

Sounding the Net: Recent Sonic Works for the Internet and Computer Networks

Peter Traub

This article examines a number of recent Internet- and network-based audio pieces with a focus on their compositional underpinnings, and the effect of the Internet and networks on the creative process. The process is discussed in excerpts from online interviews with the creators, and the technical and aesthetic goals of the pieces are described. The article concludes with a discussion of the author's own recent experience composing groundloops: for solo percussion and internet—as well as an attempt to find areas of intersection and differentiation among the approaches discussed.

Keywords: Network Music; Internet Composition; Interactive Computer Music; Collective Composition; Web-based Music

It is impossible to say, of course, what music on the Web might be like by then, or even how music may be delivered. The trip has just begun; it is too soon to know the destination. The new paradigms of Web music, like all new art forms, are self-defining, and the Internet will dictate its own styles and forms. (Duckworth, 1999)

Introduction

The Internet has now reached some degree of maturity, or at least adolescence, and Internet art can no longer simply survive on the novelty of being Internet art. While new media artists were attracted to the Internet from its early days, the growing proliferation of Internet-based artworks over the past decade has been dramatic. Some use the Internet as a medium for collaborative music making; some use it as a source of material for generative or algorithmic works; and some use the technical idiosyncrasies of the Internet as an aesthetic focus in their pieces.

While researching this article and considering the topic of Internet composition, I decided to treat the Internet as a large-scale computer network—the largest available. With this in mind, it seemed unreasonable to ignore pieces that operate using smaller

networks, when the commonalities between those pieces and Internet-based pieces are generally much greater than their differences. Thus, I decided to look at any music in which a computer network plays a primary role, and have focused on several recent pieces for this purpose.

The core of this article revolves around a number of E-mail interviews with a variety of artists, though of course this is only a small sampling. I discussed the artist's pieces with them, as well as their thoughts about larger considerations of composition, form and structure within a network environment. Since including the full interviews is beyond the scope of the article, I have created an online appendix that contains the full interview text for each of the interviewees. The appendix site is intended as a companion to this article, and can be reached at: <http://www.fictive.org/cmr/appendix/>.

Jon Thomson & Alison Craighead, *Telephony* (2000) and *Unprepared Piano* (2004)
<http://www.thomson-craighead.net/docs/telf.html>
<http://www.thomson-craighead.net/docs/pianof.html>

For over ten years, Jon Thomson and Alison Craighead have been collaborating on new media and electronic works ranging from gallery installations to online art. While most of their work utilizes visual and audio media, the two works chosen for inclusion here were of particular interest for their sonic components.

Unprepared Piano (2004) is an installation work for solo Disklavier connected to the Internet. The Disklavier is connected to a computer that is constantly searching for MIDI files. A kicked-over piano stool and a computer monitor sit by the base of the piano, with the monitor showing status information on the currently playing file. As the files are found and downloaded, a program running in Max/MSP processes them for playback on the Disklavier. A number of random operations determine which MIDI channels get played, for how long, and how much the playback accelerates or decelerates. As the Max patch has no way of knowing how the instruments of the MIDI file are assigned or what type of music the file contains, parts that are scored for drums or horns or guitars end up getting played through the piano. In the end, a MIDI file will be played back in chunks, with random combinations of channels and instruments chosen for playback, and while the resulting 'music' is quite distinct from how the original score is likely to sound, certain dynamics and traces of the original's structure remain (Thomson, 2005). The piece in both title and form has clear reference to Cage:

[T]he Cage reference is implicit. The title is both a joke (the idea of the piano being unprepared for what it is instructed to play) and a conscious connection to the idea of prepared pianos—it's just that in this instance, the preparation is by means of software—a set of rules that characterize the way the piano adapts any rendition of any given 'found' MIDI file. . . . We chose a MIDI grand piano precisely because we wanted to pit the generative playback against the authority such a traditional acoustic instrument maintains. We wanted this authority to goad listeners to 'read' these musical works within that context. (Thomson, 2005)

While *Unprepared Piano* is a network music piece, it is not a network piece in the same sense as pieces by The Hub. It is a piece that uses the Internet for a very specific purpose: as a data repository.

[T]he web in this instance is a database or library of scores that the computer picks up on and draws from. We like the openness of this connection, and that the MIDI files are 'naturally situated' online, but in our minds it's not the most important aspect that it is the net, but rather that the piano is connected to a very large body of data, from which it selects and performs. A large enough body of information to make the permutations of rendition feel almost endless. (Thomson, 2005)

It is also important to note that Thomson & Craighead consider the piece a prototype for later works in which they combine the live Internet-sampling aspect of *Unprepared Piano* with live video from the World Wide Web to create short, algorithmically generated movies in a form they call 'template cinema'.

Telephony (2000), another work by Thomson & Craighead, utilizes the mobile phone network to create a gallery-based installation. A grid of 42 Siemens mobile phones are glued to a gallery wall. The phones all ring with variations on NokiaTune, as composed and altered by Thomson & Craighead. Anyone inside or outside the gallery may dial in to any of the grid phones. A receiving phone will then call-forward to another phone in the grid, setting off a cascading chain of ring tones, until the call-forwarding ends about five to seven iterations deep. Gallery visitors may also interact with the grid by dialing out from its phones to any other phone in or out of the grid. While the grid installation is live, background 'elevator'-style music is also played into the space, with that music being an improvisation on NokiaTune.

As visitors or outside callers dial into the installation, the harmonic density of the piece increases as cascading ring tones layer upon each other with each call-forwarding iteration. The more calls that come into the grid, the denser the music becomes. While the piece undoubtedly has a strong sonic component, Thomson and Craighead's visual art background informs their thinking about the importance of the musical aspect of the piece. While composing the ring tones was important, the overall effect of the layered sound is not the aesthetic core of the piece:

We just composed them ourselves trying to take certain things into account—i.e. keeping in key, keeping the variations related to the structure of the original and making sure it related well to the background music. Because the order and timing and layering was beyond our control we kept them simple and restricted ourselves to fairly brief loops. However, in *Telephony*, the music is not what's important, it's about situating phones in a gallery and scrutinizing the way we behave towards them, and how the live network is instrumental to how the work operates, and how people experience it. It's a kind of invisible drawing in our minds. ... We enjoy the invisible material presence of the nodal points that can be identified within the networked structure of a piece like telephony and how that implies the linear connections between the points as a drawing or diagram. (Thomson, 2005)

While the nodal points of the *Telephony* grid interact primarily with gallery visitors, the installation also interacts with people outside of its physical space. This occurs when calls are made from the grid to phones outside the gallery. These intersections of the virtual network space and the real space also have interesting and unintended side effects:

[C]all registers can create odd effects. For example, if someone dials out from the grid by pressing the call button twice, it calls the last phone that called it. So we had instances where people who had used the work at the opening, let's say by dialing in on their own phone, would keep being called by the installation! Worse than that even, the second time we staged the work (some months following the first showing) we forgot to delete the call registers, so people who had visited it first time around were called again. (Thomson, 2005)

The idea of a piece existing in both physical and virtual spaces, with the virtual space sometimes colliding with the physical space at inopportune moments is compelling. It is as if the work has a ghosting effect, much like an afterimage burned on one's retina. In *Telephony*, this effect takes the form of communications from the installation long after one has left the physical space of the gallery or the mental space of engagement with the piece.

