Network Working Group   GM. Maxwell
Internet-Draft   Xiph.Org
Intended status: Informational   K. Walsh
Expires: April 18, 2013   October 15, 2012

                 Requirements for an Internet Video Codec
                 draft-maxwell-videocodec-requirements-00

Abstract

   This document provides specific requirements for an Internet video
   codec. These requirements address quality, bit-rate, loss
   robustness, application suitability, as well as other desirable
   properties.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

   This document provides requirements for a video codec designed
   specifically for use over the Internet. The requirements attempt to
   address the needs of the most common Internet video transmission
   applications and to ensure good quality when operating in conditions
   that are typical for the Internet. These requirements address
   quality, bit-rate, and loss robustness. Other desirable codec
   properties are considered as well.

2. Definitions

   Codec bit-rates in bits per second (b/s) will be considered without
   counting any overhead (IP/UDP/RTP headers, padding, ...).

3. Applications

   The following applications should be considered for Internet video
   codecs, along with their requirements:

   o Live video streaming

   o Video on demand

   o Point to point video calls

   o Video Conferencing

   o Telepresence

   o Teleoperation

   o Remote software services

   o Other applications

3.1. Point to point video calls

   Point to point calls are calls from two "standard" (fixed or mobile)
   phones, desktop, portable computers, or tablets, and implemented in
   hardware or software.

3.2. Video Conferencing

   Video Conferencing applications (which support multi-party calls)
   have additional requirements on top of the requirements for point-to-
   point calls. Conferencing systems often have greater network
   bandwidth available. The ability to vary the bit-rate (VBR) is a
   desirable feature for the codec. This not only saves bandwidth "on
   average", but it can also help conference servers make more efficient
   use of the available bandwidth by using more bandwidth for important
   video streams and less bandwidth for less important ones.

3.3. Telepresence

   Most telepresence applications can be considered to be essentially
   very high quality video conferencing environments, so all of the
   conferencing requirements also apply to telepresence.

3.4. Teleoperation and Remote Software Services

   Teleoperation applications are similar to telepresence, with the
   exception that they involve remote physical interactions. For

example, the user may be controlling a robot while receiving a real-
time video feed from that robot. The other requirements of
telepresence apply to teleoperation as well.

The requirements for remote software services are similar to those of
teleoperation. These applications include remote desktop
applications, remote virtualization, and interactive media
applications being rendered remotely (e.g. video games rendered on
central servers).

3.5. Other applications

The above list is by no means a complete list of all applications
involving interactive video transmission on the Internet. However,
it is believed that meeting the needs of all these different
applications should be sufficient to ensure that most applications
not listed will also be met.

4. Constraints Imposed by the Internet on the Codec

   The bandwidth requirements of video are a significant obstacle for
   Internet deployment. A substantial portion of the hosts on the
   Internet have connectivity sufficient to carry perceptually lossless
   audio, even in inefficient uncompressed form. However, a much
   smaller portion of hosts have connectivity sufficient for the 200+
   megabits per second required for uncompressed 30fps standard
   definition video, and expected resolutions for Internet video are
   increasing. Even the highest resolutions widely used off the
   Internet, where wide-area bandwidth is not a constraint, have not yet
   reached perceptual losslessness. In addition to increases in
   resolution, operating models which are less broadcast-oriented
   (including video on demand and video conferencing) limit the traffic
   mitigation effectiveness of CDNs and multicast, though support for
   these technologies remains essential. Because there are a few
   applications where increases in bandwidth efficiency are not
   important and many where improved efficiency is essential--such as
   delivering HD video to bandwidth-constrained network edges--it is
   important that the codec deliver competitive quality per bitrate and
   support a wide range of bandwidths.

   Packet losses are inevitable on the Internet and dealing with them is
   one of requirements for an Internet video codec. Efficient video
   compression typically uses very high gain backward prediction, which
   can result in infinite error propagation in the worst case if
   measures are not taken to mitigate it. Error propagation is usually
   mitigated in traditional file- and broadcast-oriented codecs through
   key-frames, periodic intra refresh, and constrained back-reference
   structure. While these techniques are important, and also enable
   random access, a codec designed for the Internet should also be able
   to take advantage of bidirectional communication to reduce the impact
   of loss when possible.

   In many high-latency and non-realtime applications, however, the
   relevant transport is lossless. While random access still is
   important, general error tolerance is not, and the codec may support
   modes which have very low error tolerance--including ones which
   prevent packet decode in the presence of loss, if it results in
   efficiency gains for non-realtime applications.

   For interactive applications latency is an important codec
   performance metric and many common input and output devices add
   frames of latency. To avoid adding further delays the codec must
   support operating in a mode that adds no more delay than that from
   processing a single frame at a time. Modes which permit sub-frame
   encoding may be useful but are hampered by the lack of subframe
   support in existing input and output devices.

Another important property of the Internet is that it is mostly a
best-effort network, with no guaranteed bandwidth. This means that
the codec has to be able to vary its output bit-rate dynamically (in
real-time), without requiring an out-of-band signaling mechanism, and
without causing artifacts at the bit-rate change boundaries. Because
the complete range of useful bit-rates may not be achievable at a
single resolution the codec may need to support changing resolutions
on the fly. Additional desirable features are:

o Having the possibility to use smooth bit-rate changes with high
   bit-rate resolution;

o Making it possible for a codec to adapt its bit-rate based on the
   source signal being encoded (source-controlled VBR) to maximize
   the quality for a certain _average_ bit-rate.

