High-Resolution Audio and Steps Toward Implementing DSD Streaming

Although many people enjoy listening to music on portable audio players or smartphones on a daily basis now, there perhaps aren't so many who are interested in its audio quality. However, it may have occurred to people that "high-resolution audio" seems to be popular now, after finding articles about it in newspapers and magazines, or seeing dedicated space for high-resolution audio in large electronics retail stores over the past few years. "Hi-Res Audio" is a term used to indicate high resolution music content and its compatible audio devices with sound quality in excess of CD (Compact Disc) and DAT (Digital Audio Tape). Here we will discuss what exactly changes when the resolution is increased.

■ What's Hi-Res Audio? What's DSD?

When converting audio waveforms to digital signals in the PCM (Pulse Code Modulation) format widely used for digital audio, the analog signal is sampled at set intervals (from tens of thousands of cycles per second to hundreds of thousands of cycles per second), and their size is expressed as digital code in certain digits (quantization bits number). That means you could regard the analog signal as being converted to staircase waveform information as shown in the Figure 1.

Based on the staircase waveform on the left, we can see at a glance that the staircase waveform on the right with finer time axis and step size is closest to the original analog signal. However, it is also clear that the top staircase waveform with the finer time axis only, as well as the bottom staircase waveform with the finer step size only, are closer to the original analog signal than the staircase waveform on the left. When the granularity of segmentation has been increased in this way to bring audio closer to the original analog waveform, it is said that the resolution has been raised.

Finer time axis corresponds to higher sampling frequency, while finer step size corresponds to a higher quantization bits number. Consequently, if we suppose the waveform on the left is CD or DAT, higher sampling frequency or quantization bits number above that of the CD or DAT standard raises the resolution, resulting in high-resolution audio. Conversely, anything that falls below the CD or DAT standard would not be called high-resolution audio.

People generally tend to think that 96 kHz / 24-bit would be the high-resolution audio standard, and expect a signal contains components above 20 kHz constantly, but most of the time audio signals would not contain components above 20 kHz. Because of this, it is important for the audio signal to be more precise to the original waveform rather than having higher frequency components over 20kHz, as a result of the increased sampling frequency. This is the reason why we can hear a clear difference compared with CD on high-resolution audio even when converting very old analog recorded tapes that intrinsically have no