Golan Levin, *Dialtones: A Telesymphony* (2001)

<http://www.flong.com/telesymphony/>

Golan Levin has had a deep interest in electronic music since high school. While his work has a strong visual component, music and sound are a significant focus in several of his pieces. *Dialtones: A Telesymphony* (2001) is an audiovisual work for an 'orchestra' of mobile phones. In preparing for a performance of the piece, audience members are asked to register their mobile phones at special computer kiosks. During the registration process, an audience member gives the kiosk their mobile phone number, and the kiosk in turn uploads a custom-composed ring tone to their phone, if the phone can receive such data. The kiosk then assigns that audience member a seat in the performance space.

When the piece begins, the performers, including Golan Levin, Scott Gibbons and Gregory Shakar, are on stage controlling custom software to dial out to the phones in the audience. A database stores all the registered phone numbers and the locations of their owners in the audience so the performers have control over which phones to dial and when. Layered on top of this system is a performable graphical interface that shows each audience member as a small circle in a grid, or 'audiovisual pixel' as Levin refers to them. A digital projector projects this grid down onto the audience such that when Levin selects a phone to ring, that circle or 'pixel' is illuminated in the graphical interface as well as in the audience. A large mylar mirror, placed at an angle above the audience, allows members to see the entire audience reflected from above. This lets audience members see who is illuminated in the grid, and thus whose phone is ringing.

To some extent during its development, I began to think of the *Dialtones* telephone network as a very large polyphonic synthesizer, albeit one with a lot of unpredictable quirks (especially with regard to latency). And each of the audience's phones were voices or individual oscillators in that large synth, and my job was to play the instrument by clicking on the right notes on its keyboard at the right time.

I say I 'began' to think of the phone network as a polyphonic synth, but I certainly didn't end that way. My concept of the instrument changed entirely on the night of the first performance, when we were finally able to bring a live audience into the situation. (Levin, 2005)

Musically, *Dialtones* is divided into three sections. The first section entails the dialing of the audience's phones from the stage. Up to sixty phones can be dialed at any given moment, so the sonic layering can be quite dense. In the second section, a performer (usually Scott Gibbons) performs solo with six mobile phones on stage. In the final section, the stage phones and the audience phones are performed together.

Levin conceived of the work as a conceptual sound-art piece, and determined that a piece in which the phones just rang randomly would not be very interesting.

Dialtones was always, to begin with, a kind of sound-art piece or conceptual performance artwork. I say this because the project originated from a pure concept (that of performing the audience's mobile phones), and was motivated by a curiosity to discover what it would be like—sonically, visually and socially—to experience such a concept. In this sense, I don't think it's too much to say that the project conformed well to John Cage's definition of experimental music as music that 'initiates sonic processes the outcomes of which are not known in advance'. The problem with Cage's definition, though, is that it suggests that it wouldn't have mattered whether or not the results reflected any human patterning, or that we oughtn't intervene in some way to ensure an interesting outcome. I think if *Dialtones* just sounded like a pile of 200 phones ringing on and off randomly for half an hour, people would have been really profoundly disappointed. For the project to succeed, it was necessary for us to demonstrate that we could actually tame this enormous and unruly beast—the mobile telephony network of Upper Austria—in order to bend it to more musically structured ends. (Levin, 2005)

In bending the telephony network to more structured ends, Levin encountered a number of technical constraints that affected the parameters of the composition. Latency, an issue that is common through almost all of the works discussed here, had a direct effect on the interaction between the dialer/performer on stage and the ringing phones in the audience. Primarily, there was a several second delay between the selection of a phone for ringing and the initiation of that ring, making perfect synchronization impossible.

The telephone network imposed an unpredictable latency between the time that we dialed a phone, and the time that the requested phone would begin to ring. We did some experiments and determined that the average delay was 4.74 seconds, with a standard deviation of about a second or so. In some cases, particularly when we dialed international numbers, the delay could be as long as twelve or thirteen seconds. This fact had serious compositional consequences, musically speaking, since it meant that we couldn't create precise synchronizations between rhythmic

ring tones. It also meant that any chord progressions would have to play out over a fairly long timescale in order to be reliably perceived. We ended up composing ring tone melodies which all shared the same tonal center—I think it was A-880—and adopted a more textural approach to compensate. (Levin, 2005)

The central problem of latency Levin experiences is similar to the latency issues faced by the other composers discussed here, and seems to be a central issue in much current network music. Yet latency is not necessarily viewed as a problem or a negative in this sense. Rather, it is seen by many as a property of the environment that must be engaged. As Levin states above, chord progressions needed to play out over a longer period of time in order to become apparent. In this way, the environmental qualities of the network implied a different set of priorities for the composer. In *Dialtones* and other pieces discussed below, the primary means by which latency is approached is through a de-emphasis on synchronization and rhythm, and a re-emphasis on texture and large-scale temporal changes.

Chris Chafe, *Ping* (2001) and the SoundWIRE Project

<http://ccrma.stanford.edu/groups/soundwire/ping.html>

<http://ccrma.stanford.edu/groups/soundwire/>

Chris Chafe, at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University, has done considerable research on network latency from both technological and compositional perspectives. Chafe's work on the SoundWIRE (Sound Waves on the Internet from Real-time Echoes) project since 2000 led to the creation of his network piece, *Ping* (2001). SoundWIRE was a research project funded by the National Science Foundation in which Chafe and his fellow researchers looked into ways to sonify network latency times as a means of measuring Quality of Service (QoS). The SoundWIRE technology allows one to circulate audio data between two points on a network, with the resulting sound being an indicator of the network's latency. Much like a plucked string, the higher the latency (i.e., the slower the Round Trip Time (RTT) of the packets), the lower the pitch of the resulting sound, and vice versa.

[T]he sound comes 'naturally' from sound waves created across the network. The audio stream is looping around a network path (which has a delay). It's as if the network were a real pitched instrument like a stretched string, where the pitch depends on the delay time for an excitation to loop around on the string medium. Instead it's done over the network by designing a physical model circuit (distributed over the network) and using the network delay in place of the computer memory that's typical for physical model delays. (Chafe, 2005)

Other than using audio to evaluate network QoS, the SoundWIRE project also experimented with network-based performances in which ensembles at different locations were connected via a very high bandwidth, low-latency network connection, using Internet2. These 'telepresence' experiments enabled streaming video and multi-channel audio to be shared between performance locations, and it is an area that Chafe continues to work in.

Ping is a sound installation that gets its title from the network testing tool of the same name. With the 'ping' application, network packets are sent to a remote server simply asking for a response. The RTT that results is commonly referred to as the 'ping time'. In *Ping*, the RTT to multiple servers is sonified. The installation premiered at the *010101: Art in Technological Times* exhibit at the San Francisco Museum of Modern Art in 2001. It consisted of eight speaker towers configured in a ring. The computer driving the installation sent pings to remote servers, and the RTTs for each server were sonified and played through the towers. The towers were situated outside on a balcony, and if the listener stood in the middle, he or she would hear pitched 'pings' of sound coming from each tower. The pings were tied to multiple remote servers being simultaneously pinged. The overall effect was a layering of plucked string-like pulses of different rates and pitches surrounding the listener in a multi-channel environment. An interface consisting of a horizontal metal wheel in the middle of the space allowed visitors to move the sound around the space, affecting its placement, pitch and timbre.

