

Real-Time AV1 in WebRTC

Dr. Alex - CoSMo Software



CoSMo Software



**ALLIANCE FOR
OPEN MEDIA
RESEARCH**

Symposium 2019

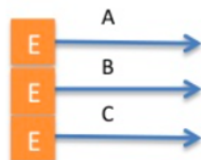
AOM :: USE CASES

- VOD, pre-recorded content, almost-live, live (3s): **CODEC ONLY**
 - All the time in the world to encode,
 - encoding / upload / storage / delivery all separated,
 - no filtering between the encoder and the decoder
 - Main cost on storage and delivery (bandwidth)
 - Only delivery and decoding is time sensitive
 - Quality is often most important than latency
 - Entire Ecosystem: cloud encoding, hardware encoders, decode-only, players,
...

AOM :: USE CASES

- Real-time (<1s): **CODEC** and **MEDIA TRANSPORT** with **SFU**
 - **Cisco Webex, Poly, Vidyo, CoSMo**, [Facebook Msg/whatsapp],
 - Latency is king,
 - Traditionally Simpler Encoder, single-frame encoding, no B-frames,
 - Encoder, Media Server, and decoder must ALL be real-time.
 - Real-Time requires end-to-end control, no storage, ...
 - Need to define the **Media Transport**, which will shoulder some of the RT properties
 - Need to define everything with **SFU** logic in mind
 - Codec / OBU is not enough.
 - Multiple deliverables: RTP payload, SDP O/A, RCTP (FIR, ...) support, SVC support
 - Non deliverable but needed IRL => BWE, CC, FEC, RED, RTX,
 - Test is challenging, since we now need end-to-end testing with SFU and filtering.

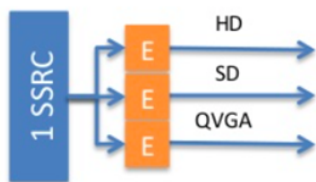
Reminder: “Multi-streams” vs Simulcast vs SVC



Multicast

- Several tracks
- Decodable separately
- Bandwidth management separated

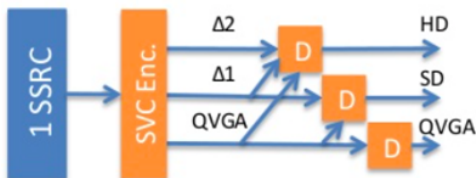
IN WEBRTC TODAY



Simulcast

- Several tracks
- Coming from the same source
- Decodable separately
- Smart bandwidth management possible

WEBRTC 1.0



SVC Encoding

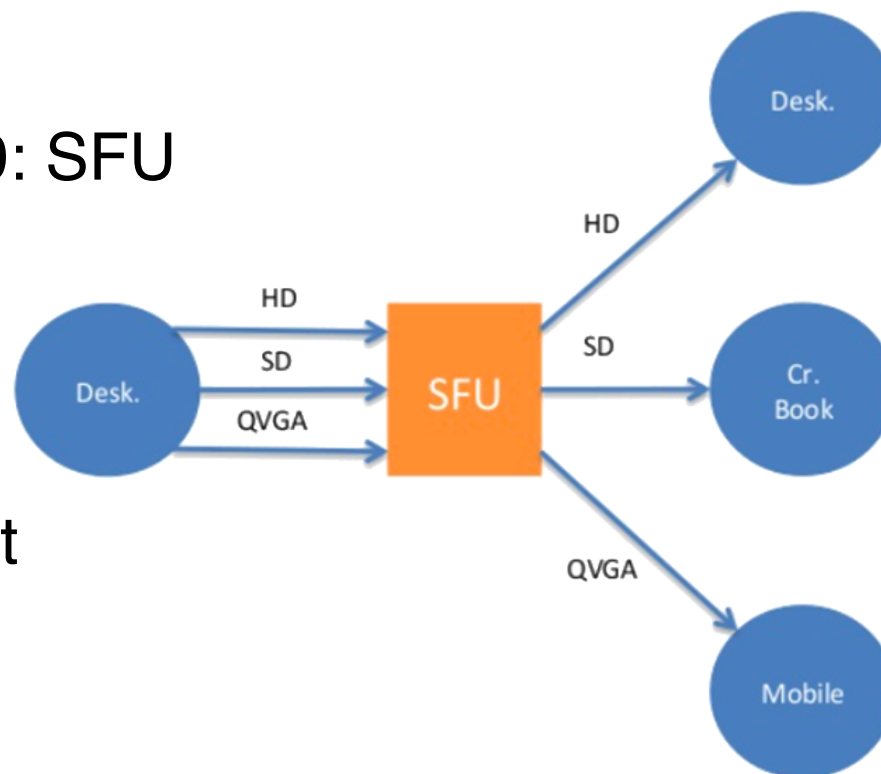
- Several tracks
- Coming from the same source
- Not Decodable separately (Except base layer)
- Smart bandwidth management mandatory
- Less bandwidth, more resilience.

