

NETVC BoF

Dallas, TX, USA

Tuesday, March 24th, 2015

0900 - 1130

Note Well

- Any submission to the IETF intended by the Contributor for publication as all or part of an IETF Internet-Draft or RFC and any statement made within the context of an IETF activity is considered an "IETF Contribution". Such statements include oral statements in IETF sessions, as well as written and electronic communications made at any time or place, which are addressed to:
 - the IETF plenary session,
 - any IETF working group or portion thereof,
 - the IESG, or any member thereof on behalf of the IESG,
 - the IAB or any member thereof on behalf of the IAB,
 - any IETF mailing list, including the IETF list itself, any working group or design team list, or any other list functioning under IETF auspices,
 - the RFC Editor or the Internet-Drafts function
- All IETF Contributions are subject to the rules of RFC 5378 and RFC 3979 (updated by RFC 4879).
- Statements made outside of an IETF session, mailing list or other function, that are clearly not intended to be input to an IETF activity, group or function, are not IETF Contributions in the context of this notice. Please consult RFC 5378 and RFC 3979 for details. Please consult RFC 3978 (and RFC 4748) for details.
- A participant in any IETF activity is deemed to accept all IETF rules of process, as documented in Best Current Practices RFCs and IESG Statements.
- A participant in any IETF activity acknowledges that written, audio be made and may be available to the public.

Administrative Tasks

- Blue Sheets
- Note Takers
- Emergency Backup Note Taker
- Jabber Scribe

Agenda

Time	Length	Discussion Leader	Topic
0900 - 0910	10 minutes	Chairs	Administrivia
0910 - 0920	10 minutes	Area Director	Introduction and Scoping of BoF
0920 - 0930	10 minutes	Chairs	Goals
0930 - 0940	10 minutes	Chairs	Progress to Date
0940 - 1000	20 minutes	Mo Zanaty	Codec Considerations
1000 - 1020	20 minutes	Timothy Terriberry	Daala Coding Tools and Progress
1020 - 1055	35 minutes	Chairs	Charter Discussion
1055 - 1125	30 minutes	Chairs	Questions to be Answered

And now a word from our AD



Goals for the Proposed WG

- Development of a video codec that is:
 - Optimized for real-time communications over the public Internet
 - Competitive with or superior to existing modern codecs
 - Viewed as having IPR licensing terms that allow for wide implementation and deployment
 - Developed under the IPR rules in BCP 78 (RFC 5378) and BCP 79 (RFCs 3979 and 4879)
- Replicate the success of the CODEC WG in producing the Opus audio codec.

Progress So Far

- Need for RF codec developed within an SDO initially became prominent during RTCWEB “mandatory-to-implement” video codec discussion.
- Work has been progressing on Daala and VP10 codecs.
- Preliminary conversations on “video-codec” mailing list, informal face-to-face meeting at IETF 90.
- Several individual drafts have been published:
 - draft-valin-videocodec-pvq
 - draft-egge-videocodec-tdlt
 - draft-terriberry-codingtools
 - draft-moffitt-netvc-requirements
 - draft-daede-netvc-testing
 - draft-terriberry-ipr-license
- Some RF license grants on file:
 - <https://datatracker.ietf.org/ipr/2389/>
 - <https://datatracker.ietf.org/ipr/2390/>

Key Considerations for an Internet Video Codec

Mo Zanaty, Cisco
IETF 92

Beyond Compression

- Compression efficiency is the primary consideration in all video codecs.
- Beyond compression, there are many more key considerations, especially for interactive use on the Internet.
 - Complexity, Parallelism, Elasticity, Fast Rate Control, Error Resilience, Scalability, Content-Specific Tools, Algorithm Agility (for IPR avoidance), etc.
- These considerations may be in the charter, requirements, evaluation/testing, or not.

Complexity

- Reasonable resource requirements
 - Compute cycles
 - Memory and memory bandwidth
- Real-time operation in SW on common HW
- Efficient implementation in new HW designs
- Evaluation methodology must include this
 - Understand compression/complexity trade-offs
 - But with very wide latitude

Parallelism

- High-level multi-core parallelism
 - Encoder and decoder operation, especially entropy encoding and decoding, should allow multiple frames or sub-frame regions (e.g. 1D slices, 2D tiles, or partitions) to be processed concurrently, either independently or with deterministic dependencies that can be efficiently pipelined.
- Low-level instruction set parallelism
 - Favor algorithms that are SIMD/GPU friendly over inherently serial algorithms.

