world class sound...

Naim NDX Network Audio Player

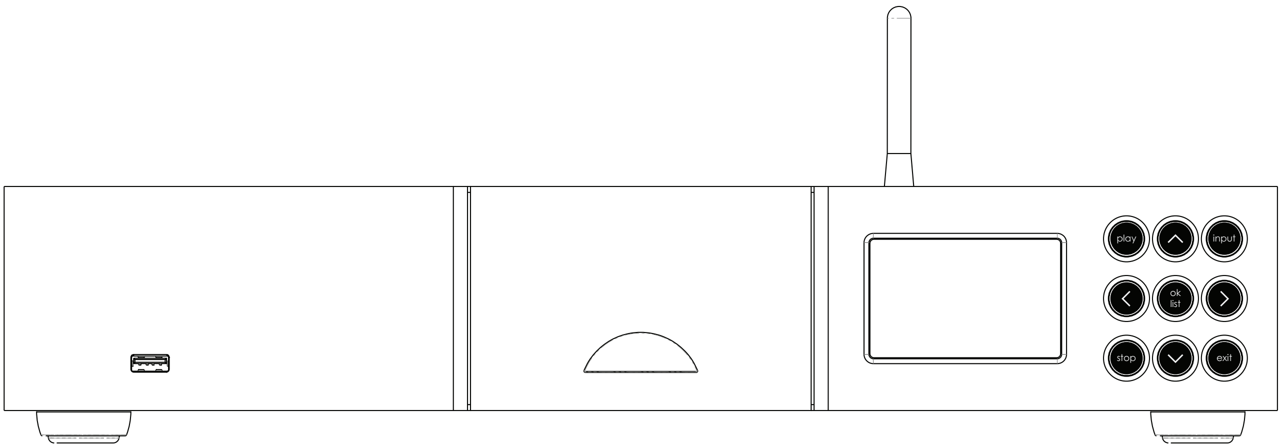
DESIGN, ENGINEERING AND TECHNOLOGY

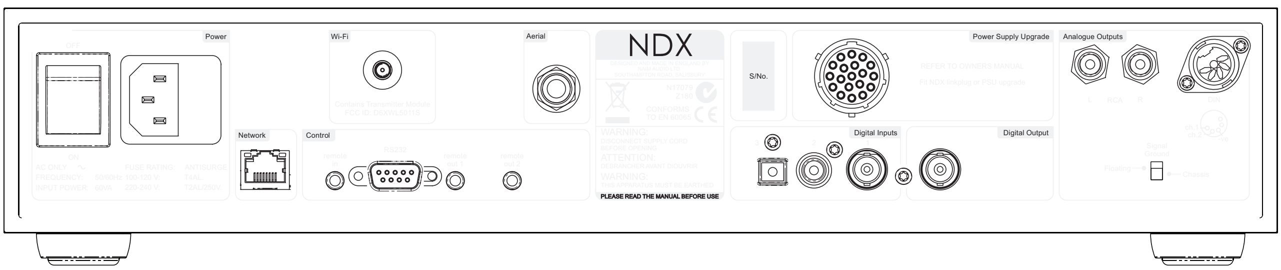
By Gary Crocker and Hjalmar Nilsson September 2010

Naim Audio Ltd · Southampton Road · Salisbury · Wilts · SP1 2LN · England

Objective

To design an upgradeable network audio player that is sonically comparable to a high-end CD player.





Functionality

Naim's NDX is an audio source product that integrates network streaming, internet radio, iPod digital, replay from USB memory stick and an optional DAB/FM tuner all in one box. It also has three S/PDIF inputs (two transformer-isolated coaxial inputs and one buffered optical input) for the connection of other digital sources.

As standard the NDX is controlled either via the front panel display and buttons or by the supplied remote control. With purchase of the Naim n-Stream application full functional control is available via hand-held devices such as the Apple iPad or iPod Touch.

