Transports for WebRTC
draft-ietf-rtcweb-transports-17

Abstract

This document describes the data transport protocols used by WebRTC, including the protocols used for interaction with intermediate boxes such as firewalls, relays and NAT boxes.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on April 29, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Introduction

WebRTC is a protocol suite aimed at real time multimedia exchange between browsers, and between browsers and other entities.

WebRTC is described in the WebRTC overview document, [I-D.ietf-rtcweb-overview], which also defines terminology used in this document, including the terms "WebRTC endpoint" and "WebRTC browser".
Terminology for RTP sources is taken from [RFC7656].

This document focuses on the data transport protocols that are used by conforming implementations, including the protocols used for interaction with intermediate boxes such as firewalls, relays and NAT boxes.

This protocol suite intends to satisfy the security considerations described in the WebRTC security documents, [I-D.ietf-rtcweb-security] and [I-D.ietf-rtcweb-security-arch].

This document describes requirements that apply to all WebRTC endpoints. When there are requirements that apply only to WebRTC browsers, this is called out explicitly.

2. Requirements language

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 [RFC2119].

3. Transport and Middlebox specification

3.1. System-provided interfaces

The protocol specifications used here assume that the following protocols are available to the implementations of the WebRTC protocols:

- UDP [RFC0768]. This is the protocol assumed by most protocol elements described.
- TCP [RFC0793]. This is used for HTTP/WebSockets, as well as for TURN/TLS and ICE-TCP.

For both protocols, IPv4 and IPv6 support is assumed.

For UDP, this specification assumes the ability to set the DSCP codepoint of the sockets opened on a per-packet basis, in order to achieve the prioritizations described in [I-D.ietf-rtcweb-qos] (see Section 4.2) when multiple media types are multiplexed. It does not assume that the DSCP codepoints will be honored, and does assume that they may be zeroed or changed, since this is a local configuration issue.

Platforms that do not give access to these interfaces will not be able to support a conforming WebRTC endpoint.
This specification does not assume that the implementation will have access to ICMP or raw IP.

The following protocols may be used, but can be implemented by a WebRTC endpoint, and are therefore not defined as "system-provided interfaces":

- TURN - Traversal Using Relays Around NAT, [RFC5766]
- STUN - Session Traversal Utilities for NAT, [RFC5389]
- ICE - Interactive Connectivity Establishment, [I-D.ietf-ice-rfc5245bis]
- TLS - Transport Layer Security, [RFC5246]
- DTLS - Datagram Transport Layer Security, [RFC6347].

### 3.2. Ability to use IPv4 and IPv6

Web applications running in a WebRTC browser MUST be able to utilize both IPv4 and IPv6 where available – that is, when two peers have only IPv4 connectivity to each other, or they have only IPv6 connectivity to each other, applications running in the WebRTC browser MUST be able to communicate.

When TURN is used, and the TURN server has IPv4 or IPv6 connectivity to the peer or the peer’s TURN server, candidates of the appropriate types MUST be supported. The "Happy Eyeballs" specification for ICE [I-D.ietf-mmusic-ice-dualstack-fairness] SHOULD be supported.

### 3.3. Usage of temporary IPv6 addresses

The IPv6 default address selection specification [RFC6724] specifies that temporary addresses [RFC4941] are to be preferred over permanent addresses. This is a change from the rules specified by [RFC3484]. For applications that select a single address, this is usually done by the IPV6_PREFER_SRC_TMP preference flag specified in [RFC5014]. However, this rule, which is intended to ensure that privacy-enhanced addresses are used in preference to static addresses, doesn’t have the right effect in ICE, where all addresses are gathered and therefore revealed to the application. Therefore, the following rule is applied instead:

When a WebRTC endpoint gathers all IPv6 addresses on its host, and both non-deprecated temporary addresses and permanent addresses of the same scope are present, the WebRTC endpoint SHOULD discard the permanent addresses before exposing addresses to the application or
using them in ICE. This is consistent with the default policy described in [RFC6724].

If some of the temporary IPv6 addresses, but not all, are marked deprecated, the WebRTC endpoint SHOULD discard the deprecated addresses, unless they are used by an ongoing connection. In an ICE restart, deprecated addresses that are currently in use MAY be retained.

