

IMPROVING
STREAMING VIDEO
PERCEPTION
WITH
BETTER AUDIO.

BY BENJAMIN SCHWARZ

SPONSORED BY UNIFIED STREAMING AND DTS

Takeaways

Booming streaming video markets are awash with competition and innovation.

When exclusive content isn't available, other unique features are needed to compete.

Better multi-channel "theatrical" sound will differentiate any video experience.

Streaming technology has stabilized with the rise of the vendor-independent DASH standard.

New audio technologies can bring a home-cinema user experience to the pocket of anyone with a pair of stereo headphones.

Commercial deployments are happening across the globe targeting both the dedicated living room and mobile devices.



CONTENT

1.A Brief history of surround sound streaming.....	3
2.The home-theatre market: from exclusive to every pocket..	5
3.Sound magic.....	6
4.Under the hood	6
5.Psychology of sound and video	7
6.Strategy: differentiation with the same content?	10
7.Real deployments from the field, issues and trends	11
8.Trends	13
9.Looking forward.....	15

A brief history of surround sound streaming

In the early days of Internet streaming, all services were distributed on a unicast basis, where the streaming server had to manage the connections with all active clients. This confined scalability to thousands rather than millions of potential viewers and many early Internet streaming innovations like the first web TVs failed.

Meanwhile IPTV evolved as a closed and fully managed network, bringing multicast streaming, usually based on the RTP protocol, fixing the scalability issue and able to deliver multiple channels reliably at high quality. But this approach that started almost 15 years ago at the start of the new Millennium requires a dedicated multicast infrastructure that is expensive to install and maintain. It was confined not just to the footprint of the service provider but often also just to households relatively close to the operator's nearest Central Office.

This led to growing demand for a scalable streaming protocol that would overcome IPTV's lack of range by using standard Internet infrastructure components like HTTP caches and CDNs. Rate adaptive HTTP-based streaming, pioneered by Move networks in 2007 and currently embodied by MPEG-DASH, fulfils these needs.

Mod_h264_streaming then began as a simple open-source Apache web server module for creating an index on the fly so that players could start playback immediately. This was called pseudo-streaming. However the MP4 limitations remained, such as the absence of adaptive bit rates and a single soundtrack. So when in 2008 Microsoft introduced Smooth Streaming (SS) in time for the 2008 Olympics in Beijing, SS was immediately added to the mod_h264_streaming module, renamed Mod-Smooth-Streaming in the process. This was described by Unified Streaming founder Dirk Griffioen as “One of the very first Smooth Streaming implementations”. In the following year, *Unified Streaming* created a commercial platform called USP (Unified Streaming Platform), adding support for Apple’s HLS and then Adobe’s HDS, before implementing the new MPEG-DASH standard.



This opened the way for specialist high definition audio firms to come into the streaming arena, one being DTS with its *DTS-HD* surround sound technology designed to bring the theatrical experience to connected devices. Elevating the consumer experience is nothing new to DTS, which started out in the theatrical space. In the preparation of 1993 blockbuster Jurassic Park, Steven Spielberg together with sound designer Gary Rydstrom and George Lucas, who supervised the sound crew, wanted more than just multi-channel sound. They wanted realistic sound. Spielberg personally invested in the creation of DTS to render the ground-breaking audio including dinosaur roars and moving bullets.

The home-theatre market: from exclusive to every pocket

over quality.

Premium multi-channel audio can now be enjoyed with either speaker or headphone based surround sound that delivers a stunning audio experience via all connected devices, irrespective of the corresponding video quality on the screen.

DTS has developed a technology to deliver the full immersive, cinematic audio experience to your headphones that you've come to expect from a surround sound movie theatre: DTS Headphone:X®.

While home theatre systems were once the preserve of audiophiles, they have become more accessible in the home with all-in-one systems and soundbars. CEA data suggests that about 18% of consumers will purchase a Home Theatre system within the first year of a TV purchase.