Fast, Fine Rate Control

- Network bandwidth can vary quickly and dramatically
- Encoder rate control must adapt fast, fine or steep
 - Adapt quantization of frames or sub-frame regions
 - Skip input frames or sub-frame regions
 - Adapt resolution (efficiently) if necessary
- Accurate rate control over time intervals relevant to transport systems often requires adapting quantization or skipping at granularities finer than a frame
 - Sub-frame quantization and skip control can be as coarse as a few fixed regions, or as fine as the smallest coding structure. With block sizes of 64x64, a row of blocks may be the minimum granularity needed.

Error Resilience

- Packet loss inevitably causes distortion
 - Decoder operation, especially entropy decoding, should be robust to loss.
 - Decode subsequent frames or sub-frame regions (e.g. slices, tiles, partitions) successfully even if distorted.
- Distortion spreads until resynchronization
 - Efficient resynchronization should be supported that reuses existing synchronized reference frames (e.g. locked, golden, or long-term reference frames) rather than requiring flushing and reinitializing them all.

Scalability

- Temporal scalability is critical
 - Effective for fast rate control
 - Effective for some degree of receivers' rate diversity
 - Can improve compression efficiency
- Spatial/resolution and quality/quantization scalability are useful but less critical
 - Rescaling reference frames may be sufficient
 - Degrades compression efficiency
 - Advantages outweigh this penalty for some applications

Content-Specific Tools

- Evaluation/testing should include several content classes, including synthetic (non-camera) content.
- RGB 4:4:4 for screen share, wireless display, remote gaming/graphics, etc.
- Different search strategies and coding tools
- More component planes, e.g. alpha, depth
- Exploiting component correlation

Algorithm Agility

- Avoidance of non-RF IPR is critical
- May require agility in tools that prove risky
- No good ideas how to handle this after a spec, implementations, and content are out
- Brilliant thoughts are welcome

Daala Coding Tools and Progress
netvc
IETF 92 (March 2015)

Daala Goals

- Two major goals
 - Better than state-of-the-art compression
 - Defensible IPR strategy

Daala Strategy

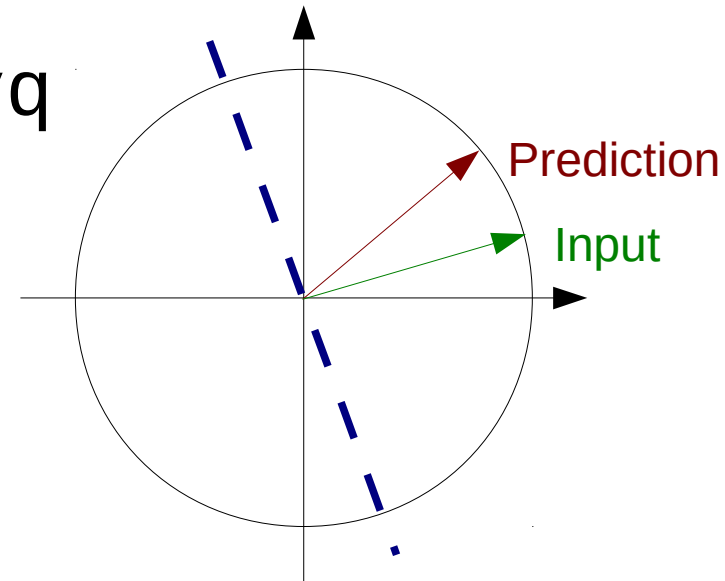
- Replace major codec building blocks with fundamentally different technology
 - Not incremental evolution
 - Higher risk/reward
- Be sufficiently different from existing approaches to avoid large swaths of patents
 - Boundaries of IPR uncertain in the best case
 - Means lawyers don't have to be perfect
 - Creates new challenges others haven't solved

