Preemptive Multiplexed HTTP Streaming of 360° Tiled Videos to Head-Mounted Displays

Shou-Cheng Yen

Advisor: Cheng-Hsin Hsu

Networking and Multimedia Systems Lab,
CS Dept., National Tsing Hua University

Outline

- Introduction
- Challenges
- System Overview
- Adaptive BitRate (ABR) Algorithms
- Evaluations
- Conclusion

Introduction

Bandwidth Requirement for VR Systems

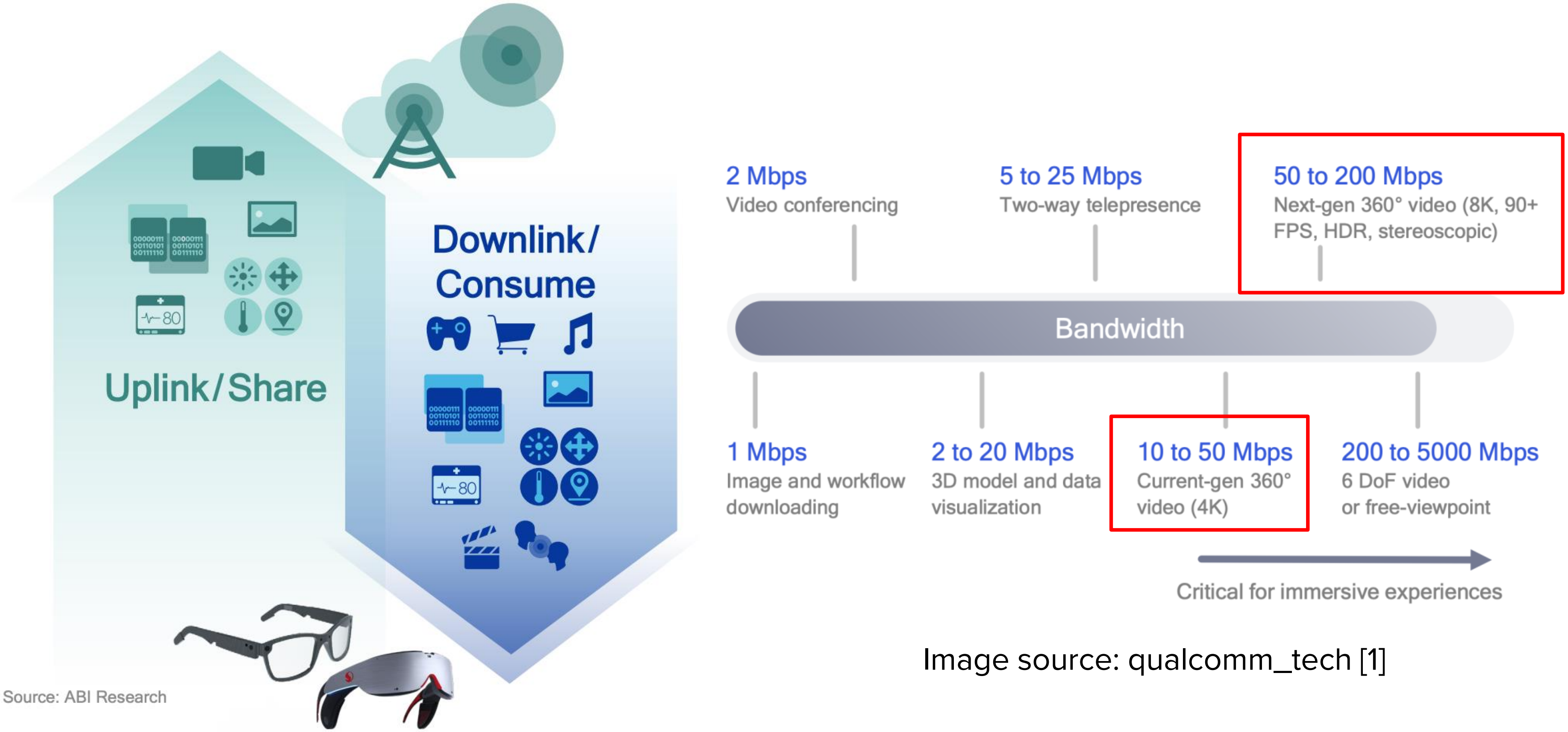


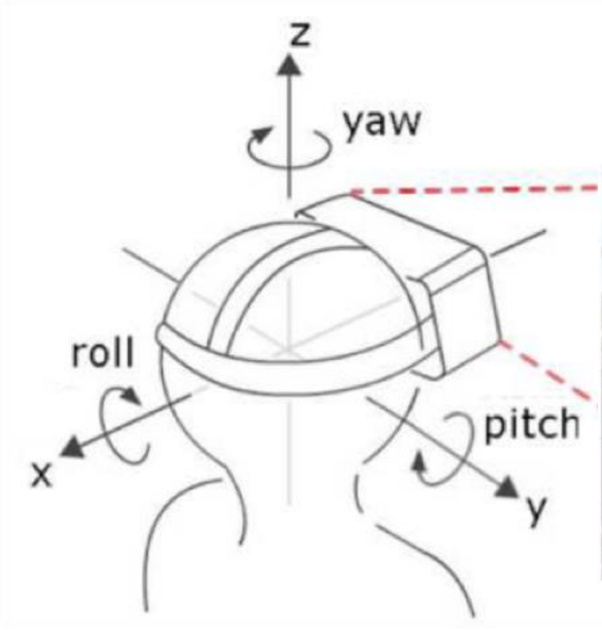
Image source: qualcomm_tech [1]

High network bandwidth is required

[1] https://www.qualcomm.com/media/documents/files/the-mobile-future-of-augmented-reality.pdf

Only Stream Field-of-View (FoV)

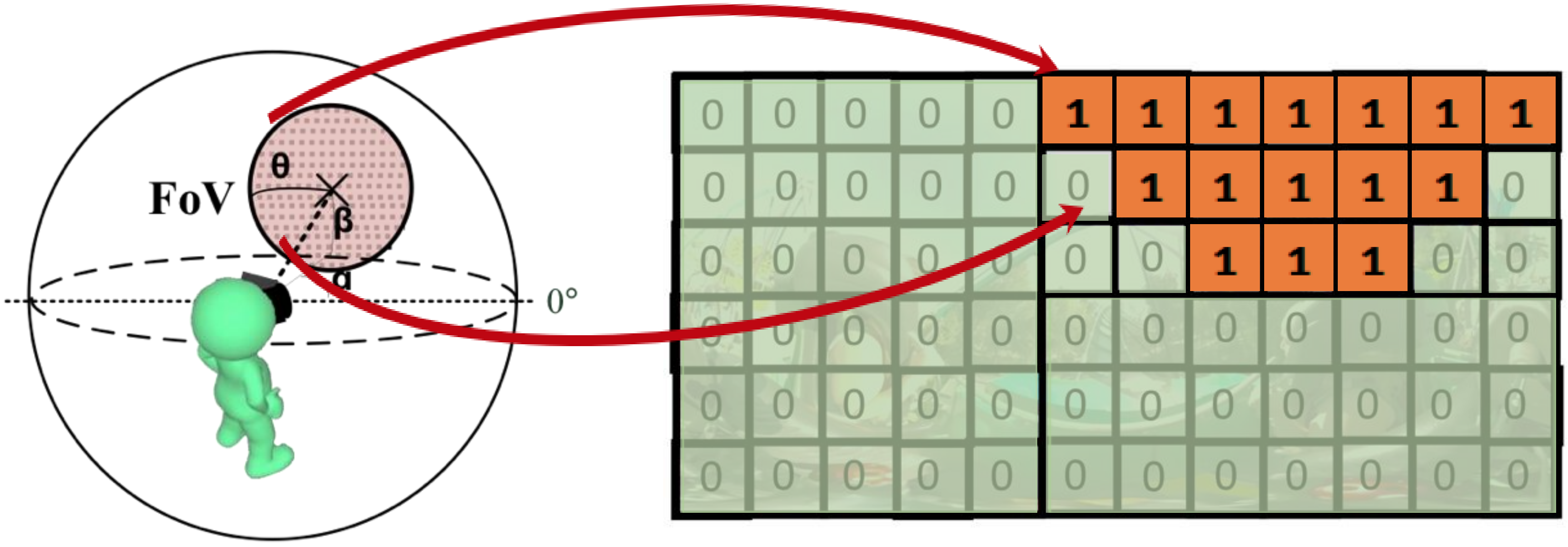
- The viewer actively changes the viewing orientation by rotating his/her head
- The HMD viewer only gets to see a small part of the whole 360˚ video (< **1/3**)



