and discard. The term Severely Concealed Seconds defines a count of seconds during which the proportion of the media stream lost through packet loss and discards a specified threshold.

2.3. Numeric formats

This report block makes use of binary fractions. The terminology used is

S X:Y

where S indicates a two's complement signed representation, X the

number of bits prior to the decimal place and Y the number of bits after the decimal place.

Hence 8:8 represents an unsigned number in the range 0.0 to 255.996 with a granularity of 0.0039. S7:8 would represent the range -128.000 to +127.996.

3. High Resolution VoIP Metrics Report Block

3.1. Block Description

This block comprises a header and a series of sub-blocks. The Map field in the header defines which sub-blocks are present.

Header sub-block

```

0               1               2               3
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      BT=N      |      Map      |      block length      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     SSRC of source         |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Duration                |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Basic Loss/Discard Metrics sub-block

```

0               1               2               3
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Loss Proportion      |      Discard Proportion      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Number of frames expected
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Burst/Gap metrics sub-block

```

0               1               2               3
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Threshold      |      Burst Duration (ms)      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Gap Duration (ms)      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Burst Loss/Disc Proportion | Gap Loss/Disc Proportion |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```


Playout metrics sub-block

0								1								2								3							
0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
On-time Playout Duration																															
On-time Active Speech Playout Duration																															
Loss Concealment Duration																															
Buffer Adjustment Concealment Duration																															

Concealed Seconds metrics sub-block

0								1								2								3							
0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
Unimpaired Seconds																															
Concealed Seconds																															
Severely Concealed Seconds																RESERVED								SCS Threshold							

Delay and PDV metrics sub-block

0								1								2								3							
0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
Network Round Trip Delay																End System Delay															
External Delay																Mean PDV															
Pos Threshold/Peak PDV																Pos PDV Percentile															
Neg Threshold/Peak PDV																Neg PDV Percentile															
PDV Type								JB/PLC config								JB nominal															
JB maximum																JB abs max															
JB high water mark																JB low water mark															

Call Quality metrics sub-block

0								1								2								3							
0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
	R-LQ								R-CQ								R-CQ								R-CQ						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
	MOS-LQ								MOS-CQ								MOS-CQ								MOS-CQ						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
	R-LQ Ext In								R-LQ Ext Out								RFC3550 Payload								Media Type						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
	RxSigLev (IP)								RxNoiseLev (IP)								Local RERL								Remote RERL						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
	RxSigLev (Ext)								RxNoiseLev(Ext)								Metric Status								Metric Status						
+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-

3.2. Header

Implementations MUST send the Header block within each High Resolution Metrics report.

3.2.1. Block type

Three High Resolution VoIP Metrics blocks are defined

mmm = HR Metrics- Cumulative

mmm+1 = HR Metrics- Interval

mmm+2 = HR Metrics- Alert

The time interval associated with these report blocks is left to the implementation. Spacing of RTCP reports should be in accordance with [RFC3550](#). The specific timing of RTCP HR reports may be determined in response to an internally derived alert such as a threshold violation however the interval between RTCP HR reports must not be less than the minimum determined according to [RFC3550](#).

Note that interval data may be derived by subtracting successive cumulative reports, which provides increased tolerance to potential loss of RTCP reports.

3.2.2. Map field

A Map field indicates the optional sub-blocks present in this report. A 1 indicates that the sub-block is present, and a 0 that the block is absent. If present, the sub-blocks must be in the sequence defined in this document. The bits have the following definitions:

- 0 Burst/Gap Metrics block
- 1 Playout Metrics block
- 2 Concealed Seconds Metrics block
- 3 Call Quality Metrics
- 4-7 Reserved, set to 0

3.2.3. Block Length

The block length indicates the length of this report in 32 bit words and includes the header.

3.2.4. SSRC

The SSRC of the stream to which this report relates. The value of this field shall follow the rules defined in [RFC3550](#) with regard to the forwarding of RTP and RTCP messages.

3.2.5. Duration

The duration of time for which this report applies expressed in milliseconds. For cumulative reports this would be the call duration. For interval reports this would be the duration of the interval.

If the measured value exceeds 0xFFFFFFFF, the value 0xFFFFFFFF SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.3. Basic Loss/ Discard Metrics

The Basic Loss/Discard Metrics sub-block MUST be present.

This block reports the proportion of frames lost by the network and the proportion of frames discarded due to jitter.

For sample-based codecs such as G.711, a frame shall be defined as an RTP frame. For endpoints that incorporate jitter buffers capable of fractional frame discard the proportion of frames discarded MAY be determined on the basis of the proportion of samples discarded. If Voice Activity Detection is used then the proportion of frames lost and discarded shall be determined based on transmitted packets, i.e. frames that contained silence and were not transmitted shall not be considered.

A frame shall be regarded as lost if it fails to arrive within an implementation-specific time window. A frame that arrives within this time window but is too early or late to be played out shall be regarded as discarded. A frame shall be classified as one of received (or OK), discarded or lost.

The Loss and Discard metrics are determined after the effects of FEC, redundancy [[RFC2198](#)] or other similar process.

3.3.1. Loss Proportion

Proportion of frames lost within the network expressed as a binary fraction in 0:16 format. Duplicate frames shall be disregarded.

If the measured value exceeds 0xFFFFD, the value 0xFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.3.2. Discard Proportion

Proportion of voice frames received but discarded due to late or early arrival, expressed as a binary fraction in 0:16 format.

If the measured value exceeds 0xFFFFD, the value 0xFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.3.3. Number of frames expected

A count of the number of frames expected, estimated if necessary. If no frames have been received then this count shall be set to zero.

If the number expected exceeds 0xFFFFFFFFD, the value 0xFFFFFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.4. Burst/Gap metrics sub-block

The Burst/Gap metrics sub-block MAY be present and if present MUST be indicated in the Map field.

This block provides information on transient IP problems and is able to represent the combined effect of packet loss and packet discard. Burst/Gap metrics are typically used in Cumulative reports however MAY be used in Interval reports.

The definition of Burst and Gap is consistent with that defined in the [[RFC3611](#)] VoIP Metrics block, with the clarification that Loss

and Discard are defined in terms of frames (as described in [Section 3.3](#) above). To accomodate the range of jitter buffer algorithms and packet discard logic that may be used by implementors, the method used to distinguish between bursts and gaps may be an equivalent method to that defined in [RFC3611](#). The method used SHOULD produce the same result as that defined in [RFC3611](#) for conditions of burst packet loss, but MAY produce different results for conditions of time varying jitter.

