No Frame Context Signaling

Thomas Daede
tdaede@mozilla.com

Requirements for Temporal Scalability

- It should be possible to determine and control which previously coded frames are dependencies of the current frame
  - E.g., allow skipping every other frame

Requirements for Error Resilience

- It should be possible to determine if the decoder is missing a frame required for decoding
  - This allows retransmission or dropping of the frame
  - Allows a decoder that never shows a broken frame

VP9

- Reference frame dependencies
  - Implicitly or explicitly signaled with picture IDs in RTP mapping
  - Up to three allowed per frame (from pool of 8)
- Frame contexts (probabilities)
  - Stores probabilities that are backwards-adapted based on data from previous frames
  - Decoder maintains four independent sets
  - Each frame signals which one to use
    - Optionally writes back to the slot it read from
  - Choice uncorrelated with references or picture IDs

Problems with Frame Contexts

- If you lose a frame, you don’t know which slot it updated
  - You no longer know if you can decode any frame
- The last frame to update the slot you’re using might not be one of your reference frames
  - Frame contexts introduce potential hidden fourth frame dependency (this is surprising)
  - RTP mapping only signals three picture IDs (could fix)
- You can’t fork probabilities and evolve them independently
  - Every layer pays cost of re-learning probabilities

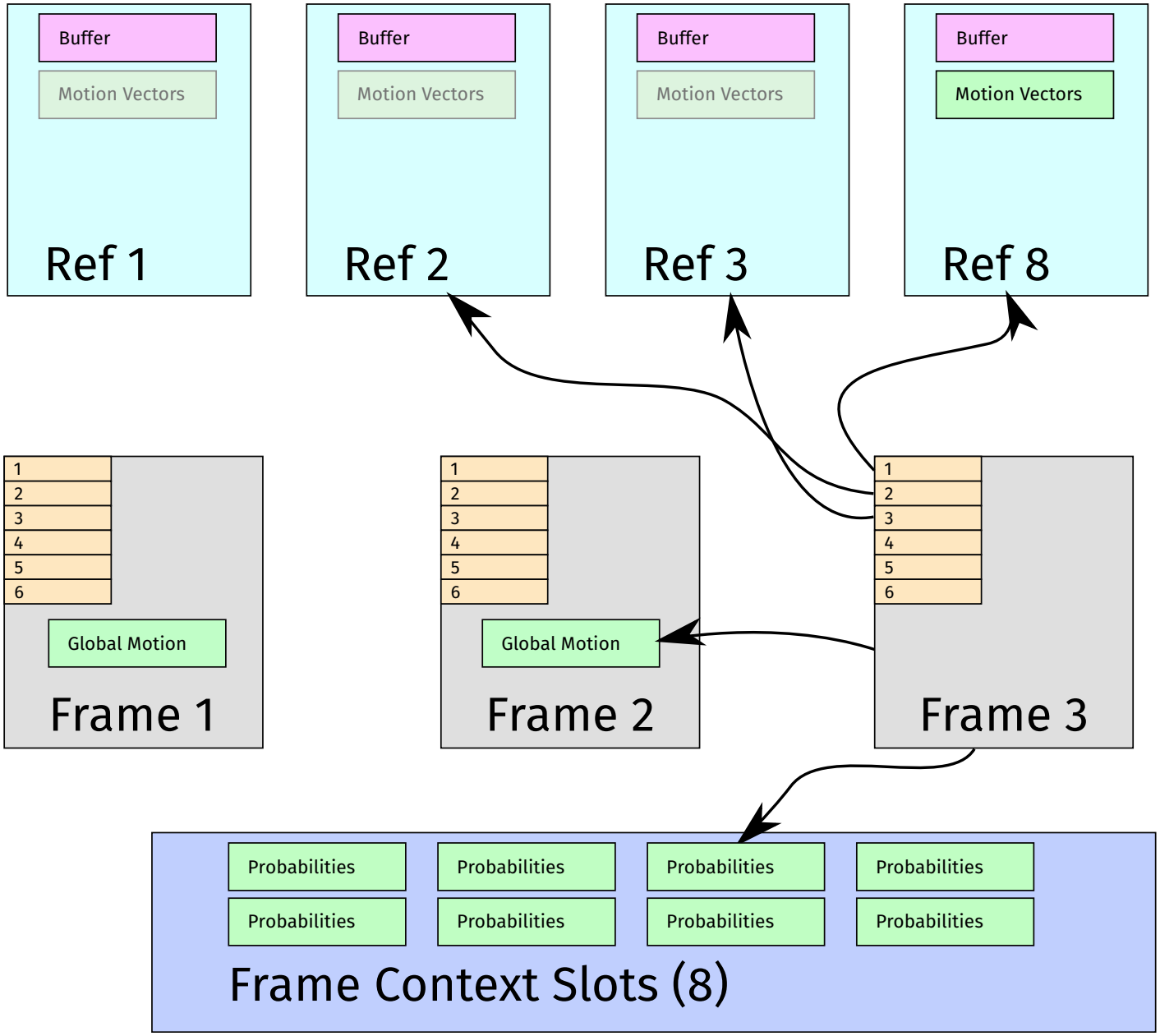
AV1

- Reference frame dependencies
  - Explicitly signaled and coded with frame IDs in codec payload
  - Up to six allowed per frame (from pool of 8)
- Probabilities
  - Currently same as VP9, but expanded to 8 slots
- Motion vectors for temporal MV prediction
  - Always from last coded frame
  - Fixed up by tempmv_signaling proposal
- Global motion data
  - Coded as deltas relative to last coded frame