Fundamentally Different

- Identified four key areas we can avoid
 - Quantizing the residual of a “Displaced Frame Difference”
 - Adaptive loop filters (deblocking)
 - Spatial prediction (“intra”)
 - Binary arithmetic coding (specifically, context modeling)

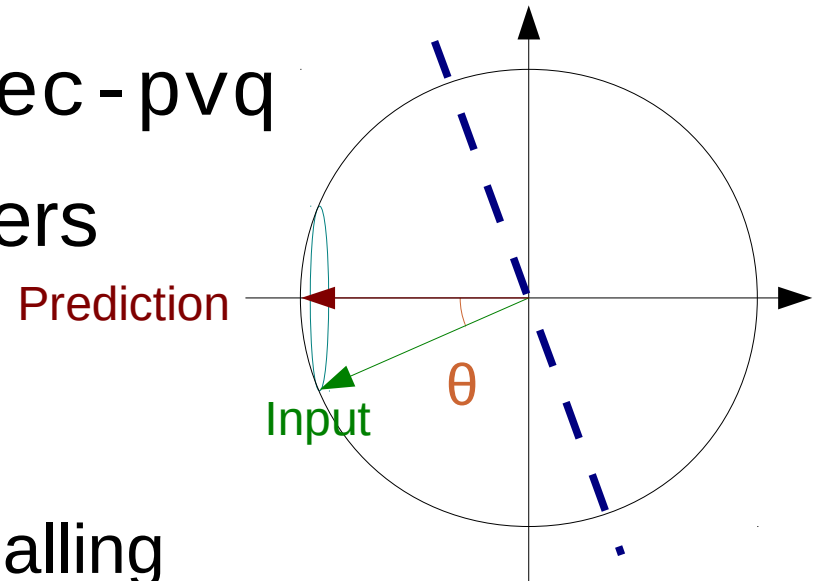
Perceptual Vector Quantization

- draft-valin-videocodect-pvq
- Simple perceptual parameters
 - energy preservation
 - prediction efficacy
 - activity masking without signalling
- Codes blocks with a predictor without subtracting and coding a residual
 - avoids anything that uses a displaced frame difference



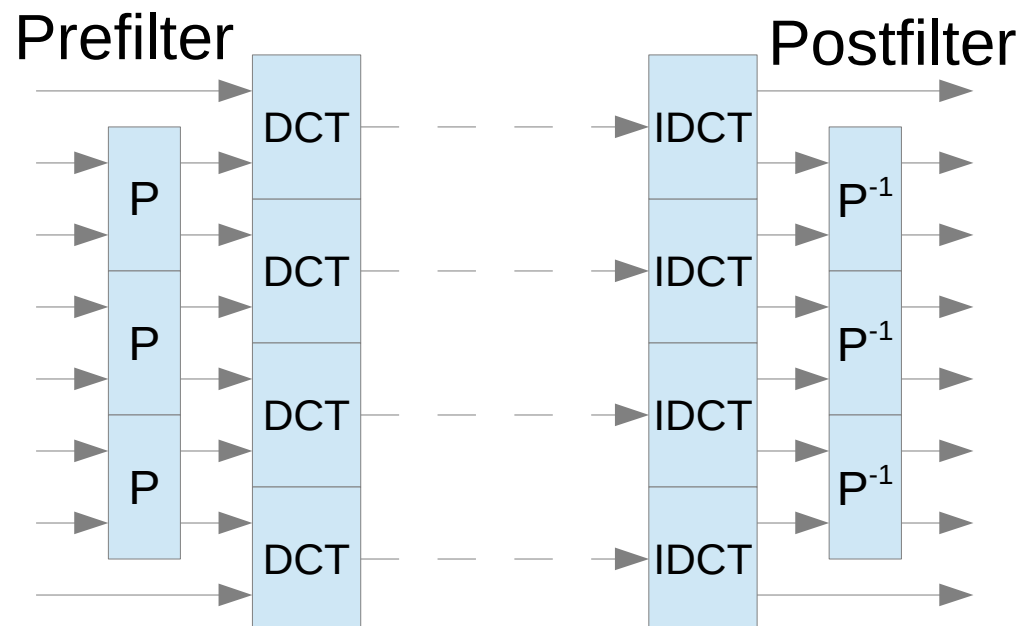
Perceptual Vector Quantization

- draft-valin-videocodect-pvq
- Simple perceptual parameters
 - energy preservation
 - prediction efficacy
 - activity masking without signalling
- Codes blocks with a predictor without subtracting and coding a residual
 - avoids anything that uses a displaced frame difference



