April 2021

HIGH QUALITY MULTI-PARTICIPANT CONFERENCE RECORDINGS USING DISTRIBUTED INTELLIGENT LOCAL BUFFERS

Aachala Singhan

Alex Taipale

Ana Lucia De Faria

Kyla Hayes

Hazim Dahir

See next page for additional authors

Inventor(s)
Aachala Singhan, Alex Taipale, Ana Lucia De Faria, Kyla Hayes, Hazim Dahir, Paul Giralt, Kaouther Abrougui, Sandhya Kerehalli, Jeff Apcar, and Geetha Anandakrishnan

HIGH QUALITY MULTI-PARTICIPANT CONFERENCE RECORDINGS USING DISTRIBUTED INTELLIGENT LOCAL BUFFERS

AUTHORS:
Aachala Singhan
Alex Taipale
Ana Lucia De Faria
Kyla Hayes
Hazim Dahir
Paul Giralt
Kaouther Abrougui
Sandhya Kerehalli
Jeff Apcar
Geetha Anandakrishnan

ABSTRACT

Online meeting audio and video recordings are typically performed in the cloud. As a result, conference call recordings may contain call quality imperfections that are caused by, for example, network impairments that may have arisen during the meeting. A recording may also be limited by, for example, any bandwidth constraints that were present between the meeting participants and the cloud. To address these types of challenges, various solutions are presented herein through several techniques. In particular, the techniques may include the distributed local recording of a meeting to allow for, among other things, high quality cloud-constructed playback of a meeting while minimizing the cost of local storage on each device. The techniques may further provide for high quality audio recordings by, for example, initiating local recording on a speaker's device when bandwidth issues arise and then through a cloud agent collecting and integrating such local recordings into a final recording that may subsequently be published and reused.

DETAILED DESCRIPTION

Most cloud-based web or video conferencing software as a service (SaaS) solutions provide some sort of mechanism for creating a recording of a meeting that, for example, is available for later playback. Such recordings are typically done either in the cloud (from an anonymous participant's perspective) or locally (from a meeting host's perspective).

While the current state of the art provides for reasonably good recordings of the audio, video, and application or desktop sharing sessions, the recordings will always be impacted by any network impairments that may occur during a meeting and which may negatively affect any of the media streams.

Additionally, a web or video conferencing participant will frequently refer to a meeting recording to, for example, catch up on a missed meeting, review some specific topics that were discussed during a meeting, etc. In some situations, due for example to network volatility, some portions of a recording which may be important may be corrupted or otherwise unintelligible.

Since a meeting recording may be performed centrally in the cloud, any corruption that may occur at an edge affects the final central recording. If participants in the meeting are speaking or presenting (and, for example, some network issues arise on their connection) there is no guarantee that their audio will be intact and understandable in the final recording.

To address challenges of the types that were described above, various solutions are provided herein through several techniques.

A first technique, as presented herein and as discussed and illustrated below, provides for the creation of meeting recordings without artifacts that may be due to network impairment because the recorded meeting data is transmitted in a non-real-time fashion. Additionally, aspects of this technique consume minimal local resources by utilizing an intelligent buffer to eliminate redundancy between the cloud and a local device.

A second technique, as presented herein and as discussed and illustrated below, may further ensure that the quality of an audio recording is maintained for each and every meeting. Additionally, aspects of this technique employ, among other things, edge agents to locally buffer (e.g., in a cyclic manner) and record the audio of speakers with bandwidth issues and a centralized cloud agent which may process and integrate (through, for example, audio stitching) the crowdsourced recorded sequences from all of the speakers into an online meeting recording thus ensuring better quality.

Under aspect of a first technique, as referenced above, all of the participants in a meeting – in addition to transmitting their audio, video, and/or sharing streams to the cloud – also save their own local copy of any cloud orchestrator requested media streams (such as, for example, audio, video, content sharing, etc.) to a local buffer on the participant's computer or endpoint. Media streams will only be recorded locally when the conference orchestrator requests a media stream from a device. Optionally, the user may be notified of which specific streams they are recording on their device.

Periodically, the cloud orchestrator may assess the quality of the received media streams through, for example, packet loss, bit rate, etc. The cloud orchestrator may downgrade the requested streaming bit rate when it observes packet loss. When a client receives a downgraded bit rate request, and there is available space on the device, a local recording will be kept for later transmission.

