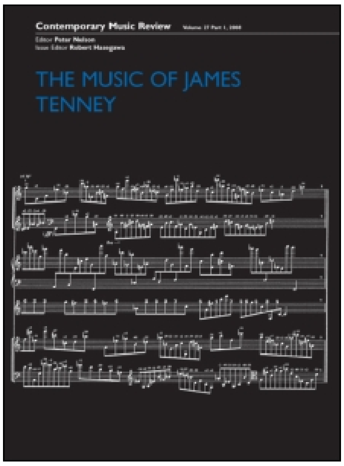

Contemporary Music Review
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713455393

Tapping into the Internet as an Acoustical/Musical Medium

Chris Chafe

Online publication date: 14 December 2009

To cite this Article Chafe, Chris(2009) 'Tapping into the Internet as an Acoustical/Musical Medium', Contemporary Music Review, 28: 4, 413 — 420

To link to this Article: DOI: 10.1080/07494460903422362

URL: http://dx.doi.org/10.1080/07494460903422362

Contemporary Music Review
Vol. 28, Nos. 4/5, August/October 2009, pp. 413–420

Tapping into the Internet as an Acoustical/Musical Medium

Chris Chafe

Network audio technology transforms advanced networks into a new kind of sound propagation medium, with its own properties compared to air, water, or solids. Three areas of research are described: 1) methods for using your ears for monitoring quality-of-service (QoS) of networks supporting high-quality, real-time, interactive, bidirectional flows; 2) new musical practices being made in the medium; and 3) a discussion of human factors affected by some unique acoustical properties of the medium.

Keywords: Network Music Performance; Time Perception; Audio Over Internet; New Media

Introduction

The fantastic world described by Francis Bacon in his *New Atlantis* imagined elements that are one and the same with computer music developments of the last half century. 'We have also sound-houses, where we practise and demonstrate all sounds and their generation. We have harmony which you have not, of quarter-sounds and lesser slides of sounds. Divers instruments of music likewise to you unknown' (Bacon, 1626). The passage's closing line forecasts worldwide sound: 'We have all means to convey sounds in trunks and pipes, in strange lines and distances'. Electronics first realized this prediction with analog telephony and then radio over a century ago. Yet from our present vantage point of rapidly expanding digital networks and media streaming, a compelling new reading emerges promising something much more.

Internet2, Canarie, Geant2, Cernet2 and their peers are scaling up to astonishing capacities. Where demonstrated real-time, interactive, uncompressed flows are currently in the 'centi-flow' range for high-definition audio and music collaboration, we can expect (soon) to significantly reinterpret Bacon's vision with the emergence of a new medium that boasts a several orders of magnitude greater number of interactive channels. Music performance in prototype 'acoustical chat rooms' already achieves a sense of co-location beyond traditional teleconferencing. Improved

ISSN 0749-4467 (print)/ISSN 1477-2256 (online) © 2009 Taylor & Francis
DOI: 10.1080/07494460903422362

collaboration schemes will transcend the present state-of-the-art from 'almost like being there' to 'better than being there'. New forms of presence will couple upgrades in raw network power and media fidelity with research in perception, synthesis and prediction.

Physics

Take the medium itself: just like in air, sound waves traveling between hosts on the Internet can bounce off edges, boundaries and obstacles. These reflections give rise to a configurable sound world of rooms with enclosing walls that contain networked and network objects which vibrate and produce sound. The world is entered from anywhere in the physical world connecting with a high-enough speed Internet connection. The relatively short time delays across next-generation Internets regionally make the apparent acoustical separation musically plausible. These path delays themselves can also be used to constitute network sound objects in a new breed of synthetic, distributed musical instruments. Recirculating echoes are used to create instrument tones whose pitches are in the musical range if fast enough. One can, in fact, 'play the network' as a waveguide instrument simulation stretched between San Francisco and Los Angeles and obtain a mid-range pitch.