WEBRTC IV
2016 at best

Simulcast / SVC: Use case for WebRTC 1.0

- Use case for WebRTC 1.0: SFU

- Browser send simulcast, does not receive simulcast (in WebRTC 1.0)



AOM :: RT :: USE CASES :: Video Conferencing

- Video Conference: e.g. Cisco Webex
 - Duplex
 - Everybody's equal
 - But the Active speaker is more interesting
 - Optimizations possible based-on voice activity detection, and Active Speaker
 - Echo cancellation mandatory,
 - Scaling is quadratic with respect to the number of users.
 - Cascading is possible but not mandatory
 - “The cisco dilemma”: supporting as much as possible existing hardware-based devices.

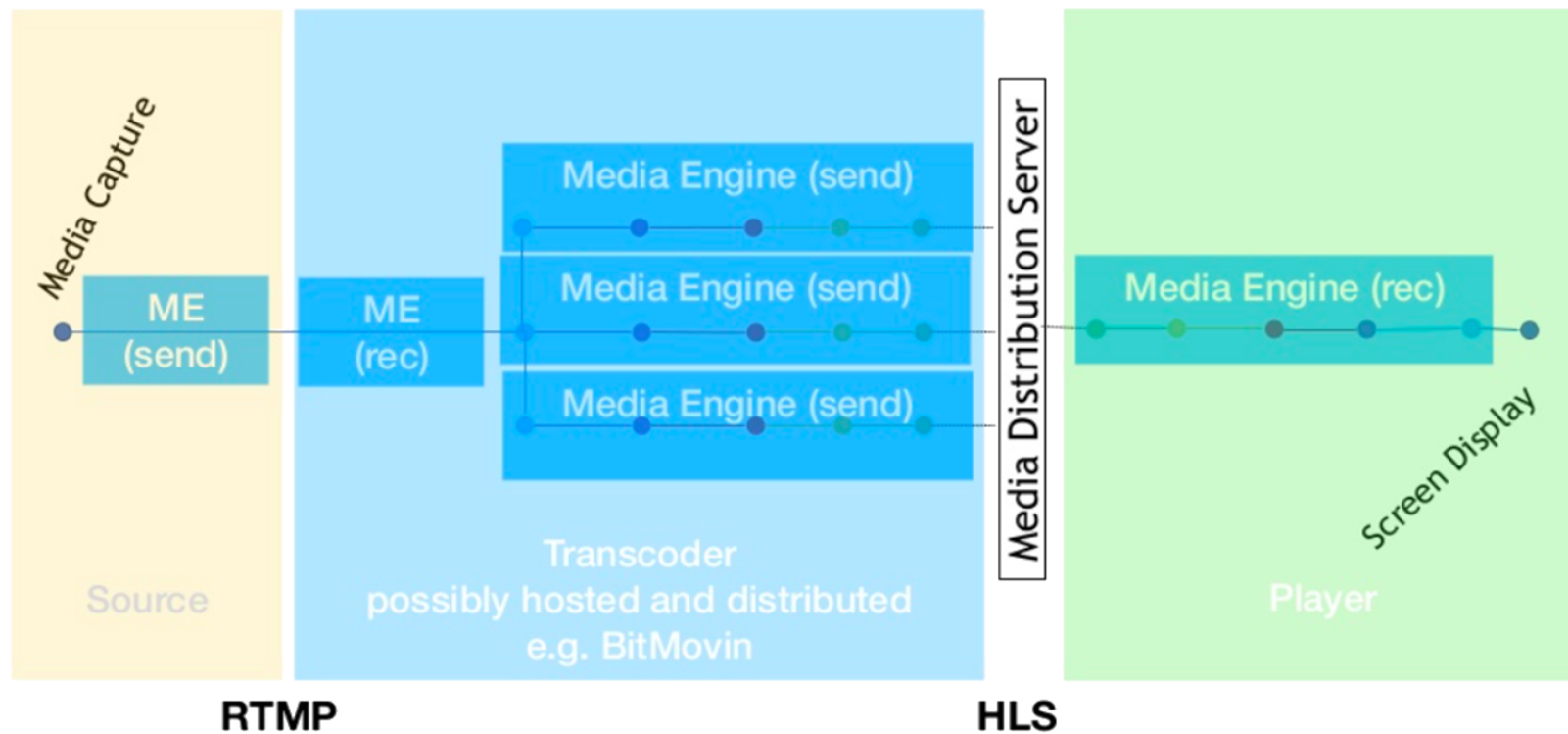
