Performance Measurements of 360◦ Video Streaming to Head-Mounted Displays Over Live 4G Cellular Networks

Wen-Chih Lo, **Ching-Ling Fan**,
Shou-Cheng Yen, and Cheng-Hsin Hsu

Department of Computer Science
National Tsing Hua University
Hsin-Chu, Taiwan

Introduction

- Using conventional displays to watch broadcast live events ⇒ passive experience

360 videos provide immersive experience

360° Videos

360° video in equirectangular projection

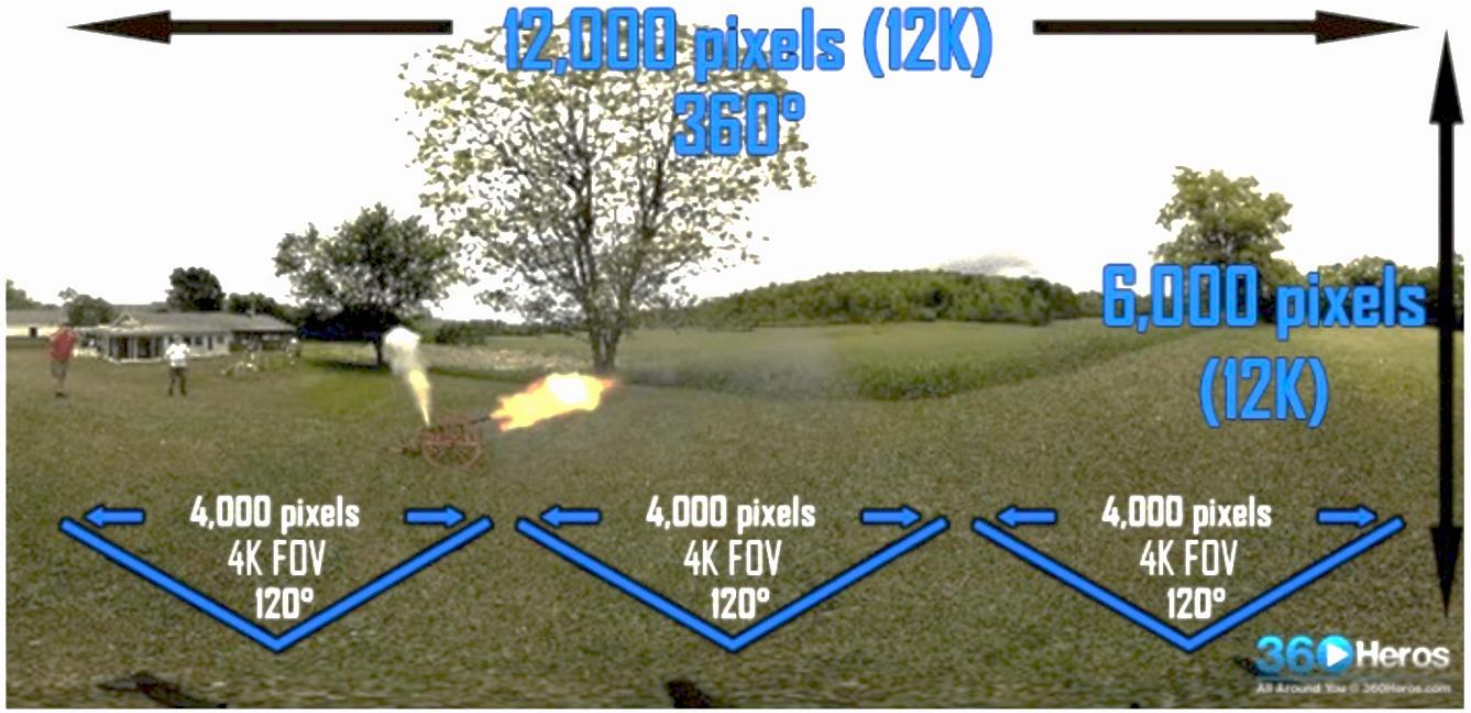
360° video in HMD view

Mobile VR

- Mobile HMDs
  - Samsung Gear VR, Carl Zeiss VR One Plus, Google Cardboard, even our smartphone
- Experience VR anywhere
  - YouTube, Facebook, Discovery, vTime

Streaming 360° Videos

- 4k resolution in HMD requires 12k resolution for the whole 360° videos (≈ 135 Mbps in HEVC )
  - extremely large file size ⇒ **insufficient bandwidth**

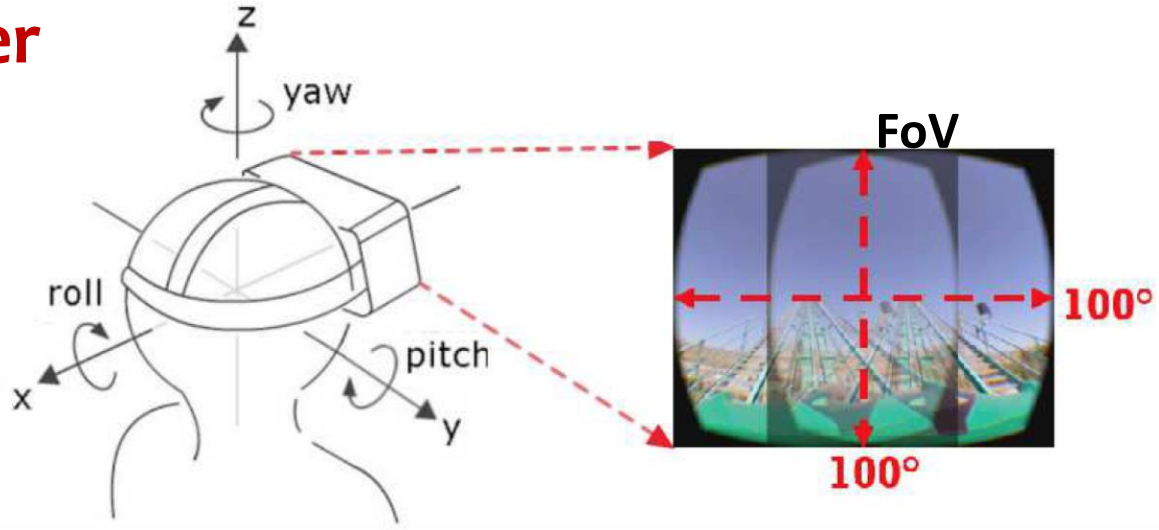


Images source: 360Heros

Opportunities and Solution Approaches

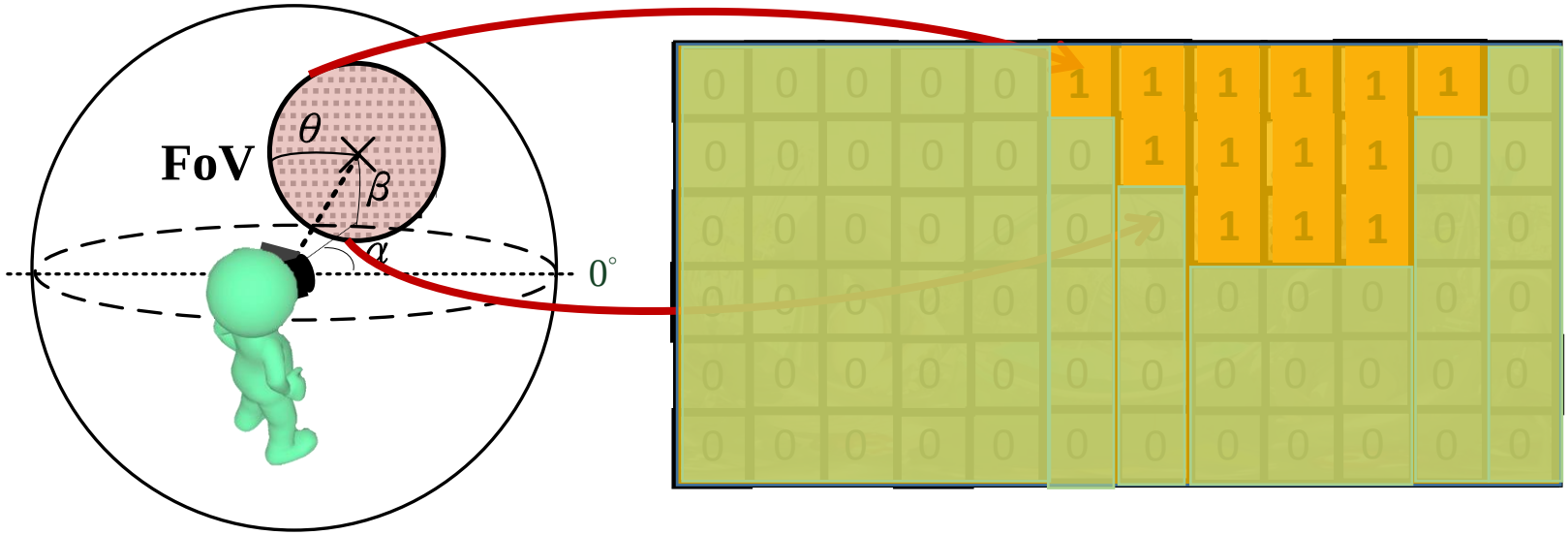
- The HMD viewer only gets to see a small part of the whole 360° video (< **1/3**)
- The viewer actively changes the viewing orientation when rotating his/her head.

