

Machine Learning and Web Media

W3C Machine Learning Workshop

Bernard Aboba

Harald Alvestrand

Jan-Ivar Bruaroey

The Pandemic of 2020: A Pivotal Moment

- What if you could be transported into the next decade to see how communications would evolve?
 - To see what will compel and captivate consumers?
 - To glimpse how businesses will reinvent themselves?
 - To see how machine learning will be used?
- Amidst the horror, an unparalleled burst of *user-driven communications innovation* has taken place, a rethinking of:
 - Politics
 - Art
 - Entertainment
 - Sports
 - And more...
- “You can observe a lot, just by watching” - Yogi Berra
 - What have ***you*** observed?
 - Here are some of the things in my scrapbook...



Tacoma Little
Theatre 2020
("Robin Hood")

Source: <https://www.youtube.com/watch?v=Z3WU6SrX3mM>



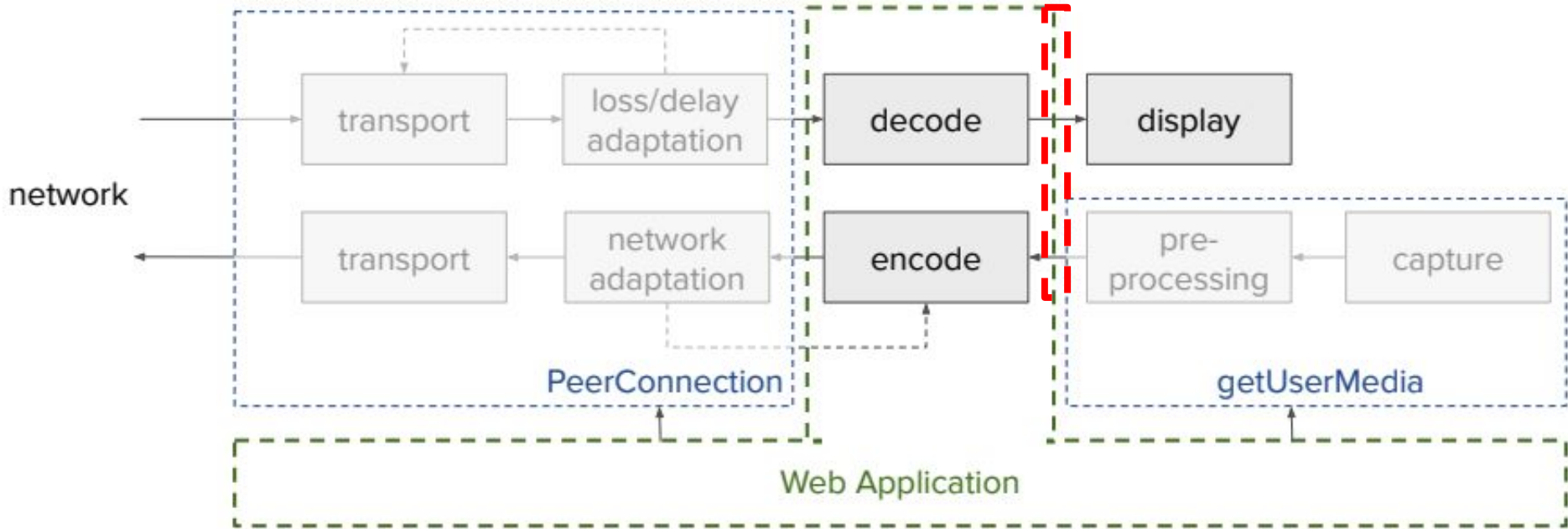
NBA “Together Mode”

Microsoft Teams at NBA arenas. | Microsoft

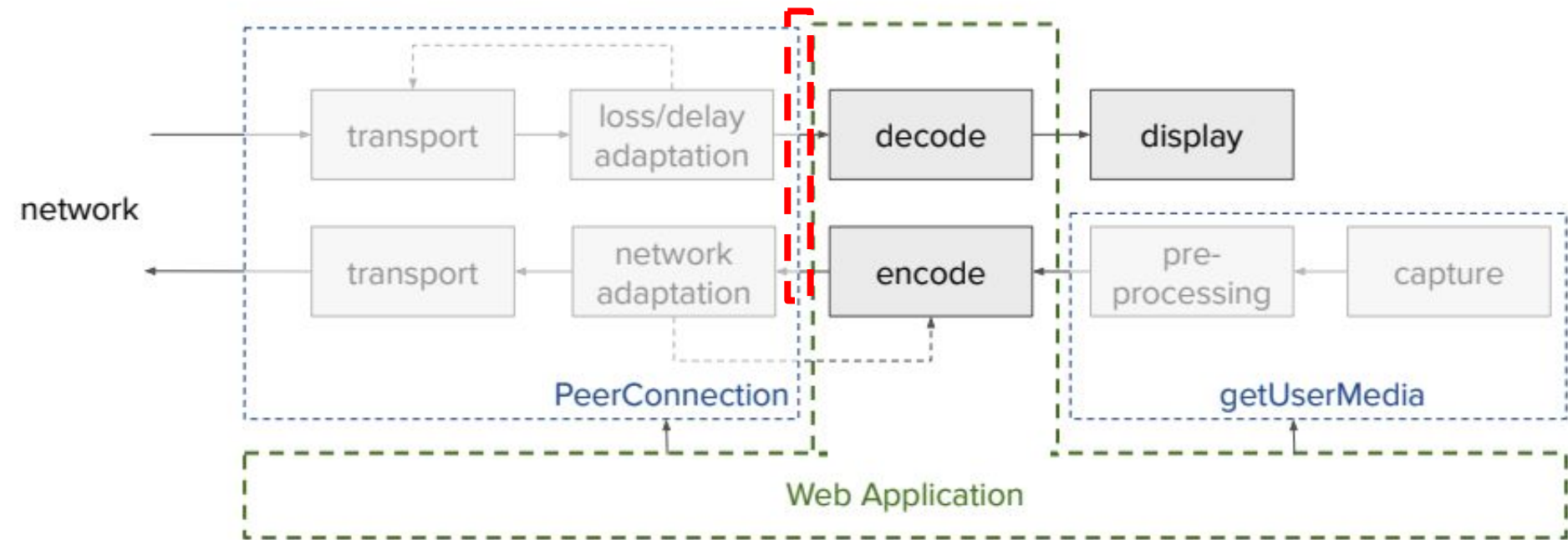
The National Basketball Association (NBA) is using Microsoft Teams' new Together Mode to place basketball fans courtside in a virtual experience during live games. Microsoft only just revealed [Together Mode for Teams earlier this month](#), and it uses AI to segment your face and shoulders and place you together with other people in a virtual space.