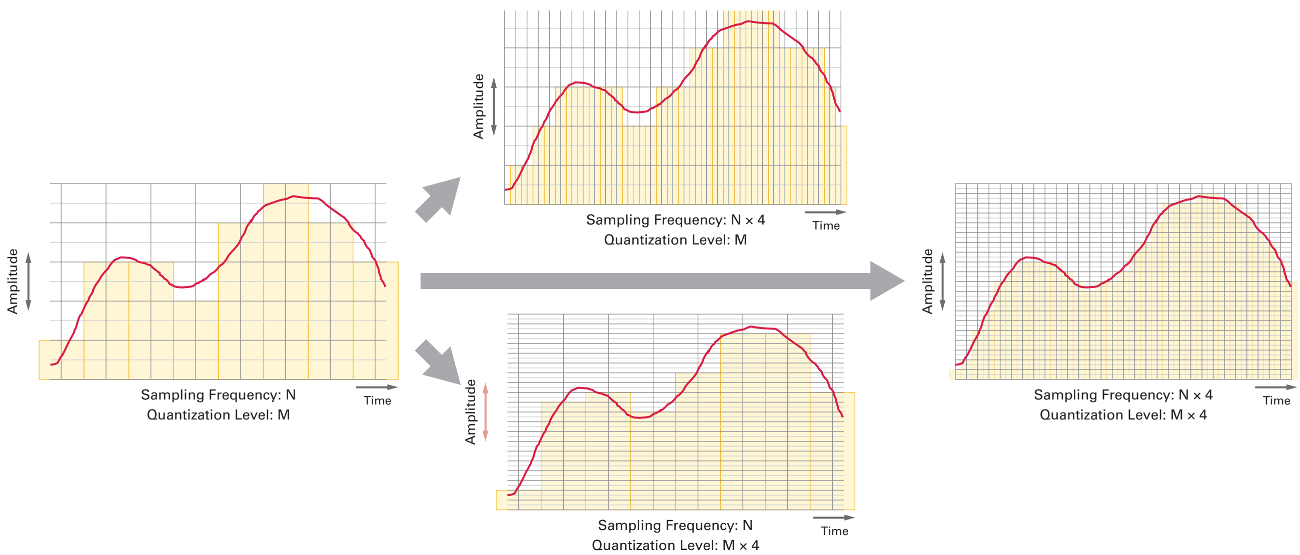

Figure 1: Waveform Variation Based on Differences in Sampling Frequency and Quantization Level

components over 20 kHz, as the audio signal from the master tape can be transferred more precisely including hiss noise*1 due to the faithful digitization. This would demonstrate that even with a sampling frequency of 44.1 kHz or 48 kHz that cannot record a signal over 20 kHz, 20-bit or 24-bit audio is a respectable high-resolution audio source.

Table 1 is a list of the types of high-resolution audio sources that are currently available on the market for digital music distribution. This includes systems such as DSD (Direct Stream Digital), which may contradict our previous explanation that resolution is raised by increasing the number of quantization bits. We will examine this DSD format next.

DSD is the name given to a system for directly recording, transmitting, and playing back a high-speed 1-bit quantized delta-sigma modulation signal. This system was adopted for the Super Audio CD (SA-CD) audio format released in May 1999. Unlike the PCM format, which converts an analog signal to staircase waveform information, DSD converts the volume of sound to dense information consisting of 1s and 0s that change at extremely fast intervals (from millions of cycles per second to tens of millions of cycles per second) (Figure 2).

Encode	Sampling Frequency	Quantization	File Format (extension)
PCM	44.1kHz	24bit (32bit float)	WAV/FLAC/AIFF/ALAC (.wav/.flac/.aiff/.m4a)
	48kHz		
	88.2kHz		
	96kHz		
	176.4kHz		
	192kHz		
	352.8kHz		WAV (DXD Digital eXtreme Definition)
DSD	2.8224MHz	1bit	DSF/DSDIFF (.dsf/.dff) (DSD Direct Stream Digital)
	5.6448MHz		
	11.2896MHz		

Table 1: High-Resolution Audio Formats Currently Available for Digital Music Distribution

You can regard this system as being similar to how inkjet printers can print vivid grayscale photos simply by changing the density of black ink dots. This system provides more freedom, allowing the frequency response of the dynamic range to be determined based on delta-sigma modulation design, without altering the sampling frequency or quantization bits number like PCM. In principle, it makes it possible to demodulate to an analog signal by merely passing the data through an analog low-pass filter (LPF)*2, meaning this digital signal has characteristics just like an analog signal.

Figure 3 shows the typical characteristics of DSD with the SA-CD sampling rate of 2.8224 MHz, which is 64 times that of CDs. This demonstrates that DSD provides flexibility through the

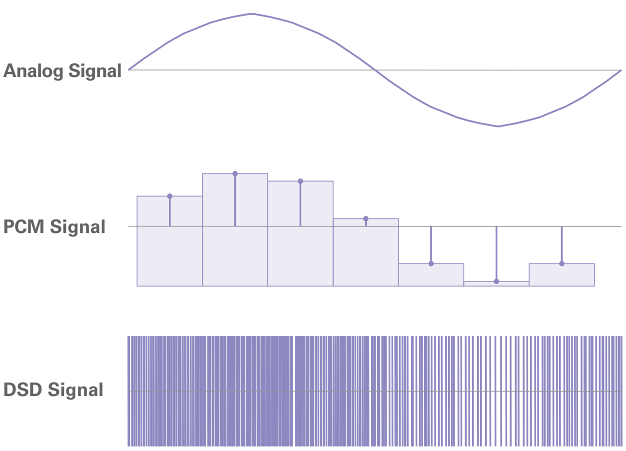

Figure 2: Differences in How the PCM Format and DSD Format Express the Waveform Information