AOM :: RT :: USE CASES :: Streaming

- Streaming: e.g. MilliCast
 - One-way
 - Source and viewers with very different logic and capacity
 - No scaling optimization possible like in VC
 - Scaling is linear with respect to the number of users.
 - Cascading of servers is almost always needed
 - Real challenges to keep quality and network resilience at scale

AOM :: RT :: USE CASES :: Testing

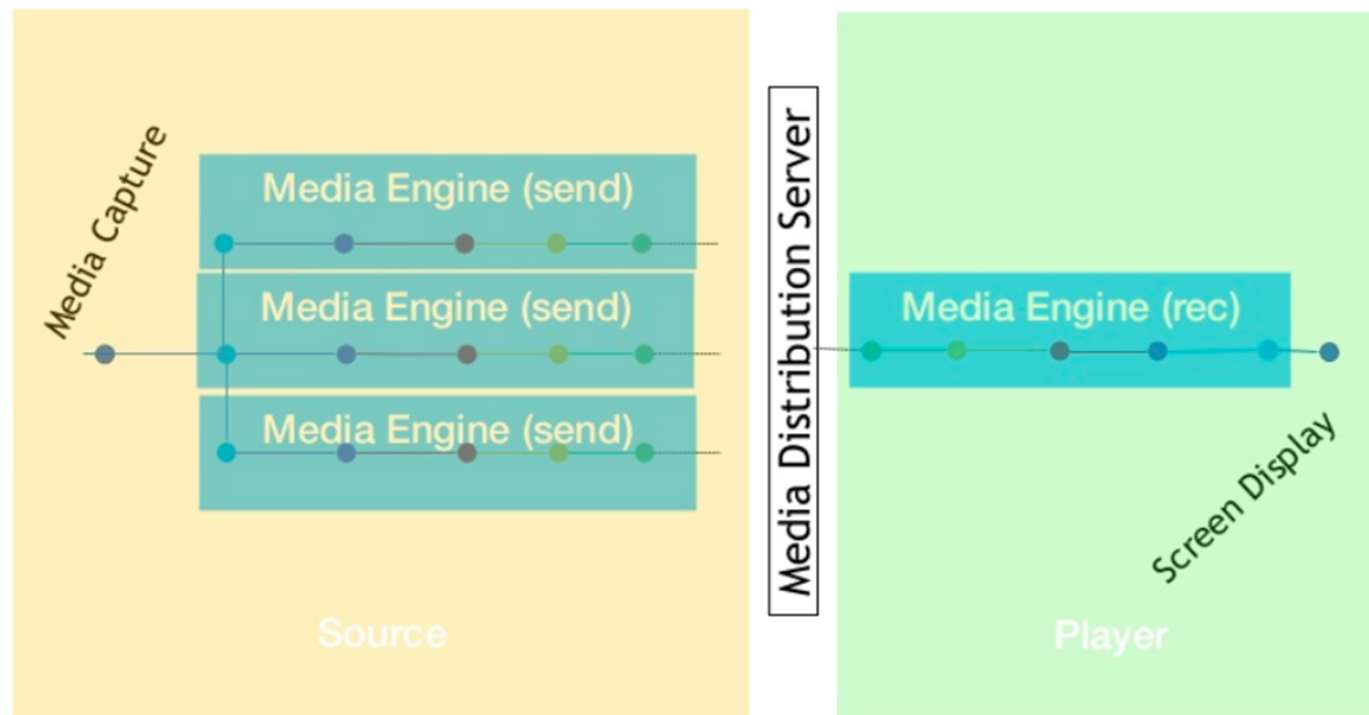
- In p2p mode, there is no difference between VC and streaming
- In 1 server mode, there is also no difference
- When you start serving more than 1,000 viewers, and/or need more than one media server in the media path, things start becoming ... interesting.