More than ever, consumers are also enjoying content on the go via mobile and tablet devices. Until recently, the mobile experience has been compromised by convenience

Sound magic under the hood

Binaural audio was one of the first innovations in sound recording and reproduction, dating back to 1881, although the first commercially produced pop record did not follow until almost a century later with Lou Reed's Street Hassie. Binaural audio was designed to work properly only with headphones in an attempt to reproduce as accurately as possible the way sound is experienced in an actual venue, allowing for the dynamics created by the human

head and aural system. Only through headphones do binaural techniques become relevant for creating an authentic theatrical experience.

So how can multi-channel audio be reproduced from a 2-channel mobile device? Many of the components remain the same:

- The actual speakers
- Basic audio signal processing and amplification
- Transport technology for delivering the sound stream to the device
- Encoding format for surround sound (typically a 5.1 format)

Modern video streaming with DTS audio sends the same data – typically 5.1 audio - to all devices. It is then up to the device rather than the delivery infrastructure to decode the audio according to its capabilities and the number of available audio channels.

A key advantage of this approach is that just one audio track will deliver the best possible experience for different setups. If only stereo (or even mono) output from a sound bar is available then that will be used. If there is a 5.1 sound system then full advantage is taken of that.

Psychology of sound and video

When the end device has a DTS decoder and processing technology like DTS Headphone:X, then a virtualizer is used to render multi-channel audio from just two headphone speakers.

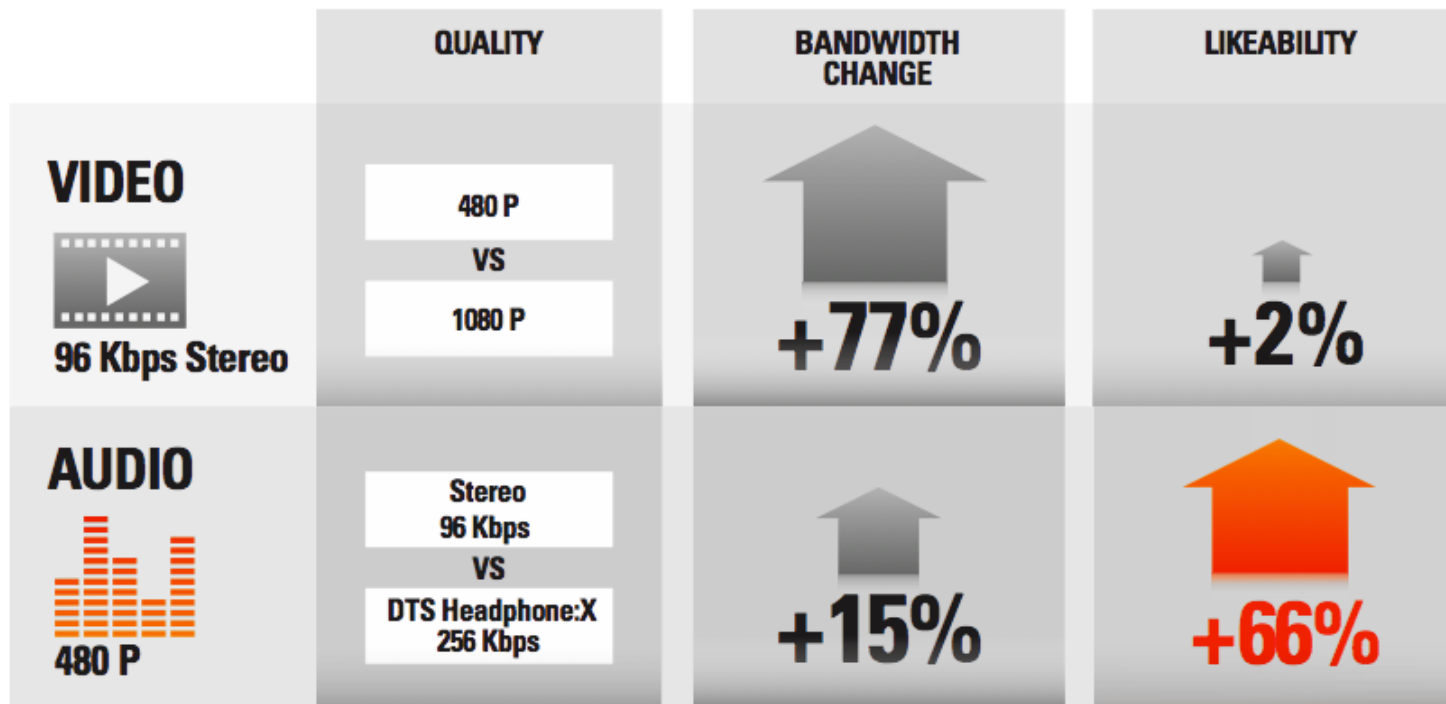
In the case of video there are several perceptible attributes that interact over their contribution to the overall experience. For example the proponents of High Dynamic Range (HDR) have conducted what they consider conclusive studies showing that by enabling much better contrast, viewers perceive more detail in the picture, even though the resolution is unchanged.

