Joint Media/Audio/WEBRTC WG Meeting

October 26, 2021
7:00 - 9:00 AM Pacific Time

W3C WG IPR Policy

- This meeting abides by the W3C Patent Policy https://www.w3.org/Consortium/Patent-Policy/
- Only people and companies listed at https://www.w3.org/groups/wg/webrtc/ipr or https://www.w3.org/groups/wg/audio/ipr or https://www.w3.org/groups/wg/media/ipr are allowed to make substantive contributions

W3C Code of Conduct

- This meeting operates under W3C Code of Ethics and Professional Conduct

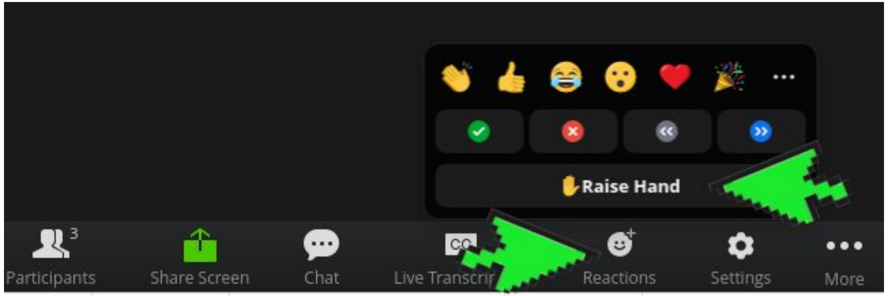
- We're all passionate about improving the Web, but let's all keep the conversations cordial and professional

About this Meeting

- Meeting info:
  - https://www.w3.org/events/meetings/81d215c7-f4c2-4697-a4b4-985efba5427e
- Link to Slides has been published on WG wiki
- Scribe? IRC http://irc.w3.org/ Channel: #webrtc
- Volunteers for note taking?

Virtual Meeting Tips

- Use raise hand (in the reactions menu) on Zoom to get into and out of the speaker queue.
- Please use headphones when speaking to avoid echo.
- Please wait for microphone access to be granted before speaking.
- Please state your full name before speaking.
- If you want to explore the current status of next generation media APIs (and live dangerously), join Zoom from the browser.



Welcome!

- Welcome to the joint meeting of the Media, Audio and WebRTC WGs at TPAC 2021!
- At this meeting, we will attempt to cover:
  - The State of Next Generation Media APIs
  - Potential gaps and problems to work on

Topics for Discussion Today

- 08:10 - 08:30 AM Next generation media APIs and potential gaps (Bernard)
- 08:30 - 08:50 AM WebRTC and WebCodecs (Harald)
- 08:50 - 09:10 AM WebCodecs Challenges (Chris Cunningham)
- 09:10 - 09:30 AM Audio Challenges (Paul)
- 09:30 - 09:50 AM Next generation audio codecs (Harald?)
- 09:50 - 10:00 AM Wrap-up and Next Steps

Time control:

- A warning will be given 2 minutes before time is up.
- Once time has elapsed we will move on to the next item.

Some ground rules...

The goal of this meeting is to identify problems that need to be solved. We will not solve them here.

Some have already been filed as Issues. Others lurk in the shadows.

Next Generation Media APIs

End Time: 8:30 AM

Once Upon a Time… Before the Pandemic

Streaming and realtime communications technologies evolved in ***silos***, with distinct APIs and protocol stacks.

Streaming offered ***large scale*** but ***high delay*** as well as ***content protection***.

Realtime communications offered ***low delay***, but ***modest scale***.

The Pandemic: 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?
- Amidst the horror of Pandemic, we have seen an unparalleled burst of *user-driven innovation in communications,* a rethinking of:
  - Politics
  - Education
  - Arts and Entertainment
  - Sports
  - And more…
- “You can observe a lot, just by watching” - Yogi Berra
  - What have ***you*** observed?
  - In the back of this deck are some of pages from my scrapbook...