If Voice Activity Detection is used the Burst and Gap Duration shall be determined as if silence frames had been sent, i.e. a period of silence in excess of Gmin frames MUST terminate a burst condition.

The Burst/Gap Metrics sub-block contains the following elements.

[3.4.1.](#) Threshold

The Threshold is equivalent to Gmin in [RFC3611](#), i.e. the number of successive frames that must be received and not discarded prior to and following a lost or discarded frame in order for this lost or discarded frame to be regarded as part of a gap.

[3.4.2.](#) Burst Duration (ms)

The average duration of a burst of lost and discarded frames.

If the measured value exceeds 0xFFFFFD, the value 0xFFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFF SHOULD be reported.

[3.4.3.](#) Gap Duration (ms)

The average duration of periods between bursts.

If the measured value exceeds 0xFFFFFDD, the value 0xFFFFFDE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFF SHOULD be reported.

[3.4.4.](#) Burst Loss/Discard Proportion

The proportion of Lost and Discarded frames during Bursts expressed as a binary fraction expressed in 0:16 format.

If the measured value exceeds 0xFFFF, the value 0xFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.4.5. Gap Loss/Discard Proportion

The proportion of Lost and Discarded frames during Gaps expressed as a binary fraction expressed in 0:16 format.

If the measured value exceeds 0xFFFD, the value 0xFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.5. Playout Metrics sub-block

The Playout Duration metrics sub-block MAY be present and if present MUST be indicated in the Map field.

At any instant, the audio output at a receiver may be classified as either 'normal' or 'concealed'. 'Normal' refers to playout of audio payload received from the remote end, and also includes locally generated signals such as announcements, tones and comfort noise. Concealment refers to playout of locally-generated signals used to mask the impact of network impairments or to reduce the audibility of jitter buffer adaptations.

This sub-block accounts for the source of the output audio, in millisecond units. The on-time and active speech playout durations allow calculation of the voice activity fraction. The on-time, and concealment durations allow calculation of concealment ratios. This sub-block distinguishes between reactive (due to effective packet loss) and proactive (due to buffer adaptation) concealment.

3.5.1. On-time Playout Duration

'On-time' playout is the uninterrupted, in-sequence playout of valid decoded audio information originating from the remote endpoint. This includes comfort noise during periods of remote talker silence, if VAD is used, and locally generated or regenerated tones and announcements.

An equivalent definition is that on-time playout is playout of any signal other than those used for concealment.

On-time playout duration MUST include both speech and silence intervals, whether VAD is used or not. This duration is reported in millisecond units.

If the measured value exceeds 0xFFFFFFFFD, the value 0xFFFFFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.5.2. On-time Active Speech Playout Duration

The duration, in milliseconds, of the on-time playout duration corresponding to playout of active speech signals, if known.

In the absence of silence suppression, on-time active speech playout equals on-time playout ([Section 3.5.1](#)).

If the measured value exceeds 0xFFFFFFFF, the value 0xFFFFFFFF SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.5.3. Loss Concealment Duration

The duration, in milliseconds, of audio playout corresponding to Loss-type concealment.

Loss-type concealment is reactive insertion or deletion of samples in the audio playout stream due to effective frame loss at the audio decoder. "Effective frame loss" is the event in which a frame of coded audio is simply not present at the audio decoder when required. In this case, substitute audio samples are generally formed, at the decoder or elsewhere, to reduce audible impairment.

Only loss-type concealment is necessary to form Concealed and Severely Concealed Seconds counts, in [Section 3.6](#).

If the measured value exceeds 0xFFFFFFFF, the value 0xFFFFFFFF SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.5.4. Buffer Adjustment Concealment Duration (optional)

The duration, in milliseconds, of audio playout corresponding to Buffer Adjustment-type concealment, if known.

If the measured value exceeds 0xFFFFFFFF, the value 0xFFFFFFFF SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

Buffer Adjustment-type concealment is proactive or controlled insertion or deletion of samples in the audio playout stream due to jitter buffer adaptation, re-sizing or re-centering decisions within the endpoint.

Because this insertion is controlled, rather than occurring randomly in response to losses, it is typically less audible than loss-type concealment ([Section 3.5.3](#)). For example, jitter buffer adaptation

events may be constrained to occur during periods of talker silence, in which case only silence duration is affected, or sophisticated time-stretching methods for insertion/deletion during favorable periods in active speech may be employed. For these reasons, buffer adjustment-type concealment MAY be exempted from inclusion in calculations of Concealed Seconds and Severely Concealed Seconds.

However, an implementation SHOULD include buffer-type concealment in counts of Concealed Seconds and Severely Concealed Seconds if the event occurs at an 'inopportune' moment, with an emergency or large, immediate adaptation during active speech, or for unsophisticated adaptation during speech without regard for the underlying signal, in which cases the assumption of low-audibility cannot hold. In other words, jitter buffer adaptation events which may be presumed to be audible SHOULD be included in Concealed Seconds and Severely Concealed Seconds counts.

Concealment events which cannot be classified as Buffer Adjustment-type MUST be classified as Loss-type.

3.6. Concealed Seconds metrics sub-block

The Concealed Seconds metrics sub-block MAY be present and if present MUST be indicated in the Map field.

This sub-block provides a description of potentially audible impairments due to lost and discarded packets at the endpoint, expressed on a time basis analogous to a traditional PSTN T1/E1 errored seconds metric.

The following metrics are based on successive one second intervals as declared by a local clock. This local clock does NOT need to be synchronized to any external time reference. The starting time of this clock is unspecified. Note that this implies that the same loss pattern could result in slightly different count values, depending on where the losses occur relative to the particular one-second demarcation points. For example, two loss events occurring 50ms apart could result in either one concealed second or two, depending on the particular 1000 ms boundaries used.

The seconds in this sub-block are not necessarily calendar seconds. At the tail end of a call, periods of time of less than 1000ms shall be incorporated into these counts if they exceed 500ms and shall be disregarded if they are less than 500ms.

3.6.1. Unimpaired Seconds

A count of the number of unimpaired Seconds that have occurred.

An unimpaired Second is defined as a continuous period of 1000ms during which no frame loss or discard due to late arrival has occurred. Every second in a call must be classified as either OK or Concealed.

Normal playout of comfort noise or other silence concealment signal during periods of talker silence, if VAD is used, shall be counted as unimpaired seconds.

If the measured value exceeds 0xFFFFFFFF, the value 0xFFFFFFFF SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.6.2. Concealed Seconds

A count of the number of Concealed Seconds that have occurred.

A Concealed Second is defined as a continuous period of 1000ms during which any frame loss or discard due to late arrival has occurred.

Equivalently, a concealed second is one in which some Loss-type concealment (defined in [Section 3.6](#)) has occurred. Buffer adjustment-type concealment SHALL not cause Concealed Seconds to be incremented, with the following exception. An implementation MAY cause Concealed Seconds to be incremented for 'emergency' buffer adjustments made during talkspurts.