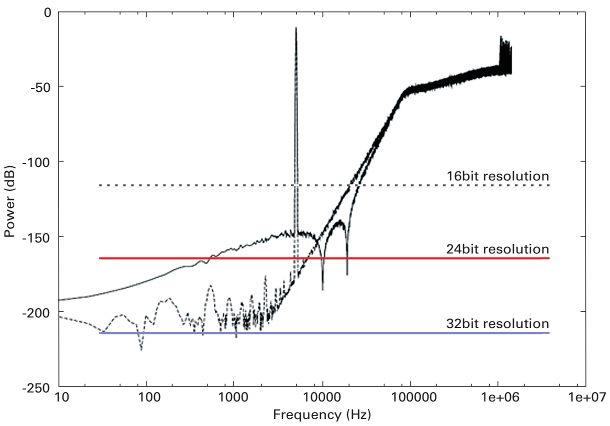

Figure 3: Typical Characteristics of a 1-Bit Quantized Delta-Sigma Modulation Signal Operating at 2.8224 MHz

*1 A continuous hiss noise generated when playing back magnetic recording tape.
*2 Low-pass filter (LPF): A filter that removes components of a frequency higher than the required frequency band. It is used to eliminate specific signals or noise.

ability to select various audio characteristics such as the fairly flat dynamic range under 20 kHz (solid line), or the aggressive dynamic range in lower audio frequencies (dashed line), based on the delta-sigma modulation design. Additionally, when we look at the red line in Figure 3 we can see that in both cases there are resolutions in excess of 24-bit in the audio frequency, and I believe the fact that resolution is not limited by quantization bits like PCM is a very important point for audio applications.

■ Transition from Physical Mediums to Digital Files

Here I would like to reflect on the history of online music distribution. SA-CD was released as a next-generation audio format and successor to the CD in May 1999, meaning that the shift to high-resolution audio for physical mediums started 17 years ago. However, commercial online music distribution also began in that same period. The “bitmusic” subsidiary of Sony Music Entertainment (which was integrated with “mora” in July 2007) was established in December 1999, coinciding with the launch of the NW-MS7 Memory Stick Walkman, and the company began online music distribution under certain master licenses. Given that dial-up access was still standard at the time, balance between the Internet access speed of general households and sound quality was taken into consideration, and music was sold as 132 kbps lossy compression format “ATRAC3” files incorporating Sony’s proprietary “OpenMG” digital rights management (DRM). However, the V.90 analog modems that were commonplace back then had a maximum downlink speed of 56 kbps (theoretical value), so even with a good connection it would take more than triple the actual playback time of music files to download. That meant when including connection fees for pay-as-you-go dial-up access, the cost was not reasonable at all. As a result, most music files on portable audio players continued to be data ripped*3 from CDs. In 2003, U.S. company Apple launched the iTunes Music Store digital music distribution site, and began selling 128 kbps lossy compression format “AAC” files incorporating their proprietary “Fairplay” DRM for 99 cents a track.

Around this time, xDSL access started gathering momentum amongst general households in developed countries, and the availability of broadband connections with effective downlink speeds of 1 Mbps resolved the issues with download time and connection fees. Then, the iTunes Music Store gradually expanded their services in the music market, becoming extremely successful. Subsequently, technological innovation in xDSL and the spread of cable modems and FTTH boosted effective downlink speeds to between a few Mbps and tens of Mbps. This began opening the way for high-resolution audio distribution, but there was still one more obstacle in its path. That obstacle was DRM, but this fact may not have been recognized widely.

