IBC2024

#ACCELERATORS2024

SCALABLE ULTRA-LOW LATENCY STREAMING FOR PREMIUM SPORTS

14 SEPTEMBER 2024

ALEX GILADI

COMCAST FELLOW.

PROBLEM STATEMENT

GOAL

Provide a streaming solution for premium sports experience

REQUIREMENTS

- **Latency** same or better than terrestrial broadcast, equivalent to social media
  - 1-3s latency needed to maintain fan engagement
  - Good enough for betting
- Tight **synchronization** across all viewers
- **Scalability**
- **Non-proprietary solution**

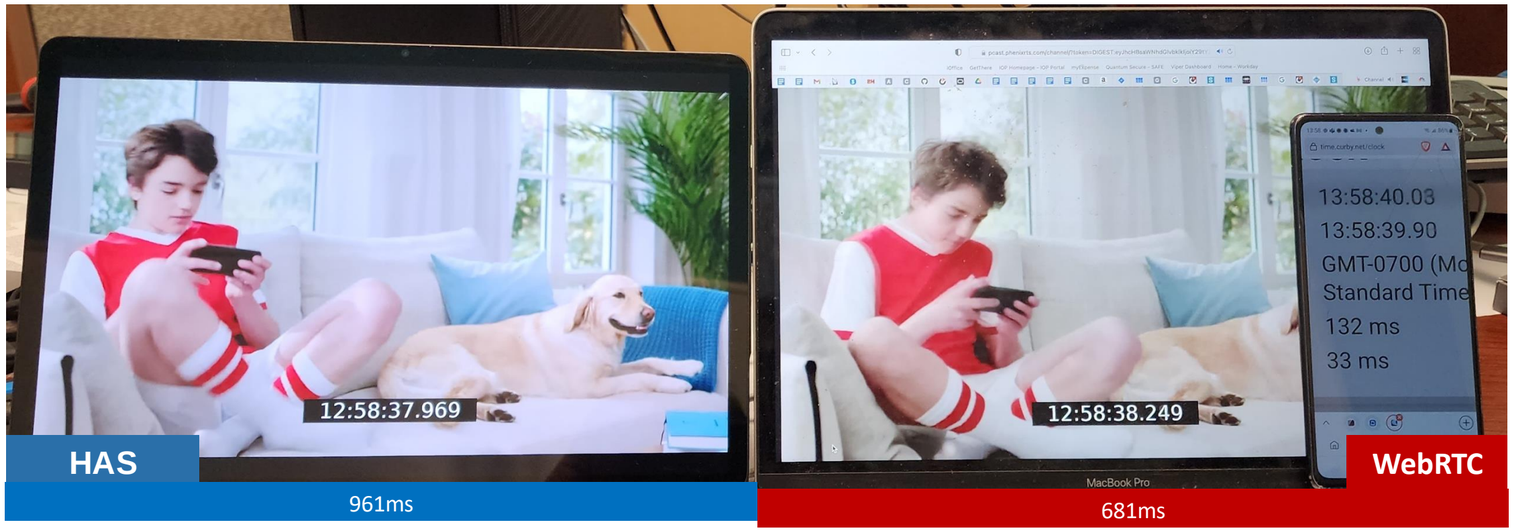
WHY HTTP STREAMING?

