

Barco

# Auro 11.1 versus object-based sound in 3D

**All aspects compared**

## **Authors**

Brian Claypool  
Senior Director Strategic Business Development, Barco NV

Wilfried Van Baelen  
Founder & CEO, Auro Technologies NV

Bert Van Daele  
R&D Manager, Auro Technologies NV

## PURPOSE OF THIS DOCUMENT

George Lucas is famous for saying that *sound* comprises at least 50% of the movie experience – sound helps create more of an emotional response in the audience than image alone. As such, the film industry has always been on the forefront of audio technology as it pushed the envelope of sound reproduction: the introduction of Stereo and later Surround Sound, the introduction of digital audio with Dolby Digital, DTS...

3D sound – an immersive sound field, or ‘sound all around’ – is the next logical step required to deliver the ultimate listening experience. Several solutions have been proposed to the market, but not all of them can fulfill the cross-media requirements of the entire multi-media space: movies, broadcast, music, events...

To help **movie exhibitors** and **post-production professionals** take this next step, this white paper compares the technology and the practical implications of the Auro 11.1 *channel-based* system to the several *object-based* or ‘*hybrid*’ (a combination of channel- and object-based) systems being proposed in the market.

## INTRODUCTION

### The importance of sound in 3 dimensions

Today’s movie theaters position speakers as arrays in a two-dimensional horizontal plane (along the x=width and y=depth axes). This current layout is known as 5.1 or 7.1 Surround, depending on the number of channels deployed.<sup>1</sup>

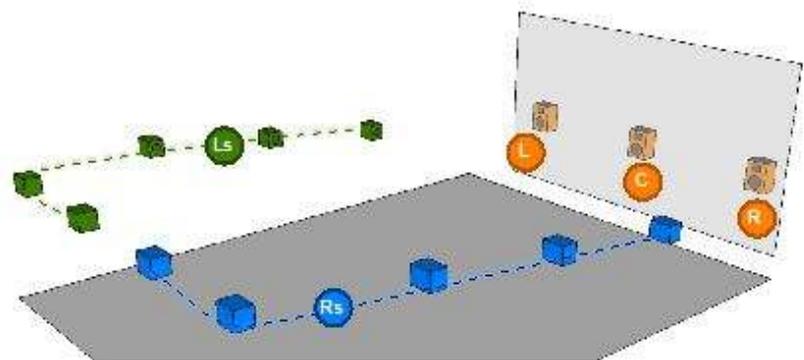


Figure 1 - Standard 5.1 Surround Sound

Auro 11.1 adds the **next, third dimension (Height)**, with speakers arrayed above the spectator. The third dimension of sound around the listener supplies the natural, ambient sounds that occur in acoustic reflections all around us.

<sup>1</sup> (Holman 1996)

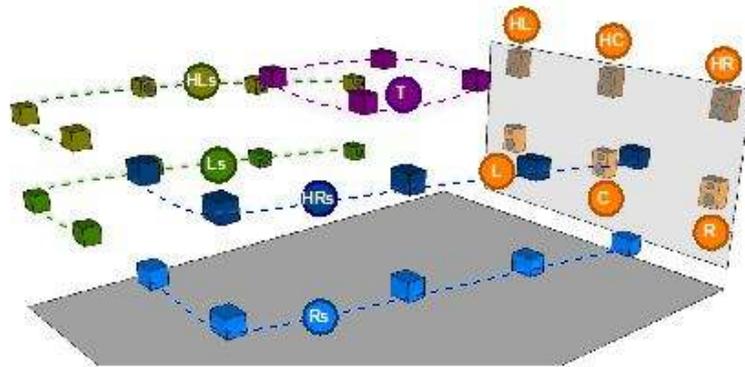


Figure 2 - Auro 11.1 with three distinct layers (surround, height, top)

Several studies have been conducted to investigate the reproduction of life-like, three-dimensional sound and the importance of reflections for spatial hearing.<sup>2,3,4,5,6</sup>

These studies have shown the importance of combining lower as well as height speakers, with an elevation angle of ca. 30° compared to ear level, to create the true immersive sound we are used to hear in everyday life, regardless of room type or size; from home theaters to dub stages and cinema venues. Just as in nature or on a busy city street, this added dimension of sound produces an authentic, fully immersive auditory experience for the moviegoer.<sup>7</sup>

In contrast, the leading Hybrid/Object-based system has no height channels at 30° and only a Stereo overhead. As a consequence, that Stereo overhead has to cover the gap between the lower Surrounds and the ceiling. Although they are suitable for various fly-over effects, ceiling overhead channels alone have difficulty reproducing that hemisphere of sound because the gap between the overhead channels and the Surround channels becomes too big and therefore feels disconnected.