With the spread of broadband access and the increase in transmission speeds, it became possible to download uncompressed CD quality (44.1 kHz / 16-bit) audio files faster than their actual playback time. Consequently, e-onkyo music*4 began the digital distribution of 96 kHz / 24-bit files as early as August 2005. However, at this time the files were encrypted through WMA lossless encoding, and it was not possible to play them on PCs other than the Windows PC that was used for the download, so this service was not well received at all even among the audiophile market. But when U.K. company LINN Records*5 started high-resolution digital music distribution without DRM in 2006, this began to catch on with audiophiles. U.S. HDtracks*6 entered the market in 2008, assembling high quality content from various music labels and distributing it as DRM-free high-resolution audio files. These services made it possible to enjoy the content without it being tied to devices. Simultaneously, U.S. audiophiles had just begun to transition to playing digital files rather than playing CDs as they are, and the number of people seeking high-resolution audio files for playback started to grow.

In response, PC applications and USB-DAC for high-resolution audio file playback were developed and introduced to the market in quick succession. In Japan, e-onkyo music commenced distribution of DRM-free high-resolution audio from July 2010. Initially, e-onkyo music also distributed CD quality (16-bit) files, but after seeing that high-resolution audio made up the lion’s share of their downloads, they shifted their focus to high-resolution audio distribution. In May 2013, they ceased distribution of CD quality files, and have since been working on increasing their lineup of music as a site specializing in the distribution of high-resolution audio.

*3 Ripping: Importing an exact copy of the digital data stored on a music CD to a PC as a disk image file, or converting it to audio data files in certain file formats that are easier to manage on PCs.

*4 e-onkyo music (http://www.e-onkyo.com/music/) (in Japanese).

*5 U.K. LINN Records (http://www.linnrecords.com).

*6 U.S. HDtracks (http://www.hdtracks.com).

The high-resolution audio files distributed on these services were PCM sound sources recorded or mastered at a high bit rate and high sampling frequency, and stored as uncompressed WAV/AIFF files or FLAC files with lossless compression, so DSD audio files were not available at that time. However, the situation changed completely in November 2010, upon the freeware release of Korg's*7 DSD-capable playback and editing application (Windows/Mac) "AudioGate V2.1," which had until then been bundled software for their own DSD-compatible recorder. AudioGate*8 enabled people to play back DSD files on PCs using its own downsampling function to PCM signal, or create DSD disks (with DSF format files recorded to DVD-R according to a specified directory structure) that could be played on the PlayStation®3 or certain SA-CD players, which gave DSD positive momentum as a distribution format.

OTOTOY*9 had already begun distributing their content in the DSD format from August 2010, and e-onkyo music entered the market for DSD format music distribution in December 2010. This led to the introduction of DSD compatible USB-DAC products to the market from various hardware manufacturers one after another from 2011. After that, the number of sites handling the digital distribution of DSD files also increased in the United States and Europe. Over the past few years, the DSD format has been becoming more and more popular on a global scale. In 2014, the appearance of native DSD music*10 specializing in DSD in Europe stunned audiophiles, and in 2015 major high-resolution music distributor HDtracks began offering DSD files.

Figure 4 is a basic summary of the developments of digital audio in the consumer market. Although there was a period of 17 years between the release of the CD in 1982 and the release of the SA-CD in 1999, it only took 11 years to transition from ATRAC3 distribution in 1999 to DSD distribution in 2010. Furthermore, in contrast with the fourfold increase in the bit rate between CD and SA-CD, DSD boasts 43 times the bit rate of ATRAC3, showing how amazingly fast the rise of high-resolution audio distribution has been. In 2013, 5.6448 MHz DSD music files that have double the sampling frequency of SA-CD came onto the market, along with 358.2 kHz / 24-bit WAV files (also known as DXD). As a result, DSD files with the same sampling rate as SA-CD that were simply known as "DSD" were redefined as "DSD 2.8M" or "DSD64," and 5.6448 MHz DSD files were given the name "DSD 5.6M"