Neither HTTP nor MP4 were ever meant for real-time video delivery

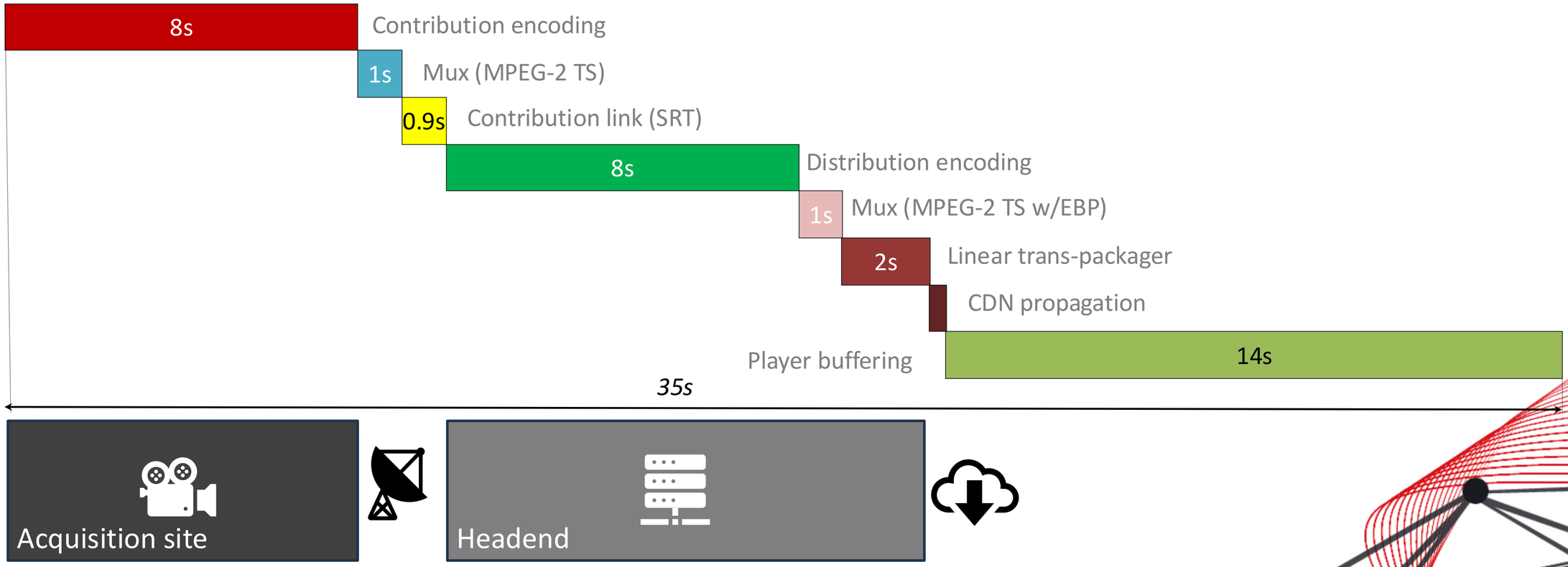
WebRTC reaches ~500ms latencies

- Harder to scale, requires new infrastructure

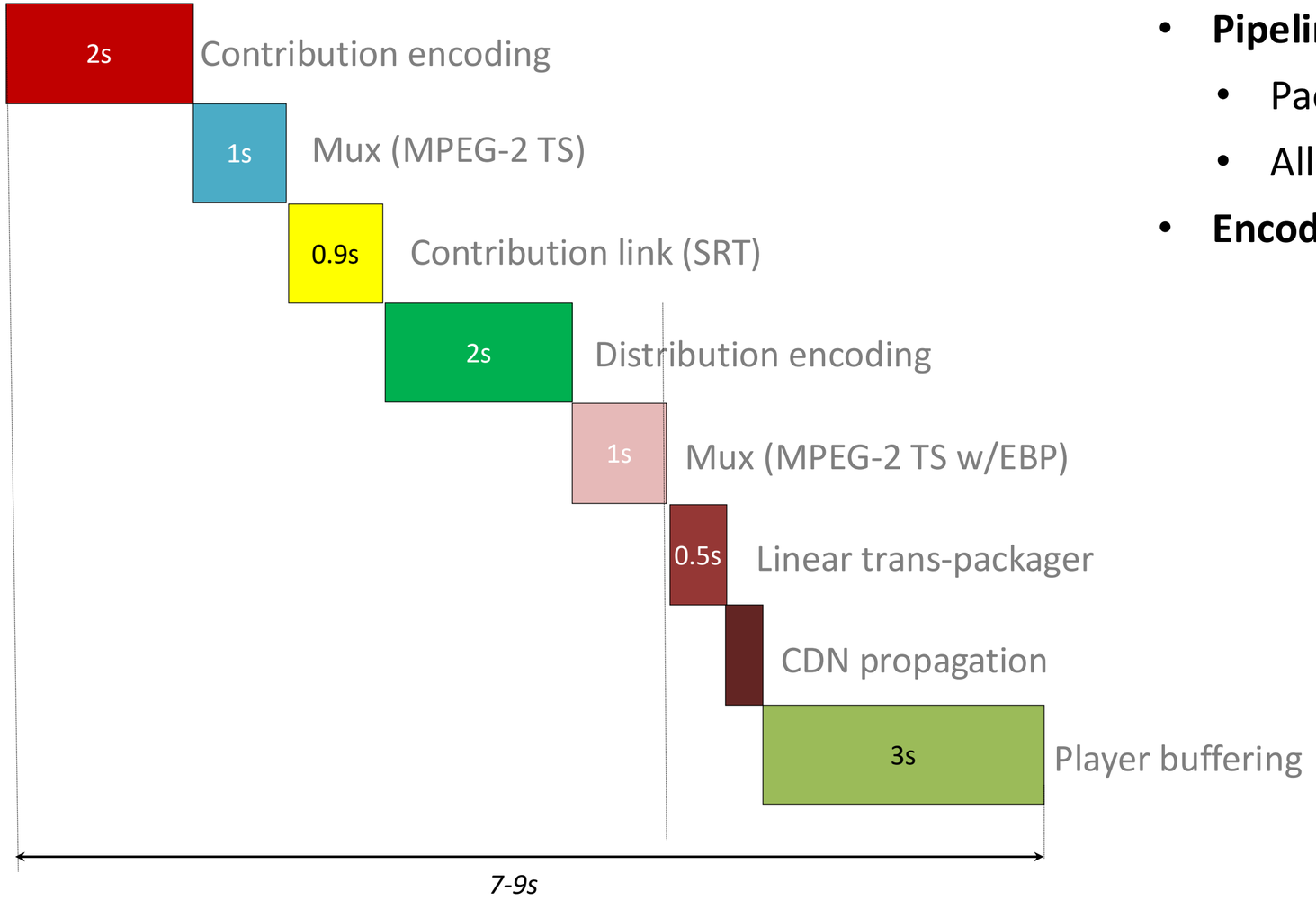
Does a difference between 700ms and 1s justify a rollout of new infrastructure and tech stack?



WHERE DO HTTP STREAMING LATENCIES COME FROM?



FIRST STEP: LL-DASH / LL-HLS



- **Pipelining delivery**
  - Packaging and transmissions in e.g. 0.5s chunks
  - Allows much shorter player buffers
- **Encoder** delays algorithmically reduced to 1.5-3s

CAN WE DO BETTER?