Listening to the Music of QoS

One application uses these tones to monitor the quality of a network connection intuitively by ear. Just as someone might clap to get a sense of the size of a darkened room or knock on an object to know its rigidity, network users can sound their connections and listen to the vibrations that result. By plucking a 'network guitar', the network's quality of service (QoS) becomes audible. Network delay itself provides the string's 'memory' in a physical model whose pitch is a function of the time between two sites. The longer the sound takes to make it back, the lower the resulting pitch. And the more constant the tone, the better the QoS and the closer it is to ideal. An 'audio ping' in this form monitors QoS at a finer granularity than the traditional network 'ping' utility and in real time. The tones often exhibit an unusual pitch wavering due to changes in the speed of sound over the network.

Heading Underwater to Play Music

Sound propagation in the network differs from sound in air, along stretched strings or through other familiar media. Among its unique aspects are jittery arrival times of sound packet data and speed asymmetries in opposite directions over a given path. Where in physical media, distance-related delay affects signal intensity, spectrum and other qualities, in the Internet the sound remains the same even having traveled around the planet. These differences are significant for behavior in musical performance. The analogy that comes closest is from our experience with underwater acoustics. Entering into these different sound worlds with our ears, the properties of water or Internet media give them a sonic imprint all their own. We know very well the sound of the former and may soon become familiar with the latter. Network performances have begun to employ its unique qualities as part of the music, as in improvisations by the 'Net vs. Net' collective (Figure 1).

The slight discrepancies in 'now' that result from delayed sound within a distributed ensemble are significant for the choice of music. Synchronized rhythms generally de-synchronize as delay increases. One reason is the ambiguity of a perceived de-acceleration in the music. Depending on whether slowing tempo is an expressive inflection or an increase in network delay, a player reacts differently. We can observe the difficulties and formulate theories of what will work and what won't, but the behavior we expect in practical situations is not always what we get. For reasons not yet apparent, ensembles sometimes adapt to delays that should be impossibly restrictive for a particular music. Trying to explain ensemble behavior only from the mindset of network engineering, we find that a mechanical concept of 'now-ness' is insufficient and something quite apart from a musician's perceptions.

Raising questions about the human factor has become an important by-product of current online explorations. Similar to research and development in music synthesis, paradoxes arise that frame questions related to perception. Attempts to replicate a sound via acoustical analysis can uncover new principles of how we actually hear. Music made with distributed music ensembles and controlled lab experiments both have yielded paradoxical results that prompt new questions relating to time in

Figure 1 http://www.flickr.com/photos/aisenamalia/2803362699/in/set-72157606980783906/. Photo: Rodrigo Cadiz. ICMC 2008, Belfast, Northern Ireland. Net vs. Net Performance. Alain Renaud and Justin Yang on stage, Juan-Pablo Caceres remotely from Stanford, California.

performance and ensemble 'production'. These theories lead back to early work in time perception, especially phenomenology, and forward into advanced languages for media computing. What aspects are needed to model ensemble behavior and could such a model be used to engineer schedulers or artificial 'music agents' that implement our understanding of human temporal experience?

Time Perception: The 'Now' of Engineering versus Perception

Temporal order and event salience would be the first layer of such a model. Husserl investigated this with phenomenological techniques of introspection, but 'as soon as we make the attempt to undertake an analysis of pure subjective time-consciousness—the phenomenological content of lived experiences of time—we are involved in the most extraordinary difficulties, contradictions and entanglements' (Husserl, 1981 [1905]). A memory-less agent has no such difficulties, it is reactive only—but a useless model. The 'pure now' approach under-performs compared to human rhythm tracking (Gurevich et al., 2004).

A 300-millisecond integration interval serves in some theories as a window of 'now-ness', which can be treated 'as a space within time itself'. Two listener/performers coupled together, but with network delay intervening, respond in ways that may help shed light on the bigger questions: 'the two facets ... of time consciousness: on the one hand, the rich texture of the present, and on the other hand, the multiscalar hierarchy of temporal registers that underlies the flow of time' (Hansen, 2004). The acoustical qualities of the Internet change music because it changes the physics, but also refract differently our micro-time human abilities in a way that helps us understand them. Other factors certainly influence the picture; reverberation, ensemble strategy, phrasing, aspects of style—all come into play when learning why some things work and some things don't.

