QoE Evaluation of Real-Time Video Streaming using Media over QUIC with Bicast in Wireless Multihomed Networks

Yamato Kashihara*, Tomoya Kawana†, Rei Nakagawa‡, Nariyoshi Yamai§
Department of Electrical Engineering and Computer Science
Tokyo University of Agriculture and Technology
Tokyo, Japan
Email: *s245216y@st.go.tuat.ac.jp, †kawana@net.cs.tuat.ac.jp, ‡rnakagawa@go.tuat.ac.jp, §nyamai@cc.tuat.ac.jp

***Abstract*—In recent years, real-time video streaming applications such as video conferencing have rapidly increased. However, in wireless networks such as Wi-Fi, packet loss is more likely to occur due to radio interference, and the transmission rate tends to decrease compared to wired networks. As a result, in real-time video streaming, retransmission of lost packets causes latency, and if retransmission cannot be completed in time, the video quality significantly degrades. To address this issue, this paper applies the bicast communication architecture over QUIC, proposed in prior work, to Media over QUIC (MoQ) and implements bicast communication using MoQ in a wireless multihomed networks. We propose a method that mitigates video quality degradation by performing bicast communication of video data from a Relay server deployed at the network edge close to the Subscriber. Furthermore, we evaluate the proposed network architecture using the Mean Opinion Score (MOS) based on the Video Call MOS (VCM). As a result, bicasting improved the MOS by approximately 10% and increased the proportion of frames achieving SSIM above 0.97 by approximately 40%.**

***Index Terms*—Media over QUIC, Multihomed Network, Bicast Communication, Quality of Experience.**

I. Introduction

In recent years, the demand for real-time video streaming services such as interactive online broadcasting including video conferencing, social networking services (SNS), e-sports, and live streaming of video games has been steadily increasing. In particular, for video conferencing, Quality of Experience (QoE) is highly sensitive to network conditions, making low latency and high throughput a critical requirement [1]. However, wireless networks such as Wi-Fi are more susceptible to packet loss due to radio interference compared to wired networks, resulting in reduced throughput [2]. Furthermore, wireless networks generally experience higher latency than wired networks. As a result, in real-time conferencing, the additional playback latency, which is the application delay before a video or audio frame is actually played back, incurred by requesting and receiving retransmissions of lost data, tends to be large. If retransmission cannot be completed within the acceptable playback latency, the playback of video or audio may be interrupted, significantly degrading QoE.

Recently, Media over QUIC (MoQ) [3] has attracted attention as a latest technology for achieving low-latency and high-quality video streaming. MoQ is built on QUIC, a transport layer protocol that is connection-oriented similar to TCP. In addition, since QUIC is implemented in software on UDP, it allows kernel independent control. QUIC also supports multiple streams within a single connection, and this multiplexing capability enables MoQ to avoid delays due to structural problems of traditional TCP, such as Head of Line Blocking. MoQ is implemented as a single protocol on QUIC, which operates in software, making it highly extensible. Moreover, MoQ provides a semantic interpretation of media content as Object, allowing flexible control. MoQ consists of Publisher, Subscriber, and Relay. The Publisher is responsible for content distribution, the Subscriber for content reception, and the Relay for content forwarding. A key feature of the Relay is its ability to cache data. Furthermore, the underlying transport protocol of MoQ, QUIC, supports multihoming through advancements such as Multipath QUIC [4]. By leveraging this capability, the Relay can flexibly utilize multiple transmission paths, which is expected to enhance QoE[1].

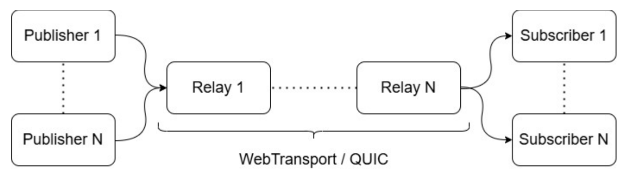


Fig. 1: MoQ network overview diagram

Our previous research proposed bicast communication in wireless multihomed networks. A multihomed network is one in which a device is simultaneously connected to multiple networks and communicates over multiple paths. To stabilize communication utilizing this network, bicast communication can be employed. Bicast communication is a method of simultaneously transmitting the same packet over multiple paths in a multihomed network. Previous studies have shown that bicast communication in wireless multihomed environments enables more stable communication compared to unicast communication [5].

