Audio Engineering Society

Convention Paper 8347

Presented at the 130th Convention
2011 May 13–16 London, UK

High quality 'radio' broadcasting over the internet

David R Errock[1]

[1] Wave Science Technology Ltd, London, UK
david.errock@wavesciencetechnology.co.uk

ABSTRACT

Broadcasting audio over the internet has dramatically improved the experience for listeners, including removing the geographic boundaries of terrestrial transmission, allowing playback on portable devices and time-shifting content.

Internet audio has primarily increased choice, with traditional broadcasters now competing alongside 'internet only' music stations. Internet distribution can also increase the audio quality beyond that of terrestrial and satellite transmission channel constraints; unlike video streaming, which is of lower perceived quality than traditional transmission.

With the sustainable data capacity of internet connections increasing, lossless audio can be streamed live to the consumer. This paper discusses the issues relating to these distribution techniques and the challenges which will be experienced by the broadcasters, the distribution network, receiver manufacturers and listeners.

1. INTRODUCTION

1.1. Background

Radio stations have been streamed over the internet to listeners for some years. The services were originally used by 'early adopters' and those who were listening to stations beyond the geographic boundaries of the traditional transmission area, including special-interest stations. Internet listening is now mainstream for both live and 'listen again' recorded audio from traditional radio stations, adopted by listeners as a viable alternative alongside analogue FM, digital radio (such as DAB) and digital TV (DSat, DTT). In the UK, the take-up of broadband connections has helped to increase the listening figures although, at the end of 2010, internet radio still only accounts for around 3%[1] of all radio listening – a 50% increase since the end of 2009.

Perceived audio quality has improved since audio streaming started, through improved codecs, higher bit rates and greater quality control in the studio centre. The playback quality through computer sound interfaces and consumer equipment has improved too. Overall, the audio performance can be impressive, so that stations streamed at 192kbps in AAC are widely perceived as equal to or exceeding the quality from FM transmission.

There is potential to improve internet audio quality further; digital satellite and terrestrial broadcasting systems do not offer this option as their capacity is finite.

Broadcasters now typically serve a number of streams of different bit rates, codec formats and 'wrappers' to suit the listeners' needs. This targeting of particular receiving situations and devices is something that was not possible with terrestrial broadcasting. This principle can be extended to offer very high quality streams for interested listeners. This is more than the choice of codec or bit rate.

Very small radio stations have the flexibility to trial new technologies and re-engineer infrastructure easily. Conversely, major broadcasters have varied programme content, many listeners, legislative restrictions and complex infrastructure such as the scale of studio centres. This paper discusses the issues, benefits and challenges of using the internet to carry high quality services to listeners faced by major broadcasters with 'high-value' programme content. It looks beyond the current practices to assess the potential for lossless streaming and for high resolution streaming with higher sample rates and bit-depths.

1.2. Possibilities from internet audio streaming

Traditional one-to-many broadcasting offers only one path to all listeners, with the broadcaster's costs linked to the geographic coverage area. If a radio station were to set up two FM services carrying identical programmes – one for a small, special-interest group such as audio quality enthusiasts and one for most listeners – the per listener costs for the minority group would be prohibitive. The regulatory authorities would also be unlikely to licence such a service!

Internet audio streaming allows a different approach. Adding streams to an existing radio station has relatively insignificant studio centre installation and operating costs and uses the same data network to deliver audio to listeners, regardless of stream bit rate. Therefore, higher quality services can be streamed to audio quality enthusiasts, with the only additional broadcaster's costs being the difference between the standard bit rate and the higher bit rate for the small listener group.

2. QUALITY

2.1. Perceived quality

The traditional broadcast transmission systems are capable of delivering high technical quality and high availability – which the listener has come to expect as the norm.

The implementation of internet broadcasting has improved since its inception. The pursuit of 'higher quality' services over the internet must not worsen the overall experience to the listener.

2.2. Technical quality

Achieving good technical audio quality within a modern digital studio centre is well understood. A standard professional digital audio path can provide technical performance which exceeds that of most listeners' playback equipment, with near perfect frequency response (including to DC!), excellent dynamic range, very low noise floor and low distortion.