### 3.4. Middle box related functions

The primary mechanism to deal with middle boxes is ICE, which is an appropriate way to deal with NAT boxes and firewalls that accept traffic from the inside, but only from the outside if it is in response to inside traffic (simple stateful firewalls).

ICE [I-D.ietf-ice-rfc5245bis] MUST be supported. The implementation MUST be a full ICE implementation, not ICE-Lite. A full ICE implementation allows interworking with both ICE and ICE-Lite implementations when they are deployed appropriately.

In order to deal with situations where both parties are behind NATs of the type that perform endpoint-dependent mapping (as defined in [RFC5128] section 2.4), TURN [RFC5766] MUST be supported.

WebRTC browsers MUST support configuration of STUN and TURN servers, both from browser configuration and from an application.

Note that there is other work around STUN and TURN sever discovery and management, including [I-D.ietch-rtcweb-return] for server discovery, as well as [I-D.ietf-tram-turn-server-discovery].

In order to deal with firewalls that block all UDP traffic, the mode of TURN that uses TCP between the WebRTC endpoint and the TURN server MUST be supported, and the mode of TURN that uses TLS over TCP between the WebRTC endpoint and the TURN server MUST be supported. See [RFC5766] section 2.1 for details.

In order to deal with situations where one party is on an IPv4 network and the other party is on an IPv6 network, TURN extensions for IPv6 [RFC6156] MUST be supported.

TURN TCP candidates, where the connection from the WebRTC endpoint’s TURN server to the peer is a TCP connection, [RFC6062] MAY be supported.

However, such candidates are not seen as providing any significant benefit, for the following reasons.
First, use of TURN TCP candidates would only be relevant in cases which both peers are required to use TCP to establish a PeerConnection.

Second, that use case is supported in a different way by both sides establishing UDP relay candidates using TURN over TCP to connect to their respective relay servers.

Third, using TCP between the WebRTC endpoint’s TURN server and the peer may result in more performance problems than using UDP, e.g. due to head of line blocking.

ICE-TCP candidates [RFC6544] MUST be supported; this may allow applications to communicate to peers with public IP addresses across UDP-blocking firewalls without using a TURN server.

If TCP connections are used, RTP framing according to [RFC4571] MUST be used for all packets. This includes the RTP packets, DTLS packets used to carry data channels, and STUN connectivity check packets.

The ALTERNATE-SERVER mechanism specified in [RFC5389] (STUN) section 11 (300 Try Alternate) MUST be supported.

The WebRTC endpoint MAY support accessing the Internet through an HTTP proxy. If it does so, it MUST include the "ALPN" header as specified in [RFC7639], and proxy authentication as described in Section 4.3.6 of [RFC7231] and [RFC7235] MUST also be supported.

3.5. Transport protocols implemented

For transport of media, secure RTP is used. The details of the profile of RTP used are described in "RTP Usage" [I-D.ietf-rtcweb-rtp-usage], which mandates the use of a circuit breaker [I-D.ietf-avtcore-rtp-circuit-breakers] and congestion control (see [I-D.ietf-rmcat-cc-requirements] for further guidance).

Key exchange MUST be done using DTLS-SRTP, as described in [I-D.ietf-rtcweb-security-arch].

For data transport over the WebRTC data channel [I-D.ietf-rtcweb-data-channel], WebRTC endpoints MUST support SCTP over DTLS over ICE. This encapsulation is specified in [I-D.ietf-tsvwg-sctp-dtls-encaps]. Negotiation of this transport in SDP is defined in [I-D.ietf-mmusic-sctp-sdp]. The SCTP extension for NDATA, [I-D.ietf-tsvwg-sctp-ndata], MUST be supported.

The setup protocol for WebRTC data channels described in [I-D.ietf-rtcweb-data-protocol] MUST be supported.
Note: DTLS-SRTP as defined in [RFC5764] section 6.7.1 defines the interaction between DTLS and ICE ([I-D.ietf-ice-rfc5245bis]). The effect of this specification is that all ICE candidate pairs associated with a single component are part of the same DTLS association. Thus, there will only be one DTLS handshake even if there are multiple valid candidate pairs.

WebRTC endpoints MUST support multiplexing of DTLS and RTP over the same port pair, as described in the DTLS-SRTP specification [RFC5764], section 5.1.2, with clarifications in [I-D.ietf-avtcore-rfc5764-mux-fixes]. All application layer protocol payloads over this DTLS connection are SCTP packets.

