Affordable Sound Field Panning in Theatre

USITT Research in Sound 2020
Christopher Plummer

Abstract:

This paper will evaluate an affordable software that allows theatre designers to work in a virtual 3D space, a sound field, instead of worrying about setting volumes on individual loudspeakers. We will evaluate the software in our production of Sarah Ruhl's Eurydice in a black box theatre with 20-30 loudspeakers above, below and around the audience. Through this production we will evaluate the transparency of actors' reinforced voices statically positioned in the system as well as static and moving sound effects and ambisonic ambiances played back in Q-Lab.

Précis:

Sound field panning, with or without wave-field synthesis, is poised to be the foundation for all theatre production work in the future. These systems allow for 2D or 3D panning across many loudspeakers from standard 7.1 surround set-ups to managing hundreds of loudspeakers and providing easy to use dynamic panning through the entire sound system. This allows designers to better utilize an advanced sound system to its full potential, quickly building complicated soundscapes and sound movements that are exceptionally tedious to build without such systems. The most capable of these systems include two features beyond easy panning: an object oriented processing system and wave field synthesis.

An object oriented panning system allows a sound field to be played back on any sound system. While this clearly has its limitations: stereo playback will not sound like a 60-speaker immersive sound installation. It does mean that if you have a 7.1.4 studio (7 surround speakers, 1 sub, and 4 ceiling or height speakers) you can program your pans as objects and then the system will translate them to the full immersive sound installation and it should be fairly close in sound. Not only that, but, with a binaural or VR system you could simulate your 3D sound field on headphones, making work on the road significantly more powerful.

Wave field synthesis refers to the ways an immersive sound engine pans through multiple speakers. In a very simple system, the speakers will be thought of as lots of stereo pairs where the sound is in one speaker or moving to a second speaker or at most triangulated between three speakers in amplitude only. This provides significant power and benefit to a designer. However, in the most advanced immersive audio systems a sound will be represented as a point in space and the waves coming from that sound will be reproduced with appropriate delay through every appropriate speaker. This modeling of the wave front rather than amplitude panning of sound opens up a whole other level of transparency and flexibility in the sound system.

Currently systems capable of this workflow are often very expensive and don't necessarily provide the full benefit of object oriented panning or wave field synthesis. This paper will provide a detailed case study of using the more affordable SPAT tools developed by IRCAM and sold through Flux Software Engineering. These tools bring the cost of object oriented sound field panning down to around $3,000 as an add-on to an existing multi-speaker sound system, significantly less if you already have a computer and audio interface. This system is capable of

both object oriented panning and several different approaches to wave field synthesis for different types of system set-ups.

To make this as valuable as possible to theatre professionals the tools will be evaluated in a black-box theatre production utilizing Q-Lab as the playback engine with sound object control in SPAT accessed through an AU plug-in. Source files will consist of traditional mono, stereo, and surround music and sound effect files, as well as Ambisonic recordings. Because SPATs internal object oriented architecture is in ambisonics this opens up the possibility of decoding ambisonic ambiances live in the theatre from 4 track Q-lab playback to any size or configuration sound system. This is particularly exciting to me since Q-Lab is not currently able to support plug-ins with different # of channels in and out, making ambisonic decoding cumbersome. And even if Q-Lab did support this, most plug-in decoders are built around an assumption of film surround systems not theatre systems.

This system is also capable of interfacing with actor trackers for auto panning and delaying wireless microphones, as well as providing software control through plug-ins on Avid consoles. We will not be fully testing these capabilities although we will have the SPAT system available to our console and will be able to apply automation via Q-lab. It is unclear right now what creative opportunities this presents for this production but we will test SPAT for static delay of actors to evaluate its basic transparency and usefulness in this area.

Overview of The System Used

Our system was made possible by our existing investment in Dante and the replacement of our computer lab Mac OS computers. For a long time we have been running our Mixer and Q-Lab computer directly into our BSS DSP system. In our search for immersive tools we had already added a reverb computer onto the Dante network to provide surround reverb that was routed in the DSP to many of the speakers and easily routed to and from both Q-Lab and our mixer. With this infrastructure all we had to add was a computer and the software.