[1]WebRTC is another technology for low-latency video delivery. However, its limited flexibility restricts in-network and multipath control.

TABLE I: Group table of OpenFlow

Group Table	Type	Action Bucket
1	all	send to Subscriber via port 1
		send to Subscriber via port 2

The objective of this study is to realize multi-path communication control at the Relay server. In this paper, as a preliminary evaluation toward this ultimate goal, all packets are transmitted using bicasting to verify the effectiveness of the proposed approach. The connection between the Subscriber and the Relay is assumed to operate in a wireless environment. Therefore, by performing bicasting of video data from the Relay server, it is expected that the degradation of QoE can be mitigated even under poor network conditions.

The contributions of this paper are as follows:

- We address the degradation of real-time video quality under poor network conditions by combining MoQ with communication multiplexing.
- The proposed system can be easily implemented without requiring modifications to the OS by leveraging the software based OVS and MoQ.
- The use of a multihomed network resulted in improved MOS values under adverse wireless conditions with high packet loss rates and high latency.

II. Related Technology

A. Media over QUIC

In recent years, QUIC has attracted attention as a promising technology for achieving low-latency communication. QUIC is a connection-oriented protocol, similar to TCP. In addition, QUIC is implemented in software on UDP, allowing flexible userland control that does not depend on the kernel. QUIC also supports multiplexing, enabling multiple data streams to be handled within a single connection. Building on QUIC, Media over QUIC (MoQ) [3] has emerged as a state-of-the-art technology for delivering low-latency, high-quality video. MoQ leverages QUIC ' s flexibility and multiplexed streams to provide a highly adaptable communication protocol suitable for real-time video streaming where both low latency and high quality are required. As illustrated in Figure 1, MoQ adopts a Pub/Sub model. In browser environments, media is transmitted via WebTransport, while in other environments, QUIC is used as the transport protocol. MoQ consists of three main components: Publisher, Subscriber, and Relay. The Publisher is the endpoint that sends media, while the Subscriber is the endpoint that receives it. The Relay simply forwards media to other Relays or Subscribers without modifying or reencoding the data. In MoQ, data is packaged into the smallest unit called an Object, a collection of Objects forms a Group, and a collection of Groups forms a Track. MoQ assigns names to each Object, enabling semantic interpretation of media content. For example, in video streaming, Objects can be named to represent key or delta frames, allowing control based on frame information. A key feature of MoQ is that such content-based control can be realized in-network through the Relay.

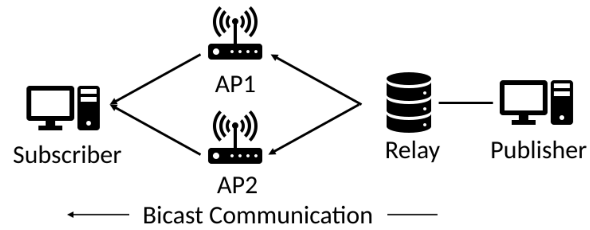


Fig. 2: Proposed network architecture

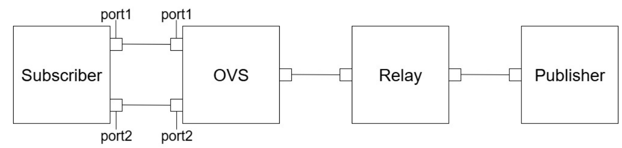


Fig. 3: Implemented network architecture

III. Proposed Architecture

A. Proposed Network Architecture

The objective of this study, as illustrated in Figure 2, is to realize in-network flexible multi-path communication control at the Relay server. In this paper, as a preliminary evaluation toward this ultimate goal, all packets are transmitted using bicasting to verify the effectiveness of the proposed approach. All communications between the Subscriber and the Publisher are performed over MoQ. Furthermore, a multihomed network with two wireless communication paths is constructed between the Subscriber and the Relay server. Packets transmitted from the Relay server are duplicated and sent over both wireless paths. This configuration is expected to suppress degradation in real-time video quality.