Chafe's interest in making imperceptible aspects of a particular environment perceptible led to his work with SoundWIRE and *Ping*. In the case of networks, the underlying nature of the data flow is not perceptible to most users without the use of specific tools.

Ping began a thread in my work with Greg Niemeyer that we've come to call Extrasensory Perceptions (there's a CD-Rom by that name with *Ping* and Oxygen Flute). The concept is simple: we can translate sensing data into sensory experience and make normally hidden aspects of the environment perceivable. Network flows and the carbon dioxide/oxygen exchange have been it, so far. That gives the substrate for these pieces, but it's finally about composing up from that to the surface. Since these are real-time systems, they become like interactive improv with the data flows. (Chafe, 2005)

Chafe's interest in networks stems from the ways in which they hover between predictability and unpredictability. This semi-predictable nature is also evident in Levin's *Dialtones* and Thomson and Craighead's *Telephony*. In *Dialtones*, the performer knows that a phone is likely to ring around a given time, but they do not know exactly when: the short-term ordering of the sonic texture is unpredictable in this way. In *Telephony*, one knows there will be a call-forwarding cascade if one calls into the grid, but not how that cascade will play out. With *Ping*, one knows that a pinged server will generally respond at a regular rate, but network traffic could affect the ping time, or the server could be down and not respond at all. This mix of both regularity and chaos is what Chafe and a number of others who work with network-based art find alluring: 'If we had perfect networks they'd sound regular and boring. The other extreme would be totally random busted networks. In between, where I find other systems to mine for music as well, is this place that's got predictability and surprise in varying balance' (Chafe, 2005). The approach Chafe describes above applies to a number of net-based pieces, including the ones described so far, in which the play between predictability and surprise forms a fundamental structural layer.

Chris Brown, *Eternal Network Music* (2003)

http://www.transjam.com/eternal/eternal_client.html

Chris Brown is among the early pioneers of network computer music. In the 1980s, he was a member of the network music group, The Hub. Consisting of John Bischoff, Chris Brown, Scot Gresham-Lancaster, Tim Perkis, Phil Stone and Mark Trayle, The Hub represented a technological step beyond its predecessor: The League of Automatic Music Composers. The League and The Hub performed live improvisations with multiple computers in various networking configurations. In League performances, ad-hoc network connections were created between individual computers, requiring much preparation time. The Hub's name was derived from the newly included use of a central computer acting as a network hub connecting all the computers. This central hub operated as a type of data mailbox, in which members of the group could deposit control data during a performance. The other members could access that data at their discretion. This configuration enabled a form of live collaboration in which each member of the group had a hand in the musical processes of the other members, such that no single performer had complete control over their output. The ways in which the network and data passing relationships were constructed varied from piece to piece, but the construction of those relationships was central to structure of The Hub's pieces (Brown & Bischoff, 2002).

During his tenure in The Hub, Brown participated in what are probably the first known instances of wide area network computer music. The initial Hub concert, held in 1987, took place in New York and connected six performers divided between two venues joined by a 300-baud modem network. The resulting sounds on either ends of the modem connection were not the same, although data was shared between the two spaces (Brown & Bischoff, 2002). Similar to Chafe's interest in the predictability/unpredictability balance exhibited by complex networks, Brown, both in his work with The Hub as well as his more recent work on *Eternal Network Music* (2003), focuses on the ways in which that complexity can be mined for artistic ends.

I came into electronic music from the Tudor/Mumma tradition where the object was always to bring the system/instrument to life, which meant to foreground and celebrate the surprising and unexpected aspects of its behavior. It's always been for me about a balance between in-control and out-of-control, with the performer as instigator and navigator of a electroacoustic space—most of the time in non-network musics, this is very much about the physical space itself, which electronics allows you to explore in very different ways than acoustic instruments do—with network music, the system itself takes over that spatial dimension. (Brown, 2005)

Like Chafe, Brown's interest is not in having a perfect network, but in having a system that provides for multiple interesting outcomes given a certain set of parameters. The artists discussed here explore this characteristic of network music in varying ways. Some, like Jason Freeman, explore it through the failure of an inaccurate search algorithm, albeit a failure in a positive aesthetic sense in which the output of the failing network is considered more interesting than the output of one

that does not fail. In Tim Perkis's Hub composition, *Waxlips* (1991), bursts of MIDI data from a lead performer were sent over the Hub network, triggering clusters of events and creating an audible change in sonic density as MIDI messages were dropped, or failed to transmit between Hub performers. *Waxlips* exposed potential failings of the Hub MIDI network, albeit failings that ultimately were the result of programming errors. While the Hub did not compose directly to expose the potential failures or shortcomings of the network, they were nevertheless interested in them: 'All of us were always interested and accepting of the ways in which they crept into the music. We never expected or wanted the network to be a perfect vehicle for the data-sharing ideas, but were interested in the interactive collision of imperfect instruments with the listening and taste of the performers' (Brown, 2005).

Latency, much like the inaccuracy and data loss issues mentioned above, is a network characteristic that can be approached in a number of ways, one of which is as an interesting point of failure. As discussed earlier, latency is perhaps the most common issue composers of network-based music encounter, especially if it is real-time music. Whether or not to call latency an issue depends on the artist and, to a more specific degree, the nature of the piece that encounters it. In Chafe's work with the SoundWIRE project, it was dealt with as an issue to be overcome or minimized to the greatest degree possible. With *Ping*, however, latency was a necessary condition for the work to function. Others see it not as an asset or an obstacle, but as a quality of the environment that must be taken into account in the structuring of a piece.

For computer musicians a slow-response time is tolerable because we're mostly starting and stopping processes rather than trying to inflect momentary gestural activity. My solution to that limitation is just to add those gestures locally, and not worry that the music is different here than it is there. . . . In short, latency is always present as an aspect of this music, and you design different pieces around how you're planning to handle it. (Brown, 2005)

In *Eternal Network Music*, latency is not a critical issue. Like Max Neuhaus's *Auracle*, it is an online interactive installation that allows people of any musical or non-musical ability to jam together using a Java-based interface. *Eternal Network Music's* synthesis engine was developed using Phil Burk's JSyn software, a Java-based synthesis toolkit geared toward online and network-based sound. When entering the installation, a user visits the *Eternal Network Music* website, signs on as a user and is presented with the Java-based graphical interface (see Figure 1). The interface consists of a black rectangular space, populated by several small circles of different colors. Each circle represents a synthesis voice, and the user/performer grabs them and drags them around the space to change their amplitude, pitch and timbre. It is possible to play the piece solo, but more fun to play with others. As other people log in and play, the user will see the colored circles moving independently around the space as the other participants drag them about. Participants can drop the circles they are moving and pick up inactive ones at any point.

