Internet Engineering Task Force
Internet-Draft
Expires: 29 August 2006

A. Clark
Telchemy Incorporated
A. Pendleton
Nortel
R. Kumar
K. Connor
Cisco Systems
March 2006

RTCP HR - High Resolution VoIP Metrics Report Blocks
draft-clark-avt-rtcphr-01

Status of this Memo

Copyright Notice

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

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.

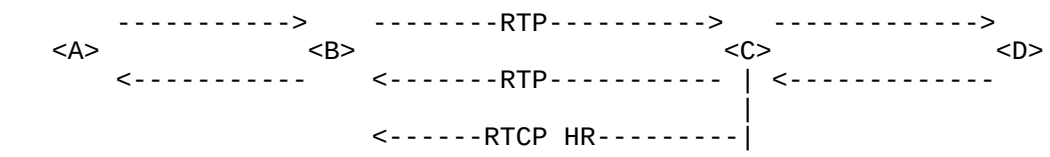
For certain types of VoIP service it is desirable to report VoIP performance metrics to a higher resolution than provided in the RFC3611 VoIP Metrics block or RFC3550 Receiver Reports. The report blocks described in this section provide both interval based and cumulative metrics with a higher resolution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

2. Definitions

2.1 Local and Remote IP Endpoints

A report block produced per this draft is normally produced by the endpoint of an RTP stream, and relates to the quality of the received RTP stream and impairments that may affect the RTP payload. The diagram below illustrates the potential end and mid points that may be involved in this process. Within the context of this example, endpoint "C" is the reporting endpoint and the RTCP HR report relates to the RTP stream from "B" to "C". The other points "A", "B" and "D" could potentially be generating RTCP XR or RTCP HR reports.



With respect to RTCP HR report blocks generated by "C" in relation to the RTP stream from "B" to "C":

(i) The term "External" is used to relate to the network connected to the other side of "this" endpoint (i.e. to the connection from "C" to "D").

(ii) The Local IP Endpoint is "this" endpoint (i.e. "C").

(iii) The Remote IP Endpoint is the source for the RTP stream terminated at this endpoint, and for which packet/frame related metrics apply (i.e. "B").

(iv) The Remote External Endpoint is the remote endpoint on the "external" side of this endpoint (i.e. "D").

(v) The endpoint on the external side of the Remote IP Endpoint (i.e. "A") does not have a specific term applied to it, however note that some metrics may apply to the "A" to "B" connection.

If this model is applied to a trunking gateway, then the "B" to "C" connection would be the IP network or Voice over IP connection whereas the "C" to "D" connection would typically be the PCM trunk.

If this model is applied to a conference bridge or session/border controller then the "B" to "C" connection would be a Voice over IP connection and the "C" to "D" connection may also be Voice over IP. In this scenario, "C" could be generating RTCP HR or XR reports in both directions, i.e. sent to "D" for the "D" to "C" RTP stream and sent to "B" for the "B" to "C" RTP stream.

2.2 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.

2.3 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 time was lost through packet loss and discard. The term Severely Concealed Seconds defines a count of seconds during which the proportion of time lost through packet loss and discard exceeds a specified threshold.

2.4 Numeric formats

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

S X:Y, where S indicates a signed representation,
X the number of bits prior to the decimal place and
Y the number of bits after the decimal place.

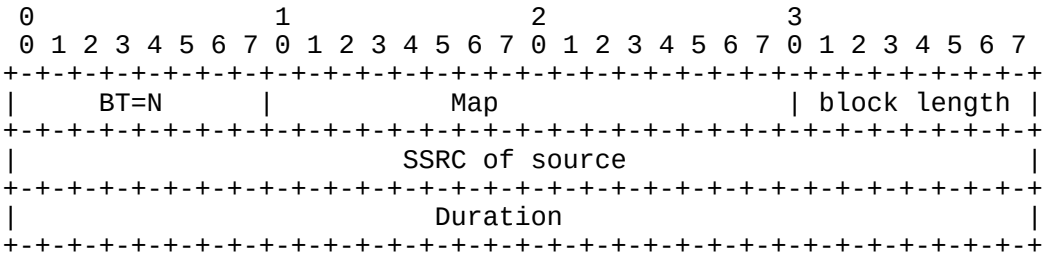
Hence 8:8 represents an unsigned number in the range 0.0, 0.0039 to 255.996. S7:8 would represent the range -127.996 to +127.996.