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

Next generation Web media APIs

Enable multi-threaded applications to deliver ***both*** low-latency ***and*** large scale, through low-level access to building blocks:

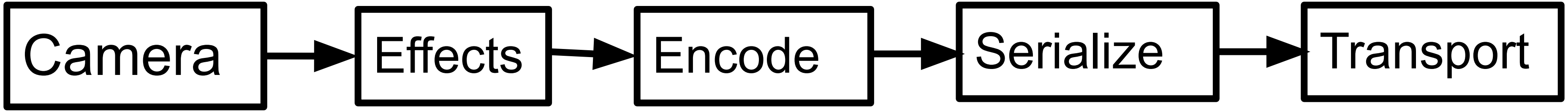
- Capture
- Encode/Decode
- Transport
- Rendering

Next generation Web media APIs

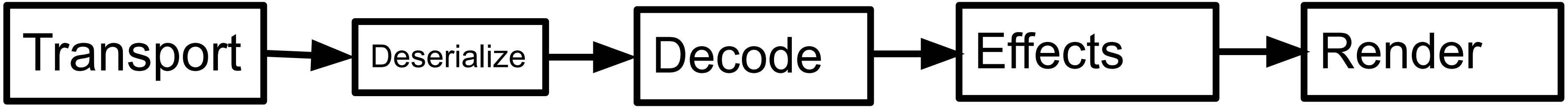
- Capture
  - Media Capture and Streams Extensions
  - Mediacapture-transform
- Encode/decode
  - WebCodecs
  - MSEv2
- Transport
  - WebTransport (HTTP/3 over QUIC)
  - WebRTC data channel in Workers (SCTP/DTLS/UDP)
- Framework
  - WHATWG Streams
  - Web Assembly

The “Pipeline” Model

- Send



- Receive



The “Pipeline” Model in Javascript

- Send

```
inputStream
    .pipeThrough(SpecialEffects())
    .pipeThrough(EncodeVideoStream(config))
    .pipeThrough(Serialize())
    .pipeTo(SendStream(transport))
```

- Receive

```
ReceiveStream(transport)
    .pipeThrough(Deserialize())
    .pipeThrough(DecodeVideoStream())
    .pipeTo(outputStream)
```

Transport on the Web

<table>
<tr><th></th><th>Client-Server</th><th>Peer-to-peer</th></tr>
<tr><td>Reliable and ordered</td><td>WebSocket
(also WebTransport!)</td><td rowspan="3">RTCDataChannel
(SCTP/DTLS/UDP)</td></tr>
<tr><td>Reliable but unordered</td><td rowspan="2">WebTransport
(HTTP3/QUIC)</td></tr>
<tr><td>Unreliable and unordered</td></tr>
</table>

Does “The Narrative” Hang Together?

Or are we missing a ***little*** something?

Or perhaps some ***big*** somethings?

Little Somethings (Workers)

Worker support across the ***entire*** media pipeline is not yet a reality.

No browser currently supports ***both*** MSEv2 and RTCDataChannel in workers.
But we are working on the specifications.

Is there a fundamental gap in our worker story? Do we need a meta-specification?

Little Somethings (Testing)

If low latency is a goal, we need to ***specify*** and ***test*** for it.

Example: “low latency MSE” is ***neither*** well specified, ***nor*** tested in WPT.

To expand WPT test coverage, WebTransport and WebRTC WGs are adopting an ***Echo Test framework***.

Would this make sense for WebCodecs?

Little Somethings (Performance)

- Seams between APIs (e.g. copies)
- Issues with WHATWG Streams APIs for media processing

We have more ***opinions*** on these problems than ***data***, at the moment.

Do we have the performance requirements, data and processes (e.g. joint meetings) to follow up and bring issues to closure?