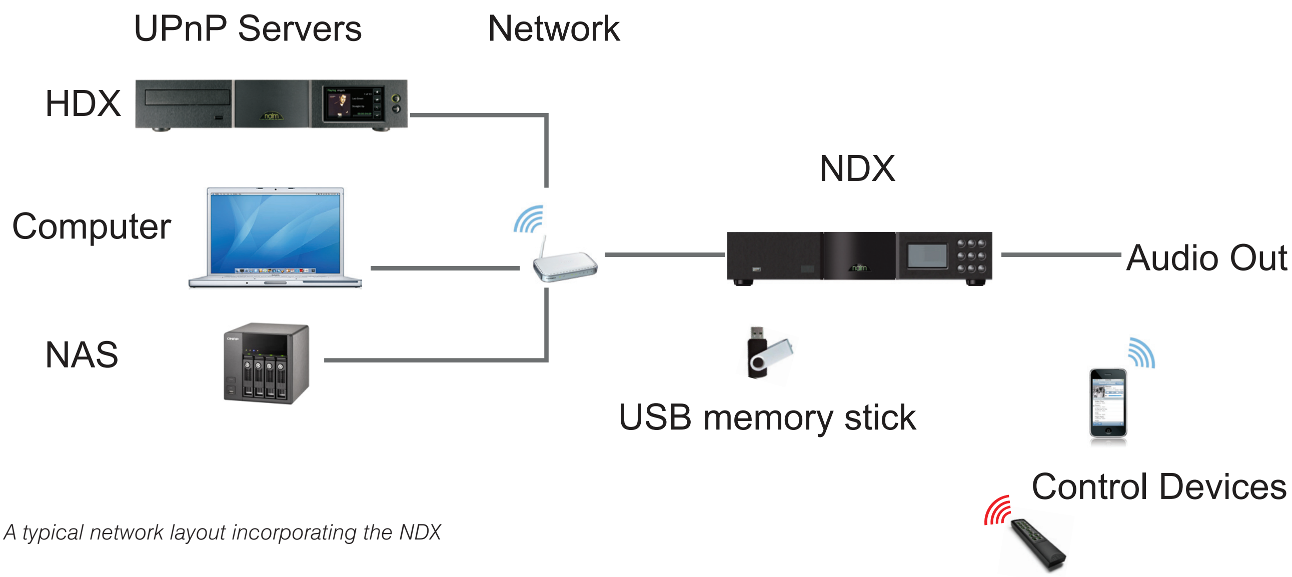
Network Streaming

Using the UPnP™ network protocol the NDX plays music stored on a hard drive across a standard ethernet network. With today's trend of downloading or ripping music to hard drives the ability to achieve audiophile sound quality from music data stored on an HDD or NAS is paramount. Good engineering techniques and principles were utilised to realise this goal with the NDX, as with all Naim products, but this project posed some new and interesting challenges.

The first requirement for achieving high quality sound reproduction from a hard drive is that music data ripped from CD must be of the highest possible quality. So the ripping engine developed as part of Naim's HDX hard disk player became the first building block of the NDX.

