BITMOVIN
Video Infrastructure for the Web

TODAY'S AND FUTURE CHALLENGES WITH NEW FORMS OF CONTENT LIKE 360°, AR, AND VR

Stefan Lederer
Bitmovin, Inc.

+1 650 4438956
stefan.lederer@bitmovin.com
@bitmovin

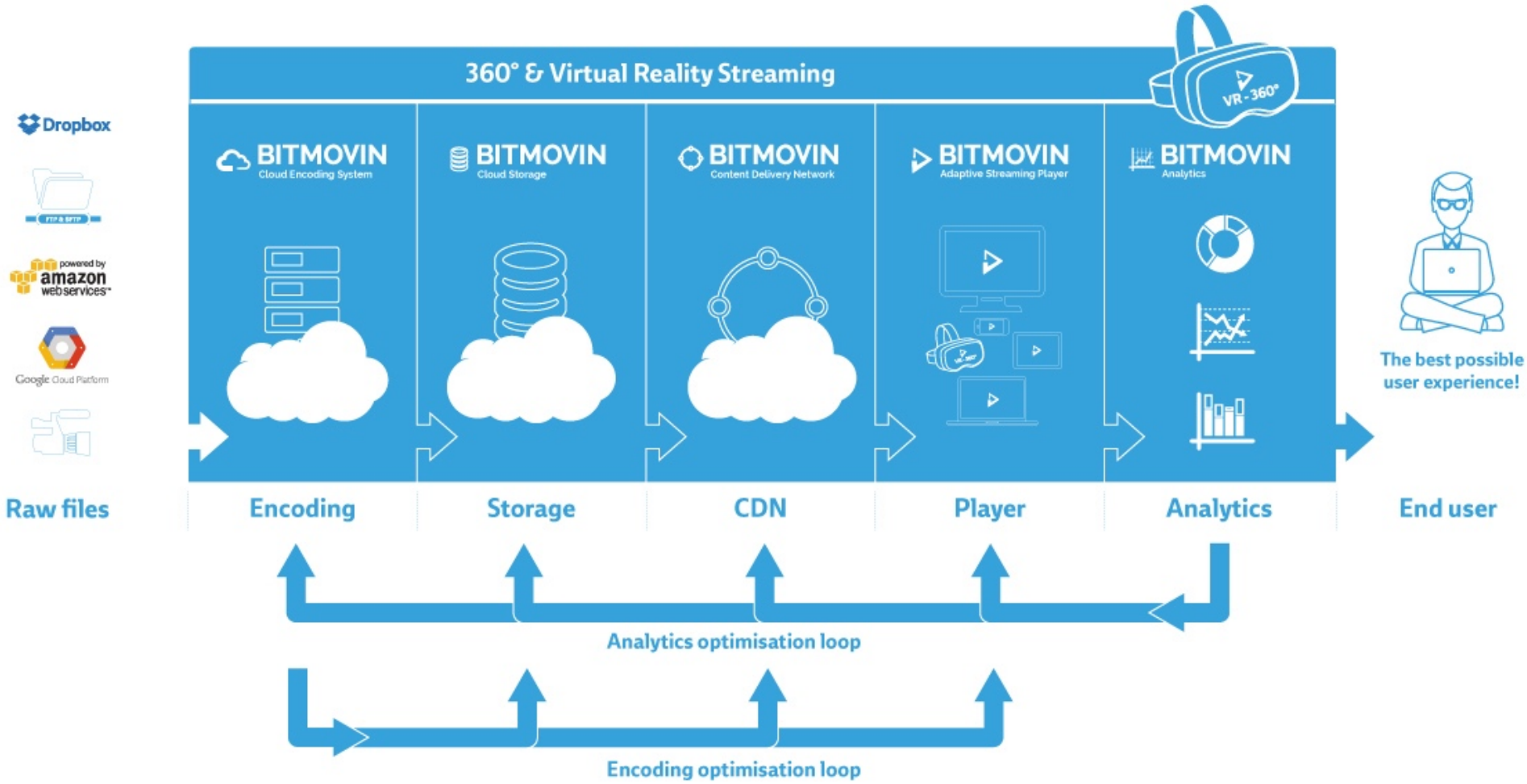
www.bitmovin.com

301 Howard Street, Suite 1800 | San Francisco | CA 94105 | USA
Schleppeplatz 7 | 9020 Klagenfurt | Austria | Europe

MPEG WORKSHOP: GLOBAL MEDIA TECHNOLOGY STANDARDS FOR AN IMMERSIVE AGE

FULL-STACK VIDEO INFRASTRUCTURE

BITMOVIN
Video Infrastructure for the Web

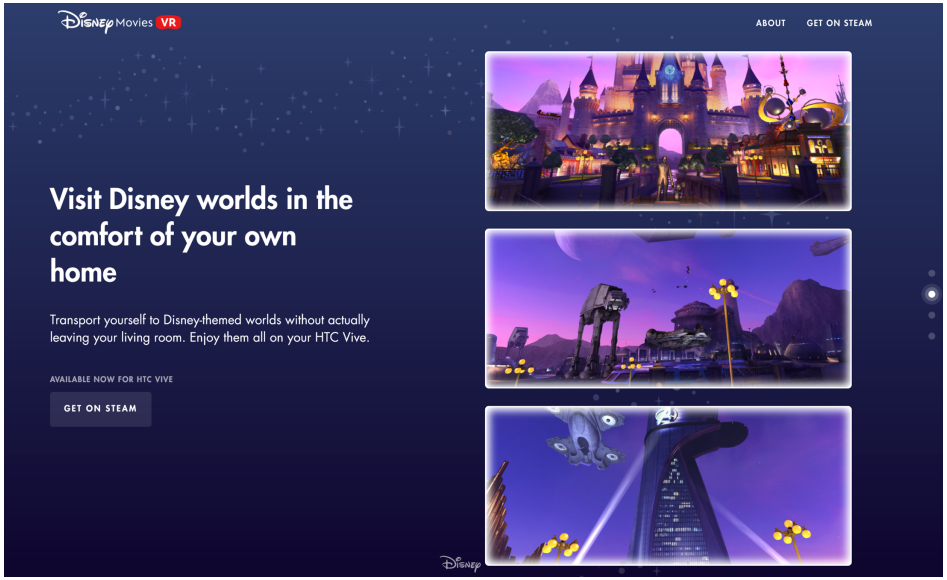


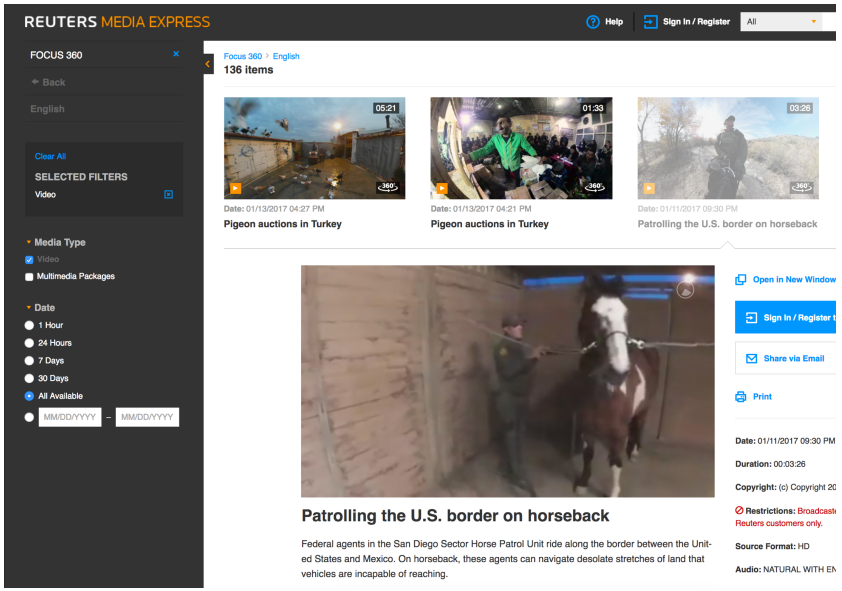
END-TO-END 360° INFRASTRUCTURE

BITMOVIN
Video Infrastructure for the Web

CUSTOMER USE CASES

BITMOVIN
Video Infrastructure for the Web





VIDEO CODING FOR 360° VIDEO?