Little Somethings (Transport)

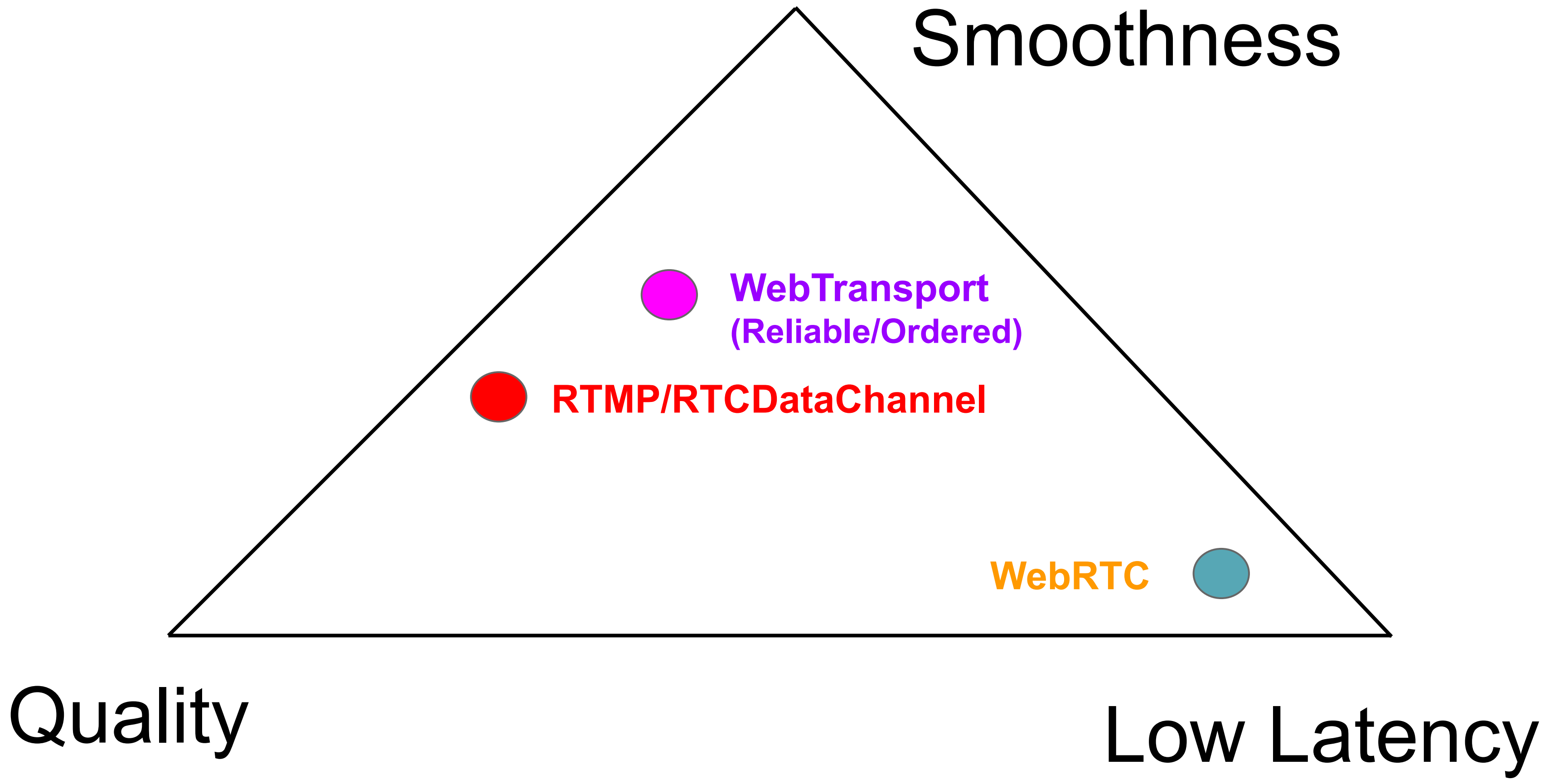
How do we transport WebCodecs with low latency?

a. Access to WebRTC RTP transport? (See Harald's slides)
b. Support for low latency congestion control for WebTransport or RTCDataChannel?