For clarification, the count of Concealed Seconds MUST include the count of Severely Concealed Seconds.

If the measured value exceeds 0xFFFFFFFF, the value 0xFFFFFFFF SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFFFFFF SHOULD be reported.

3.6.3. Severely Concealed Seconds

A count of the number of Severely Concealed Seconds.

A Severely Concealed Second is defined as a non-overlapping period of 1000 ms during which the cumulative amount of time that has been subject to frame loss or discard due to late arrival, exceeds the SCS Threshold.

If the measured value exceeds 0xFFFF, the value 0xFFFF SHOULD be

reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.6.4. SCS Threshold

The SCS Threshold defines the amount of time corresponding to lost or discarded frames that must occur within a one second period in order for the second to be classified as a Severely Concealed Second. This is expressed in milliseconds and hence can represent a range of 0.1 to 25.5 percent loss or discard.

A default threshold of 50ms (5% effective frame loss per second) is suggested.

3.7. Delay and Packet Delay Variation (PDV) metrics sub-block

The Delay and PDV metrics sub-block MUST be present. This sub-block contains a number of parameters related to overall delay (latency), delay variation and the current jitter buffer configuration.

3.7.1. Network Round Trip Delay (ms)

The Network Round Trip Delay is the most recently measured value of the RTP-to-RTP interface round trip delay, typically determined using RTCP SR/RR.

If the measured value exceeds 0xFFFFD, the value 0xFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.7.2. End System Delay (ms)

The End System Delay is the internal round trip delay within the reporting endpoint, calculated using the nominal value of the jitter buffer delay plus the accumulation/ encoding and decoding / playout delay associated with the codec being used.

If the measured or estimated value exceeds 0xFFFFD, the value 0xFFFFE SHOULD be reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.7.3. External Delay (ms)

The External Network Delay parameter indicates external network round trip delay through cellular, satellite or other types of network with significant delay impact, if known.

If the measured value exceeds 0xFFFFD, the value 0xFFFFE SHOULD be

reported to indicate an over-range measurement. If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

If the external network is IP based then this parameter is typically determined using RTCP SR/RR. If the external network delay is known and does not vary materially then this value may be provisioned.

Where there is any ambiguity in assigning a delay contribution to one of the three metrics of Network Round Trip Delay, End System Delay, and External Delay, the following guidance is provided. The objective is that the sum of the three metrics SHOULD approximate as closely as possible to the sum of the delays "mouth to ear". Each significant source of delay SHOULD be counted in one, and only one, of the three metrics.

For slow links where packet serialisation delays are significant, delays should be referenced to the same point within the packet for both send and receive interfaces, e.g. the delay should be measured from the time at which the first bit of a packet leaves the send interface, to the time at which the first bit of a packet arrives at the receive interface. Definitions of delay which use different reference points on the packet at different interfaces, e.g. "first bit sent to last bit received", are likely to lead to errors from double-counting the serialisation delay when adding contributions.

3.7.4. PDV/Jitter Metrics

Jitter metrics defined are:

3.7.4.1. Mean PDV

For MAPDV this value is generated according to ITU-T G.1020. For interval reports the MAPDV value is reset at the start of the interval.

For PPDV the value reported is the value of $J(i)$ calculated according to [RFC3550](#) at the time the report is generated.

(16 bit, S11:4 format) expressed in milliseconds

If the measured value is more negative than -2047.9375 (the value which would be coded as 0x8001), the value 0x8000 SHOULD be reported to indicate an over-range negative measurement. If the measured value is more positive than +2047.8125 (the value which would be coded as 0x7FFD), the value 0x7FFE SHOULD be reported to indicate an over-range positive measurement. If the measurement is unavailable, the value 0x7FFF SHOULD be reported.

3.7.4.2. Positive Threshold/Peak PDV

The PDV associated with the Positive PDV percentile (16 bit, S11:4 format) expressed in milliseconds. The term Positive is associated with packets arriving later than the expected time.

If the measured value is more negative than -2047.9375 (the value which would be coded as 0x8001), the value 0x8000 SHOULD be reported to indicate an over-range negative measurement. If the measured value is more positive than +2047.8125 (the value which would be coded as 0x7FFD), the value 0x7FFE SHOULD be reported to indicate an over-range positive measurement. If the measurement is unavailable, the value 0x7FFF SHOULD be reported.

3.7.4.3. Negative Threshold/Peak PDV

The PDV associated with the Negative PDV percentile (16 bit, S11:4 format) expressed in milliseconds. The term Negative is associated with packets arriving earlier than the expected time.

If the measured value is more negative than -2047.9375 (the value which would be coded as 0x8001), the value 0x8000 SHOULD be reported to indicate an over-range negative measurement. If the measured value is more positive than +2047.8125 (the value which would be coded as 0x7FFD), the value 0x7FFE SHOULD be reported to indicate an over-range positive measurement. If the measurement is unavailable, the value 0x7FFF SHOULD be reported.

3.7.4.4. Positive PDV Percentile

The percentage of packets on the call for which individual packet delays were less than the Positive Threshold PDV expressed in 8:8 format.

If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.7.4.5. Negative PDV Percentile

The percentage of packets on the call for which individual packet delays were more than the Negative Threshold PDV expressed in 8:8 format.

If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

If the PDV Type indicated is IPDV and the Positive and Negative PDV Percentiles are set to 100.0 then the Positive and Negative

Threshold/Peak PDV values are the peak values measured during the reporting interval (which may be from the start of the call for cumulative reports). In this case, the difference between the Positive and Negative Threshold/Peak values defines the range of IPDV.

3.7.5. PDV Type

Indicates the type of algorithm used to calculate PDV:

0: PPDV according to [[RFC3550](#)],

1: MAPDV according to [[G.1020](#)],

2: IPDV according to [[Y.1540](#)]

Other values reserved

For example:-

(a) To report PPDV ([RFC3550](#)):

Threshold PDV = FFFF (Undefined); PDV Percentile = FFFF (Undefined);
PDV type = 0 (PPDV)

causes PPDV to be reported in the Mean PDV field.

(b) To report MAPDV ([G.1020](#)):

Pos Threshold PDV = 50.0; Pos PDV Percentile = 95.3; Neg Threshold PDV = 50.0 (note - implies -50ms); Neg PDV Percentile = 98.4; PDV type = 1 (MAPDV)

causes average MAPDV to be reported in the Mean PDV field.

Note that implementations may either fix the reported percentile and calculate the associated PDV level OR may fix a threshold PDV level and calculate the associated percentile. From a practical implementation perspective it is simpler to use the second of these approaches (except of course in the extreme case of a 100% percentile).

IPDV, according to [Y.1540](#) is the difference in delay between the i-th packet and the first packet of the stream. If the sending and receiving clocks are not synchronized, this metric includes the effect of relative timing drift.

3.7.6. Jitter Buffer / PLC Configuration

Indicates the configuration of the jitter buffer and the type of PLC algorithm in use.

bits 0-3

- 0 = silence insertion
- 1 = simple replay, no attenuation
- 2 = simple replay, with attenuation
- 3 = enhanced
- Other values reserved

bits 4-7

- 0 = Fixed jitter buffer
- 1 = Adaptive jitter buffer
- Other values reserved

3.7.7. Jitter Buffer Size parameters

Current nominal, maximum and absolute maximum jitter buffer size expressed in milliseconds, as defined in [[RFC3611](#)].