BITMOVIN
Video Infrastructure for the Web

No special techniques for coding in spherical domain are widely available yet

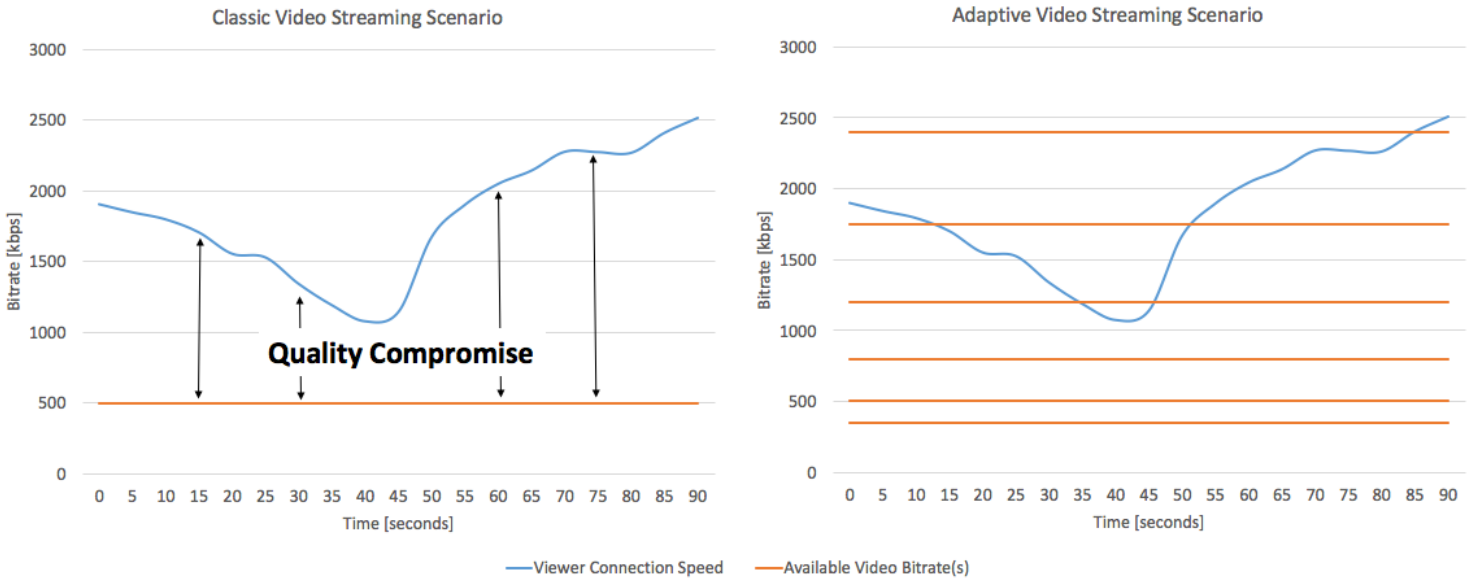
- Encoding in rectangular domain
- Therefore we need the described projections to rectangular layouts
  - Equirectangular projection
  - Cube projection
  - Pyramid, Equal-Area projection and more . . .
- Traditional video codecs are used:
  AVC / HEVC (/ VP8 / VP9)

360° VIDEO – WHAT'S OUT THERE?

BITMOVIN
Video Infrastructure for the Web

Primarily: Progressive MP4 for 360°

- 1080p to 4K Videos, using H.264 (some VP9)
- Why?
  - Browser/OS restrictions, e.g., on iPhone
  - Application has no access to frames
- Without adaptive streaming, the result is buffering and poor QoE



360° VIDEO – WHAT'S OUT THERE?