Today, low-latency RTCDataChannel streaming from server -> client can be achieved by replacing RTCDataChannel congestion control on the sender side.

But what about use cases like low latency video ingestion (client -> server) or bi-directional communications (e.g. video conferencing)?

Transport Tradeoffs



Little Somethings (Congestion Control/Resilience)

1. Encoder “Average bitrate targets” overshoot (keyframe) and undershoot (delta) the target.
2. After loss, the “average bitrate target” may be reduced, but if keyframe(s) are required to recover, a bandwidth spike will result.
3. Scalable Video Coding (SVC) can reduce the need for recovery (e.g. discardable frames need not be retransmitted).
    a. This assumes that decoders can carry on without discardable frames, which may require decode target filtering.

Bigger Somethings (Missing Pieces)

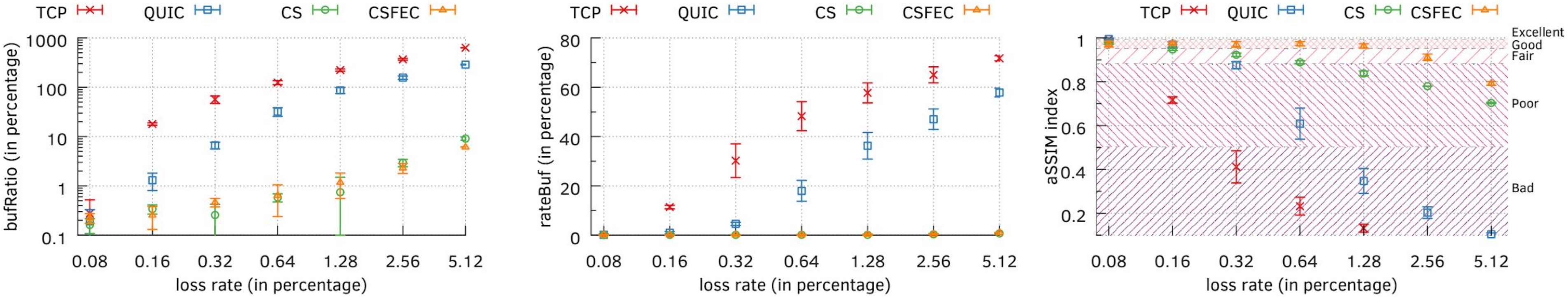
A new generation of SFUs.

Bigger Somethings (Missing Pieces)

A new generation of streaming and ingestion protocols.

Some Promising Results...

- Next generation streaming video based on QUIC datagrams
  - Avoids retransmission of frames that subsequent ones don't depend on
  - Result: lower latency, fewer video freezes
- Paper: "The QUIC Fix for Optimal Video Streaming": https://arxiv.org/pdf/1809.10270.pdf



(a) Playback time spent in buffering.
(b) Relative frequency of stall events.
(c) Adjusted SSIM of reference & received frames.

Figure 3: ***ClipStream (CS) and ClipStreamFEC (CSFEC) outperform TCP and QUIC across a wide range of loss values.***

Bigger Somethings (Missing Pieces)

WebCodecs and Content Protection.

a. How do opaque Audio/Video Frames work?
b. What format is sent over the wire?

See Chris’s slides.

WebRTC and WebCodecs (Harald Alvestrand)

End Time: 8:50 AM

WebRTC and WebCodecs - basics

Note: All scenarios described apply equally to outgoing and incoming data. Brevity rules.