Modern lossy codecs are efficient and, if fed with clean audio, can produce excellent technical results. Basic steps can be followed to gain the best quality from the encoding process. Sufficient audio headroom must be maintained within the encoder to allow for potential clipping in the listeners' playback equipment. It is important to feed the broadcaster's encoder directly from the digital audio infrastructure in the studio centre and that a minimum of other processes affect the audio, which may in turn reduce the perceived quality or the efficiency of the encoder. For example, audio watermarking encoders effectively generate 'spurious' frequency components; loudness processing, as well as changing the dynamic range, also by its nature produces 'distortion' components.

2.3. Quality of delivery

The quality of delivery is important for audio. Listeners have come to expect their 'radio' to deliver continuous, silence-free and glitch-free audio into the home. This is achieved with traditional broadcasting transmission by systems which offer very high reliability, with availability in excess of 99.98%. There is built-in resilience of all aspects of the transmission chain, from backup power supplies and redundancy of critical components to instantaneous switchover between diverse incoming audio circuits – all designed to provide seamless audio during failures.

Broadcasters cannot achieve the same level of availability with internet radio. The studio centre systems can be built to similar levels of resilience, but the distribution is, to some degree, outside of the broadcaster's control as no one organisation can control all parts of the network or can guarantee a Quality of Service level. The data transport mechanisms are tailored to packet transfer, not to audio's continuous transmission. If any of the individual links in the network becomes congested, this will affect the delivery of audio to the end user.

Internet radio is, therefore, statistically more likely to suffer from silences than a dedicated transmission network. Any premium internet audio service will require higher bit rates to be streamed, which will make these effects more likely.

3. BACKGROUND TO STREAMING

3.1. Introduction

The basic principles of streaming are straightforward. In the most common 'unicast' model, the listener's playback device opens a control dialogue with a streaming server and negotiates for a unique audio stream to be sent to the listener's playback device, either as raw audio or contained within a 'wrapper'.

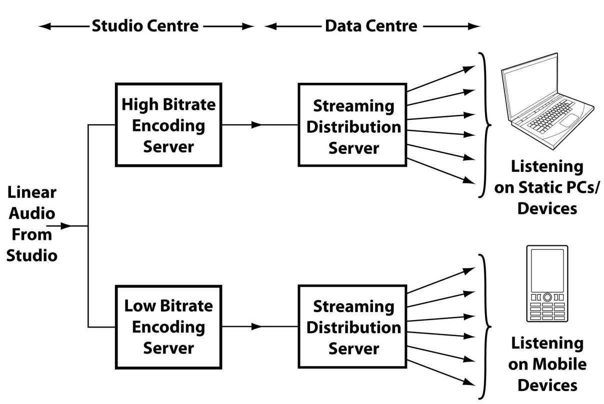

Figure 1 Simple streaming model

Audio is buffered to accommodate some network errors. While all parts of the network can continue to sustain the chosen bit rate – the link from the studio centre to the streaming datacentre, the datacentre's link to the Internet Service Provider (ISP) and the ISP's link to the listener – the audio stream will continue seamlessly.

3.2. Unicast versus Multicast streaming

Traditional broadcasting transmissions make their signals 'freely' available to any number of receivers, without restriction. Internet delivery of audio to listeners is fundamentally different, with the listener's 'receiver' requesting a stream to be delivered.

Currently, all main internet radio services use unicast streaming, in which the broadcaster must originate one stream directly for each listener and therefore the peak number of listeners is a critical factor in the distribution system design. This current configuration of internet radio distribution as a point-to-point unicast model, although feasible for low-volume delivery of radio, is a constraint to high-volume use, both in terms of the number of simultaneous listeners for 'standard' bit rate services and of the total cost of streaming for higher bit rate services.

A 'multicast' infrastructure would allow the same stream of data from the broadcaster's encoder to be delivered to many users simultaneously, effectively with the audio data being replicated in network switches. Multicast streaming is theoretically possible, but not yet implemented sufficiently widely in the internet to be practical for many users. If multicast protocols were adopted by internet providers, it would remove the limits of peak audience demand suffered by unicast streaming, making internet radio more like traditional broadcasting transmission.