BITMOVIN
Video Infrastructure for the Web

Current trends? MPEG-DASH / HLS

- Logical next step to use adaptive streaming
- Especially for high bitrate/resolution content

Challenges:

- Device coverage & issues
  - Desktop browsers, mobile Web, smartphone apps, VR headsets, TVs, casting devices, etc.
  - 360° rendering and access to frames is different on all platforms
  - Lack of frame access, DASH/HLS support, etc.
- Overlays and ads
  - How to position/communicate/integrate different types of ads
- DRM protection
  - No access to decoded frame
  - See also next slides

PROBLEMS: DRM

BITMOVIN
Video Infrastructure for the Web

Content owners have to protect their content

E.g. in Browsers: HTML5 Encrypted Media Extensions

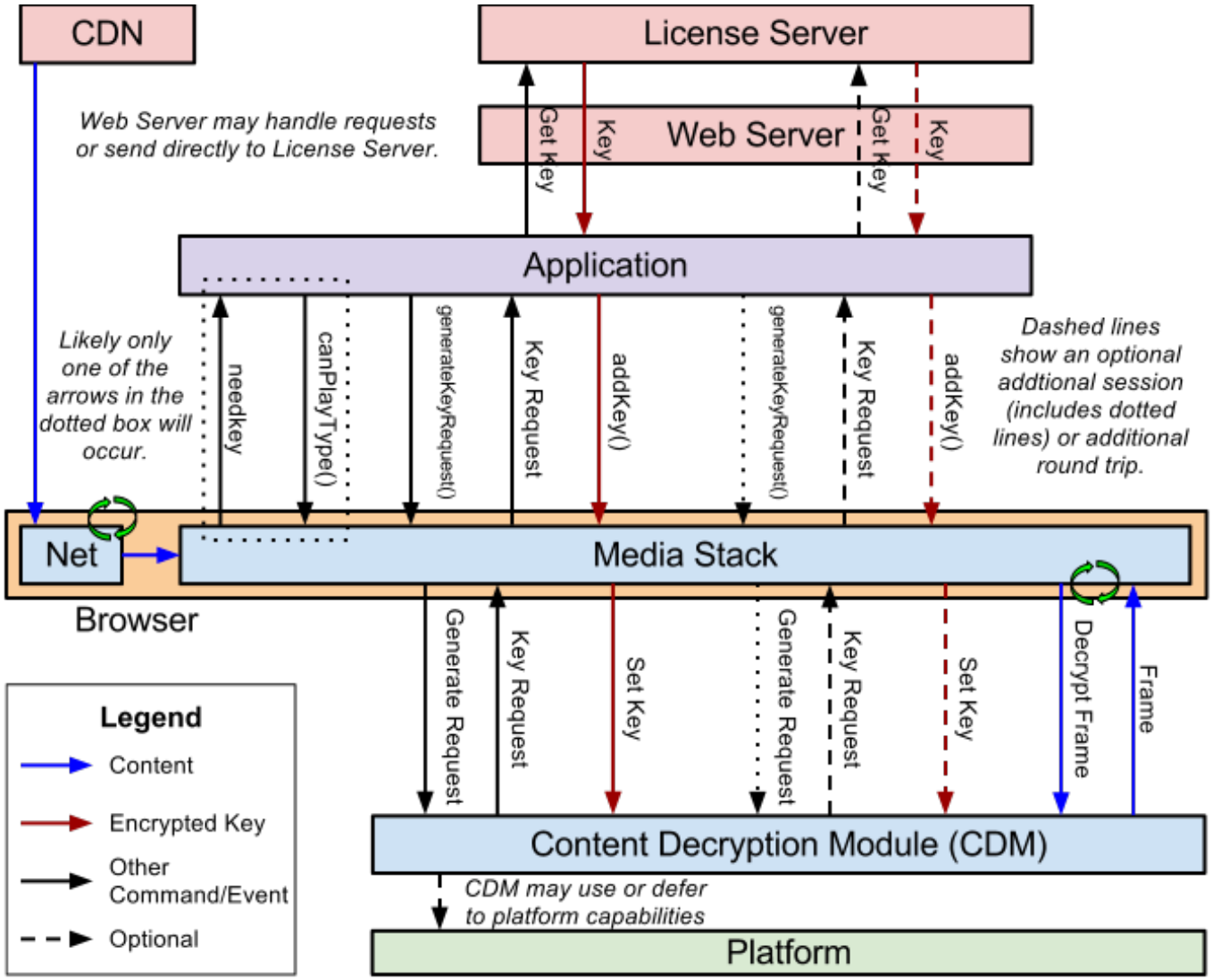
› DRM protection for adaptive streaming in HTML5
› Support for DASH / CENC content in HTML5

Where is the problem?

› 360°/VR content needs to be rendered in JavaScript
› Equirectangular to viewpoint rendering
› However: touching the frame for rendering breaks the DRM security principles (one could store the unprotected frames)

PROBLEMS: HTML5 EME + 360°

BITMOVIN
Video Infrastructure for the Web



PROBLEMS: DRM

BITMOVIN
Video Infrastructure for the Web

This problem needs to be solved to enable business cases

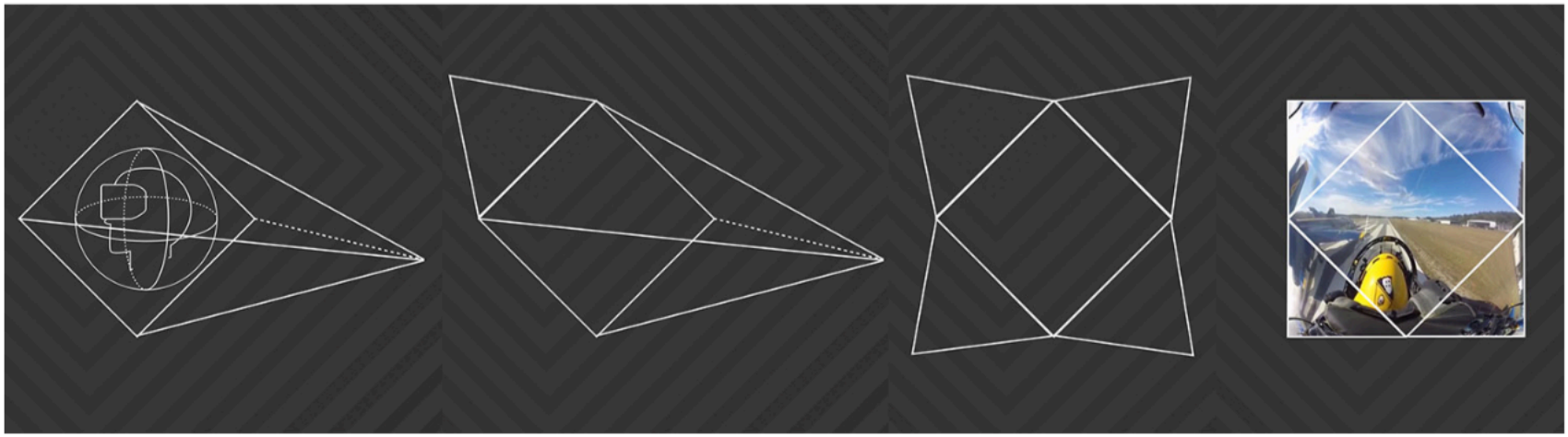
By whom?