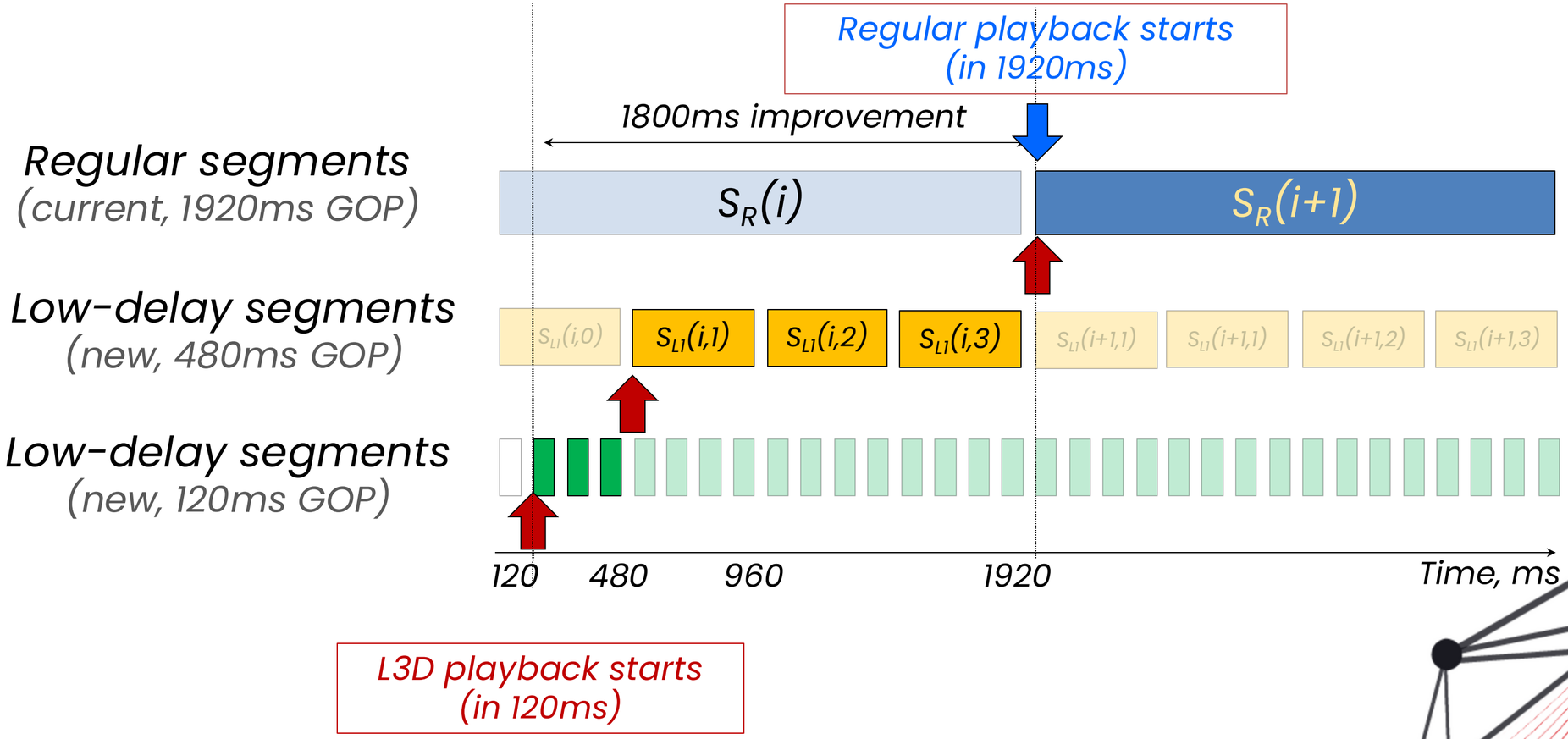
GOAL
Go beyond what has been done in typical LL-DASH/LL-HLS and “trim some fat”

WHERE DO WE TRIM?

- **Eliminate mezzanine feeds**
  - Trade non-essential encoder delay for extra operational complexity
  - Remove MPEG-2 TS as a transmission intermediary
- **Reduce encoder delay**
  - Most delay is needed for compression efficiency, but we can sacrifice a little bit
- **Reduce partial segment (chunk) size**
  - Lower packaging delay, *may* translate into lower buffering delay
- **Reduce video start time**
  - Faster than GOP boundary start-up
- **Give the player more time**
  - Some of the time saved should be given to player buffer to improve stability

