Research Report on Real-time Multi-channel Audio Transmission Systems

Taebin Yoo 2025. 10. 21

1. Research Objectives

The objective of this research was to implement a system capable of stable, real-time transmission and reception of DAW-based multi-channel audio over a **Wide Area Network (WAN)**. Specifically, this study explored and experimented with audio transmission methods that satisfy both **low latency** and **sample accuracy**, assuming a remote collaboration scenario where a live performance is recorded or mixed in real-time at a remote studio.

2. Methodological Approaches and Limitations

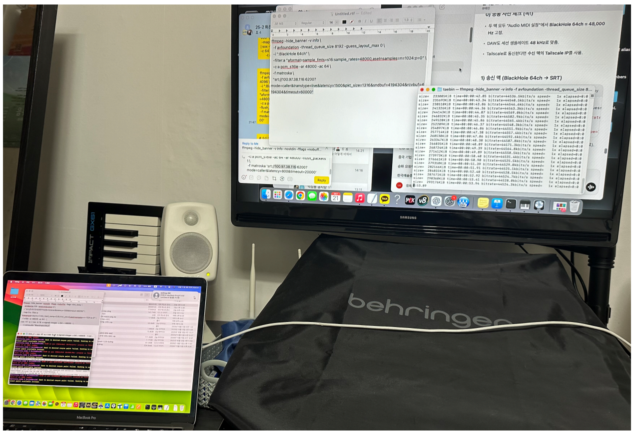
(1) **Commercial Protocols (AES67 / Dante / AVB**)
Results: While hardware-based clock synchronization offers superior audio quality and stability, these were excluded from the study due to their lack of native WAN support, L2/L3 network constraints, and high dependency on proprietary hardware/infrastructure.

(2) **Server-based Relay System (SRS)**
Results: Audio transmission was stable, but the media-stream-centric architecture was unsuitable for multi-channel DAW audio required in performance environments and could not guarantee sample accuracy.

(3) **Peer-to-Peer SRS LiveTransit (Non-server)**
Results: Since the audio format and timing of the sender and receiver could not be forcibly synchronized, tick noise occurred due to word clock mismatches between DAWs.

(4) **Tailscale-based P2P + ffmpeg**
Results: Successfully bypassed NAT and established a WAN connection. However, because ffmpeg processes audio based on media frames, it could not synchronize with the audio interface's hardware clock. This resulted in **clock drift**and intermittent tick noise during operation.

(5) **JACK Audio + JackTrip**
Results: Attempted multi-channel real-time transmission using **JackTrip**, which supports sample-level audio transfer. Although the network connection was successful, a common word clock (sample timing reference) could not be enforced in a WAN environment. This led to persistent **XRUNs (buffer overruns/underruns)** and unresolved tick noise.



3. Analysis of Failure Causes

The core reason for the failures identified in this study was not network bandwidth or general latency, but rather the **limitations of the audio clock synchronization architecture**.

In a WAN environment, there is no common "Master Clock" between the sending and receiving DAWs. While JackTrip and JACK Audio can deliver audio samples over a network, they do not guarantee the synchronization of hardware word clocks. Consequently, **sample drift** accumulates over time, inevitably leading to XRUNs and audible digital artifacts (tick noise).

4. Conclusion and Significance

This research experimentally confirmed that the primary constraint of real-time network audio is not merely latency or bandwidth, but rather **system architecture and clock synchronization**.

By sequentially validating SRS, ffmpeg, and JackTrip, the study concluded that real-time multi-channel audio between DAWs over a WAN is practically unsustainable without a dedicated, independent clock synchronization structure. This confirms that future system designs must incorporate a synchronization mechanism that functions independently of the network transmission layer.