Similarly, although it may not seem intuitive initially, better sound also improves the perceived quality of the video. This is particularly the case for video viewed on mobile devices with small screens. This implies that sound quality is absolutely integral to the next generation TV experience being ushered in under the umbrella of Ultra HD.

In a 2014 study, DTS used the following methodology to test brain activity at varying levels of sound and video quality.

- 107 subjects watched video on mobile devices in a lab setting
- 28 sensors placed on their heads measured real-time brain activity
- subjects viewed 4 pieces of content
- audio and video quality was randomly ordered to remove bias
- Video resolution varied between 240p, 480p and 1080p
- Half the subjects (54) watched video with ordinary stereo 96kbps audio; half with DTS' Headphone:X at 256kbps.

The results are very telling. Increasing resolution from SD to HD improved likeability by only 2%, whereas moving from stereo to true surround-sound raised likeability by 66%. Surround sound creates a more immersive effect.



This study is conducted by Neuro Insight in conjunction with DTS, January 2014.

Strategy:

how can you differentiate with the same content?

In a competitive market where multiple service providers deliver the same content, enhancing the user experience is the only way to stand out.

- Content is still king, but in many markets the same titles are available across different services
- Content packaging used to be a differentiator, however many households have moved away from packaged media
- Pricing can be a differentiator, but many services have similar pricing models

Adding theatrical surround sound to video is a straightforward way to differentiate. Immersive multi channel audio with premium video is an opportunity to be seized now for streaming operators that want to increase engagement on mobile platforms.

Real deployments from the field, issues and trends

Combining state-of-the-art streaming with stereo-based surround sound is now very easy. We have spoken to a few operators that are using both DTS and Unified Streaming solutions to deliver a premium streaming experience.

- US based FandangoNow (formerly M-GO) has launched with DTS-HD premium surround sound on both Samsung and LG UltraHD and Smart TV platforms
- Nowtillus prepares content for leading European streaming services, and has ingested over 4000 movie titles with DTS-HD surround sound.
- GoPrimeTime is a successful Internet-based video streaming service from Thailand. With over 300,000 subs they have had excellent customer feedback regarding the 5.1 sound feature. The only two features mentioned on their homepage are UHD and DTS 5.1. Kasidit Kolasastraseni, the company owner, told me “what impresses

people the most is to get full surround-sound with their regular headphones”. He went on to tell me about the enthusiastic response from his CPE partners, noting that “all smart TVs sold by LG and all Samsung Galaxy phones in Thailand come with our App”. The GoPrimeTime app is based on Berlin-based CastLabs’ app framework.

- Alibaba is Daniel Zhang ‘s renowned online Chinese platform that streams premium VoD to its 10 million strong customer base. The Chinese service has a dedicated “DTS channel” where users can find content exclusively with DTS multi-channel audio.
- CinemaNow in the US, Canada & UK is another example of a service offering DTS multi-channel surround sound. The service is a leader on the Xbox One platform, and also available on Samsung Smart TV’s and Blu-ray players.
- Hardware devices are starting to embed specific technology to render multi-channel audio. Turtle Beach gaming headphones for example now incorporate Headphone:X.