The circles always generate sound, even when they are dropped and inactive, thus the 'eternal' nature of the music. Brown's compositional concern in developing the

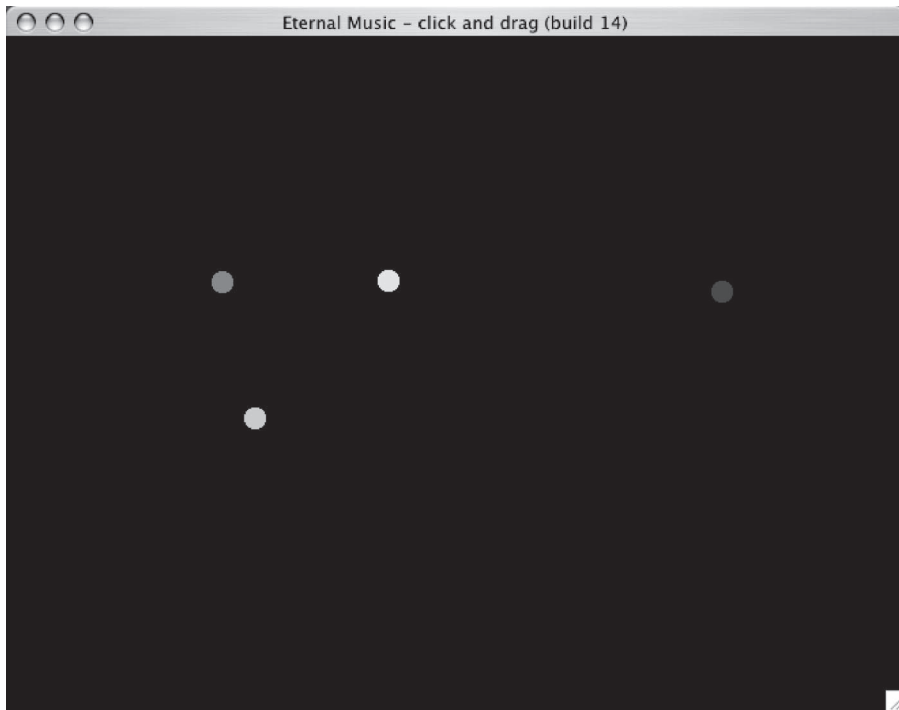


Figure 1 The interface for Chris Brown's *Eternal Network Music*. Each circle controls a different synthesis voice.

piece was, like many interactive works directed at the general public, focused on creating an intuitive environment in which the exclusivity of musical performance is opened up to non-musicians.

The idea was to make the simplest possible gestures result in a rich sonic texture. I didn't want the players to have to feel like they needed to learn how to control the music, but that they could explore gesture and response intuitively, perhaps discovering the simple linkages behind the instrument. This is one of those pieces where the instrument IS the composition—the sound it makes is very specific and could result in many different kinds of performances, but I think that the identity of the piece is always pretty clear. (Brown, 2005)

Like Neuhaus's *Auracle* (discussed below), *Eternal Network Music* does not transmit any audio data between players. Rather, it transfers control information between players that is used in each user's local synthesis process. This saves a tremendous amount of bandwidth and makes the work more accessible to users without high-speed connections. As stated before, latency is not as much of an issue in this piece as each player only hears the output of their own synthesis engine. However, as there is still some latency during the transfer of data between players, it is difficult to synchronize gestures. In playing the piece, one seems to focus on larger, less time-sensitive gestures.

In allowing users to have so much control over the work, *Eternal Network Music* joins a class of interactive pieces often termed ‘sound toys’. While many who compose for networks or the Internet do so from a more performance-privileged position, in which an entry barrier of performance or technical ability is necessary to participate, there is a thread among Internet composers involving the egalitarian nature of the medium to include non-musicians in the music-making process. Most commonly, these efforts are expressed through the creation of interactive works like Brown’s. In fact, soundtoys.net (<http://www.soundtoys.net>) is a website dedicated to this endeavor. Within these efforts, artists like Brown are also concerned with finding uses for the network other than just commerce and email:

Networking technologies are getting easier to employ, but there still hasn’t been much success in getting people to play music with each other in the data networks in which they increasingly spend more and more of their lives. This is sociologically problematic—if our social lives take place within data networks, we need music to be part of it, and trading sound-files is a very primitive form of interaction! (Brown, 2005)

Max Neuhaus, *Auracle* (2004)

<http://www.auracle.org>

Perhaps the first known example of network music is Max Neuhaus’s *Public Supply*. Composed in the mid-1960s, *Public Supply* used a local New York radio station and the public phone network to create a two-way aural space in which listeners were encouraged to call in and add their sounds to the space. Neuhaus mixed between ten phone lines coming into the station, and listeners were asked to keep their radios on in order to create feedback effects. In the 1970s, *Radio Net* built upon this concept by using National Public Radio (NPR) and the public phone network to create a 1,500 mile-long and 3,000 mile-wide feedback-based sound processor. *Radio Net* invited listeners to whistle into their phones to contribute to the nationally broadcast live composition. Neuhaus controlled the piece from a custom-designed mixer at the NPR studios in Washington, DC (Neuhaus, 2005).

These works can be seen as precursors to the current crop of network-based music and network-based feedback pieces. In many ways, the aesthetic considerations are very similar. These pieces purport to be egalitarian, using publicly accessible resources to invite musicians and non-musicians alike to participate in the music-making process. They also use the public resources of the phone and radio networks towards artistic ends for which they were not intentionally designed. Both the radio and phone networks were and are used primarily as a means of transmitting communications, bi-directionally in the case of the of the phone network, and uni-directionally in the case of the radio network. By combining these two networks, Neuhaus created a hybrid network topography in which any participant could put information into the network, and all participants could get that information out of the network. When Neuhaus proposed the creation of *Auracle* (2004), an interactive Internet-based instrument that

lets participants play with their voices, the goal of the work was much the same as with his earlier network pieces: to give the public access to the collaborative music-making experience, and in so doing, encourage ‘new ways of thinking about music’ (Neuhaus, 2005).

Like Brown’s *Eternal Network Music*, *Auracle* users jam together through the interface; except in this piece the voice is the primary controller. The technical requirements are thus slightly higher, as a working microphone is needed as the controller input. Once a participant engages the instrument, a moving waveform display shows their vocal input as mapped by the synthesis engine (see Figure 2). A map of the Earth at the top of the display shows where participants in the current session are located. As users experiment with vocal sounds, *Auracle* outputs synthetic tones based on the mappings of the voice to various control parameters. *Auracle* does not actually use the vocal sound itself as part of its audio output; rather, the voice is used as a controller for a JSyn-based synthesis engine. This is particularly useful for sharing data between

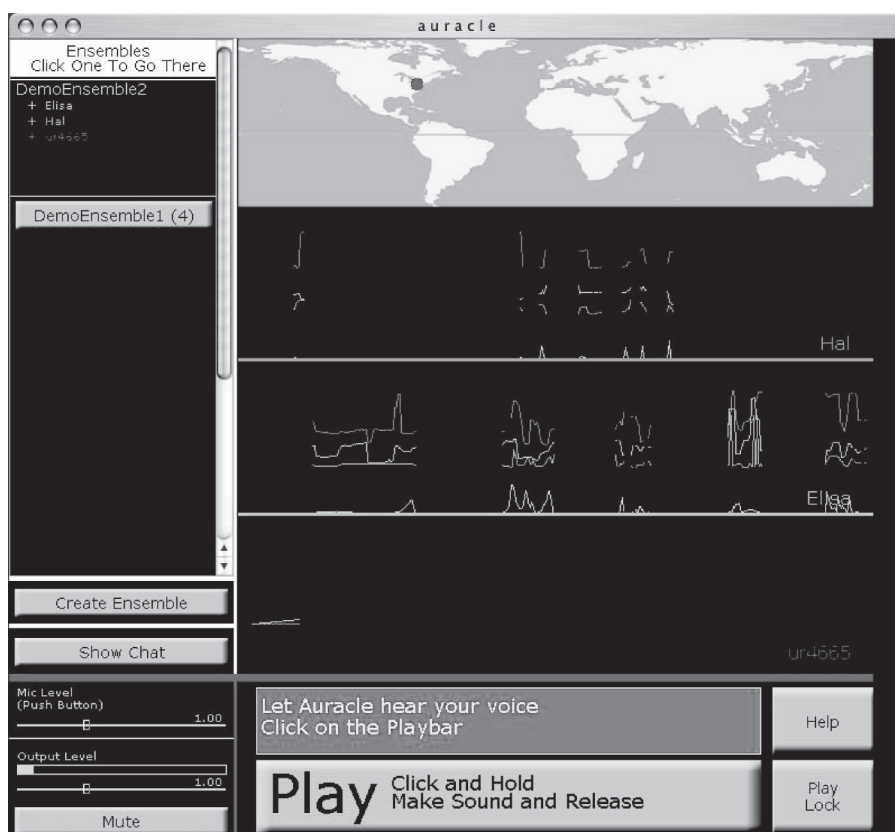


Figure 2 The *Auracle* interface, showing control and waveform data from three different users.