3.8. Call Quality Metrics sub-block

The Call Quality Metrics sub-block MAY be present and if present MUST be indicated in the Header Map field. This sub-block reports call quality metrics and estimates of signal, noise and echo levels.

Signal, noise and echo metrics should be long term averages and should not be instantaneous values.

3.8.1. Listening and Conversation Quality R Factors - R-LQ, R-CQ

Expresses listening and conversational quality in terms of R factor, a 0-120 scaled parameter in 8:8 format. The algorithm used to calculate R factor MAY be defined in the RTCP HR Configuration block (see [Section 4](#)).

If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.8.2. Listening and Conversation Quality MOS - MOS-LQ, MOS-CQ

Expresses listening and conversational quality in terms of MOS, a 1-5 scaled parameter in 8:8 format. The algorithm used to calculate MOS MAY be defined in the RTCP HR Configuration block.

Note that R factors and MOS scores may be defined for both narrow and wide-band VoIP calls. R Factors are continuous for narrow and wideband, hence the R factor for a wideband call may be higher than that for a narrowband call. MOS scores are scaled relative to reference conditions and hence both narrow and wideband MOS occupy the same 1-5 scale; this can lead to a wideband MOS being lower than a narrowband MOS even though the listening quality may be higher.

If the measurement is unavailable, the value 0xFFFF SHOULD be reported.

3.8.3. R-LQ Ext In and Out

These parameters provide call quality information for external networks - for example an external PCM or cellular network - or for a reporting call quality from the "other" side of a transcoding device or mixer - for example a conference bridge.

R-LQ Ext In - measured by this endpoint for incoming connection on "other" side of this endpoint. A 0-120 scaled parameter in 7:1 format. If the measurement is unavailable, the value 0xFF SHOULD be reported.

R-LQ Ext Out - copied from RTCP XR message received from remote endpoint on "other" side of this endpoint A 0-120 scaled parameter in 7:1 format. If the measurement is unavailable, the value 0xFF SHOULD be reported.

e.g. Phone A <---> Bridge <---> Phone B

In XR message from Bridge to Phone A:-

- R-LQ = quality for PhoneA ----> Bridge path
- R-LQ-ExtIn = quality for Bridge <---- Phone B path
- R-LQ-ExtOut = quality for Bridge -----> Phone B path

This allows PhoneA to assess

(i) received quality from the combination of

R-LQ measured at A and R-LQ-ExtIn reported by the Bridge to A

(ii) remote endpoint quality from the combination of

R-LQ reported by the Bridge and R-LQ-ExtOut reported by the Bridge

3.8.4. RFC3550 RTP Payload Type

The RTP Payload type field - as per [RFC3551](#) and <http://www.iana.org/assignments/rtp-parameters>. Where payload type is dynamically assigned, the correlation tag mechanism ([Section 4.2](#)) may be used to find signalling-layer information which binds the Payload Type to a specific codec.

3.8.5. Media Type

Media type -

- 0 = No media present
- 1 = Narrowband audio
- 2 = Wideband audio

3.8.6. Received Signal and Noise Levels - IP side

The received signal level during talkspurts and the noise level expressed in dBm0, for the decoded packet stream. Expressed in S7 format.

If the measured value is more negative than -127 dBm0 (the value which would be coded as 0x81), the value 0x8000 SHOULD be reported to indicate an over-range negative measurement. If the measured value is more positive than +125 dBm0 (the value which would be coded as 0x7D), the value 0x7E SHOULD be reported to indicate an over-range positive measurement. If the measurement is unavailable, the value 0x7F SHOULD be reported. Either over-range is extremely unlikely for such a power measurement.

3.8.7. Received Signal and Noise Levels - External

The received signal level during talkspurts and the noise level expressed in dBm0, for the PCM side of a gateway, audio input from a handset or decoded packet stream for an IP-to-IP gateway. Expressed in S7 format.

If the measured value is more negative than -127 dBm0 (the value which would be coded as 0x81), the value 0x8000 SHOULD be reported to indicate an over-range negative measurement. If the measured value is more positive than +125 dBm0 (the value which would be coded as 0x7D), the value 0x7E SHOULD be reported to indicate an over-range positive measurement. If the measurement is unavailable, the value 0x7F SHOULD be reported. Either over-range is extremely unlikely for such a power measurement.

3.8.8. Local and Remote Residual Echo Return Loss

The Local and Remote Residual Echo Return Loss (RERL) expressed in dB. The Local RERL is the echo level that would be reflected into the IP path due to line echo on the circuit switched element side of this IP endpoint if a gateway or acoustic echo if a handset or wireless terminal. Expressed in S7 format.

The Remote RERL is the echo level that would be reflected into the remote IP endpoint from the network "behind" it, and would typically be measured at and reported from the remote endpoint. This value is included as it may be used in calculating the R-CQ and MOS-CQ values expressed in this report block. Expressed in S7 format.

If the measured RERL value is more negative than -127 dB (the value which would be coded as 0x81), the value 0x8000 SHOULD be reported to indicate an over-range negative measurement. If the measured value is more positive than +125 dB (the value which would be coded as 0x7D), the value 0x7E SHOULD be reported to indicate an over-range positive measurement. If the measurement is unavailable, the value 0x7F SHOULD be reported. Either over-range is extremely unlikely for such a power ratio measurement.

3.8.9. Metric Status

Indicates the source of parameter values used in call quality calculation:

Bit	Description	Source
0-1	Local IP side Signal/Noise Levels measured on the incoming decoded VoIP stream to this endpoint	
	00 = assumed	
	01 = measured for this call	
	10 = measured across multiple calls on this port	
	11 = measured across multiple ports	
2-3	Remote IP side Signal/Noise Levels reported by the remote IP endpoint through RTCP XR or equivalent	
	00 = assumed	
	01 = measured for this call	
	10 = measured across multiple calls on this port	
	11 = measured across multiple ports	

- 4-5 Local Trunk side Signal/Noise Levels measured on the incoming PCM, Audio or non-IP side of this endpoint

00 = assumed
01 = measured for this call
10 = measured across multiple calls on this port
11 = measured across multiple ports

- 6-7 Local Echo level measured in the incoming line/ trunk/ handset direction at this endpoint after the effects of echo cancellation

00 = assumed
01 = measured for this call
10 = measured across multiple calls on this port
11 = measured across multiple ports

- 8 Remote Echo level measured in the incoming line/ trunk/ handset direction at the remote endpoint after the effects of echo cancellation and reported to this endpoint via RTCP XR or equivalent.

0 = assumed
1 = reported from remote endpoint

9-15 Reserved

For example, if this endpoint is "C" in the diagram below then the following definitions would apply.

Endpoint B <-----RTP-----> Gateway C <-----PCM-----> D
"Remote" "Local" "Trunk/PCM/External"

Reporting endpoint is "C"

Local IP side signal/noise metrics relate to signal/noise levels from decoded RTP packets received by C from B

Remote IP side signal/noise metrics relate to signal/noise levels from decoded RTP packets received by B from C, and reported by B to C through RTCP XR or RTCP HR VoIP Metrics blocks

Local Trunk side signal/noise metrics relate to signal/noise levels from the PCM signal received by C from D

Local Echo level relates to the proportion of the signal passing from B to C to D that is reflected back to C at some point between C and D or on the far side of D. This would typically be electrical echo or acoustic echo.

Remote Echo level relates to the proportion of the signal passing from D to C to B that is reflected back to B at some point between B and the user. This echo level is typically measured at B and reported to C via RTCP XR or RTCP HR VoIP Metrics blocks.

4. RTCP HR Configuration Block

This block type provides a flexible means to describe the algorithms used for call quality calculation and other data. This block need only be exchanged occasionally, for example sent once at the start of a call.

Header sub-block