This is why a speaker arrangement comprised of 3 layers is necessary to reproduce a natural three-dimensional sound field.

<sup>2</sup> (Hamasaki, The minimum number of loudspeakers and its arrangement for reproducing the spatial impression of diffuse sound field 2002)

<sup>3</sup> (Hamasaki, Advanced Multichannel Audio Systems with Superior Impression of Presence and Reality 2004)

<sup>4</sup> (Blauert 1997)

<sup>5</sup> (Kim, Lee and Pulkki 2010)

<sup>6</sup> (Theile and Wittek 2011)

<sup>7</sup> (Gerzon 1973)

## Discrete versus ambient sounds

Three-dimensional, immersive sound is not only about discrete, individual sounds that originate from around and above the listener (such as birds, airplanes, etc.), but more important is the ability of the sound system to deliver the sensation of being immersed by sound that originates from the combination of both Lower and Height channels – accurately reproducing the ambient sounds such as concert hall reflections, the sound of leaves rustling in the forest or the ambience of a busy market square. As such, Surround Sound combined with overhead speakers alone cannot reproduce the full 3D auditory experience.

Auro 11.1's Channel-based solution is based on three layers of loudspeakers (Lower Surround, Height and Overhead) to achieve a full three-dimensional sound spread capable of reproducing natural acoustical space with a minimum number of channels and speakers.

Additionally, human hearing is very sensitive to distinct sounds originating from the rear – a survival instinct, sometimes referred to as the 'Tiger Gene', warning us of potential danger creeping up from behind. This subconscious mechanism increases the sensitivity for discrete sounds from the rear by ca. 1.5dB, immediately drawing all attention to these alarm triggers. Because of this, the use of such effects in the Surround channels is limited to a minimum by most film producers<sup>8</sup>. Otherwise, these sounds tend to draw too much attention to themselves, distracting the audience from the action on-screen. In other words, sounds defined as objects, travelling to the back, risk compromising the immersive illusion that film producers strive to achieve. This fact restricts the added value of such systems to a limited number of use cases.<sup>9</sup>

## CHANNEL-BASED VERSUS OBJECT- BASED SOUND

In a film, the sound the audience hears is a combination of many individual sound elements, brought together during the mixing process at the post-production stage. These sounds are recorded on the set (e.g. dialogue) or in the studio (e.g. music) or are taken from vast sound libraries. It is not unusual for large film productions to end up with several hundreds of such audio elements.

---

<sup>8</sup> (Allen 1991)

<sup>9</sup> (Holman, 5.1 Surround Sound, Up and Running 2000)

### Principles of Channel-based and Object-based audio

Up until recently, all these sound elements were distributed over several audio ‘tracks’ and mixed in a reference studio using huge mixing consoles and digital audio workstations compliant with a standardized speaker setup. The end result of this process is a set of ‘channels’, each channel representing the content of one speaker or speaker-array (for the Surround-channels). This is called the ‘Channel-based’ approach and is in use by virtually all production facilities throughout the world.

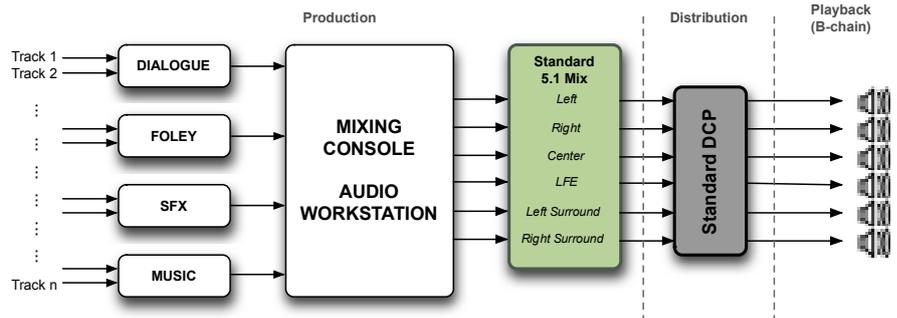


Figure 3 - Conventional, channel based approach

In an ‘Object-based’ system, the different sound elements are bundled with metadata to form audio objects. This metadata then describes how the sound element should be reproduced in the theater, by defining its position in the 3D-field with a vector, level, etc...<sup>10</sup> The playback system then needs to render that information in real-time for each object related to the playback system’s speaker layout. As such, Object-based systems always need special, dedicated hardware (rendering-stations), depending on the system of choice.