Bandwidth Adaptive Media Streaming Pipeline in practice - the usual



Bandwidth Adaptive Media Streaming Pipeline in practice - the usual

No storage
1 less enc/dec
= 50% load



Recent history of AV1 with focus on Real Time

- 03 2018, [AOMedia announced](#) the release of AV1 along with its reference implementation: libaom.
- 09 2018, [chrome 70](#) and Firefox nightly had added some kind of support for decoding / playing AV1.
- 10 2018, [CoSMo Software announced](#) the first AV1 integration in RTP and WebRTC. Not real-time, no SVC support.
- 12 2018, AOMedia Sponsored dav1d encoder has been released. It is included e.g. in [Firefox67](#),

- 01 2019, [CoSMo Software joins AOMedia](#).
- 04 2019, INTEL and NETFLIX, [announced their collaboration around the SVT-AV1 open-source codec](#).
- 04 2019, Allegro DVT [announced its AL-E210 multi-format video encoder hardware IP](#), the 1st (?) hardware AV1 encoder.
- 05 2019, [Realtek announced](#) the RTD2893, its first integrated circuit with AV1 decoding, up to 8K.
- 06 2019, [it announced](#) the RTD1311 SoC for set-top boxes with an integrated AV1 decoder.
- 06 2019, [Cisco demoed](#) of the first Real-Time AV1 integration in RTP and WebRTC (webex). **No SVC, not open-source.**
- 07 2019, [CoSMo Software released](#) a demo of Real-Time AV1 integration in RTP and WebRTC. **No SVC.** Open source.
- 08 2019, AV1 Availability in [MilliCast](#) is announced at IBC, along with [RealTime SSAI](#) (see next presentation)

<http://webrtcbydralex.com/index.php/2019/07/09/real-time-av1-in-webrtc-is-now-production-ready/>

II. AV1 as a payload for the Real-Time Protocol (RTP)

- AV1 OBUs \Leftrightarrow RTP packets: Easy
 - Modes
 - Fragmentation
 - Reconstruction
- RTP + (SVC + SFU) for bandwidth adaptation + E2EME: Hard(er)
 - Extend AV1 modes to be a better CPU / Network citizen: K-SVC
 - Simplify the decoding / filtering: “DTI” Decoding Target Information
 - Help Filtering without reading payload,
 - Manage Encrypted payloads (E2EME)

II. AV1 RTP Payload

- AV1 OBUs \Leftrightarrow RTP packets
 - Modes
 - Fragmentation
 - Reconstruction
- **Extend AV1 modes: K-SVC**
- Simplify the decoding / filtering: “DTI” Decoding Target Information
- Help Filtering without reading payload, Manage Encrypted payloads (EEME)

Benefits Of Shifted Temporal Prediction Structures

- More balanced bit distribution on the wire (reduced congestion/delay)
 - Rate distribution with concurrent TL0's: 54%, 15%, 16%, 15%
 - Rate distribution with shifted TL0's: 36%, 21%, 28%, 15% (variance ratio 4.5:1)