It is not recommended to use Dante Virtual Sound Card with SPAT. It does work and could yield acceptable results if you are only using Q-Lab. However, it has a significant amount of latency that can be a problem if you are reinforcing live sound sources. It also puts more load on the computer processor than a hardware Dante interface, thus limiting the complexity you can achieve with SPAT. If you don't have a network audio system, you can still utilize the SPAT system you just need an audio interface with enough ins and outs to go between your sound sources and your loudspeaker system. If you are using this system with an external reverb computer, that computer can easily use Dante Virtual Sound card, as the latency just limits the lowest pre-delay available on your reverbs and that is usually acceptable or at least workable.

Attached to the end of this document is a simplified version of our patching. I haven't provided a lot of detail on the output side because that will be heavily dependent on the space and production. I have provided some information about ways we found it useful to work quickly with SPAT. Truth be told, as we experimented with the system, we ended up using almost all of the 64 channels of in and out available. One of the things we set-up as a safety was a one-to-one

patch so that sounds could be sent to individual speakers through the SPAT engine. We never used this, and because of the panning mode we selected, it didn't really work anyway[1].

The one item of interest on the output side is that we had a recording computer and fed it with a 7.1 output generated by SPAT as well as a binaural feed. These worked very well. We got a great spatial immersive mix of our sound effects, music, and actor reinforcement all in a natural sound field. This recording was far better than 90% of theatre recordings I have heard. There was more actor voice in the surround speakers than I would have mixed but overall if you want an archival recording this works very well and is crazy easy. These recordings were made as a technical test and are not available.

The computer we ran SPAT on was a 2013 Mac Pro with 32GB of RAM and a 4 core 3.7MHz Xeon E5-1620 v2 processor, also known as a "trash can". This computer with a USB3 RME Digiface gave us 5.65ms of delay while processing 45 inputs to 47 outputs. We were working to keep the latency down and 5.65ms of latency was the lowest possible. This is an older, although still powerful, computer. I'm guessing that newer i7 based iMacs will provide similar performance. SPAT does appear to use multiple processors fairly well, so a system with higher multiprocessor performance as opposed to single processor performance is probably recommended.[2]

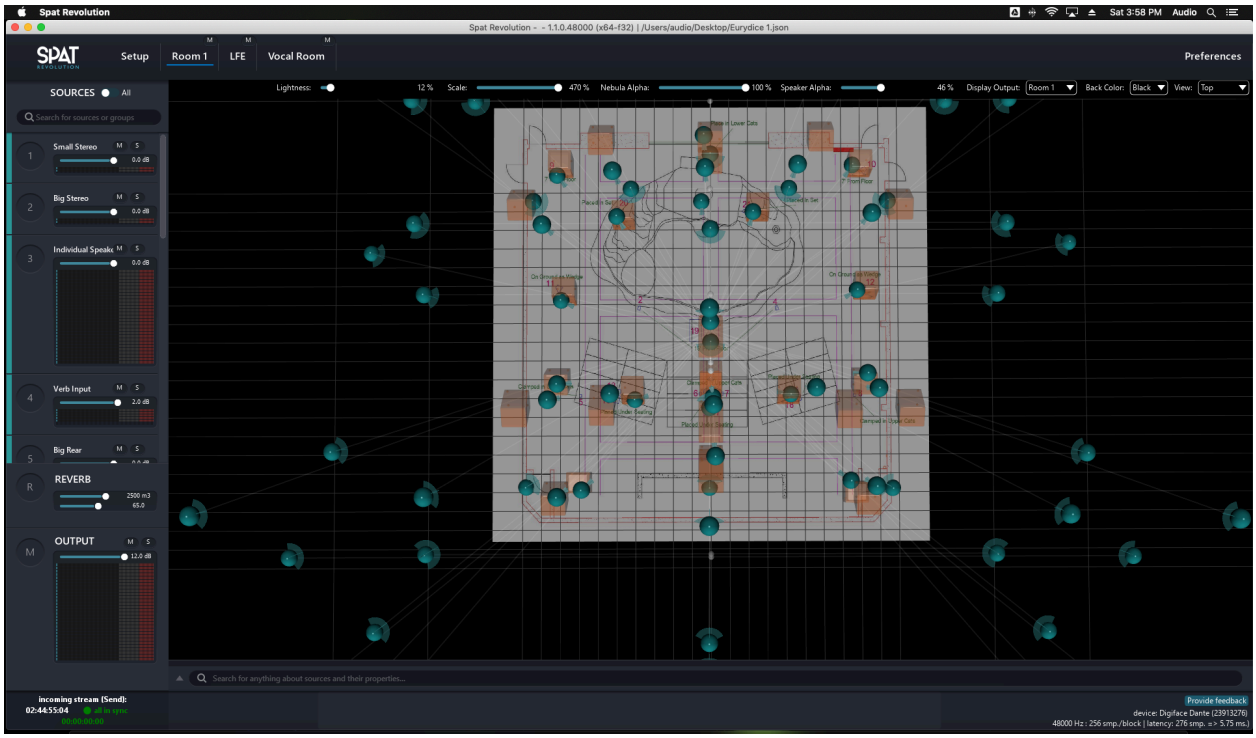
In addition to the panning spatialization, SPAT also has a built-in multichannel reverb processor and we used that on several inputs. Having the built-in reverb is very useful and powerful, especially since it can theoretically provide uncorrelated sound to each loudspeaker. In practice it was a pretty good sounding reverb that required a bit of tweaking to get right. Most of our reverb use was from our external 5.1 Altiverb reverb that was then up-mixed in SPAT to our 37-loudspeaker system.