⇒ **Only stream the current Field-of-View (FoV) of the viewer**



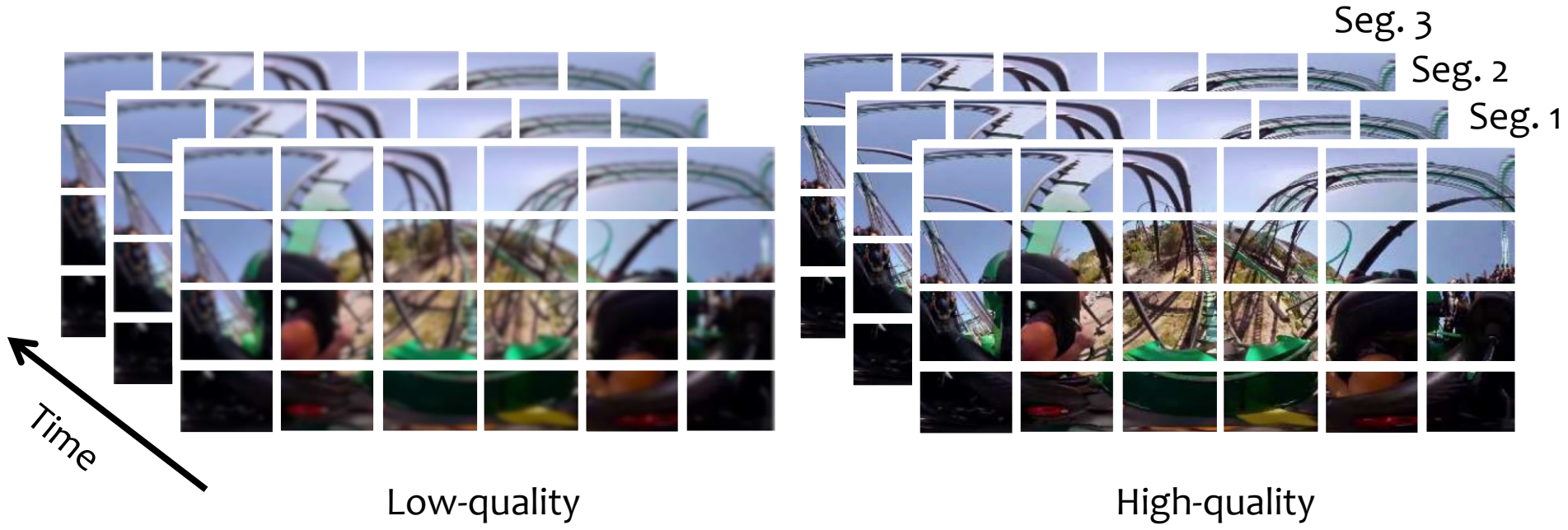
Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- 360° video is split into tiles of sub-videos (spatial) and independently encoded
  - Only the tiles overlapped with the viewer's FoV are streamed to the client



Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments



Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments

Selected tiles

0°, 0°

Viewer Orientation

Available Bandwidth

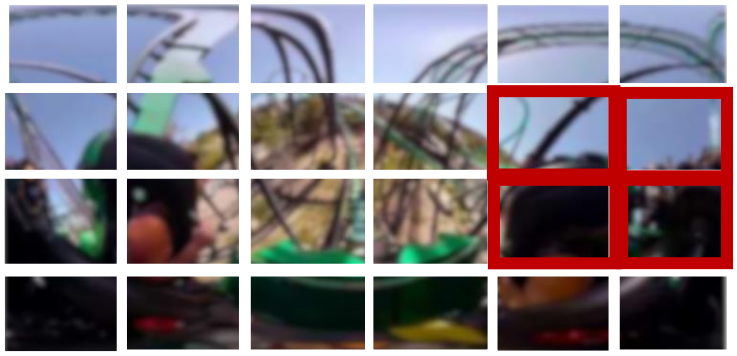
Time

Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments

Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments



Seg. 3

Time

Low-quality

High-quality

Selected tiles

89°, 0°

Viewer Orientation

Available Bandwidth

Time

Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments

Selected tiles

0°, 0°

Viewer Orientation

Available Bandwidth

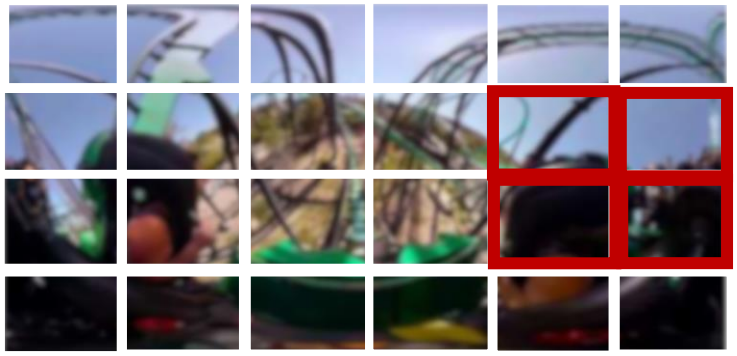
Time

Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments

Tiling with MPEG Dynamic Adaptive Streaming over HTTP (DASH)

- Tiles are split into temporal segments
  - tiles and qualities can change in every segments



Seg. 3

Time

Low-quality

High-quality

Selected tiles

89°, 0°

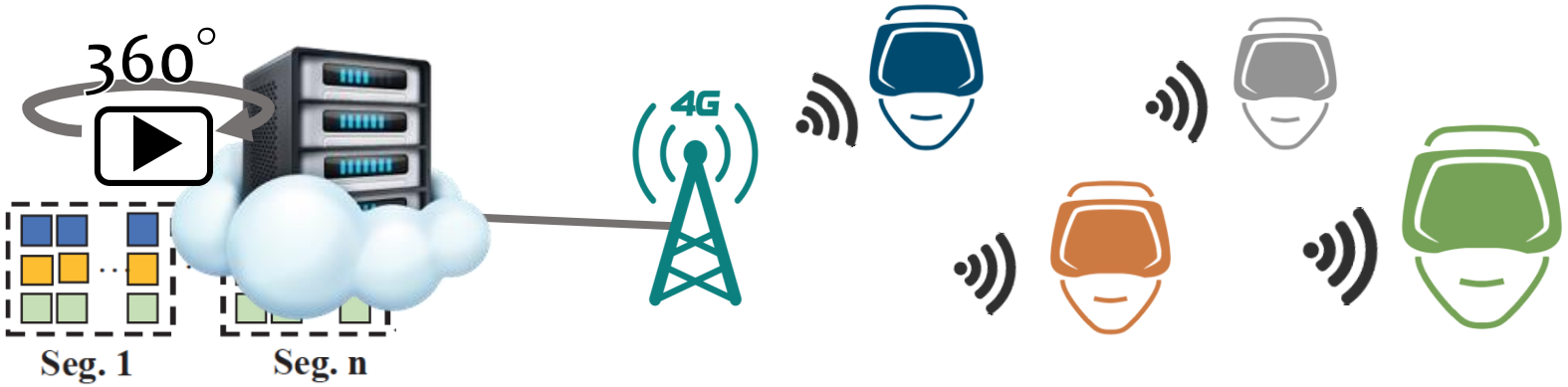
Viewer Orientation

Available Bandwidth

Time

Before Streaming 360° Videos over 4G Networks

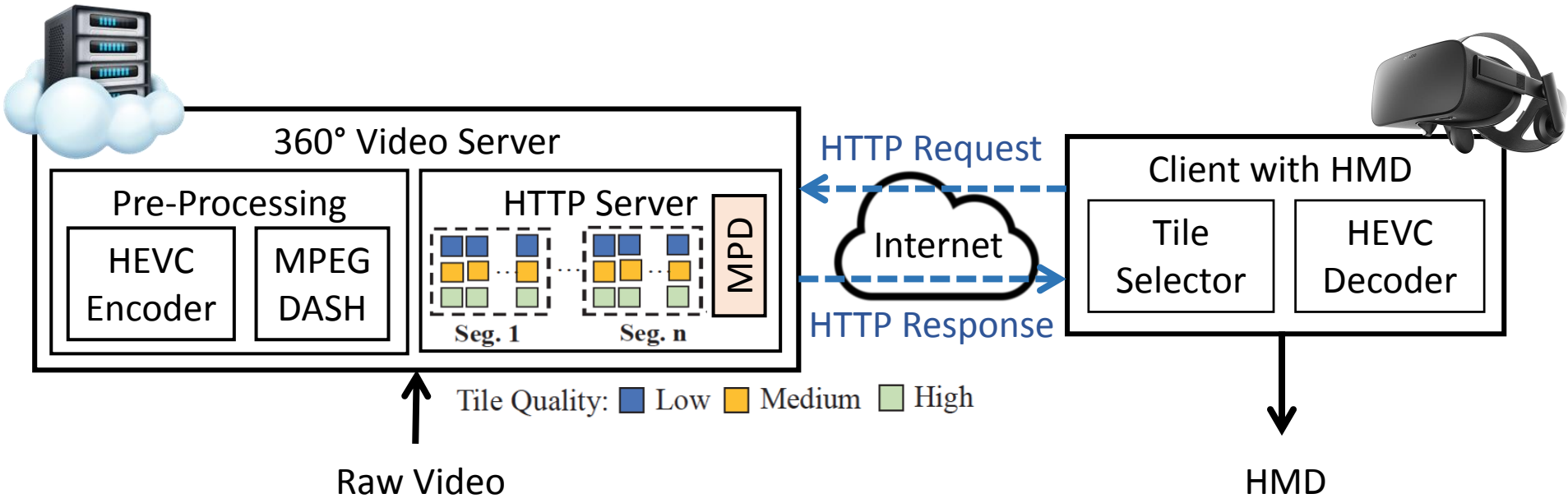
- Tiling with MPEG DASH -> reduce bandwidth
- Questions to answer
  - How tile size affects the streaming system performance?
  - How much bandwidth saving can we get by selectively requesting useful tiles?
  - How many users can be supported in one 4G cell?