Section B explains the structure of the multihomed network and the implementation method used in this study.

B. Bicasting Communication in Multihomed Network

A multihomed network is a network that connects to multiple networks simultaneously and utilizes multiple paths for communication. A multihomied network can maintain continuous communication by switching to alternative networks when one path becomes unavailable. Therefore, multihoming improves the overall availability of the network. Furthermore, bicast communication is a method used to stabilize communication in multihomed networks. Bicast communication involves transmitting the same packet simultaneously over multiple paths in a multihomed network. This approach reduces the packet loss rate and prevents degradation of transmission speed. In previous work [5], a multihomed network was constructed using QUIC and Software Defined Networking (SDN) with OpenFlow, and the download time was compared between unicast and bicast under varying packet loss rates. The results demonstrated that bicast communication achieves more stable transmission than unicast, even in unstable network environments.

To implement the multihomed network, SDN with OpenFlow is utilized. OpenFlow adopts an architecture that separates the control and forwarding functions of network devices [6]. The control function is executed by the OpenFlow controller, which manages routing and flow control. The forwarding function is executed by the OpenFlow switch, which processes received packets based on the flow table. However, the flow table can define actions for only a single port. By using the group table feature, it becomes possible to define actions for multiple ports simultaneously. By utilizing the group table feature, bicast communication is implemented by sending the same packet through multiple ports. Table I shows the group table used for bicast communication.

Based on the above, the implemented network is shown in Figure 3. Open vSwitch (OVS) is employed as the OpenFlow switch. OVS is a software based, OpenFlow compliant virtual switch. By configuring OpenFlow as shown in Table I, packets from the Relay can be duplicated and sent from OVS port 1 and port 2 to Subscriber ports 1 and 2, thereby enabling bicast communication.

IV. EVALUATION

A. Test Method

In this experiment, virtual machines functioning as the Publisher, Relay, OVS, and Subscriber were constructed using VMware Workstation [7], and network connections were configured to implement the network architecture shown in Figure 3. For the implementation of MoQ, we employed moq [8], which is implemented in the Rust programming language. Using this network, we compared the Mean Opinion Score (MOS) between unicast and bicast communication. In the experiment, the link between the Subscriber and the Relay was assumed to be a wireless network. Accordingly, the packet loss rate on this link was varied among 0%, 3%, 5%, and 10%, and the round-trip time (RTT) was varied among 0 ms, 25 ms, 50 ms, 100 ms, and 200 ms. To measure the MOS of the video received at the Subscriber from the Publisher under each network condition, we employed the Video Call MOS (VCM) model [9].

VCM is a model that estimates the MOS by comparing a degraded video with a reference video. A key feature of VCM is that it considers temporal distortions, such as frame freezes and frame skips. In addition to comparing the image quality of each frame between the degraded and reference videos, VCM uses markers embedded in the video to detect temporal distortions, which are then employed as features. This approach allows the calculation of MOS that accounts for the perceptual impact of temporal distortions specific to video conferencing based on quantitative metrics.

In this study, we used ref_01, one of the reference videos included in the VCM dataset, and conducted 10 experiments each for unicast and bicast communications, calculating the average MOS for each case.

B. VCM metorics

The metrics used for estimating MOS with VCM are described below.

Additive Distortion Measurement (ADM) [10] separates image degradation into detail loss and additive distortion components, making it a metric that more closely correlates with subjective evaluation than conventional metrics.