Lapped Transforms

- draft-egge-videocodec-tdlt
- *Non-adaptive, invertible* deblocking post-filter
- Encoding applies inverse (a “blocking” filter)



Non-spatial Intra Prediction

- We can't copy pixels until we undo the lapping
 - We can't undo the lapping until we've predicted those pixels
- Don't copy pixels: copy transform coefficients
 - Currently just horizontal and vertical directions for luma
 - Chroma predicted from luma
- Not as good as spatial intra prediction, but lapping itself helps make up the difference
 - Keeps us from doing really badly (50% gains on specially constructed clips)
 - Much *cheaper* than spatial prediction (does not require full reconstruction, better hardware pipelining)

Non-binary Arithmetic Coding

- `draft-terriber-codingtools`
- Code up to 16 possible values per symbol
 - Equivalent to 4 binary decisions
 - Better throughput/cycle
- Avoids binary context modeling
- Things we use instead:
 - Frequency counts
 - Explicit Laplace/exponential models
 - Parameterized by expected value
 - “Generic Encoder” (to be replaced by more specific models later)

We need better metrics than PSNR

- We are not tuning for PSNR
 - Many of our changes actively hurt it
- Who are you going to believe? Metrics, or your lying eyes?

Current MTI Codec Example

0.537 bpp, PSNR = 33.04 dB



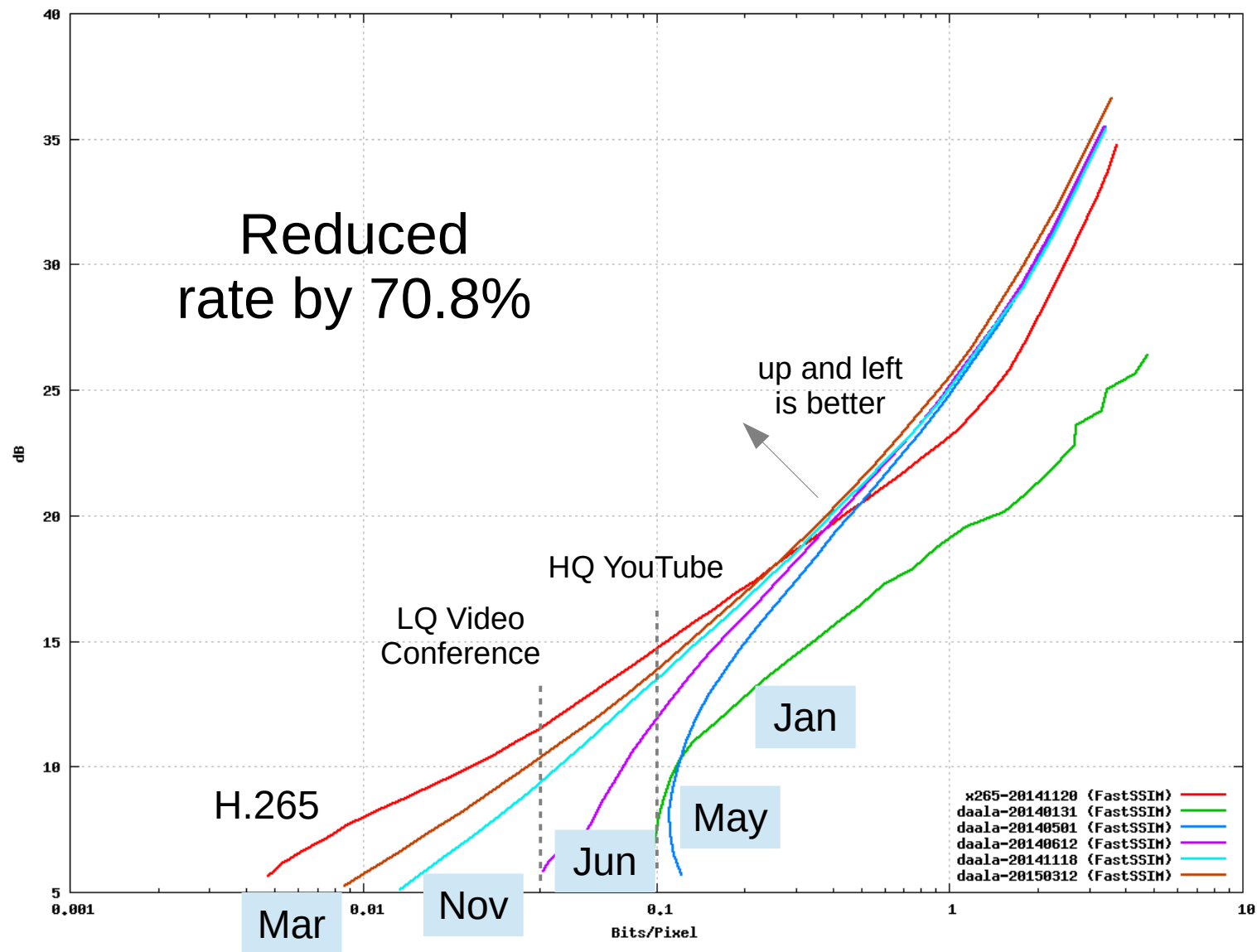
Daala Example

0.531 bpp, PSNR = 30.89 dB



Daala Progress

January 2014 to March 2015



Contributors (37)

Andreas Gal <andreas.gal@gmail.com>

Basar Koc <bkoc@mozilla.com>

Ben Brittain <ben@brittain.org>

Benjamin M. Schwartz <bemasc@google.com>

Brendan Long <self@brendanlong.com>

Brooss <brooss.teambb@gmail.com>

Cullen Jennings <fluffy@iii.ca>

David Richards <kradradio@gmail.com>

