Video Over RTP Enhancements and Testing Methods

Huy Nguyen   Sebastian Tran

Advisors: Dr. Amiruzzaman & Dr. Ngo
West Chester University of Pennsylvania

Abstract

RTP, or real time protocol, is designed for internet transmission with minimal latency. Applications that require near millisecond latency will implement RTP to achieve such goals. Common applications for RTP include real time video streams, online conferencing applications, and voice over IP. To enable additional enhancements and modifications, RTP allows for extensions to the header in order to transmit extra data with the video and audio streams. This study explores the feasibility of various modifications of this extension header and how this impacts data transfer performance.

First we attempt to implement sending and receiving Video QOS metrics inside of the RTP packet rather than through the RTP control stream. Second, we add a role based video streaming to ensure video is directed only to authorized users. In addition, we design a testing scheme to measure latency and other QOS metrics in order to compare base protocol designs against other modifications.

RTP Packet Header

Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	V		P	X		CC			M		Payload Type						Sequence Number															
32	Timestamp																															
64	SSRC identifier																															
96	CSRC identifier(s)																															
...	Profile-specific Extension Header Identifier																Extension Header Length															
...	Extension Data																															

RTP Packet Header Layout

V (Version): Indicates the version of the protocol
P (Padding): Tells the receiver that padding is enabled.
X (Extension): Tells the receiver that extension modifications to the header are enabled.
CC (CSRC Count): Stores the number of CSRC identifiers
M (Marker): Used to indicate to the receiver this is special packet
Payload Type: Indicates the type of data in the packet
Sequence Number: The packet number. Used to correct out of order transmission and detect packet loss
Timestamp: Used for synchronization of different streams
SSRC identifier: Determines who has sent the packet
CSRC identifier(s): Lists who has contributed data to the packet
Profile Specific Extension Header Identifier: Extension specific headers
Extension Header Length: Length of the extension data, including header
Extension Data: Extra profile specific data

Traditional Video and Audio Synchronization Over RTP

Generally, multiplexing video and audio RTP streams over the same session is not recommended. Therefore, audio and video streams are separated and sent as different sessions. Streams may also be sampled at different rates, which requires synchronization. This is done via occasionally announcing the common world clock time $T_{S_{SR}}$ with the corresponding RTP time stamp M_{SR}.

The time T_S an RTP packet was created can then be estimated as:

$$T_S = \frac{T_{S_{SR}} + (M - M_{SR})}{R}$$

Where R is the sampling frequency of the media form in hertz. A delay D between an audio and video stream is then simply:

$$D = D_{\text{audio}} - D_{\text{video}}$$

For the stream that is ahead, a synchronization delay is added to match the delay of the other stream. The offset, measured in media timestamp units, shifts playback for that amount, and the receiver will make these packets reside in the buffer to compensate for the delay.

$$\text{offset} = D * R$$

Testing Metrics

To test the efficacy of RTP transmissions under different network conditions, we measure these network quality metrics. Each has a unique effect on the quality of an RTP stream. To eliminate overhead of analyzing results during the stream, data will instead be collected and logged to be analyzed later.
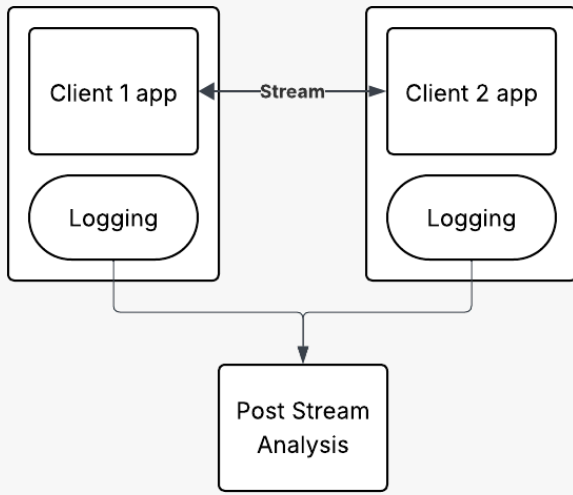
Transmission Delay

Transmission delay measures the time needed for the sender to push all bits onto the network.

$$t = \frac{L}{R}$$

where L is size of packet in bits, and R is bandwidth, measured in bits per second

Transmission delay has a critical role in real time systems, as it propogates itself as higher latency.