3 High Resolution VoIP Metrics Report Block

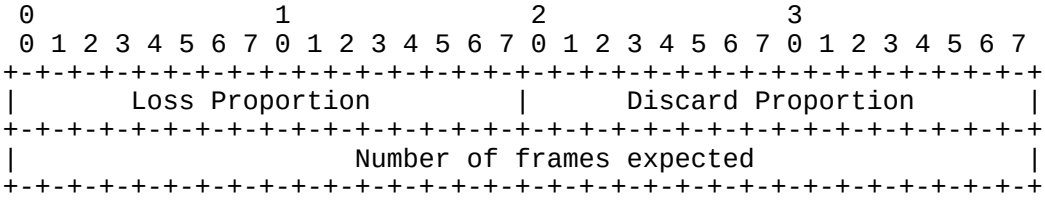
3.1 Block Description

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

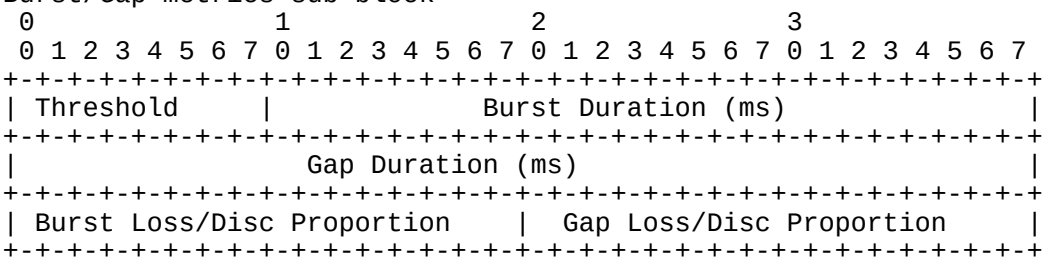
Header sub-block



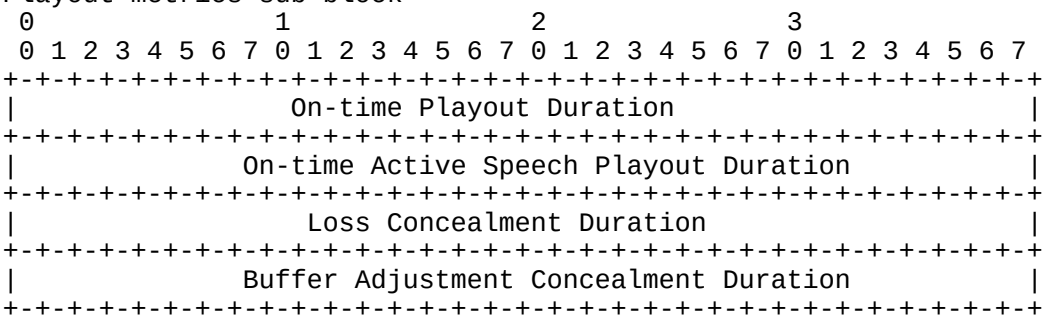
Basic Loss/Discard Metrics sub-block



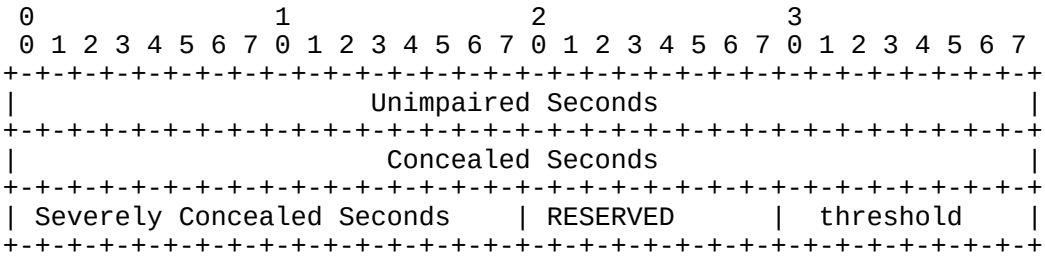
Burst/Gap metrics sub-block



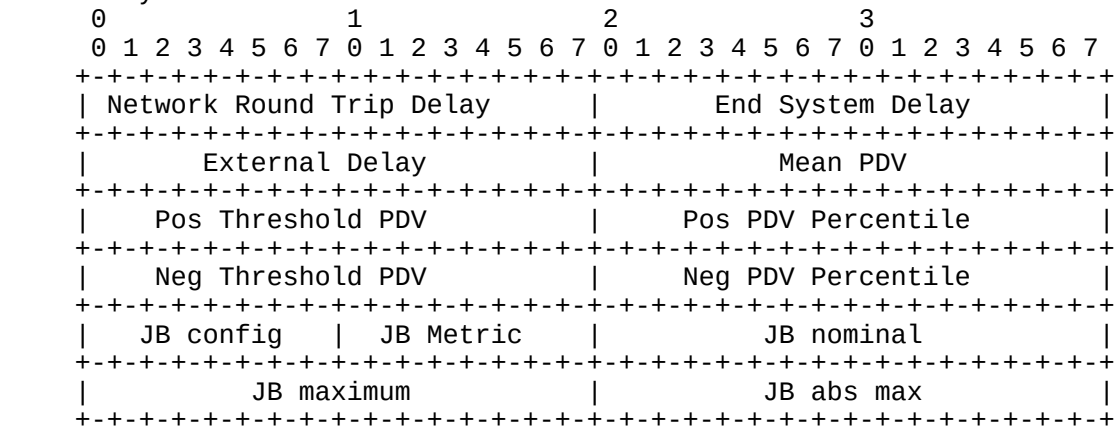
Playout metrics sub-block



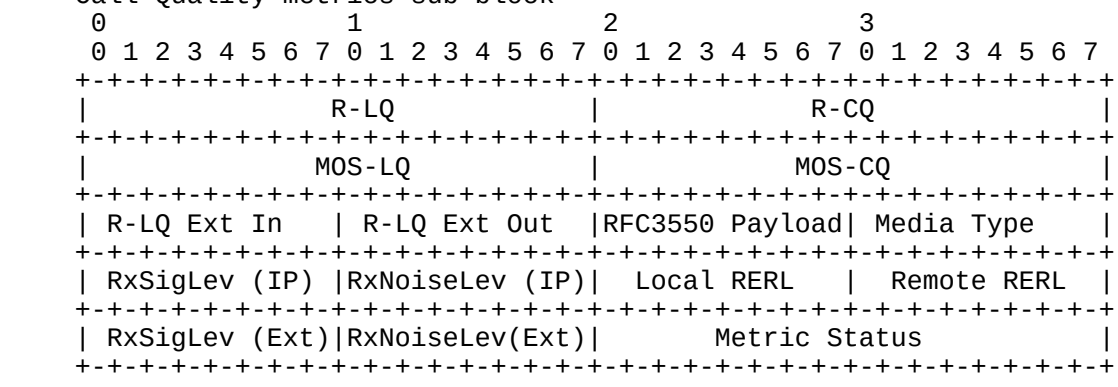
Concealed Seconds metrics sub-block



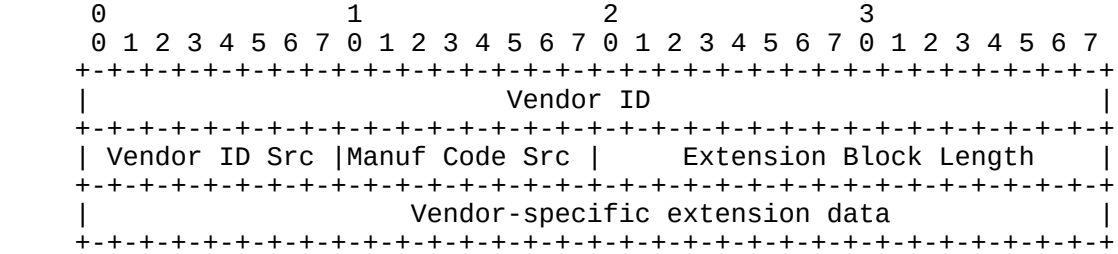
Delay and PDV metrics sub-block



Call Quality metrics sub-block



Vendor specific extension sub-block



3.2 Header

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

3.2.1 Block type

Nine High Resolution VoIP Metrics blocks are defined

mmm = HR Metrics- Cumulative, Locally Generated
mmm+1 = HR Metrics- Cumulative, Relayed from Remote IP Endpoint
mmm+2 = HR Metrics- Cumulative, Relayed from Remote Ext Endpoint
mmm+3 = HR Metrics- Cumulative,TBD

mmm+4 = HR Metrics- Interval, Locally Generated
mmm+5 = HR Metrics- Interval, Relayed from Remote IP Endpoint
mmm+6 = HR Metrics- Interval, Relayed from Remote Ext Endpoint
mmm+7 = HR Metrics- Interval, TBD
mmm+7 = HR Metrics- Alert, Locally Generated
mmm+9 = HR Metrics- Alert, Relayed from Remote IP Endpoint
mmm+10 = HR Metrics- Alert, Relayed from Remote Ext Endpoint
mmm+11 = HR Metrics- Alert, TBD

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 spacing of RTCP HR reports must not exceed that defined in 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 Vendor specific extension block
5-15 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 and any extension octets.