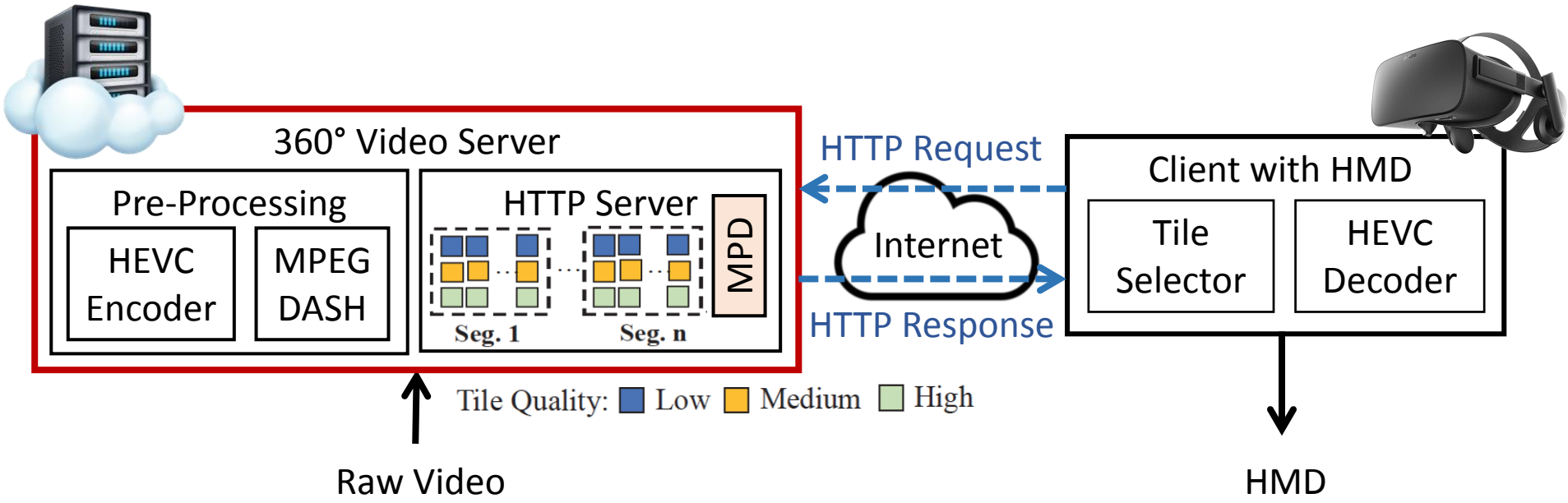
Contributions

- We design and implement an end-to-end 360° video streaming system to Head Mounted Displays
- We evaluate our system's performance over a real 4G cellular network to answer the three questions
- We collect and share (upon request) the dataset collected with our system

Overview of the Tile-Based Streaming Platform

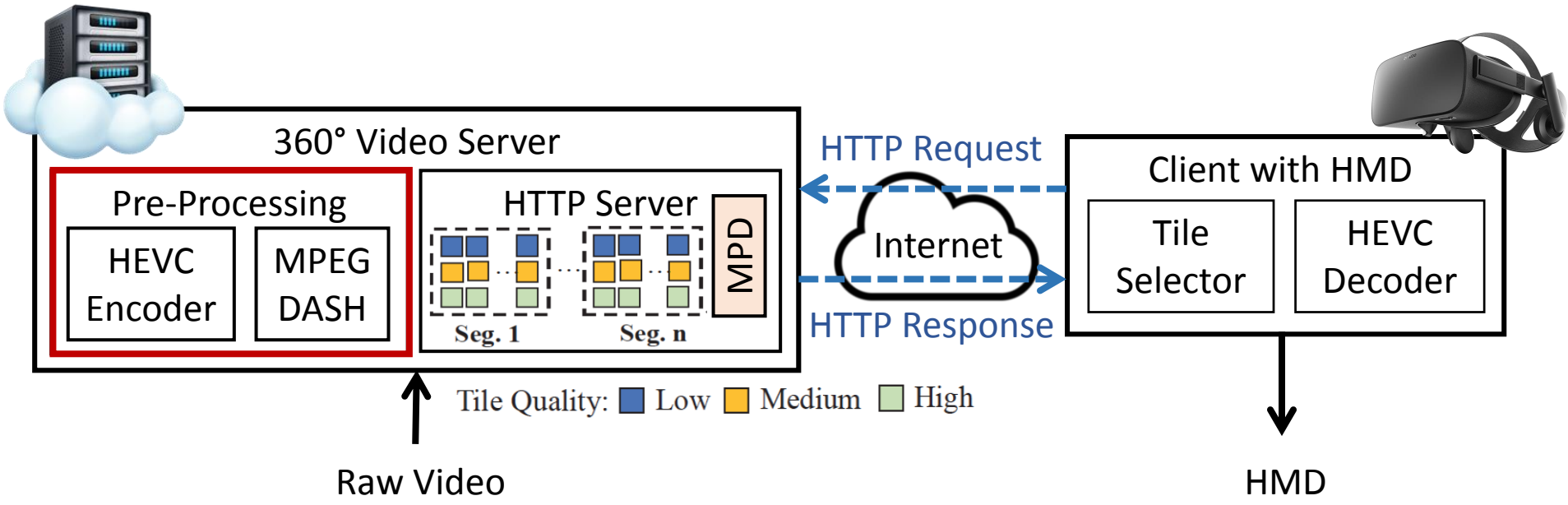


Overview of the Tile-Based Streaming Platform



- 360° Video Server

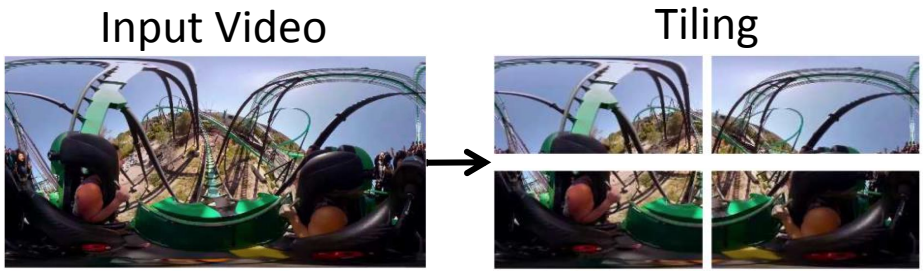
Overview of the Tile-Based Streaming Platform



- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator

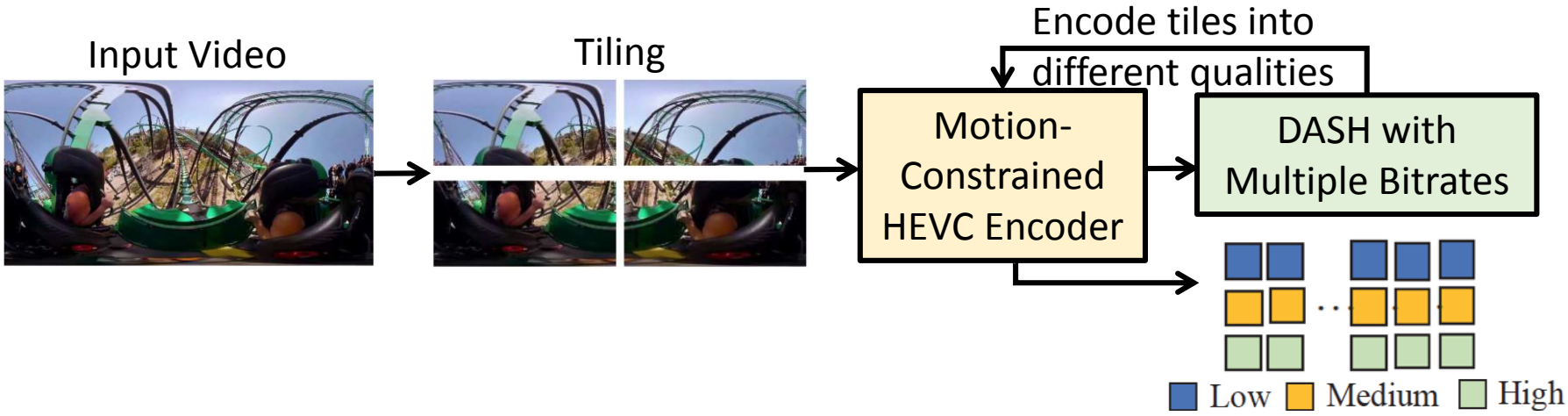
[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

The Pre-Processing Procedure



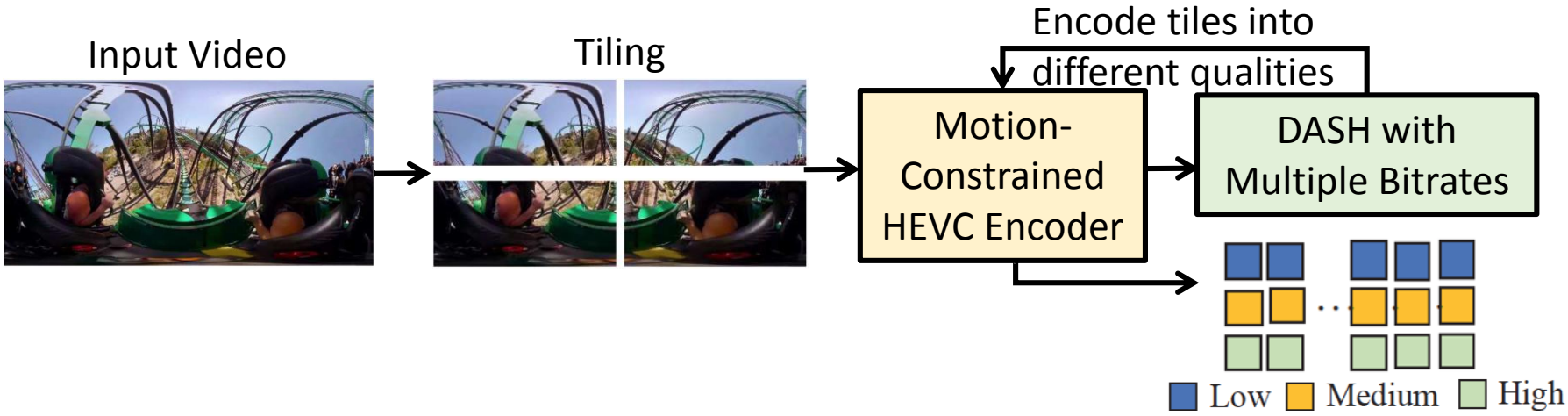
- Split the videos into tiles of sub-videos