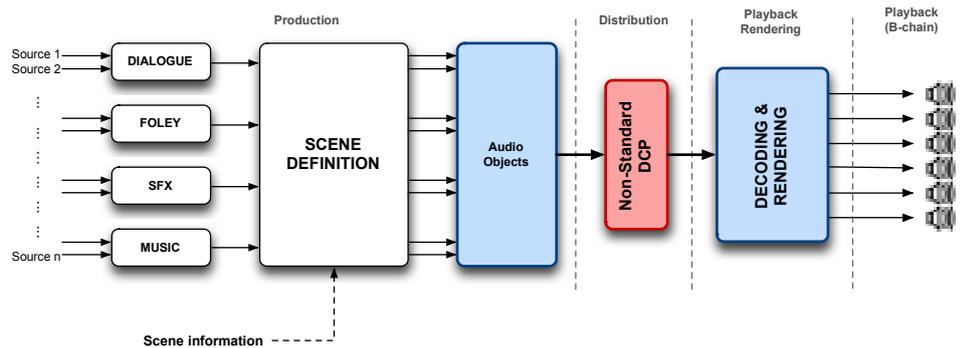


Figure 4 - Object-based approach

<sup>10</sup> (Jang, et al. 2003)

## 'Dry' versus 'wet' sound

Based on the principles described above, it can be understood that the Object-based approach is especially powerful for the reproduction of discrete sound sources that require precise localization as well as movement. Examples are flying bullets, cars passing by, a knock on a door... However, this only works best if these sources are recorded 'dry', without reverberation or reflections.

In reality, however, such distinct, dry sources are the exception rather than the rule; they are almost always accompanied by 'wet' or ambient sounds coming from the object's surroundings. This kind of diffuse sounds is best handled by Channel-based systems. That is also the reason why some systems apply a hybrid approach where ambient sounds as well as music are Channel-based, while some discrete sounds are then handled as objects.

In these cases, the audio objects are often recorded dry, while the necessary reflections and reverberation to make the sound life-like are added as pre-mixed channels. This technique creates the potential risk that the source and its reflections seem detached from one another, creating an unnatural (less immersive) experience.

## Movement of sounds

The main advantage of Object-based systems is their capability of maintaining a pre-defined localization or movement of a sound source, regardless of the reproduction system in the theater.

Consider the following trajectory as described by the audio object's vector.

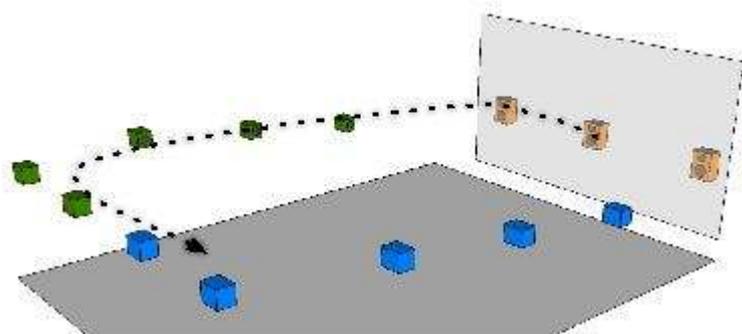


Figure 5 – Described trajectory of a sound object

In this example the sound's trajectory starts at the Center speaker and travels over the Left and Left Surround speakers around the auditorium towards the right back corner.

If in the target speaker system, all the speakers are wired and amplified individually, the trajectory can be rendered with much detail, traveling through each speaker individually.

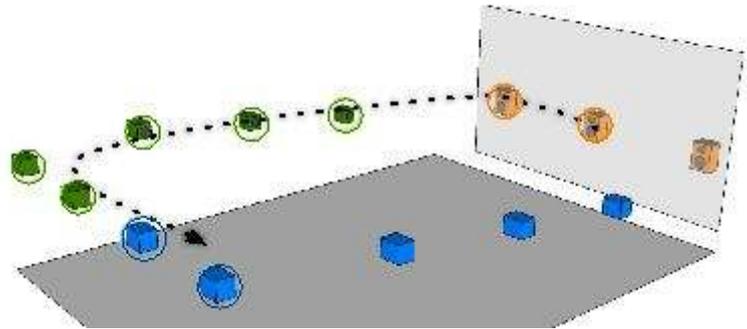


Figure 6 - Trajectory when rendered on fully equipped theater system

When rendered in an auditorium equipped with a standard 5.1 surround system, the trajectory will be changed since the sound will not see the same smooth transitions in the room and will also end up on the right side wall.

