Democratizing Interactive Music Production over the Internet

Trond Engum
Norwegian University of Science and Technology
7491 Trondheim
trond.engum@ntnu.no

Otto J Wittner
UNINETT
Trondheim, Norway
otto.wittner@uninett.no

ABSTRACT

This paper describes an ongoing research project which address challenges and opportunities when collaborating interactively in real time in a "virtual" sound studio with several partners in different locations. "Virtual" in this context referring to an interconnected and inter-domain studio environment consisting of several local production systems connected to public and private networks. This paper reports experiences and challenges related to two different production scenarios conducted in 2017.

Author Keywords

Audio over IP, Collaborative music production, Networked music production, Networked music, Music production

CCS Concepts

•**Human-centered computing** → *Human computer interaction (HCI);* •**Networks** → *Overlay and other logical network structures;* •**Information systems** → *Multimedia streaming;*

1. INTRODUCTION

In the academic institutions there is a long history of real-time networked musical performances (e.g. [11, 7]). Many of these institutions use non commercial protocols for transmission of sound and video [9].

As a parallel there has been a vast development within audio over IP for professional and semi professional music production systems the last decade [5, 4, 2].

However when viewing music production in a real-time networked context, the above described development has not been integrated in a day to day practice, and has therefore not reached a wider group of producers and mixers inside the professional music recording industry, nor the large community of "bedroom producers".

2. BACKGROUND

Before describing experiments with two production scenarios we will start with a short summary of the context.

2.1 Music production

In the last two decades since the introduction of digital sound technology a dramatic democratization and increase in access to sound production devices and systems has taken place. This has pushed the boundaries of what can be defined as a sound studio when it comes to mobility, costs, functionality and efficiency compared to its predecessors. Introducing real-time interactive music production in this context could expand the bedroom producers production system by giving them access to numerous acoustically adapted spaces, and further enable them to share sound equipment and approach music production as a real-time collaborative process.

2.2 Networked music

Even though one of the earliest examples of networked music can be traced back to the 1950's [6] it was not until the early 90's that the first experiments with networked music performance over the Internet took place [10]

The steady increase in capacity in Internet backbone networks over the last two decades has enabled an average user to install a fixed Internet access connection with 10-100Mbps capacity [3].Hence professional music studios and concert venues may now interconnect and initiate music collaboration.

The large community of bedroom producers will soon also be able to interconnect professionally and gain real time access to high end resources from e.g. professional studios or online music equipment rental companies.

2.3 Imagining a framework

One of our main questions has been whether it is possible to integrate existing production practices and contemporary technical solutions across different DAWs into a context where several producers and musicians collaborate interactively with real-time music production.

The premises to ensure enough usage flexibility such that both real-time musical interaction and real-time sharing of acoustical spaces and processing equipment is enabled, has been to enable low latency (<50ms if possible) audio and video connections. An inclusive system infrastructure is required to enable exchange of audio and control data between production systems such that real-time interactive music production scenarios may be maintained, scenarios where involved participants may still offer their expertise by applying their work flow of choice through their own production "instrument".

We have established a comprehensive "virtual" studio across Trondheim city consisting of several buildings acoustically treated for musical performance and production based on available technology.

To ensure an inclusive infrastructure we have chosen to experiment with an extended use of LANs via Virtual LAN technology (VLAN) [1] and VxLAN tunneling [8].

3. TWO PRODUCTION SCENARIOS

NIME'18, June 3-6, 2018, Blacksburg, Virginia, USA.

This section describes two different production scenarios completed within this infrastructure during 2017.

3.1 Interconnecting two professional studios

This session was focused around a live recording of a band split between two different locations 4km apart using a private VxLAN network, DANTE as audio protocol and Eucon protocol for control signals. Skype was used to connect multiple cameras for visual communication between locations. The session was carried out to highlight the possibility to expand the number of musicians and acoustical spaces by sharing them between studios, the possibility to share and combine analogue and digital effects, and the possibility to work with several producers in real-time.