HMD: **Oculus**

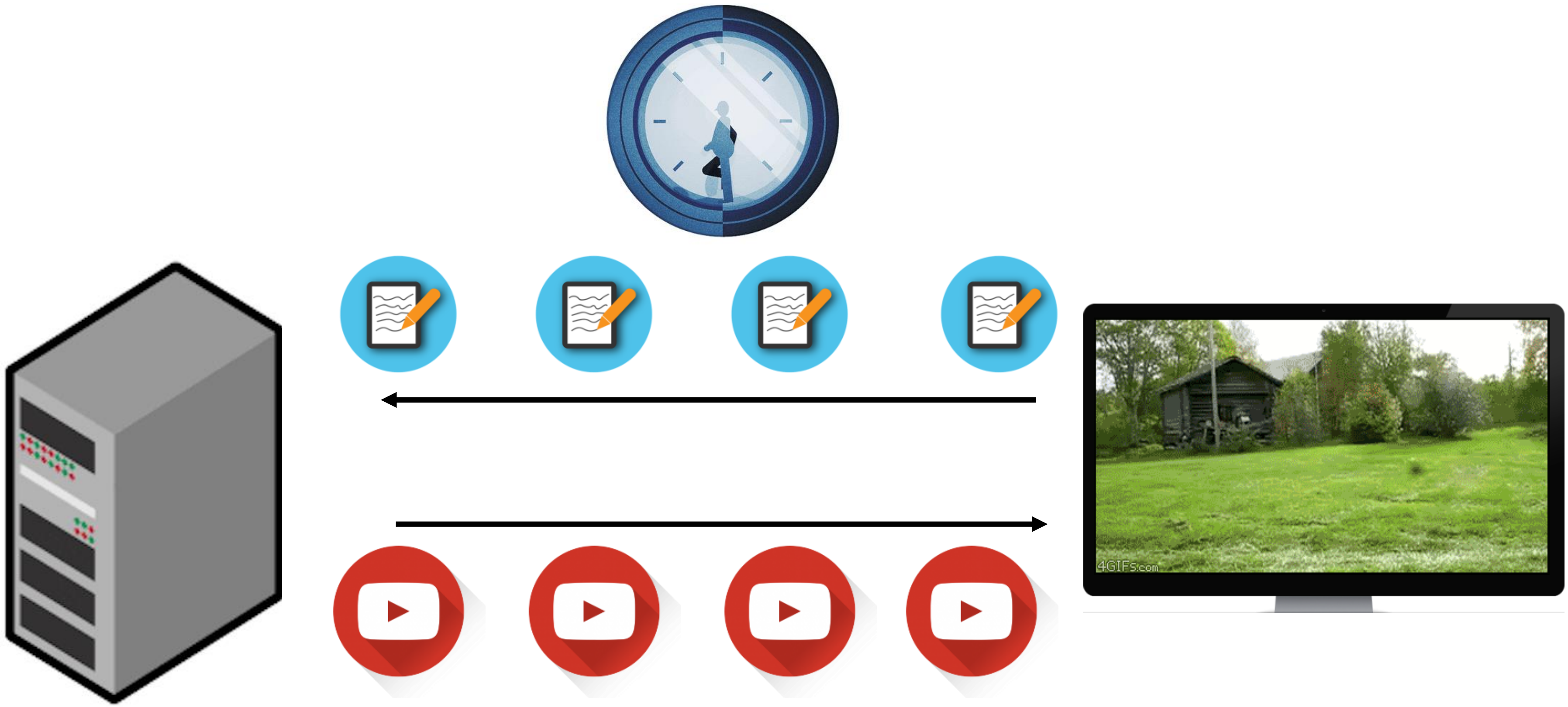
Tiled Streaming

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



Challenges

Impacts on Network Latency



FOV solution needs to rely on

fixation prediction algorithms

Viewer Dynamics and Imperfect Fixation Prediction Algorithms

Ideal

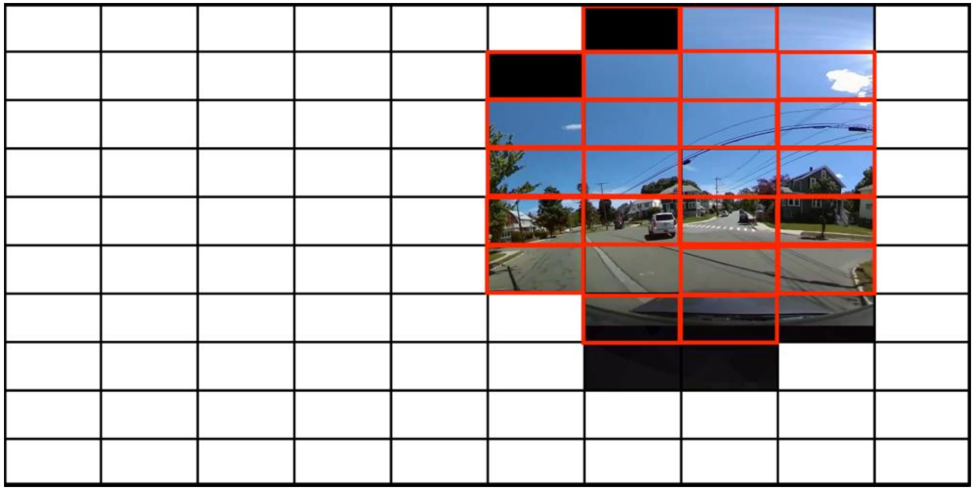
Imperfect prediction

Missing tiles (black holes) in viewports

Fixation Prediction Algorithms

Accuracy of fixation prediction algorithms

# of overlapped tiles / # of ground truth tiles



accuracy: 13/15 = 86.6 %
missing ratio: 2/15 = 13.4 %

Fixation Prediction Algorithms

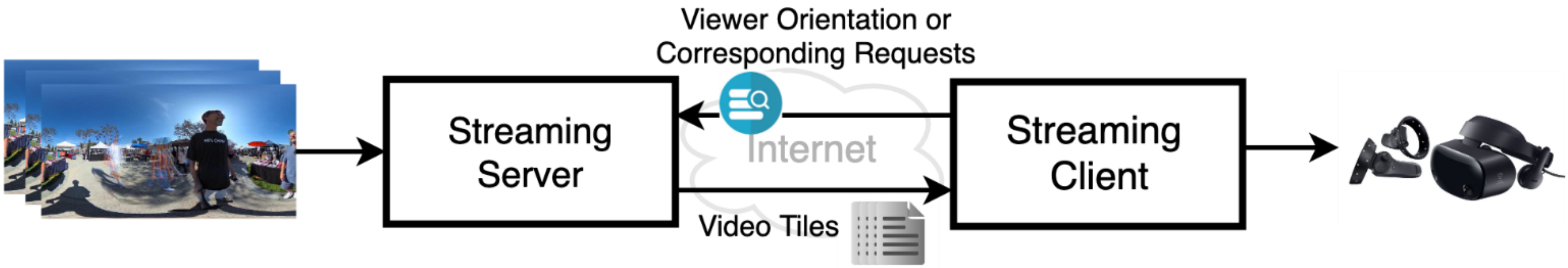
Literature	Approach	Accuracy
Fan et al. [1]	LSTM	86.35%
Nguyen et al. [2]	LSTM	82%
Monroy et al. [3]	CNN	80.8%
Mavlankar et al. [4]	Dead Reckoning	78.8%

★ **Accuracy of fixation prediction algorithms is only around 80%**

[1] Ching-Ling Fan, Jean Lee, Wen-Chih Lo, Chun-Ying Huang, Kuan-Ta Chen, and Cheng-Hsin Hsu. 2017. Fixation Prediction for 360° Video Streaming in Head- Mounted Virtual Reality. In Proc. of ACM Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'17). Taipei, Taiwan, 67– 72.
[2] Anh Nguyen, Zhisheng Yan, and Klara Nahrstedt. 2018. Your Attention is Unique: Detecting 360-Degree Video Saliency in Head-Mounted Display for Head Movement Prediction. In Proc. of ACM International Conference on Multimedia (MM'18). Seoul, South Korea, 1190–1198
[3] Rafael Monroy, Sebastian Lutz, Tejo Chalasani, and Aljoscha Smolic. 2017. Sal- Net360: Saliency Maps for omni-directional images with CNN. Signal Processing: Image Communication (September 2017).
[4] Aditya Mavlankar and Bernd Girod. 2010. Video streaming with interactive pan/tilt/zoom. In High-Quality Visual Experience. Springer, 431–455.

Streaming Protocols

- **DASH** employs HTTP over TCP
    - Suitable for **presentational** and **unidirectional** video streaming
- Real-time Transport Protocol (RTP)
    - Unreliable transmission
    - Network congestion
- 360° tiled video streaming is more **interactive**



Naively applying DASH for 360° tiled video streaming leads to suboptimal streaming quality

DASH: Dynamic Adaptive Streaming over HTTP

Streaming Protocol Comparisions

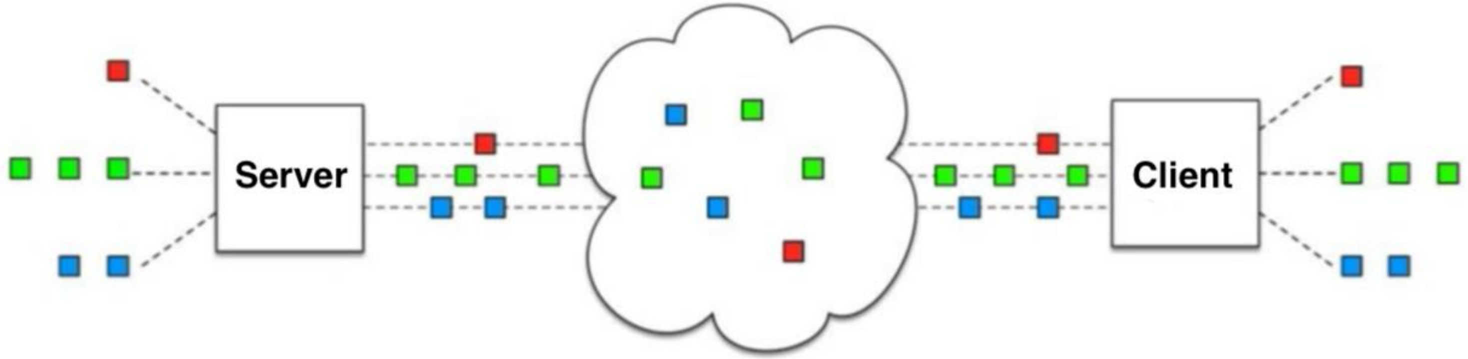
Streaming Protocol	Archi.	Networks	Transport Protocol	NAT Traversal Issue	Prioritized Stream	Multiplexed Stream	Stream Termination	Server Push
MMT	Push based	IP/ broadcast	MMT	-	-	✓	-	✓
RTP	Push based	IP	UDP	✓	-	✓	-	✓
DASH over HTTP/1.1	Pull based	IP	TCP	-	-	-	-	-
DASH over HTTP/2	Pull based	IP	TCP	-	✓	✓	✓	✓
DASH over QUIC	Pull based	IP	UDP	-	✓	✓	✓	✓

Table source: Fan et al. [1]

[1] Ching-Ling Fan, Wen-Chih Lo, Yu-Tung Bai and Cheng-Hsin Hsu. A Survey on 360° Video Streaming: Acquisition, Transmission, and Display. Accepted to appear at ACM Computing Surveys, 2019.

Multiplexed and Prioritized Streams

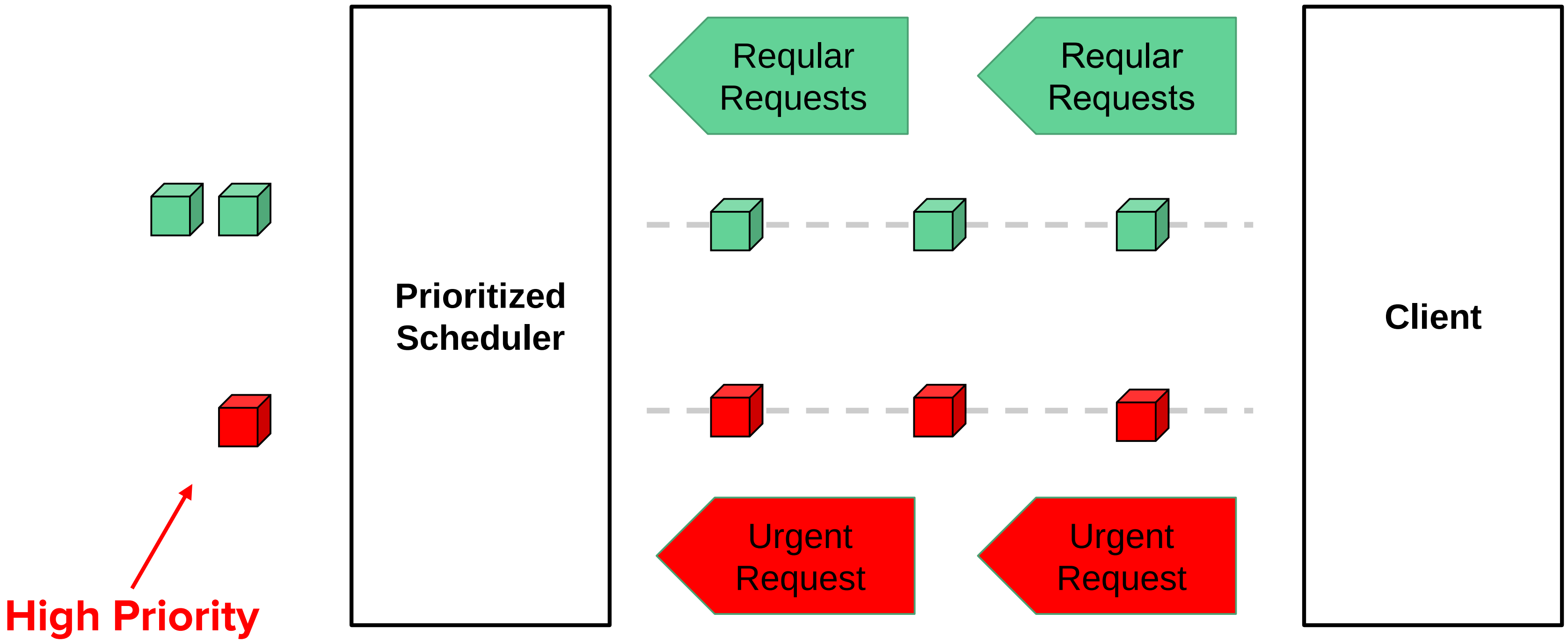
- Multiplexed streams
  - Multiple streams in one connection
- Prioritized streams
  - Each stream has different transmission speed/order
- Prioritized scheduler
  - Strict priority scheduler
    - High priority stream is transmitted first
  - Weighted priority scheduler
    - Allocate bandwidth resource among streams