WebRTC sending pipeline

- MediaStreamTrack
- RTCRtpSender
    - Encoder
    - Packetizer
    - Transmitter
- Feedback from congestion control to encoder modifies data rate dynamically
- Keep-the-video-rolling is paramount

WebCodecs via Breakout Box*

- MediaStreamTrack
- Stream<VideoFrame>
- WebCodec encoder
- Non-WebRTC packetization and transmission

(*) BreakoutBox is a proposal, not a W3C TR

Encoded data via Insertable Streams

- Encoding with RTCRtpSender
- Post-processing of data as a Stream<NotQuiteTheSameEncodedFrame>
- Transmission back into RTCRtpSender

Attempts to make more creative connections won’t work.

Why Real Time and Stored Differ

Keep media flowing

- Congestion: Drop frames, reduce resolution/FR/quality
- More BW: Increase
- SVC allows reduction without resync

Deliver stored media

- Congestion: Send slowly (and spin)
- Possibly switch to a different (encoded) source
- More BW: Catch up, switch back
- No SVC

Some Desirable Patterns

- Connect incoming encoded stream to outgoing encoded stream
- Distribute one encoded stream to multiple destinations
- Mix and match WebTransport and WebRTC RTP transport as needed
- Control delay & reactions to packet loss

WebCodecs and Configuration

- The configuration language of WebRTC is SDP
- We would like to not spread SDP further
  - It is an old, clunky and not well-loved language
- Expressive power equivalence is desirable
  - If we could do it in SDP, we still want to do it (sort of)
  - The concept of “negotiate envelope, control instant” is one that bears preserving
- Investigation has not yet started

Conclusions

- WebCodecs and WebRTC are powerful tools
- What we need depends on the purpose.
- Some patterns fit together. Some don’t.
- More learning needed!

Discussion

WebCodecs Challenges (Chris Cunningham)

End Time: 9:10 AM

Containers

- FAQ: how do I (de)mux?
- Answer: find a JavaScript / WASM library

I like the answer, but "find" part isn't great.

Exploring libavformat -> WASM this Q.

Will assess viability, do a write up.

Reclamation

- The UA can reclaim your codec to ensure codec availability for "foreground" apps
- Challenge: identifying when apps are semantically in the foreground is tricky.
- ex: low fps apps broken by screenshare
- ex: background video production "rendering"
- We're exploring solutions (hints, locks, ...)

Content Protection

- https://github.com/w3c/webcodecs/issues/41
- A few feature requests. Makes sense (low latency streaming will often want).
- Proposal: EME, not SFrame
- Challenge: rendering.
  - "write only" Canvas could work for software CDM.
  - maybe use generated MediaStream for hardware CDM

Discussion

Audio Challenges (Paul Adenot)

End Time: 9:30 AM

Audio is hard real-time

Symmetry is nice, but audio doesn’t have the same challenges as video.

Audio data must only ever touch real-time threads, after being decoded. Any other setup will lead to resilience issues under load.

The only missing bit

It’s better to buffer on the real-time thread side, and not mix the push and pull models.

The bit of information that is missing is to tell an `AudioWorkletProcessor`’s `process()` method that there was an input underrun (when using `MediaStream`s).

Strawman

```
class processor extends AudioWorkletProcessor {
...
  process(inputs, output, params) {
    // inputs[0][0] is the first channel of the first input
    // of this AudioWorkletProcessor
    if (inputs[0][0].length == 0) { … input underrun }
  }
}
```

Useful not only for WebRTC: HTMLMediaElement playing via MSE or HTTP can underrun too !

Web Audio + Web Codecs = OK

Decoder:

https://github.com/chcunningham/wc-talk/blob/main/audio_renderer.js

Renderer:

https://github.com/chcunningham/wc-talk/blob/main/audiosink.js

`SharedArrayBuffer`-based SPSC ring buffer, tuneable low-latency playback, underrun detection, etc.

Discussion

Next Generation Audio Codecs (Harald Alvestrand)

End Time: 09:50 AM

Why New Audio Codecs?