Multicasting could help to enable higher bit rate radio services, by dramatically lowering costs for the broadcasters, but there are no business drivers to encourage the implementation of this technology. The investment impact for internet providers would be negative – the return for their investment would be a reduction in revenue. The most commonly cited potential for multicasting is for the delivery of live (broadcast) TV over broadband networks, but it could very well be used to deliver broadcast radio as well.

3.3. Distribution options

With unicast streaming where a unique stream is served to each listener, broadcasters have two choices: to build a distribution network in-house or to pay a third party distribution agent service – a Content Delivery Network (CDN) – or to use a mix of both. These concurrent unicast streams place demands on the distribution hardware and on the throughput capacity of the data links from the broadcaster to the network edge, making a reliable in-house distribution network a major investment. Catastrophic failures are also possible – for example due to unexpected peak demands – which must be avoided by any professional radio station.

CDNs can offer good Quality of Service by using their own private long-haul backbone network to feed directly into the ISPs' 'edge' hubs. CDNs can serve a seemingly infinite number of streams, as server and network capacity is shared across many clients, geographic boundaries and time zones. Therefore, a peak demand for one client is unlikely to disrupt another's service.

Distribution networks demonstrate, every day, that they can sustain – and the user's broadband connection can support – rates of 1.5Mbps, which suggests that higher bit rate audio streaming is possible.

3.4. Costs

Traditional broadcasting systems such as FM can be received by any number of radios, without any restriction or cost implications for the broadcasters.

Internet radio delivery is essentially a 'one-to-one' medium and so the charges from a CDN are linked directly to the popularity of the station. A CDN charges for the total volume of data distributed, which will depend on the number of listeners, the total listening hours and the encoded bit rate of the transmission to each listener. (CDN pricing is currently competitive, but even so, if all UK radio listening were through the internet, it would consume 54 petabytes of data every week and would be many times the cost of traditional transmitter networks.)

As the broadcasters' costs are directly linked to the number of listeners and the bit rate, this acts as a business driver to keep the bit rate at an acceptable minimum level, at least for popular stations.

Serving different formats and wrappers has virtually no cost penalty to the broadcaster. While the number of listeners to a higher bit rate service is a small percentage, the additional cost penalty is relatively insignificant. To keep costs low, it would be important to force the interested listener to choose positively a higher quality stream instead of the standard service. This cost penalty may deter smaller stations from offering a higher bit rate service.

3.5. Format and bit rate choice

People can now listen to internet radio on a number of devices, including personal computers, mobile phones (cell phones), games consoles and internet connected hi-fi systems and 'radios'. Broadcasters can choose to support these different devices with appropriate wrappers, audio formats and bit rates. For example, targeting a low bit rate high-efficiency audio codec such as HE-AAC v2 to mobile phones using a simple wrapper, and a 96 to 192kbps stream to a broadband connected device.

There is an understanding that consumers expect to see a bit rate figure of "128kb/s" to represent high quality stereo streaming so, in the UK, this has tended to become the typical bit rate, even though it is likely to deliver better audio quality than listeners can hear on low-cost PC speakers.

CDNs tend to offer more competitive pricing for streaming popular, stable and commercially supported wrappers and formats. The popularity of formats and wrappers has changed over time and, so far, internet broadcasting has not reflected the same stability as the FM transmission standard!

4. THE CASE FOR HIGHER QUALITY

4.1. Higher quality audio

In Summer 2010 the UK's public service broadcaster, the BBC, launched a trial of streaming their internationally renowned "Proms" concerts in 320kbps AAC coding, with audio fed directly from the studio, bypassing all loudness processing. The comments received and uptake were sufficiently strong for the BBC to launch a permanent high bit rate stream of its classical music station, Radio 3 [2].

Similarly, whilst music sales data are commercially sensitive, it is clear from the increase in the marketplace of music websites offering high resolution downloads (ranging from 44.1kS/s 24bit through to 192kS/s 24bit) that there are sufficient consumers who will pay a premium for higher quality sound. This and the continued sales of hi-fi products with internet radio capability, suggest that broadcasters offering high bit rate services will be taken up by some listeners. Further tests and trials will be needed over time as internet connections improve to understand the true appetite for these services. Internet radio's clear advantage over traditional broadcasting of being able to offer additional streams to special-interest listener groups will allow these trials and tests to take place.