Strict Priority Scheduler

Regular tiles : low priority

Urgent tiles : high priority



Streaming Protocol Comparisions

Streaming Protocol	Archi.	Networks	Transport Protocol	NAT Traversal Issue	Prioritized Stream	Multiplexed Stream	Stream Termination	Server Push
MMT	Push based	IP/ broadcast	MMT	-	-	-	-	✔
RTP	Push based	IP	UDP	✔	-	-	-	✔
DASH over HTTP/1.1	Pull based	IP	TCP	-	-	-	-	-
DASH over HTTP/2	Pull based	IP	TCP	-	✔	✔	✔	✔
DASH over QUIC	Pull based	IP	UDP	-	✔	✔	✔	✔

Table source: Fan et al. [1]

[1] Ching-Ling Fan, Wen-Chih Lo, Yu-Tung Bai and Cheng-Hsin Hsu. A Survey on 360° Video Streaming: Acquisition, Transmission, and Display. Accepted to appear at ACM Computing Surveys, 2019.

QUIC Protocol

- Created by Google and has been adopted as an IETF standard
- QUIC runs on **UDP** and has three main features
  - **Secured communications**
  - **Multiplexed streams with prioritized schedulers**
  - **Low latency**

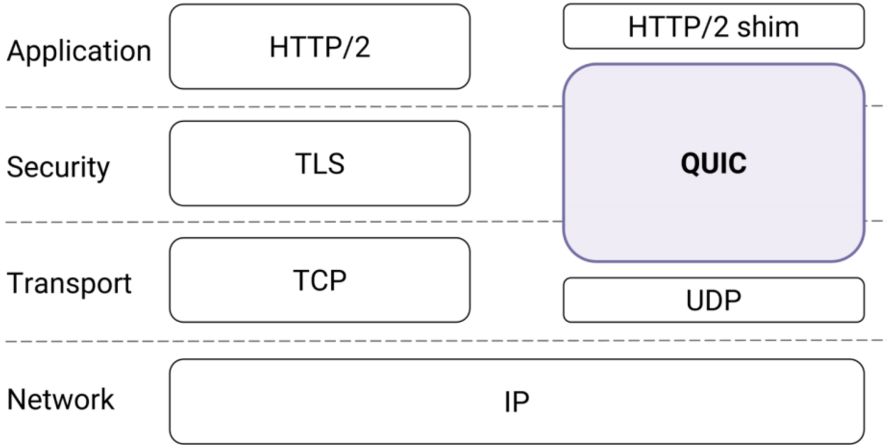


Image source: [1]

[1] Adam Langley, Alistair Riddoch, Alyssa Wilk, Antonio Vicente,. 2017. The QUIC Transport Protocol: Design and Internet-Scale Deployment. In Proc. of the Conference of the ACM Special Interest Group on Data Communi- cation (SIGCOMM'17). Los Angeles, CA, 183–196.

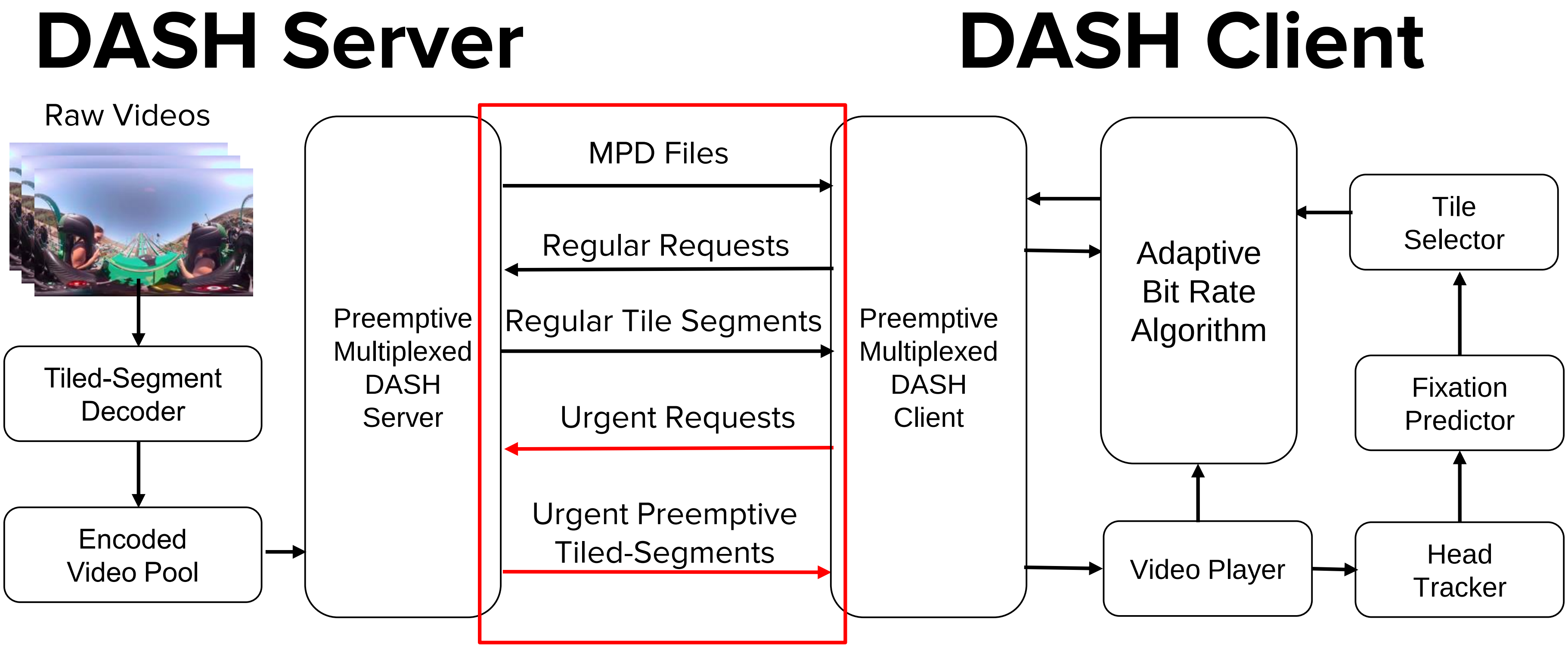
Contributions

- Propose QUIC-based 360° tiled video DASH streaming system
- Design and implement our proposed system on existing open-sourced projects
- Optimize the proposed system by realizing three key components
  - **Fixation predictor**
  - **Tile selector**
  - **ABR Algorithms**

- Evaluate our system through real experiments driven by a public HMD viewer dataset

