Real-time Transport Protocol

From Wikipedia, the free encyclopedia

The **Real-time Transport Protocol** (**RTP**) is a network protocol for delivering audio and video over IP networks. RTP is used in communication and entertainment systems that involve streaming media, such as telephony, video teleconference applications including WebRTC, television services and web-based push-to-talk features.

RTP typically runs over User Datagram Protocol (UDP). RTP is used in conjunction with the RTP Control Protocol (RTCP). While RTP carries the media streams (e.g., audio and video), RTCP is used to monitor transmission statistics and quality of service (QoS) and aids synchronization of multiple streams. RTP is one of the technical foundations of Voice over IP and in this context is often used in conjunction with a signaling protocol such as the Session Initiation Protocol (SIP) which establishes connections across the network.

RTP was developed by the Audio-Video Transport Working Group of the Internet Engineering Task Force (IETF) and first published in 1996 as RFC 1889 which was then superseded by RFC 3550 in 2003.

Overview

RTP is designed for end-to-end, real-time transfer of streaming media. The protocol provides facilities for jitter compensation and detection of packet loss and out-of-order delivery, which are common especially during UDP transmissions on an IP network. RTP allows data transfer to multiple destinations through IP multicast. RTP is regarded as the primary standard for audio/video transport in IP networks and is used with an associated profile and payload format. The design of RTP is based on the architectural principle known as application-layer framing where protocol functions are implemented in the application as opposed to in the operating system's protocol stack.

Real-time multimedia streaming applications require timely delivery of information and often can tolerate some packet loss to achieve this goal. For example, loss of a packet in audio application may result in loss of a fraction of a second of audio data, which can be made unnoticeable with suitable error concealment algorithms. The Transmission Control Protocol (TCP), although standardized for RTP use, is not normally used in RTP applications because TCP favors reliability over timeliness. Instead the majority of the RTP implementations are built on the User Datagram Protocol (UDP). Other transport protocols specifically designed for multimedia sessions are SCTP and DCCP, although, as of 2012, they are not in widespread use.

RTP was developed by the Audio/Video Transport working group of the IETF standards organization. RTP is used in conjunction with other protocols such as H.323 and RTSP. The RTP specification describes two protocols: RTP and RTCP. RTP is used for the transfer of multimedia data, and the RTCP is used to periodically send control information and QoS parameters.

The data transfer protocol, RTP, carries real-time data. Information provided by this protocol includes timestamps (for synchronization), sequence numbers (for packet loss and reordering detection) and the payload format which indicates the encoded format of the data. The control protocol, RTCP, is used for quality of service (QoS) feedback and synchronization between the media streams. The bandwidth of RTCP traffic compared to RTP is small, typically around 5%.

RTP sessions are typically initiated between communicating peers using a signaling protocol, such as H.323, the Session Initiation Protocol (SIP), RTSP, or Jingle (XMPP). These protocols may use the Session Description Protocol to specify the parameters for the sessions.

An RTP session is established for each multimedia stream. Audio and video streams may use separate RTP sessions, enabling a receiver to selectively receive components of a particular stream. The specification recommends that RTP port numbers be chosen to be even and that each associated RTCP port be the next higher odd number. However, a single port is chosen for RTP and RTCP in applications that multiplex the protocols. RTP and RTCP typically use unprivileged UDP ports (1024 to 65535), but may also use other transport protocols, most notably, SCTP and DCCP, as the protocol design is transport independent.

RTP is used by real-time multimedia applications such as voice over IP, audio over IP, WebRTC and Internet Protocol television.

RTP Control Protocol

The **RTP Control Protocol** (**RTCP**) is a sister protocol of the Real-time Transport Protocol (RTP). Its basic functionality and packet structure is defined in RFC 3550. RTCP provides out-of-band statistics and control information for an RTP session. It partners with RTP in the delivery and packaging of multimedia data, but does not transport any media data itself.

The primary function of RTCP is to provide feedback on the quality of service (QoS) in media distribution by periodically sending statistics information such as transmitted octet and packet counts, packet loss, packet delay variation, and round-trip delay time to participants in a streaming multimedia session. An application may use this information to control quality of service parameters, perhaps by limiting flow, or using a different codec.

Protocol functions

Typically RTP will be sent on an even-numbered UDP port, with RTCP messages being sent over the next higher odd-numbered port.

RTCP itself does not provide any flow encryption or authentication methods. Such mechanisms may be implemented, for example, with the Secure Real-time Transport Protocol (SRTP) defined in RFC 3711.

RTCP provides basic functions expected to be implemented in all RTP sessions:

- The primary function of RTCP is to gather statistics on quality aspects of the media distribution during a session and transmit this data to the session media source and other session participants. Such information may be used by the source for adaptive media encoding (codec) and detection of transmission faults. If the session is carried over a multicast network, this permits non-intrusive session quality monitoring.

- RTCP provides canonical end-point identifiers (CNAME) to all session participants. Although a source identifier (SSRC) of an RTP stream is expected to be unique, the instantaneous binding of source identifiers to end-points may change during a session. The CNAME establishes unique identification of end-points across an application instance (multiple use of media tools) and for third-party monitoring.
- Provisioning of session control functions. RTCP is a convenient means to reach all session participants, whereas RTP itself is not. RTP is only transmitted by a media source.

RTCP reports are expected to be sent by all participants, even in a multicast session which may involve thousands of recipients. Such traffic will increase proportionally with the number of participants. Thus, to avoid network congestion, the protocol must include session bandwidth management. This is achieved by dynamically controlling the frequency of report transmissions. RTCP bandwidth usage should generally not exceed 5% of total session bandwidth. Furthermore, 25% of the RTCP bandwidth should be reserved to media sources at all times, so that in large conferences new participants can receive the CNAME identifiers of the senders without excessive delay.

The RTCP reporting interval is randomized to prevent unintended synchronization of reporting. The recommended minimum RTCP report interval per station is 5 seconds. Stations should not transmit RTCP reports more often than once every 5 seconds.