TABLE II: MOS measurement results

	Bicast				Unciast			
	0%	3%	5%	10%	0%	3%	5%	10%
0ms	4.30	4.29	4.29	4.29	4.29	4.26	4.29	4.22
25ms	4.29	4.29	4.27	4.25	4.29	4.22	4.22	4.19
50ms	4.30	4.26	4.23	4.19	4.29	4.19	4.17	4.15
100ms	4.31	4.19	4.15	4.12	4.30	4.14	4.09	4.06
200ms	4.27	4.11	4.05	4.06	4.24	4.00	3.95	3.76

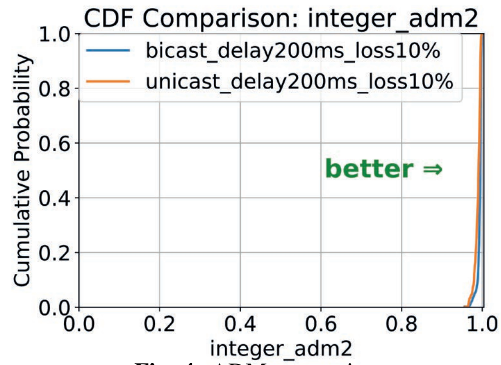


Fig. 4: ADM comparison

Motion [11] is a temporal metric that measures video motion based on the mean absolute pixel difference of the luminance component between adjacent frames. Here, integer_motion represents the comparison between the current frame and the previous frame, while integer_motion2 compares the current frame with the previous two frames and selects the smaller motion. Motion is used as a weighting factor in VMAF scores and is therefore not a degradation metric by itself.

Visual Information Fidelity (VIF) [12] evaluates how faithfully visual information is preserved between a reference image and a distorted image. Scale 0 corresponds to the original resolution, and scale 1 and above correspond to progressively downsampled images.

Structural Similarity Index Measure (SSIM) [13] assesses the similarity of luminance, contrast, and structure between a reference image and a distorted image.

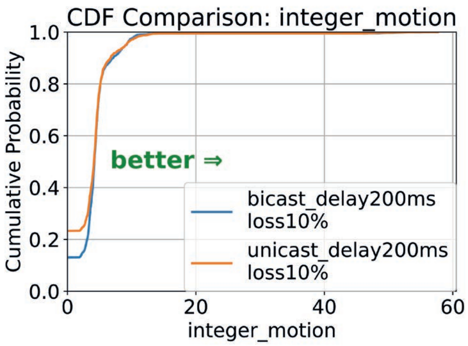
Peak Signal-to-Noise Ratio (PSNR) quantifies the ratio of maximum signal power to noise between a reference image and a distorted image in decibels. PSNR can be decomposed into luminance and chrominance components, where PSNR_y represents the luminance component, PSNR_cr the red chrominance, and PSNR_cb the blue chrominance.

Video Multi-Method Assessment Fusion (VMAF) [14] is a metric developed by Netflix to predict subjective video quality. VMAF combines multiple image quality metrics such as VIF, ADM, and Motion to compute a composite score.

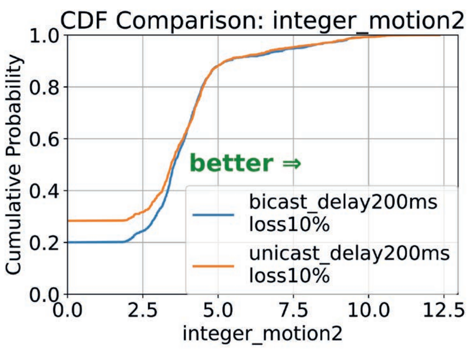
Frame Freeze indicates the number of consecutive frames that remain static before the display changes to the next reference frame.

Frame Skip represents the number of frames skipped when transitioning from one reference frame to the next.

VCM model estimates the MOS by inputting these metrics as features into a Long Short-Term Memory network.