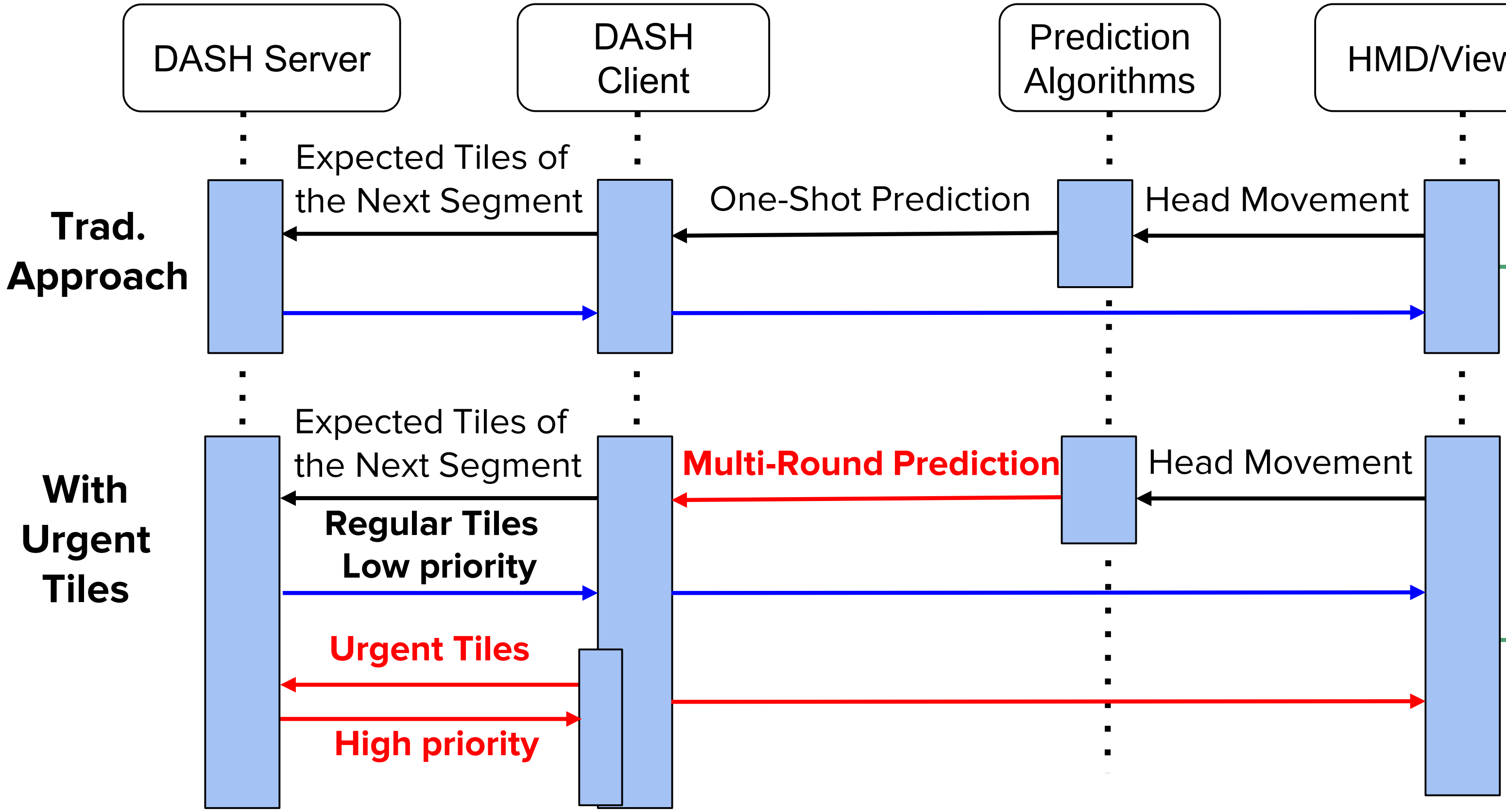
System Architecture

System Overview



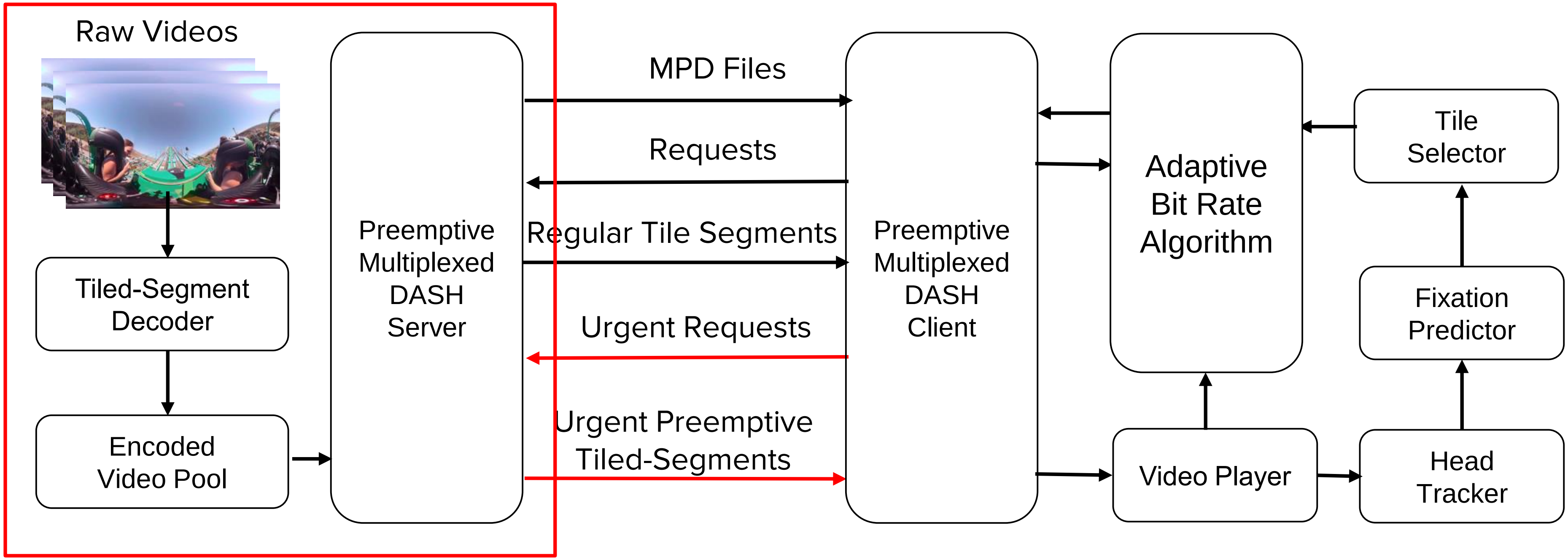
Leverage Urgent Tiles with High Priority

Perform multi-round predictions and preemptively request urgent (missing) tiles



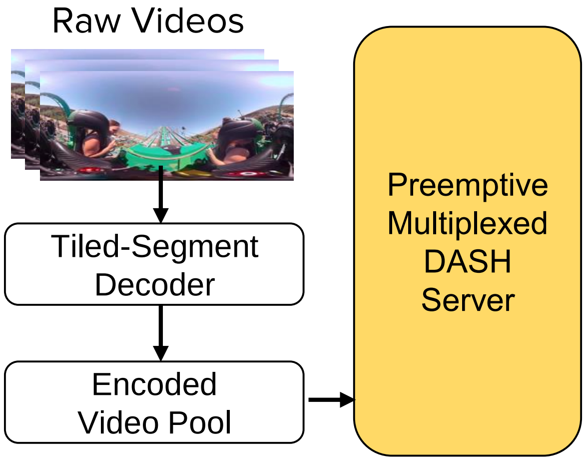
System Overview

DASH Server

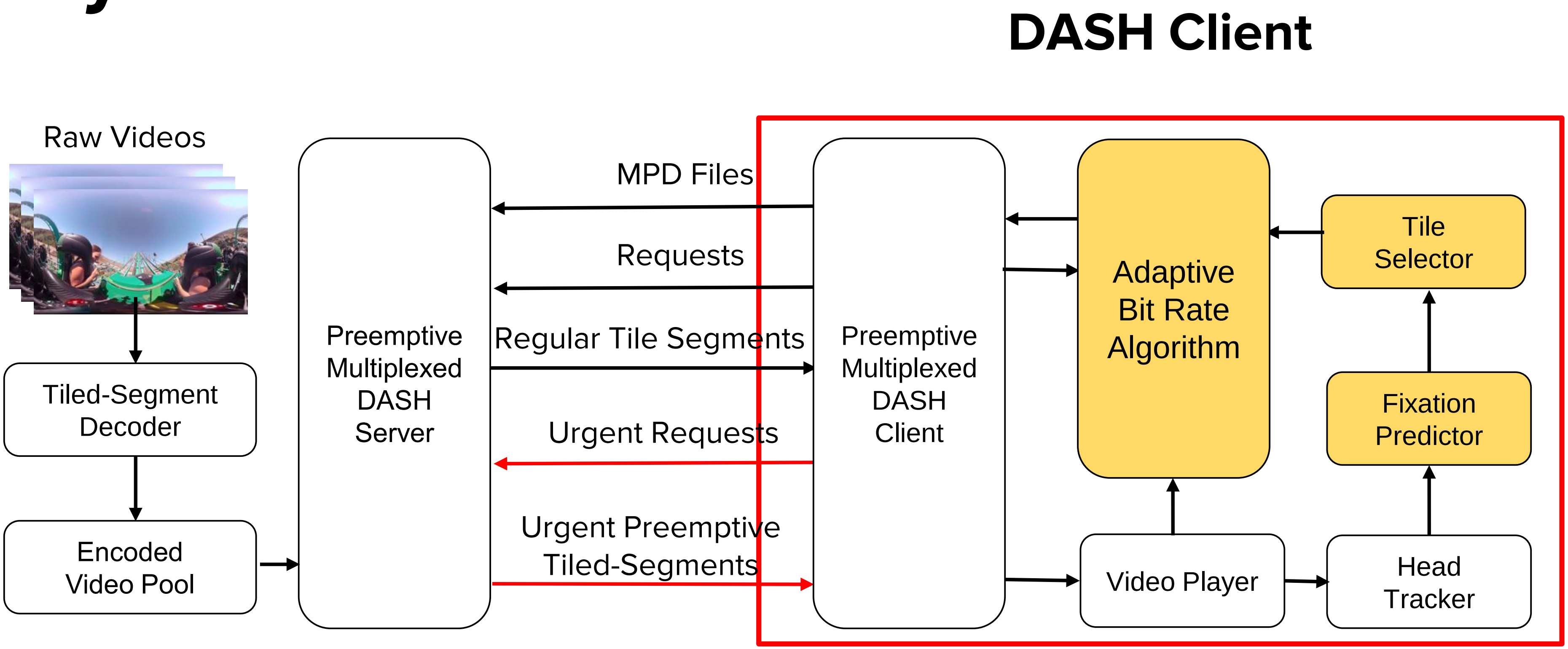


DASH Server

- Preemptive Multiplexed DASH server
  - Supports HTTPS/QUIC
  - Sends prioritized tiled segments over multiplexed streams
  - Adopts strict priority scheduler
    - Regular tile : low priority
    - Urgent tile : high priority



System Overview



Fixation Predictor

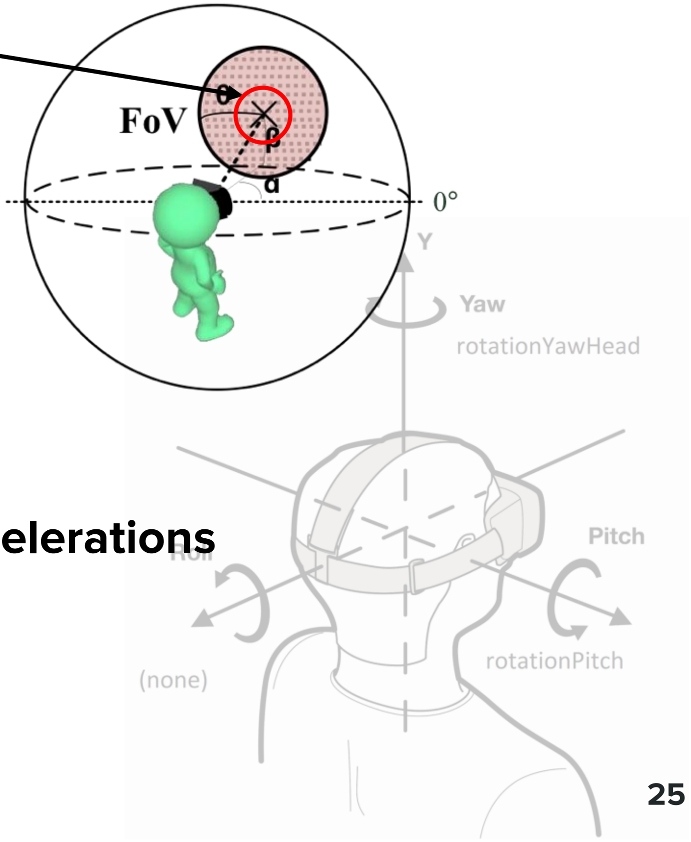
Dead-Reckoning (DR) are employed to predict future viewports