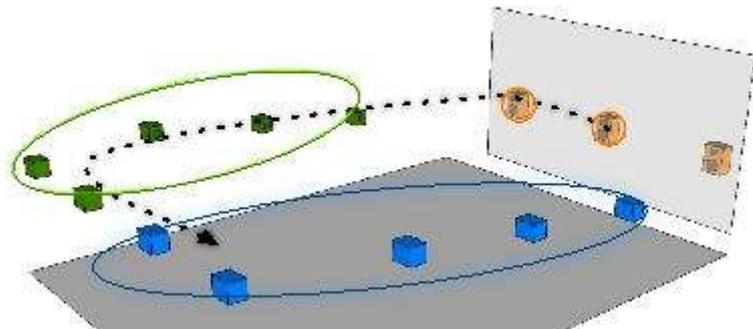


Figure 7 - Trajectory when rendered on standard 5.1 theater system

While in a Channel-based system the precise localization and movement will indeed be limited to the capabilities of the standardized speaker system, the mixing engineers can take this into account from the start and optimize the mix accordingly, circumventing these limitations.

One of the strengths of the Height layer in Auro 11.1 is that it reproduces important reflections from above, greatly improving the localization of sounds in the lower Surround channels, even without the use of Object-based technologies. More often than not, objects coming from the screen into the audience are fast moving objects. The difference between a Channel-based or Object-based system will be small for the audience since there is almost no time to experience the exact

location of such objects. Object-based systems especially perform better with very slow moving objects. But those are first of all not common and secondly risk drawing attention away from the screen, which results in a less immersive cinema experience.

### Sweet spot and power handling

A Channel-based system – like Auro 11.1 – *clusters* speakers into channels, so that, for example, a 42-speaker array will be clustered into 12 channels. This is especially done with the Surround channel and enlarges the system's 'sweet spot'; the sound is delivered as consistent as possible to moviegoers sitting throughout in the theater.

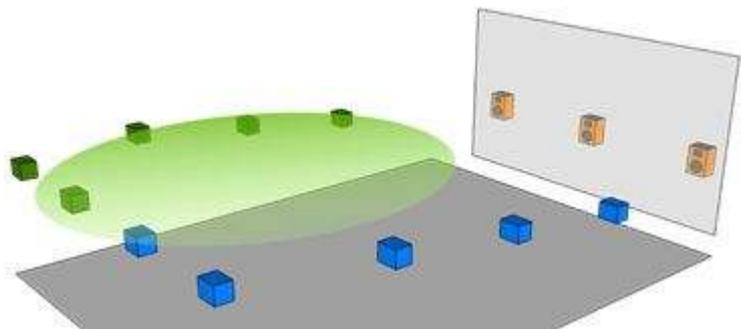


Figure 8 - Sound distribution of Surround arrays

On the other hand, while in an Object-based system the resolution can be improved by adding more individually amplified speakers, this also leads to a different experience for every seat in the house (i.e. it has a smaller sweet spot). On top of that, the arrangement and coordination of the speaker array to accurately reproduce the sounds – including volume scalability and phase issues arising from the divergence of sounds – can be extremely complex.

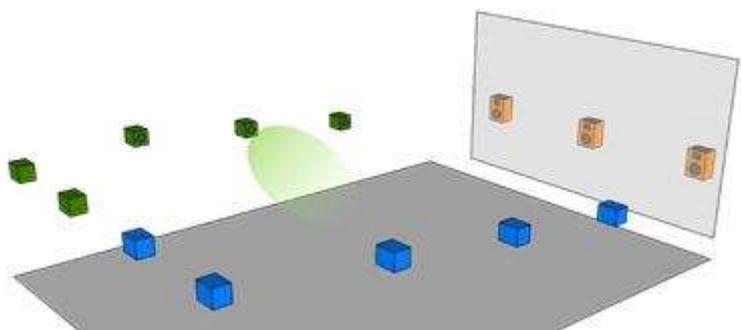


Figure 9 - Sound coming from a single surround speaker

For instance, if a single full-range Surround speaker is not able to reproduce the power level needed to cope with audience coverage, the audio object will be spread over multiple speakers to increase the output power, (effectively going back to the effect and sound when reproducing the sound over a regular surround array as used for the Surround channels in a channel-based system). For this to work properly, the speakers must be positioned close enough to one another in order to be coupled and deliver sufficient power without compromising the spot source effect with phase and coloration artifacts.