players, as audio data does not need to be transmitted over the network. Only control data is transmitted, requiring far less bandwidth. In choosing to use vocal sounds as the controller input for *Auracle*, Neuhaus desired control data with a great degree of nuance that all players could produce, and the human voice fit this criteria:

I wanted to give lay players the most facility to control the synthesis of sounds. Our almost built in fine control over inflection does this. Preventing conversation also moves *Auracle* beyond a forum for just talking and opens it up to one for inventing musics. The fact that verbal language is no longer there also surmounts the barriers to making it cross cultural. (Neuhaus, 2005)

In creating *Auracle*, Neuhaus wrote a constitution for the piece that is intended to guide its future development. One amendment to this constitution is that the voice will always be the source of control data. The synthesis engine, however, has been opened up to third-party development, so that future software additions could radically change the way the instrument sounds.

Our opening up the design of its instruments to third parties is the first step in this direction. *Auracle's* constitution is the only thing that is fixed. It defines what *Auracle* is and keeps it distinct. How it is actually implemented I hope will always evolve with future advances in technology, of course, but most importantly with new ways of thinking about music. (Neuhaus, 2005)

Even with the opening of development to third parties, another amendment specifies that all of *Auracle's* sounds will not sound synthetic, such that they 'must convince [one] that they could actually be made by objects in the natural, non-electronic world' (Neuhaus, 2005).

Auracle, like *Eternal Network Music*, also encounters the latency issue. It is currently not possible for multiple players to closely synchronize their gestures, and so a focus on other aspects of the sound and interaction takes precedence. Neuhaus views the latency factor as inherent to the nature of the system, but not necessarily as a compositional barrier:

When one begins working with any new medium it is important to understand what its intrinsic strengths and limitations are. For me, it makes no sense to try and make it do things outside its nature. For example the tight synchronization in time of traditional music is not available on the Internet today. This does not mean that one cannot make music with it. It just means one has to think about possibilities for music without this particular parameter. (Neuhaus, 2005)

For Neuhaus, the network offers the possibility of a cultural common ground: a place where people can meet and engage in communal activities that have, in large part, been subjugated by uni-directional mass media such as television or radio. The network allows for the creation of a communal space in which players engage in the type of music-making in which participation and community take precedence over exclusivity and virtuosity:

My Networks propose the self-evolution of new musics. Their premise is a form of music making which remains now only in societies untouched by modern man. Rather than something to be listened to, music in these cultures is an activity open to the public at large—a dialogue with sound rather than a performance. I believe this to be the original impulse for music in mankind. The Internet and the broadcast/telephone networks I invented previously offered me a contemporary means of reinstating this practice. (Neuhaus, 2005)

Jason Freeman, *N.A.G. (Network Auralization for Gnutella)* (2003)
<http://turbulence.org/Works/freeman/>

Jason Freeman, besides being a primary developer for *Auracle*, has also created his own intriguing network-based installation. *N.A.G.* (2003), or *Network Auralization for Gnutella*—a generative software installation that uses the Gnutella file-sharing network to create sound collages. Unlike the other interactive pieces discussed above that use browser-based software to provide the interface, *N.A.G.* comes in the form of a downloadable desktop application. When connected to the Gnutella network, the application offers the user the ability to search for any term, much like they would search for an audio or video file through more utilitarian sharing applications.

Once a search term is entered, the application tries to pull in results of the search that appear to be mp3 files. The user can choose whether or not to manually select which search results are downloaded, or *N.A.G.* can decide. The search results are matched based on the file name and metadata. Mp3s begin to playback in the application shortly after they start downloading, meaning that one starts to hear sound well before file downloads are complete. At this point, a number of choices are given to the user, all configurable in real-time. The user may choose how many songs *N.A.G.* tries to play at once (two to eight), how much to vary the playback speed and volume of the songs, whether or not to loop the currently downloaded material before a song download is complete, and how quickly to move among downloading songs. For example, ten songs may be downloading while *N.A.G.* skips around playing four songs simultaneously. Apart from the optional looping, speed and volume changes, *N.A.G.* does not perform any other processing on the files (see Figure 3). This is partially due to Freeman's concern over processor limitations, but also because he wants users to clearly hear the downloaded material in order to appreciate the juxtapositions that occur.

Freeman's piece deals with two network issues. One is the ability of the network, and the almost limitless data resource that the Internet has become, to provide for serendipitous events and experiences through semi-random sampling of the data. The other is the use of the Internet as a file-sharing medium, often in breach of copyright laws. *N.A.G.* engages both of these issues, but focuses more on the serendipitous event through the chance combinations of various songs. Freeman considers serendipity an important part of both his instrumental and electronic work, but thinks about it differently depending on the compositional context:

I think that serendipity plays a role in just about everything I write, but it manifests itself in a very different way in pieces like *N.A.G.* than in my instrumental music.

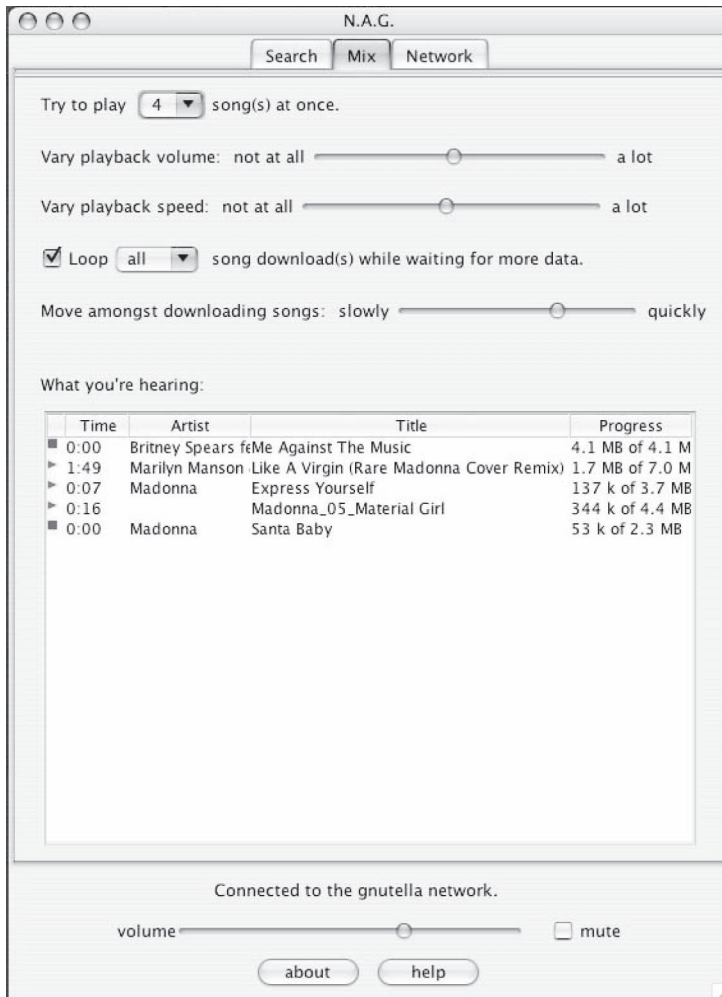


Figure 3 The mix screen of Jason Freeman's *N.A.G.* after a search for 'madonna'.