David Schleef <ds@schleef.org>

Derek Buitenhuis <derek.buitenhuis@gmail.com>

EKR <ekr@rtfm.com>

Felipe Rojo <felipe.rojod@gmail.com>

Gregory Maxwell <greg@xiph.org>

Guillaume Martres <smarter3@gmail.com>

Jack Moffitt <jack@metajack.im>

Jean-Marc Valin <jmvalin@jmvalin.ca>

Josh Aas <joshmoz@gmail.com>

Martin Olsson <martin@minimum.se>

Monty Montgomery <xiphmont@gmail.com>

Nathan E. Egge <negge@dggl.org>

Nick Desaulniers <ndesaulniers@mozilla.com>

Philip Jägenstedt <philip@foolip.org>

Ralph Giles <giles@xiph.org>

RI <u-wp1p@aetey.se>

Ron <ron@debian.org>

Sam Laane <silverdev@ymail.com>

Scott Anderson <ascent12@hotmail.com>

Sean Silva <chisophugis@gmail.com>

Sebastian Dröge <slomo@circular-chaos.org>

Steinar Midtskogen <stemidts@cisco.com>

Suhas Nandakumar <snandaku@cisco.com>

Thomas Daede <daede003@umn.edu>

Thomas Szymczak <thomasthekitty@hotmail.com>

Timothy B. Terriberry <tterribe@xiph.org>

Tristan Matthews <tmatth@videolan.org>

Vittorio Giovara <vittorio.giovara@gmail.com>

Yushin Cho <ycho@mozilla.com>

Lots of work to do

- These results are with
 - No B-frames or altref equivalents
 - No intra mode in our motion search
 - No motion compensation blocks larger than 16x16
 - No transforms larger than 32x32
 - No deringing filter (pending)

Summary

- Daala is making good progress
- We would like to contribute it as a potential candidate for NETVC

Proposed Charter Text

NETVC

Proposed Charter

Objectives

This WG is chartered to produce a high-quality video codec that meets the following conditions:

1. Is competitive with current video codecs in widespread use.
2. Is optimized for use in interactive web applications.
3. Is viewed as having IPR licensing terms that allow it to be widely implemented and deployed.