The result is an array composed of more speakers than a standard Surround or even an Auro 11.1 configuration. This introduction of additional speakers in the Surround layer causes further difficulties: more phase problems can occur during playback of the channel-based sounds. This means that, when standard 5.1 or 7.1 content is played back through such an object-based system, the sound can be inferior to that from a standard Surround system because the use of more speakers has added more phase issues.<sup>11</sup>

## COMPATIBILITY CONSIDERATIONS

### Audio Format

With the introduction of Digital Cinema, it became possible to distribute the sound with the audio quality originally created in the post-production facility.<sup>12</sup> The SMPTE<sup>13,14</sup> and DCI standards defined that up to 16 channels of uncompressed, high-quality audio in the PCM-format could be included in a Digital Cinema Package (DCP). These 16 tracks may only contain one single reproduction format and need to include the special audio tracks for the Hearing as well as the Visually Impaired.

With the creation of the Auro 11.1 codec, all these limitations were taken into account; the result is a system that remains within the limitations of the existing standards by staying in the PCM domain for the delivery of both the standard 5.1 surround and Auro 11.1 formats, without any concession in the audio quality of both.

<sup>11</sup> (Holman, New Factors in Sound for Cinema and Television 1991)

<sup>12</sup> (Digital Cinema Initiatives, LLC 2008)

<sup>13</sup> (Society of Motion Pictures and Television Engineers n.d.)

<sup>14</sup> (Society of Motion Pictures and Television Engineers n.d.)

## Speaker configurations and design

Some object-based or hybrid systems may require the existing Surround speakers to be mounted higher than the current standard.<sup>15</sup> This means that the reproduction of standard 5.1 or 7.1 content over such system will create a substantially different sound field than originally intended.

Auro 11.1 can easily be installed in existing theaters, in most cases using components from the existing audio system for the lower layer. Depending on the speaker layout, simply adding height speakers will bring realistic sound in 3D to the theater (in the majority of the cases the existing surrounds do not need to be moved at all). Moreover, Auro 11.1 sound processing is similar to, and compatible with, existing B-chain set-ups.

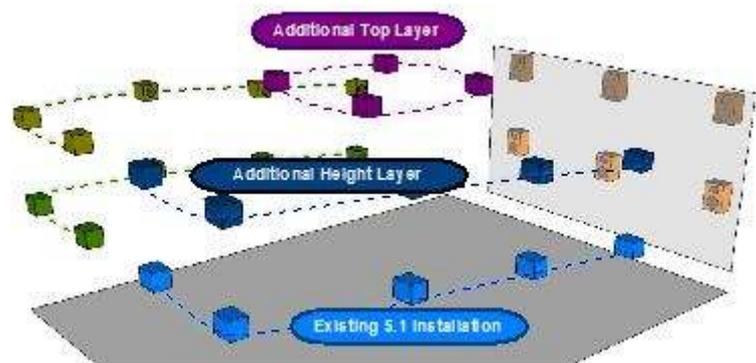


Figure 10 - Auro 11.1 builds on existing 5.1 installation

## DCI-compliance - Watermarking

An important aspect of the Digital Cinema Initiative (DCI) is the use of forensic watermark technology to protect content from being copied from media device hard discs and deter piracy in theaters. Therefore, a unique real-time watermarking system on each output channel is required for any theatrical sound technology. So when it comes out of the media block (digital or analog), this watermarking is continuously present in the signal.

As the Auro 11.1 Decoder runs on the same electronics where the content is decrypted and watermarking is produced, all decoded tracks of the Auro 11.1 PCM streams are provided with the required watermarking following the DCI specs. It is not yet clear whether object-based systems will provide watermarking on all output channels (an expensive

<sup>15</sup> (Uggelberg 1991)

proposition) in order to be DCI compliant. So far, Auro 11.1 is the only known DCI-compliant solution for sound in 3D.

### **Media Servers**

A major advantage of the Auro 11.1 system is that the decoder is a simple software upgrade that can be implemented in existing media servers. Most new media servers will even offer the decoder by default.

However, this is not the case with Object-based systems, which we can assume will require a new object-based media server as current media blocks can only play back a total of 16 channels.

## **PRODUCTION WORKFLOW**

Audio for film has a single workflow that should remain unchanged by adding either a third dimension or Object-based functionalities. The key issue thus becomes the ease with which the extra dimension can be implemented with as little impact on the workflow as possible. This is more readily achieved by continuing in a Channel-based environment than by introducing an Object-based system (or creating a Hybrid system).

### **Number of sources/sound elements**

Object-based systems require a powerful rendering station that calculates in real-time all the vector information and metadata related to each object.

As stated above, large movie productions, such as action films, use several hundreds of simultaneous sound elements. To reduce the complexity of such productions, stems and pre-mixes are created in smaller teams before the final post-production stage.

