

VP9 Payload Format

uberti-payload-vp9-01

VP9 Overview

- Much better compression than VP8/H.264 (40%)
- Up to 8 reference frames
 - Can be different resolution than target frame
 - Allows temporal + spatial layering
- Higher layers can be removed without affecting decode of lower layers

Terminology

- **Layer Frame:** single frame for a specific spatial layer
- **Super Frame:** combo of all layer frames for a given input picture

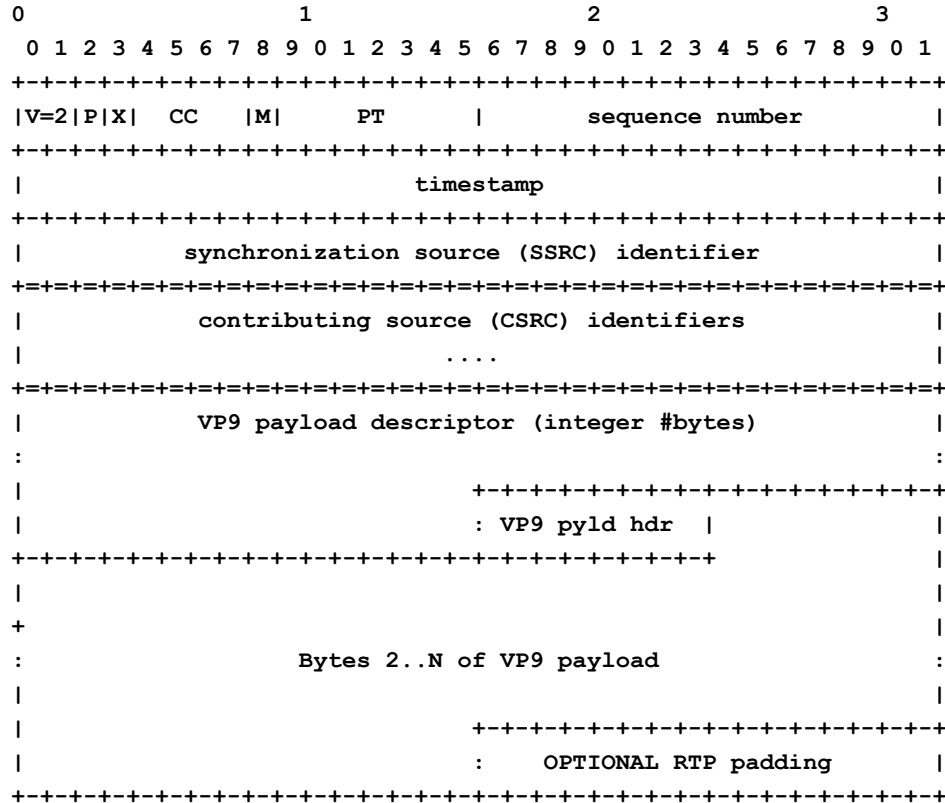
Packetization Features

- Specifies spatial and temporal indices
- Specifies scalability structure so receivers know how many layers are being sent
- Dependencies are explicitly indicated, either using specific indices, or a previously specified structure (SS = Scalability Structure)

Flexible vs Non-Flexible Mode

- **Flexible Mode:** each packet contains up to 3 indices of reference frames, which can be set arbitrarily
- **Non-Flexible Mode:** the reference indices of a group of frames (GOF) **MUST** be pre-specified in the scalability structure; each frame will have an index that refers to one of these pre-specified references.

Packetization



Flexible Mode Packetization

```

    0 1 2 3 4 5 6 7
+---+---+---+---+---+
|I|P|L|F|B|E|V|-| (REQUIRED)
+---+---+---+---+---+
I:  |M| PICTURE ID  | (RECOMMENDED)
+---+---+---+---+---+
M:  | EXTENDED PID | (RECOMMENDED)
+---+---+---+---+---+
L:  | T |U| S |D| (CONDITIONALLY RECOMMENDED)
+---+---+---+---+---+
P,F: | P_DIFF      |X|N| (CONDITIONALLY RECOMMENDED)
+---+---+---+---+---+
X:  |EXTENDED P_DIFF| (OPTIONAL)
+---+---+---+---+---+
V:  | SS           |
    | ..           |
+---+---+---+---+---+
                                     -\
                                     .
                                     . - up to 3 times
                                     .
                                     -/
```

Non-Flexible Mode Packetization

```

    0 1 2 3 4 5 6 7
+---+---+---+---+---+---+
|I|P|L|F|B|E|V|-| (REQUIRED)
+---+---+---+---+---+---+
I:  |M| PICTURE ID  | (RECOMMENDED)
+---+---+---+---+---+---+
M:  | EXTENDED PID | (RECOMMENDED)
+---+---+---+---+---+---+
L:  |GOF_IDX|  S  |D| (CONDITIONALLY RECOMMENDED)
+---+---+---+---+---+---+
    |  TLOPICIDX  | (CONDITIONALLY REQUIRED)
+---+---+---+---+---+---+
V:  | SS          |
    | ..          |
+---+---+---+---+---+---+
```


Usages

- **No layering:**
only initial byte used; only B/E bits used
- **Temporal layering (non-flex):**
PICID, GOF_IDX, TLOPICIDX (+SS sent with key)
- **Spatial + temporal layering (non-flex):**
PICID, GOF_IDX + S + D, TLOPICIDX (+SS as above)

Scalability Structure (SS) Format

```

      0 1 2 3 4 5 6 7
+---+---+---+---+---+---+
V:  | N_S |Y|  N_G  |
+---+---+---+---+---+
Y:  |          WIDTH          | (OPTIONAL)  . -\
+          +                      .
|          | (OPTIONAL)          .
+---+---+---+---+---+          . - N_S + 1 times
|          HEIGHT          | (OPTIONAL)    .
+          +                      .
|          | (OPTIONAL)          .
+---+---+---+---+---+          -/
N_G: |  T  |U| R  |-|-| (OPTIONAL)         .
+---+---+---+---+---+          -\
|          P_DIFF          | (OPTIONAL)    . - N_G + 1 times
+---+---+---+---+---+          -/
|          | (OPTIONAL)          . - R times  .
+---+---+---+---+---+          -/

```