Aspects of the above are illustrated in Figure 1, below.

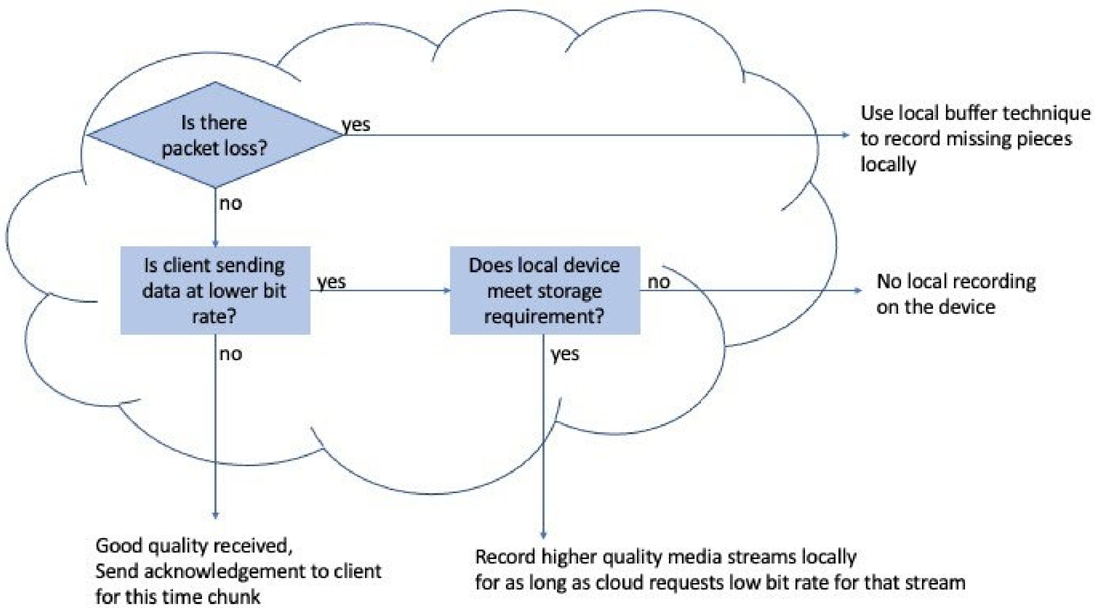

Figure 1: Illustrative Quality Assessment

If the evaluation of the media streams for a particular time interval is good, an acknowledgment may be sent to the device for which the media stream is being requested to notify it that any local recordings for that interval may be deleted from the local buffer. Otherwise, the local recordings will remain in the buffer until (a) they receive an acknowledgement from the cloud, (b) the first-in, first-out (FIFO) buffer is almost full and the oldest recordings are written to local storage, or (c) the call ends and all recordings are written to local storage. Aspects of this process are illustrated in Figure 2, below.

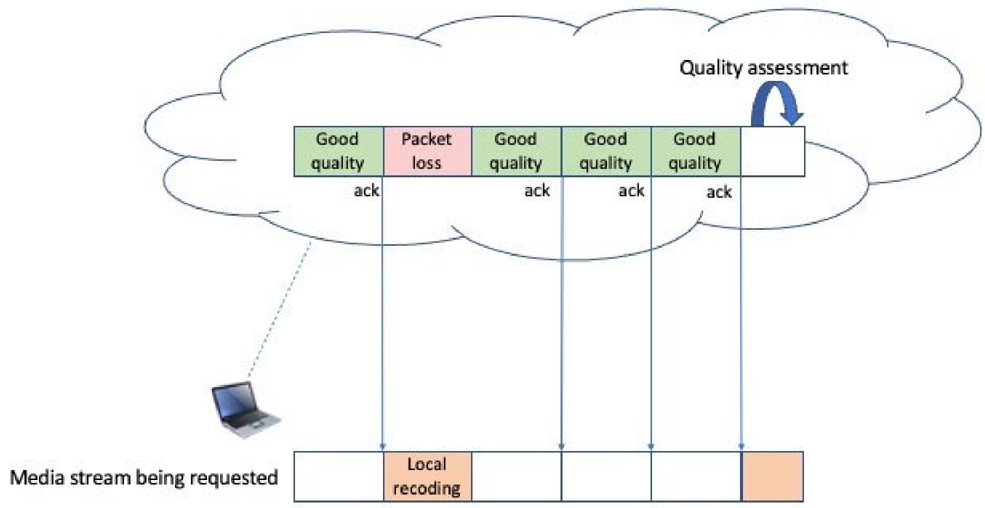

Figure 2: Exemplary Media Stream Quality Evaluation

It is important to note that the use of a local buffer to keep minimal parts of a local recording for later reliable transport is a key element of the first technique as described above.