4.2. Streaming data rates

As mentioned earlier in this paper, video streaming in the UK has shown that higher bit rates can be sustained over the internet to consumers, up to 1.5Mbps for standard definition and at 3.2Mbps for 'high definition' video streams.

These data rates are sufficient to allow lossless audio encoding at 44.1/48kS/s and even high resolution audio at 88.2/96kS/s with 24 bit depth.

Codecs are available. Development would be required to carry these streams in a cost effective way through a unicast streaming model, via a CDN. CDN costs may always be higher per gigabyte if the format is unusual and requires special support.

4.3. Paying for a premium service

Costs to the broadcasters will clearly increase if higher bit rate audio is streamed. The level of CDN network delivery pricing paid by large users such as major broadcasters is confidential, but taking typical estimates suggests that for a person listening to a high bit rate stream for an average 3 hours every day, the cost increase to the broadcaster over a typical 128kbps lossy codec stream would be in the region of:

- 50 to 100 USD per annum for 44.1kS/s 24bit streams
- 100 to 200 USD for 96kS/s 24bit streams.

Market testing would be needed to ascertain if special-interest groups would pay for this quality of streaming. If payments are made, and it is difficult to see at these rates how the service could be commercially sustained in a unicast model without the listener contributing to the cost, then a method of restricting access to the streams is essential. Encryption could be applied to the stream, although it can be argued that this is an almost pointless engineering burden. Streams could be targeted at special receivers ('hi-fi' products) employing unique codecs but, as with encryption, this approach would be relatively easy to break given the power of the internet user groups. At a simple level, streams could be served only to defined IP addresses, which would in turn require customers to have broadband connections with static IP addresses. None of these solutions is ideal.

A further problem exists. The broadcasters with the most high value content – typically public service broadcasters – may not be permitted by legislation to charge for a premium service.

Multicast streaming would remove these issues as the cost increase should be near to zero for the equivalent level of service.

5. LISTENERS' ISSUES

5.1. 'Fair Use' policies

ISPs have seen significantly increased data throughputs due to streaming, for example in the UK due to the success of the BBC's iPlayer radio and television streaming service, and music downloads.

The ISPs impose a "Fair Use" policy total bandwidth transfer cap on their customers. These current limits, such as 60GB per month, would restrict the possibilities for delivering higher bit rate streaming to the home. The example in Section 4.3 of listening to a 96kS/s 24 bit stream for 3 hours a day would exceed the monthly cap by over 30GB. This assumes that no other person in the household watches any streamed video or also listens to higher bit rate streams.

5.2. Streaming faults

Live audio must be delivered on a timely basis, which is not a requirement for typical 'web' traffic. Even though the audio is buffered in the receiver, dropouts can occur due to peak traffic elsewhere. Listeners to high bit rate services will suffer more breaks than from traditional transmission.

The listening enjoyment is more likely to be disturbed by glitches in audio streams than by the equivalent pauses in video. If listeners are required to pay for a premium service, they may be even less tolerant of such faults.

6. BROADCASTERS' ISSUES

6.1. Introduction

Section 4 has discussed the possibility of streaming the equivalent of 'high resolution' audio into the home. High resolution commercial recordings are often captured with high quality microphones, converters and minimal equipment in the audio chain.

To maintain flexibility, broadcasting cannot follow the same approach. Many more processes and equipment are present in the chain between a microphone and the listener at home (See Figure 2). Some of these will have an effect on the perceived quality.

6.2. Content & sampling limitations

The main programme content broadcast by radio stations can be summarised as:

- News and documentaries
- Commercially recorded and published music (such as pop, rock, classical)
- Live music events and local recordings
- Drama, book readings and entertainment shows.

The sample rate/bit-depth of most music output will be limited by the original recordings. The only real scope for increasing the quality above that of 'CD' is for broadcasters to improve their own production of live events, music and drama. These are also the most likely to offer the highest content value to listeners.