Predicts the viewport center **τ** s later than the current time **t**

$$\hat{\theta}^c_{t+\tau} = \theta^c_t + v^{\theta^c}_t \tau + (1/2) a^{\theta^c}_t \tau^2;$$

$$\hat{\phi}^c_{t+\tau} = \phi^c_t + v^{\phi^c}_t \tau + (1/2) a^{\phi^c}_t \tau^2.$$

viewport center **angular speeds** **angular accelerations**



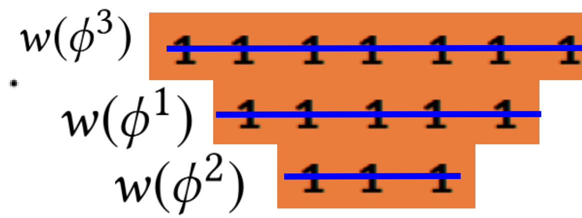
Tile Selector

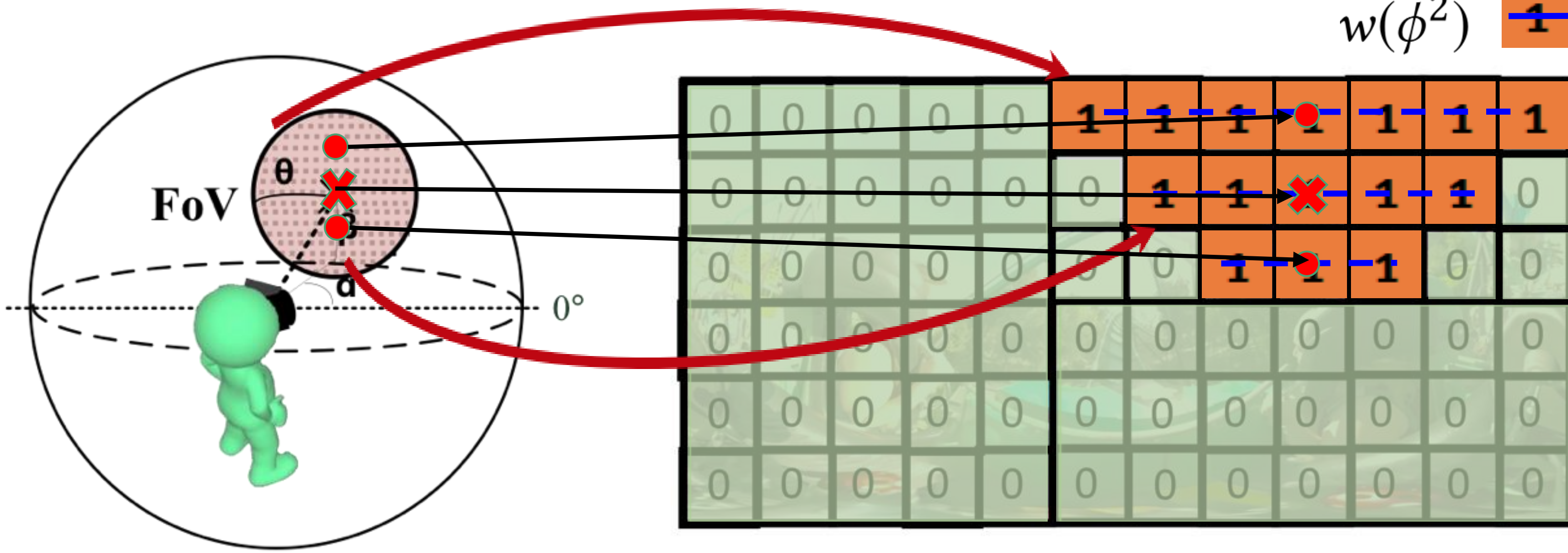
Traditional: 6.32 **s**
Our approach: 0.06 **ms**

Key idea:

approximate the viewport on the video content with the center as (x^c, y^c) **and the width as a function of the pitch value** ϕ^ϵ

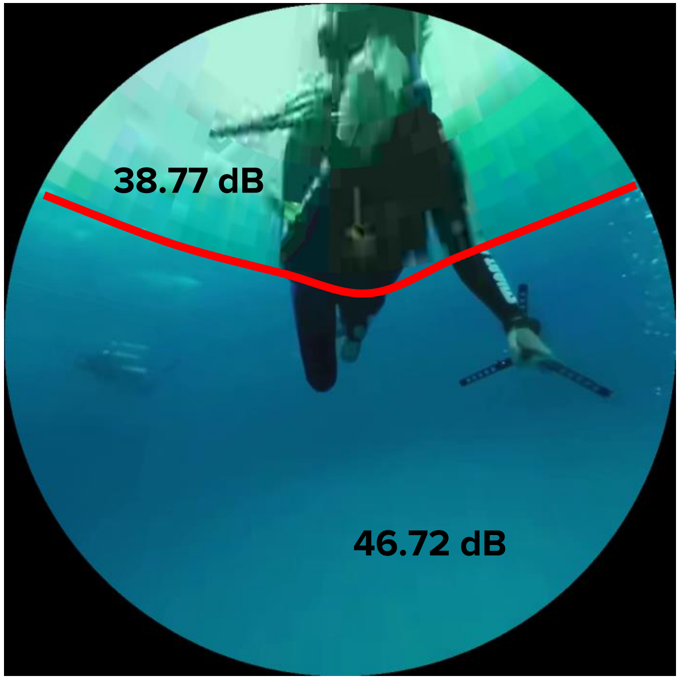
$$w(\phi^\epsilon) = \frac{2r_v\sqrt{1-\cot^2 r_v \cdot \tan^2(\phi^\epsilon - \phi^c)}}{\cos\phi^\epsilon} \cdot \frac{W}{360^\circ}.$$





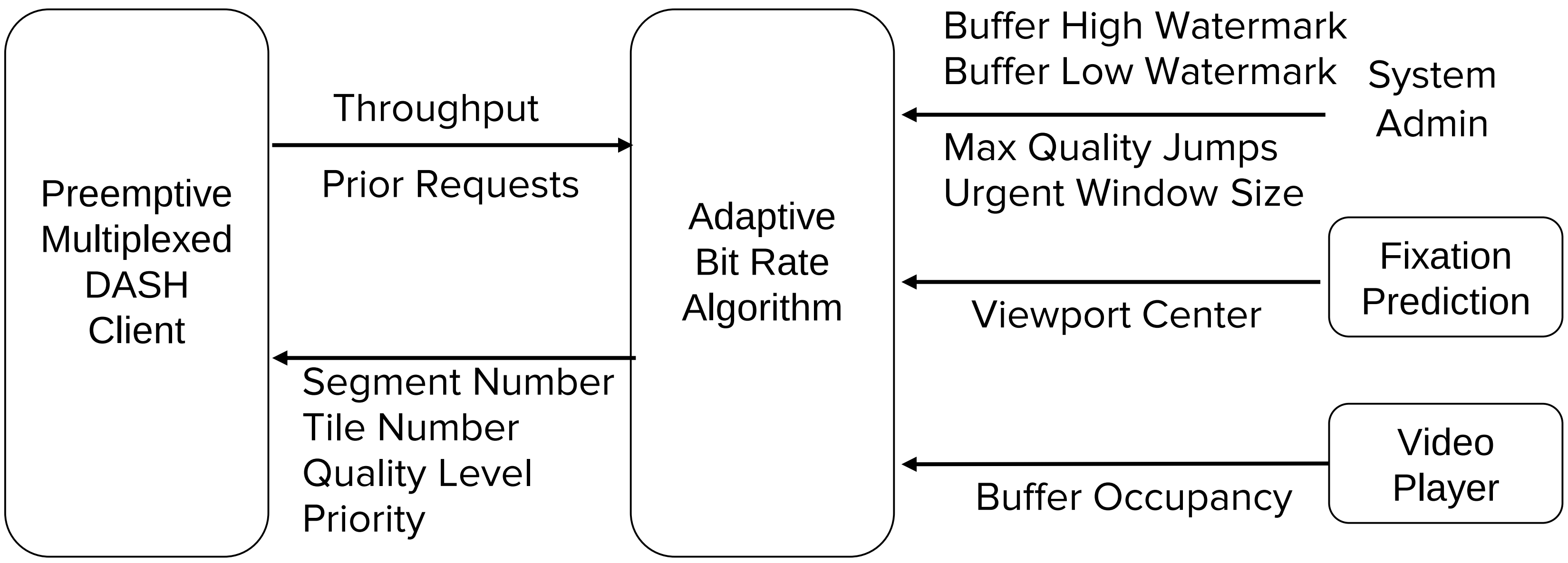
Preemptive Multiplexed Adaptive Bit Rate Algorithms

Design Criteria



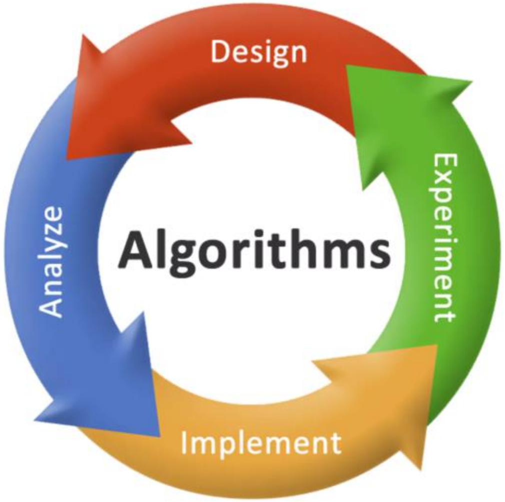
- **Achieve high average video quality**
  - Crucial to the visual quality experience
- **Avoid large quality jumps**
  - Large quality jumps (spatial or temporal) negatively affect the viewing experience
- **Avoid buffer under-run**
  - Buffer under-run leads to playout stalls or black holes

Inputs and Outputs



Designed Algorithms

- Augmented Existing (AE) ABR Algorithm [PV'19]
  - Leverage reliable design
  - Verify the effectiveness of urgent tiles



We proposed a new approach (from scratch) to further optimize the results

- Preemptive Multiplexed (PM) ABR Algorithm [mm'19 under review]
  - An ABR algorithm for a preemptive multiplexed streaming system
  - Enhance overall user viewing experience

1. AE ABR Algorithm

Buffer-Based ABR Algorithm - NETFLIX

Select video rate based on **buffer occupancy**

- Avoid unnecessary rebuffering
  - As long as $C(t) \geq R_{\min}$ adpot $f(B) \to R_{\min}$ as $B \to 0$
- Maximize average video rate
  - The average video rate matches the average capacity when $R_{min} < C(t) < R_{max}$