After a meeting has concluded all of the locally recorded content that was kept for each participant may be uploaded to the cloud using a reliable transport so that any packet loss does not affect the quality of the uploaded data. The cloud will stitch together the media stream as it is received during the meeting, along with any missing pieces that were kept locally and retransmitted later, to form an artifact-free version of the media stream which can then be combined into the full meeting recording bundle along with other media streams that were part of the original recording.

Any network impairments that may have inhibited one participant from hearing or seeing another participant speaking are thus eliminated in the final product. The result is a meeting recording that is potentially of a higher quality than the original live meeting.

Turning to the second technique, as referenced above, aspects of this technique support, among other things, a video conference recording that allows participants or collaborators who were not able to attend the meeting to review the content of the meeting at a later time. Additionally, aspects of this technique also allow participants with bandwidth issues to replay the recording at a later time. Having a good quality recording with clear audio of the speakers is important. During an online meeting everything may be centrally recorded. Hence a recording will reflect any bandwidth issues, audio dropouts, etc. that a speaking participant might experience. It is important to note that while the narrative that is presented below focuses principally on audio recordings, aspects of this technique may be applicable to, for example, issues of video drop outs and corruption.

Aspects of the second technique overcome the different limitations that were described above by locally buffering (e.g., in a cyclic manner) and recording the audio of speakers with bandwidth issues and then publishing such recordings to a centralized cloud agent. The cloud agent may then process and integrate (through, for example, audio stitching) the crowdsourced recorded sequences from all of the speakers into an online meeting recording to ensure better quality. Note that audio stitching is only necessary if the system detects bandwidth issues at a participant end that is speaking.

Aspects of the second technique augment a recording that was generated by an online meeting facility as it has mechanisms at the edge and in the cloud to collect and process local recordings from all of the speakers with a bandwidth issue and then integrate the collected and processed recording to generate one full recording having good quality.

The second technique comprises three main components, including:

1. Participants.
2. Edge agents.
3. Cloud agent.

The first component, participants, encompasses, for example, speakers or listeners. At any one time a participant in a meeting is either listening or speaking (or in some cases perhaps interrupting).

The second component encompasses edge agents. Each meeting participant has, possibly among other things, an edge agent. Such an edge agent has two main roles, including:

1. Maintaining a cyclic buffer of the last x seconds of a conversation (x being a configurable number that is expressed in, for example, seconds and which is larger than the time that it takes to detect a problem). For example, one possible setting for x may be 60 seconds.

2. Monitoring the participant's bandwidth (e.g., when speaking) and taking one of two actions, including:
   a. If a bandwidth issue is detected, then the edge agent may record the audio until the bandwidth issue is resolved. The edge agent creates audio sequences by prepending the buffered audio and the recorded audio while the bandwidth was affected. Then, the edge agent may add timestamps to the beginning and the ending of each sequence and publish the sequences to the cloud agent.
   b. If no bandwidth issues are detected, then the edge agent may continue to buffer an audio content of the last (e.g., 60 seconds) in a cyclic manner.

Aspects of the above description of the second component, edge agents, are illustrated in Figure 3, below.

