How SyncStage™ supports Multi-Territory & Multi-Network Music Collaboration using the 5G Future Forums' Edge Discovery APIs

March 20, 2023

Abstract

This paper outlines how Open Sesame Media, Inc.'s (Open Sesame) digital audio platform (SyncStage™) was used to enable three guitarists to remotely jam together synchronously across three different locations (Canada, United States, & United Kingdom) and across three different mobile telephone networks (Verizon, Rogers, & Vodafone) using the Edge Discovery Service (EDS) APIs created by the 5G Future Forum (5GFF).

Minimizing audio latency (or audio delay) between a group of musicians is critical for musicians to hear, play, and react to each other's performances. Musicians who play together in the same physical location, experience little to no audio latency between each other.

Musicians who have attempted to remotely perform online with widely used digital collaboration solutions such as video and audio meeting applications and audio-only communications applications have experienced challenges with synchronously performing together due to the significant audio latency within these solutions (+250 ms).

That changes today with the launch of Open Sesame's SyncStage and the 5GFF's Edge Discovery Service (EDS) APIs. This paper describes how ultra-low latency synchronous audio collaboration and experiences are now available for musicians to perform together online. 5GFF and Open Sesame showcased this new enablement at Mobile World Congress 2023.

Problem Statement

Audio synchronization is critical between a group of musicians for them to successfully play music together whether they are in the same physical location or are performing remotely.

This paper addresses the key audio communication challenge that affects a group of musicians' ability to remain within a synchronous performance:

Latency - how fast musicians hear one another's performances in order to react and play with each other in a synchronized collaborative session.

According to research, musicians require less than 66ms of latency to synchronously perform together.

Background

Audio latency benchmark parameters have identified the following collaborative latency thresholds and the performance capabilities within each range:

- **55-66ms: Deterioration** (playing & singing accuracy rapidly falls off)
- **25-55ms: Deceleration** (Musicians experience some delay compensation)
- **8-25ms: BEST SYNCHRONICITY** (stable tempos of performers)
- **0-8ms: Acceleration** (ability to anticipate another musician's performance)

There are additional considerations that may further impact latency parameters such as:

- Musical performance tempos (the beats-per-minute for the song eg. 80 BPM for Ballads; 120 BPM for Pop; 200 BPM for Punk),
- Musician experience levels (beginner, semi-pro, to expert)
- Music instruments required to be synchronized (guitars only, guitars & drums, etc.)

The Latency Issue Solution

In order to achieve these latency parameters to enable musicians to perform remotely, SyncStage has addressed two important latency challenges to deliver users an ultra-low latency synchronized audio collaboration experience:

1) **Processing latency** - the time needed to capture, process, and play back the audio feed. SyncStage has maximized the processing latency challenge by maximizing the low-latency audio processing capabilities within the smartphones.

2) **Network latency** - the total time the audio data packets travel between musicians located at different physical locations. Usually, the greater the distance between users; the higher the network latency.

To solve the Network Latency challenge of connecting three guitarists across three countries and between three telecom networks, the 5GFF team with Open Sesame explored several architecture solutions to reduce the network latency so that three guitarists could jam together in a live remote musical session.

Diagram A (Option 1) illustrates the attempt to use a Client Server Model with a Public Cloud.