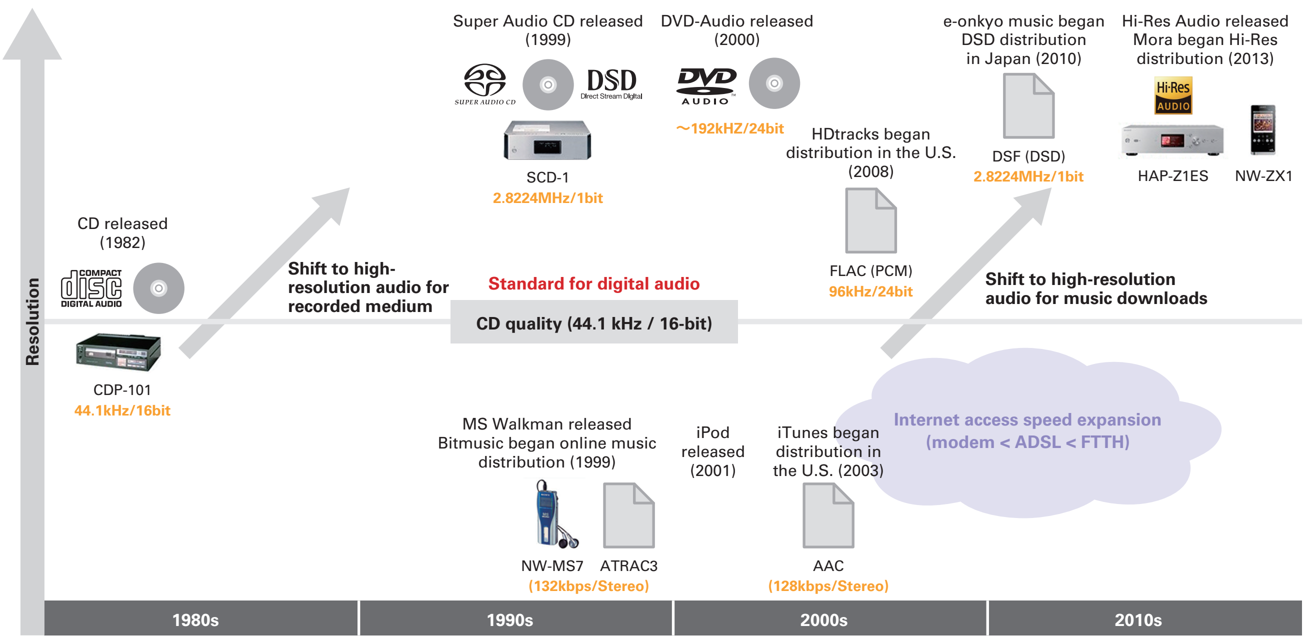

Figure 4: History of the Developments of Digital Audio in the Consumer Market

*7 KORG (http://www.korg.com/jp/) (in Japanese).
*8 AudioGate (http://www.korg.com/jp/products/audio/audiogate4/).
*9 OTOTOY (http://ototoy.jp/top/) (in Japanese).
*10 native DSD music (https://www.nativedsd.com).

or “DSD128.” Then in 2015, the distribution of DSD music files with a sampling rate of 11.2896 MHz (DSD 11.2M or DSD 256) even made an appearance, offering a wide dynamic range approaching 100 kHz. DSD is now liberated from the SA-CD medium, and has benefited significantly from the expansion of broadband access.

■ The Current Situation of Music Streaming Distribution

With the advances that have been made, you may expect that a shift to high-resolution audio would be on the horizon for music streaming distribution next. However, unlike music download distribution in which music files are downloaded at home where there is a good Internet environment, and then enjoyed either at home or while out, music streaming distribution is mainly targeted at mobile device users, so lossy compression audio between 128 kbps and 320 kbps is used for stable playback over mobile networks.