```

0           1           2           3
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      BT=N      |      Map      |      block length      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     SSRC of source         |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Correlation Tag sub-block

```

0           1           2           3
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Tag Type      | Tag length  | Correlation Tag...      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     ... Correlation Tag   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Algorithm sub-block

```

0           1           2           3
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Alg type      | Descriptor len| Algorithm descriptor... |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     ... Algorithm descriptor |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

4.1. Header

Implementations MUST send the Header block within each RTCP HR Configuration report.

4.1.1. Block type

One RTCP HR Configuration block is defined

mmm+3 = RTCP HR Configuration Block

The time interval associated with these report blocks is left to the

implementation. Spacing of RTCP reports should be in accordance with [RFC3550](#) however the specific timing of RTCP HR reports may be determined in response to an internally derived alert such as a threshold crossing.

[4.1.2.](#) Map field

A Map field indicates the optional sub-blocks present in this report. A '1' indicates that the sub-block is present, and a '0' that the block is absent. If present, the sub-blocks must be in the sequence defined in this document. The bits have the following definitions:

- 0 Correlation Tag
- 1 Algorithm Descriptor 1
- 2 Algorithm Descriptor 2
- 3 Algorithm Descriptor 3
- 4 Algorithm Descriptor 4
- 5-7 Reserved, set to '0'

[4.1.3.](#) Block Length

The block length indicates the length of this report in 32 bit words and includes the header and any extension octets.

