Adaptive-ptime Field in RTCConfiguration - Explainer

Status: Draft
Authors: eladalon@chromium.org
Last Updated: 2019-09-25

Summary

We propose a mechanism by which the creator of an RTCPeerConnection may indicate the desire for the audio frame length to be adaptive. That is, the frame length is allowed to change during the call, so as to optimize the trade-off between bandwidth consumption and latency, according to the characteristics of the network. The proposed mechanism only indicates that this may be done, not how it should be done; it's up to the implementation to optimize the aforementioned trade-off.

The WebRTC spec does not mandate a constant frame length be used. This flag we are proposing helps avoid interop concerns. The reason this is implemented before it was approved in the WebRTC spec, is that we wish to begin experimenting with it on select websites, without risking breaking the Web at large.

Motivation

The default audio codec in WebRTC is Opus. By default, 20ms frames are produced, meaning 50 frames per second. Since each is transmitted in its own RTP packet, which in turn takes up its own UDP datagram, in its own IP frame, in its own L2 packet, substantial bandwidth is expended on overhead.

To illustrate, assume IPv4 over Ethernet. Assume 14B for Ethernet, 20B for IPv4, 8B for UDP, and 24B for RTP. This sums up to 66B per frame, or 50 * 66 * 8 = 26.4kbps.

This overhead can be reduced by increasing the frame length; however, this comes at the cost of audio latency, due to increased buffering on the sender. Whether the trade-off favors longer or shorter frames, depends on the available bandwidth - a dynamic property.

Specifics

- We will add a flag to the RTCConfiguration dictionary.
- That flag will be called “adaptivePtime”. Its type will be an enum called “RTCAdaptivePtime”. It will be allowed to assume one of three states - “off”, “on” and “default”.
- Initially, in Chrome, “default” will be mapped to “off”, but this can be changed in the future. Other implementations are free to map “default” however they see fit.
- When ptime adaptivity is turned on, the implementation may change the frame length of the active audio codec at any time. It’s up to the implementation to decide the logic by which this is done. Implementations may support this behavior for some audio codecs and ignore it for others.
- The limit imposed by any "maxptime" attribute, as defined in [RFC4566], Section 6, must still be respected.

Risks

Interoperability and Compatibility

The main risk is that an audio stream would be produced by a client and sent to another client that does not support receiving it. It is up to the creator of the RTCPeerConnection to avoid this. For example, a website which connects its visitors to telephony hardware, must either not turn ptime adaptivity on, or do so only if the relevant hardware can handle such a stream.

In practice, the latest versions of Chrome and Firefox have been tested using Opus at frame-lengths varying dynamically between 20ms and 120ms, and both have found to support that. We expect that all solutions based on a relatively recent version of the webrtc.org implementation of the WebRTC standard will support receiving Opus at adaptive frame lengths.

Ergonomics

This may be used in tandem with “maxptime”; see the relevant bullet-point under Specifics.

Activation

The feature may be used immediately, as-is. It is important to bear in mind, however, that it’s up to the implementation to decide how and when the frame length is to be adapted, if at all.