The major music streaming distribution services include those with a kind of personalized Internet radio format such as Pandora*[11] and the U.K.’s Last.fm*[12], which appeared in the early 2000s. These are basically free, with ad revenue serving as their main revenue source. The first real music streaming distribution service is said to have been Spotify*[13], which was established in Sweden in 2006, and launched its service in October 2008. Spotify has unfortunately not launched its service in Japan yet, but in 2015 Apple Music*[14], AWA*[15], and LINE MUSIC*[16] appeared in quick succession and started services, and most people would probably agree that the migration to music streaming distribution has truly begun.

These are still currently stuck as lossy compression audio between 128 kbps and 320 kbps, but in 2014 the Tidal*[17] service run by Swedish/Norwegian company Aspiro and France’s Deezer*[18] service began the paid music streaming distribution of CD quality audio as their top tier services using lossless FLAC compression. The compression ratio of FLAC format files varies depending on

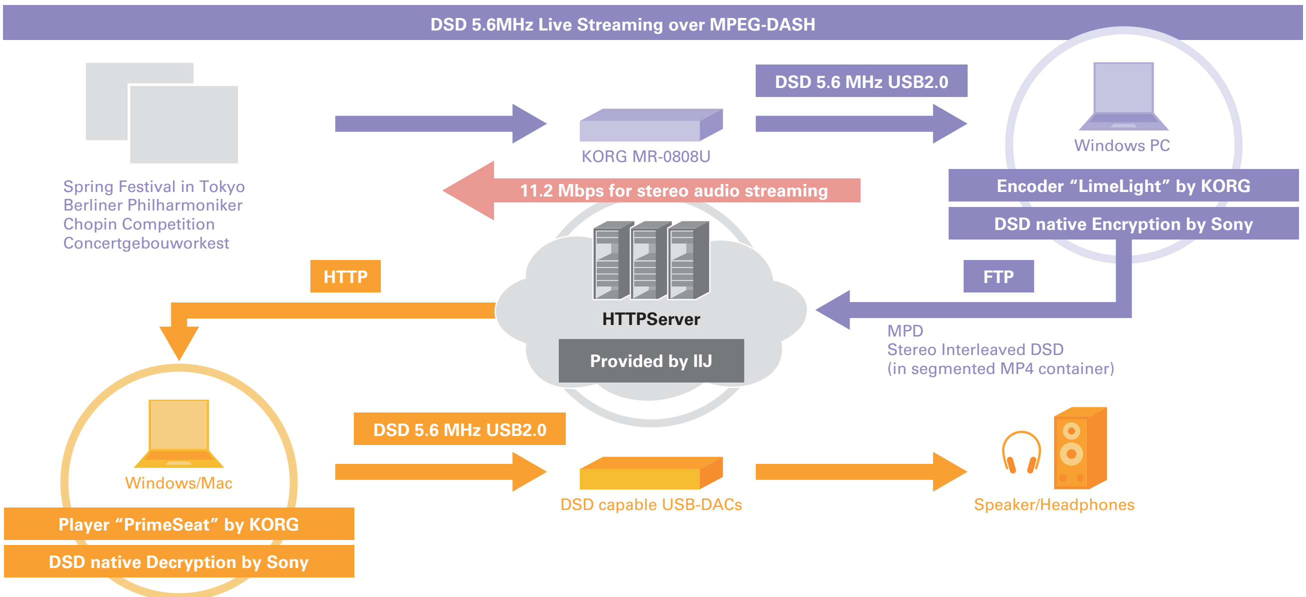

Figure 5: An Overview of the DSD Live Streaming System

*11 U.S. Pandora (http://www.pandora.com).
*12 U.K. Last.fm (http://Last.fm).
*13 Spotify (https://www.spotify.com).
*14 Apple Music (http://www.apple.com/jp/music/) (in Japanese).
*15 AWA (https://awa.fm).
*16 LINE MUSIC (https://music.line.me/trend) (in Japanese).
*17 Tidal (http://tidal.com/jp).
*18 French Deezer (http://www.deezer.com/).