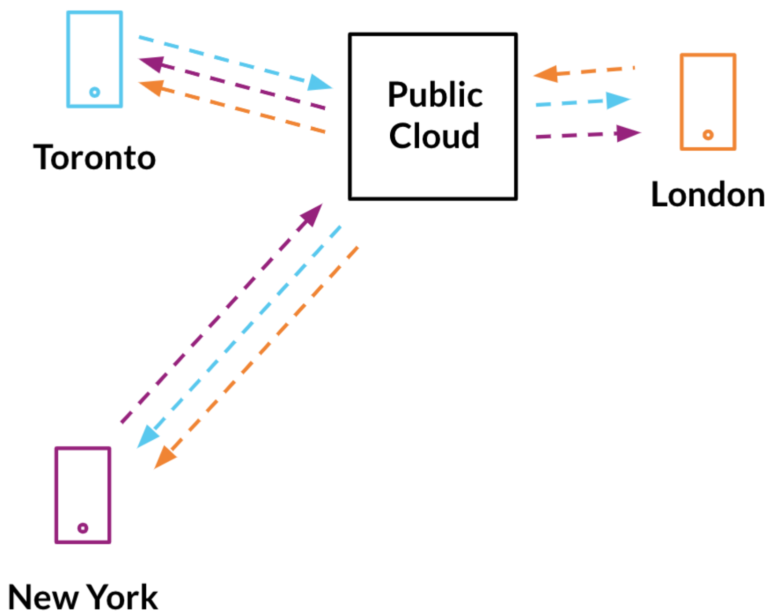

Diagram A (Option 1) Client-Server Model Utilizing the Public Cloud

The Network Latency challenge with this solution resulted in unreliable and inconsistent latency needed to enable synchronized collaboration of the musicians. The standard Public Cloud created geographical distance from one or more users and thus resulted in higher network latency than what was needed.

Diagram B (Option 2) illustrates the attempt to use a Client Server Model on the MEC.

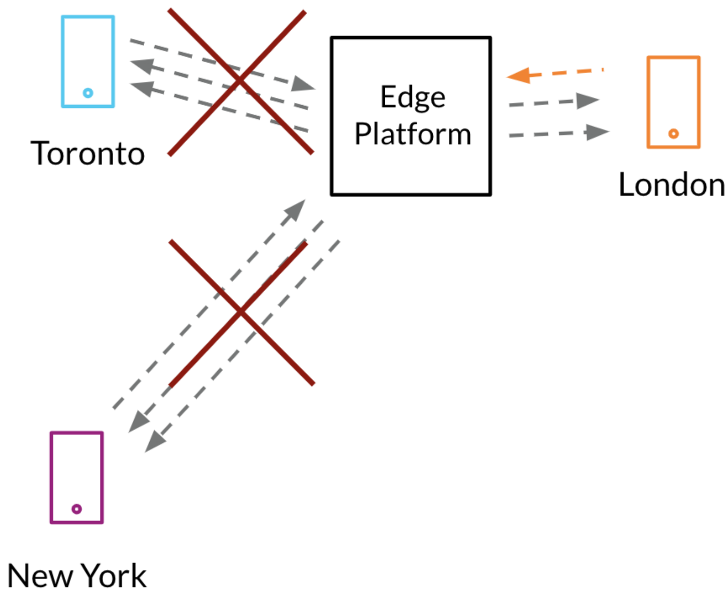

Diagram B (Option 2) Client-Server Model on the MEC

Interoperability across countries - musicians are likely to be connecting to a session from different networks and also from different countries. With current telecom multi-access edge computing architecture, they can connect only to an edge location specific to that telecom operator in that country. Therefore, connections between users on different carriers can't be established.

Solution - Selected Technical Architecture for Remote Music Collaboration over Significant Distances and Across Three Telecom Networks