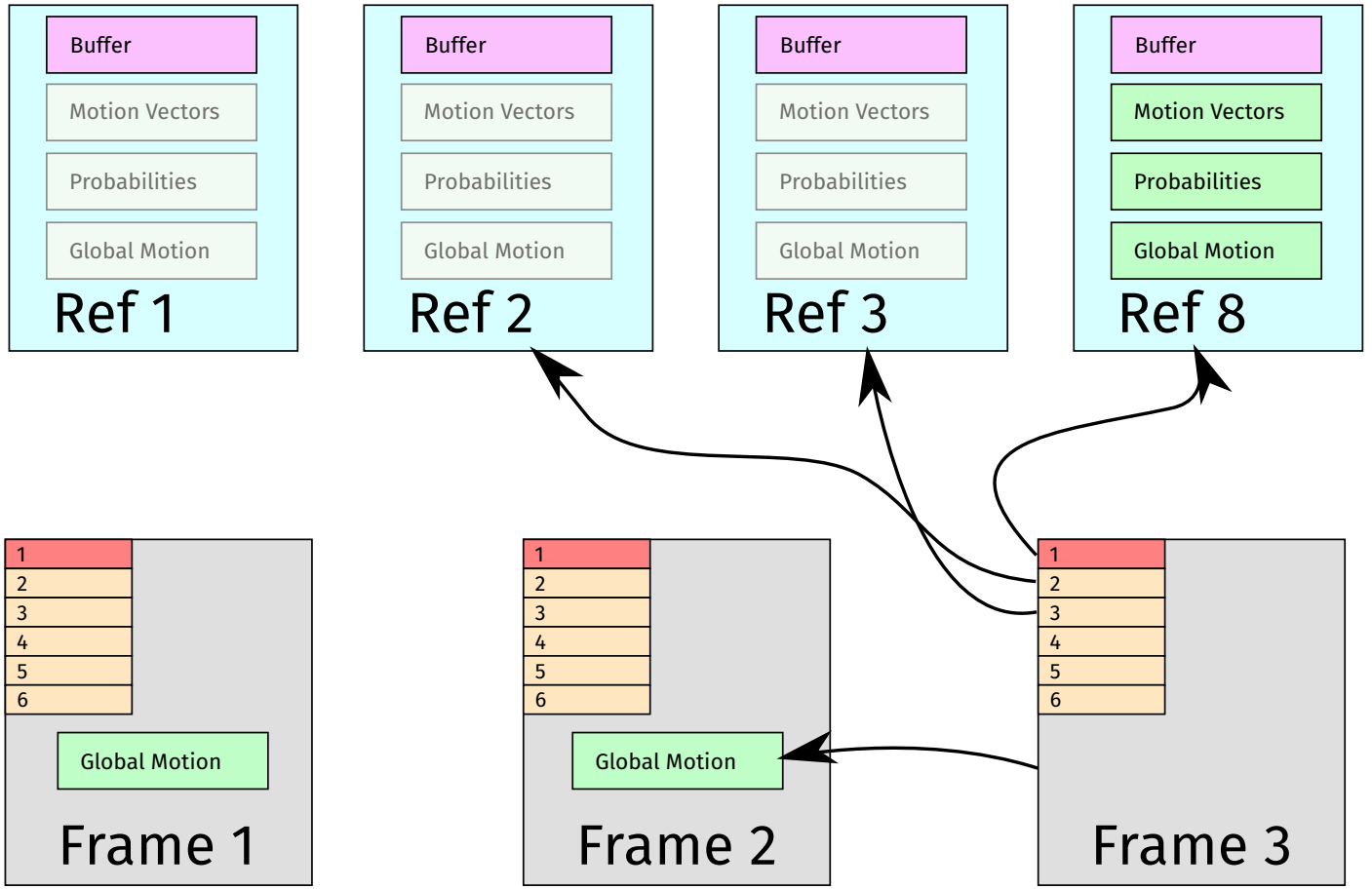
Proposal

- Make all dependencies between frames track with reference frame structure

Before


Buffer
Motion Vectors
Ref 1
Buffer
Motion Vectors
Ref 2
Buffer
Motion Vectors
Ref 3
Buffer
Motion Vectors
Ref 8
1
2
3
4
5
6
Global Motion
Frame 1
1
2
3
4
5
6
Global Motion
Frame 2
1
2
3
4
5
6
Frame 3
Probabilities
Probabilities
Probabilities
Probabilities
Probabilities
Probabilities
Probabilities
Probabilities
Frame Context Slots (8)

After


Buffer
Motion Vectors
Probabilities
Global Motion
Ref 1
Buffer
Motion Vectors
Probabilities
Global Motion
Ref 2
Buffer
Motion Vectors
Probabilities
Global Motion
Ref 3
Buffer
Motion Vectors
Probabilities
Global Motion
Ref 8
1
2
3
4
5
6
Global Motion
Frame 1
1
2
3
4
5
6
Global Motion
Frame 2
1
2
3
4
5
6
Frame 3

Details

- Remove frame context slots
- Remove all syntax elements for saving and restoring frame contexts
- Instead, always save frame context with reference buffers
  - When storing a reference frame, store updated probabilities, temporal MVs, etc., too
- No more syntax to reset frame contexts
  - Implicit on a keyframe

Complexities

- Interaction between reference number and function
  - In current encoder, first reference is always last frame from same layer
  - Need to re-order reference list to use probabilities from long-term reference (golden frame), alt-ref, etc.
- Using a previous frame context for intra-only frames
  - Currently not supported (same as VP9)
- Using probabilities from a non-reference frame
  - No longer supported
  - With up to 6 of 8 references per frame, impact seems low

TODO

- Still some things to move to frame context
  - Global motion
  - Frame size prediction