» Standardization bodies (W3C, MPEG, etc.)
» Platform providers (Browsers, mobile OS vendors)
» DRM systems (Fairplay, Playready, Widevine, etc.)

NEW APPROACHES: ADAPTIVE PYRAMID MAPS

BITMOVIN
Video Infrastructure for the Web

- Base of pyramid represents
- Field of View: **high quality**
  Sides of pyramid contain rest of whole panorama: **low quality**
- File size **reduced by 80 percent** against the original



NEW APPROACHES: ADAPTIVE PYRAMID MAPS

BITMOVIN
Video Infrastructure for the Web

- Approach was developed by Facebook
- No public available scientific demos, evaluations or sources so far
- When the viewer shifts perspective, a new pyramid is streamed and rendered
- In total **30 viewports** covering the sphere, separated by about 30 degrees
- **5 different resolutions** for each stream
- Results in **150 different versions** of same video, thus **significantly increase storage and encoding** requirements

WHAT ARE THE REAL NEEDS FOR 360°?

BITMOVIN
Video Infrastructure for the Web

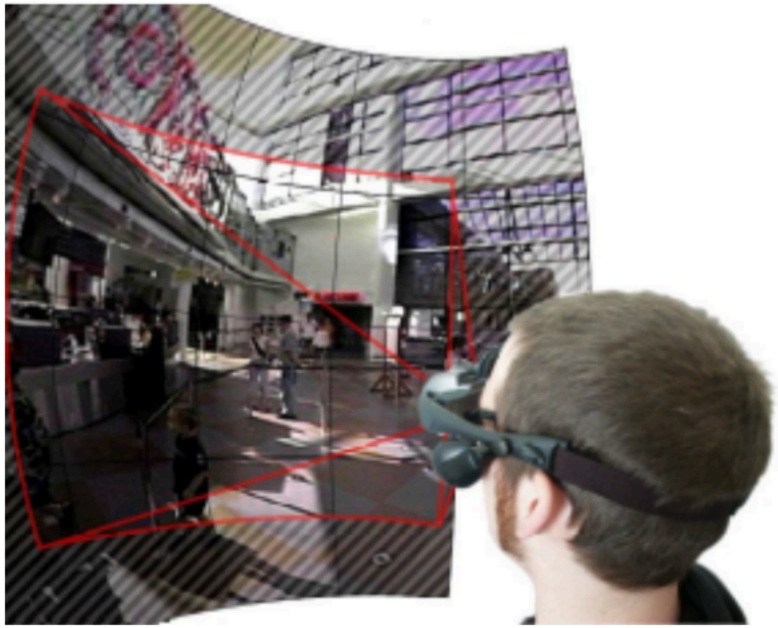
- 360° content allows the user to change the viewing orientation while watching a video
  - Large areas of the video are not seen
- Full 360° area has to be stored
  - Large file sizes
  - Challenge of streaming of such content
- → Increase streaming efficiency
  - Choose the “right” content representation
  - Utilize new delivery approaches

WE NEED MORE BANDWIDTH EFFICIENT APPROACHES

BITMOVIN
Video Infrastructure for the Web

The state of the art streaming technology is to encode the video with each frame containing the full panorama as a whole in one uniform quality