One system limits the number of objects to 128, which also contain the channels used for the beds placed in the channel-based sounds, effectively defining each channel as one audio object.

However, if this number is too restrictive, engineers could be obliged to continuously make decisions whether a sound needs to be treated as Object-based or Channel-based, which leads to a confusing workflow arrangement.

### **Tools**

Since Auro 11.1 uses the standard toolsets and way of working (based on plug-ins in standard Digital Audio Workstations), the

number of sound elements will only be limited by the total power of the workstations distributed over all stages.

In an Object-based workflow, each stage in the production, from sound-design to post-production, will need a specific, powerful workstation (i.e. a rendering station) to create the Object-based metadata files representing each stem. These files will then need to be combined and processed again at the final stage to create the total mix.

It can be understood that it will be a serious investment for post-production facilities to install such systems at each and every stage of the movie production (e.g. in each sound design room). A solution could be to postpone the complete mix to the final stage, which introduces a serious change in the workflow and increases the complexity and cost to handle the project.

Auro 11.1 productions make use of the standard tools already installed in every studio, without introducing any change to the creation of pre- and sub-mixes as done in current workflows.

When compared to the Auro 11.1 system, the Object-based or Hybrid system's hardware requirements introduce yet another hurdle for producers and sound engineers to overcome.

### **Quality control**

The idea behind Object-based (or Hybrid) systems is to create a single mix, with metadata for each audio object that can be rendered towards any reproduction system. It is clear that the final listening experience in theaters will depend heavily on the reproduction system and that, as a consequence, creatives such as directors, producers, and mixing engineers will never be able to control the exact sonic design of the film in every theater. Film producers, however, want to have full control over each format released to the market - they do not want to rely on an 'automatic conversion' to render this data to mainstream formats like 5.1, 7.1 and future formats like Auro 9.1, Auro 11.1, etc... The method used to create the Auro 11.1 mix and derived versions during mixdown, simplifies this process while maintaining artistic control.

Many discrete audio "objects" will require different level- and panning-settings for each standard, Channel-based speaker-configuration (5.1, 7.1, Auro 11.1...) in order to create the perfect mix for each format. It is not clear how Object-based systems will offer this possibility and may be dramatically underestimating the importance of maintaining this level of control in the final mix.

In a Channel-based system, the mixes are always made for standardized configurations where the production team knows

exactly what to expect in a theater – it will therefore be reproduced 100% as created. The Auro Codec has a unique feature that allows the mixing engineer to dynamically control the downmix-levels to create the 5.1 version from the original Auro 11.1 mix, optimizing the experience in both versions. These dynamic volume-changes are then reversed during the decoding process, revealing all the original signals of the Auro 11.1 version at their original level.

## **SINGLE INVENTORY DISTRIBUTION**

As a production company or exhibitor, costs and risks need to be kept to a minimum. For this reason, the goal is to deliver a single DCP of the film containing all audio formats in one package.

The Auro 11.1 codec allows all audio standards (Stereo, 5.1 and Auro 11.1) to be integrated into a single, standard (uncompressed) 5.1 PCM stream that will play back on every 5.1 Surround system.

When a movie is mastered with the Auro 11.1 codec, exhibitors who do not have an Auro 11.1 system will still be able to reproduce the 5.1 track at exactly the same quality as the standard 5.1 Surround. In the case of exhibitors with the Auro 11.1 system, the decoder will simply extract the supplementary sound information from the PCM signal and play it back in its original Auro 11.1 format with no audible compromise.

### **A single DCP, a single KDM**

Auro 11.1 has only one audio file on a DCP (digital cinema package), which is identical to the 5.1 file and contains the PCM Surround master carrying all of the information for the Auro 11.1 playback. This means that only one DCP and just 1 KDM key must be produced. In addition to the advantages this represents for distribution costs, no handling mistakes can be made due to extra requirements as is the case with the object-based master that requires a KDM for the 5.1 Surround and a different KDM for the object-based mix, increasing handling and distribution costs. This also means every auditorium with an object-based or hybrid system must be noted in the TDL (Trusted Device List) – which is not necessary for Auro 11.1.