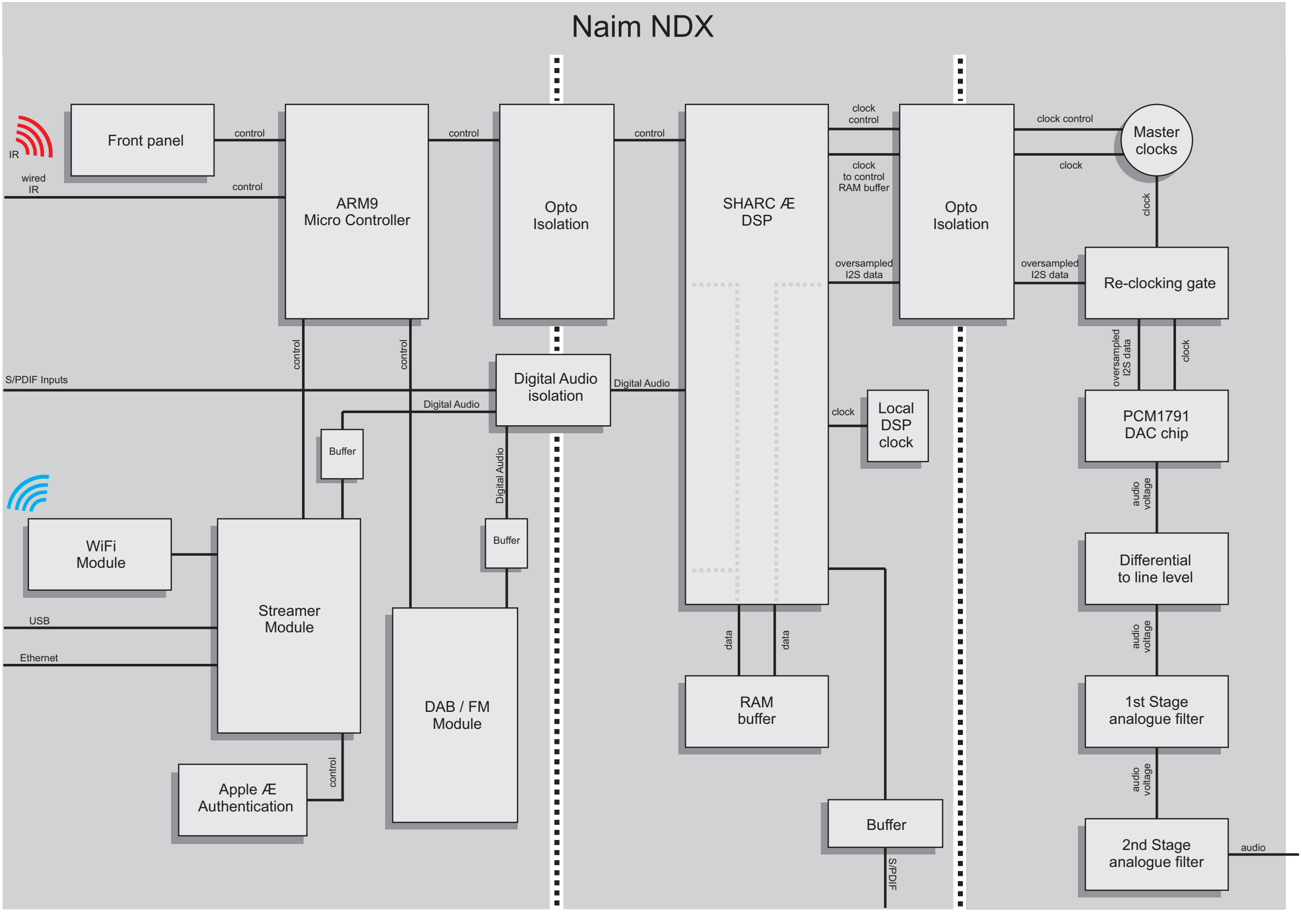
Although the HDX can also play audio files from network devices, it and the NDX work differently. The HDX is a hard disk audio player where the music is played from the internal hard drive or from music files found on the network. The HDX scans the network to find playable media which are then added to the database and can be played. By contrast, the NDX uses the UPnP™ network protocol to find UPnP™ servers on the network and play audio files served by these servers. As part of the NDX's development Naim has tested many UPnP™ servers and we have written our own UPnP™ server for Naim's own server products. So now the HDX and the new UnitiServe are not only hard disk players but also UPnP™ servers, from which the NDX can stream audio.

Naim's UPnP™ servers deliver the uncompressed audio data ripped from CD using the Naim ripping engine to ensure the best quality reproduction. Uncompressed audio data will always give better results than compressed. Even lossless compression may not reproduce audio with equivalent quality to the uncompressed original as the processing required to uncompress the data increases the computational load. This raises the power supply noise floor, which detracts from the sound quality.



A typical network layout incorporating the NDX world class sound...

NDX Internal Architecture



Dataflow

When the NDX requests a music file, the UPnP™ server concerned will packetize the file into a TCP (Transmission Control Protocol) stream for transport across the network. The NDX accepts this TCP stream, unpacks it and buffers the data in memory. This buffering takes up the inevitable variations in network latency which could otherwise cause audible clicks and pops or even drop-outs. From this buffer, located immediately after the Streamer Module, the data is clocked out to the DSP as a digital audio stream, using differential transmission to reduce radiated electrical noise from this fast digital signal. This differential digital stream is then galvanically isolated from the DSP section using a high-speed pulse transformer. The full internal functions of the DSP are explained later in this document. Buffered I²S data output from the DSP is opto-isolated before the DAC section, where it is converted to an analogue signal. The analogue signal is then filtered to remove the out-of-band artefacts that are a by-product of the digital to analogue conversion process.

Isolation

To reduce noise transfer to the analogue domain from its digital circuits and connections the NDX incorporates galvanic isolation between key sections of its circuit. The digital audio data source is isolated from the DSP and, in turn, the DSP is isolated from the master clock and DAC. In this way the maximum isolation is achieved between digital sources and analogue output. Power supplies for each section of the signal path are also independent, as explained later.

DSP Section

The NDX shares some of the technology developed for the Naim DAC, as the three following sections ('RAM buffer jitter removal', 'Oversampling and analogue filtering' and 'Digital filtering using Naim algorithms') explain.