(a) integer motion comparison



(b) integer motion2 comparison

Fig. 5: Motion comparison

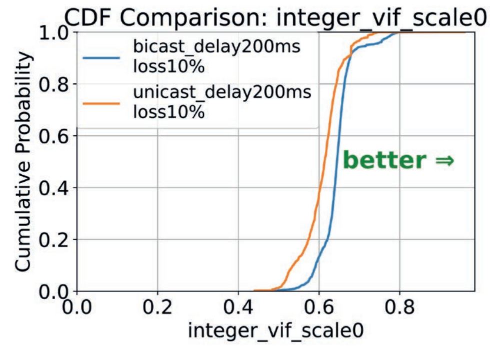
C. Result

Table II presents the measured MOS results. From Table II, it can be observed that bicast communication consistently achieves higher MOS values than unicast communication under all conditions. In particular, under the condition of a RTT of 200 ms and a Packet Loss Rate of 10%, a difference of approximately 0.3 was observed in the MOS between the two methods, demonstrating an improvement of approximately 8%. These results indicate that the advantage of bicast becomes more pronounced as network conditions deteriorate.

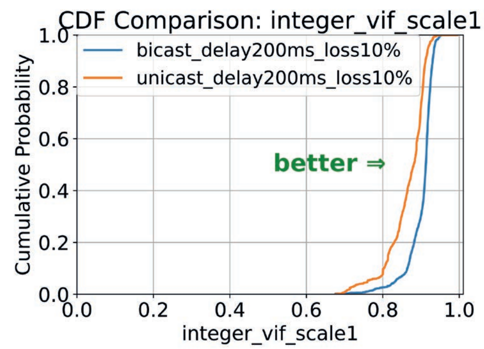
Next, we provide a detailed analysis of the results under the most pronounced condition in this experiment: an RTT of 200 ms and a packet loss rate of 10%. Under this condition, the cumulative distribution functions (CDF) of the individual metrics that constitute the VCM were examined for both unicast and bicast communication, and the results were compared.

First, we compared the CDF of ADM. Bicasting method shows improvements for ADM in all periods, as compared to unicasting method. As shown in Figure 4, the proportion of frames with ADM values exceeding 0.99 was found to be approximately 30% higher in bicasting method compared to unicasting method. These results suggest that bicast reduces packet loss, thereby increasing the proportion of frames that can be accurately reconstructed, which leads to a higher number of frames with elevated ADM values.

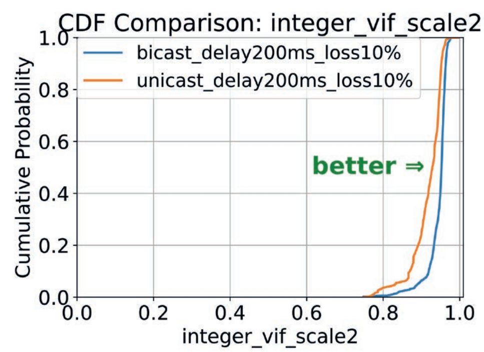
Next, we compare the CDF of Motion. as can be seen in both Figure 5a and Figure 5b, bicasting shows approximately 10% fewer zero compared to unicasting. These results suggest that bicast reduces the packet loss rate, which in turn lowers the frequency of frame freezes. Further analysis of frame freeze is discussed in a later section.



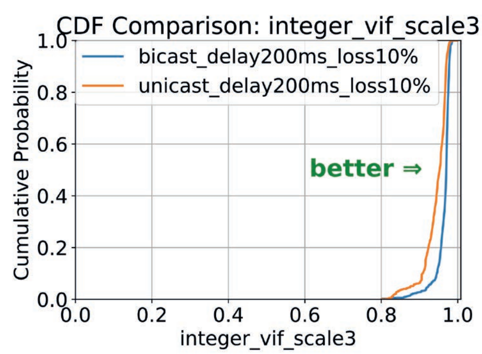
(a) scale0 comparison



(b) scale1 comparison



(c) scale2 comparison



(d) scale3 comparison

Fig. 6: VIF comparison

Next, we compare the CDFs of VIF. Bicasting method shows improvements for VIF scales in all periods, as compared

