



北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

# Interactive an Immersive Sound Over IP Networks





深圳  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

# LiveVideoStackCon 2019 深圳

2019.12.13-14



成为讲师: [speaker@livevideostack.com](mailto:speaker@livevideostack.com)

成为志愿者: [volunteer@livevideostack.com](mailto:volunteer@livevideostack.com)

赞助、商务合作: [kathy@livevideostack.com](mailto:kathy@livevideostack.com)

出品:  **LiveVideoStack**   
—— 音视频技术社区 ——



北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

1. Who is Fraunhofer IIS?
2. Fraunhofer IIS' Audio Technologies
3. China 3D Audio vs. MPEG-H 3D Audio
4. Merits of MPEG-H 3D Audio
5. Streaming Signal Chain with MPEG-H 3D Audio
6. Case Studies

# Fraunhofer Association

## Europe's Largest Applied Research Organization

- Founded 1949 in Munich, Germany
- 69 Institutes at 40 locations
- 26,000 staff
- 2.6 billion € annual research budget
- Research centers and representative offices in Europe, USA, Asia and in the Middle East



# Fraunhofer IIS

## Enabling True Audio & Multimedia Experiences



Main inventors of

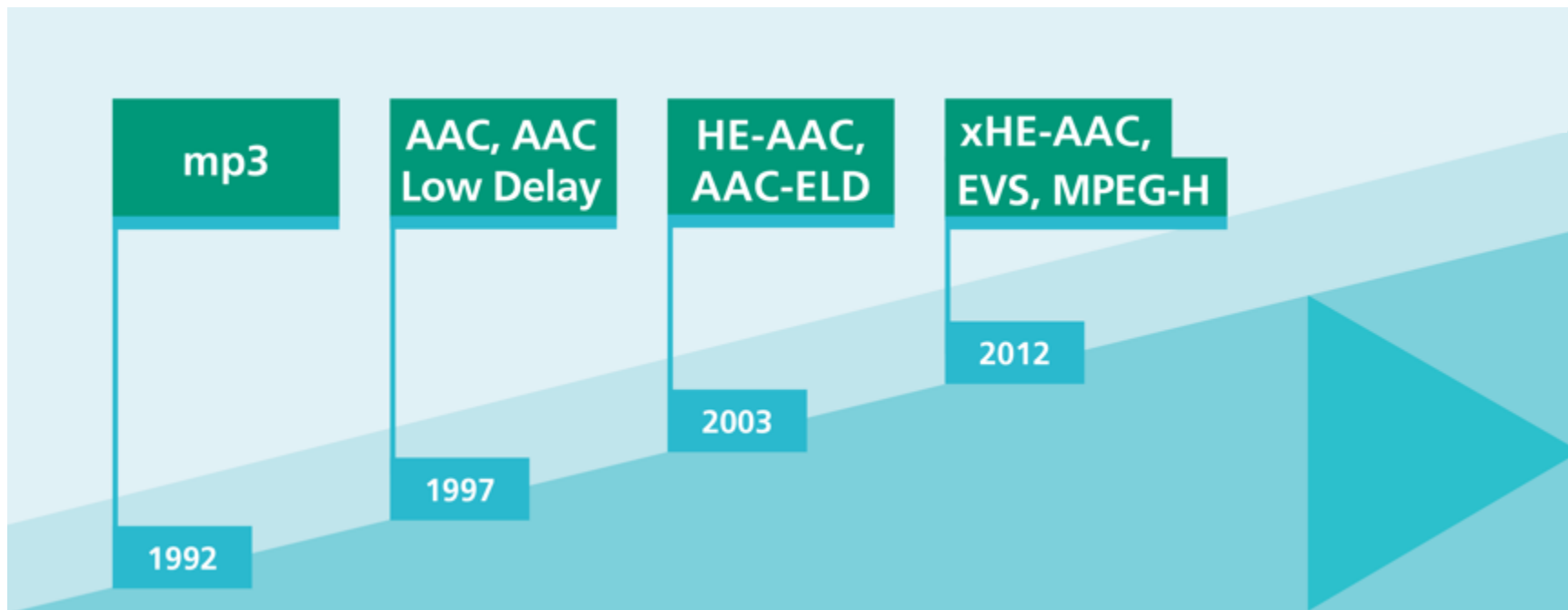


Co-developers of AAC

- Fraunhofer IIS is the worldwide leading center of competence in the field of audio and multimedia codecs and processing
- More than 1000 software customers
- More than 12 billion devices
- More than two decades of experience