[4.1.4.](#) SSRC

The SSRC of the stream to which this report relates.

[4.2.](#) Correlation Tag

The Correlation Tag sub-block MAY be present and if present MUST be indicated in the map field. This tag facilitates the correlation of the high resolution VoIP metrics report blocks with other call-related data, session-related data or endpoint data.

An example use case is for an endpoint to convey its version of a call identifier or a global call identifier via this tag. A flow measurement tool (sniffer) that is not call-aware can then forward the RTCP-HR reports along with this correlation tag to network management. Network management can then use this tag to correlate this report with other diagnostic information such as call detail records.

The Tag Type indicates the use of the correlation tag. The following values are defined:

0: IMS Charging Identity (ICID) subfield of the P-Charging-Vector header specified in [[RFC3455](#)].

1: Globally unique ID as specified in ITU-T H.225.0 (Table 20/H.225.0) [[H.225.0](#)].

2: Conference Identifier, per ITU-T H.225.0 (Table 20/H.225.0 [[H.225.0](#)]).

3: SIP Call-ID as defined in [[RFC3261](#)].

4: PacketCable Billing Call ID (BCID) [[PKTMMS](#)].

5: Text string using the US-ASCII character set [[ASCII](#)].

6: Octet string.

7-255: Future growth.

Although the intent of this RFC is to list all currently known values of usable correlation tags, it is possible that new values may be defined in the future. An IANA registry of correlation tags is recommended.

The tag length indicates the overall length of the sub-block in 32 bit words and includes the tag type and length fields.

[4.3.](#) Algorithm descriptor

The Algorithm Type sub-block MAY be present and if present MUST be indicated in the map field

The Algorithm Type is a bit field which indicates which algorithm is being described. The bits are defined as:-

Bit 0:	MOS-LQ Algorithm
Bit 1:	MOS-CQ Algorithm
Bit 2:	R-LQ Algorithm
Bit 3:	R-CQ Algorithm
Bit 4-7:	Reserved and set to '0'

The descriptor length gives the overall length of the descriptor in 32 bit words and includes the algorithm descriptor and length fields.

The algorithm descriptor is a text field that contains the description or name of the algorithm. If the algorithm name is shorter than the length of the field then the trailing octets must be set to 0x00.

For example, an implementation may report:

Algorithm descriptor = 0xF0 - R and MOS algorithms
Descriptor length = 3 - 3 words
Descriptor = "P.564" 0x00 - description

Call quality estimation algorithms may be defined for listening or conversational quality MOS or R factor.

5. SDP Signalling

This section defines Session Description Protocol (SDP) [RFC4566] signalling for RTCP HR that can be employed by applications that utilize SDP. The approach follows the design pattern established for RTCP XR in [RFC3611], with modifications arising from the use in RTCP HR of multiple sub-blocks of a single RTCP XR block, rather than the multiple top-level RTCP XR blocks as used in RTCP XR. This SDP signalling is defined to be used either by applications that implement the SDP Offer/Answer model [RFC3264] or by applications that use SDP to describe media and transport configurations in connection with such protocols as the Session Announcement Protocol (SAP) [RFC2974] or the Real Time Streaming Protocol (RTSP) [RFC2326]. There exist other potential signalling methods that are not defined here. RTCP HR blocks MAY be used without prior signalling. This is consistent with the rules governing other RTCP packet types, as described in [RFC3550]. An example in which signalling would not be used is an application that always requires the use of RTCP HR. However, for applications that are configured at session initiation, the use of some type of signalling is recommended.

Note that, although the use of SDP signalling for RTCP HR may be optional, if used, it MUST be used as defined here. If SDP signalling is used in an environment where RTCP HR is only implemented by some fraction of the participants, the ones not implementing RTCP HR will ignore the SDP attribute.

5.1. The SDP Attributes

This section defines two new SDP attributes "rtcp-hr-span" and "rtcp-hr-subblk" that can be used to signal participants in a media session how they should use RTCP HR.

The two SDP attributes are defined below in Augmented Backus-Naur Form (ABNF) [RFC5234]. They are both session and media level attributes. When specified at session level, they apply to all media level blocks in the session. Any media level specification MUST replace a session level specification, if one is present, for that media block.

```
rtcp-hr-span-attr = "a=rtcp-hr-span:"  
                  [hr-span-format *(SP hr-span-format)] CRLF  
hr-span-format = "cumulative"  
               / "interval"  
               / "alert"
```



```

rtcp-hr-subblk-attrb = "a=rtcp-hr-subblk:" hr-subblk-formats
hr-subblk-formats = [hr-subblk-format *(SP hr-subblk-format)] CRLF
hr-subblk-format =   loss
                    / burst-gap
                    / playout
                    / conceal
                    / delay
                    / quality

```

```

loss           = "loss"
burst-gap      = "burst-gap"
playout        = "playout"
conceal        = "conceal" ["=" thresh]
delay          = "delay" [ "," pdvtype ] [ "," nspec "," pspec ]
quality        = "quality"

```

```

thresh         = 1*DIGIT           ; threshold for SCS (ms)
pdvtype        = "pdv=" 0           ; ppdv RFC 3550
                  / 1             ; mapdv ITU-T G.1020
                  / 2             ; ipdv  ITU-T Y.1540
nspec          = "nthr=" fixpoint   ; negative threshold PDV (ms)
                  / "npc=" fixpoint ; negative PDV percentile
pspec          = "pthr=" fixpoint   ; positive threshold PDV (ms)
                  / "ppc=" fixpoint ; positive PDV percentile

```

```

fixpoint       = 1*DIGIT "." 1*DIGIT ; fixed point decimal
DIGIT          = %x30-39
CRLF           = %d13.10

```

The Header sub-block is mandatory in an RTCP HR report block, and hence does not appear as a possible value for "hr-subblk-format". The Basic Loss/Delay sub-block is also mandatory. However if SDP requests that an RTCP HR report block should be sent, then the value "loss" MUST be present in the attribute list of hr-subblk-format, in order to avoid potential ambiguity in the meaning of an empty list. The Delay and Packet Delay Variation (PDV) Metrics sub-block is also mandatory, but it requires parameters to control its behaviour. If SDP requests that an RTCP HR report block should be sent, the value "delay" MUST appear in the list of hr-subblk-format, together with its parameters.