IP Packet Delay Variation (IPDV)

IPDV is the can be measured as the root mean square of packet variation samples.

A sample, k, measures the average of the absolute difference in latency between a successive n packets. Where R_i and S_i are the arrival times and sending time of the packet respectively

$$IPDV(k) = \frac{\sum_{i=n*(k-1)+2}^{k*n+1} |(R_i - R_{i-1}) - (S_i - S_{i-1})|}{n}$$

Applying the absolute value, rather than squaring, makes the sample average resistant to a small number of outliers, which are largely unnoticeable to the user.

The root mean square of the samples can then be calculated as:

$$IPDV_{RMS}(I) = \sqrt{\frac{\sum_{k=N*(I-1)+1}^{I*N} IPDV^2(K)}{N}}$$

Root mean squares provides a measure of variation as a magnitude, unlike a standard deviation, which measures variation as a distance from the mean.

IPDV is noticeable in congested networks. Real time apps compete with other network services for bandwidth. This manifests itself as high packet variation, resulting in a stream that buffers often and offers a lower quality of experience for the user.
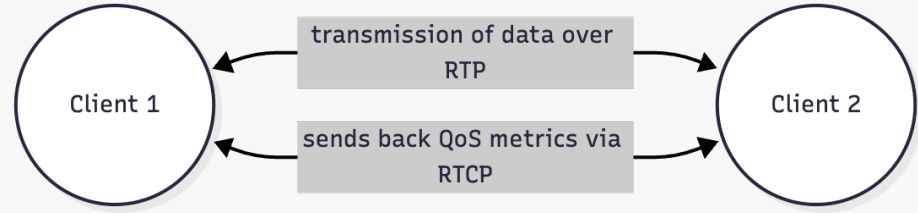
Packet Loss

Packet loss is the percentage of packets that are successfully received by a client

$$PL = \frac{\text{Packets Sent} - \text{Packets Received}}{\text{Packets Sent}} * 100\,\%$$

High packet loss causes RTP streams to be choppy and unresponsive. Poor network conditions and failed routing are the main causes of packet loss

Extensions

In a traditional RTP setup, data and QoS transmissions are separated into different transmissions on separate ports

With the extension header, it is possible to transmit both types of data in the header. This adds complexity to RTP in encoding and decoding, but has the benefit of reducing the overhead when separate protocols are used.

Jung and Seo have demonstrated it is possible to achieve synchronization of RTP streams without the need for the separate RTCP protocol. Their results have shown a significant decrease in synchronization time compared to traditional methods.

We highlight several RTCP functions that can be integrated with RTP, while still maintaining the adaptability of the protocol to wide use cases as designed.

Primarily, we will store NTP, number of packet lost, and inter-arrival jitter from the RTCP packet header sender report, which provides the necessary data to make QoE adjustments.

NTP (Network Time Protocol): Used as a standard reference time for synchronization of RTP streams.

Number of Packets Lost: Measures the number of RTP packets that had not been successfully delivered to the receiver. The sender can use this number to estimate the packet loss percent, and make adjustments.

Interarrival Jitter: Provides an estimate of the RTP packet data arrival time. True jitter will be calculated in our tests, but an estimate can provide a gauge of network quality.

The delay between packets, where S is the timestamp of the packet, and R is the received time, is:

$$D(i, j) = (R_j - R_i) - (S_j - S_i)$$

Jitter of a packet is estimated as then:

$$J(i) = J(i - 1) + (|D(i - 1, i)| - J(i - 1))/16$$

References

[1] Leopoldo Angrisani, Domenico Capriglione, Luigi Ferrigno, and Gianfranco Miele. Internet protocol packet delay variation measurements in communication networks: How to evaluate measurement uncertainty? *Measurement*, 46(7):2099–2109, 2013.

[2] Ankur Chadda. *Quality of service testing methodology*. PhD thesis, University of New Hampshire (Department of Computer Science), 2004.

[3] Tae-jun Jung and Kwang-deok Seo. A client-driven media synchronization mechanism for rtp packet-based video streaming. *Journal of Real-Time Image Processing*, 12(2):455–464, 2016.

[4] Colin Perkins. *Rtp: audio and video for the internet*. Addison-Wesley Professional, first edition, 2003.

[5] Henning Schulzrinne, Steven Casner, R Frederick, and Van Jacobson. Rfc3550: Rtp: A transport protocol for real-time applications, 2003.