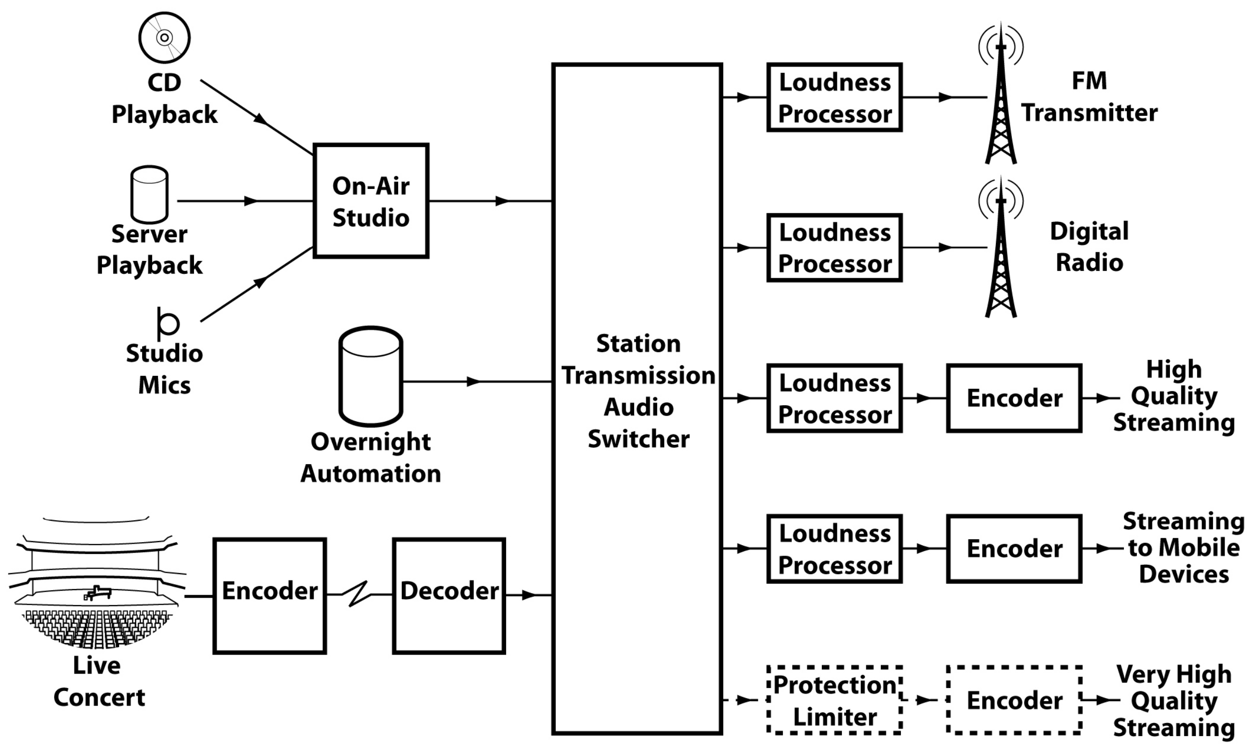

Figure 2 Broadcaster's infrastructure example

Even for lossy audio streaming, the codec choice and bit rate are not the only factors controlling the perceived quality. Good industry practice for the digital audio acquisition before encoding is important.

6.3. Copyright

Music copyright owners have been protective of their content and may react against audio being carried in 'full CD quality' to the listener. Lossless delivery may also reduce the (perceived) financial gain to the broadcasters from sales of programme recordings.

6.4. Studio Centre infrastructure issues

6.4.1. Dynamic processing and operational practice

All broadcasters employ protection limiters to maintain signal levels within acceptable limits for final transmission. It has become common practice to use loudness processors with multiband processing to automatically manage the dynamic range and to create a 'station sound'.

Removing these devices from the programme chain, with their inherent dynamic range distortion, can make a significant perceived improvement to audio quality on a premium service.

However, the long-term trend away from carefully-crafted audio quality to station loudness has allowed sound level balancing in the studio to become less important. The loudness processors are left to manage the production balance, with some sound engineers even compensating in the studio for the processors' effect. So the special-interest listener will have to tolerate a wider dynamic range – not a problem for a single piece of music, but one which can be irritating if levels jump between tracks and contributors.

