mlcodec
Internet-Draft
Updates: 6716 (if approved)
Intended status: Standards Track
Expires: 24 January 2026

T. Terriberry
Xiph.Org
JM. Valin
Google

23 July 2025

# Extension Formatting for the Opus Codec
## draft-ietf-mlcodec-opus-extension-04

## Abstract

This document updates RFC6716 to extend the Opus codec (RFC6716) in a way that maintains interoperability, while adding optional functionality.

## 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 https://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 24 January 2026.

## Copyright Notice

Table of Contents



## 1. Introduction

This document updates RFC6716 to extend the Opus codec (RFC6716) in a way that maintains interoperability, while adding optional functionality.

### 1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

## 2. Extension Format

The Opus padding mechanism provides a safe way to extend the Opus codec while preserving interoperability and without having to transmit any extra packets. [RFC6716] specifies that all padding bytes "MUST be set to zero" by the encoder, while the decoder "MUST accept any value for the padding bytes". In that way, any non-zero padding will indicate to an extended decoder that extensions are present and can be processed. On the other hand, for any all-zero padding, the decoder will just discard the padding like any non-extended decoder. A non-extended decoder receiving a packet with extensions will simply discard the extensions and proceed as if none were present.

An instance of an extension is composed of an "extension ID byte" and an optional payload, which may be prefixed by an optional length indicator, followed by 0 or more bytes of extension data. Although there is only one padding region per Opus packet, each extension instance is tied to an underlying Opus frame, of which there may be more than one per packet. Extension instances are grouped by the corresponding Opus frame they are extending, starting from the first frame, with frame separator extensions (Section 2.2) delineating the boundaries between the extensions for each frame.

There are three types of extensions:

* Structural extensions (IDs 0, 1, and 2), which control extension parsing, but do not inherently change the behavior of a decoder themselves.

* Short extensions (IDs 3 through 31), which have either 0 or 1 bytes of extension data.

* Long extensions (IDs 32 through 127), which can have an arbitrary number of extension data bytes.

An extension instance starts with an "extension ID byte" that contains a 7-bit ID, as well as a binary flag L for length signaling. For short extensions, L=0 means that no data follows the extension ID byte, whereas L=1 means that exactly one byte of extension data follows. For long extensions, L=0 signals that the extension data takes up the rest of the padding. In any given packet, this signal cannot appear more than once. Conversely, L=1 in a long extension signals that a length indicator follows. The following byte contains a length value from 0 to 254, or the special value 255, indicating that the length is 255 plus the length signaled from the next byte. The 255 case MAY repeat as long as the size of the padding is not exceeded. Also, any extension signaled with a length that would cause the decoder to read beyond the bounds of the packet MUST be ignored by the decoder.

For ID 0 (Original Padding), L=0 has the same meaning as for long extensions, but L=1 signals a length of zero (no length indicator or extension data follows). For ID 1 (Frame Separator), the L flag has the same meaning as for short extensions. For ID 2 (Repeat These Extensions), the extension itself has no payload (for either L=0 or L=1), but is used to signal that previously coded extensions are to be repeated for subsequent Opus frames. The payloads of the repeated extensions follow immediately after. See Section 2.3 for the details of this process and for how the L flag is to be interpreted.

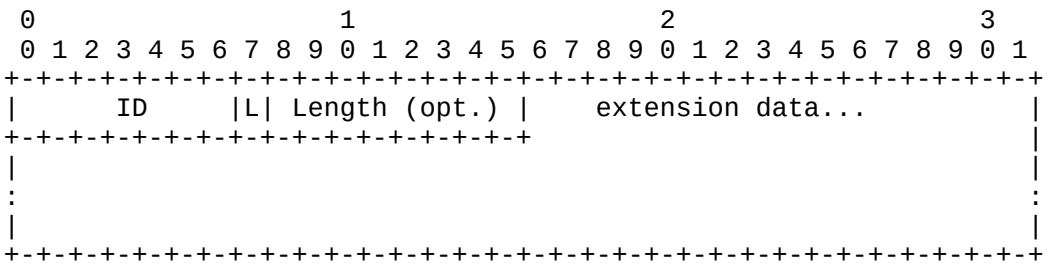


Figure 1: Extension framing

A decoder MUST ignore any extension it does not support, decoding the rest of the packet as if the extension was not present. Additionally, a decoder MAY ignore any other extension even if it technically supports it. An encoder MUST NOT alter the way it encodes the non-extension part of an Opus packet in such a way as to noticeably reduce its quality when decoded with a non-extended decoder.