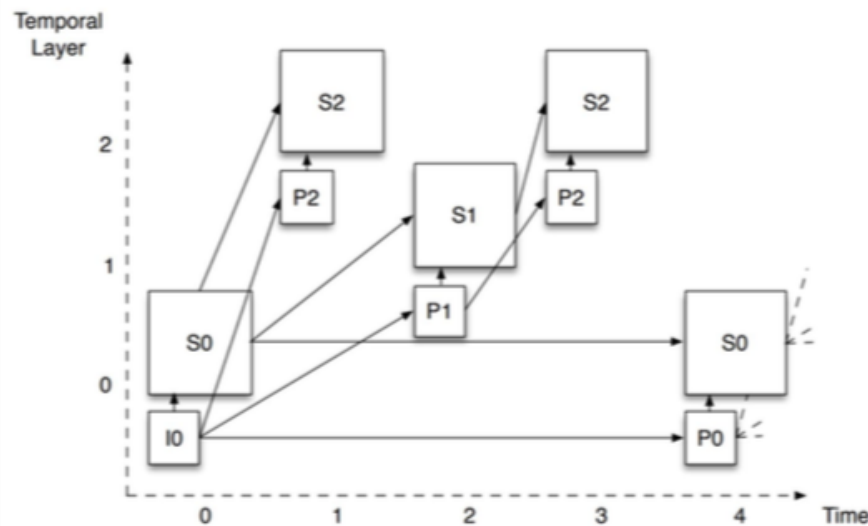
- More balanced CPU usage at encoder

II. AV1 RTP Payload

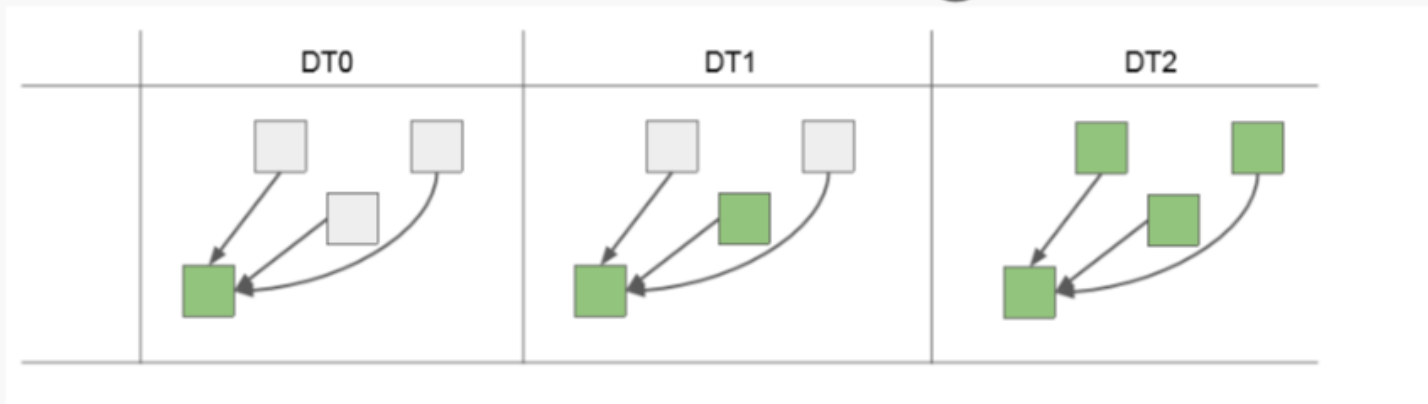
- AV1 OBUs \Leftrightarrow RTP packets
 - Modes
 - Fragmentation
 - Reconstruction
- Extend AV1 modes: K-SVC
- **Simplify the decoding / filtering: “DTI” Decoding Target Information**
- Help Filtering without reading payload, Manage Encrypted payloads (EEME)

Example: L2T3

- At Time 3, SFU decides to move from full resolution quarter framerate operating point to full resolution, full frame rate.
- Problem: P2 cannot be forwarded if P1 wasn't forwarded, S2 cannot be forwarded if S1 wasn't forwarded. Need to wait for S0 switching up point ("B" flag set).
- Can we determine this from the framemarking header alone (e.g. no access to operating point bitmask or scalability structure)?
- "Discardable" marking for frame P1 depends on the operating point, which is not provided.