When I'm 'through-composing' instrumental music, chance relationships between sounds play themselves out during the compositional process, not during the performance. ... On the other hand, serendipity in *N.A.G.* plays a role not so much in the development of the software (i.e., 'composition') as in the moment of its use (i.e., 'performance'). (Freeman, 2005)

The serendipitous possibilities offered by the Internet search process were one of the factors that motivated Freeman to create *N.A.G.* The inaccuracy of the search process leads to situations in which one can 'search for "madonna" and end up not only with "Borderline", but also with "Lady Madonna"' (Freeman, 2005).

I really wanted to tap into the fascination with network searches that I (and just about everyone else) had upon first discovering the World Wide Web. But since the Web is dominated by commercial outfits these days ... I turned to a network (Gnutella) which is not dominated by commercial interests, and which doesn't have as effective search tools, as a way to revisit the excitement of serendipity in network searching—where the process is a goal in and of itself. (Freeman, 2005)

Freeman's piece also engages notions of cultural appropriation for artistic use in the way it accesses a public network to acquire files that, according to the Recording Industry Association of America (RIAA), are illegally shared. When end users download free music through Gnutella, copyright owners generally consider the act copyright infringement. When *N.A.G.* downloads these songs, the issue becomes a little more complicated. Because the songs are downloaded as the process of a public, non-profit art work, whether or not that act is a violation of copyright law is unclear—Freeman believes it to be legal under fair use. While he sees this potential issue being raised by his piece, he did not conceive of it to confront the issue:

I didn't create *N.A.G.* with the intention of commenting on the legality of file-sharing, or Gnutella, or the RIAA. I was simply interesting in exploring things about Gnutella which I found fascinating. I do believe that *N.A.G.* falls under fair use, while many other uses of Gnutella do not, though I realize that others may disagree with my opinions. (Freeman, 2005)

Like Chafe and Brown, Freeman is interested in the complexity offered by networks, but more importantly in their propensity for failure. The failure he is interested in, however, is treated as a structural necessity, and is the type of failure that is exploited in *N.A.G.* when the inaccurate search algorithm produces the types of results described above for a search on 'madonna'. Were this type of failure not present, the juxtapositions that *N.A.G.* plays with could not exist.

What differentiates networks for me is not their complexity, but rather their fallibility. I am attracted to networks which fail to accomplish the goals they have set out: they are often unreliable or inefficient or inaccurate in the tasks they perform. And it's the chaotic, unpredictable process by which they sometimes utterly fail, sometimes triumphantly succeed, and often fall somewhere in between, that makes them interesting to me. (Freeman, 2005)

Jesse Gilbert, *interaXis* (2002) and *Conformed_Bits* (2002)

<http://www.turbulence.org/Works/interaxis/>

http://im.calarts.edu/conformed_bits/

Having worked with streaming media since 1997, Jesse Gilbert has composed and collaborated on numerous works involving the streaming of audio and graphical data over the Internet. The streaming has been directed at end users, or used to join physically separated performance spaces. Much of Gilbert's work uses software he writes himself, and so he is able to create systems for streaming and performance that are highly tailored to his needs. The two pieces discussed here reflect Gilbert's

different interests. One, *interaXis* (2002), is an example of his work with streaming, spatialization and live performance. The other, *Conformed_Bits* (2002), uses streaming and compression technologies to create visual and audio feedback loops that gradually degrade digital data through the compression process.

InteraXis was a collaborative piece premiered in November 2002. It connected ensembles at two locations in Los Angeles and New York via the Internet. Audiences at each location heard their local ensemble play with the other ensemble at the remote location. The music itself was primarily improvised, with guidance provided by a computer-generated graphical score controlled from a central location. Unlike the telepresence experiments with the SoundWIRE project, minimal latency was not a goal in *interaXis*. Rather, latency was accepted as a characteristic of the environment and as something that is intrinsic to collaborative music made over the Internet. Performers did not focus on creating tightly synchronized vertical structures, but instead worked towards shaping the form of the piece in a way that was reflective of the temporal dislocation caused by the Internet connection.

No matter how low it can be taken a network interaction can never be said to be real-time in the way that a traditional concert can, and frankly I'm not sure that's the point. For a performer or audience member, the moment of the performance is the real-time moment, and latent events become creative components that can be layered with the live. I think accepting that working on the network involves the representation of asynchronous moments in time is much more interesting than trying to obsessively re-create traditional modes or methods of performance. (Gilbert, 2005)

Part of Gilbert's custom-written software for *interaXis* included a spatialization system with which he could control the placement of the networked audio within the performance spaces on either end of the continent. The system worked in such a way that the placement of sound from each ensemble in the other ensemble's space was mirrored to the greatest possible degree. If a sound in New York was shifted to the rear left of the room in Los Angeles, then the sound from the latter would also be shifted to the rear left of the room in the former.

I am less interested in trying to reproduce a wholly 'accurate' reproduction of sonic reality between the sites than I am in using spatial processing to indicate the impact of the network on a visceral level to an audience. I find spatial systems intriguing as a means of 'unbalancing' a listener's ears, taking them out of the comfort zone of familiar ways of hearing. For many reasons, this seems to me a good parallel to the experience of performing across the network, where the temporal dislocations and the non-presence of the body force performers to reach for each other across data space. This is not always comfortable for a performer that is used to a certain level of reinforcement or feedback from their collaborating musicians, but it can produce very interesting results. (Gilbert, 2005)

In his more recent pieces, including *interaXis*, Gilbert has moved away from using the network as a replacement for more traditional forms of transmission, and has focused instead on the network as a space with distinct qualities that enable

the creation of a unique type of experience not possible with other types of broadcast media.

I think that *interaXis* reflects an increasing desire on my part to move away from looking to Internet streaming as a replacement for traditional broadcast media (which presently it does quite poorly) and moving more towards a notion of constructing site-specific installations that are network enabled to create interactions between people in those spaces that are accessible to a live audience. Not only do I feel that the work can be presented more fully in a designed environment (as opposed to the undefined at-home experience where there may not even be dedicated speakers to listen through), but that the audience for on-site events is potentially much greater. I am still quite interested in the network, but less for what it might represent than for what it can actually do. (Gilbert, 2005)

In addition to his work with remote performance pieces like *interaXis*, Gilbert has also been working on a project that uses streaming technology to explore the aesthetic possibilities of digital compression. *Conformed_Bits* is an ongoing collaboration with several other artists in which the effects of compressed, network-based feedback on images and sounds are explored. In order to stream moving images or audio over the Internet, a compression program called a 'codec' is commonly employed to remove enough data from the source material to enable it to stream smoothly over limited bandwidth. Streaming Internet radio (usually in mp3 format) and streaming video, such as the RealPlayer format, are examples of this. In the process of compression, large amounts of data are removed from the source file, invariably leaving behind visual or acoustic artifacts. The higher the compression ratio, the more apparent the artifacts become.