Setting up the SPAT software

The quality of the SPAT spatialization is dependent on it knowing the location of the audience and the precise location of each speaker. With 37 loudspeakers to enter, this wasn't the part of the job I was most excited about. I had dreams of sitting at my center audio position and shooting each speaker with a laser distance measure that would also give me the angle to the speaker. This

[1] See the section on Panning options for an explanation of why this didn't work in our case.
[2] Geekbench <https://browser.geekbench.com/mac-benchmarks> is a good location to get a general sense of computer performance when buying a new computer. The Mac we used had a multiprocessor score of: 3256 and a single processor score of 813. According to Geekbench a new i3 based Mac mini has similar power to this computer.

would have mostly worked although speakers that are obscured by catwalks or fabric or whatever else will not be measurable this way.



What I actually did was way easier. I used the laser to measure one speaker and placed that in SPAT. The I imported my speaker plot into SPAT and scaled it to properly locate that speaker and my listening position. Once that was in place, I just placed speakers on the speaker plot, measured their height, and I was done. It went very quickly and all of my outputs were complete.

The inputs are all based on what you want creatively and what will allow you to work quickly. To start with, I created a couple of stereo sound sources in SPAT added a few of my favorite music tracks to Q-Lab and started moving things around to see what this software did. I immediately fell in love. By creating just a few stereo sound sources in various locations:

- a narrow pair on stage with the actors
- a wider pair at the upstage edge of the stage
- a very wide pair spread beyond the sides of the theatre
- a stereo pair to spread wide to either side of the listening position
- a wide pair behind the audience spread wide

This provided basic locations that I could very quickly drop different ambiance tracks on and have an immersive environment so much better than anything I could create in a similar amount of time in Q-Lab. It just worked. Then, we added some specific mono locations for the elevator, a few off stage sound locations, and some very far away sounds. These again just dropped in and sounded great. The distant sounds are so awesome. We actually added one distant location for when an actor goes off stage and we still wanted him reinforced. This is often a situation with

very undesirable sound quality having neither the offstage sound nor the vocal clarity needed. We got both by clicking a couple of boxes and then having the mixer just pull one fader down and another up on the sound console. In this situation it easily provided a very believable sense of offstage sound, yet it was still clear and intelligible. Magic, I tell you.

System Tuning

I often spend many hours getting the delays "just so" on my system, especially when we get above twenty loudspeakers and I want a highly level of clarity from the sound system in addition to solid localization to the stage. For this system I spent no time on delays. I did a quick EQ of each speaker, insured they all produced the same SPL at the listening position, and then let SPAT do all the rest. This process cut my tuning time in more than half, removed a significant amount of math, and also removed the part of the job I most often make mistakes in. For this show I was happy with the results. I do plan on investigating this in more detail to see what is really going on with the delays and timing accuracy in the system. The way we set the system up resulted in a little bit woolier sound that I usually get.