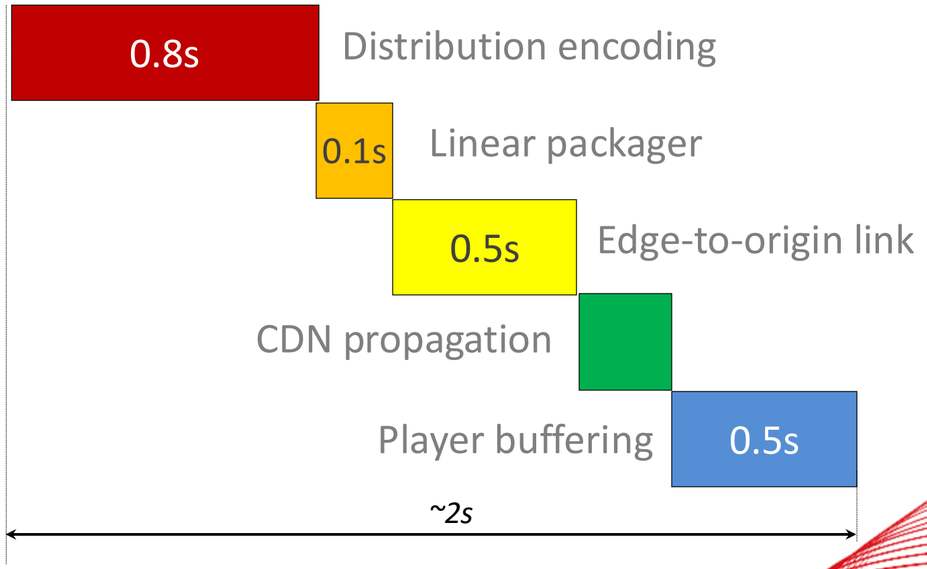
L3D-DASH: NEAR-INSTANT VIDEO PLAYBACK

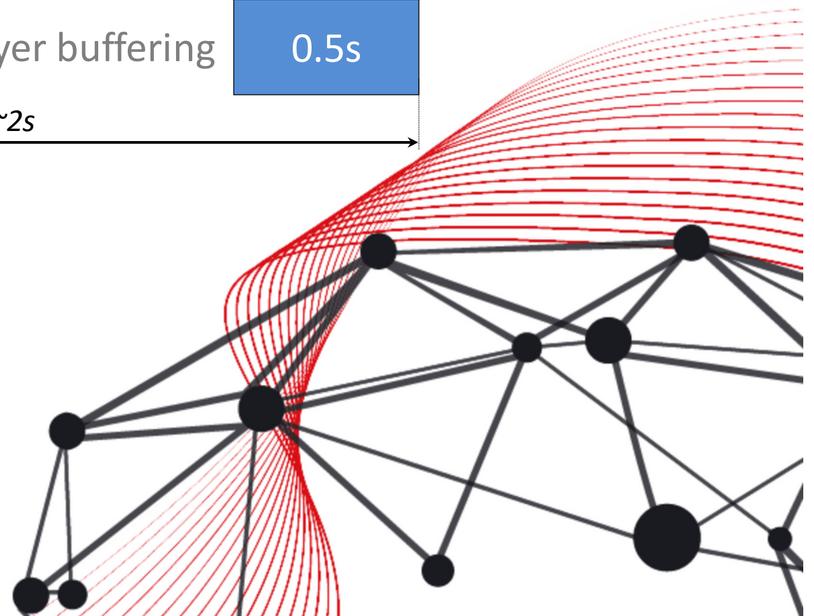
- **Low-latency low-delay** extensions of MPEG DASH (upcoming 6th edition)
- **Fine-grain random access** (possibly frame-accurate)



EDGE ENCODING (“SINGLE-GEN”)

- **Edge encoding** (“single-gen”) positions **distribution encoders next to playout**
  - Reduction in latency
  - Improvement in quality
    - multiple rounds of transcoding reduce quality
  - Less bandwidth on the edge-to-origin hop
- **DASH-IF Live Ingest**
  - Transfers CMAF chunks using HTTP POST
  - Eliminates muxing delay
    - TR 101 290 –style rate constancy not needed
    - Flow control handled by TCP or QUIC