The interesting aspect to me as a composer is that a codec represents a culmination of research in psycho-acoustics, particularly around the sensitivities of the human ear, and which frequencies are most important in establishing the essential character of a given sound. Through this a generic picture of the ear is developed, a profile of sorts that the codec uses to make decisions about what frequencies to discard and what to emphasize. The compression process, like any process of conversion for transmission, introduces characteristic noise into the signal that the broadcaster hopes is ignored by the listener, as the transmitted media is 'close enough' to the original to be accepted as a reasonable facsimile. (Gilbert, 2005)

In *Conformed_Bits*, the source material is put into a feedback loop through the compression and streaming process. The combination of network latency plus data removal ends up dramatically transforming the source material as particular pieces of data—such as frequencies in a sound or colors in an image—are emphasized or de-emphasized through the feedback process. This process has some interesting parallels to Chafe's work in which the invisible or unperceivable aspects of the network medium are made visible through a reinforcing algorithmic process. In *Conformed_Bits*, the noise—or artifacts—introduced by the codec are almost like a fingerprint or signature of the compression process, hidden at the edge of one's perception if the bit-rate is high enough. Repeated passes of material through the codec reinforce artifacts,

bringing the hidden audio or visual fingerprint of the process into the perceptible realm. What is gradually revealed, in a similar process to Alvin Lucier's *I am Sitting in a Room* (1970), is the acoustic or visual coloration of the codec.

The feedback structure of *Conformed_Bits* amplifies this noise over time, and also brings the codec to its limits by reinforcing the selected frequencies. Thus we can use the codec as a variable filter that 'rings' or resonates in certain characteristic ways as the original material is introduced and ultimately destroyed by repeated processing. I find this interesting to experience over time. I tend to think of this as representing the breakdown of a way of thinking about the world that posits that anything can be represented in discrete digital form, or that human perception can be quantified as an 'average' result that fits us all. (Gilbert, 2005)

Peter Traub: *ground loops: for solo percussion and internet* (2005)

I have had a personal interest in network-based art and music since I created my first piece of Internet art, *bits & pieces* (<http://www.fictive.org/bits>), in 1999. I collaborated on two more works after *bits*, before taking a several year hiatus from composition. A new piece of mine, *ground loops: for solo percussion and internet*, was premiered at the University of Virginia on 3 May 2005. It was my first piece for live performer and Internet, and was the beginning of what I hope will be a long-term focus on such efforts (see Figure 4).

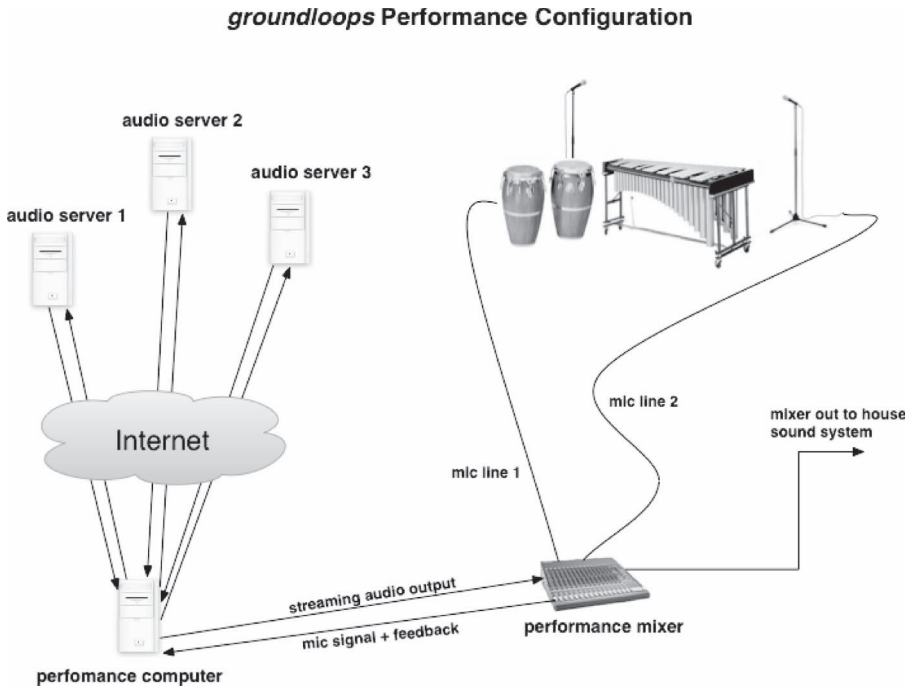


Figure 4 The performance configuration for *ground loops*.

ground loops involves guided improvisation for a solo percussionist, in which the sound from the performer is mixed into feedback loops to several servers, not unlike those in Chafe's SoundWIRE project or Gilbert's *Conformed_Bits*. It is less technically sophisticated than those projects, but the audible results are nevertheless compelling. Using a combination of out-of-the-box software packages on Linux, I set up a series of streaming feedback loops to three remote servers dispersed around the country. I use the Ices streaming audio relay software to send audio from the performance space out to three IceCast streaming audio servers. The XMMS (X Multimedia System) player is then used to pull the streams back into the local performance computer. A percentage of the audio brought back in is mixed with the live percussion audio and looped back into the streaming system, thus putting it through the same encoding process repeatedly. In this case, the Ogg Vorbis codec is used to compress the outgoing audio. The processed audio is played out into the performance space along with the live percussion audio. In the score, the technical implementation is left up to the programmer/performer, with three primary requirements: the outgoing audio must be compressed with a codec; there must be three external servers, each accessed sequentially through the piece; and the incoming audio from each external server must be fed into the outgoing audio to the other two servers.

The percussionist begins the piece by choosing gestures from five available instruments and/or timbres: wood, skin, cymbal, struck vibraphone and bowed vibraphone. The score of *ground loops* is structured in such a way that instruments are sequentially removed from use as the percussionist plays through the piece. The performer is left playing only one instrument by the end. As this progressive instrument subtraction occurs, the number of feedback loops is gradually increased from one to three. Each addition of a loop causes a considerable increase in the density of the feedback, and further reinforces particular frequencies and timbres through the repeated passes through the encoder. The loops, besides increasing in number, also feed back into each other, so that sound looping between Virginia and California is also fed into the loop between Virginia and New Hampshire, and vice versa. The result is that the performer and the system both contribute to a filtering process in which the percussion and the feedback reinforce the same sonority.

As this is only my first experiment with a live performer and network, there are many areas that will need to be improved upon in subsequent works. The out-of-the-box software is a quick way to set up feedback loops, but there are a number of drawbacks that prevent me from using the network with the level of control I desire. Since all the loops feed back into a single instance of the XMMS player, it is not possible to separate the loops from each other spatially. The loop times also approach extremes of anywhere from five to ten seconds. The reason for this is not network speed, but rather buffering occurring in multiple places throughout the feedback chain, such as in the XMMS player and the IceCast server. Finally, one must take into account the effects of the external analog mixer on the loops, and the degree to which its EQ and frequency characteristics affect what is emphasized in the feedback

process. A fully digital mixing process, preferably all internal and software-based, would solve this problem.