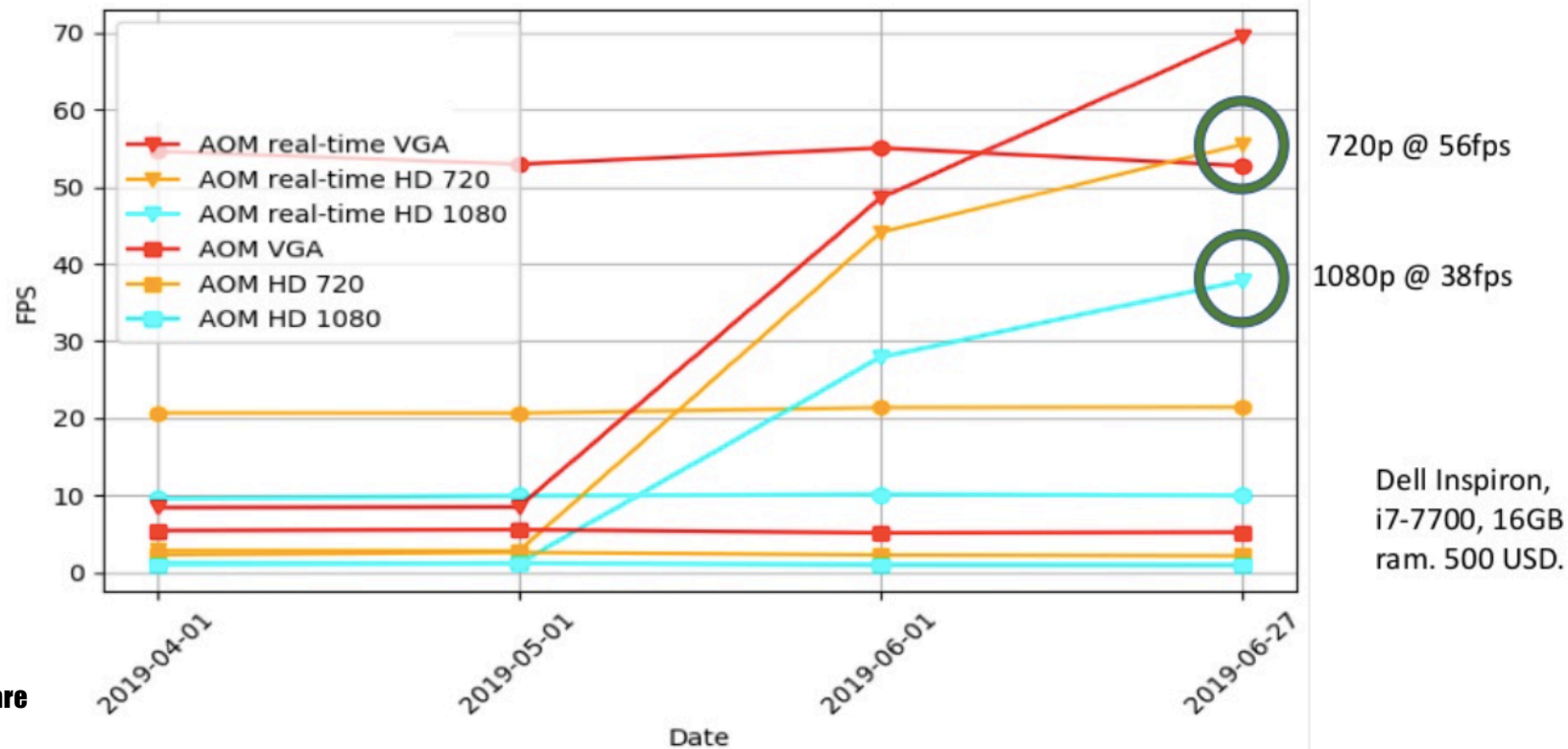
To the rescue: Decode Target Information



- Decode target: The set of frames needed to decode a coded video sequence at a given spatial and temporal fidelity.
- Decode Target Information (DTI): Describes the relationship of a frame to a Decode target.
 - 'not present': The frame does not belong to the Decode target.
 - 'discardable': The frame will not be a referred frame for any frame belonging to that Decode target.
 - 'switch indication': all subsequent frames for that Decode target will be decodable if the the frame containing the indication is decodable.
 - 'required': The frame belongs to the Decode target and has neither a Discardable nor a Switch indication. A frame belonging to more than one Decode target may be Required for one