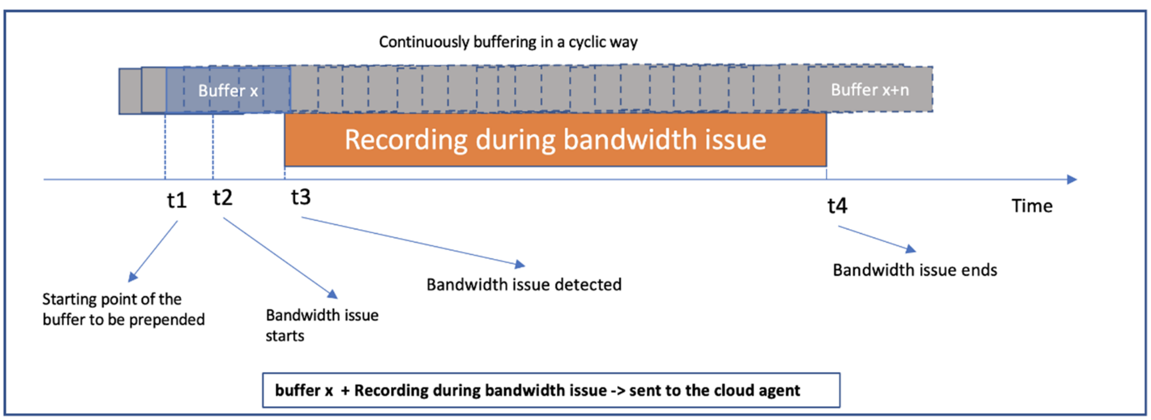

Figure 3: Illustrative Buffer

The second technique's third component encompasses a cloud agent. A cloud agent synchronizes the time for all of the edge agents in order to be able to correctly patch a recording that was generated by an online meeting facility. Such a cloud agent receives all of the time stamped recordings from all of the edge agents of speakers, uses these crowdsourced sequences to patch an online meeting facility's recording, and then publishes the processed recording. Aspects of a cloud agent are presented in Figure 4, below.

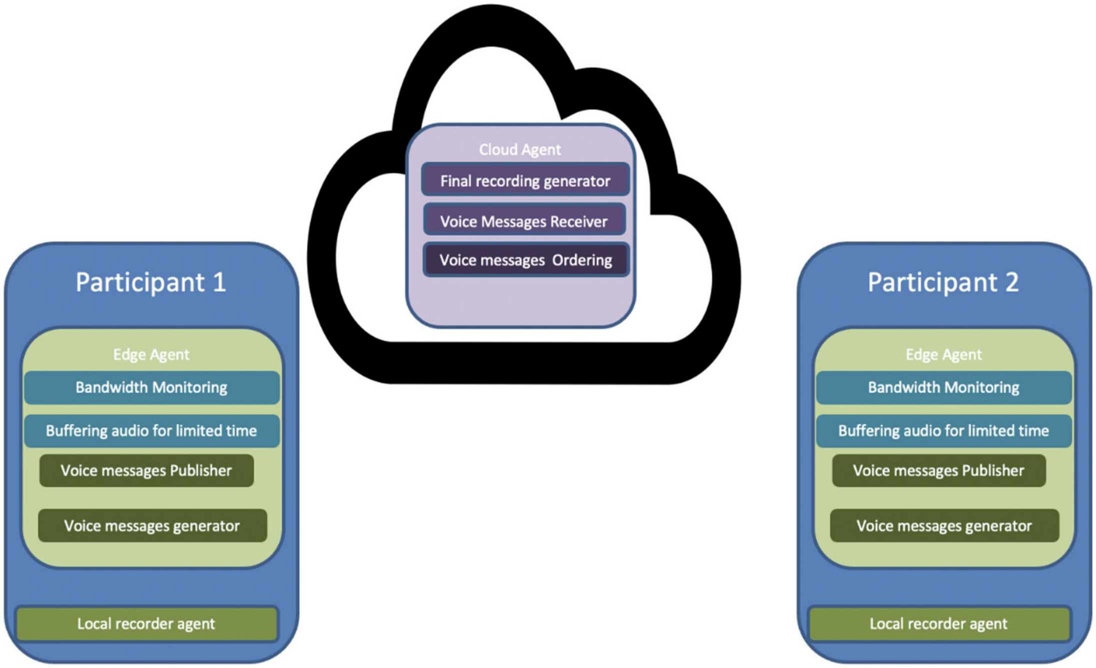

Figure 4: Illustrative Cloud Agent

The workflow steps that may be associated with the second technique, as described above, may include, for example:

1. Global time synchronization is provided for of all of the edge agents.
2. Each edge agent monitors the bandwidth at their corresponding local device.
3. Each edge agent cyclically buffers the last, for example, 60 seconds of audio of the speaker.
4. If any bandwidth issues are detected at the speaker level:
    a. The edge agent uses a device's local recorder to record the speaker during the live collaboration event.
    b. The edge agent creates recordings leveraging the buffered and recorded audio sequences.
    c. Recordings from a speaker having bandwidth issues are pushed to the cloud agent.
5. The cloud agent collects all of the recordings from affected speakers' edge agents.