To elaborate, this video codec will need to be commercially interesting to implement by being competitive with the video codecs in widespread use at the time it is finalized.

This video codec will need to be optimized for the real-world conditions of the public, best-effort Internet. It should include, but may not be limited to, the ability to support fast and flexible congestion control and rate adaptation, the ability to quickly join broadcast streams and the ability to be optimized for captures of content typically shared in interactive communications.

The objective is to produce a video codec that can be implemented, distributed, and deployed by open source and closed source software as well as implemented in specialized hardware.

The WG will prefer algorithms or tools where there are verifiable reasons to believe they are RF over algorithms or tools where there is RF uncertainty or known active IPR claims with royalty liability potential. The codec specification will document why it believes that each part is likely to be RF, which will help adoption of the codec. This can include references to old prior art and/or patent research information.

Process

The core technical considerations for such a codec include, but are not necessarily limited to, the following:

1. High compression efficiency that is competitive with existing popular video codecs.
2. Reasonable computational complexity that permits real-time operation on existing, popular hardware, and efficient

implementation in new hardware designs.

3. Use in interactive applications, such as point-to-point video calls, multi-party video conferencing, telepresence, teleoperation, and in-game video chat.
4. Resilient in the real-world transport conditions of the Internet, such as the flexibility to rapidly respond to changing bandwidth availability and loss rates, etc.
5. Integratable with common Internet applications and Web APIs (e.g., the HTML5 <video> tag and WebRTC API, live streaming, adaptive streaming, and common media-related APIs without depending on any particular API.).

The working group shall heed the preference stated in BCP 79: "In general, IETF working groups prefer technologies with no known IPR claims or, for technologies with claims against them, an offer of royalty-free licensing." This preference cannot guarantee that the working group will produce an IPR unencumbered codec.

Non-Goals

Optimizing for very low bit rates (typically below 256 kbps) is out of scope because such work might necessitate specialized optimizations.

It is explicitly not a goal of the working group to produce a codec that will be mandated for implementation across the entire IETF or Internet community.

Based on the working group's analysis of the design space, the working group might determine that it needs to produce a codec with multiple modes of operation. The WG may produce a codec that is highly configurable, operating in many different modes with the ability to smoothly be extended with new modes in the future.

Collaboration

In completing its work, the working group will liaise with other relevant IETF working groups and SDOs, including PAYLOAD, RMCAT, RTCWEB, MMUSIC, and other IETF WGs that make use of or handle negotiation of codecs; W3C working groups including HTML, Device APIs and WebRTC; and ITU-T (Study

group 16) and ISO/IEC (JTC1/SC29 WG11).

It is expected that an open source reference version of the codec will be developed in parallel with the working group.

The WG will accept and consider in its decision process input received from external parties concerning IPR risk associated with proposed algorithms.

Deliverables

1. A document that outlines the IPR terms the working group wishes contributors to the specifications would use to license their IPR.
2. A set of technical requirements and evaluation metrics. The WG may choose to pursue publication of these in an RFC if it deems that to be beneficial.
3. Proposed Standard specification of an encoder and decoder where the normative algorithms are described in English text and not as code.
4. Specification of a storage format for file transfer of the encoded video as an elementary stream compatible with existing, popular container formats to support non-interactive (HTTP) streaming, including live encoding and both progressive and large-chunk downloads. The WG will not develop a new container format.
5. A collection of test results, either from tests conducted by the working group or made publicly available elsewhere, characterizing the performance of the codec. This document shall be informational.

Goals and Milestones

TBD IPR licensing terms goals (Informational)

TBD Requirements to IESG, if the WG so chooses (Informational)

TBD Submit codec specification to IESG (Standards Track)

TBD Submit storage format specification to IESG (Standards Track)

TBD Testing document to IESG (Informational)

Charter: Objectives (1/2)

This WG is chartered to produce a high-quality video codec that meets the following conditions:

1. Is competitive with current video codecs in widespread use.
2. Is optimized for use in interactive web applications.
3. Is viewed as having IPR licensing terms that allow it to be widely implemented and deployed.

Charter: Objectives (2/2)

To elaborate, this video codec will need to be commercially interesting to implement by being competitive with the video codecs in widespread use at the time it is finalized.