In the first performance of *ground loops*, I configured the encoder to resample the audio at 22.05 KHz and use as low a bit-rate as it would allow. The result was streaming audio in the 10–15 Kbps range, making the compression artifacts apparent from the first delayed return of a live sound. The audience could clearly hear the dramatic quality change to the live percussion sounds as they were processed repeatedly through the encoder. The long delay times of the loops and the gradual quality decay of the sounds created a hypnotic texture, becoming a very slow pulse that built as the piece progressed and the feedback was increased.

Conclusion

The artists described above are part of a growing field that uses the Internet and networks in ways that have only begun to be explored. When Neuhaus premiered his radio works, and The Hub first began making computer network music, the barriers to using the network as a medium were high. As the technology has become faster, cheaper, and more widely available, the number of people experimenting with networks for art has grown. But what does it mean to make music on the Internet? The implications of the technology are so rich, varied and dynamic that it is difficult to say with certainty what the trajectory of the Internet or networks as musical mediums might be. Because of their distributed nature, the notion of audience changes, and with that, the notion of what constitutes a composition or a performance.

The Internet allows for a greater blurring of the line between composer, performer and listener than has been possible in the past. In pieces like Brown's *Eternal Network Music* and Neuhaus's *Auracle*, the composers do not have control over individual sound events, or even the shape of a particular performance. On the other hand, they do have control over the general performance and musical environment. In this way, the Internet allows composers to shift the focus of their activity within the music-making, performing and listening process in ways that were previously unavailable to them.

[W]e may have to think about a more open way to define the notions of the composer, the performer, the listener. The composer would (still) be a person who sets a scope and lays connections of ideas, concepts and atmospheres in music. The performer would be the person to mediate a primarily auditory event, albeit following different rules than in the past. Finally, the listener would be the one to put it to a final order, to receive it at a final destination and interpret his or her personal version of it. Whatever will happen, the main effect is that the range of roles for each grows. (Föllmer, 2001)

Aside from challenging more established notions of composer, performer and audience, networks (and the Internet, in particular) challenge notions of listening. Where does sound exist on the Internet? As the space of the network is virtually vast

yet physically non-existent, the meaning of distance and location changes. A sound coming over the network from a server 1,000 miles away may take longer to reach a user than one coming over the Internet from 3,000 miles away. Both of them may still arrive within a very short amount of time, but even then, what does the nature of that distance mean to the listener, if anything?

The space of the network, as both acoustic realm and transmission structure, goes a step further, completely challenging the notion of any locus of hearing. By removing the concept of source and site in general, and shattering it across a multitude of non-places, we no longer comprehend the source, locale or temporal origin of sound. Listening to sound from the [Internet], we have no sense of from where a sound emanates, or why it is there in the first place. Sound rests in the ultimate non-place of the networked world: the territory of the server. (Salter, 2001)

The network has redefined many of the spaces that human beings previously encountered in the physical world: economic spaces, social spaces, listening spaces and communication spaces to name a few. Yet these new virtual spaces are in a constant state of flux, through continuing technological development and cultural reassessment. As such, art that is made to exist in these spaces is also in a state of flux and constant development.

The network environment is in the process of replacing the former social spaces in which mass media and daily life were separated. In a world in which mediated communication and exchange are at the center of everyday life, the artist must seek out a manifestation of her or his work in a different way. The digitized world, especially the world of digitized sound, is one of endless possibilities and nonhierarchical structures. This creates a freedom that allows for all known truths and traditions to be reconstructed, re-invented and rebuilt as if from scratch. (Bosma, 2001)

With the freedom Bosma describes above, comes the necessity for those who listen to, view or interact with network art to re-evaluate what the experience of engaging with music and art means. According to Golo Föllmer (2001), music on the net is 'hardly a music of its own yet'. Four years after he wrote that, this is still the case. New art forms take a long time to develop, and because of that, the ways in which people respond to these new forms also takes time to develop.

It took over 50 years of experimentation with recording technology (the gramophone) to lead to acknowledged and influential musical results with *Musique Concrète*. Similarly, with networked music, we presently are talking about potentials of music which appears sometimes strange, sometimes awkward, because it requires a new musical state of mind in order to be made and to be heard. (Föllmer, 2001)

While Internet music is not quite 'a music of its own yet', it is possible to see several trends or directions as evidenced by the works covered in this article: telepresence works (collaboration over a network), sound toys for non-musicians, feedback works and/or works that expose the nature of the medium, and sampling

works that use the Internet as a data source—these are but few of the many approaches that artists are taking to this new medium. Composers like Chris Chafe (2005) ‘sense some new forms on the horizon’, and it is likely, given the nature of the technology, that there are many forms yet to be invented that will use the Internet and computer networks in novel ways. As a composer and technologist, the prospect of witnessing the birth of these new forms, or even helping to create them, is an exciting one indeed.

References

- Barbosa, Á. (2003). Displaced soundscapes: A survey of network systems for music and sonic art creation. *Leonardo Music Journal*, 13, 53–59.
- Bischoff, J., Gold, R. & Horton, J. (1978). Music for an interactive network of microcomputers. *Computer Music Journal*, 2(3), 24–28.
- Bosma, J. (2001). *Music and the Net: Musaïc*. Available online at: <http://crossfade.walkerart.org> (accessed 1 March 2005).
- Brown, C. (2005). E-mail interview with the author.
- Brown, C. & Bischoff, J. (2002). *Indigenous to the Net: Early network music bands in the San Francisco Bay area*. Available online at: <http://crossfade.walkerart.org/brownbischoff/IndigenoustotheNetPrint.html> (accessed 1 March 2005).
- Chafe, C. (2005). E-mail interview with the author.
- Chafe, C., Wilson, S. & Walling, D. (2002). *Physical model synthesis with application to Internet acoustics*. Paper presented at the International Conference on Acoustics, Speech and Signal Processing (ICASSP), Orlando, FL.
- Duckworth, W. (1999). Making music on the Web. *Leonardo Music Journal*, 9, 13–18.
- Follmer, G. (2001). *Soft Music*. Available online at: http://crossfade.walkerart.org/foellmer/text_print.html (accessed 1 March 2005).
- Freeman, J. (2005). E-mail interview with the author.
- Freeman, J., Ramakrishnan, C. & Varnik, K. (2004). *The architecture of Auracle: A real-time, distributed, collaborative instrument*. Paper presented at the New Interfaces for Musical Expression (NIME) conference, Hamamatsu, Japan.
- Gilbert, J. (2005). E-mail interview with the author.
- Gresham-Lancaster, S. (1998). The aesthetics and history of The Hub: The effects of changing technology on network computer music. *Leonardo Music Journal*, 8, 39–44.
- Levin, G. (2005). E-mail interview with the author.
- Neuhaus, M. (n.d.). *The Broadcast Works and Audium*. Available online at: http://www.auracle.org/docs/Neuhaus_Networks.pdf (accessed 1 March 2005).
- Neuhaus, M. (2005). E-mail interview with the author.
- Salter, C. (2001). *The architecture of listening*. Available online at: <http://crossfade.walkerart.org> (accessed 1 March 2005).
- Thomson, J. (2005). E-mail interview with the author.