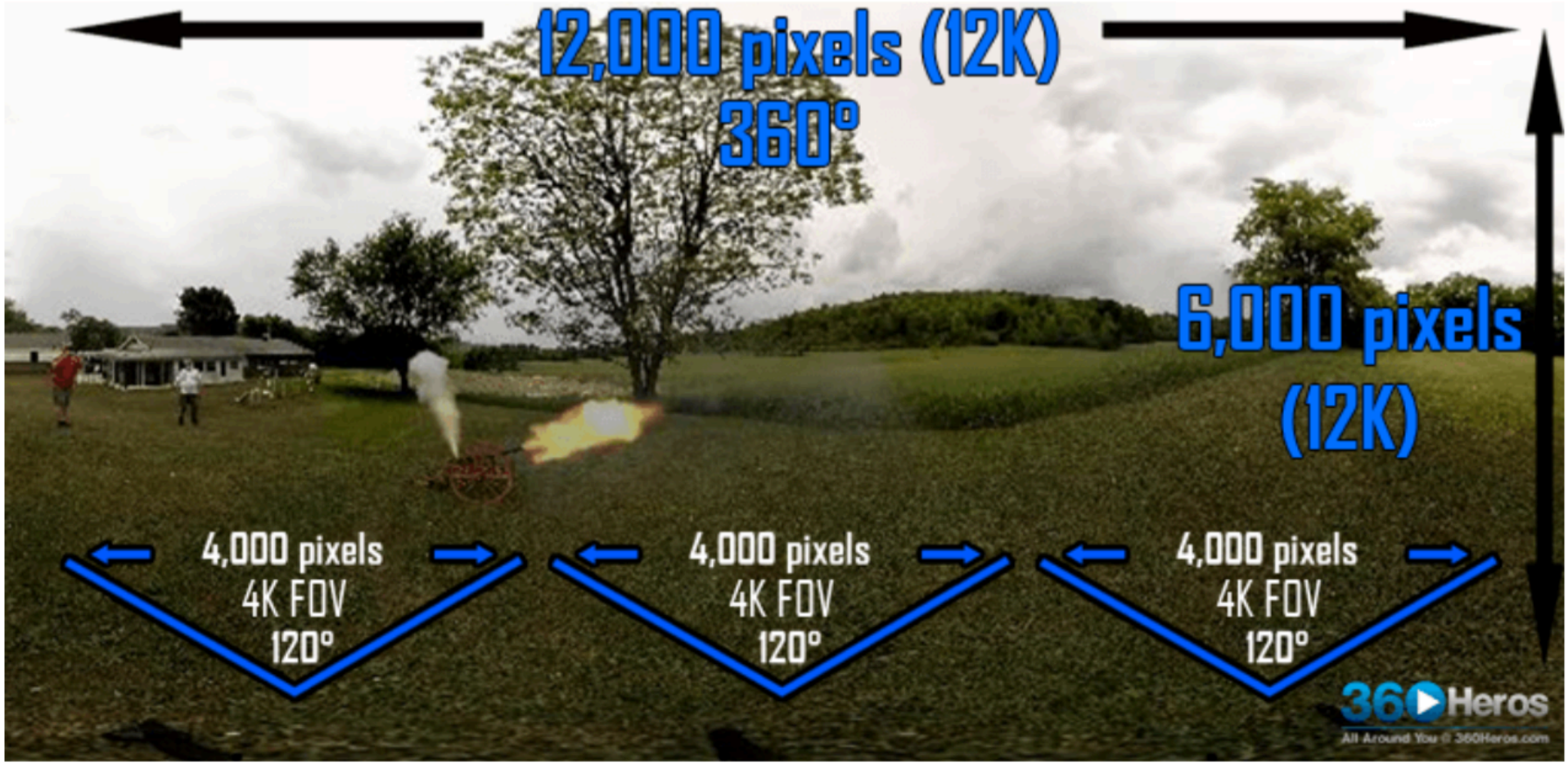
- This is inefficient because
- only a part of each video frame is presented to the user based on the current viewing angle
- But all parts are streamed in same high quality



WHY WE NEED MORE BANDWIDTH EFFICIENT APPROACHES

BITMOVIN
Video Infrastructure for the Web

Needed resolution of panoramic video for achieving 4k field of view resolution:



MOTIVATION

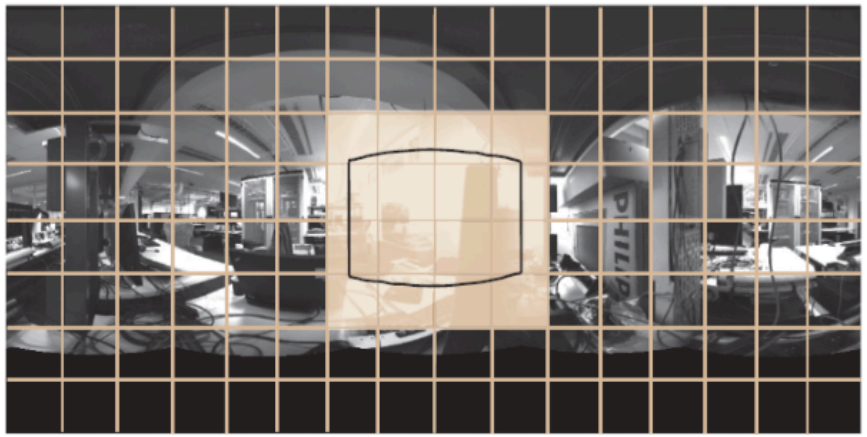
BITMOVIN
Video Infrastructure for the Web

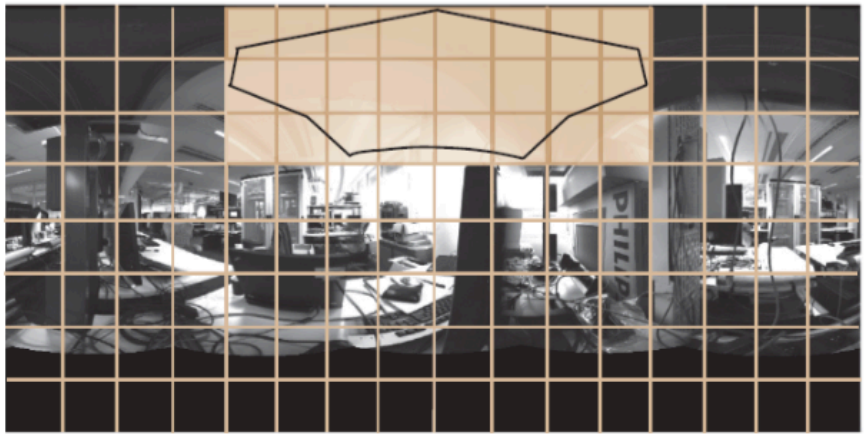
- Field of view << full panorama
  - E.g., 90° → only a fourth of the full panorama displayed
- Render FoV in 2K → full panorama should be 8K
  - Won't work well with traditional streaming approaches
- Traditional codecs (AVC/HEVC) need a rectangular two-dimensional video representation to work well
  - Equirectangular, cubic, pyramid, etc. projections
  - Traditional packaging and storing technologies can be used
  - Streaming technologies like DASH and HLS can be used
- Special video player is needed
  - Must know which projection is used
  - Must provide interaction possibilities

NEW APPROACHES: TILED BASED STREAMING

BITMOVIN
Video Infrastructure for the Web

- Video is sliced into tiles
- Tiles are encoded in several qualities
- Only needed tiles are streamed in high quality
- Tiles outside FoV are streamed in descending quality, lower quality or not at all