This video codec will need to be optimized for the real-world conditions of the public, best-effort Internet. It should include, but may not be limited to, the ability to support fast and flexible congestion control and rate adaptation, the ability to quickly join broadcast streams and the ability to be optimized for captures of content typically shared in interactive communications.

The objective is to produce a video codec that can be implemented, distributed, and deployed by open source and closed source software as well as implemented in specialized hardware.

The WG will prefer algorithms or tools where there are verifiable reasons to believe they are RF over algorithms or tools where there is RF uncertainty or known active IPR claims with royalty liability potential. The codec specification will document why it believes that each part is likely to be RF, which will help adoption of the codec. This can include references to old prior art and/or patent research information.

Charter: Process (1/2)

The core technical considerations for such a codec include, but are not necessarily limited to, the following:

1. High compression efficiency that is competitive with existing popular video codecs.
2. Reasonable computational complexity that permits real-time operation on existing, popular hardware, and efficient implementation in new hardware designs.
3. Use in interactive applications, such as point-to-point video calls, multi-party video conferencing, telepresence, teleoperation, and in-game video chat.
4. Resilient in the real-world transport conditions of the Internet, such as the flexibility to rapidly respond to changing bandwidth availability and loss rates, etc.
5. Integratable with common Internet applications and Web APIs (e.g., the HTML5 <video> tag and WebRTC API, live streaming, adaptive streaming, and common media-related APIs without depending on any particular API.).

Charter: Process (2/2)

The working group shall heed the preference stated in BCP 79: “In general, IETF working groups prefer technologies with no known IPR claims or, for technologies with claims against them, an offer of royalty-free licensing.” This preference cannot guarantee that the working group will produce an IPR unencumbered codec.

Charter: Non-Goals

Optimizing for very low bit rates (typically below 256 kbps) is out of scope because such work might necessitate specialized optimizations.

It is explicitly not a goal of the working group to produce a codec that will be mandated for implementation across the entire IETF or Internet community.

Based on the working group's analysis of the design space, the working group might determine that it needs to produce a codec with multiple modes of operation. The WG may produce a codec that is highly configurable, operating in many different modes with the ability to smoothly be extended with new modes in the future.

Charter: Collaboration

In completing its work, the working group will liaise with other relevant IETF working groups and SDOs, including PAYLOAD, RMCAT, RTCWEB, MMUSIC, and other IETF WGs that make use of or handle negotiation of codecs; W3C working groups including HTML, Device APIs and WebRTC; and ITU-T (Study group 16) and ISO/IEC (JTC1/SC29 WG11).

It is expected that an open source reference version of the codec will be developed in parallel with the working group.

The WG will accept and consider in its decision process input received from external parties concerning IPR risk associated with proposed algorithms.

Charter: Deliverables

1. A document that outlines the IPR terms the working group wishes contributors to the specifications would use to license their IPR.
2. A set of technical requirements and evaluation metrics. The WG may choose to pursue publication of these in an RFC if it deems that to be beneficial.
3. Proposed Standard specification of an encoder and decoder where the normative algorithms are described in English text and not as code.
4. Specification of a storage format for file transfer of the encoded video as an elementary stream compatible with existing, popular container formats to support non-interactive (HTTP) streaming, including live encoding and both progressive and large-chunk downloads. The WG will not develop a new container format.
5. A collection of test results, either from tests conducted by the working group or made publicly available elsewhere, characterizing the performance of the codec. This document shall be informational.

Charter: Milestones (Dates TBD)

- IPR licensing terms goals (Informational)
- Submit requirements to IESG, if the WG so chooses (Informational)
- Submit codec specification to IESG (Standards Track)
- Submit storage format specification to IESG (Standards Track)
- Testing document to IESG (Informational)

Questions for the Community

NETVC

Question 1

Is there a
problem that
needs solving?

Question 2

Is the scope of the problem well defined and understood? Is there agreement on what a WCG would need to deliver?

Question 3

Are there people willing to do the work?

- Who will write the drafts?
- Who will review the drafts?
- Who will implement, test, and characterize a reference implementation?

Question 4

How many people
feel that a WG
should not be
formed at the IETF?