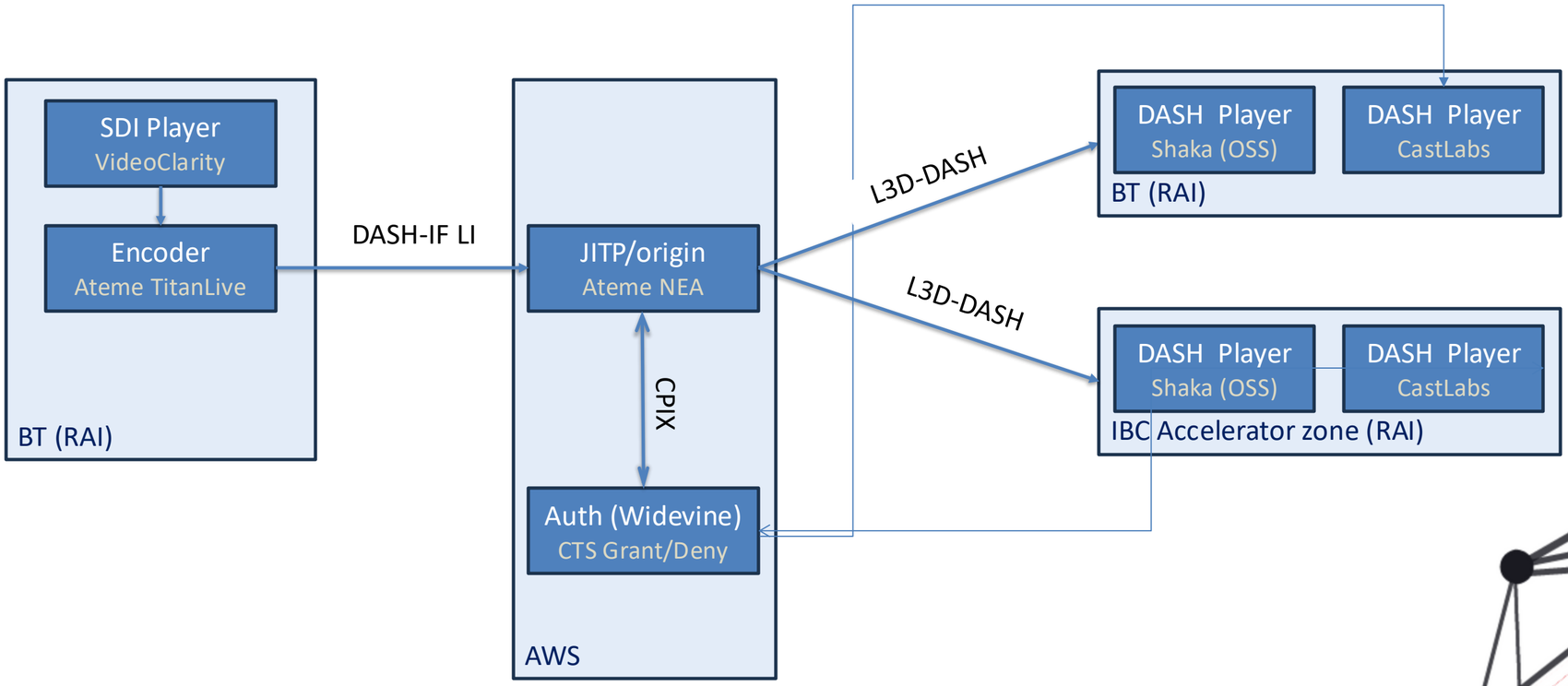
DEMO

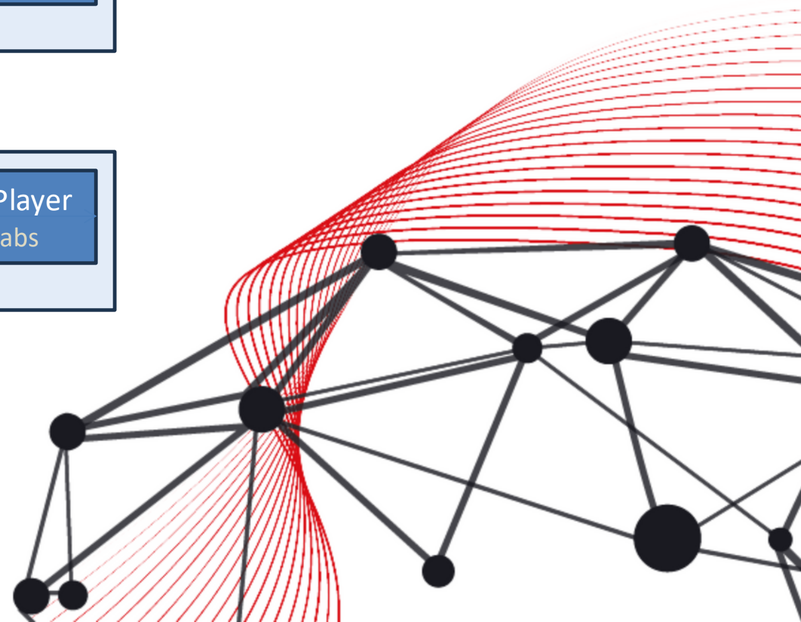
HIGHLIGHTS

- Single-gen encoding architecture
- DASH-IF Live Ingest
- L3D-DASH

PERFORMANCE

- 1.8s contribution-to-glass latency (across US)
- ~1s average improvement in tune-in and zapping times





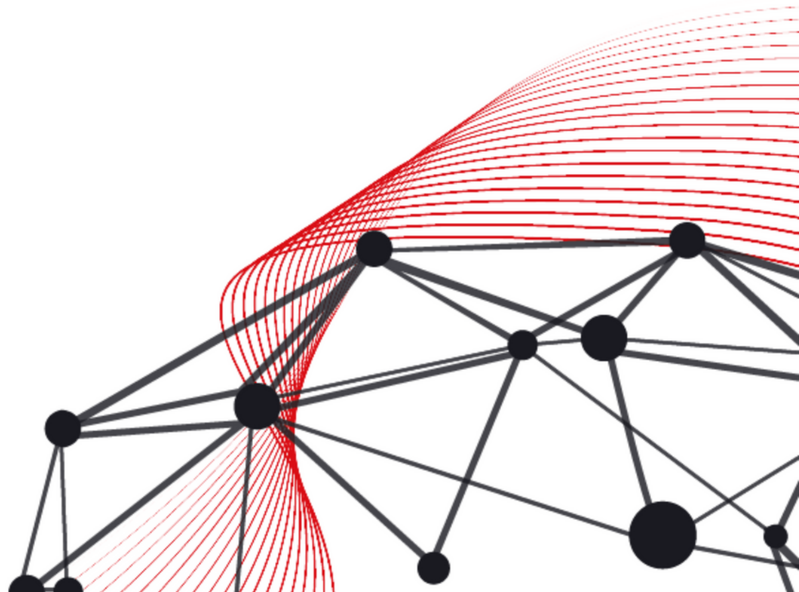
NEXT: OPTIMIZING TRANSPORT

FROM TCP TO QUIC

- **Parallelized multiplexed connections**
  - Beneficial for both DASH-IF Live Ingest and L3D-DASH
  - >7 parallel streams up, 3-4 streams down, one connection
- **Priorities** allow getting manifests and audio faster
- **HTTP/3 server implementations immature**
  - Initial results show 50-100ms improvement over HTTP