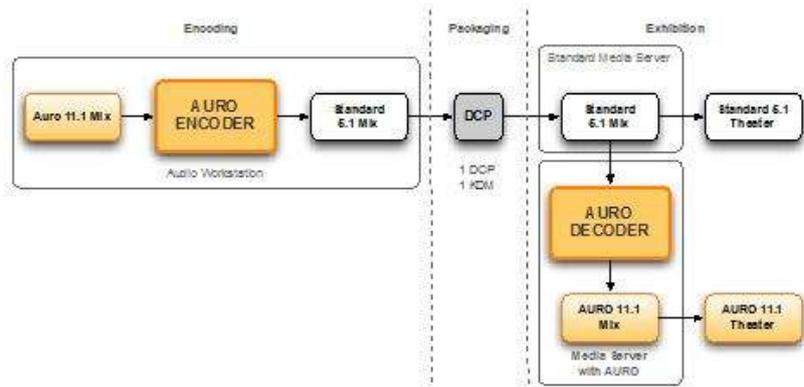


Figure 11 - Cinema Audio Workflow

Note: The Auro Encoder is a software plug-in for most existing Digital Audio Workstations (e.g. Pro Tools, Nuendo...) on Mac as well as PC. The Auro Decoder is available as a firmware upgrade for existing media servers and will be available in most new servers and media blocks.

### Multiple markets

The Auro 11.1 format is designed to deliver all kinds of content from all potential sources (film, games, broadcast, music, events, etc) to the exhibitor in the existing 5.1 Surround format – allowing the content to be transported everywhere the 5.1 mix is distributed.

The same technology and tools will also be used in the consumer-oriented markets (including automotive, mobile and gaming), effectively creating a unified distribution method for the whole multi-media market.

Hybrid and Object-based systems are typically focused on to the professional cinema market and require significant additions to the DCP and changes to the wrapping specifications to accommodate the Object-based files that were not anticipated in the current packaging specifications.

Even if Object-based content would be delivered to the consumer market, the advantage over Channel-based sound will be limited when the reproduction systems remain limited to standard speaker configurations.

### COST

As with the introduction of any other innovative technology, the balance between *cost* and *experience* is another aspect carrying significant importance. Auro 11.1 provides a simple system at a viable cost, without any recurring licensing fees.

### **Costs for content developers**

To limit the cost to content developers as much as possible, the Auro 11.1 solution maintains compatibility with today's 5.1 and 7.1 Surround systems with its speaker layout and its Channel-based technology. The current question is whether all of the metadata for each object - that must be programmed during the Object-based mix to produce the different formats - will not be prohibitively costly in terms of issues and effort.

### **Costs for post-production houses**

Auro 11.1 only requires a standard Channel-based expansion of the speaker array (height + ceiling) and software-based plug-in modules for the final encoding and panning logic. In contrast, to experience the full capability of Object-based (or Hybrid) systems, one amplifier-channel per speaker is required in each room (sound design room, pre-dub room, dubbing theater, etc.) plus at least one rendering station in all of these rooms as well as a specific processor to allow for the distribution of these channels. This can be at a substantial additional cost.

The Auro 11.1 workflow is designed in such a way that it maximizes existing workflows kept in the system. Having all stems running at the same time as an output master is just one aspect of that. It is not yet clear how this will work when Channel-based and Object-based systems are used together - but it is likely to be much more complex.

The Auro 11.1 concept is based on building and controlling the Auro 11.1 mix and the 5.1 Surround mix simultaneously: while mixing the Auro 11.1 format, the engineer checks the compatibility with the 5.1 Surround mix. However, the two sound experiences differ so much that engineers may be inspired to enhance the 5.1 Surround mix, which is easily controlled with the Auro codec plug-in. This means that engineers can prepare both the Auro 11.1 and the 5.1 Surround formats in the same time it takes to prepare the 5.1 Surround mix alone, much in the same manner engineers check compatibility with the stereo mix on stages today.

### **Cost in distribution and exhibition**

Since the content of DCPs in Object-based systems will be heavily dependent on the chosen technology, each format will require its own set of DCPs, increasing the cost for handling, KDMs, etc.

Auro 11.1 does not add to the distribution cost as it is delivered on the same DCP as currently used for all standard 5.1 content, not requiring any extra KDM. While all 3D Sound systems require additional speakers, most object-based and hybrid systems require more speakers and have the extra requirement that they should be individually amplified, potentially going up to 64 channels – this increases the cost of the B-chain for post-processing and amplification, whereas Auro 11.1 is limited to 12 channels.

### **Data storage and transmission**

Since Object-based and Hybrid systems add a significant amount of additional audio data (up to a factor of 10) to the DCP, the time to ingest a movie will be increased. transmission over satellite links will become longer and thus more expensive and handling multiple versions will become even more laborious for distribution.

Auro 11.1 does not add any audio data at all, the audio track in the DCP remains within the 6 tracks of standard 5.1 audio.