Because the Internet transmits data in bytes, a codec should produce
compressed data in integer numbers of bytes. In general, the codec
design should take into consideration explicit congestion
notification (ECN) and multicast and may include features that would
improve the quality of an ECN or multicast enabled deployment.

The IETF has defined a set of application-layer protocols to be used
for transmitting real-time transport of multimedia data, including
video. It is thus important for the resulting codec to be easy to
use with these protocols. For example, it must be possible to create
an [RTP] payload format that conforms to BCP 36 [PAYLOADS]. If any
codec parameters need to be negotiated between end-points, the
negotiation should be as easy as possible to carry over SIP
[RFC3261]/SDP [RFC4566] or alternatively over XMPP [RFC6120]/Jingle
[XEP-0167].

5. Detailed Basic Requirements

   This section summarizes all the constraints imposed by the target
   applications and by the Internet into a set of actual requirements
   for codec development.

5.1. Quality and bit-rate

   The quality of a codec is directly linked to the bit-rate, so these
   two must be considered jointly. When comparing the bit-rate of
   codecs, the overhead of IP/UDP/RTP headers should not be considered,
   but any additional bits required in the RTP payload format after the
   header (e.g. required signaling) should be considered. In terms of
   quality vs bit-rate, the codec to be developed must be better than
   the following codecs, that are generally considered as royalty-free:

   o VP8

   o Theora

   It is desirable for the codecs to support source-controlled variable
   bit-rate (VBR) to take advantage from the fact that different inputs
   require a different bitrate to achieve the same quality.

5.2. Computational resources

   The resulting codec should be implementable on a wide range of
   devices, and should not have a design which gratuitously complicates
   low power ASIC implementations. While the codec must not depend on
   special hardware features or instructions, the codec design should
   allow implementations to take full advantage of hardware accelerators
   and vector instructions where available. Complexity should generally
   scale with resolution, and it is also desirable to support multiple
   encoder and decoder complexity levels via mechanisms other than
   resolution, in order to achieve the best possible bitrate/quality
   trade-off available across many kinds of devices without unduly
   constraining resolution. The codec should also be able to take
   advantages of advances in computer speed and the deployment of
   hardware accelerators which would allow the use of higher complexity
   modes in a broader set of applications.

   In addition to computational complexity, dynamic memory for reference
   storage is a significant resource constraint for video codecs. It is
   desirable that the codec support different memory usage tradeoffs to
   fit on more devices, and that the codec not require implementations
   to utilize more memory without reasonable efficiency gains.

6. Additional considerations

   There are additional features or characteristics that may be
   desirable under some circumstances, but should not be part of the
   strict requirements. The benefit of meeting these considerations
   should be weighted against the associated cost.

7. Encoder side potential for improvement

   In most video codecs, it is possible to improve the quality by
   improving the encoder without breaking compatibility (i.e. without
   changing the decoder). Potential for improvement varies from one
   codec to another. All things being equal, being able to improve a
   codec after the bit-stream is a desirable property. However, this
   should not be done at the expense of quality in a straight-forward
   encoder.

8. Bit error robustness

   The vast majority of Internet-based applications do not need to be
   robust to bit errors because packets either arrive unaltered, or do
   not arrive at all. Considering that, the emphasis should be on
   packet loss robustness and packet loss concealment. That being said,
   it is often the case that extra robustness to bit errors can be
   achieved at no cost at all (i.e. no increase in size, complexity or
   bit-rate, no decrease in quality or packet loss robustness, ...). In
   those cases then it is useful to make a change that increases the
   robustness to bit errors. This can be useful for applications that
   use UDP Lite transmission (e.g. over a wireless LAN). Robustness to
   packet loss should *never* be sacrificed to achieve higher bit error
   robustness.

9. Legacy compatibility

   In order to create the best possible codec for the Internet, there is
   no general requirement for compatibility with legacy Internet codecs.
   However, compatibility with commonly used video color formats is
   desirable.

10. Security Considerations

   Although this document itself does not have security considerations,
   this section describes the security requirements for the codec.

   Just like for any protocol to be used over the Internet, security is
   a very important aspect to consider. This goes beyond the obvious
   considerations of preventing buffer overflows and similar attacks
   that can lead to denial-of-service or remote code execution. One
   very important security aspect is to make sure that the decoders have
   a bounded and reasonable worst case complexity. This prevents an
   attacker from causing a DoS by sending packets that are specially
   crafted to take a very long (or infinite) time to decode.

11. IANA Considerations

   This document has no actions for IANA.

12. Informative References

   [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston,
              A., Peterson, J., Sparks, R., Handley, M., and E.
              Schooler, "SIP: Session Initiation Protocol", RFC 3261,
              June 2002.

   [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session
              Description Protocol", RFC 4566, July 2006.

   [RFC6120] Saint-Andre, P., "Extensible Messaging and Presence
              Protocol (XMPP): Core", RFC 6120, March 2011.

   [XEP-0167]
              Ludwig, S., Saint-Andre, P., Egan, S., McQueen, R., and D.
              Cionoiu, "Jingle RTP Sessions", XSF XEP 0167,
              December 2009.

   [PAYLOADS]
              Handley, M. and C. Perkins, "Guidelines for Writers of RTP
              Payload Format Specifications", RFC 2736, BCP 36.

   [RTP]   Schulzrinne, H., Casner, S., Frederick, R., and V.
              Jacobson, "RTP: A Transport Protocol for real-time
              applications", RFC 3550.

Authors' Addresses

   Gregory Maxwell
   Xiph.Org

   Email: greg@xiph.org

   Kat Walsh

   Email: kat@mindspillage.org