By Zefir Kurtisi, Xiaoyuan Gu, *and* Lars Wolf

ENABLING NETWORK-CENTRIC MUSIC PERFORMANCE IN WIDE-AREA NETWORKS

NMP on the Internet is not only possible, its delay bounds can satisfy the tight requirements involved in human perception while generating high-end audio quality for musicians and listeners alike.

The ubiquitous availability of broadband Internet connectivity at home is driving the use of the Internet for entertainment and other forms of recreation. Paving the way for many demanding multimedia applications over IP, it has accelerated the emergence of new ideas for network-centric collaborative work that was impossible only a few years ago for both technical and economic reasons.

Among many new types of networked entertainment genres, network-centric music performance, or NMP, [3] represents a vision of multi-party musical performance delivered through cyberspace that strives to overcome the inherent limitations in conventional rehearsals and concerts. NMP refers to a system that allows musicians who are physically separated, even over vast geographical distances, to participate in rehearsals and concerts across the Internet with bounded delay and acceptable audio quality. Similar work is exemplified by the Stanford SoundWire Project [2] and the Conductor Driven Scheme [1].

In [3], we focused on building a proof-of-concept NMP system prototype in a LAN. We also investigated delay and audio quality as they are affected by end systems. To make it possible for NMP to go beyond being only a laboratory tool, we've had to address a number of technological challenges. Above all, we've found that meeting its extreme stringent delay bound is a critical prerequisite. Here, we describe the NMP application boundary, presenting some of our evaluation results on NMP operation in wide-area networks.

The figure outlines a set-up of an NMP system consisting of a centralized server acting as the mixer and one or more clients connected through the Internet. The client(s) produce and send audio packets to the server. The server puts packets from each client into a separate queue, applies mixing, and returns the mixed packets to all clients where the final contents are played out through the users' audio playback devices.

In order for the musicians in an NMP session to interact with one another in a natural way, the end-to-end delay must be kept below human perception; 30msec [4] is a widely recognized bound. End-to-end delay here means the total delay (such as between hitting a piano key and hearing the time-synchro-