The "rtcp-hr-subblk" attributes parameter list MAY be empty. This is useful in cases in which an application needs to signal that it understands the SDP signalling but does not wish to avail itself of RTCP HR functionality. For example, an application in a SIP

controlled session could signal that it wishes to stop using all HR subblocks by removing all applicable SDP parameters in a re-INVITE message that it sends. If HR subblocks are not to be used at all from the beginning of a session, it is RECOMMENDED that none of the "rtcp-hr" attributes be supplied.

When the "rtcp-hr-subblk" attribute is present but not populated with any parameters, even those for mandatory sub-blocks ("loss", "delay"), participants SHOULD NOT send any RTCP HR information. This means that inclusion of an "rtcp-hr-subblk" attribute without any parameters tells a participant that it SHOULD NOT send any optional RTCP HR subblocks at all. The purpose is to conserve bandwidth. There are, however, contexts in which it makes sense to send an RTCP HR block in the absence of a parameter signalling its use. For instance, an application might be designed so as to send certain report blocks without negotiation, while using SDP signalling to negotiate the use of other blocks.

When the "rtcp-hr-subblk" attribute is present and populated with at least the parameters for mandatory sub-blocks ("loss" and "delay") participants SHOULD send mandatory sub-blocks but SHOULD NOT send optional RTCP HR subblocks other than the ones indicated by the parameters.

5.2. Usage in Offer/Answer

In the Offer/Answer context [[RFC3264](#)], the interpretation of SDP signalling for RTCP HR packets depends upon the direction attribute that is signaled: "recvonly", "sendrecv", or "sendonly" [[RFC4566](#)]. If no direction attribute is supplied, then "sendrecv" is assumed. This section applies only to unicast media streams, except where noted.

For "sendonly" and "sendrecv" media stream offers, the answerer SHOULD send the corresponding RTCP HR subblocks. For "sendrecv" offers, the answerer MAY include the attributes in its response, and specify any parameters in order to request that the offerer send the corresponding XR blocks. The offerer SHOULD send these blocks. For "recvonly" media stream offers, the offerer's use of the "rtcp-hr-" attributes indicates that the offerer is capable of sending the corresponding RTCP HR sub-blocks. If the answerer responds with the set of two "rtcp-hr-" attributes, the offerer SHOULD send RTCP HR subblocks. For multicast media streams, the inclusion of "rtcp-hr-" attributes means that every media recipient SHOULD send the corresponding HR sub-blocks. If a participant receives an SDP offer and understands the "rtcp-hr-" attributes but does not wish to implement RTCP HR functionality offered, its answer SHOULD include "rtcp-hr-" attributes without parameters. By doing so, the party

declares that, at a minimum, it is capable of understanding the signalling.

5.3. Usage Outside of Offer/Answer

SDP can be employed outside of the Offer/Answer context, for instance for multimedia sessions that are announced through the Session Announcement Protocol (SAP) [[RFC2974](#)], or streamed through the Real Time Streaming Protocol (RTSP) [[RFC2326](#)]. The signalling model is simpler, as the sender does not negotiate parameters, but the functionality expected from specifying the "rtcp-hr-" attributes is the same as in Offer/Answer.

When a parameter is specified for the "rtcp-hr-subblk" attribute associated with a media stream, the receiver of that stream SHOULD send the corresponding RTCP HR block.

6. Practical Applications

6.1. Overview

The objective of this section is to identify a number of cases in which there could potentially be some ambiguity in the application of the report blocks defined above or some exceptions to the defined operation of the metrics.

6.2. Supplementary Services: Call Hold and Transfer

6.2.1. General

Supplementary services are under control of call/session control protocols like SIP. Such signalling protocols are acting also as "non-RTP means" (definition see clause 3 of [[RFC3550](#)]) in such service scenarios.

The "northbound" served user instance for RTCP HR data is typically "co-located" to the served user instance of the call/session control protocol controlling the supplementary service. This allows to correlate in principle supplementary service control events with RTCP HR measurements in such network elements (like a SIP UA, SIP proxy, application server, etc.).

Thus, the correlation between RTP/RTCP session control and supplementary service control allows basically the minimization of potential ambiguity.

Below sub-clause providing some additional notes dependent on specific supplementary services.

6.2.2. Supplementary Service: Call Transfer

A successful call transfer means that an initial call between A and B is transferred to a call between C and B. This means that the RTP end system A is "replaced" by RTP end system C, accompanied by all correspondent changes in a RTP/RTCP endpoint (e.g., SSRC for A "replaced" by SSRC for B).

In the scope of RTCP HR, it is therefore recommended to consider the two call phases (1st phase: call A-B, 2nd phase: call C-B) as separate measurement phases. Separate measurement phases could be e.g. based on interval metrics and the derivation of call phase-individual cumulative metrics by the "northbound" served user instance of RTCP HR, or by "resetting" the cumulative metrics in each call phase.

6.2.3. Supplementary Service: Call Hold

Call hold enables the served (holding) user A to put user B (with whom user A has an active call) into a hold condition (held user) and subsequently to retrieve that user again. During this hold condition, user B may be provided with media on hold (MoH). The served (holding) user A may perform other actions while user B is being held, e.g. consulting with another user C.

In the scope of RTCP HR, it is recommended to consider the different call phases firstly as separate measurement phases (see also 8.2.2).

6.3. Bitrate efficiency improvements: VAD/Silence Suppression based on Voice Activity Detection (VAD) Elimination

A VoIP call is either enabled or disabled for silence suppression. This is typically a call-individual configuration parameter, negotiated during call establishment phase, and not changed anymore during the remaining call phase.