TILED STREAMING

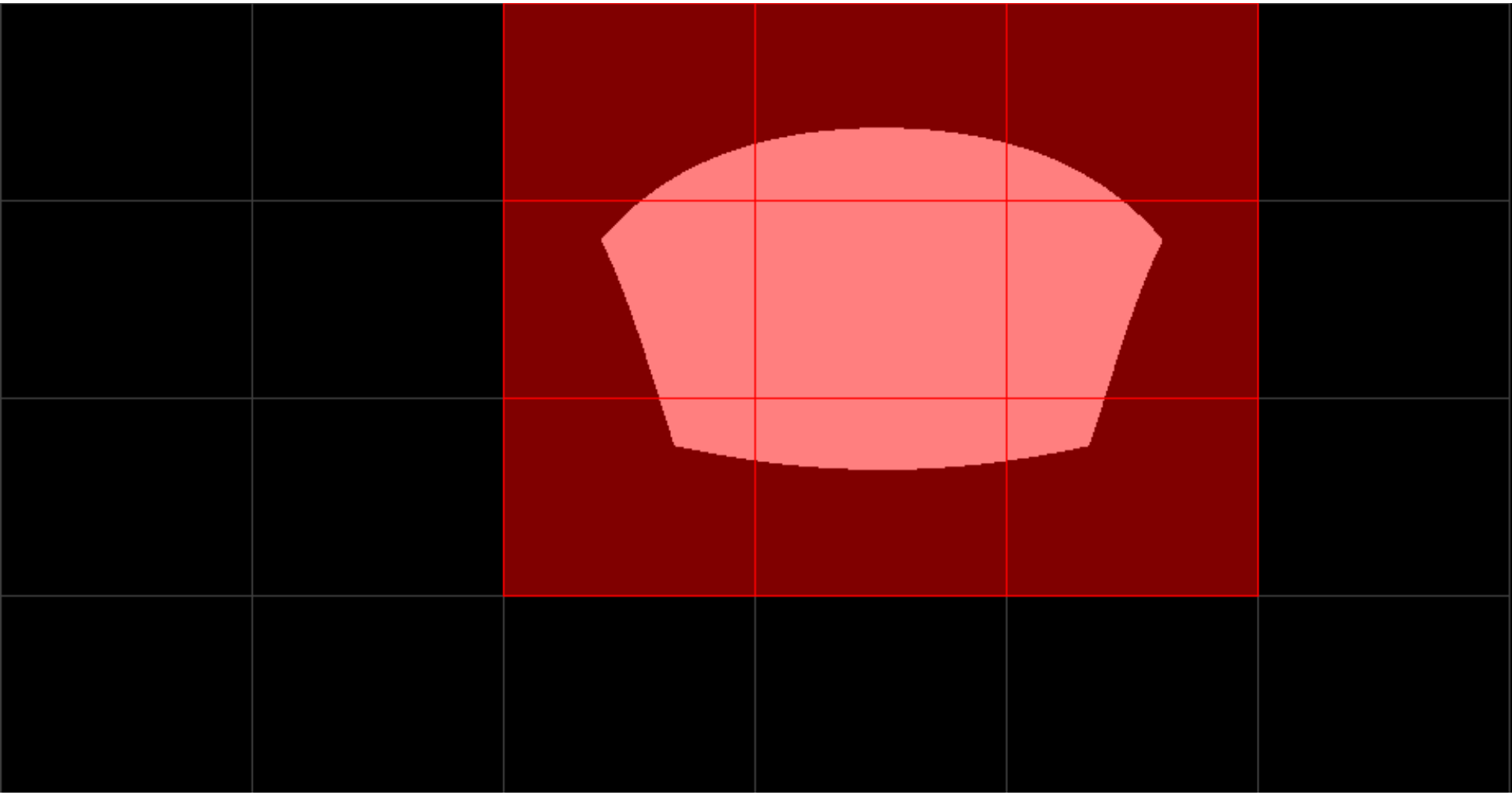
BITMOVIN
Video Infrastructure for the Web

- Each tile can be encoded separately and independently streamed, decoded, and rendered
  - Each tile can have different bitrates and resolutions
  - Adapt quality of each tile independently
- H.265/HEVC provides tile support
  - Unfortunately not H.264
  - Need for a single decoder (enabled by HEVC), as usage of multiple decoders is not (yet) possible on consumer devices
- Can be used to realize a compressed domain mixing of tiles
  - Implementation as network services
  - Client side implementation