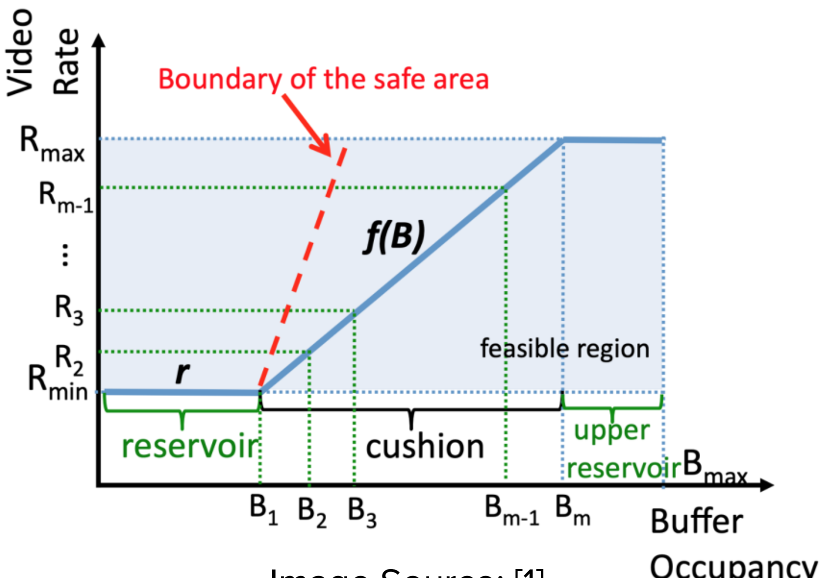
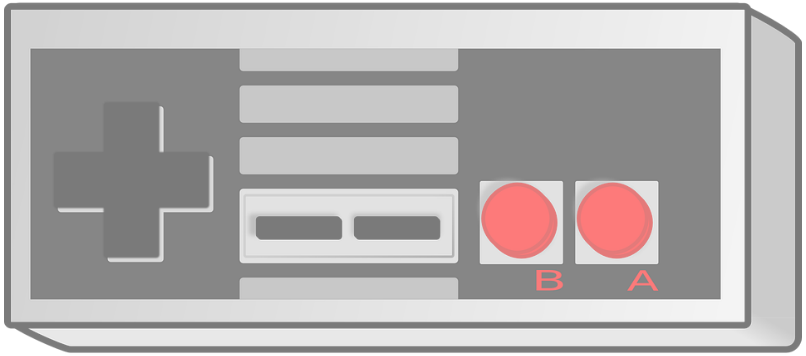


Image Source: [1]

[1] Te-Yuan Huang, Ramesh Johari, Nick McKeown, Matthew Trunnell, and Mark Watson. 2014. A Buffer-based Approach to Rate Adaptation: Evidence from a Large Video Streaming Service. In Proc. of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM'14). Chicago, IL, 187–198.

Proposed AE Algorithm



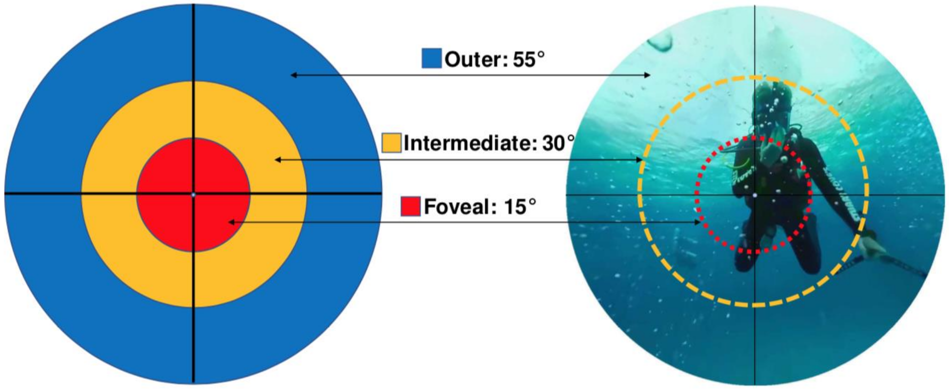
Two paralles and independent flows:

- Regular flow (low priority)
  - Event triggered
  - Maintain buffer occupancy to prevent buffer under-run
  - Take transmission time used by urgent flow into consideration when deciding the suitable quality level

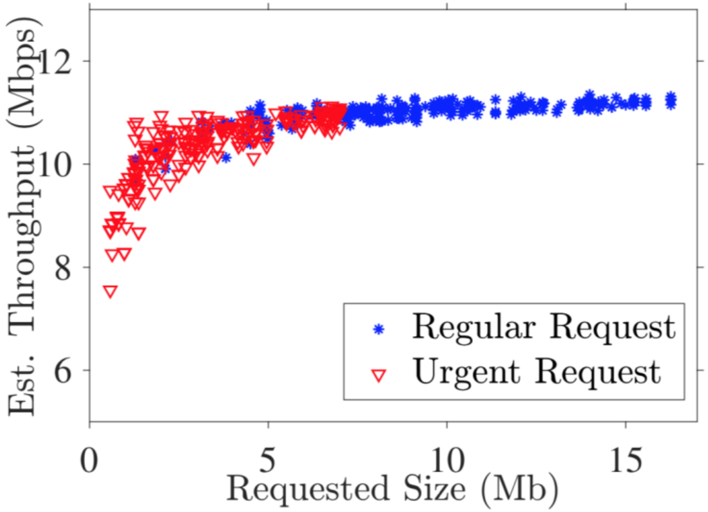
- Urgent flow (high priority)
  - Time triggered (every urgent window)
  - Kicks in when some tiles in the future viewports are bound to be missing
  - Consider tiled segments in urgent window when selecting quality levels

2. PM ABR Algorithm

Design Decisions



- Diverse tile priority levels
  - Foveated streaming
  - Visual acuity decreases as the radius increases
- Parallel and independent flows
  - Regular flows (requests)
  - Urgent flows (requests)
- Preemptive multiplexed streams
  - Ensure timely delivery of urgent tiles
- Estimating the network throughput with regular requests only [1]
  - Need sufficient workload for a reliable throughput measurement



Under **12** Mbps Bandwidth

[1] Cong Wang, Amr Rizk, and Michael Zink. 2016. SQUAD: A Spectrum-based Quality Adaptation for Dynamic Adaptive Streaming over HTTP. In Proc. of International Conference on Multimedia Systems (MMsys'16). Klagenfurt, Austria, 1–112.

PM ABR Algorithm - Regular Flow

Algorithm 1 Regular Flow

Input: T, J, V_c, B_o, i

Output: S_i^r, R_i^r

$S_i^r \leftarrow S(V_c, r_v)$

$T_b \leftarrow T \times (B_o - D - B_t)$

$Q_i \leftarrow \max_{q \in [0:Q_{N-1}+J]} q$ s.t. $\sum_{j \in S_i^r} b_j^q \leq T_b$

$R_i^r = \{(j, Q_i, 4) \, \forall j \in S_i^r\}$

$T_b \leftarrow T_b - \sum_{j \in S_i^r} b_j^{Q_i}$

while $T_b > 0$ **do**

  $r_v \leftarrow r_v + 5°$

  $S_t \leftarrow S(V_c, r_v) \setminus S_i^r$

  $Q^t \leftarrow \max_{q \in [Q_i - J:Q_i]} q$ s.t. $\sum_{j \in S_i^r} b_j^q \leq T_b$

  if Q^t exits **then**

    $S_i^r \leftarrow S_i^r \cup S_t$

    $R_i^r = R_i^r \cup \{(j, Q^t, 4) \, \forall j \in S_t\}$

    $T_b \leftarrow T_b - \sum_{j \in S_i^r} b_j^{Q^t}$

  else

    break

  end if

end while

Prevent high temporal/spatial quality jumps

1. Obtain the tile set overlapped with the viewport
2. Decide suitable quality levels considering throughput and buffer occupancy
3. Assign quality levels and stream priorities to the tiles
4. Augment viewport sizes to accommodate more tiles to prevent missing tiles

PM ABR Algorithm - Urgent Flow

Algorithm 2 Urgent Flow

Input: $T, i, J, V_c, Q_i, U, S_i^r$
Output: S_i^u, R_i^u
$F(i)$, tile sets of foveal (0), intermediate (1), and outer (2)
$foveal = 15°$, $intermediate = 30°$, and $outer = 55°$

$F(0) \leftarrow S(V_c, foveal) \setminus S_i^r$
$F(1) \leftarrow S(V_c, intermediate) \setminus (S_i^r \cup F(0))$
$F(2) \leftarrow S(V_c, outer) \setminus (S_i^r \cup F(0) \cup F(1))$
$S_i^u \leftarrow F(0) \cup F(1) \cup F(2)$
$R_i^u \leftarrow \{(j, Q_i - J, n + 1) \, \forall j \in F(n) \text{ and } n \in \{0, 1, 2\}\}$
$T_b \leftarrow T \times U - \sum_{j \in S_i^u} b_j^{Q_i - J}$
for n in $\{0, 1, 2\}$ **do**
    $Q_t \leftarrow \max_{q \in [Q_i - J : Q_i]} q$ s.t. $\sum_{j \in F(n)} b_j^q \leq T_b$
    if Q^t exits **then**
        $T_b \leftarrow T_b - \sum_{j \in F(n)} b_j^{Q^t}$
        Update $R_i^u \leftarrow \{(j, Q_t, n + 1) \, \forall j \in F(n) \text{ and } n \in \{0, 1, 2\}\}$
    else
        break
    end if
end for

Limit the quality jump

1. Obtain overlapped tile set of each area
2. Assign tiles with initial quality levels and priorities
3. Upgrade quality level from foveal to outer areas

Evaluations

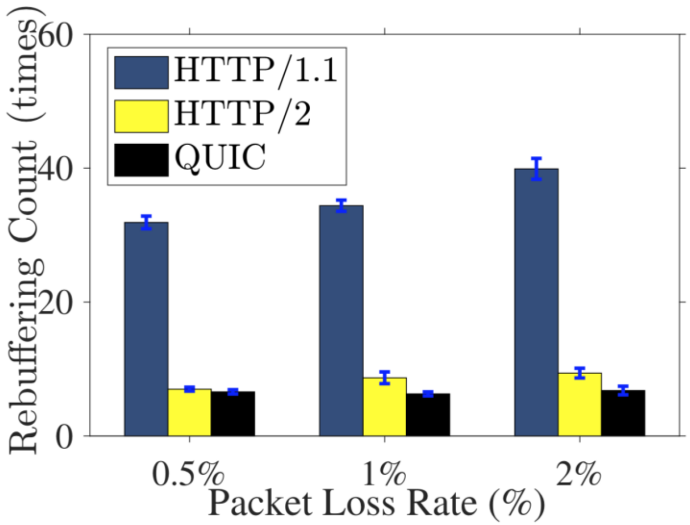
Experiment Setup

- Environment
  - DASH server, Intel i7 CPU desktop with 12 GB RAM
  - DASH client, Intel i7 CPU desktop with 16 GB RAM
- Tiling/DASH
  - No. tiles = {10x10}
  - DASH segment length = {1} sec
  - Video bitrate = {15, 14, 12, 11, 10, 9, 6, 8, 5} Mbps
  - FoV size = {55} degree radius (ground truth = 50 degree radius)
- Viewers
  - Randomly select 10 user and 10 video traces from our dataset
- Baselines
  - Buffer-based ABR algorithm (NETFLIX)
  - Petrangeli et al. [1] (PSHD17)

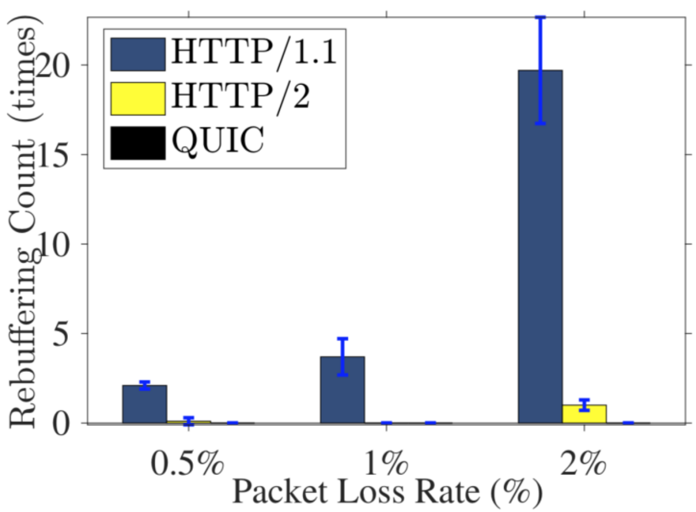
[1] Stefano Petrangeli, Viswanathan Swaminathan, Mohammad Hosseini, and Filip De Turck. 2017. An HTTP/2-Based Adaptive Streaming Framework for 360° Virtual Reality Videos. In Proc. of ACM International Conference on Multimedia (MM'17). Mountain View, California, 306–314.