Panning Options

SPAT has 8 different panning algorithms and the differences between them are sometimes very audible. To narrow this down and understand what these are, I read the manual. It turns out that two of the systems are for backward compatibility only, so I avoided them. No sense getting excited about something that is probably going away. I then noted that a few of the systems are really designed for a small sweet spot, only a couple of seats in the middle of the speaker array with equidistant speakers. This was not what I had. SPAT does provide some tools to make that work with an irregular speaker array, but I decided for this situation to focus on panning algorithms that the manual said worked best for irregular speaker arrays and wide, oddly positioned audiences. The manual recommends both Distance Based Angular Panning (DBAP) and K Nearest Neighbor (KNN). As we listen to the different algorithms, one of the things that become clear was that some created a bigger more diffuse and enveloping sound while others created much more accurate points of sound localization. We settled on DBAP and an immersive and somewhat fuzzy localization as it worked for the majority of our content for this show.

It is possible with SPAT to create multiple rooms with different panning algorithms. This way you could send your ambiances to a diffuse and enveloping room, and your spot effects to a precise room. I didn't get the opportunity to test this hypothesis in my production process, but it is on the list for the next production.

One of the things that the diffuse algorithms do is spread the sound over more speakers. This has the possible drawback of giving a less precise sound image, but it also means that individual speakers are having to push less of the SPL and this then increases your headroom. Interestingly we also sent our vocals through this system and the feedback performance was pretty astounding. It is especially unbelievable because I had planned on creating a separate vocal room with fewer speakers. However, it seems that each of your rooms has to have the same speakers in order to sum them together to a set of outputs. I had a few work-arounds to this but lacked the time to implement them. What this meant was that our stage vocals were not only going to the vocal reinforcement speakers but a little bit to every speaker in the room. This included the spot effect speakers in the set and reverberation speakers behind the performers. This should have been a

feedback nightmare, but it wasn't a problem at all. Well, except for the one-time during rehearsal an actor bent over and placed their head just a foot or two in front of an onstage speaker. We made a slight adjustment of speaker location to give us an additional couple feet of safety and the issue never happened again. This diffuse, use-all-the-speakers approach also means that any one speaker's location isn't all that important as it is part of this large sound field.

As I re-read parts of the manual it seems that, with the KNN panning method, you get an extra control that allows you to specify how many Nearest Neighbors a sound is panned across. This could provide an amazing morphing between the power being diffusely spread over many speakers and getting a sharper focused sound that is coming from a specific location. Experiments for next time.

Playback Workflow and Results

If Q-Lab is your playback system of choice, you need this software or something similar. It gives you so much more power and control over the sound and allows you to work so much faster. This brings object oriented panning and design to the Theatre. It allows a team to create a sound system in which the system designer can create a stable foundation upon which SPAT creates an artistic canvas. Then when you are at the design table, all you have to worry about is your canvas. You don't care which speaker is where. You only care where you want the sound to come from.

As you create your canvas, think about what sources you are working with. Sound effects are often in mono, stereo, and now ambisonic. All of these sounds can be routed to appropriate objects in SPAT and placed precisely where you want the sound to be. The time savings of being able to decode ambisonic in the theatre is amazing. In the past I had written apple scripts that processed my sounds into three surround files (low, head height, and high) for playback on multichannel systems. So far that system using the Harpex plug-in does result in better quality sound but it is also a lot of work, and the scriptable editor I was using is no longer available having been apparently shut down by a hedge fund company camping on intellectual rights. I am hopeful, since I haven't explored many of the ambisonic native panning methods, that I will be able to get better quality sound along with the speed and flexibility of working live in the theatre.