RAM buffer jitter removal

Naim's buffer or memory method of jitter removal relies on a simple concept: the audio data is clocked into the memory at the incoming, inconsistently timed rate and is then clocked out of the memory and into the DAC chips using a precise clock. The rate at which the memory fills and empties is controlled by selecting the master clock that best matches the average incoming clock frequency. In this way the data entering the DAC chips is completely isolated from the incoming jitter. Only in rare cases will none of the Naim NDX's selectable master clocks be closely enough matched to the incoming data rate. To cope with this eventuality we have also implemented, as a backup, an asynchronous sample rate converter (ASRC) that will accept any sample rate from 32kHz −10% to 192kHz +10%.

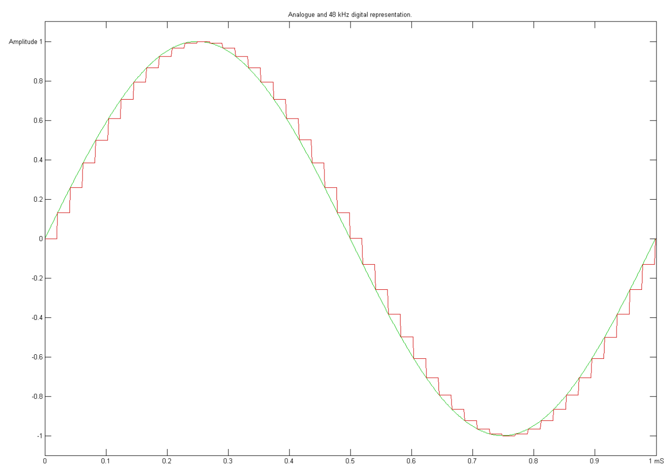
Oversampling and analogue filtering

When an analogue signal is converted to digital form it is no longer continuously variable: it is now a discrete representation of the original. This means that the signal amplitude is only known at certain, regularly spaced discrete time intervals determined by the sample rate. For CD the sample rate is 44.1kHz (44,100 samples per second) and therefore the time interval between each sample is 22.7μs (microseconds). To recreate the analogue signal, the DAC chip holds each sample value until the next arrives, resulting in a staircase waveform rather than the smooth, continuous original. An example is shown below.

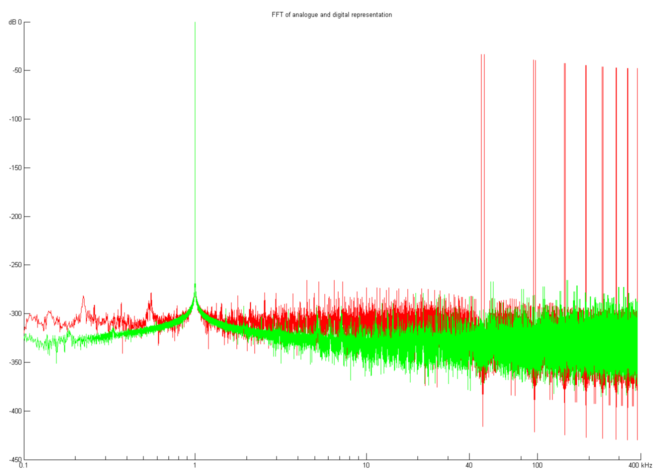
Looking at the staircase signal in the analogue time domain, as below, it may not look too bad. But if we look instead at the frequency spectrum of this signal (below right) we see that instead of containing just the single frequency of the original analogue signal, the staircase waveform contains a lot more due to its sharp high frequency steps.
If the analogue signal is a 1kHz sine wave (below right), then the frequency spectrum of the staircase equivalent will also display a peak at 1kHz. But it additionally has peaks at every multiple of the sample rate plus or minus 1kHz (*ie* 47 and 49kHz, 95 and 97kHz, 143 and 145kHz, etc). In order for the DAC output to be as close as possible to the analogue original, these extra frequency components have to be removed by filtering out everything above half the sample rate. Then only those frequency components that were present in the original signal are left.

This requires a very steep low-pass filter. Achieving adequate performance using an analogue filter is extremely difficult. It requires the use of costly, high-precision components, and even then the filter performance may change with temperature, loading, etc. One way of relaxing the constraints on the analogue filter is to increase the frequency space between the audible band (up to 22.05kHz at 44.1kHz sample rate) and the first of the unwanted frequency components. One popular method of achieving this is called oversampling and is described below. By oversampling we can use a relatively simple analogue filter to remove the remaining high frequency components that the DAC introduces.