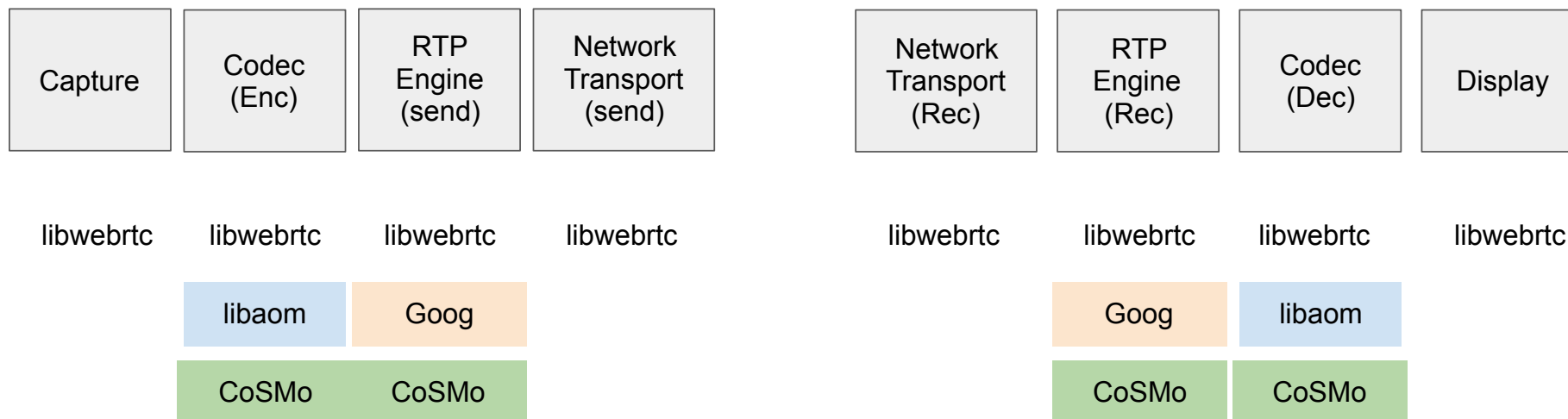
III. Open-Source Implementation

In may 2019, libaom team added a real-time mode, and improve performance greatly

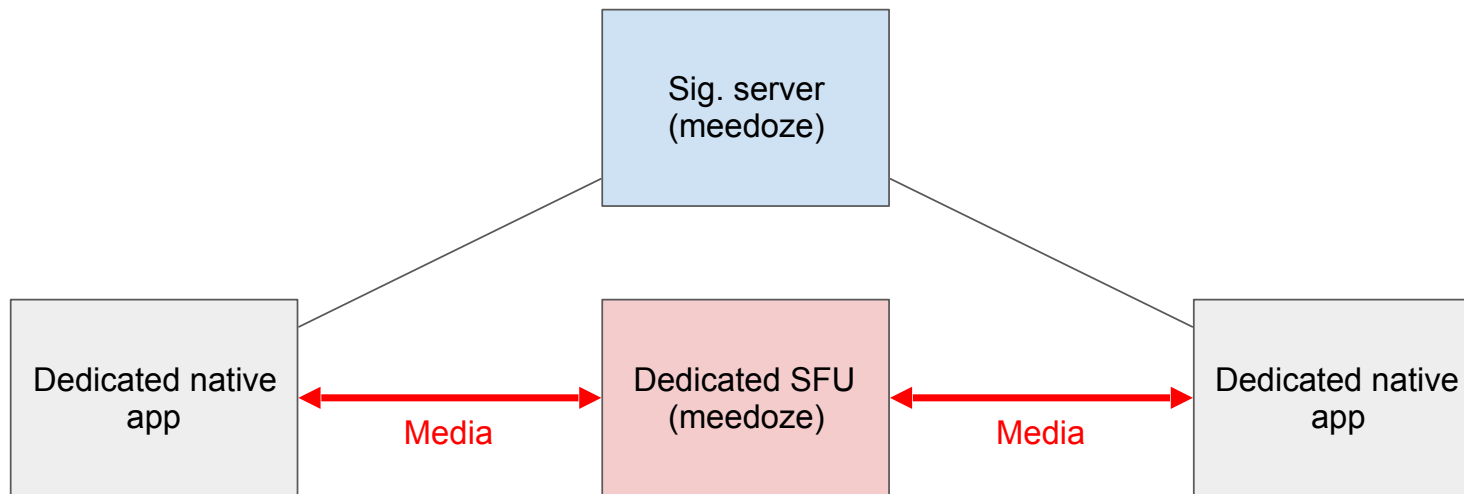


AV1 Minimum Open Source System (p2p)

- LibWebRTC is the webrtc stack implementation used in all modern browsers. If you use it, upstreaming to browsers for interoperability is but one patch (and some google goodwill) away ..