6. After a meeting is finished, and the recording is stopped, the cloud agent integrates with the current online meeting recording by patching some parts of an online meeting facility's recording with the chunks that were pushed to the cloud agent based on timestamps. The cloud agent does not need to do any processing other than patching the received recording, based on timestamps, into online meeting recording.
7. The post-processed recording with better quality may be published.

Aspects of the above discussion of workflow steps are illustrated in Figure 5, below.

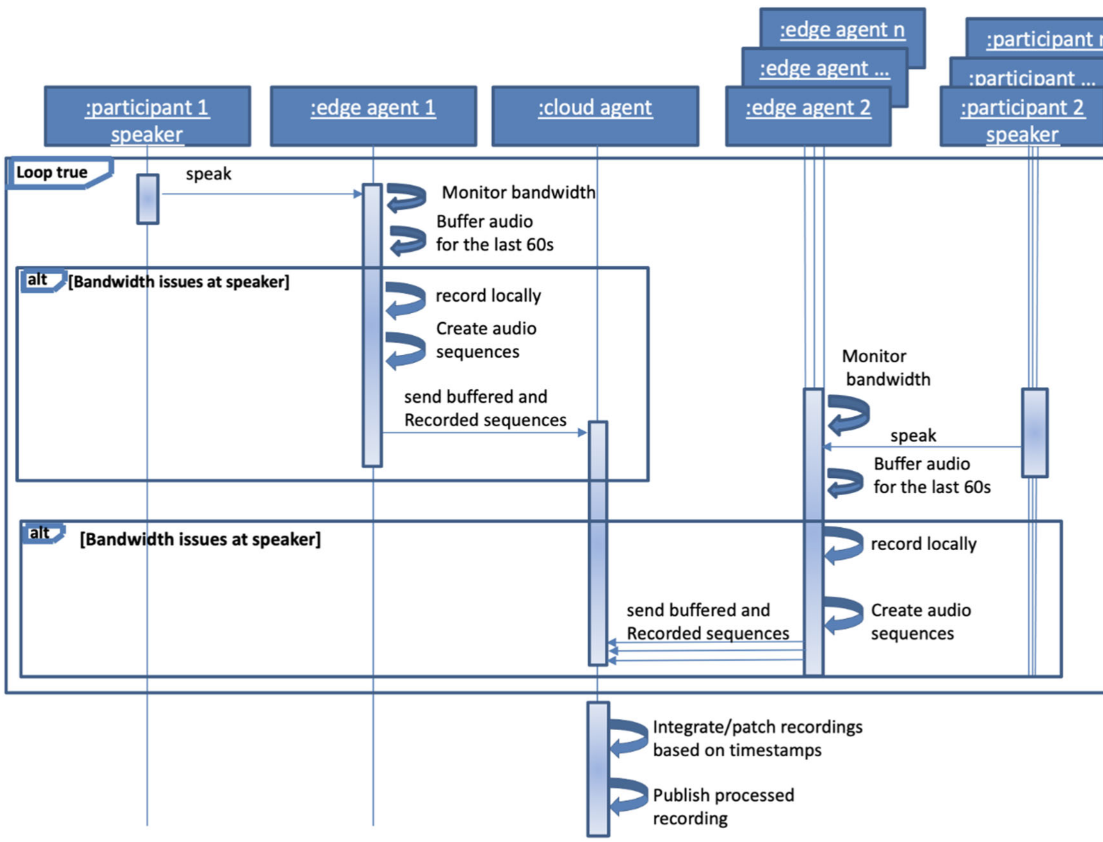

Figure 5: Exemplary Workflow Steps

In summary, techniques have been presented herein that advantageously provide for the distributed local recording of an online meeting to allow for, among other things, high quality cloud-constructed playback of a meeting while minimizing the cost of local storage on each device. The techniques that have been presented herein may further provide for the use of edge agents to locally buffer (e.g., in a cyclic manner) and record the audio of speakers with bandwidth issues and publish such recordings to a centralized cloud agent where such a cloud agent may process and integrate (through, for example, audio stitching) the crowdsourced recorded sequences from all of the speakers into an online meeting recording to ensure better quality.