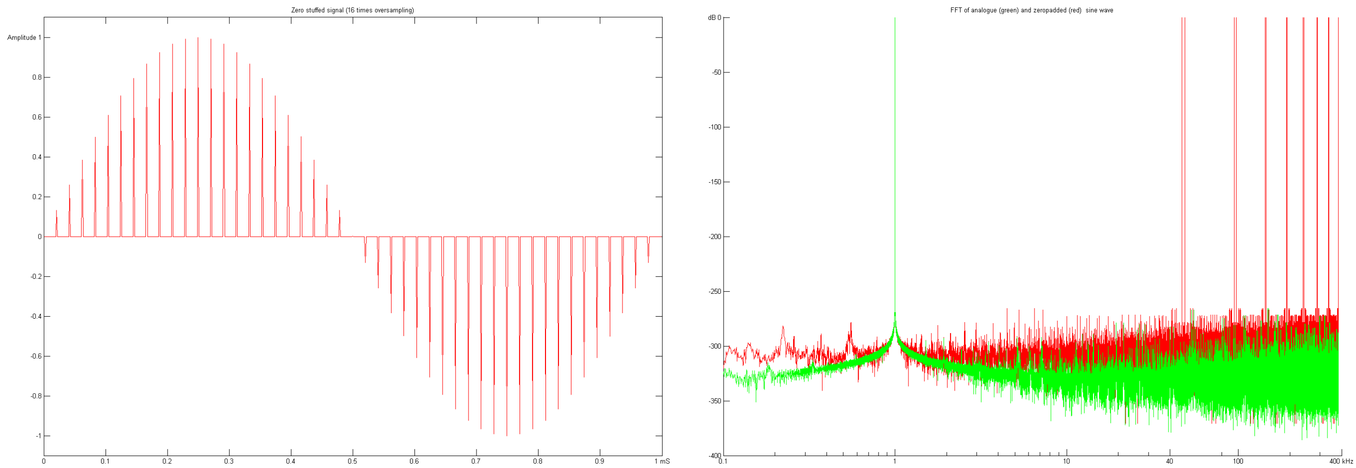
To increase the sample rate we have to insert additional samples between the original samples. If we want to double the sample rate then we need to insert one extra sample between each two original samples, if we want to quadruple the sample rate then we need to insert three extra samples between each two original samples, and so on. But what sample values do we put there? If we oversample by a factor of two we could perform linear interpolation, so that each additional sample has a value half way between that of the original samples on either side – but that simplistic approach will create a lot of unwanted frequency components. In fact, linear interpolation is just a very basic low-pass filter. We can do better than that!



1kHz sine wave (green trace) and the staircase equivalent generated by a DAC chip (red trace). Sample rate is 48kHz



Spectral analysis of the original signal (green trace) and the unprocessed DAC output (red trace).
Note that the peak in the green trace at 1kHz overlays an equivalent peak in the red trace. Sample rate is 48kHz



The same 1kHz signal as before (left) and its spectrum (right), sampled originally at 48kHz but now zero-stuffed with 15 zero-valued samples between each adjacent pair of original samples. The sample rate is now (16 x 48 =) 768kHz

Fortunately there is a smarter way of generating the new samples. First we 'zero stuff' the digital signals – in other words, insert the required extra samples but give them all a value of zero. When we do the result is a waveform and spectrum like those shown above.

Note that the spectrum is very similar to that shown previously but all the out-of-band peaks are now of the same height as the in-band peak (0dB). If we can remove all the extra tones (above 24kHz in the example below) then we will create a DAC output signal that has smaller steps – provided, of course, that the DAC supports the new higher sample rate. That in turn means that the remaining unwanted frequency components will be higher in frequency, allowing them to be removed using a simpler low-pass analogue filter.

Digital filtering using Naim algorithms

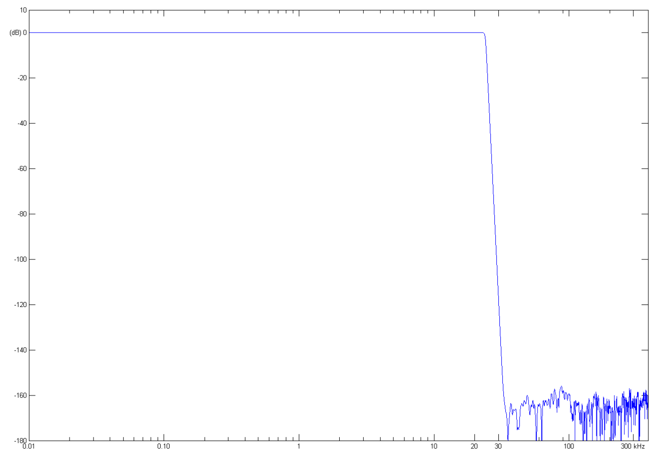
If we insert 15 zeros between each pair of input samples as the first stage of 16x oversampling, the frequency spectrum of the zero-stuffed signal will contain lots of unwanted components, as shown previously.