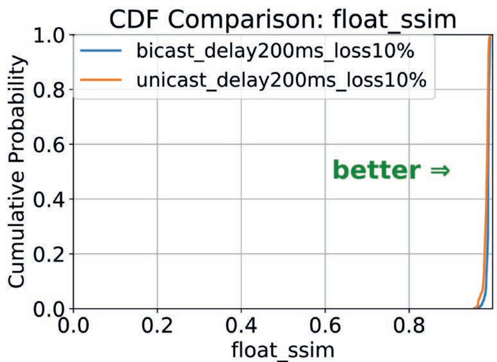


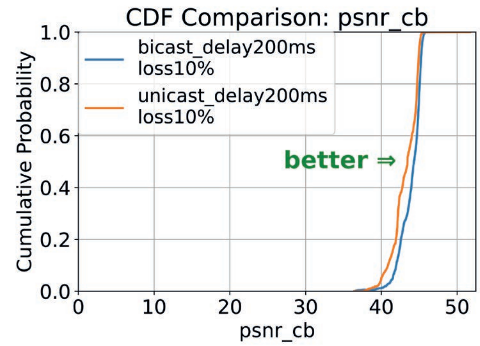
Fig. 7: SSIM comparison

to unicasting method. At a cumulative probability of 0.1, scale0 showed an improvement of approximately 0.06 in absolute value, corresponding to about 10% (Figure 6a); scale1 showed an improvement of approximately 0.06, corresponding to about 8% (Figure 6b); scale2 showed an improvement of approximately 0.05, corresponding to about 5% (Figure 6c); and scale3 showed an improvement of approximately 0.04, corresponding to about 4% (Figure 6d). These results indicate that bicasting reduces packet loss, allowing more visual information to be transmitted and improving VIF values.

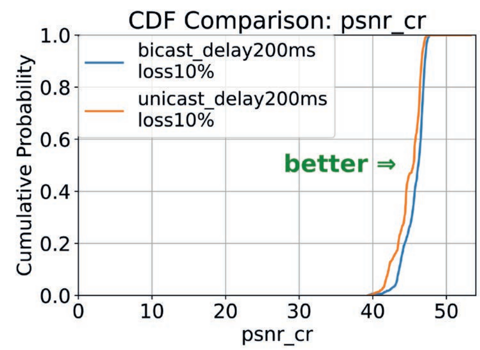
Next, we compared the CDF of SSIM. Bicasting method shows improvements in SSIM in all periods, as compared to unicasting method. As shown in Figure 7, the proportion of frames with SSIM values exceeding 0.97 was approximately 40% higher for bicasting than for unicasting. These results suggest that the reduction in packet loss through bicasting suppressed local frame noise, leading to improvements in SSIM values.

Next, we compared the CDFs of PSNR. Bicasting method shows improvements in PSNR in all periods, as compared to unicasting method. As shown in Figure 8a, for PSNR_cb, at a cumulative probability of 0.1, bicasting shows an improvement of approximately 1.0 in absolute value, corresponding to about 3%. Additionally, for both PSNR_cr and PSNR_y, as shown in Figure 8b and Figure 8c, an improvement of approximately 1.5 in absolute value, corresponding to about 4%, is observed at a cumulative probability of 0.1. These results suggest that bicasting reduces packet loss, suppresses noise, and consequently leads to an improvement in PSNR values.