The Pre-Processing Procedure



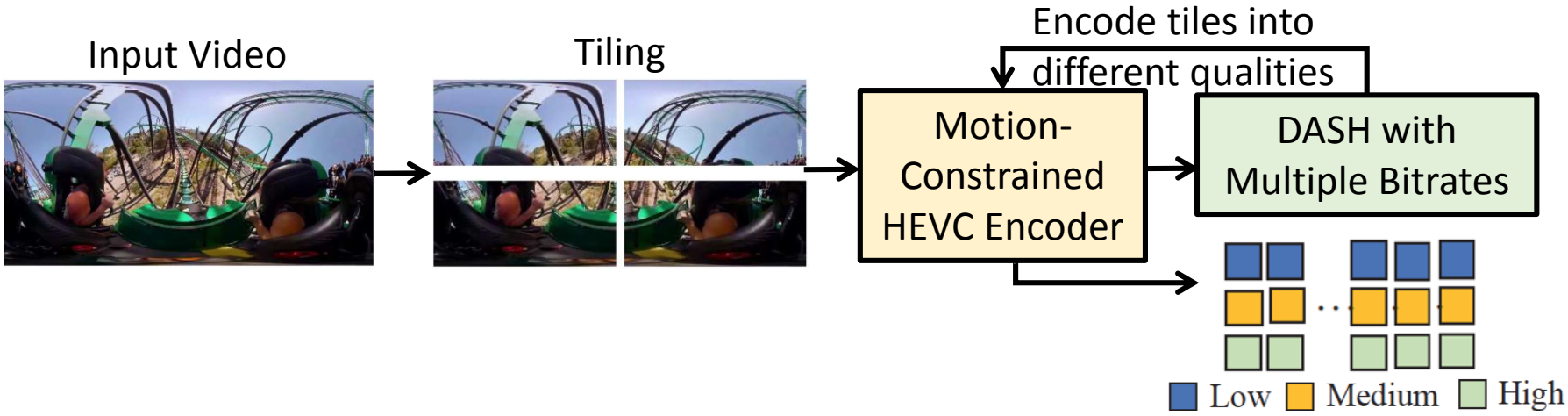
- Split the videos into tiles of sub-videos
- Encode the tiles using motion-constrained HEVC encoder with different bitrates (qualities)

The Pre-Processing Procedure



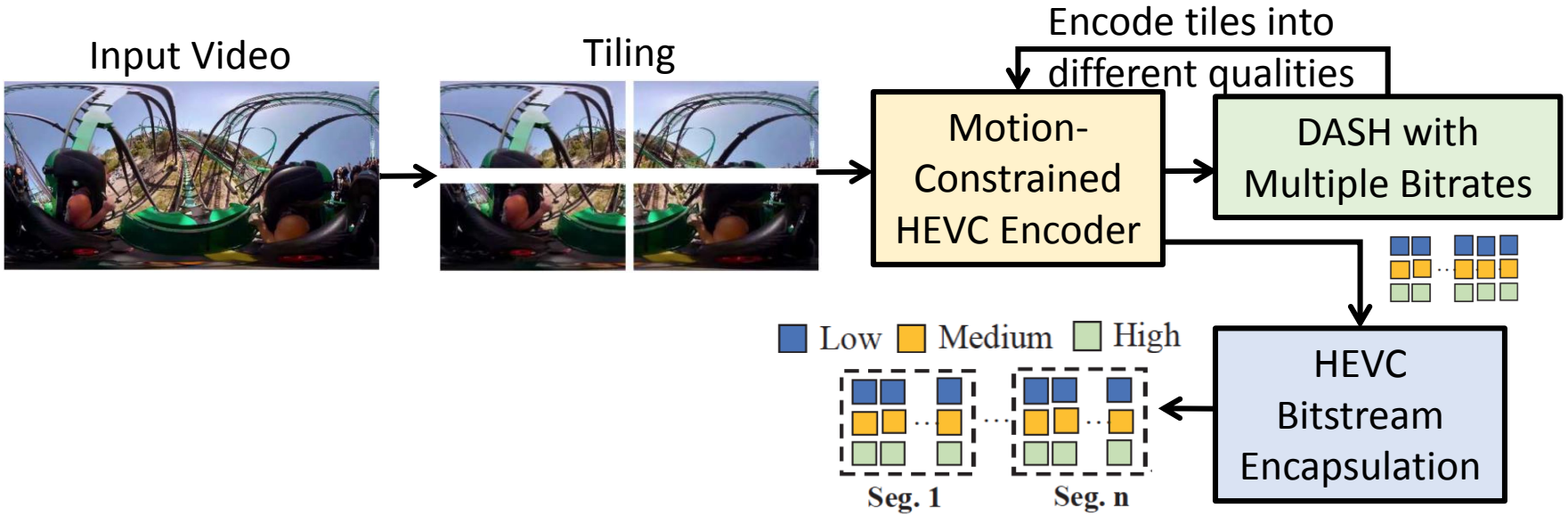
- Split the videos into tiles of sub-videos
- Encode the tiles using **motion-constrained** HEVC encoder with different bitrates (qualities)
  - constrain the tiles encoding so that each tile only refers to the same tiles in previous or future frames -> avoid decoding glitches

The Pre-Processing Procedure



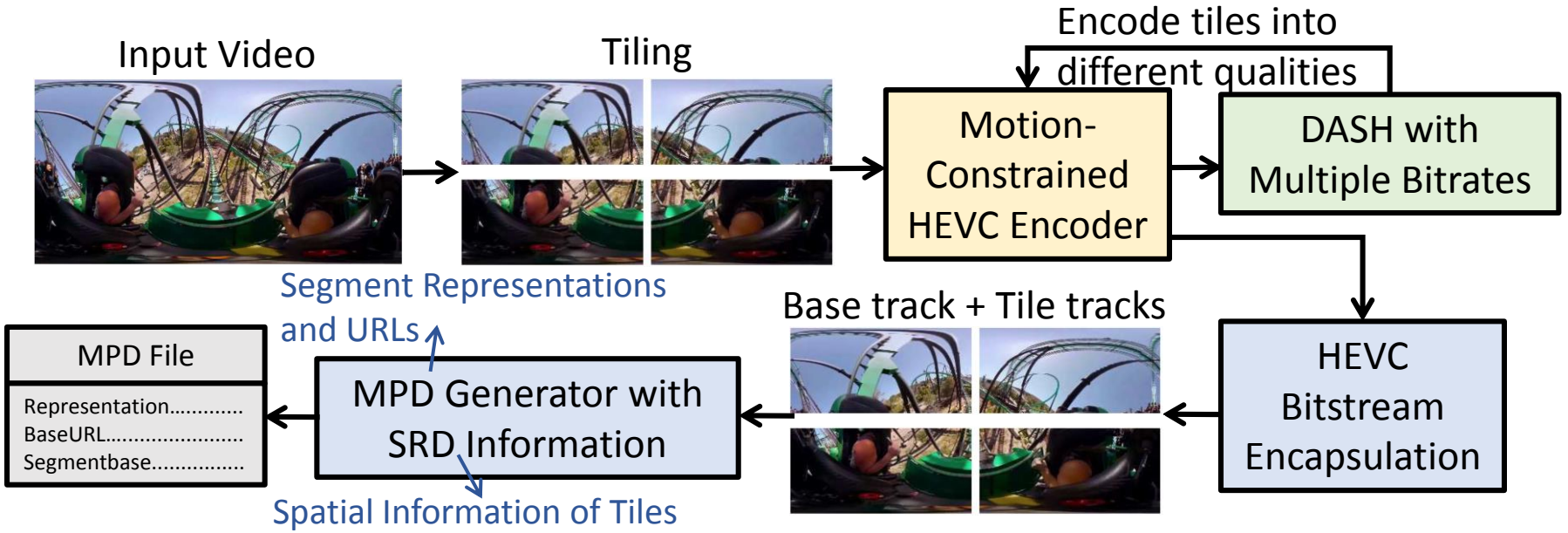
- Split the videos into tiles of sub-videos
- Encode the tiles using motion-constrained HEVC encoder with **different bitrates (qualities)**
  - constrain the tiles encoding so that each tile only refers to the same tiles in previous or future frames -> avoid decoding glitches
  - adapt to network condition

The Pre-Processing Procedure



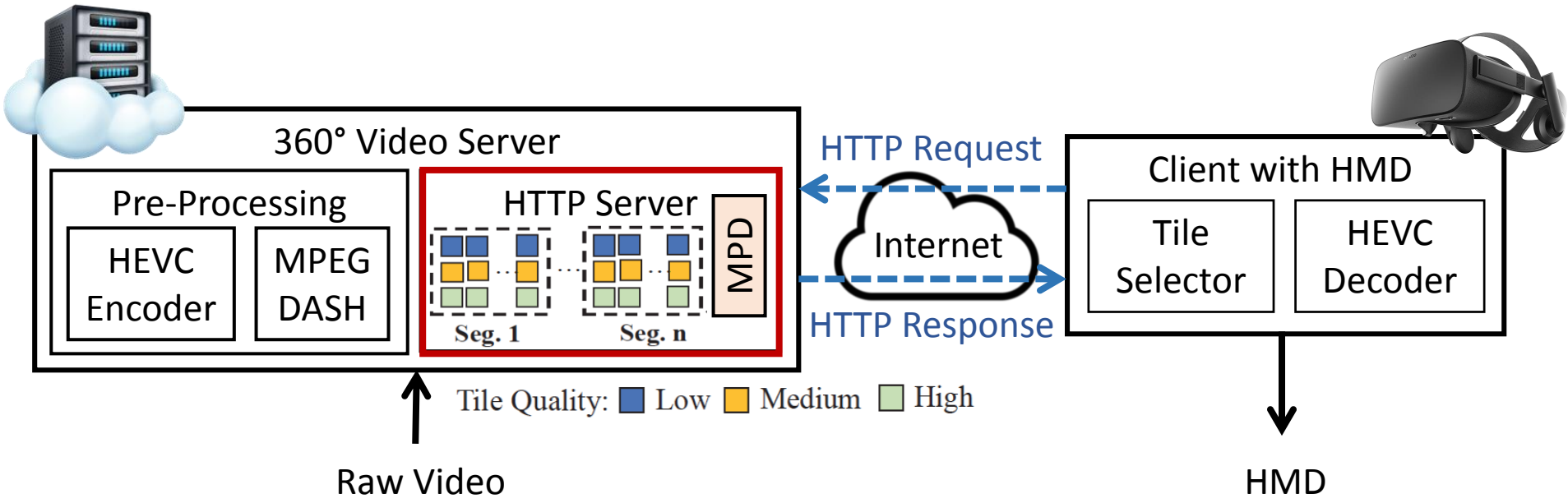
- Split the videos into tiles of sub-videos
- Encode the tiles using motion-constrained HEVC encoder with different bitrates (qualities)
- Encapsulate tiles into single HEVC bitstream