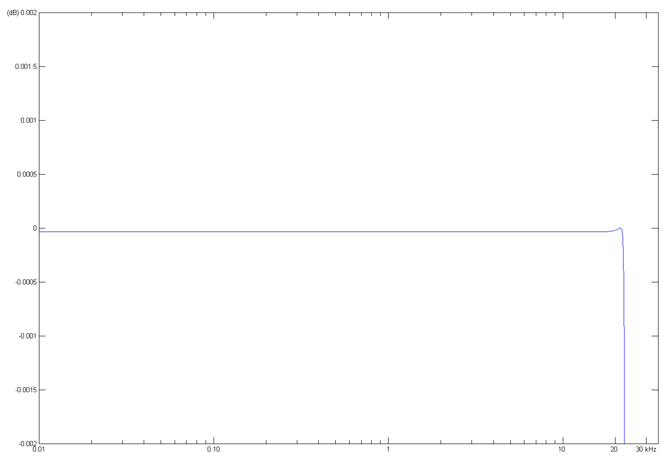
All the frequency content occurs above half the sample rate, ie above 22.05kHz in the case of CD, and must be removed by the digital filter. While it is fairly easy to design a digital filter that applies a high degree of attenuation to everything above 22.05kHz while having minimal effect on frequency response below 20kHz, we need to ensure that we use sufficient mathematical precision to keep the filter's arithmetic noise below –144dB (the theoretical signal-to-noise ratio of 24-bit PCM audio) at all times.

As the incoming audio can have up to 24-bit resolution, 24-bit processing will not be sufficiently accurate as it will incur rounding and/or truncation errors of the same resolution as the incoming audio. The normal solution is to use 32-bit floating point processing instead but in 32-bit IEEE floating point arithmetic the mantissa is 24 bits long, so if we have sample values close to 1 – where the exponent will be 0 – we will again achieve only 24-bit precision. (For readers unfamiliar with floating point arithmetic, numbers are represented in the basic form 1.23456E+5, where the first part of the number – 1.23456 in this case – is termed the mantissa and the second part, here +5, is the exponent, ie the power of 10 by which the mantissa is multiplied. In conventional integer arithmetic 1.23456E+5 is written as 123,456. Note that IEEE floating point numbers use base 2 (binary) rather than base 10 representation, but the structure is the same.)

In reality it is very hard to estimate the resulting precision using floating point numbers. The best way to assess it in the context of filter designs is to use real-world signals, send them through the filter and check the outcome to see what precision was achieved. A quick test along these lines verifies that 32-bit floating point processing does not achieve the precision we require. Fortunately, more powerful processors from Analog Devices support a new type of floating point number, 40-bit floating point. Instead of having a 24-bit mantissa these use a 32-bit mantissa with 7 bits of exponent. This gives the resolution we require but keeps memory requirements fairly low (only one more byte compared to 32 bits). If we send music signals through our implementation of a filter with 40-bit floating point and analyse the results we see that arithmetic noise is at about –180dB, which is well below the noise floor of the DAC chips.

When it comes to choosing how to implement the low-pass filter, it is all a matter of taste and how well you can implement the method you choose. Having listened to IIR (infinite impulse response) and FIR (finite impulse response) implementations of filters with identical amplitude and phase responses, we found that we prefer the IIR implementation. Even though, in a perfect world, FIR should be able to implement any IIR response to the resolution of the DACs, we found that processing loads, arithmetic noise, etc had greater influence on sound quality than the phase errors that IIR filters inherently introduce.





The Naim DSP filter response and passband ripple, input sample rate 48kHz

The NDX uses the same 16x oversampling filter as the Naim DAC, which is implemented in the SHARC processor. The chosen filter is a modified Butterworth to which additional poles have been added to prevent too much phase shift occurring within the audio band. To ensure both low arithmetic noise (fewer additions and multiplications that cause rounding) and low power supply noise (since the DSP draws less current when it isn't calculating) the filter is implemented as efficiently as possible, using only five lines of assembly code.