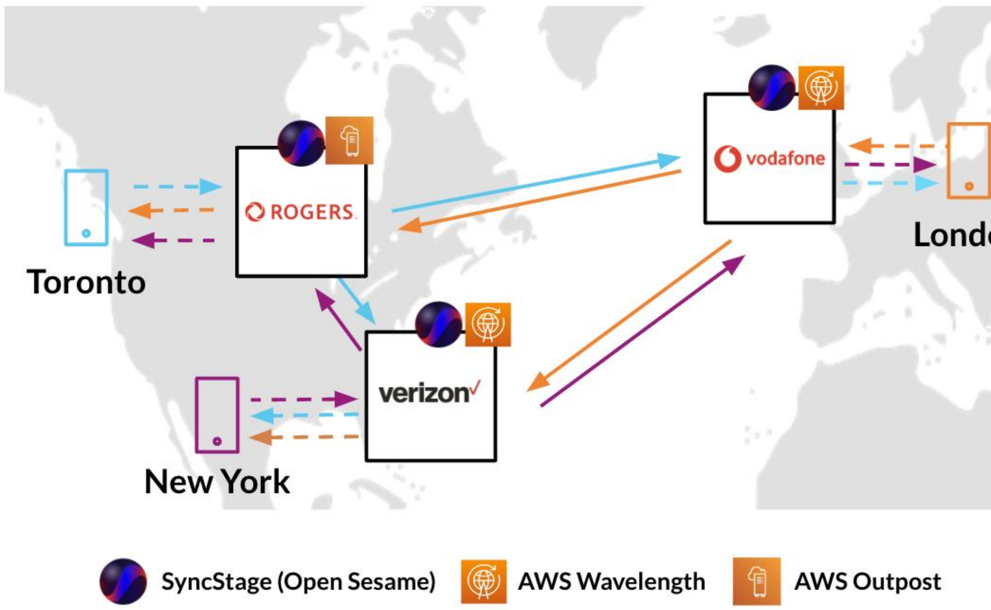

Diagram C (Option 3 - SyncWorld) - Selected Technical Architecture for 5GFF Music Festival

The finalized technical architecture was configured to minimize latency of the musicians for them to jam together in 'real-time'. Each musician was uniquely connected to the edge location and network that was the closest to their physical location: Toronto's musician connecting to Rogers network edge location; New York's musician connecting to Verizon network edge location; and London's musician connecting to Vodafone network edge location. Each edge location then is interconnected with one another.

1. **SyncStage deployed on AWS infrastructure across Verizon, Rogers, and Vodafone.**
2. **Each performance location within each country was connected to the closest edge location utilizing the 5GFF Edge Discovery Service API.** This architecture minimized the distance to the carrier backbone network thus providing reduced latency, jitter, and enabled audio synchronization.
3. **Interoperability across multiple networks in multiple territories** - enabling the connection of the 3 guitarists to perform together digitally across networks and countries.

This architecture then maximizes the two variables we discussed (full methodologies around testing methodology appear in the Appendix).

Latency

It is important to note this architecture still must deal with the geographic distance between locations. However, what the SyncWorld architecture - utilizing the 5GFF's Edge Discovery APIs - allows for a reduction in latency between the musician's device and the local edge location in that territory, when compared to local Public Cloud: a 14% decrease in latency in New York; a 25% decrease in latency in London; and a 21% decrease in latency in Toronto, when compared to local Public Cloud (Diagram C):

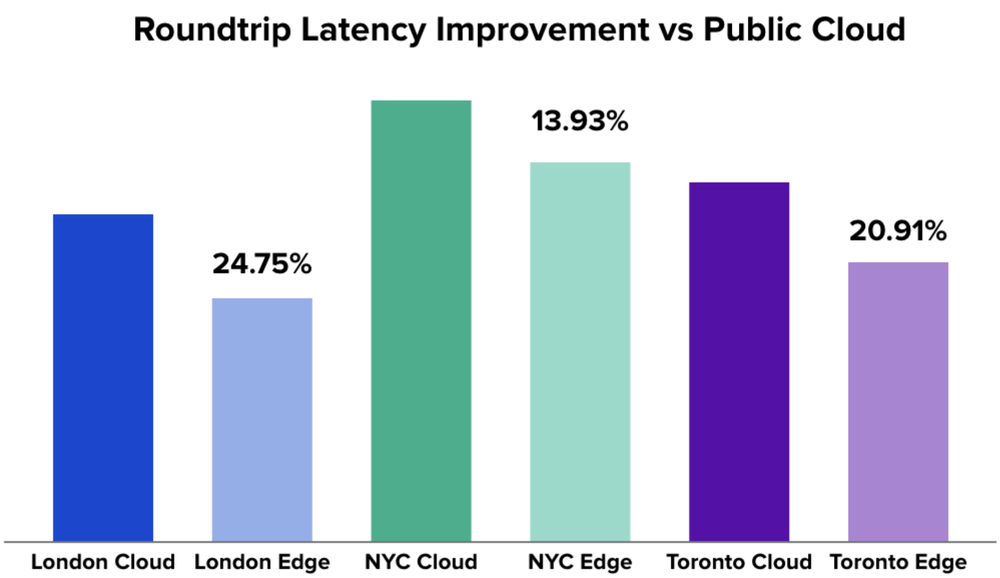

Diagram C: Latency between the musician's client device in New York, Toronto, London connecting to their local studio server using SyncStage on Public Cloud vs. 5GFF Edge Discovery Service APIs recommended Edge location.