An interesting point to come out of all these platforms was that all services chose to offer premium audio to all subscribers. In the field, as there is no marginal costs involved, providers are so far choosing to use multi-channel audio as a differentiator rather than an ARPU generator. For now, all these service are contained within their original geographies.

In Europe and North America most major online services are either launching premium audio streaming services directly with MPEG-DASH or moving to MPEG-DASH from existing Smooth Streaming deployments. This became a matter of urgency for many when Google removed Microsoft Silverlight support from its Chrome products. Delivering the best user experience to mobile devices is a key concern, as is the availability of apps for connected platforms such as TVs, game consoles and dedicated streaming devices like the Roku, Amazon fire TV, Apple TV or Google Chrome.

Trends

Content security is another major concern in the market and the DRM approach we see in Europe to handle the Microsoft / Google support issue in particular is '**PIFF2CENC**', which works well whether content is already DRM-encrypted on disk or requires DRM only for transport and play-out.

Many existing setups use either mp4 or fragmented mp4 (ismv) and encrypt on-the-fly with PlayReady, but this approach will no longer be supported in Chrome. Using DASH as the streaming format is more future-proof, with Widevine Modular DRM for content protection.

Recent products such as Unified Streaming's Unified Origin can ingest (fragmented) mp4 and play-out to PlayReady protected Smooth Streaming or DASH as well as Widevine-protected DASH. As packaging is done on-the-fly there is no need to store other formats beyond the (fragmented) mp4.

Of course the Origin server can also use CENC-encrypted content.

Apple's FairPlay DRM still has traction in the market, but there is a trend for service providers to become as independent as possible from DRM vendors. It is possible to use a "[TransDRM](#)" function with some Origin servers such as the Unified Streaming one, in order to output to other content protection schemes.

PlayReady and Widevine Modular have been the key DRMs for all the pioneering deployments. Meanwhile Common Encryption (CENC) is being discussed more and more across all customers and market segments as a core part of the transition to a multi-DRM and multiplatform OTT world.

In Asia there is a trend towards the use of HLS as well as customized variations of HLS. Support for mobile devices remains a key priority although several set-top-box only deployments are bringing premium audio streaming to the home. In some markets multiple set-top-boxes are used in a single household, with each content service using its own device for streaming or digital download delivery. For encryption AES and Intertrust are common.

Looking forward

As household bandwidth increases and audio/video codec efficiency improves, the streaming market is poised to bring new and exciting features to the end consumer. In the audio realm, adding height with up-firing speakers and soundbars is well on its way to delivering new levels of immersion to connected consumers. In the video space, 4K UHD and Virtual or Augmented Reality present new experiences for the streaming market.

Streaming ecosystems that use DTS surround sound and Headphone:X are already integrating new UHD features beyond multi-channel sound and 4K resolution. For example, FandangoNow streaming and download include High Dynamic Range (HDR) for a better viewing experience.

In the coming years, technologies such as Object-based sound will enhance both interactivity and immersion. An example is the DTS demo of a football match, which has fully immersive 3D sound by default, but lets the viewer or listener independently control volume levels for six different sources (Announcer, referee, Coach, Goalie, crowd and Field). Sources can also be switched so that one can choose between say multiple announcers. Object-based sound also allows a sound-author to position the audio of say a fly anywhere in the viewer's perceptual space with a Multi Dimensional Audio (MDA) plugin for their editing software.

Live content has an intrinsically immersive aspect as viewers are drawn literally into the moment. So such an immersive technology as surround-sound delivered live through state-of-the-art streaming complements this and the “wow” effect can be even greater with say a live football match. Imagine a penalty shot where the thud of the kick seems to emerge out of a deep silence right from the centre of your perceptual field, then when the ball hits the bar on one side the sound faithfully comes from that spot and finally the crowd roars from hundreds of meters all around you as the ball goes in ... and all this is experienced on your mobile phone.