In addition, all code that runs in the SHARC processor has been tuned by listening tests to maximise sound quality. With every increase in number crunching the DSP consumes more power, and the more power the DSP consumes the greater the power supply noise. By optimising the DSP algorithms and controlling the way the data is buffered, power supply noise is kept to a minimum – to the benefit of audio performance.

Master clock

The stability or noise of a DAC's master clock has a direct influence on the audio output. For example, if the clock frequency were to increase momentarily by one per cent then the analogue output frequency would increase by the same amount. So the NDX's master clock has been designed to oscillate with extremely low noise. There are many types of clock frequency instability with many different causes but by careful design these influences can be minimised.

Phase noise is random and can be caused by internal and external influences on the clock. Power supply noise, for instance, can influence its oscillation, so in the NDX the master clock has its own voltage regulator to prevent external noise influencing stability. But poor printed circuit board (PCB) layout can introduce power supply noise even when a regulator is used. The NDX PCB has six layers and is laid out such that circulating currents are kept local. All decoupling capacitors are connected with short, wide tracks to ensure low impedance and less HF noise on the power supply lines.

When the NDX is playing audio via S/PDIF, the source of the S/PDIF signal is the master and the NDX's master clock must, on average, match the frequency of the source clock. Some products achieve this either by using an asynchronous sample rate converter (ASRC) or a voltage-controlled crystal oscillator (VCXO) together with a Phase Locked Loop (PLL).

ASRCs can work well but they rely on considerable mathematic processing, which will cause problems if not implemented very carefully. There are no ASRC chips on the market that can convert arbitrary input sample frequency to high enough frequencies that the analogue output filter can be kept fairly simple. This means that you still have to oversample the data before the DAC, so you might as well do that properly from the outset and avoid the ASRC entirely. Moreover, since ASRC relies on averaging the incoming sample rate, not all jitter will be removed. The NDX uses the built-in ASRC in the SHARC DSP only when the incoming data rate is outside the S/PDIF spec. The SHARC's ASRC has an arithmetic noise level of –128dB and so it is only used in these exceptional circumstances.

VCXOs are a good solution too but since a VCXO's output frequency is controlled by a voltage, if noise is present on the control line then the VCXO's output will exhibit phase noise. Normally the control line to the VCXO is feed via a PLL which means that it is constantly changing to match the incoming clock rate.

The NDX approaches the problem of matching its clock frequency to the source in a quite different way. Its master oscillator offers ten switchable fixed frequencies which are selected so as to keep the average clock frequency the same as the source's. The master clock in the NDX is a VCXO controlled by two 9-bit DACs connected to its control line to give 10-bit resolution (1024 discrete steps). This method was chosen to give greater resolution, ie smaller step changes, in the master clock frequency. The 9-bit DACs are controlled by the DSP in software so the rate at which the master clock is altered can also be controlled. This is important because if the master clock is changed too quickly this can cause audible artefacts, ie introduce jitter. The SHARC DSP monitors the rate at which the RAM buffer is either filling or emptying and changes the clock frequency only if the buffer is going to either overflow or underflow. This way the incoming clock jitter is completely isolated from the NDX master clock. When the system has settled, it will only modify the master clock frequency every 10 to 15 minutes.

DAC

The DAC chip in the NDX is a Burr-Brown PCM1791A used in external oversampling mode. This is the same Delta-Sigma DAC used in other high quality Naim products but in the NDX we switch off the internal oversampling and digital filter. The oversampling and digital filtering is instead done externally by the Naim-written code in the DSP. At 44.1kHz and 48kHz sample rate the oversampling factor is 16x, which runs the DAC section of the PCM1791A at its maximum sample rate of 768kHz. The internal Delta-Sigma modulator in the DAC oversamples the incoming data by a further factor of four, so the effective oversampling is 64x.

Analogue output filter

The DAC's output signal has to be filtered to remove the remaining unwanted frequency components using an analogue low-pass filter. In the NDX this is multi-stage seven-pole design using Burr-Brown OPA42 op-amps; as these are single-chip devices a total of six are required. This op-amp was chosen from the plethora of available audio op-amps after many hours of listening tests as having the finest sound quality. A combination of Sallen-Key and multiple feedback low-pass filters are used to implement the seven-pole filter.

The highest quality through-hole components are used in the signal path rather than surface-mount equivalents. They are tested and matched into tight tolerance groups to ensure correct frequency response.