Protocol identification MUST be supplied as part of the DTLS handshake, as specified in [I-D.ietf-rtcweb-alpn].

4. Media Prioritization

The WebRTC prioritization model is that the application tells the WebRTC endpoint about the priority of media and data that is controlled from the API.

In this context, a "flow" is used for the units that are given a specific priority through the WebRTC API.

For media, a "media flow", which can be an "audio flow" or a "video flow", is what [RFC7656] calls a "media source", which results in a "source RTP stream" and one or more "redundancy RTP streams". This specification does not describe prioritization between the RTP streams that come from a single "media source".

All media flows in WebRTC are assumed to be interactive, as defined in [RFC4594]; there is no browser API support for indicating whether media is interactive or non-interactive.

A "data flow" is the outgoing data on a single WebRTC data channel.

The priority associated with a media flow or data flow is classified as "very-low", "low", "medium" or "high". There are only four priority levels at the API.

The priority settings affect two pieces of behavior: Packet send sequence decisions and packet markings. Each is described in its own section below.
4.1. Local prioritization

Local prioritization is applied at the local node, before the packet is sent. This means that the prioritization has full access to the data about the individual packets, and can choose differing treatment based on the stream a packet belongs to.

When an WebRTC endpoint has packets to send on multiple streams that are congestion-controlled under the same congestion control regime, the WebRTC endpoint SHOULD cause data to be emitted in such a way that each stream at each level of priority is being given approximately twice the transmission capacity (measured in payload bytes) of the level below.

Thus, when congestion occurs, a "high" priority flow will have the ability to send 8 times as much data as a "very-low" priority flow if both have data to send. This prioritization is independent of the media type. The details of which packet to send first are implementation defined.

For example: If there is a high priority audio flow sending 100 byte packets, and a low priority video flow sending 1000 byte packets, and outgoing capacity exists for sending >5000 payload bytes, it would be appropriate to send 4000 bytes (40 packets) of audio and 1000 bytes (one packet) of video as the result of a single pass of sending decisions.

Conversely, if the audio flow is marked low priority and the video flow is marked high priority, the scheduler may decide to send 2 video packets (2000 bytes) and 5 audio packets (500 bytes) when outgoing capacity exists for sending > 2500 payload bytes.

If there are two high priority audio flows, each will be able to send 4000 bytes in the same period where a low priority video flow is able to send 1000 bytes.

Two example implementation strategies are:

- o When the available bandwidth is known from the congestion control algorithm, configure each codec and each data channel with a target send rate that is appropriate to its share of the available bandwidth.

- o When congestion control indicates that a specified number of packets can be sent, send packets that are available to send using a weighted round robin scheme across the connections.
Any combination of these, or other schemes that have the same effect, is valid, as long as the distribution of transmission capacity is approximately correct.

For media, it is usually inappropriate to use deep queues for sending; it is more useful to, for instance, skip intermediate frames that have no dependencies on them in order to achieve a lower bitrate. For reliable data, queues are useful.

Note that this specification doesn’t dictate when disparate streams are to be "congestion controlled under the same congestion control regime". The issue of coupling congestion controllers is explored further in [I-D.ietf-rmcat-coupled-cc].

4.2. Usage of Quality of Service - DSCP and Multiplexing

When the packet is sent, the network will make decisions about queueing and/or discarding the packet that can affect the quality of the communication. The sender can attempt to set the DSCP field of the packet to influence these decisions.

Implementations SHOULD attempt to set QoS on the packets sent, according to the guidelines in [I-D.ietf-tsvwg-rtcweb-qos]. It is appropriate to depart from this recommendation when running on platforms where QoS marking is not implemented.

The implementation MAY turn off use of DSCP markings if it detects symptoms of unexpected behaviour like priority inversion or blocking of packets with certain DSCP markings. Some examples of such behaviors are described in [ANRW16]. The detection of these conditions is implementation dependent.

A particularly hard problem is when one media transport uses multiple DSCP code points, where one may be blocked and another may be allowed. This is allowed even within a single media flow for video in [I-D.ietf-tsvwg-rtcweb-qos]. Implementations need to diagnose this scenario; one possible implementation is to send initial ICE probes with DSCP 0, and send ICE probes on all the DSCP code points that are intended to be used once a candidate pair has been selected. If one or more of the DSCP-marked probes fail, the sender will switch the media type to using DSCP 0. This can be carried out simultaneously with the initial media traffic; on failure, the initial data may need to be resent. This switch will of course invalidate any congestion information gathered up to that point.