A given extension ID MAY appear multiple times and the ordering of extension instances within each Opus frame is significant (see Section 2.2). A particular extension ID definition MAY place further restrictions on count and ordering of these extensions instances (see Section 3). Reordering of extension instances between Opus frames caused by the repeat mechanism is not significant and an extended decoder MUST treat repeated extensions as equivalent to the same extensions coded individually (see Section 2.3).

## 2.1. ID 0: Original Padding

For compatibility reasons, an ID of 0 means that the remaining content of the padding is actual padding, as originally defined in [RFC6716]. As in its original definition, the padding bytes MUST be set to zero by the encoder, while the decoder MUST ignore any non-zero padding. In the case where the L flag is set, the extension ID byte (0x01) is simply skipped and extension decoding continues from the next byte. This allows inserting padding one byte at a time in a way that would not be possible if an explicit padding length were coded instead (that would make L=1 padding require at least two bytes, and appending a L=0 padding might require signaling a multi-byte length indicator for a preceding long extension that would not otherwise be necessary, both causing the packet size to increase by more than one byte).

## 2.2. ID 1: Separator

In the case where multiple Opus frames are packed inside the same packet, frame separators specify which extension instance(s) are associated with which frames. An extension instance with ID=1 acts as a separator between extension instances from different Opus frames.

By default, extension instances are associated with the first Opus frame in the packet (frame 0). When parsing sequentially, any time a separator with L=0 is encountered, the associated frame index is incremented by one. If L=1 is used, the following payload byte indicates the amount by which to increment the frame index. The frame index MUST NOT exceed the number of frames in the packet minus one (i.e., indexing starts at zero), regardless of how it is incremented. The decoder MUST ignore all extension instances associated with an out-of-bounds frame index.

## 2.3. ID 2: Repeat These Extensions (RTE)

In the case where multiple Opus frames are packed inside the same packet, the Repeat These Extensions (RTE) extension can reduce the overhead of coding extension IDs and frame separators when the extensions in the current frame also appear in every subsequent frame (albeit, possibly with different payloads). An extension with ID=2 acts as a signal to repeat all of the non-padding (ID=0) extensions following the most recent of

* The start of the packet, or

* A frame separator (ID=1) with a non-zero increment, or

* A preceding RTE extension (ID=2), if any.

Padding extensions are not repeated, nor is any frame separator with an increment of 0 (which acts as another form of padding). Extensions preceding a frame separator with an increment of zero do get repeated, as they still belong to the current frame. An RTE extension MAY appear multiple times in the same frame. Only the extensions which follow the most recent RTE (if any) are repeated by a subsequent RTE.

The RTE extension itself has no payload, but it is immediately followed by the payloads of new instances of the repeated extensions. The payloads for all of the repeated extensions for the next frame come first, followed by those of the frame after, etc. The extension ID byte corresponding to each payload is implicit and not coded: only the length (if needed) and the subsequent extension data are coded. An RTE extension MAY appear in the last frame. In this case, no extensions are repeated.

For short extensions, the repeated extension payloads use the same L flag as the instance of the extension being repeated. If the length of a short extension needs to change between frames, this repeat mechanism cannot be used to signal that. All repeated long extension payloads except the final instance of the last repeated long extension in the last frame are coded as if with L=1 (using an explicit length indicator). The final repeated long extension payload is coded with the L flag specified by the RTE extension. In the case that the RTE extension specifies L=0, and the last repeated long extension is followed by one or more repeated short extensions with a payload, then the final long extension does not consume the rest of the padding as normal, but leaves enough room for the payloads of the repeated short extensions that follow. If there is not enough room for the repeated short extensions that follow, even if the length of the final long extension were set to zero, then that extension instance and all remaining padding data MUST be ignored by the decoder.

If the RTE extension uses L=1, then extension coding continues afterwards with the same frame index as the RTE extension. This allows a frame to contain both repeated and non-repeated extensions. This also means that the complete collection of extension instances for a given frame might not all be contiguous in the packet. If the RTE extension uses L=0, but the repeated extensions did not contain a long extension, then extension coding continues afterwards with the frame index following that of the RTE extension (as if an L=0 separator had been coded). If an RTE extension with L=0 appears in the last frame, then the rest of the padding (if any) MUST be set to zero by the encoder, and the decoder MUST ignore any additional non-zero padding.