The Pre-Processing Procedure



- Split the videos into tiles of sub-videos
- Encode the tiles using motion-constrained HEVC encoder with different bitrates (qualities)
- Encapsulate tiles into single HEVC bitstream
- Integrate with DASH for spatial index generation (MPD and SRD)

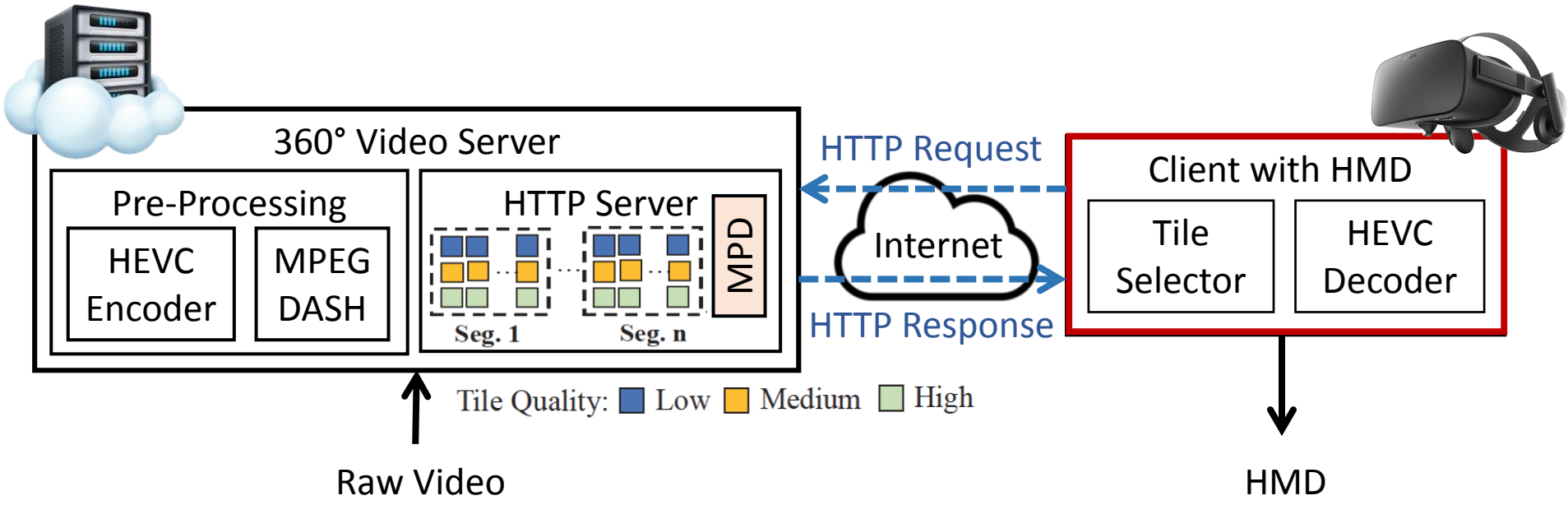
Overview of the Tile-Based Streaming Platform



- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator
  - HTTP Server

[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

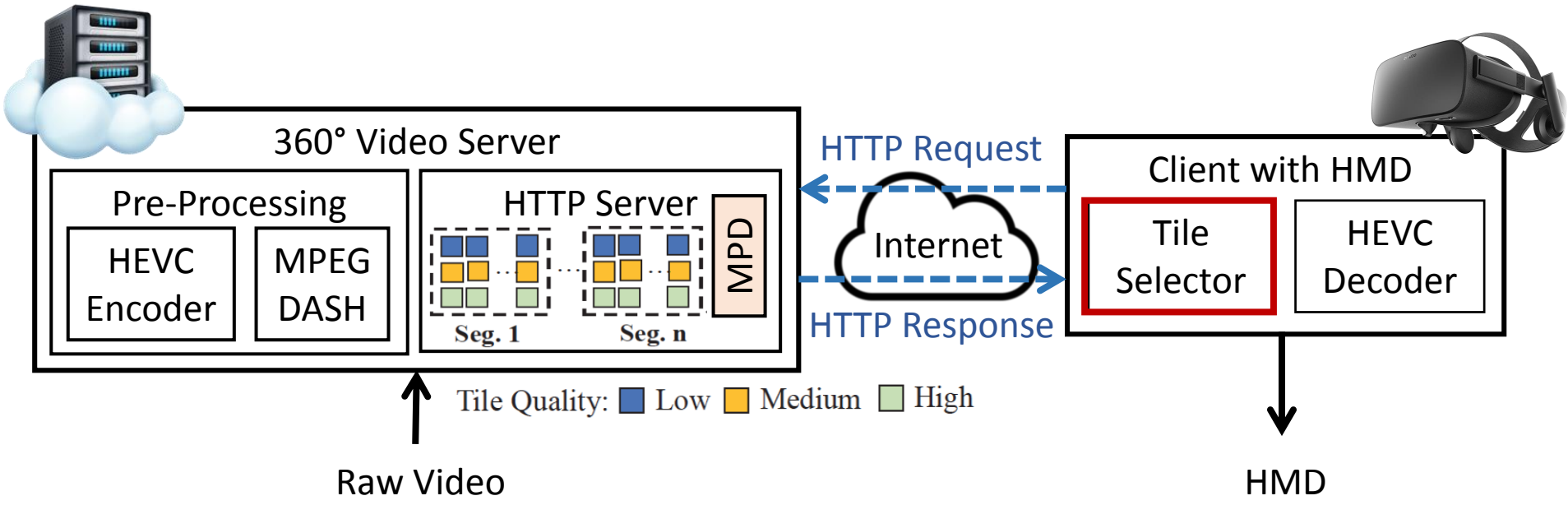
Overview of the Tile-Based Streaming Platform



- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator
  - HTTP Server
- Client with HMD

[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

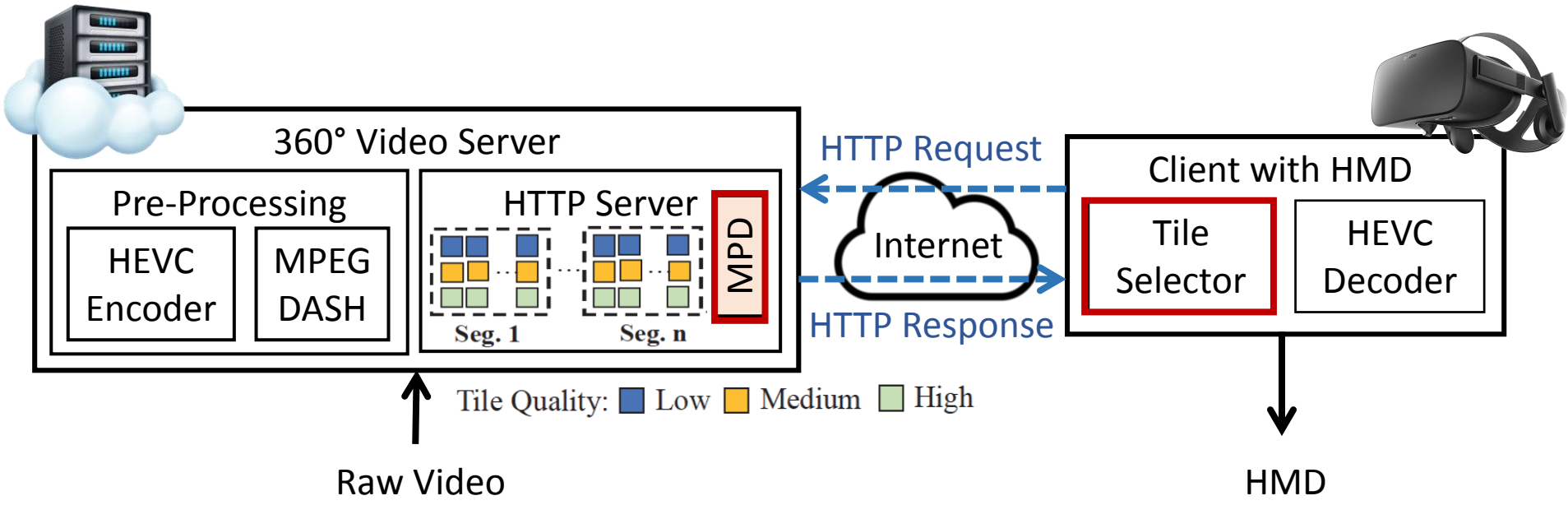
Overview of the Tile-Based Streaming Platform



- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator
  - HTTP Server
- Client with HMD
  - Tile selector