the song, but is said to average about 70%, so a CD bit rate of 1.4 Mbps would be streamed at approximately 1 Mbps. This value is easily achievable under FTTH connection speeds, so it should also be possible to stream higher bit rate audio sources, or in other words high-resolution audio. However, considering the heavy load on distribution servers and network access, it is difficult to break into the market without strong infrastructure support, for business models that involve offering an extensive collection of music and streaming them to a wide variety of people. I also believe that it would be very hard to achieve this when relying on external resources for this infrastructure.

■ Activities of IIJ Aiming at High-Resolution Audio Streaming Distribution

In this market environment, IIJ launched a collaborative project between four parties, IIJ, Korg, Sony, and Saidera Mastering, to verify the feasibility of implementing high-resolution audio streaming distribution. We carried out high-resolution audio live streaming with DSD in the form of open experiments in April 2015. On April 5, we streamed the live performance of the "Marathon Concert at the Spring Festival in Tokyo" from the Tokyo Bunka Kaikan in Ueno, followed by the live performance of "Berliner Philharmoniker" from the Philharmonie in Berlin on April 11. As shown in Figure 5, DSD Live Streaming uses the MPEG-DASH method, splitting a DSD signal into segmented MP4 containers that are then uploaded in order. Streaming distribution is achieved by distributing these to the client from an HTTP server along with an MPD file.

For the proof-of-concept tests, we prepared two MR-8080U and LimeLight on the upload side to transmit both DSD 5.6M and DSD 2.8M simultaneously. Users with DSD capable USB-DACs (Korg and Sony products only) were able to listen using the PrimeSeat*[19] client software that was distributed for free in advance. Apart from some users who did not get a good listening experience due to last-one-mile instability, they could enjoy DSD Live Streaming in areas ranging from Japan, to the United States, Europe, and Asia. It could be said that this would be possible due to the infrastructure of IIJ, as we already have backbone networks in place and are able to optimize our Content Delivery Network (CDN) by ourselves.

In 2015, we performed DSD Live Streaming of the "Prize-Winners Concerts" of the 17th International Fryderyk Chopin Piano Competition in Warsaw, Poland on October 22 and 23, and the live performance of the Royal Concertgebouw Orchestra in Amsterdam, Netherlands on October 30 as well. For this streaming, we used the installed internet access of each performance venue. At around the same time there was the release of PrimeSeat v1.2, which enabled connection of DSD-capable USB-DACs other than those from Korg and Sony, and simplified listening without USB-DACs via real-time PCM conversion. Through this we verified that the hurdles for high-resolution audio streaming distribution and playback could certainly be overcome.

There is great significance in proving that stable high-resolution audio streaming is possible using the DSD format which is the most challenging of all high-resolution audio sources. Since April of last year we have received comments expressing surprise and anticipation from people all over the world who did not expect they would be able to listen to high-resolution audio streaming so early. On December 23 last year we launched the PrimeSeat service, which provides the world's first DSD Internet radio in addition to live and on-demand streaming distribution. We continue to make further progress toward building this business, with the aim of enhancing our high-resolution streaming services. There will be more developments in the PrimeSeat service to look forward to in the future.

Author:
Ayataka Nishio
Mr. Nishio is a Senior Engineer in the Operation & Engineering Section of the Application Development Department of the IIJ Product Division.
In his former job he was mainly engaged in the development of professional audio equipment and signal processing, before taking on his current position from October 2015.
He is now striving to make high-resolution audio streaming distribution a reality, putting to use over 25 years of experience in music production, as well as SA-CD, DSD, and high-resolution audio development.

*19 PrimeSeat (http://primeseat.net/en/).