Failures can also start happening during the lifetime of the call; this case is expected to be rarer, and can be handled by the normal mechanisms for transport failure, which may involve an ICE restart.
Note that when a DSCP code point causes non-delivery, one has to switch the whole media flow to DSCP 0, since all traffic for a single media flow needs to be on the same queue for congestion control purposes. Other flows on the same transport, using different DSCP code points, don’t need to change.

All packets carrying data from the SCTP association supporting the data channels MUST use a single DSCP code point. The code point used SHOULD be that recommended by [I-D.ietf-tsvwg-rtcweb-qos] for the highest priority data channel carried. Note that this means that all data packets, no matter what their relative priority is, will be treated the same by the network.

All packets on one TCP connection, no matter what it carries, MUST use a single DSCP code point.

More advice on the use of DSCP code points with RTP and on the relationship between DSCP and congestion control is given in [RFC7657].

There exist a number of schemes for achieving quality of service that do not depend solely on DSCP code points. Some of these schemes depend on classifying the traffic into flows based on 5-tuple (source address, source port, protocol, destination address, destination port) or 6-tuple (5-tuple + DSCP code point). Under differing conditions, it may therefore make sense for a sending application to choose any of the configurations:

- Each media stream carried on its own 5-tuple
- Media streams grouped by media type into 5-tuples (such as carrying all audio on one 5-tuple)
- All media sent over a single 5-tuple, with or without differentiation into 6-tuples based on DSCP code points

In each of the configurations mentioned, data channels may be carried in its own 5-tuple, or multiplexed together with one of the media flows.

More complex configurations, such as sending a high priority video stream on one 5-tuple and sending all other video streams multiplexed together over another 5-tuple, can also be envisioned. More information on mapping media flows to 5-tuples can be found in [I-D.ietf-rtcweb-rtp-usage].

A sending implementation MUST be able to support the following configurations:
0. Multiplex all media and data on a single 5-tuple (fully bundled)
0. Send each media stream on its own 5-tuple and data on its own 5-tuple (fully unbundled)

It MAY choose to support other configurations, such as bundling each media type (audio, video or data) into its own 5-tuple (bundling by media type).

Sending data channel data over multiple 5-tuples is not supported.

A receiving implementation MUST be able to receive media and data in all these configurations.

5. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

6. Security Considerations

RTCWEB security considerations are enumerated in [I-D.ietf-rtcweb-security].

Security considerations pertaining to the use of DSCP are enumerated in [I-D.ietf-tsvwg-rtcweb-qos].

7. Acknowledgements

This document is based on earlier versions embedded in [I-D.ietf-rtcweb-overview], which were the results of contributions from many RTCWEB WG members.

Special thanks for reviews of earlier versions of this draft go to Eduardo Gueiros, Magnus Westerlund, Markus Isomaki and Dan Wing; the contributions from Andrew Hutton also deserve special mention.

8. References

8.1. Normative References


[I-D.ietf-rtcweb-data-protocol]

[I-D.ietf-rtcweb-overview]

[I-D.ietf-rtcweb-rtp-usage]

[I-D.ietf-rtcweb-security]

[I-D.ietf-rtcweb-security-arch]

[I-D.ietf-tsvwg-rtcweb-qos]

[I-D.ietf-tsvwg-sctp-dtls-encaps]

[I-D.ietf-tsvwg-sctp-ndata]

[RFC0768]

[RFC0793]


8.2. Informative References


Appendix A. Change log

This section should be removed before publication as an RFC.

A.1. Changes from -00 to -01

- Clarified DSCP requirements, with reference to -qos-
- Clarified "symmetric NAT" -> "NATs which perform endpoint-dependent mapping"
- Made support of TURN over TCP mandatory
- Made support of TURN over TLS a MAY, and added open question
- Added an informative reference to -firewalls-
- Called out that we don’t make requirements on HTTP proxy interaction (yet)

A.2. Changes from -01 to -02

- Required support for 300 Alternate Server from STUN.
- Separated the ICE-TCP candidate requirement from the TURN-TCP requirement.
- Added new sections on using QoS functions, and on multiplexing considerations.
- Removed all mention of RTP profiles. Those are the business of the RTP usage draft, not this one.
- Required support for TURN IPv6 extensions.
- Removed reference to the TURN URI scheme, as it was unnecessary.
- Made an explicit statement that multiplexing (or not) is an application matter.
A.3. Changes from -02 to -03