If we then include the processing latency of SyncStage software, we see overall one-way audio latency - from client device A -> edge location provided by EDS -> to client device B in a similar location - allows for increasing audio synchronicity, and improves upon Public Cloud; a 7.6% decrease in latency in New York; a 11.7% decrease in latency in London; and a 10.5% decrease in latency in Toronto, when compared to Public Cloud.

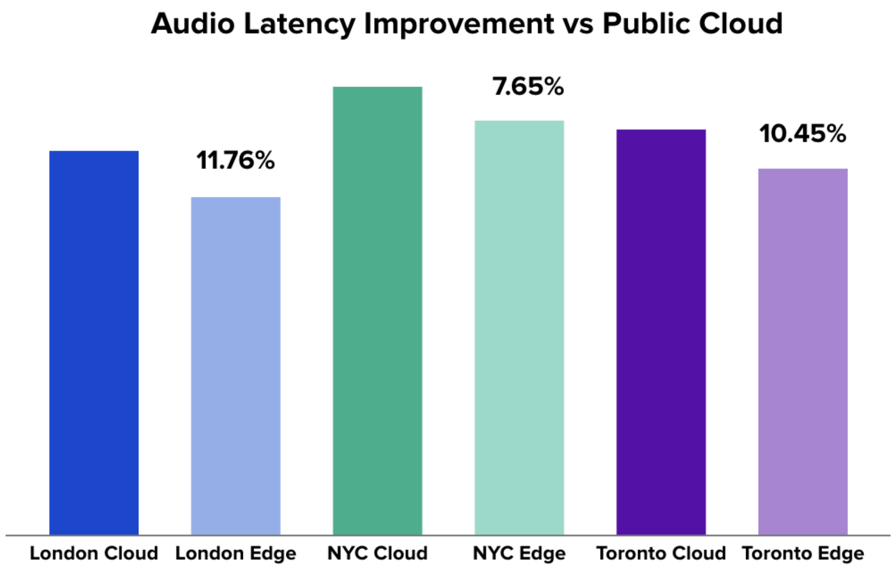

Diagram D: One way audio latency between the musician's client device in New York, Toronto, London connecting to their local studio server using SyncStage to the other musician's device in a similar location on Public Cloud vs. 5GFF Edge Discovery Service APIs recommended Edge location

Conclusion

SyncStage's audio pipeline delivers better audio synchronization to a group of musicians performing remotely online using the 5GFF's Edge Discovery Service (EDS) APIs.

By testing SyncStage running on Public Cloud against SyncStage running on the 5GFF EDS APIs MEC infrastructure we found:

- SyncStage's SyncWorld infrastructure delivers a 14-25% increase performance in round trip latency to the edge location when compared to the Public Cloud Client Server framework between musicians in similar locations

- SyncStage's SyncWorld infrastructure remained within the audio synchronization parameters for users in similar locations to deliver a 7-12% performance increase and enhancement over the Public Cloud Client Server framework (given similar benchmarked network conditions)

- SyncStage's SyncWorld infrastructure allows for a multi-network, multi-territory environment, where local users can benefit from audio synchronization and, at the same time, take part in a low latency audio session with their remotely located users and distant peers - see Diagram E.