[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

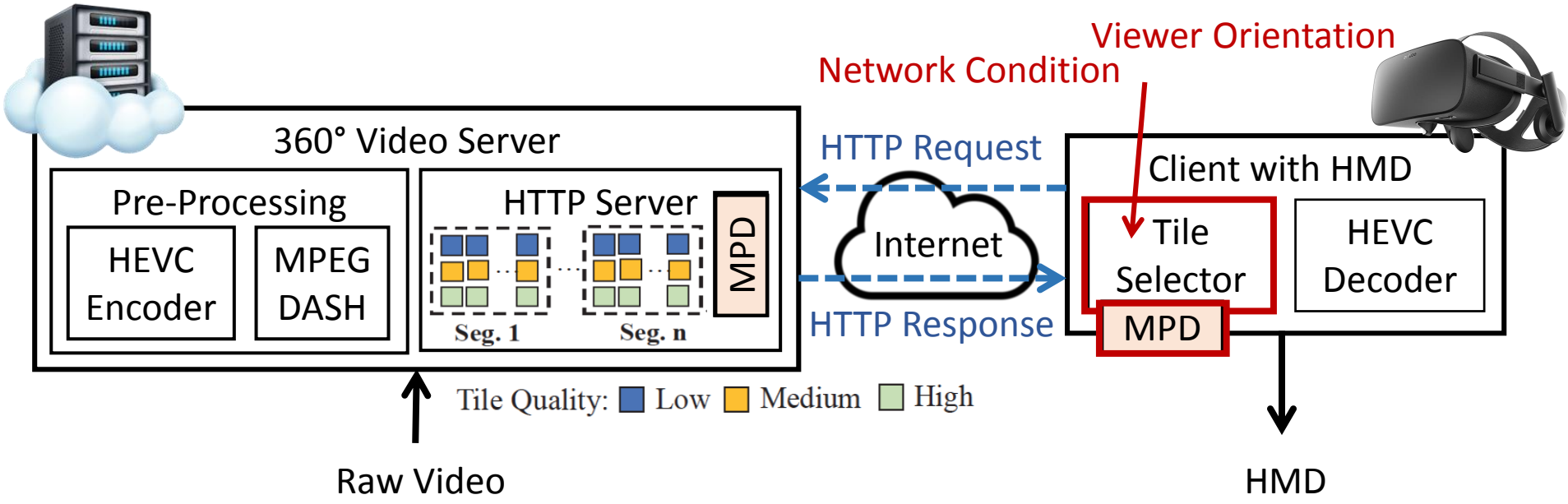
Overview of the Tile-Based Streaming Platform



- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator
  - HTTP Server
- Client with HMD
  - Tile selector

[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

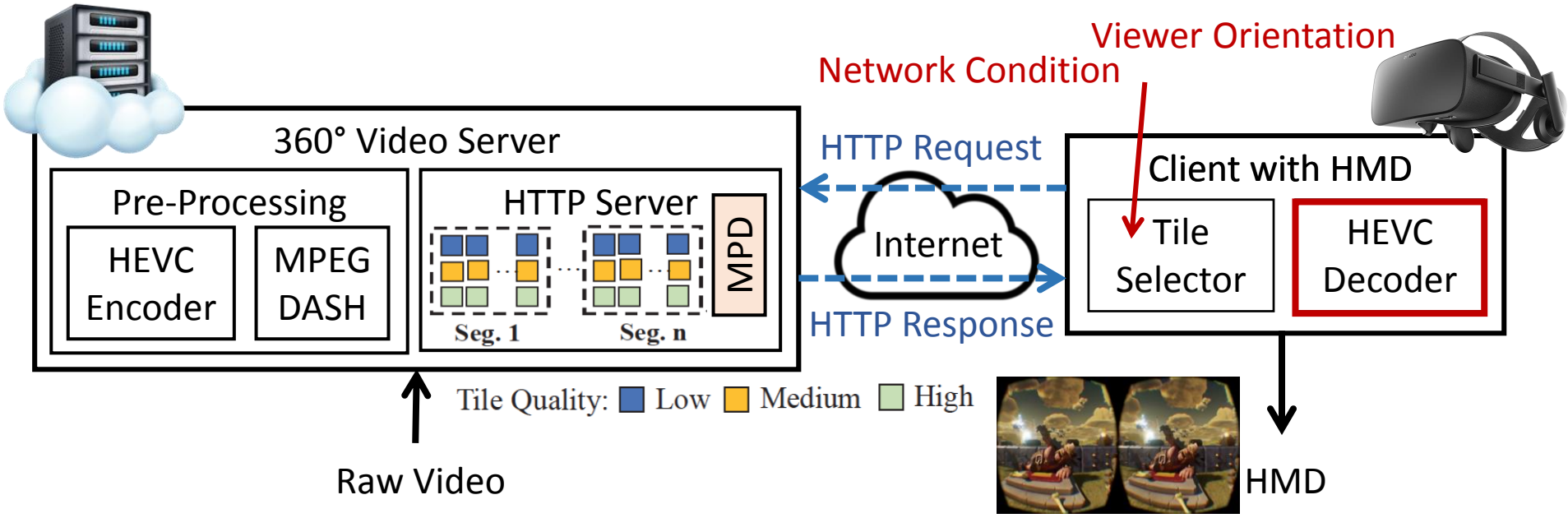
Overview of the Tile-Based Streaming Platform



- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator
  - HTTP Server
- Client with HMD
  - Tile selector

[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

Overview of the Tile-Based Streaming Platform

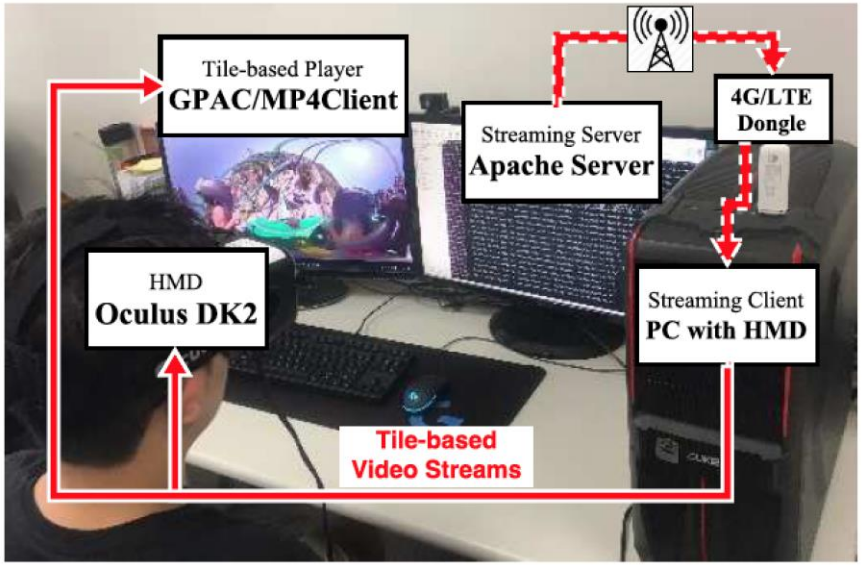


- 360° Video Server
  - HEVC[1] encoder
  - MPEG DASH[2] content generator
  - HTTP Server
- Client with HMD
  - Tile selector
  - HEVC decoder

[1] G. Sullivan et al. "Overview of the high efficiency video coding (HEVC) standard." Sullivan, Gary J., et al. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology* 22 (12), 2012, 1649-1668.
[2] ISO/IEC DIS 23009-1.2 Dynamic adaptive streaming over HTTP (DASH)

Testbed

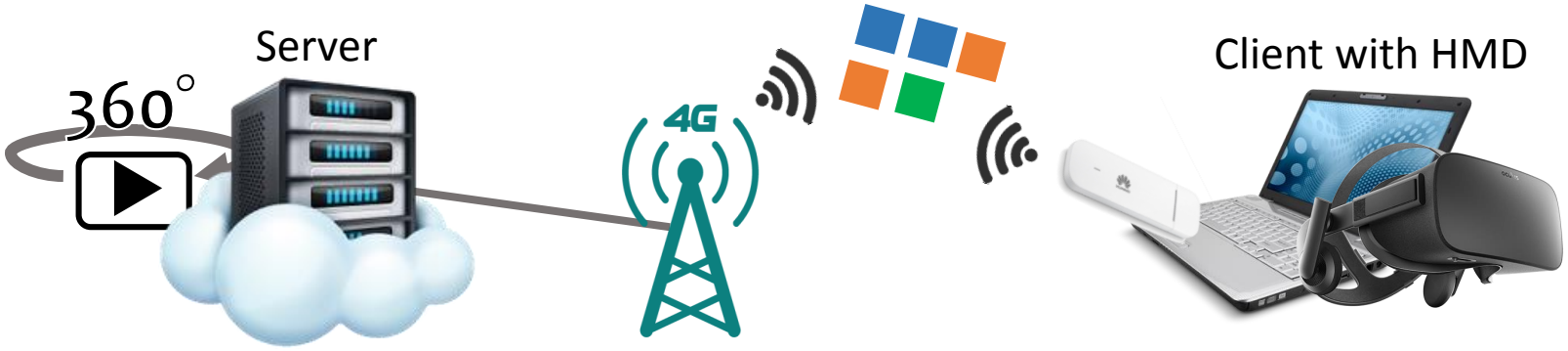
- Server
  - Kvazaar[1] HEVC encoder
  - MP4Box[2] DASH content generator
  - Apache HTTP server[3]
- Client
  - Oculus Rift DK2[4]
  - MP4Client[5]
- 4G LTE Network
  - RBS 6601 base station
  - HUAWEI E3267 4G dongle



[1] Kavazaar, an opensource HEVC encoder. https://github.com/ultravideo/kvazaar
[2] MP4Box , a multimedia packager. https://gpac.wp.imt.fr/mp4box/
[3] Apache HTTP server. https://httpd.apache.org/
[4] Oculus Rift Development Kit 2 (DK2). https://www3.oculus.com/en-us/dk2/
[5] MP4Client, an opensource multimedia player. https://gpac.wp.imt.fr/player/

Experiment Setup

- Number of tiles: {1x1, 3x3, 5x5, 7x7, 9x9}
- DASH segment length: {1, 4, 10} secs
- Bitrates: {3, 6, 12} Mbps
- Viewer's FoV is randomly chosen from dataset[1]
- Perform 3 times and report the average



[1] W. Lo, C. Fan, J. Lee, C. Huang, K. Chen, and C. Hsu, "360◦ video viewing dataset in head-mounted Virtual Reality," in Proc. of ACM MMSys'17