6.4.2. Sample rate

There are issues in the studio centre which will limit the ideal implementation of higher quality streaming. A large studio centre will lock all digital audio sources and systems to a common reference clock. An audio-only studio centre could choose to use a reference which matches most of the commercially available recordings: 44.1kS/s. But, a larger studio centre will typically include video production too and so this will dictate that the sample rate is set to 48kS/s. This implies that sample rate conversion will be in-circuit on commercially recorded music, at least.

With little mainstream commercial music being available in high resolution sample rates, it is unlikely that high resolution audio sample rates (96kS/s) would be chosen for the studio centre. This would double the storage capacity used for in-house recording, playout and automation systems and so would carry a further cost penalty.

6.4.3. Reference clocks and sample rate converters

The stability of digital audio clocks for audio playback is recognised by some sections of the audio industry. Inserting sample rate converters running as synchronisers has a similar effect to varying clocks at analogue playback. It is theoretically possible to construct a studio centre's infrastructure which is locked to the sample rate of commercial audio recordings (44.1kS/s) and has a clean path to the encoder. This arrangement could, with a lossless codec, deliver sample-accurate audio directly to the listeners' playback devices. Should a listener choose to invest in an off-air time reference such as a GPS-locked clock, matching the broadcaster's own reference source, then no synchronisation processes would be needed anywhere in the chain from the original recording to the listener's loudspeakers, as the 'internet buffer' of a few seconds will swamp the minor fluctuations in clock timing.

The same arrangement could be used to carry live concerts, again sample-accurate, from the venue's mixing desk to the home.

6.5. Live events and concerts

To carry high resolution audio from a concert venue, the broadcasters will need to invest in linear audio links back to the studio centre, at a quality level which at least matches the highest quality path to the internet listener. Alternatively, special events could be streamed directly from the venue, although this would require very high performance data links from venues, which may not always be available.

7. MANUFACTURERS' ISSUES

7.1. Stability

Listeners could use a computer with a good quality sound card to hear higher bit rate streams through a link to a hi-fi system. As more audio products are appearing with internet 'receivers' built in, listeners will gain most benefit from a wide dynamic range audio source through an acoustically-silent receiver unit, rather than through a PC with a cooling fan and rotating hard-disk. This is a positive opportunity for the manufacturing industry.

The habit of the 'internet' to continually evolve does not match the manufacturers' or consumers' views of the stability of a 'radio' design, for example, compared with FM which has remained relatively unchanged for over fifty years. Products can be upgraded, but often not as flexibly as a computer's software can be changed to accept the latest codec.

7.2. Buffering

To deliver a premium service to the listener in a unicast environment, the receiver may need to have large audio buffers to manage, with an appropriate streaming protocol, the resending of missing packets which become more likely with higher bit rate streams.

7.3. Standards

Streaming linear audio may allow the manufacturers more freedom to choose a licence-free lossless codec and an appropriately simple wrapper.

A common, open approach will be essential if manufacturers are to persuade broadcasters to consider lossless streaming. If several different formats, wrappers and lossless codecs are used, then the broadcasters will need to support one or more and the manufacturers may be required to support all, making the process difficult to manage.

The hi-fi receiver industry may need to lead a trial with broadcasters to ensure that there is agreement on standards and a common purpose.

8. CONCLUSION

Data rates through the internet to the home can support lossless audio streaming, but the ISPs' typical fair use policies will limit the number of hours of listening each month. This and the dramatically increased costs imposed on the broadcasters by the unicast streaming model, coupled with the inherent audio quality limitations within a studio centre, suggest that, for the moment, lossless streaming may only be practical for special events and live concerts.

The business arguments are strongly against lossless or high sample rate lossless audio streaming, particularly with a unicast model. However, the internet is a rapidly changing environment and, as the throughput capacity of broadband connections to the home increases and if ISPs enable multicast streaming, then lossless services could become an everyday reality.

Internet streaming offers for the first time in radio's history the ability to offer special-interest groups their own quality version of mainstream radio stations. In the short-term at least, this may just be restricted to higher than average bit rate streams from lossy codecs, but with dynamic range processing removed. This has the potential to deliver live concerts, special recordings and drama productions directly into the home in a quality previously only heard in the broadcast studio.

9. REFERENCES

[1] Rajar - All UK Radio Listening statistics by platform – 2010 Quarter 4

[2] BBC Press Release "BBC launches HD Sound for radio" 18 October 2010