Low noise power supply

Reducing power supply (PSU) noise has long been part of Naim's design philosophy. To increase perceived and measured dynamic range, PSU noise in the NDX has been reduced to an extremely low level. The NDX can be powered either from its internal PSU or, as an upgrade, from either the XPS or NS 555PS external PSUs.

The internal PSU begins with a custom-designed toroidal transformer. Toroidal transformers have very low magnetic leakage, which ensures that electromagnetically induced mains noise is low. The NDX transformer has four isolated secondary windings, feeding four sets of rectifiers and reservoir capacitors:

1) Digital (front panel; ARM9 microcontroller; Wi-Fi module; streamer module; optional DAB/FM module)
2) DSP (SHARC DSP; RAM buffer; local DSP clock)
3) Audio clock (master clocks; master clock control circuits; re-clocking gate; digital section of the DAC)
4) Audio filters (differential to line level; first-stage analogue filter; second-stage analogue filter; output relays)

Reservoir capacitors are larger than those typically used, in order to reduce the unregulated voltage noise and provide increased short-term current capability. The four separate PSUs form part of the electrical isolation of the digital circuits from the DAC chips and analogue circuits. Low noise LM317/337 regulators smooth the unregulated voltage from the reservoir capacitors. Voltage supplies to many of the digital circuits are double- and in some cases triple-regulated to reduce noise still further.

External PSU upgrade

When the PSU upgrade option is used with the NDX, power supply separation is maintained as the external power supply also has independent power supplies and ground connections. In addition it provides a larger toroidal transformer and bigger reservoir capacitors to reduce noise still further. Outputs of the external PSU are regulated using selected low noise LM317/337 regulators to achieve a lower noise floor ahead of the internal voltage regulation.

Other sonic influences

As with all Naim designs, the influence of vibration – induced microphonic noise – has been minimised in the NDX. The starting point is a rigid chassis made from high-grade aluminium. Once assembled in its aluminium sleeve the NDX has a 7.5mm-thick base, giving the electronics an extremely rigid foundation. To minimise the transmission of vibration to the FR4 fibreglass printed circuit boards, the PCBs are screwed to the chassis only at carefully selected points.

Analogue stage filter capacitors that work in the audible band are prone to microphony, so susceptible components are lifted from being in direct contact with the board to minimise this effect.

PCB design for digital circuits is as critical to performance as it is for a high-end power amplifier – possibly more so – and different techniques are required for digital and analogue board design. There are high-speed DSPs, high-resolution digital circuits and low noise analogue circuits present, all of which require different design approaches. In the NDX the DSPs are placed well within the boundaries of a six-layer PCB. Multiple power pins are decoupled using multiple decoupling capacitors mounted on the bottom of the board, and the DSP also relies on inter-plane capacitance to provide low inductance decoupling. External peripheral devices such as the SDRAM are placed as close as possible to the DSP ports to minimise the loop area of high speed circulating currents. All high speed traces are measured and resistively damped accordingly to ensure that digital signal edges are not subject to ringing since ringing radiates high frequency electromagnetic energy that could affect the analogue circuits.

The master clock drivers and timing gate are decoupled differently to slow-speed PCB design. For best performance the decoupling capacitors are placed topside to avoid using the inductive PCB vias. Decoupling traces dominate the topside, being as wide and short as possible since every millimetre of PCB trace length has a significant effect when clock signals are at 24MHz with nanosecond rise times. All high frequency traces are sandwiched on the central layers of the PCB where they are shielded by the ground and power planes that create the high frequency return current paths. Outer layers are used for low frequency signal and clock decoupling.

In the analogue circuits, different techniques are required again. Here large planed areas of conductor are detrimental to sound quality and are replaced with the star ground techniques that can be found in all Naim analogue products.

Gapless playback

The NDX supports gapless audio playback with both compressed and uncompressed audio formats. In this mode there is no two-second gap between tracks, which can be so annoying when listening to a live recording or a classical music recording where the microphones remain 'open' between movements.

Digital output

A single S/PDIF output is provided on the NDX via a BNC connector, chosen for best sound quality and impedance matching. The S/PDIF output is both buffered and isolated, the isolation being achieved using a high quality pulse transformer. The buffer, which is used to drive the S/PDIF currents through the pulse transformer to the digital receiver, isolates the DSP from the high transient currents required for S/PDIF output. As the S/PDIF signal is RAM buffered in the same way as the digital data going to the internal DAC chip, the same level of jitter removal is applied whether the output is analogue or digital. Digital output is switchable and is should be disabled when the analogue output is being used.