Where's There?

How close are we to adopting/adapting to this new medium? This may seem a bit of a stretch at this point in the game, but perhaps there will come a time when it seems less usual and even a bit special to congregate for music face-to-face. Music-making will take place increasingly in the new medium because general trends in communication run towards lower energy expenditure, higher content. Networked music performance does reduce travel and does seem poised to raise the 'channel content' (if we consider one's daily musical life as a channel)—similar to how email had already infested the early years of the Stanford AI Lab (CCRMA's first home). 'Electronic mail' was generally unknown in the 1970s but we were a small group who had access to such facilities. I remember a local newspaper story featuring oddities of lab life where the reporter noted how its denizens had an addiction to reading electronic mail first thing in the morning. Potentially, in the same way in which email has become a global energy-saving, content-increasing phenomenon, present network performance experiences are also a harbinger of change to come. After having worked in 'multi-site mode' for most concerts this past year, there was indeed a shift in returning to an entirely non-network event ('Hey, look—we're all in the same room!'). Concerning one recent network concert a critic wrote, 'what happened Tuesday night was about more than the music. It raised basic questions: What does it mean to "be here", when here is there, and there is here?' For that event, we were musicians and audiences co-located in Stanford and Beijing—6,000 miles apart (Figure 2).

Faster than Light in the Future?

Our signals flow through fiberoptic cables lacing the cities around the globe. At their transmission speed (roughly 70% of the speed of light), that is 190 milliseconds to circumnavigate it.[1] Given that some music is adversely de-synchronized by one-way delays on the order of 25 milliseconds (e.g., classical duos at medium tempo), will planetary-scale sessions ever be possible?

One experiment that is taking shape uses *prediction* to beat time delay. At least two groups (CCRMA and MIT) are in the early stages of demonstrating automata that can play their music slightly ahead of performers they are mimicking (e.g., mimicking a performer on the remote side). The basic idea is 'performance-guided synthesis' in which a very convincing model takes its cues in real time from a performer. In a more familiar application, this simply runs behind real time like an echo of the performer.

Figure 2 http://ccrma.stanford.edu/~cc/biz/StanfordBeijing.png. Photo: Enrique Aguirre. 29 April 2008. Pacific Rim of Wire at Stanford Pan Asian Music Festival, Dinkelspiel Auditorium, Stanford Laptop Orchestra (Slork) with group from Beijing University.

Pitch trackers driving synthetic voice doubling work this way. As we close the gap, getting closer to real time, this gets much tougher: pitch trackers do not have enough information or cannot run their algorithms fast enough to analyze the input. And if we go further, crossing ahead of real time, we need algorithms that can anticipate the pitch that will come next. If we can solve that, the 190-millisecond global transit time is not such an obstacle (at 60 beats per minute, that is less than a sixteenth-note). If our algorithm could predict the next note accurately, and synthesize it with convincing qualities, global lag would disappear.

Models for style, for piece, for performer and instrument are the ingredients from which these automata are built (Sarkar & Vercoe, 2007). An early study at CCRMA used a statistical model to predict the performance of a pianist after being trained on several weeks of rehearsal (Chafe & O'Modhrain, 1996). The soloist's raw timings and velocities were captured via Disklavier for a passage in the score being studied. The model was built around expressive nuances that varied from performance to performance. It was sufficiently explanatory of the performer's habits to be able to predict the next note's timing and velocity in a live performance context (for that study, it was guided in real-time by a listener creating a performance via a haptic device). Such real-time style-piece-performer-instrument models are becoming more common in computer music, apart from their deployment across a network. Recent improvisations by composer/instrumentalist Roberto Morales[2] have shown the expressive musical power that can be gained from statistical models running as real-time predictors (Figure 3).