3.2.4 SSRC

The SSRC of the stream to which this report relates.

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.

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

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

3.3.3 Dead connection detection
If no RTP, SID or RTP no-op packets have been received for a pre-specified time interval(for example ten seconds) then an RTCP HR dead connection indication MUST be sent. This time interval may be changed during the connection, for example being longer at the start of a call. Dead connection detection may be temporarily disabled during silence periods if VAD is used.
A dead connection is indicated by setting both the frames lost and frames discarded fields above to 0xFFFF (equivalent to indicating 99.99% of packets have been both lost and discarded). Receipt of an RTCP HR block with either field set to a value other than 0xFFFF shall indicate that the remote endpoint HAS received some valid RTP packets.

3.3.4 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, however if a dead connection is indicated then this count shall be regarded as undefined.

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 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.

3.4.3 Gap Duration (ms)
The average duration of periods between bursts.

3.4.4 Burst Loss/Discard Rate
The proportion of Lost and Discarded frames during Bursts expressed as a binary fraction expressed in 0:16 format.

3.4.5 Gap Loss/Discard Rate
The proportion of Lost and Discarded frames during Gaps expressed as a binary fraction expressed in 0:16 format.

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.

3.5.2 On-time Active Speech Playout Duration (optional)

The duration, in milliseconds, of the on-time playout duration corresponding to playout of active speech signals, if known. If not known, then this field is set to all ones (0x FFFF FFFF).

In the absence of silence suppression, on-time active speech playout equals on-time playout (section 3.5.1).

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.

3.5.4 Buffer Adjustment Concealment Duration (optional)

The duration, in milliseconds, of audio playout corresponding to Buffer Adjustment-type concealment, if known. If not known, then this field is set to all ones (0x FFFF FFFF).

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 on this call.

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.

3.6.2 Concealed Seconds

A count of the number of Concealed Seconds that have occurred on this call.

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.

3.6.3 Severely Concealed Seconds
A count of the number of Severely Concealed Seconds that have occurred on this call.
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.

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/ 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.

[Editor's note - need to add high and low water marks]

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 no measured delay is available then this field shall be set to 0xFFFF

[Note - potentially add one way delay]

3.7.2 End System Delay (ms)
The End System Delay is the internal round trip delay within the 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.

3.7.3 External Network 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. A value of 0xFFFF shall indicate that the delay is unknown.
If the external network is VoIP 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.

3.7.4 PDV/ Jitter Metrics
Jitter metrics defined are:

(i) Mean PDV - for cumulative reports this is a running average of PDV and for interval reports this is an interval average (16 bit, S11:4 format) expressed in milliseconds

(ii) Threshold PDV - the PDV associated with the PDV percentile (16 bit, S11:4 format) expressed in milliseconds, both positive and negative thresholds are given

(iii) PDV Percentile - the percentage of packets on the call for which individual packet delays were outside the Threshold PDV, expressed in 8:8 format. Both positive and negative percentiles are given.

3.7.5 PDV Type
Indicates the type of algorithm used to calculate PDV:

PPDV (0) according to RFC3550,
MAPDV (1) according to ITU-T G.1020,
IPDV (2) according to ITU-T Y.1540
Other values reserved

For example:-

(a) To report PPDV (RFC3550):
Mean PDV = PPDV
Threshold PDV = Undefined (FF FF)
PDV Percentile = Undefined (FFF)
PDV type = 0 PPDV

(b) To report 95th percentile MAPDV (G.1020):
Mean PDV = average MAPDV
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

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.

3.7.6 Jitter Buffer and 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
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 Voice/Data Quality Metric Algorithm block.

3.8.2 Listening and Conversation Quality MOS Scores
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 Voice/Data Quality Metric Algorithm block (see Section 4).

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.

3.8.3 R-LQ External In and Out.
R-LQ External In - measured by this endpoint for incoming connection on "other" side of this endpoint

R-LQ External Out - copied from RTCP XR message received from remote endpoint on "other" side of this endpoint

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

If the RFC3550 Payload type indicates a dynamic payload then it is strongly recommended that the Extended Payload Descriptor sub-block (Section 4.4) be sent in order that mid-stream management systems can identify the codec type used.