NEW APPROACHES: TILED BASED STREAMING

BITMOVIN
Video Infrastructure for the Web

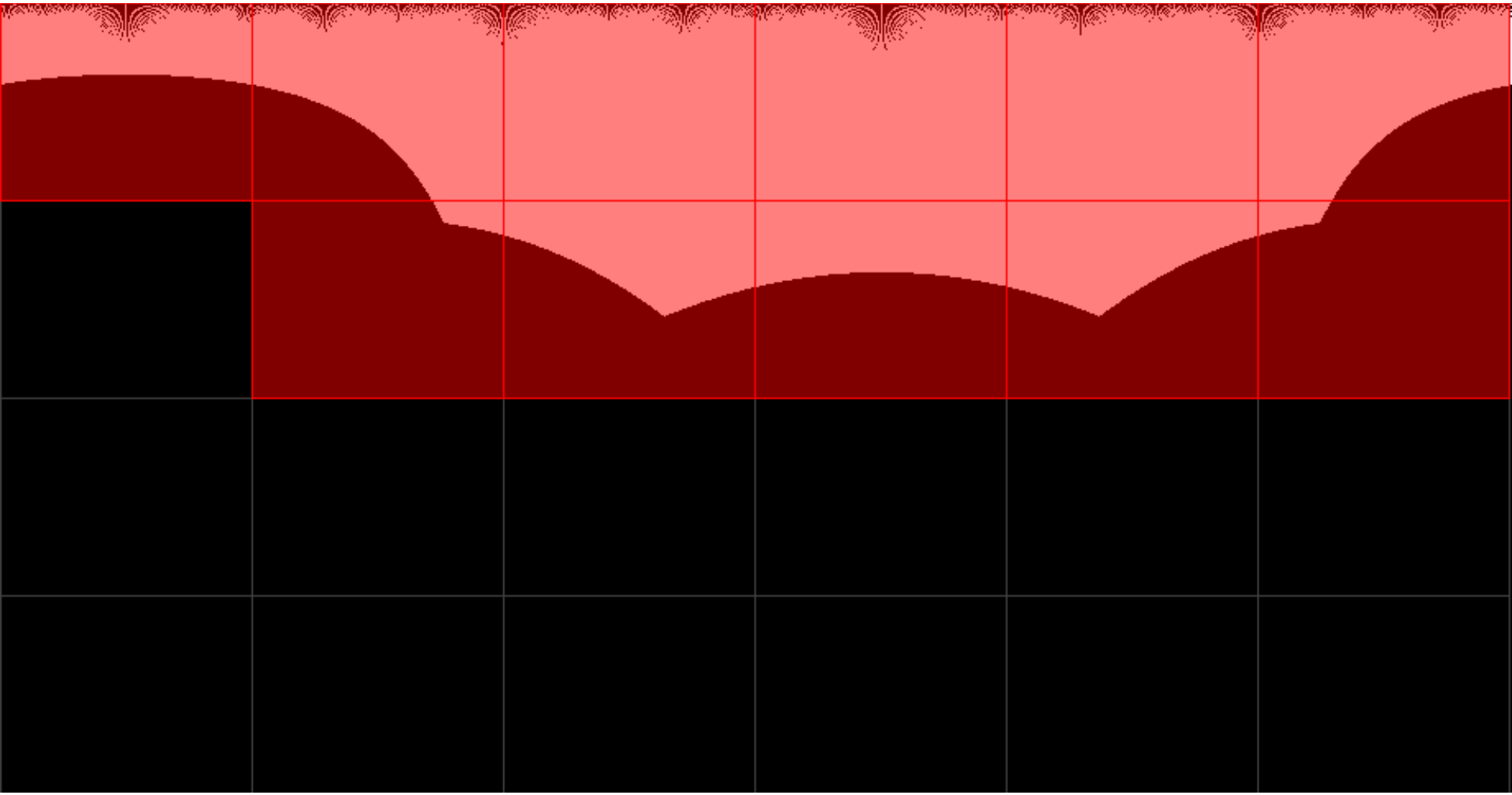
Front view, 6x4 tiles:



NEW APPROACHES: TILED BASED STREAMING

BITMOVIN
Video Infrastructure for the Web

Top view, 6x4 tiles:



NEW APPROACHES: TILED BASED STREAMING

BITMOVIN
Video Infrastructure for the Web

Front view, 3x2 tiles:

COMBINING STANDARDS: DASH + HTML5

BITMOVIN
Video Infrastructure for the Web

- Client-based approach, using a single decoder
  - Written in JavaScript
  - Using Media Source Extensions (MSE)
- Different need to be mixed into a single video stream
  - Only one decoder is required
  - No need for synchronization
- Three different quality levels for each tile
- 6 x 4 tiling scheme[1]

1- Where a higher number of tiles in general might lead to an increased user experience, but decreases the efficiency of the codec

DASH + HTML5 EXAMPLE

BITMOVIN
Video Infrastructure for the Web

play pause

Tiling scheme selector

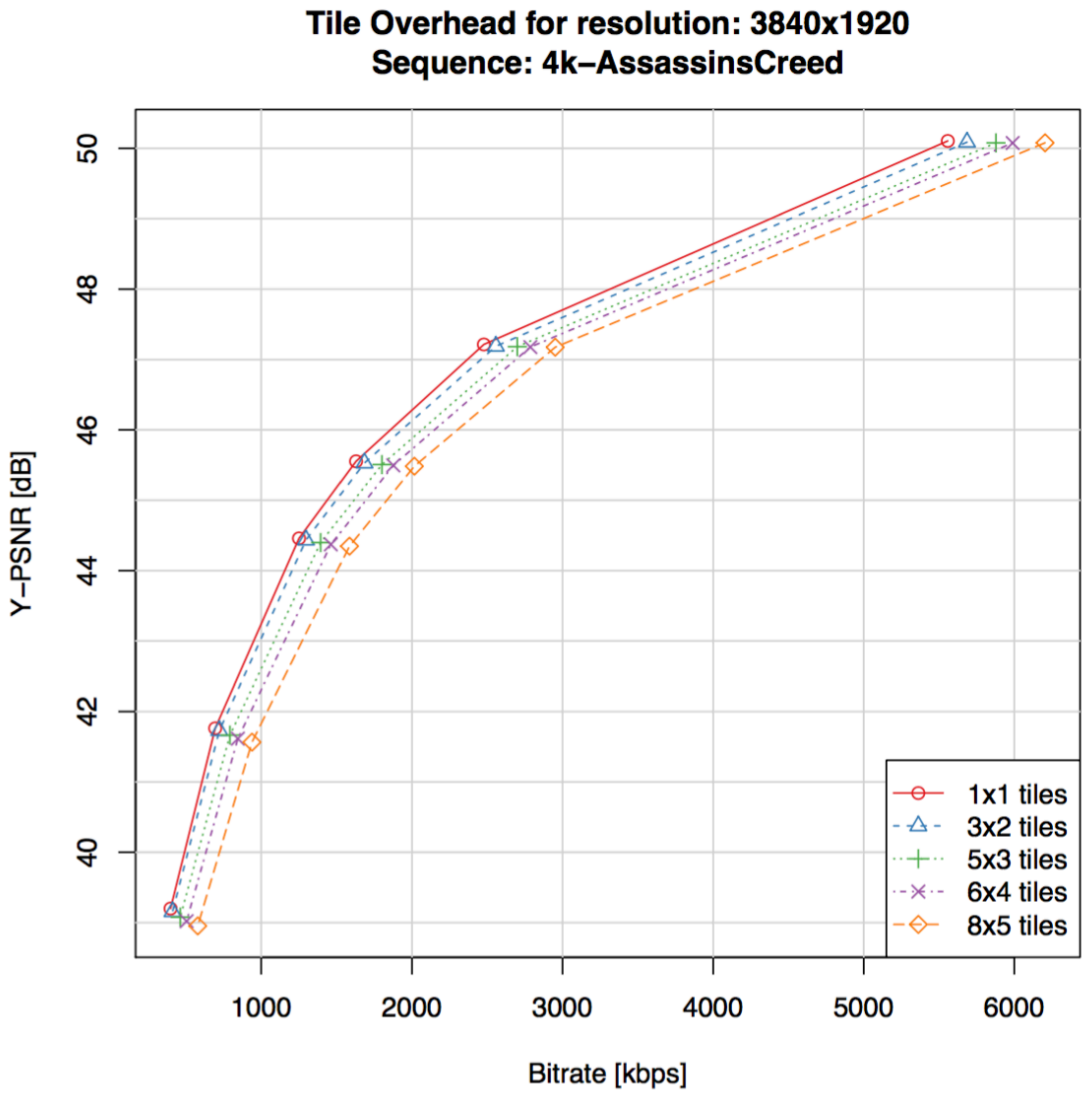
0	0	0	0	0	0
0	1	2	2	1	0
0	1	2	2	1	0
0	0	0	0	0	0