Measurement Design

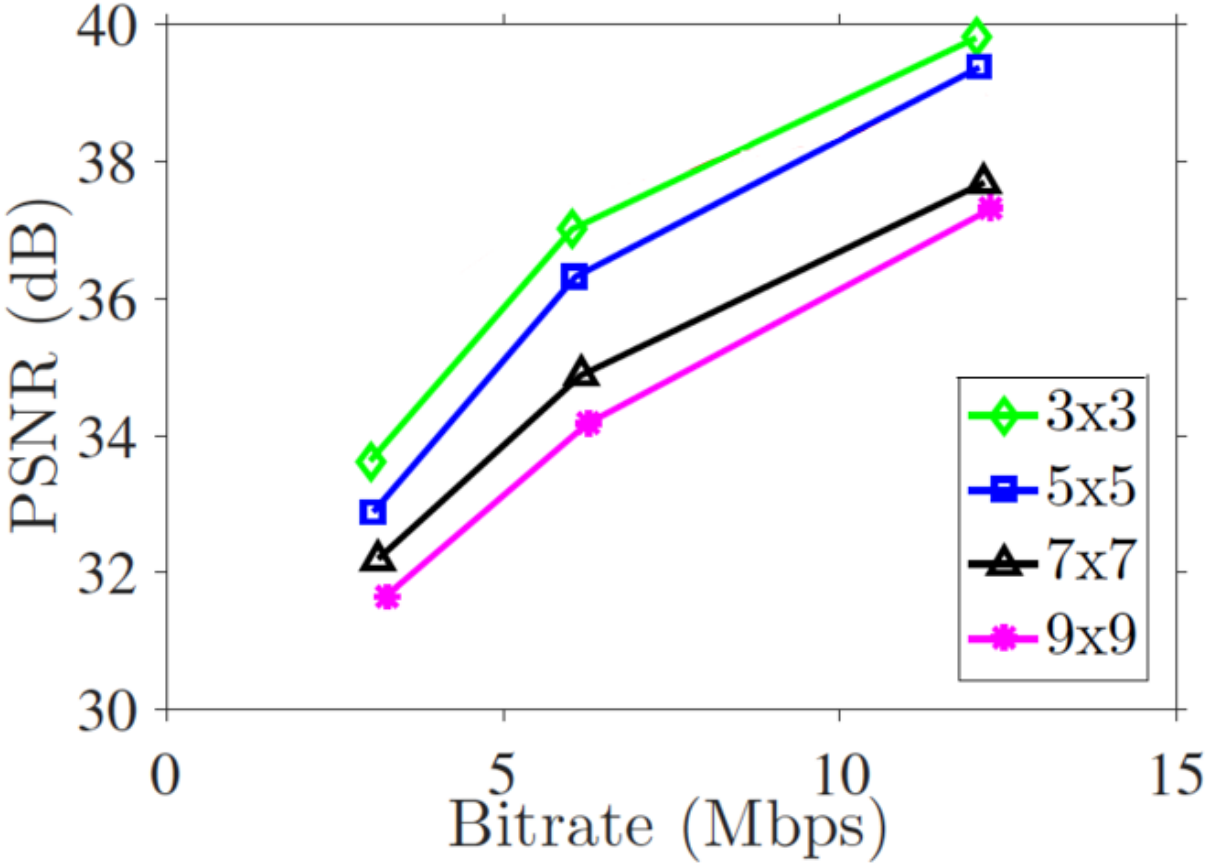
- How tile size affects the streaming system performance?
- How much bandwidth saving can we get by selectively requesting useful tiles?
- How many users can be supported in one cell?

Measurement Design

- How tile size affects the streaming system performance?
  - vary the number of tiles and the bitrates
- How much bandwidth saving can we get by selectively requesting useful tiles?
- How many users can be supported in one cell?

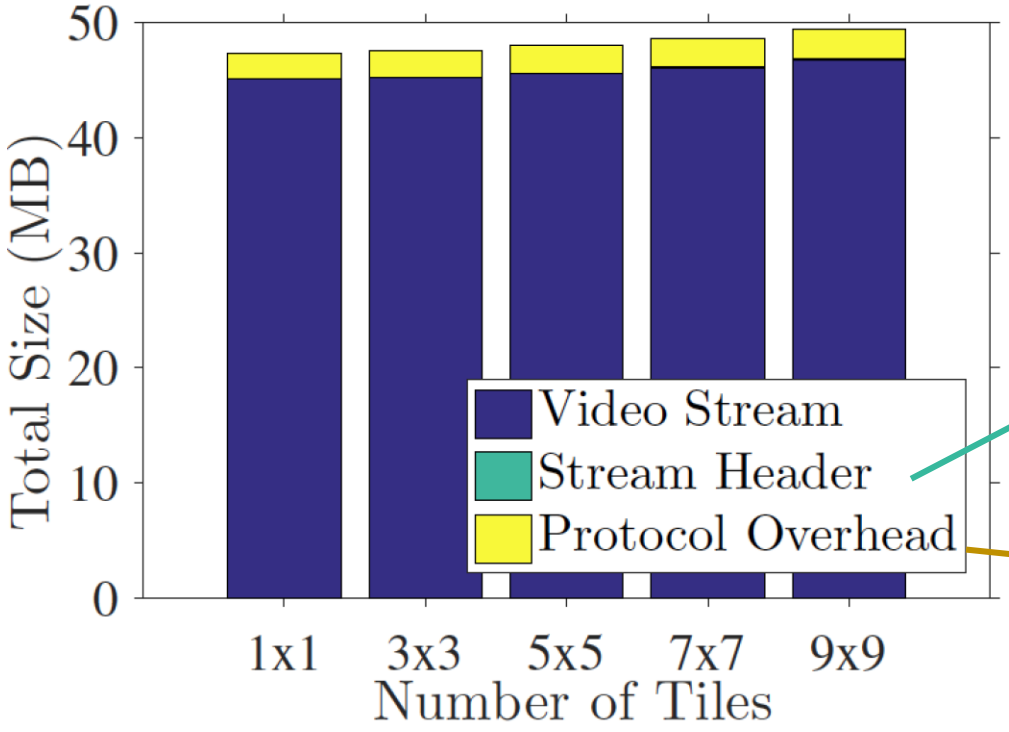
The Number of Tiles ↑ ⇒ Coding Efficiency ↓

- Due to motion constraints among tiles



Number of Tiles ↑, Protocol Overhead ↑

- The majority of the streamed tiles are videos
- The protocol overhead is always less than 3%