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.

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.

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.

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.

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 "B"

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.

3.9 Vendor Extension sub-block

One or more Vendor Extension sub-blocks MAY be present. Their presence MUST be indicated in the Header Map field. Note that the map field does not indicate the number of vendor extension sub-blocks. This must be deduced from the length of the overall report block and the lengths of the Vendor Extension sub-blocks.

Each vendor extension sub-block consists of an extension header and vendor-specific extension data. The extension header has the form Vendor ID, Vendor ID Source and Extension Block Length. The Extension Block Length is defined as including these extension header and extension data octets but does not include any subsequent vendor extension sub-blocks. An implementation can skip over a vendor extension sub-block that it does not understand.

The Vendor ID Source field indicates whether the four-octet Vendor ID is based on ITU T.35 or is an IANA private enterprise number. The designated values for these options are 1 and 2 respectively. If the Vendor ID Source field is assigned a value of 0, then all of the fields in the vendor extension sub-block with the exception of the Vendor ID Source field and the Extension Block Length field have proprietary definitions.

If the Vendor ID is based on ITU-T Recommendation T.35, its first two octets either contain a country code from Annex A of ITU-T Rec. T.35 followed by 0x00, or an escape code of 0xFF followed by a country code from Annex B of ITU-T Rec. T.35. The next two octets comprise a terminal provider code allocated by a national assignment authority (http://people.itu.int/~campos/t35/t35db.htm). This field is padded with leading zeros if necessary. If a vendor has multiple terminal provider codes in different registries (e.g. the H-series, T-series and V-series registries for the USA), then this provenance shall be indicated in the Manufacturer Code Source field. Possible values of the Manufacturer Code Source field are:

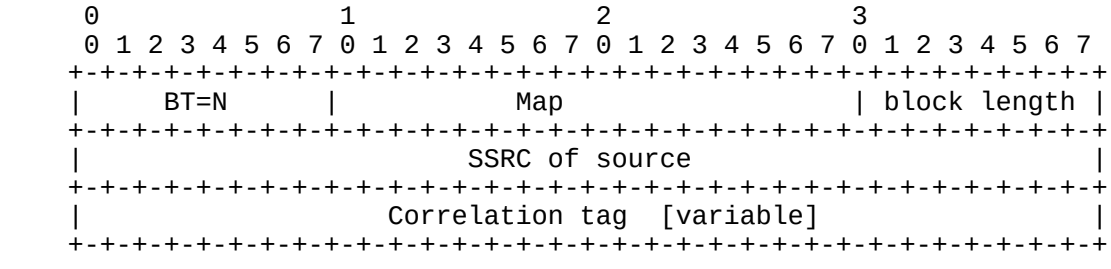
0 Unspecified/don't care (may be used if there is no conflict)
1 G-series registry
2 H-series registry
3 T-series registry
4 V-series registry
5-255 Reserved.

If the four-octet Vendor ID is an IANA private enterprise number, then it is padded with leading zeros as necessary. Current private enterprise numbers (www.iana.org/assignments/enterprise-numbers) can be accommodated within two octets. The additional two octets provide for future growth.

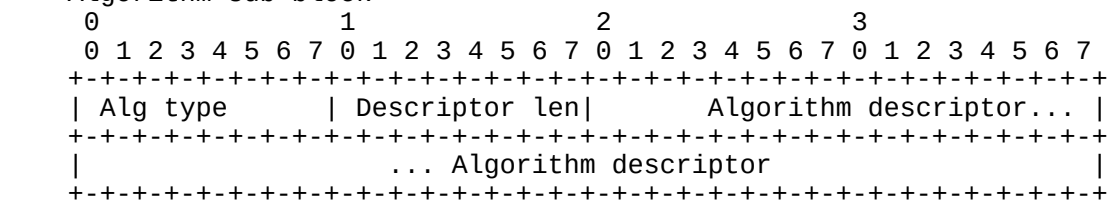
4. RTCP HR Configuration Block

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.