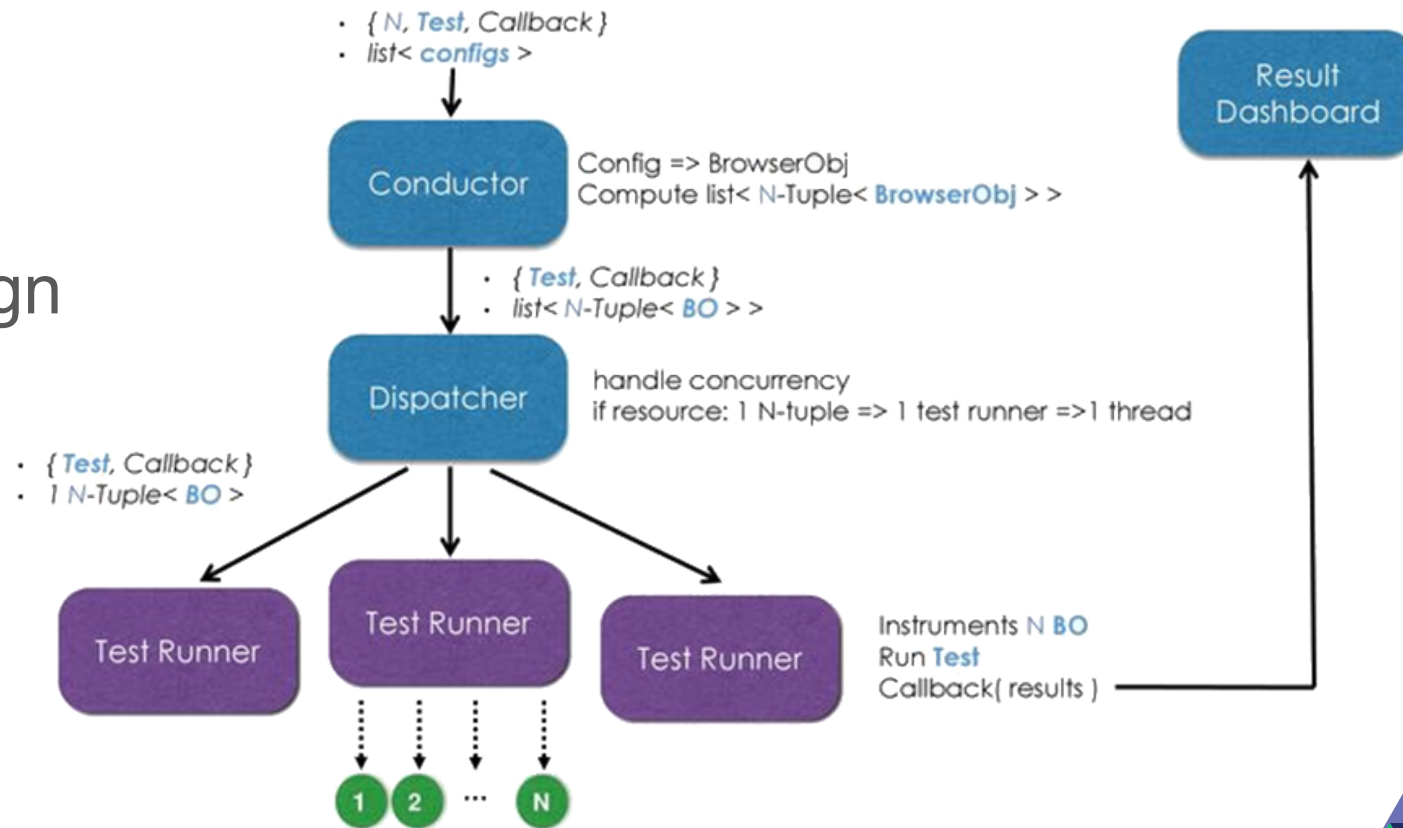
System Under tests (SFU)



Test SFU based filtering, layer changing, etc ...

KITE: Test automation for Communication Apps

KITE Design



Test Scenarios

- Start with existing webrtc test suite and adapt
 - Simulcast / SVC
 - Layer switching logic
- ...
- Could we test the matrix of all possible filtering (28 modes + k-SVC modes) exhaustively in reasonable time?

Questions?

Dr. Alex - CoSMo Software



**ALLIANCE FOR
OPEN MEDIA
RESEARCH**

Symposium 2019



CoSMo Software