3.2 Mastering over the public Internet

Traditionally audio mastering is the last stage of a production chronology, and is done offline in a studio specially constructed for this task. This means that musicians, technicians and mixing engineers often are left out during this process. Feedback may be given and received only between each track version, as complete tracks are sent back and forth. The session was conducted over the public Internet between Trondheim and London. During the session the attendants brought their own multi-track mixes, sending high quality audio using Jacktrip to a professional mastering studio in London . Skype connected to multiple cameras were used for communication between locations. This session was focused around the possibility for attendants to give and receive direct feedback to the mastering engineer during the session. Sending audio feed directly from the multi-track also enabled real time adjustments in the original mix before, during and after feeding into the mastering chain. Processing effects could also be shared and combined between the studios, and arranged in a mastering effect chain.

4. DISCUSSION AND CONCLUSION

For the first scenario, extending a dedicate LAN for music collaboration (by VLAN and VxLAN technology) has proven to be possible with limited investments. However it remains to be seen if this approach can scale further. Crossing administrative domains with such technology has also proven to be challenging due to conservative management regimes within IT network departments at the institutions involved. But when operative, the system provides numerous channels of low latency audio between locations, and is flexible enough to meet most conventional production scenarios. Even though the idea of several producers working collectively seems compelling our experiment revealed new challenges none-existent in conventional studio production. Timely and precise interpersonal communication between producers is required as well as division of labor with respect to technical control between the locations. .

The second example, run of the public Internet, proved to be more unstable. IP network, especially public, are inherently bursty of nature. The longer the network stretch in question is the more likely it is that a stream is hampered by a burst of cross traffic. In addition, when operating on a public network tight synchronization of end system becomes challenging due to delay and data loss. Recovery functionality, including discovering who your current peers are and/or if they have shut down/recovered, is more complicated since broadcast services cannot be utilized as elegantly as in private LANs. Latency on the other hand, is not as crucial in mastering session as for real-time recording session since the process is based on already stored media. In future experiments latency will be added to smooth out potential bursts transparently for the audio transmission tool. Due to the absence of live musicians the interpersonal communication between producers as well as the division of labour between them were experienced more precise. Expanding this experiment with multi-track mixing between locations will be a natural next step.

The two production scenarios presented and discussed show that there are numerous steps to walk and obstacles to cross when moving towards a democratization of interactive music production over the Internet. There are currently many directions which need to be narrowed down both concerning technical facilitation and interpersonal communication. There are also several factors, apart from physical distance, emerging when applying the public Internet, especially in terms of latency and stability. Despite this we believe that we have constructed a robust framework for further experimentation. The framework may help bring out new questions and provide answers about the design of technological solutions, about aspects of audiovisual communication and about interpersonal interactivity in distributed music production.

5. REFERENCES

[1] Ieee standards for local and metropolitan area networks: Virtual bridged local area networks. *IEEE Std 802.1Q-1998*, pages i–, 1999.

[2] AES67-2015: AES standard for audio applications of networks - high-performance streaming audio-over-IP interoperability. Technical report, Audio Engineering Society, 2015.

[3] ICT facts and figures 2017. Technical report, International Telecommunication Union (ITU), 2017. `https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf`.

[4] ALC NetworX GmbH. Ravenna: The open standard for real-time media over IP. `https://www.ravenna-network.com/`, jan 2018.

[5] Audinate Pty Ltd. Dante overview. `https://www.audinate.com/solutions/dante-overview`, jan 2018.

[6] J. Cage. Imaginary landscape no. 4. *CF Peters. Madacy Records*, 6179:1961, 1960.

[7] Gruppo per l'Armonizzazione delle Reti della Ricerca (GARR). Music and Art Communities. `https://www.garr.it/en/communities/music-and-art`, april 2018.

[8] M. Mahalingam, D. Dutt, K. Duda, P. Agarwal, L. Kreeger, T. Sridhar, M. Bursell, C. Wright, and W. R. Sutherland. Virtual extensible local area network (vxlan): A framework for overlaying virtualized layer 2 networks over layer 3 networks. RFC 7348, Internet Engineering Task Force, Aug. 2014.

[9] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti. An overview on networked music performance technologies. *IEEE Access*, 4:8823–8843, 2016.

[10] A. A. Sawchuk, E. Chew, R. Zimmermann, C. Papadopoulos, and C. Kyriakakis. From remote media immersion to distributed immersive performance. In *Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence*, ETP '03, pages 110–120, New York, NY, USA, 2003. ACM.

[11] Standford Center for Computer Research in Music and Acoustics. Soundwire - Sound Waves on the Internet from Real-time Echoes. `https://ccrma.stanford.edu/groups/soundwire/`, april 2018.