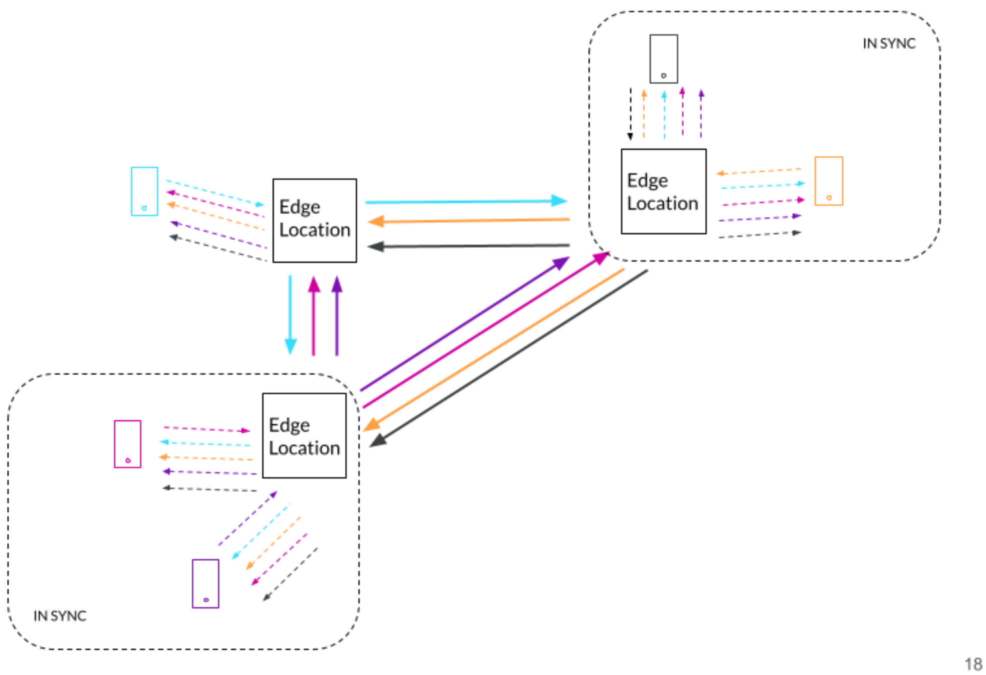

Diagram E: SyncWorld architecture extension utilizing SyncStage and 5GFF EDS APIs in New York, Toronto, London connecting to their local studio server using SyncStage for education, music collaboration and metaverse use cases

Potential Future Use Cases

- ➔ Music Collaboration
- ➔ Music Education
- ➔ Online Gaming Communications
- ➔ Metaverse Virtual Concerts

Appendix A - Latency Testing Methodology

(See Diagram F, below) We measure latency by making a SyncStage voice call between two smartphones. Within that call, we placed a pulse sound on one smartphone (App A) and captured the pulse time. We then record the output of the other smartphone (App B) and record the pulse time. We then compare both noted pulse times - on both App A and App B - to determine the delay caused by transmitting audio from one smartphone to another, and therefore determine the delay of a voice call. We then ran the same scenario using SyncStage in a Public Cloud and SyncWorld Architecture.

Testing Setup

- iPhone 13 (2 smartphones)
- Network 5G
- SyncStage 0.24.0 (studio servers in London, New York, Toronto)
    - Noise filtering off