For mono files, plan out the specific locations you want spot effects and create locations inside SPAT for those. Most importantly though is to create a solid two channel "stereo" location that outputs from channels 1-2 from Q-Lab, this allows you to very quickly throw up sound and make it sound good. This helps you work so much faster and put up something that can be immersive and have the correct relationship with the playing space immediately. What made sense to me, my first time doing this, was to create a stereo sound source that was an intimate, closely-spaced left and right image inside the acting area. Then I created a huge left right at the upstage edge of the playing area and the left and right sources being at or slightly beyond the walls of the theatre. These instantly gave me two contrasting sounds that both occupied the full theatre but did so in different ways, at different scales, and with different relationships to the action on stage. I then created stereo sources to the sides and behind the audience. This got us up and running very quickly and also provided a sense of the power of the system. As we layered more sounds, we could easily add sound sources in SPAT to create more locations and thereby further differentiate the sound scape. Adding a location for each instrument in a sampled orchestral piece could

greatly help with the realness as well as the differentiation of musical parts. So many things to try.

Vocal Workflow and Results

We used SPAT not only for our music, ambiance, and spot effect playback, but also our vocal reinforcement. In this situation we created a mono vocal source slightly upstage of center stage. Using the diffuse sound of DBAP, the reinforcements system blended into the room. It provided significant gain before feedback and, even at higher levels of support, maintained very good transparency. Neither a smearing of articulation nor a pulling of focus away from the actor were problems. The system did a great job making the actors louder and easier to understand without ever pulling focus or creating a situation that was louder but less clear.

I was very happy with the system for vocal use, especially with the DBAP panning. However, I do need to work out how to not have the vocals come through the effects speakers on stage as that is just waiting to be nightmare. I know I can solve this with multiple SPAT systems or by taking multiple outputs from one SPAT system and summing them in a DSP, but I'm hoping to find an elegant way to handle this inside SPAT or to convince the developers to add such functionality.

Future Exploration

I'm looking forward to diving deeper into the different panning algorithms and how to control them and what is really going on under the hood. I'm very interested in how, or even if, SPAT is doing much in the time domain. Because SPAT is working with sound sources in 360 degrees, and not just form the stage, it would be far better for delay times to be generated based on the programmed sound locations rather than speaker locations as I have been doing with traditional DSP. Ideally a sound source is defined not just through amplitude in each speaker but also through time of arrival. I have always been much more interested in using time to pan sounds than amplitude and in theatre we never need to worry about mono-compatibility. I have not yet had the opportunity to test the different panning systems to see how they handle time of arrival

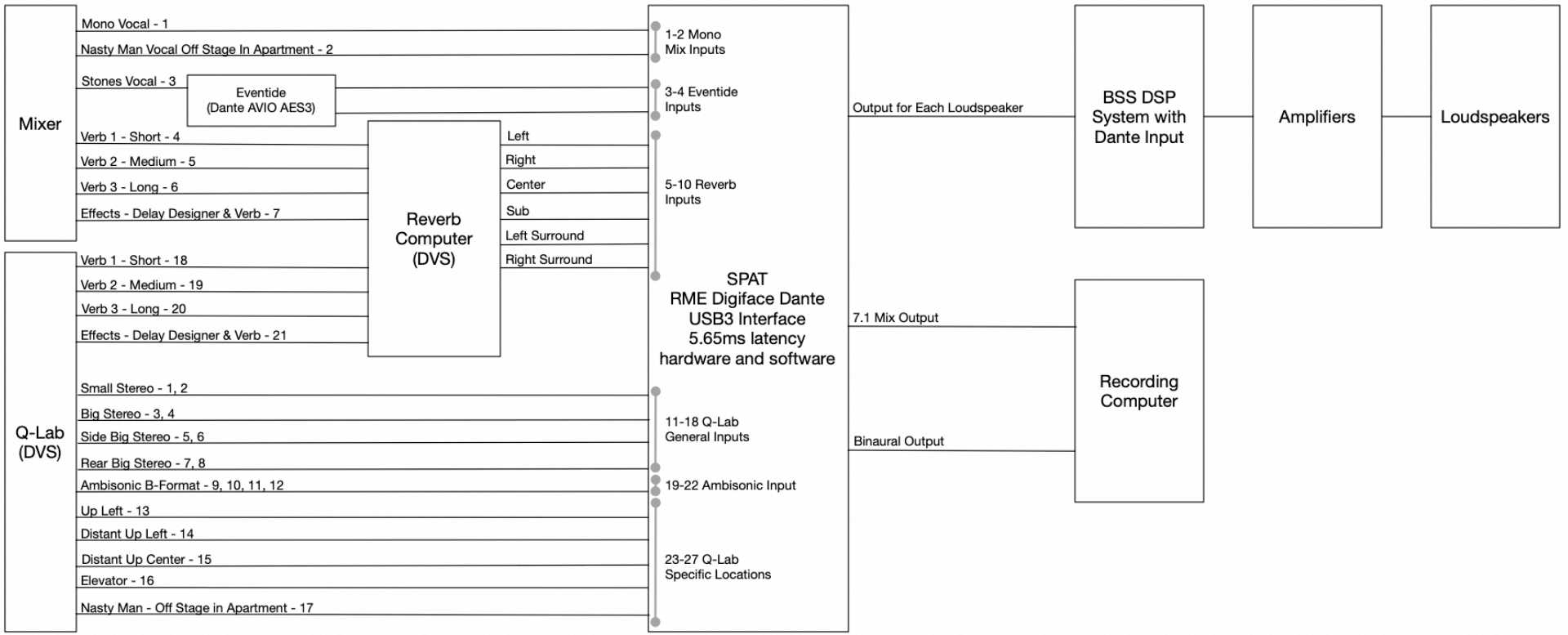
I also need to explore the possible ways that multiple rooms can be used. This seems like it could be a powerful way to utilize different panning algorithms for different artistic goals within one system. I also hope it will hold the key to having playback and live SPAT processing where the reinforcement SPAT system isn't able to use certain loudspeakers.