Figure 3 http://ccrma.stanford.edu/ ~ cc/biz/GuanajuatoStanford.jpg. Photo: June Holtz. 21 October 2008. Encuentros Simultaneos at Festival Cervantino, Guanajuato, MX. Outdoor stage, Roberto Morales, piano, Roscoe Mitchell, saxophone, Chris Chafe, celletto, with CCRMA, Stanford University, Fernando Lopez-Lezcano, synth, Juan-Pablo Caceres, synth.

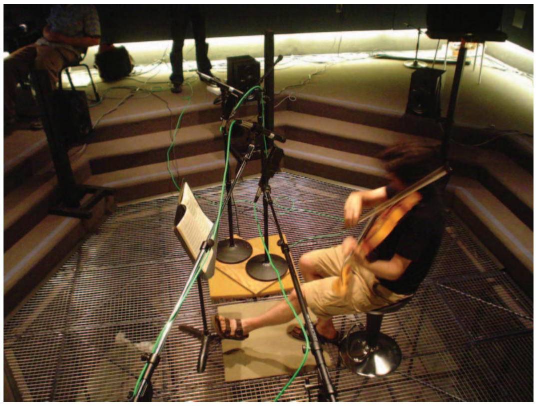

Figure 4 http://ccrma.stanford.edu/~cc/biz/GuanajuatoStanford.jpg. Photo: Juan-Pablo Caceres. 20 June 2006. Test of Classical Repertoire, Mozart Gm String Quintet (K 516) at The Banff Centre, Alberta, St. Lawrence String Quartet with CCRMA, Stanford University, Barry Shiffman, viola.

Conclusion

Conjoined stages/spaces, acoustical media with new properties, artificial intelligence running in the connections, these cannot but help spur new musical approaches. For the present, we learn when we contrast the results (of networked performances) with known scenarios. A typical setup reproduces the ensemble's acoustical arrangement, inserting loudspeakers at the locations where a remote 'phantom' player should be (Figure 4).[3] Having ensembles involved who are able to describe what they hear and feel is essential to gaining new insights.

One of these 'uncanny' moments happened in a trial between Norway and California, with a one-way delay of over 100 milliseconds. The score indicated an accelerando. This focused the ensemble's attention away from the difficulties of playing synchronously under such temporal separation and onto the shape of the phrase. The result was a perfectly synchronized arrival at the end of the phrase. The trajectory was supported by the speedup in tempo. Longer time shapes (from phrasing, arrivals, cadences, etc.) seem to have less trouble synchronizing, just as slower rhythms also seem to be less affected by difficult delay conditions. We learned in that moment that despite frustrating rhythmic deterioration, concentrating on the phrase seemed to be a way to keep the music very much together. I am likely to use that in a piece ...

Notes

[1] The mean circumference of the earth (m)/fiber speed (m/s) is 40041470/(299792458 * 0.7).

[2] I have performed numerous concerts with Morales and Morales' 'robot' pianist—an accompanist to my cello improvisations. http://ccrma.stanford.edu/~cc/pub/wav/stereo/chafeMoralesDuoBanffRolston13-Mar-2009.wav.

[3] Soundclips of the St Lawrence String Quartet participating in trials between The Banff Centre and CCRMA can be found online at: http://ccrma.stanford.edu/groups/soundwire/research/.

References

Bacon, F. (1626). *New Atlantis*. Whitefish, MT: Kessinger Press.

Chafe, C. & O'Modhrain, S. (1996). Musical muscle memory and the haptic display of performance nuance. Paper presented at the International Computer Music Conference, Hong Kong.

Gurevich, M. et al. (2004). Simulation of networked ensemble performance with varying time delays: Characterization of ensemble accuracy. Paper presented at the International Computer Music Conference, Miami.

Hansen, M. (2004). *New philosophy for new media*. Cambridge, MA: MIT Press.

Husserl, E. (1981). The lectures on internal time consciousness from the year 1905. Trans. J. Churchill. In P. McCormick & D. Elliston (Eds), *Shorter works*. Notre Dame, IN: University of Notre Dame Press.

Sarkar, M. & Vercoe, B. (2007). Recognition and prediction in a network music performance system for Indian percussion. Paper presented at the International Conference on New Interfaces for Musical Expression (NIME 2007), New York City.