- OPUS works great for many purposes
- Other codecs *may* work better
  - Extreme low bandwidth
  - Packet loss tolerance
  - Meaning-based encodings
- But how do we deploy them?

How to deploy a new audio codec

- Obtain the licenses
- Run the experiments
- Integrate into open-source codebases
- Make generally available on all platforms
- Win

…. We’d like to have step 5 right after step 2….

A New Codec Deployment Paradigm

- Performant interfaces to raw and encoded data
  - Need precise timing guarantees on playout!
  - Minimize underruns. All else is timestamps.
- Anything that looks like a codec can be a codec
- WASM for deployment and experimentation
- Integration when value is proved

Discussion

Pandemic Scrapbook

America’s Got Talent

Source:
https://www.usatoday.com/story/entertainment/tv/2020/08/11/americas-got-talent-simon-cowell-live-show-injuries/3350676001/

Politics

https://www.youtube.com/watch?v=K-tWw9UsLwI

Art Basel 2019
(Miami Beach)

“Virtual Window”
Jerusalem

(Art Basel 2019)

Art Basel 2020

Bloomberg

Visiting Art Basel From the Hamptons Will Test Online Model

Katya Kazakina · 2 days ago

(Bloomberg) -- For the past 25 years, Christophe Van de Weghe never missed the event that transforms a quiet Swiss town into the art world's epicenter each June.

Art Basel has become a destination for collectors, with more than $3 billion of modern and contemporary blue-chip works offered. As an exhibitor, Van de Weghe spends as much as $200,000 to ship works by Pablo Picasso, Alexander Calder and Jean-Michel Basquiat from New York, insure them, pay for the booth, host client dinners at Chez Donati and stay at the Three Kings hotel. On the eve of each VIP opening, he jumps into the frigid Rhine for a 3-mile swim.

Plunging Revenue

Some galleries that recently reopened are using their spaces to exhibit works offered at Art Basel's online viewing rooms. The stakes are high. Online platforms became a lifeline for galleries and auction houses during the lockdown. More than 150 U.S. galleries expect second-quarter revenue to plunge 73% from a year earlier, according to an Art Dealers Association of America survey.

Wealthy buyers snapped up works by emerging talents during the lockdown, asking for discounts of as much as 30%, according to dealers and advisers. But pieces above $5 million were a much harder sell because collectors usually want to see them in the flesh.

"Do you think anyone will buy an $8 million painting online?" Van de Weghe said. "I don't. You need to see the presence of it."

Art Basel will test this notion.

About five hours into the fair's VIP opening, David Zwirner gallery said it found a buyer for a new $8 million sculpture by Jeff Koons, "Balloon Venus Lespugue (Red)," which was offered only online. Skarstedt gallery will offer a Willem de Kooning painting for $8.5 million online and at its new East Hampton branch.

In London, Richard Nagy hung works by Kees van Dongen, Egon Schiele and Henri Matisse at his gallery's reopening on Monday. They'll be featured on Art Basel's website, with prices as high as $5 million.

Source: https://www.msn.com/en-us/finance/markets/visiting-art-basel-from-the-hamptons-will-test-new-online-model/ar-BB15C9zT

Virtual Concerts

POP

Here Are All the Livestreams & Virtual Concerts to Watch During Coronavirus Crisis (Updating)

6/23/2020 by Billboard Staff

Big Hit Entertainment

BTS

ENTERTAINMENT

All of the Concerts You Can Watch From Home Right Now

By Esther Zuckerman, Sadie Bell, and Dan Jackson Updated on 3/30/2020 at 4:10 PM

Source:
https://www.billboard.com/articles/columns/pop/9335531/coronavirus-quarantine-music-events-online-streams

Garth Brooks Concerts

Source: https://www.youtube.com/watch?v=XQvCS3gMASw
https://www.youtube.com/watch?v=fWI_q3cAIdM

Thank you

Special thanks to:

WG Participants, Editors & Chairs