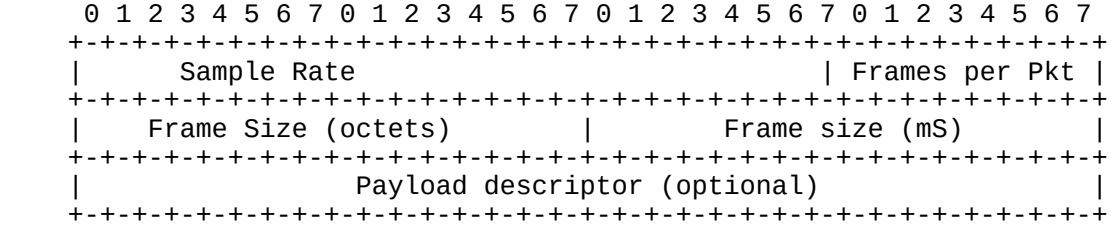
Header sub-block



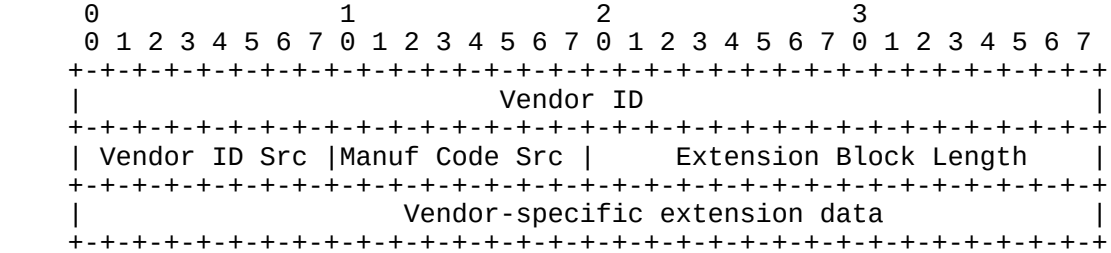
Algorithm sub-block



Extended Payload Descriptor sub-block



Vendor Specific Extensions sub-block



4.1 Header

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

4.2.1 Block type
One RTCP HR Configuration blocks is defined

mmm+12 = 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.

4.2.2 Map field
An 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 Algorithm Descriptor 1
1 Algorithm Descriptor 2
2 Algorithm Descriptor 3
3 Algorithm Descriptor 4
4 Algorithm Descriptor 5
5 Algorithm Descriptor 6
6 Algorithm Descriptor 7
7 Algorithm Descriptor 8
8 Payload Descriptor
9 Vendor Specific Extension
10-15 Reserved, set to '0'

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

4.2.4 SSRC
The SSRC of the stream to which this report relates.

4.2.5 Correlation tag
The correlation tag facilitates the correlation of this report block with other call or session related data or endpoint data.
[Note - make this field variable length]

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 = "Alg X" 0x00     - description
```

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

4.4 Extended payload description field

The Extended Payload Description sub-block MAY be present and if present it MUST be indicated in the Header Map field. This sub -block provides a detailed description of the payload format used.

4.3.1 Sample Rate (Hz)
The sample rate used for this codec.

4.3.2 Frames per Packet
The number of RTP frames incorporated into an IP packet

4.3.3 Frame Size (octets)
The size of the RTP payload expressed in octets

4.3.4 Frame Size (ms)
The time interval represented by each RTP payload.

4.3.5 Payload descriptor
A textual representation of the type of codec, intended to assist test and management systems.

4.4 Vendor Extension field

One or more vendor specific extension blocks may be added, as defined in Section 3.10

5. Practical Applications

5.1 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.

5.2 Call Hold and Transfer

5.3 VAD/Silence Elimination

5.4 Endpoint configuration changes mid-call

5.5 SSRC changes mid-call

6. Summary

7. IANA Considerations

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, Kevin Connor, Paul Jones, Claus Dahm, Bob Biskner, Mohamed Mostafa, Tom Hock, Albert Higashi, Shane Holthaus, Amit Arora, Bruce Adams, Geoff Hunt, Albrecht Schwarz, Keith Lantz and Randy Ethier.

10. Informative References

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

[2] Schulzrinne, H., Casner, S., Frederick, R. and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[3] Friedman, T., Caceres, R. and A. Clark, "RTP Control Protocol Extended Reports (RTCP XR)", RFC 3611, November 2003.

[4] "Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks" ITU-T Rec. G.1020, November 2003

[5] Annex A to G.1020: "VoIP Gateway specific reference points and performance parameters", Amendment 1 to ITU-T Rec. G.1020 May 2004

[6] "Internet protocol data communication service - IP packet transfer and availability performance parameters:, ITU-T Rec. Y.1540, December 2002

Authors' Addresses

Alan Clark
Telchemy Incorporated
3360 Martins Farm Road, Suite 200
Suwanee, GA 30024
Email: alan@telchemy.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

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.