An enabled silence suppression mode is basically affecting almost all high resolution VoIP metrics.

The "northbound" served user instance of RTCP HR may require access to the information, whether silence suppression was enabled or disabled for that call, in order to indicate that mode of operation in the VoIP measurement data.

6.4. Endpoint configuration changes mid-call

An endpoint relates to an RTP end system, which can be either a) located in VoIP user/terminal equipment (e.g. SIP UA), or b) located in VoIP gateway equipment (e.g. PSTN-to-RTP H.248 media gateway), or c) located in VoIP media server equipment.

6.4.1. Changes due to mid-call transitions between different voice codec types

Voice codec type changes are reflected in RTP payload type changes, which are visible in the Call Quality metrics sub-block ([Section 3.8.4](#)).

6.4.2. Changes due to mid-call transitions from VoIP to RTP-based VBDoIP

There might be mid-call transitions from VoIP to dedicated modes of operation for voiceband data services support in case that at least one RTP end system is located in type (b) equipment. Mode

transitions should be again reflected in RTP payload type changes in case of RTP-based VBD transport (e.g. like ITU-T Rec. V.152 [[V.152](#)] for VBDoIP).

Details are for further study.

[6.4.3.](#) Changes due to mid-call transitions from VoIP to non-RTP -based VBDoIP

UDPTL/UDP based realtime facsimile according ITU-T Rec. T.38 is an example for RTP-less transport of facsimile/modem signals. Any mid-call transition to T.38 would inherently terminated the RTP/RTCP session, thus the measurement phase.

Details are for further study.

[6.5.](#) SSRC changes mid-call

An SSRC change may be e.g. the consequence of a mid-call transport address change.

Details are for further study.

7. IANA Considerations

This document defines a series of new RTCP Extended Report (XR) block types within the existing Internet Assigned Numbers Authority (IANA) registry of RTP RTCP XR block types. In addition, this document defines the need for an IANA registry of correlation tag types ([Section 4.3](#))

8. Security Considerations

RTCP reports can contain sensitive information since they can provide information about the nature and duration of a session established between two endpoints. As a result, any third party wishing to obtain this information should be properly authenticated and the information transferred securely.

9. Contributors

The authors gratefully acknowledge the comments and contributions made by Jim Frauenthal, Mike Ramalho, Paul Jones, Claus Dahm, Bob Biskner, Mohamed Mostafa, Tom Hock, Albert Higashi, Shane Holthaus, Amit Arora, Bruce Adams, Albrecht Schwarz, Keith Lantz, Randy Ethier, Philip Arden, Ravi Raviraj and Hideaki Yamada.

10. Informative References

- [ASCII] ANSI, "Coded Character Set--7-Bit American Standard Code for Information Interchange, ANSI X3.4-1986", 1986.
- [G.1020] ITU-T, "ITU-T Rec. G.1020, Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks", November 2003.
- [G.1020AnnexA] ITU-T, "Amendment 1 to ITU-T Rec G.1020 Annex A, VoIP Gateway specific reference points and performance parameters", May 2004.
- [H.225.0] ITU-T, "ITU-T Rec. H.225.0, Call signalling protocols and media stream packetization for packet-based multimedia communication systems", July 2003.
- [PKTMMS] PacketCable(TM), "PacketCable(TM) Multimedia Specification, PKT-SP-MM-I02-040930", September 2004.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [RFC 2119](#), [BCP 14](#), March 1997.
- [RFC2198] Perkins, C., "RTP Payload for Redundant Audio Data", [RFC 2198](#), September 1997.
- [RFC2326] Schulzrinne, H., "Real Time Streaming Protocol (RTSP)", [RFC 2326](#), April 1998.
- [RFC2974] Handley, M., "Session Announcement Protocol", [RFC 2974](#), October 2000.
- [RFC3261] Rosenberg, J., "SIP: Session Initiation Protocol", [RFC 3261](#), June 2002.
- [RFC3264] Rosenberg, J., "An Offer/Answer Model with the Session Description Protocol (SDP)", [RFC 3264](#), June 2002.
- [RFC3455] Garcia-Martin, M., "Private Header (P-Header) Extensions to the Session Initiation Protocol (SIP) for the 3rd-Generation Partnership Project (3GPP)", [RFC 3455](#), January 2003.
- [RFC3550] Schulzrinne, H., "RTP: A Transport Protocol for Real-Time Applications", [RFC 3550](#), July 2003.
- [RFC3611] Friedman, T., "RTP Control Protocol Extended Reports (RTCP

XR)", [RFC 3611](#), November 2003.

- [RFC4566] Handley, M., "SDP: Session Description Protocol", [RFC 4566](#), July 2006.
- [RFC5234] Crocker, D., "Augmented BNF for Syntax Specifications: ABNF", [RFC 5234](#), STD 68, January 2008.
- [V.152] ITU-T, "ITU-T Rec. V.152, Procedures for supporting voice-band data over IP networks", January 2005.
- [Y.1540] ITU-T, "ITU-T Rec. Y.1540, Internet protocol data communication service -- IP packet transfer and availability performance parameters", December 2002.

Authors' Addresses

Alan Clark
Telchemy Incorporated
2905 Premiere Parkway, Suite 280
Adastral Park
Martlesham Heath
Duluth, GA 30097

Email: alan@telchemy.com

Geoff Hunt
BT
Orion 1 PP9
Adastral Park
Martlesham Heath
Ipswich, Suffolk IP5 3RE
United Kingdom

Phone: +44 1473 608325
Email: geoff.hunt@bt.com

Amy Pendleton
Nortel
2380 Performance Drive
Richardson, TX 75081

Email: aspen@nortel.com

Rajesh Kumar
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134

Email: rkumar@cisco.com

Kevin Connor
Cisco Systems
5590 Whitehorn Way
Blaine, WA 98230

Email: kconnor@cisco.com

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

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

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

Intellectual Property

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

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

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

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).