DASH + HTML5 EXAMPLE

BITMOVIN
Video Infrastructure for the Web

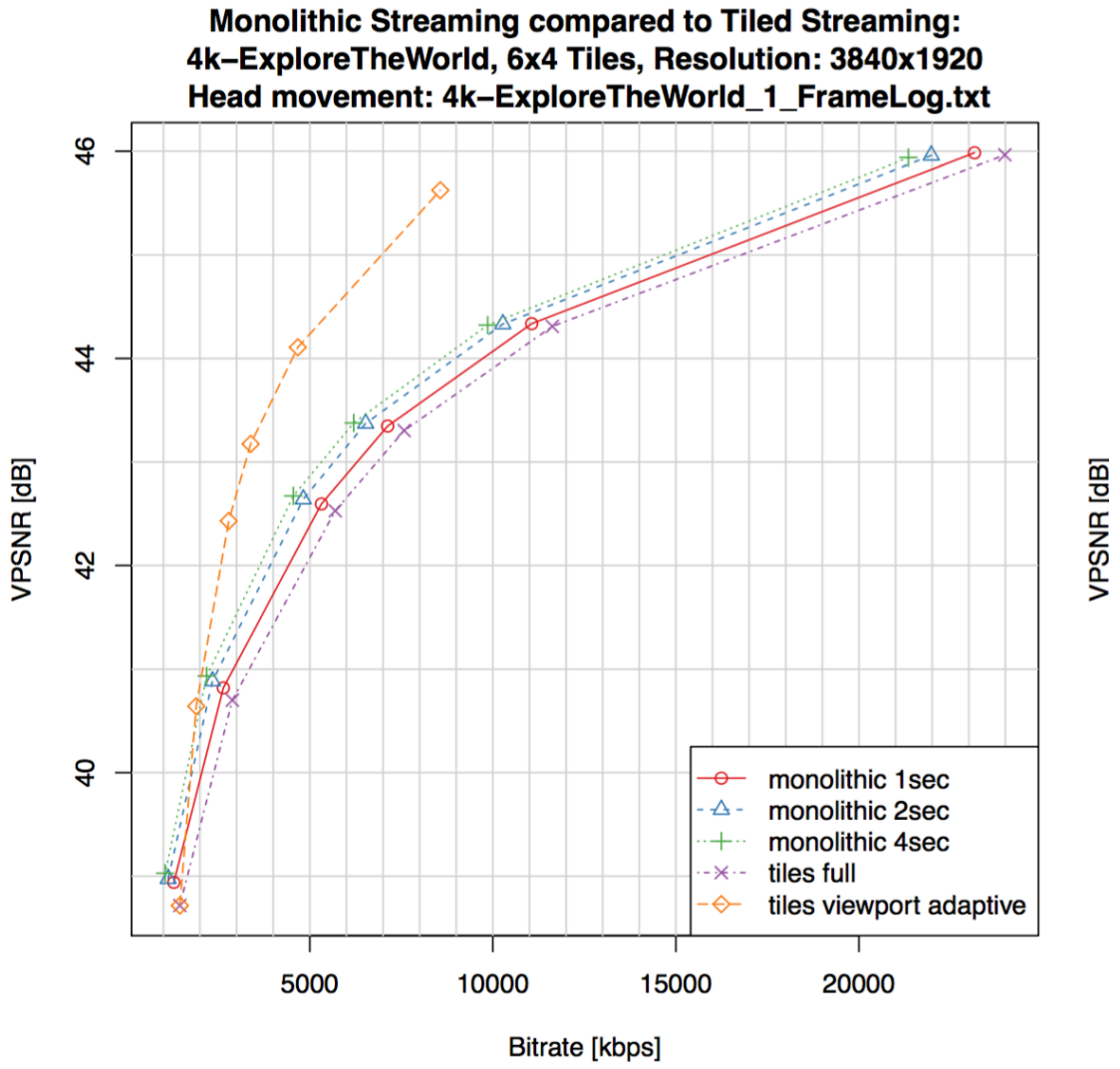
OVERHEAD USING TILES IN HEVC

BITMOVIN
Video Infrastructure for the Web



60%+ BANDWIDTH REDUCTION

BITMOVIN
Video Infrastructure for the Web



WHAT WE NEED IN THE FUTURE?

BITMOVIN
Video Infrastructure for the Web

- 6 degrees of freedom
  - Light field, point cloud, multi-view, etc. approaches
  - Codecs, streaming standards, platform support, tools, etc.

- Good demos so far:

http://8i.com/experience

BITMOVIN
Video Infrastructure for the Web

TODAY'S AND FUTURE CHALLENGES WITH NEW FORMS OF CONTENT LIKE 360°, AR, AND VR

Stefan Lederer
Bitmovin, Inc.

+1 650 4438956
stefan.lederer@bitmovin.com
@bitmovin

www.bitmovin.com

301 Howard Street, Suite 1800 | San Francisco | CA 94105 | USA
Schleppeplatz 7 | 9020 Klagenfurt | Austria | Europe

MPEG WORKSHOP: GLOBAL MEDIA TECHNOLOGY STANDARDS FOR AN IMMERSIVE AGE