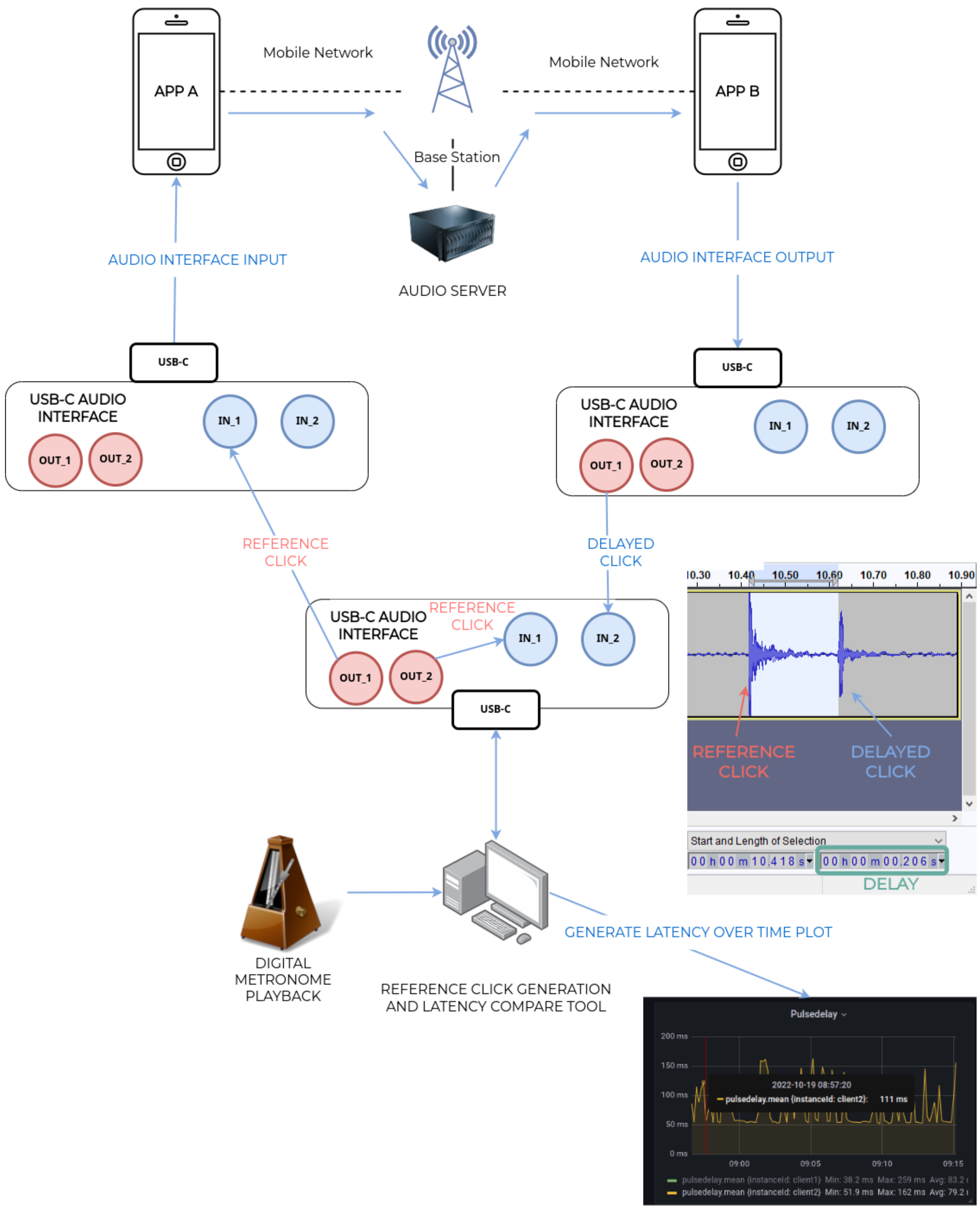

Diagram F: Latency testing methodology for comparative study.

References

1. Chafe C, Cáceres J, Gurevich M **Effect of temporal separation on synchronization in rhythmic performance Perception 2010**;39(7):982-92
2. Farner, S; Solvang, A; Sæbo, A; Svensson, U. **Ensemble Hand-Clapping Experiments under the Influence of Delay and Various Acoustic Environments Centre for Quantifiable Quality of Service in Communication Systems** JAES Volume 57 Issue 12 pp. 1028-1041; December 2009
3. Hupke R; Beyer L; Nophut M; Preihs, S; Peissig J **Effect of a Global Metronome on Ensemble Accuracy in Networked Music Performance Audio** Engineering Society Convention: 147 (October 2019) Paper Number: 10218

What is SyncStage?

SyncStage is an audio pipeline that is optimized for 5G and made available to application developers via a Software Development Kit (SDK) to integrate into their applications. SyncStage provides end-users of various audio experiences:

- Synchronized audio connections between multiple users
- Ultra-low latency connections between multiple users
- High quality audio output that is heard between all users

Application developers integrate the SyncStage Android, iOS, and/or Unity SDKs into their application and whenever a synchronized audio session is requested, Open Sesame's audio infrastructure powers the audio connections and communications between a group of synchronized audio users.

Learn More

To learn more about this capability, contact the team at Info@OpenSesame.Media and visit them at Open Sesame and SyncStage.