- Added required support for draft-ietf-tsvwg-sctp-ndata
- Removed discussion of multiplexing, since this is present in rtp-usage.
- Added RFC 4571 reference for framing RTP packets over TCP.
- Downgraded TURN TCP candidates from SHOULD to MAY, and added more language discussing TCP usage.
- Added language on IPv6 temporary addresses.
- Added language describing multiplexing choices.
- Added a separate section detailing what it means when we say that an WebRTC implementation MUST support both IPv4 and IPv6.

A.4. Changes from -03 to -04

- Added a section on prioritization, moved the DSCP section into it, and added a section on local prioritization, giving a specific algorithm for interpreting "priority" in local prioritization.
- ICE-TCP candidates was changed from MAY to MUST, in recognition of the sense of the room at the London IETF.

A.5. Changes from -04 to -05

- Reworded introduction
- Removed all references to "WebRTC". It now uses only the term RTCWEB.
- Addressed a number of clarity / language comments
- Rewrote the prioritization to cover data channels and to describe multiple ways of prioritizing flows
- Made explicit reference to "MUST do DTLS-SRTP", and referred to security-arch for details

A.6. Changes from -05 to -06

- Changed all references to "RTCWEB" to "WebRTC", except one reference to the working group
o Added reference to the httpbis "connect" protocol (being adopted by HTTPBIS)
o Added reference to the ALPN header (being adopted by RTCWEB)
o Added reference to the DART RTP document
o Said explicitly that SCTP for data channels has a single DSCP codepoint

A.7. Changes from -06 to -07

o Updated references
o Removed reference to draft-hutton-rtcweb-nat-firewall-considerations

A.8. Changes from -07 to -08

o Updated references
o Deleted "bundle each media type (audio, video or data) into its own 5-tuple (bundling by media type)" from MUST support configuration, since JSEP does not have a means to negotiate this configuration

A.9. Changes from -08 to -09

o Added a clarifying note about DTLS-SRTP and ICE interaction.

A.10. Changes from -09 to -10

o Re-added references to proxy authentication lost in 07-08 transition (Bug #5)

o Rearranged and rephrased text in section 4 about prioritization to reflect discussions in TSVWG.

o Changed the "Connect" header to "ALPN", and updated reference. (Bug #6)

A.11. Changes from -10 to -11

o Added a definition of the term "flow" used in the prioritization chapter

o Changed the names of the four priority levels to conform to other specs.
A.12. Changes from -11 to -12

- Added a SHOULD NOT about using deprecated temporary IPv6 addresses.
- Updated draft-ietf-dart-dscp-rtp reference to RFC 7657

A.13. Changes from -12 to -13

- Clarify that the ALPN header needs to be sent.
- Mentioned that RFC 7657 also talks about congestion control

A.14. Changes from -13 to -14

- Add note about non-support for marking flows as interactive or non-interactive.

A.15. Changes from -14 to -15

- Various text clarifications based on comments in Last Call and IESG review
- Clarified that only non-deprecated IPv6 addresses are used
- Described handling of downgrading of DSCP markings when blackholes are detected
- Expanded acronyms in a new protocol list

A.16. Changes from -15 to -16

These changes are done post IESG approval, and address IESG comments and other late comments. Issue numbers refer to https://github.com/rtcweb-wg/rtcweb-transport/issues.

- Moved RFC 4594, 7656 and -overview to normative (issue #28)
- Changed the terms "client", "WebRTC implementation" and "WebRTC device" to consistently be "WebRTC endpoint", as defined in -overview. (issue #40)
- Added a note mentioning TURN service discovery and RETURN (issue #42)
- Added a note mentioning that rtp-usage requires circuit breaker and congestion control (issue #43)
o Added mention of the "don’t discard temporary IPv6 addresses that are in use" (issue #44)

o Added a reference to draft-ietf-rmcat-coupled-cc (issue #46)

A.17. Changes from -16 to -17

o Added an informative reference to the "DSCP blackholing" paper

o Changed the reference for ICE from RFC 5245 to draft-ietf-ice-rfc5245bis

Author’s Address

Harald Alvestrand
Google

Email: harald@alvestrand.no