Benefit from QUIC Protocol: Less Rebuffering

- Request the tiles overlapping with the ground-truth viewports at the highest quality level
- QUIC results in lower rebuffering counts (time)
- HTTP/2 is more sensitive to packet loss



Under **5** Mbps Bandwidth



Under **8** Mbps Bandwidth

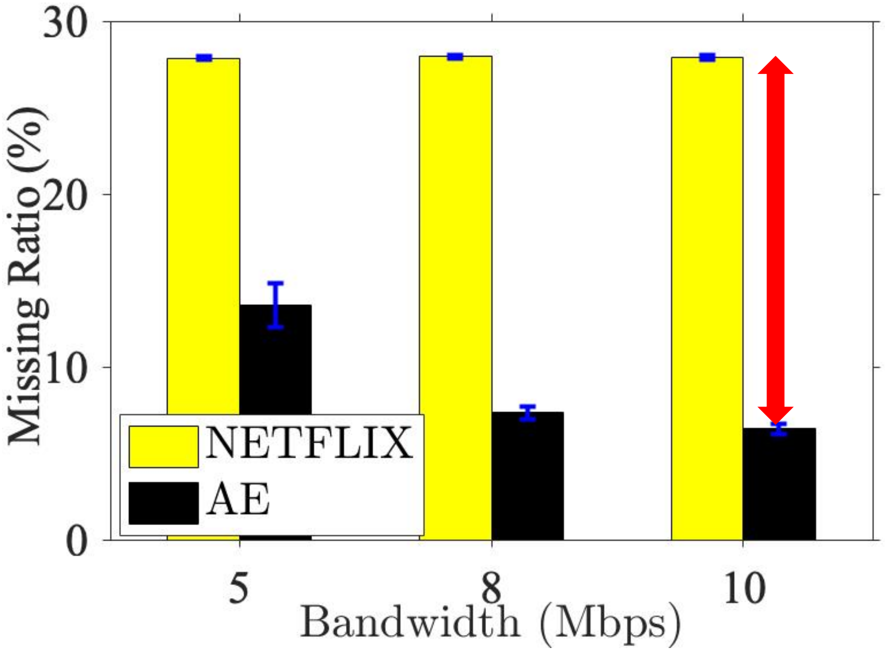
Illustrations of Rebuffering Events

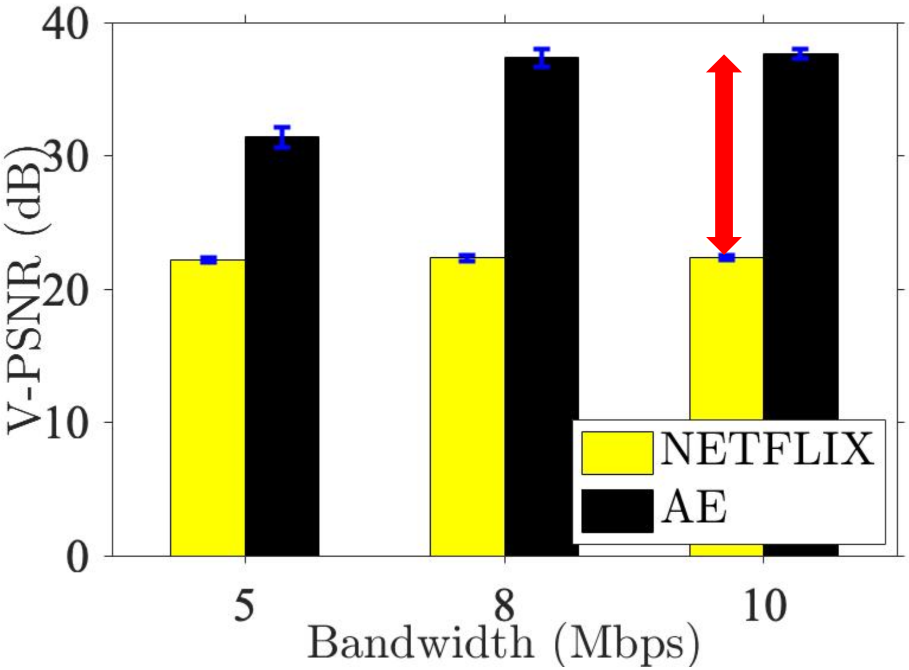
Ideal
QUIC

Rebuffering
HTTP

Effectiveness of Urgent Tile Streams

- Streaming performance with (AE) and without (NETFLIX) urgent tile streams from 1 user

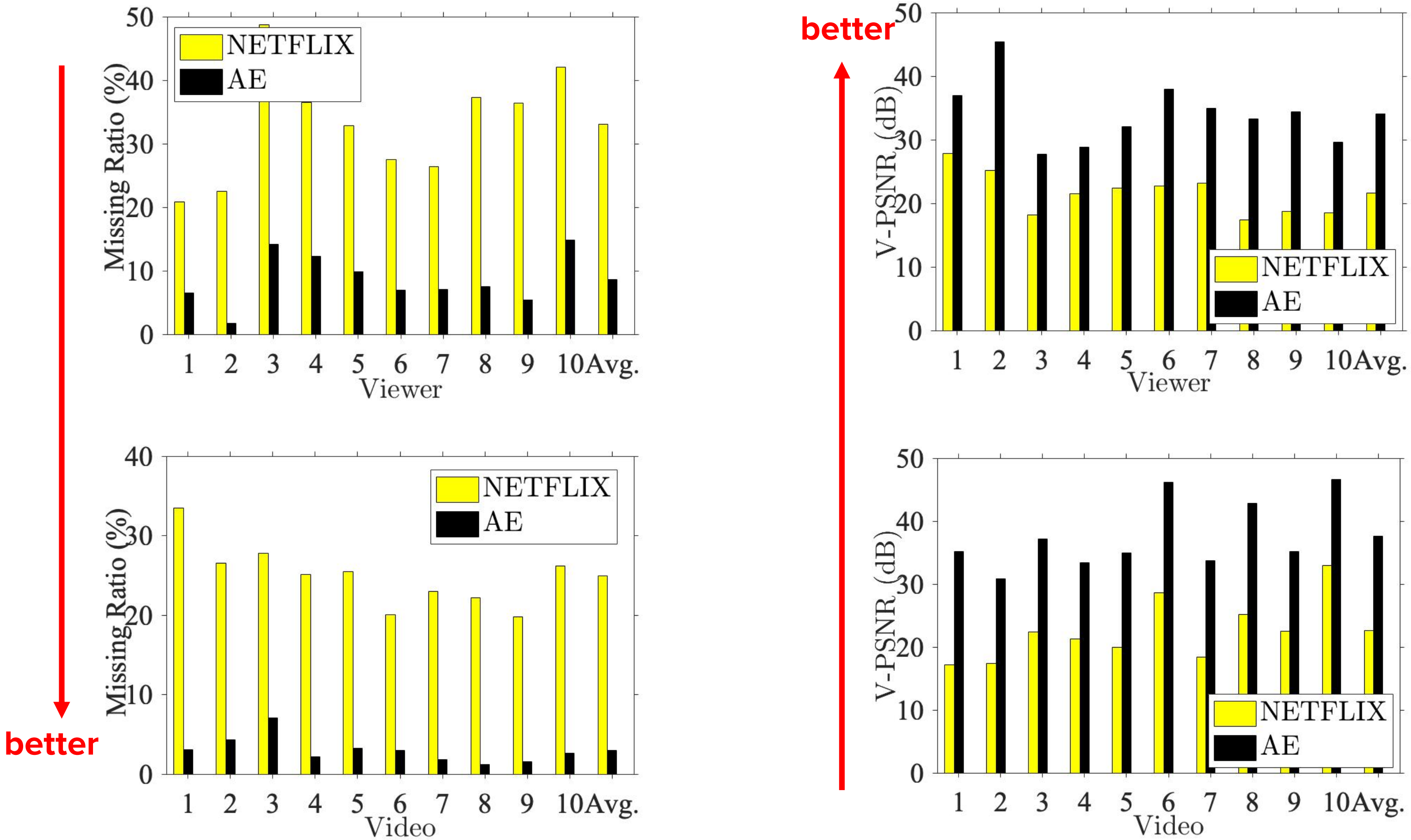




Up to 21.5% missing ratio reduction and 15.28dB video quality improvement

Effectiveness of Urgent Tile Streams

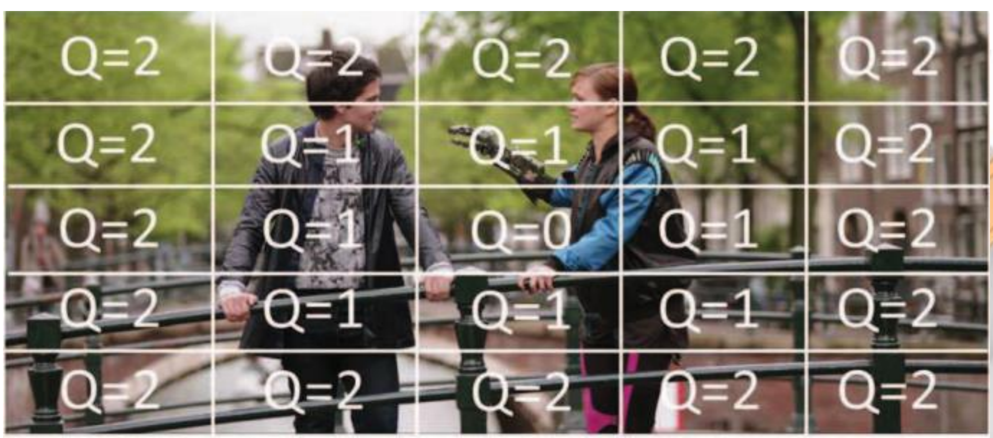
- Urgent tile streams are effective for various video types and diverse viewers