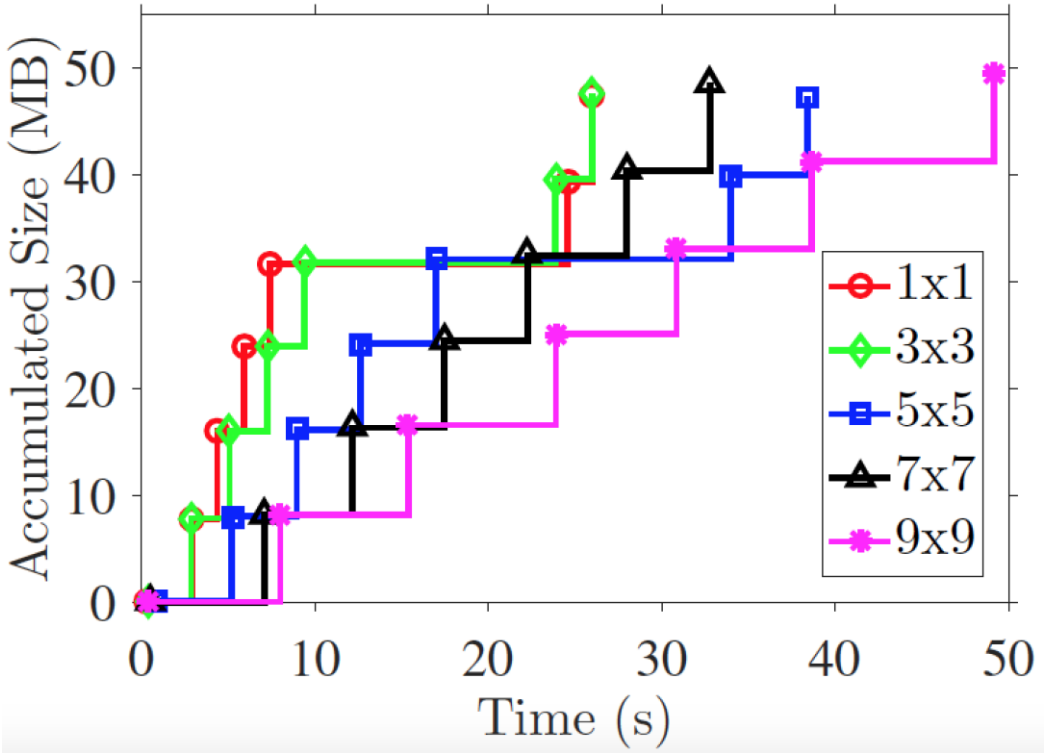


initial header files and MPD files

generated during streaming

More Tiles Consume More Time to Download

- Reasons
  - Sequentially download
  - Tile encapsulation overhead

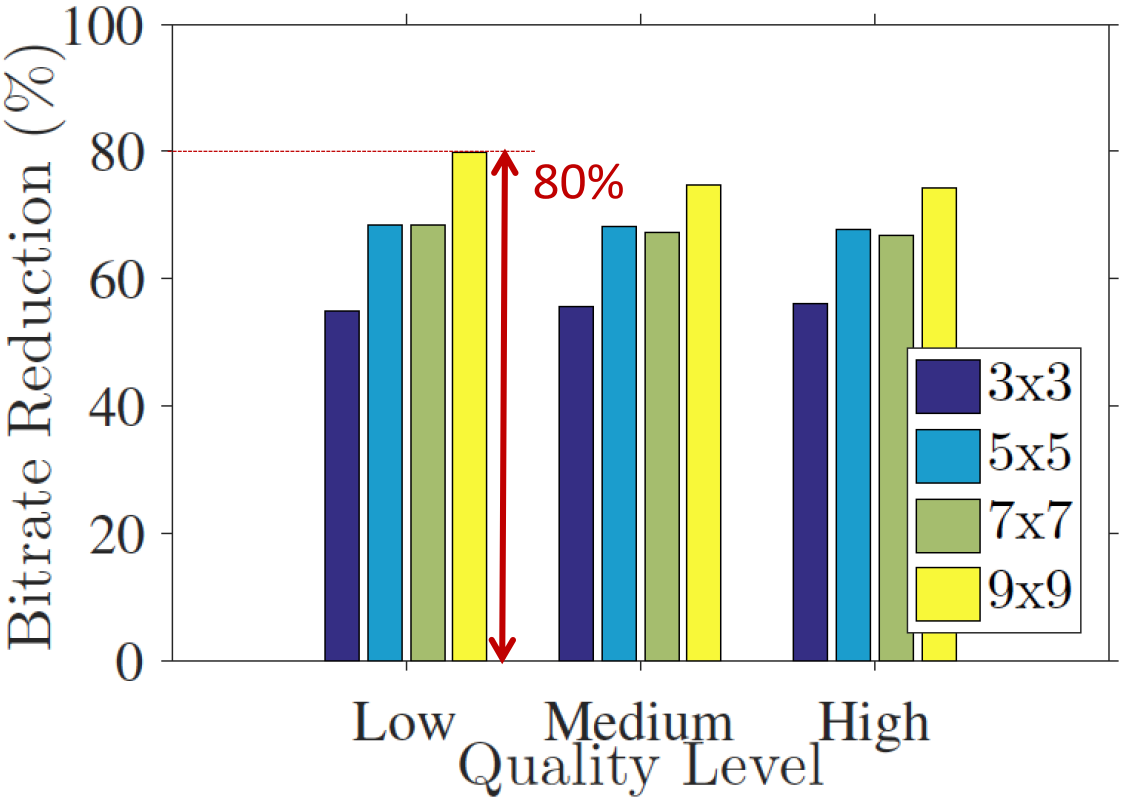


Measurement Design

- How tile size affects the streaming system performance?
  - vary the number of tiles and the bitrates
- How much bandwidth saving can we get by selectively requesting useful tiles?
  - modify the client to only request the tiles that will be watched by viewers
- How many users can be supported in one cell?

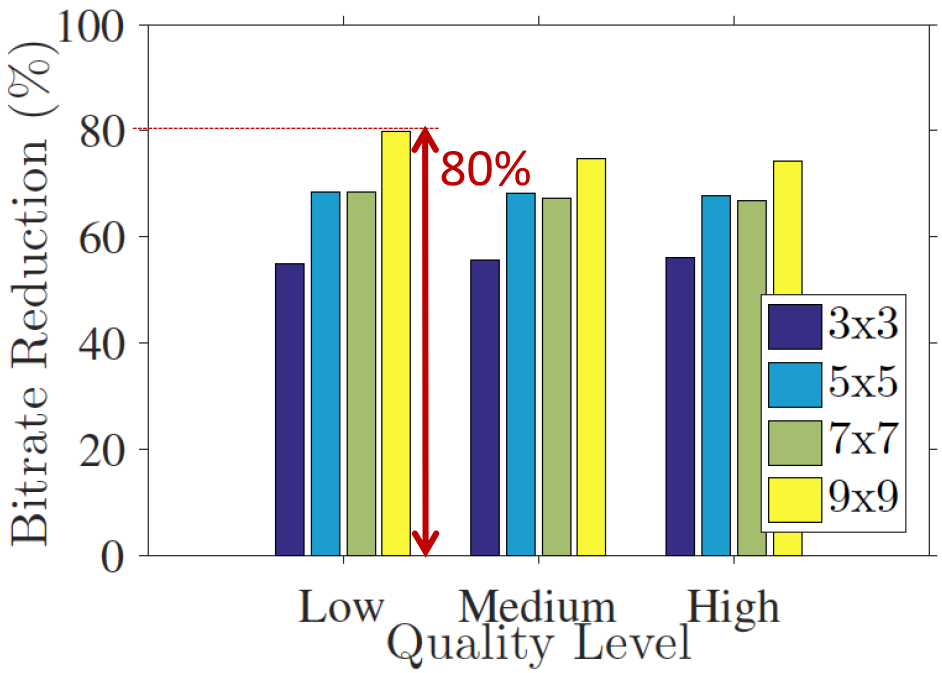
The Bandwidth Saving of FoV-Based Streaming

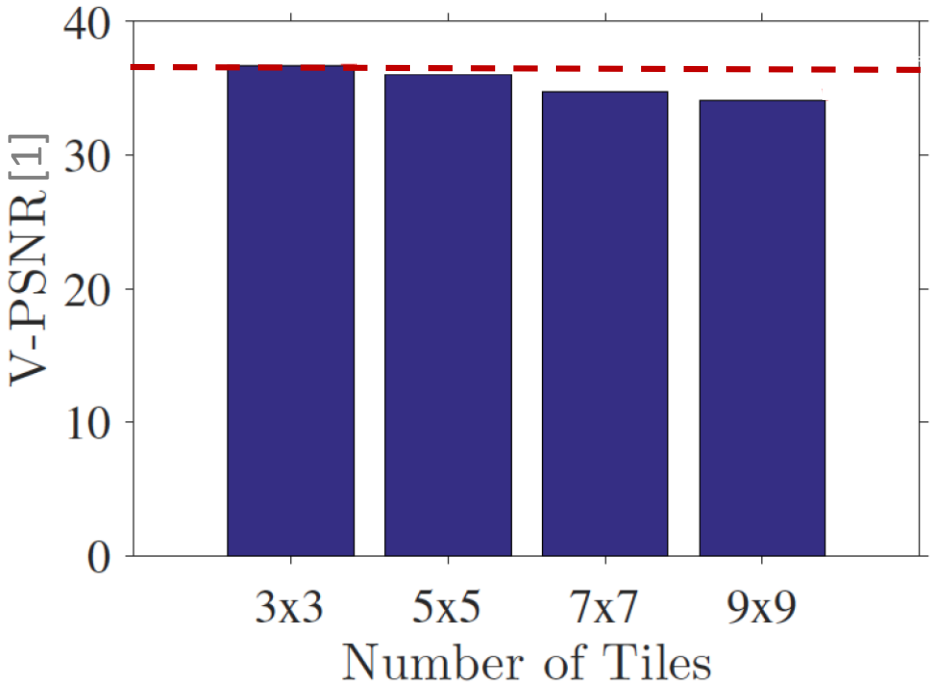
- Skipping tiles based on viewer's FoV saves the bandwidth by up to 80%



More Tiles Incurs Minor Video Quality Drop

- Reduces the amount of data and incurs minor video quality drop





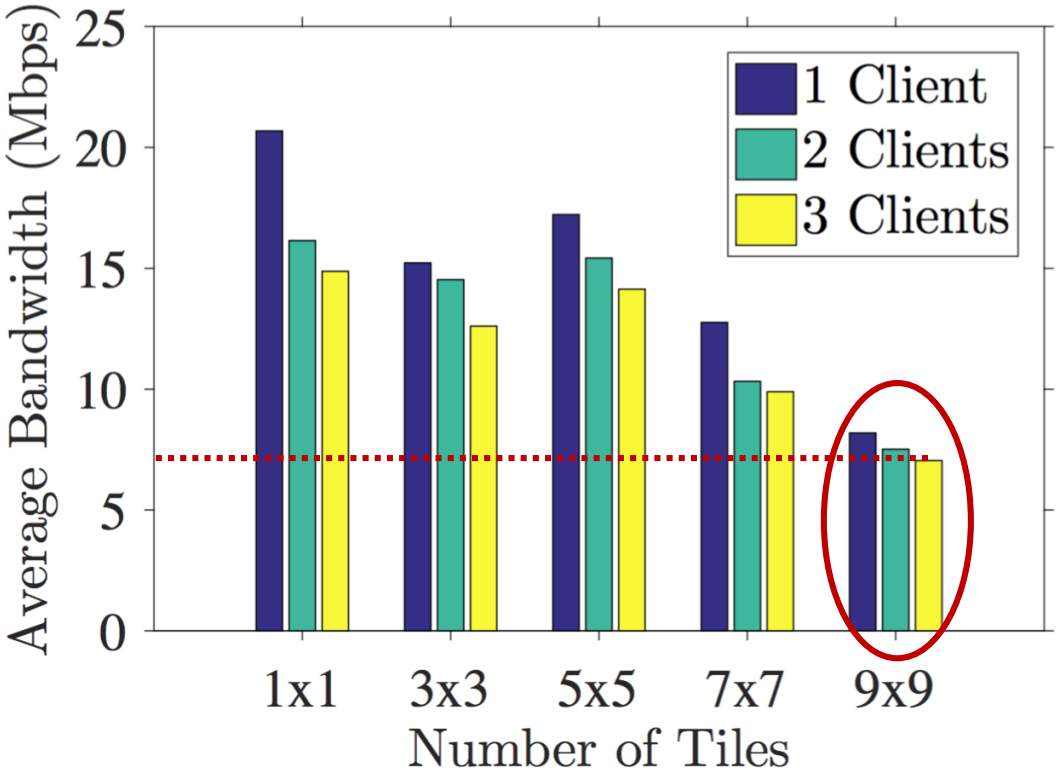
[1] M. Yu, H. Lakshman, and B. Girod, "A framework to evaluate omnidirectional video coding schemes," in Proc. of IEEE International Symposium on Mixed and Augmented Reality (ISMAR'15)

Measurement Design

- How tile size affects the streaming system performance?
  - vary the number of tiles and the bitrates
- How much bandwidth saving can we get by selectively requesting useful tiles?
  - modify the client to only request the tiles that will be watched by viewers
- How many users can be supported in one cell?
  - repeat the above experiments but with more clients

Scalability

- Our streaming testbed can support at least 3 clients
- More number of tiles -> smaller average bandwidth



Conclusion

- We design measurement experiments to quantify the performance of VR streaming over cellular networks
- We build a streaming testbed and conduct extensive experiments using real user trace[1]
  - FoV-based video streaming can save up to **80%** in bandwidth consumption
  - More tiles suffer from lower coding efficiency and late segments

[1] W. Lo, C. Fan, J. Lee, C. Huang, K. Chen, and C. Hsu, "360◦ video viewing dataset in head-mounted Virtual Reality," in Proc. of ACM MMSys'17

Future Work

- This work can be extended to optimal bitrate allocation for mobile AR/VR systems with HMDs

Q&A

ch.ling.fan@gmail.com