### 2.4. IDs 3-119: Unassigned

These extensions are to be defined in their own respective documents and the IDs are to be assigned by IANA. The meaning of the L flag is already defined for all of these unassigned IDs because a decoder must know how to skip extensions it does not support. Due to potential for interaction between extensions, new extensions are to be assigned with the "Standards Action" policy defined by [RFC8126].

### 2.5. IDs 120-126: Experimental

We reserve these 7 IDs for experimental extensions, such that extensions defined in Internet-Drafts can be tested before publication as an RFC without causing possible interoperability issues should their bitstream definitions change. When using an experimental ID, it is RECOMMENDED to use a two-byte prefix that attempts to encode an experiment number (first byte) and a version number (second byte). Experimental extension documents SHOULD attempt to choose an experiment number that does not collide with other ongoing experiments.

### 2.6. ID 127: Extended Extensions

The last ID is reserved for future extensions to the extension mechanism. As with all other long extensions, the meaning of the L flag is pre-defined to ensure decoders can skip extended extensions they do not support. The contents of the payload for this extension will be defined by a future specification.

## 3. IANA Considerations

This document defines a new registry "Opus Extension IDs" in a new "Opus" group, that allocates individual IDs to individual extensions to be defined in the future. The existing "Opus Channel Mapping Families" registry will also be moved to the newly created "Opus" group. Moreover, this document already defines the following IDs:

| Extension ID | Description | Reference |
|---|---|---|
| 0 | Original Padding | Defined in Section 2.1 |
| 1 | Frame Separator | Defined in Section 2.2. |
| 2 | Repeat These Extensions | Defined in Section 2.3. |
| 3-119 | Unassigned | To be assigned with the "Standards Action" policy [RFC8126] |
| 120-126 | Experimental | Defined in Section 2.5, following the "Experimental Use" policy [RFC8126] |
| 127 | Extended Extensions | Reserved in Section 2.6 |

Table 1

For forward compatibility, any extension MUST use the definition of the L flag dictated by its ID value (see Section 2). Extension definitions MUST specify whether or not it is permitted for the extension to appear multiple times for a given Opus frame within the packet.

### 3.1. Opus Media Type Update

This document updates the audio/opus media type registration [RFC7587] to add the following two optional parameters:

extensions: specifies a comma-separated list of supported extension IDs on the receiver side.

sprop-extensions: specifies a comma-separated list of supported extension IDs on the sender side.

extN-*: To facilitate parameter forwarding, extension document that require receiver extension parameters SHOULD name them "ext", followed by the extension ID, a hyphen, and the parameter name.

sprop-extN-*: Extension-specific sender-side parameters defined similarly as above.

All names starting with "ext" and "sprop-ext" are reserved for use by Opus extensions.

Structural extensions (IDs 0, 1, and 2) MUST be supported by any receiver that recognizes Opus extensions, and do not need to be included in the extensions or sprop-extensions lists.

### 3.2. Mapping to SDP Parameters

The media type parameters described above map to declarative SDP and SDP offer-answer in the same way as other optional parameters in [RFC7587]. As per [RFC5576] Section 6.3, media-level format parameters MUST be explicitly specified and MUST NOT be carried over blindly from another offer or answer. Regardless of any a=fmtp SDP attribute specified, the receiver MUST be capable of receiving any signal.

## 4. Security Considerations

This document does not add security considerations beyond those already documented in [RFC6716]. Future Opus extensions may have their own security implications.

## 5. References

### 5.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC6716] Valin, JM., Vos, K., and T. Terriberry, "Definition of the Opus Audio Codec", RFC 6716, DOI 10.17487/RFC6716, September 2012, <https://www.rfc-editor.org/info/rfc6716>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC7587] Spittka, J., Vos, K., and JM. Valin, "RTP Payload Format for the Opus Speech and Audio Codec", RFC 7587, DOI 10.17487/RFC7587, June 2015, <https://www.rfc-editor.org/info/rfc7587>.

[RFC5576] Lennox, J., Ott, J., and T. Schierl, "Source-Specific Media Attributes in the Session Description Protocol (SDP)", RFC 5576, DOI 10.17487/RFC5576, June 2009, <https://www.rfc-editor.org/info/rfc5576>.

## Authors' Addresses

Timothy B. Terriberry
Xiph.Org
United States of America
Email: tterribe@xiph.org

Jean-Marc Valin
Google
Canada
Email: jeanmarcv@google.com