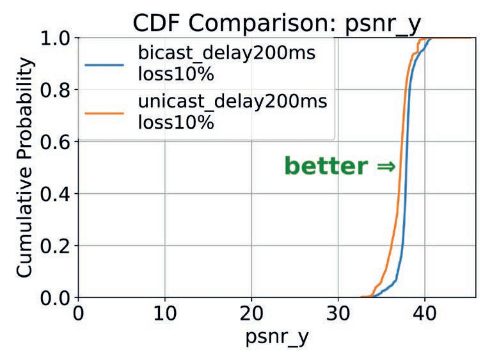
Next, the CDF of VMAF were compared. As shown in Figure 9, bicasting method shows improvements in VMAF in all periods, as compared to unicasting method. When focusing on high-quality video with VMAF scores exceeding 90, approximately 55% of unicasting method meet this threshold, whereas about 90% of bicasting method achieve it. At a cumulative probability of 0.1, an improvement of approximately 7 in absolute value, corresponding to about 9%, was observed. These results confirm that bicasting method maintains higher video quality more consistently than unicasting medho. These results suggest that the improvements in video quality achieved by bicasting, as indicated by ADM, VIF, SSIM, and PSNR, are also reflected in the VMAF scores.



(a) PSNR_cb comparison



(b) PSNR_cr comparison



(c) PSNR_y comparison

Fig. 8: PSNR comparison

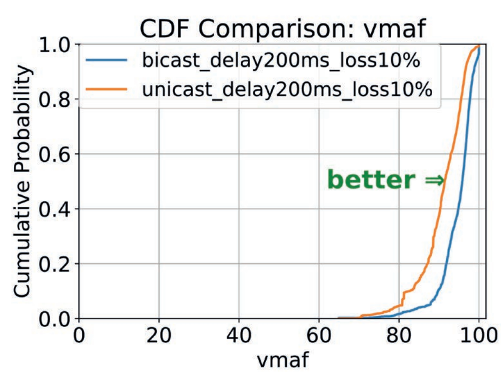


Fig. 9: VMAF comparison

Next, we compared the CDF of frame freeze. As shown in Figure 10, bicasting method shows improvements for frame freeze in all periods, as compared to unicasting method. At a cumulative probability of 0.9, an absolute difference of

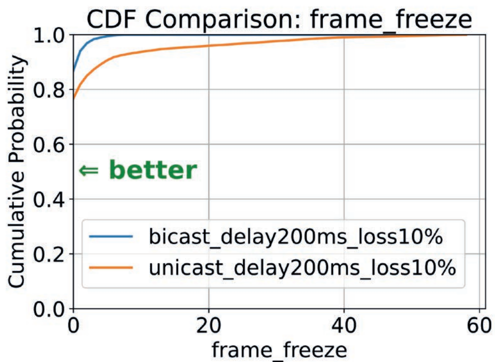


Fig. 10: Frame freeze comparison

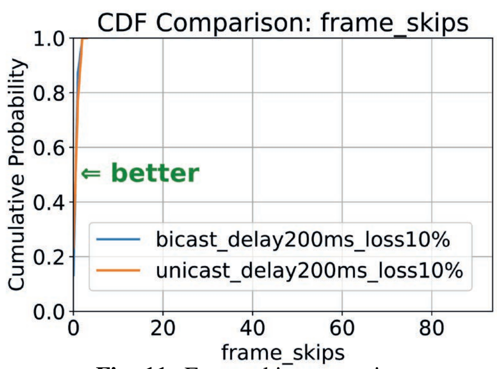


Fig. 11: Frame skip comparison

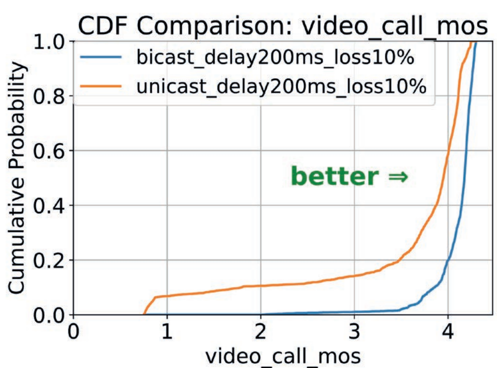


Fig. 12: Video Call MOS comparison

approximately 4 is observed, corresponding to an improvement of about 80%. These results suggest that bicasting reduced packet loss, thereby enabling smoother video playback.