HTTP/1.1 IN FEC TUNNEL

- Major reduction in rebuffering at low latency



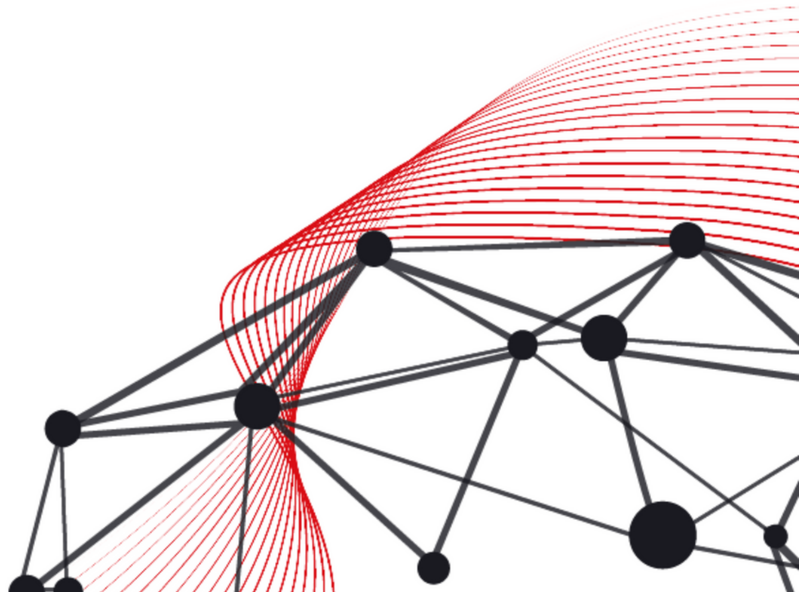
NEXT: SPECIAL-PURPOSE ORIGIN

BLOCKING REQUESTS

- Origin blocks when segment unavailable but expected
  - Current behaviour -- returning an error
  - Functionality used in LL-HLS for playlist version updates
- Segments can be requested before they are available
- Lower latency

POST-TO-GET PIPELINE

- DASH-IF Live Ingest IF-2
  - POST of MPDs and segments
- Forward blocks of POST’ed resource to blocked GET
  - Latency reduction
  - Bespoke functionality



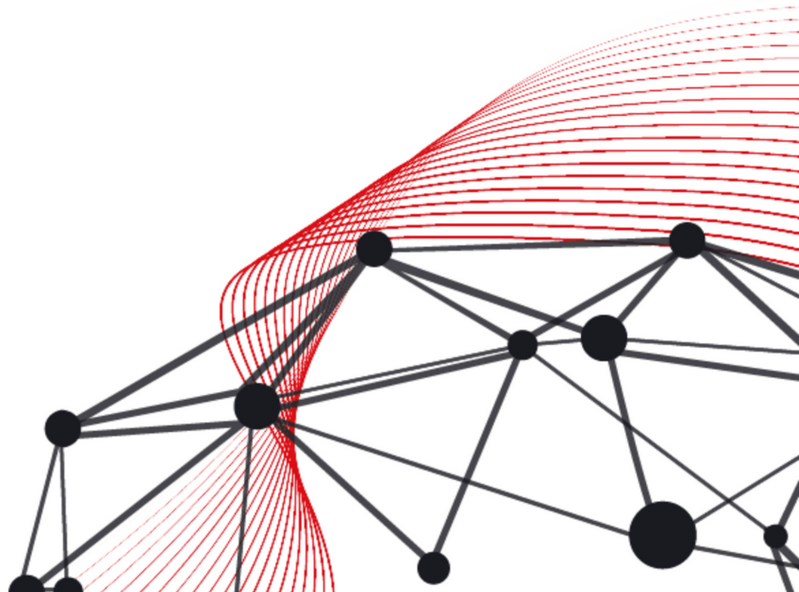
TO BE CONTINUED…

ENCOURAGING RESULTS

- Low latency
- Fast start-up

FURTHER WORK: SEE US AT MILE-HIGH VIDEO 2025

- February 18-20, 2025
- Denver CO, USA



IBC2024

#ACCELERATORS2024

SCALABLE ULTRA-LOW LATENCY STREAMING FOR PREMIUM SPORTS

14 SEPTEMBER 2024

ALEX GILADI

COMCAST FELLOW.