## COMPARISON CHART

### Quick reference comparison of channel-based and object-based systems

Channel-based	Object-based/Hybrid
<ul style="list-style-type: none"> <li>○ 5.1, 7.1, Auro 11.1 are all Channel-based formats</li> <li>○ All objects, ambience, special FX, etc. are 'mixed' in to a predetermined number of channels by the mixer on a dub stage and played back in a theater through the corresponding channels</li> <li>○ Sounds can be positioned where speakers are located or panned between the channels as 'phantom sources'</li> <li>○ Can be easily mixed down to fewer channels or encoded</li> <li>○ More compatible with existing production audio equipment and workflows</li> <li>○ More capable of handling 'Wet' sounds (natural ambient sounds and effects captured in the native channel count)</li> <li>○ Advanced mixing and mastering tools (Plug-ins for Avid Protools, Nuendo) are already available (AMS/Neve DFC consoles with 3D routing and panning)</li> <li>○ Enables single inventory distribution and uses no 'lossy' compression system</li> <li>○ Auro-encoded masters are standard linear PCM.</li> <li>○ DCI Compliant</li> </ul>	<ul style="list-style-type: none"> <li>○ Unique sounds or objects are 'steered' around the room based on vector metadata that accompanies the additional audio file in the distribution, which the playback system must render in real-time for each object related to the installed speaker layout.</li> <li>○ Hybrid: 5.1 or 7.1 base layer accompanies the object sound file for sounds that are not output by the rendering station and contains the channel-based stems (or beds) for dialogue, music and other sounds that do not 'move' or require steering</li> <li>○ Unique distribution, separate KDM/CPL</li> <li>○ Requires more unique channels and speakers...24-64</li> <li>○ Good for 'Dry' sounds – but 'Wet' sounds (natural ambient sounds or sounds rich in reflections like an orchestra or crowd) cannot easily be made immersive</li> <li>○ New workflows and content creation tools are required</li> <li>○ DCI Compliance unknown</li> <li>○ Increased immersive audio-visual experience not guaranteed. The opposite might be the case.</li> </ul>

## BIBLIOGRAPHY

- Blauert, Jens. *Spatial Hearing: The psychophysics of human sound localization*. Revised Edition. MIT Press, 1997.
- Digital Cinema Initiatives, LLC. *Digital Cinema Specification*. v1.2. DCI, March 7, 2008.
- Gerzon, Michael. "Periphony - With-Height Sound Reproduction." *AES Journal* (Audio Engineering Society) 21, no. 1 (Jan/Feb 1973): 2-10.
- Hamasaki, K. "Advanced Multichannel Audio Systems with Superior Impression of Presence and Reality." *116th AES Convention*. Berlin, Germany: Audio Engineering Society, 2004.
- —. "The minimum number of loudspeakers and its arrangement for reproducing the spatial impression of diffuse sound field." *113th AES Convention*. Los Angeles, CA: Audio Engineering Society, 2002.
- Holman, Tomlinson. *5.1 Surround Sound: Up and Running*. Focal Press, 2000.
- Holman, Tomlinson. "New Factors in Sound for Cinema and Television." *AES Journal* (Audio Engineering Society) 39, no. 7/8 (Jul/Aug 1991): 529-529.
- —. "The Number of Channels." *100th AES Convention*. Copenhagen: Audio Engineering Society, 1996.
- Ioan, Allen. "Matching to Sound to Picture." *AES 9th International Conference: Television Sound Today And Tomorrow*. Detroit, MI: Audio Engineering Society, 1991.
- Jang, Dae-young, Jeongil Seo, Kyeongok Kang, and Hoe-Kyung Jung. "Object-based 3D Audio Scene Representation." *115th AES Convention*. New York, NY: Audio Engineering Society, 2003.
- Kim, Sunmin, Young Woo Lee, and Ville Pulkki. "New 10.2-channel Vertical Surround System (10.2-VSS); comparison study of perceived audio aulity in various multichannel sound systems with height loudspeakers." *129th AES Convention*. San Francisco, CA: Audio Engineering Society, 2010.
- Society of Motion Pictures and Television Engineers. *D-Cinema Distribution Master - Audio Characteristics*.
- Society of Motion Pictures and Television Engineers. *D-Cinema Distribution Master Audio Channel Mapping and Channel Labelling*.
- Theile, G., and H. Wittek. "Surround Recordings with Height." *130th AES Convention*. London, UK: Audio Engineering Society, 2011.
- Uggelberg, Anders. "Design Considerations for Cinema Surround Speakers." *91st AES Convention*. Audio Engineering Society, 1991.