Next, we compared the CDF of frame skips. As shown in Figure 11, bicasting method shows improvements for frame skips, as compared to unicasting method. Moreover, the proportion of frames with a value of 0 is approximately 10% lower for bicasting than for unicasting, indicating a reduction in the frequency of frame skips. These results suggest that bicast reduces packet loss, allowing more frames to be successfully played back.

Finally, the CDF of MOS estimated by VCM were compared. As shown in Figure 12, bicasting method shows improvements for MOS in all periods, as compared to unicasting method. Furthermore, when examining the proportion of MOS values exceeding 4, which is considered high quality, about 40% of the values exceed 4 in unicasting method, whereas bicasting method reaches approximately 80%. At a cumulative probability of 0.1, bicasting method shows an improvement of about 2 in absolute value, corresponding to approximately 120%. These results indicate that bicast maintains higher quality video more stably compared to unicast. This improvement is attributed not only to the enhancement in video quality observed in VMAF scores but also to a reduction in the number of frame freezes and skips, which decreased the occurrence of low MOS.

V. CONCLUSION

In this paper, we proposed and evaluated a bicast communication method in a wireless multihomed network using MoQ. The experiments confirmed that bicast communication can suppress QoE degradation under adverse network conditions. However, transmitting all packets via bicast increases the network load. In future work, we plan to investigate selective bicasting to extend MoQ Relay ’ s content-based in-network control. With this method, key frames are transmitted via bicast, whereas the remaining frames are sent via unicast.

REFERENCES

[1] Phil Cluff, “The Low Latency Live Streaming Landscape in 2019,” https://www.mux.com/blog/the-low-latency-live-streaming-landscape-in-2019.

[2] A. Mahanti, N. Carlsson, C. Williamson, and M. Arlitt, “Ambient Interference Effects in Wi-Fi Networks,” 9th International IFIP TC 6 Conference on Networking (NETWORKING), Lecture Notes in Computer Science, vol. 6091 (2010), pp. 160-173.

[3] L. Curley, K. Pugin, S. Nandakumar, V. Vasiliev and I. Swett, Ed, “Media over QUIC Transport,” draft-ietf-moq-transport-14, IETF, September 2025.

[4] Yanmei Liu, Yunfei Ma, Quentin De Coninck, Olivier Bonaventure, Christian Huitema, Mirja Kühlewind, “Managing multiple paths for a QUIC connection,” draft-ietf-quic-multipath-17, IETF, October 2025.

[5] T. Kawana, R. Nakagawa and N. Yamai, “Communication Multiplexing of Server with QUIC and SDN in Multihomed Networks,” 2024 International Conference on Information Networking (ICOIN), pp. 345-350, Ho Chi Minh City, Vietnam, 2024.

[6] ONF, “OpenFlow - Open Networking Foundation,” https://opennetworking.org/.

[7] VMware, “Fusion and Workstation,” https://www.vmware.com/products/desktop-hypervisor/workstation-and-fusion.

[8] Github, “Rust library for Media over QUIC,” https://github.com/kixelated/moq.

[9] G. Mittag, B. Naderi, V. Gopal and R. Cutler, “LSTM-Based Video Quality Prediction Accounting for Temporal Distortions in Videoconferencing Calls,” ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, pp. 1-5 2023.

[10] Songnan Li, Fan Zhang, Lin Ma, and King Ngi Ngan, “Image quality assessment by separately evaluating detail losses and additive impairments,” IEEE Transactions on Multimedia, 2011.

[11] Zhi Li, Anne Aaron, Ioannis Katsavounidis, Anush Moorthy, and Megha Manohara, “Toward a practical perceptual video quality metric,” Tech. Rep., Netflix, 2016.

[12] Hamid R. Sheikh and Alan Conrad Bovik, “Image information and visual quality,” 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[14] Zhi Li, Anne Aaron, Ioannis Katsavounidis, Anush Moorthy, and Megha Manohara, “Toward a practical perceptual video quality metric,” Tech. Rep., Netflix, 2016.