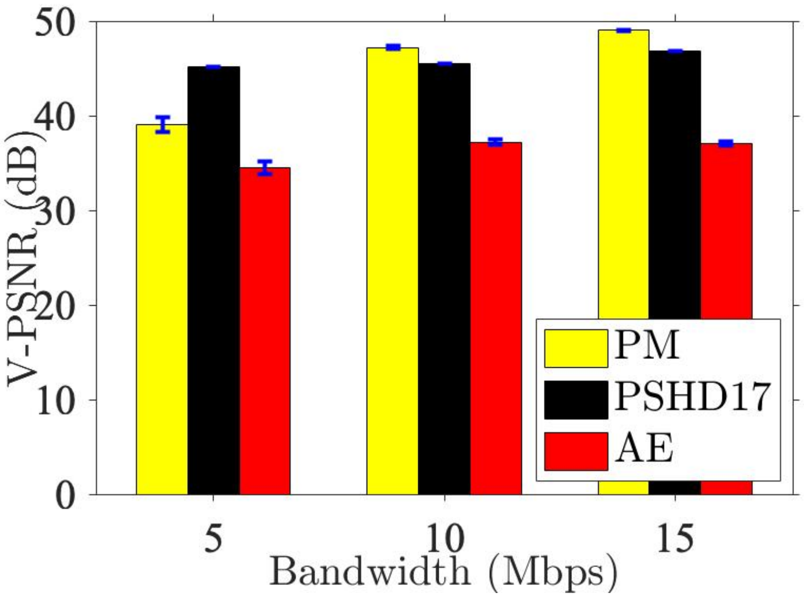
Observation from the Results

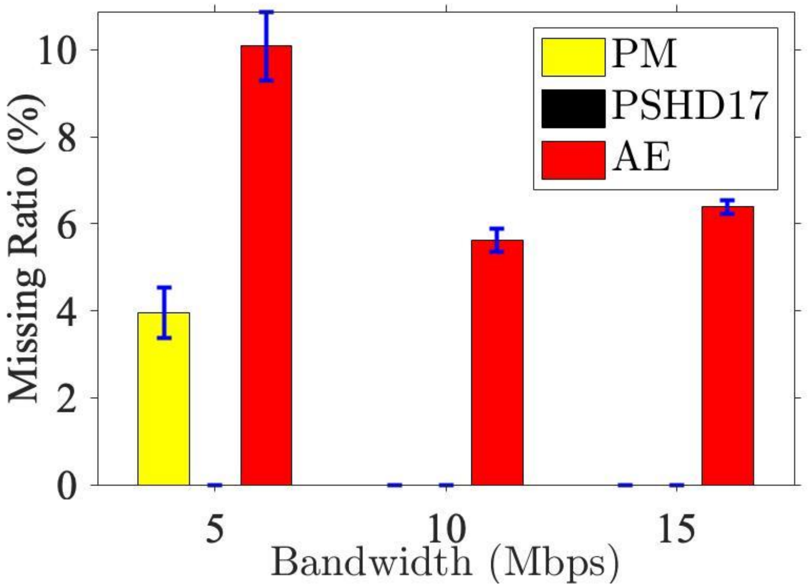
- Urgent tile streams successfully reduce the missing ratio and enhance the video quality
- Missing ratio is about 10% (21.5% reduction)
  V-PSNR is about 39 dB (15.28 dB improvment)

We next compare PM, PSHD17, and AE



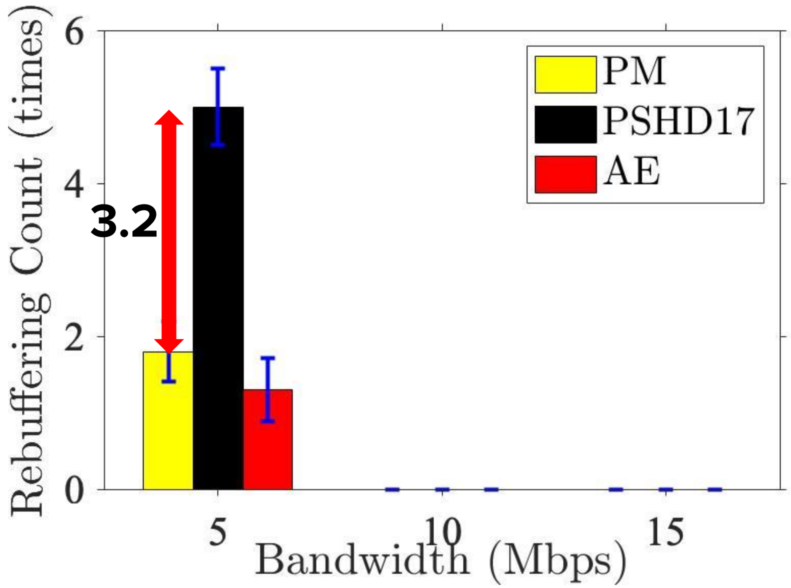
PM Improves the Video Quality

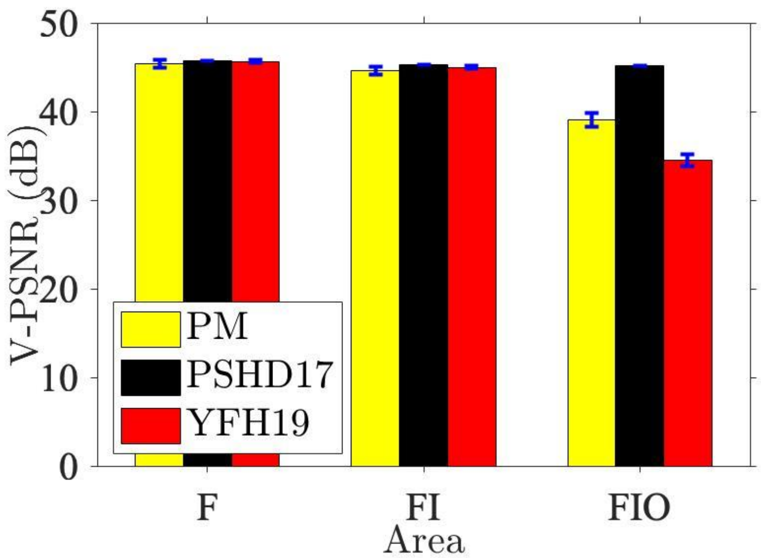




- PM algorithm leads to good video quality unless the network bandwidth is highly constrained
- Although PSHD17 leads to higher V-PSNR, but...

Efficiently Allocate the Available Bandwidth Around Viewport Center



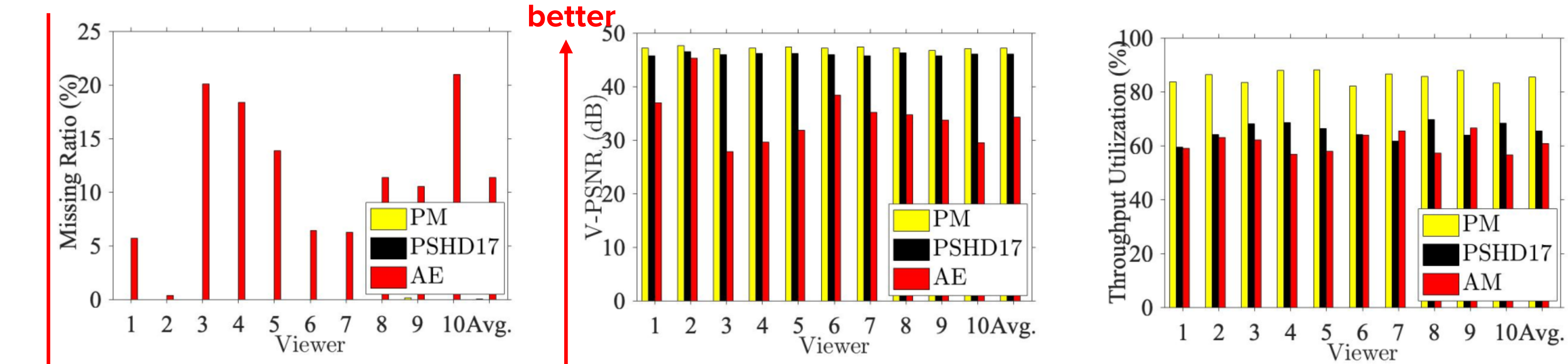


Under **5** Mbps Bandwidth

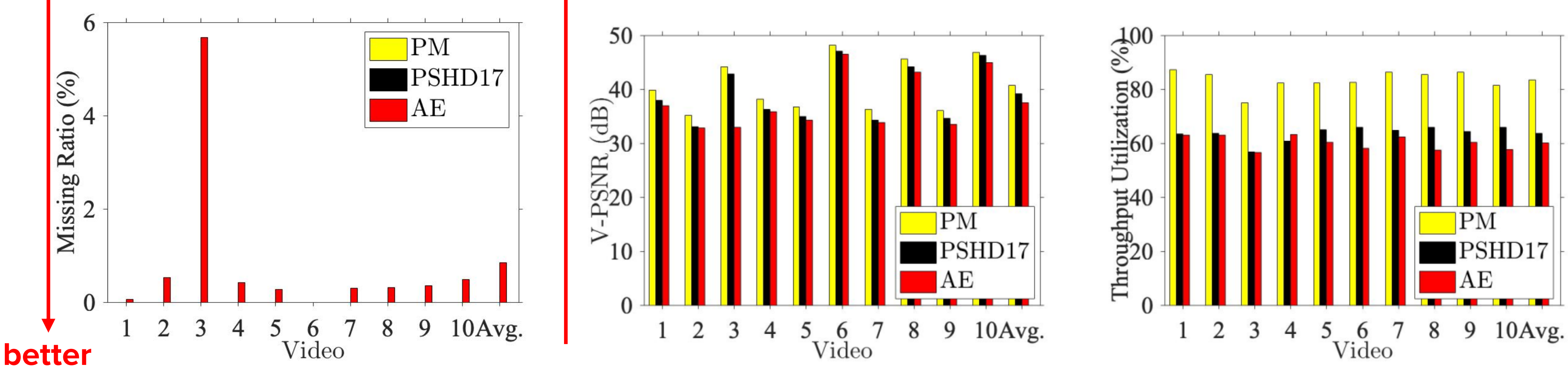
- PM has lower rebuffering events than PSHD17
- The available bandwidth is mostly distributed to the more important areas

Adapt to Various Videos and Viewers

10 Viewers



10 Videos



Under **10** Mbps Network Bandwidth

Conclusion

Conclusion

- Proposed QUIC-based 360° tiled video DASH streaming system
- Modified an existing ABR algorithm and designed an ABR algorithm for 360° tiled videos leveraging preemptive multiplexed streams
- Designed and implemented our proposed system for evaluations

Compared to the state-of-the-art algorithms (NETFLIX and PSHD17) algorithm, our proposed algorithms:

- Reduces the re-buffering counts by up to 3.2 and rebuffering time by up to 2.54 s
- Achieves higher bandwidth utilization at most 40.02%
- Delivers good average V-PSNR at 39–49 dB under 5–15 Mbps bandwidth

Limitation and Future Directions

- The adaptation algorithms for the system parameters
  - Crucial for deploying our solution in live networks
- The performance comparisons among protocols
  - Diverse priority schedulers in QUIC and HTTP/2
- Multipath-QUIC
  - Enables hosts to exchange data over multiple networks over a single connection

Thanks for listening