I'm currently testing SPAT with our reverb computer for using area mics to change the acoustics of a space. I recently added reverb to a small chamber group using just a stereo pair of microphones not very close to the performers. The gain before feedback was significantly higher than needed and the sound enhancement was pretty transparent and smooth. I'm really happy with the way SPAT takes the 5.1 output of Altiverb and expands it over all the systems speakers making it all the more believable as the natural reverberation of the space.

Final Thoughts

I have been waiting for this software my entire career. While there are situations where this software would add work, I don't see that being the case on any of the shows I am working on in the near future. The enhancements to sound quality, work flow, and how much time it can save

me in filling a space with sound are so valuable to me that it will be pretty rare for me to not use this software. And not only am I excited about the theatre possibilities, but I'm also excited for the concert possibilities. I'm going to do an initial demo with the music faculty of what we can do with this technology. In most cases this will be a transparent shifting of the space to provide resonance, presences, and envelopment to a musical performance. But I also hope that we explore the shifting of space electronically as a compositional opportunity, a tool to alter parts and add new ways to create and manipulate musical experiences.


Mixer
Mono Vocal - 1
Nasty Man Vocal Off Stage In Apartment - 2
Stones Vocal - 3
Eventide
(Dante AVIO AES3)
Verb 1 - Short - 4
Verb 2 - Medium - 5
Verb 3 - Long - 6
Effects - Delay Designer & Verb - 7
Reverb
Computer
(DVS)
Left
Right
Center
Sub
Left Surround
Right Surround
Q-Lab
(DVS)
Verb 1 - Short - 18
Verb 2 - Medium - 19
Verb 3 - Long - 20
Effects - Delay Designer & Verb - 21
Small Stereo - 1, 2
Big Stereo - 3, 4
Side Big Stereo - 5, 6
Rear Big Stereo - 7, 8
Ambisonic B-Format - 9, 10, 11, 12
Up Left - 13
Distant Up Left - 14
Distant Up Center - 15
Elevator - 16
Nasty Man - Off Stage in Apartment - 17
1-2 Mono
Mix Inputs
3-4 Eventide
Inputs
5-10 Reverb
Inputs
SPAT
RME Digiface Dante
USB3 Interface
5.65ms latency
hardware and software
11-18 Q-Lab
General Inputs
19-22 Ambisonic Input
23-27 Q-Lab
Specific Locations
Output for Each Loudspeaker
7.1 Mix Output
Binaural Output
BSS DSP
System with
Dante Input
Recording
Computer
Amplifiers
Loudspeakers