Source: <https://www.theverge.com/2020/7/24/21337326/nba-microsoft-teams-together-mode-basketball-virtual-experience-fans>

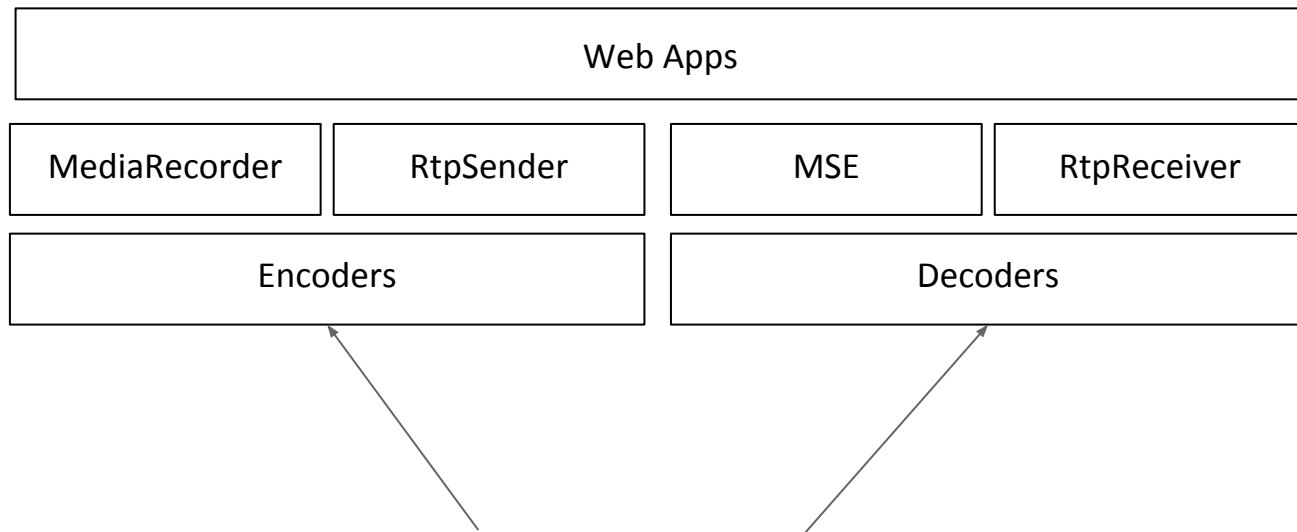
Machine Learning and the Media Pipeline



Insertable Streams - Encoded Media



The Current State of Encode/Decode APIs



These are currently hidden.
WebCodecs fixes that.

Insertable Streams - Relation to other efforts

WebCodec

- Aims to reuse “VideoFrame” and “AudioFrame” types
- Experience will be fed back to that effort

TransferableStreams

- Allows processing in WebWorkers
- Uses unmodified proposal
- Origin trial will also enable TS

AV1 Support

- AV1 is often described as “the codec of the future”.
- What if it were possible to leverage AV1 **sooner**?
 - [AV1 integrated in libwebrtc](#) (test harness under development)
- What if you could:
 - Build a Selective Forwarding Unit (SFU) that could [forward e2e encrypted payloads from any codec without parsing the payload?](#) (AV1 Dependency Descriptor)
 - Dramatically lower bandwidth for [screen content coding](#)? (“text” content-hint)
 - Selectively utilize AV1, such as for [\(low latency\) decode](#)? (RTCRtpReceiver.playoutDelay, getCapabilities(kind))
 - Support [mixed-codec simulcast](#): send AV1 in a low bitrate stream, while using other codecs (H.264/VP8/VP9) at higher bitrates.
 - Support [temporal and spatial scalability](#)

Thank you!