# Main Inventor of mp3 and Co-developer of (xHE-)AAC, MPEG-H Audio and EVS

## Four Generations of Best-in-Class Audio Coding Technologies





# Fraunhofer IIS Audio Technology Across Consumer Electronic Products

## Common Denominator in Broadcast, Internet and Mobile Applications



# Fraunhofer IIS Supports Open Standards

## Specified by Major Standards and Integrated in Widely Used Platforms



native support for HE-AAC 5.1  
in iOS and Android



EVS mandatory  
(expected soon)







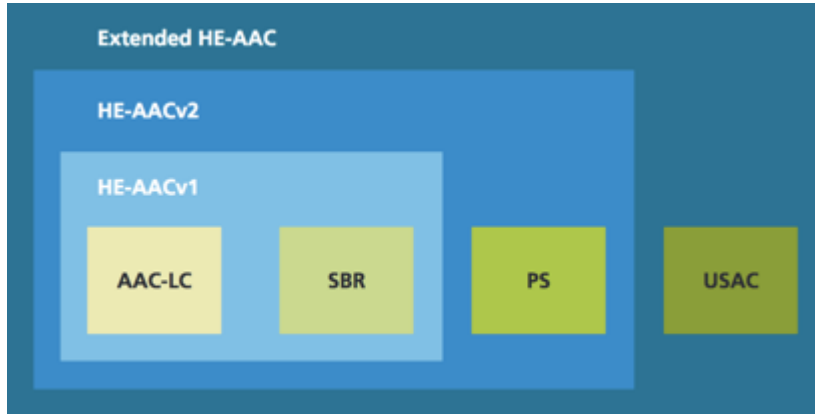
北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

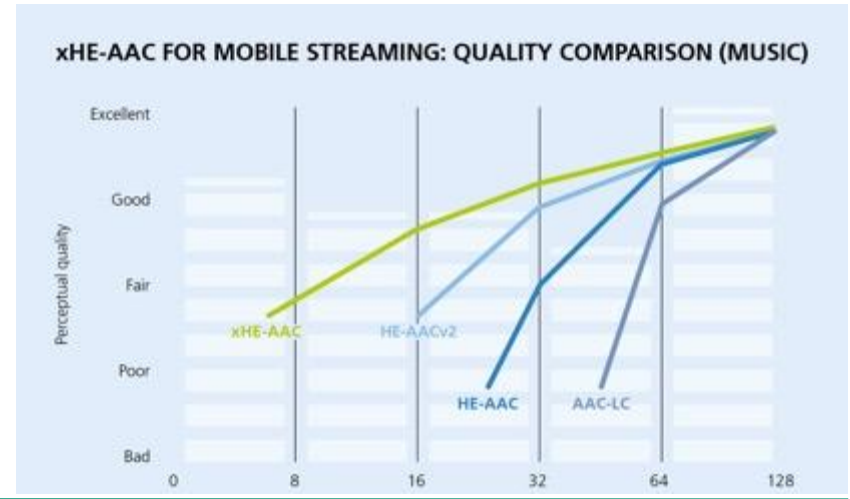
1. Who is Fraunhofer IIS?
2. Fraunhofer IIS' Audio Technologies
3. China 3D Audio vs. MPEG-H 3D Audio
4. Merits of MPEG-H 3D Audio
5. Streaming Signal Chain with MPEG-H 3D Audio
6. Case Studies

# What is xHE-AAC?

## Latest Addition to the AAC Family of Codecs



- Optimized for lowest bitrates
- Demand from major content providers and „2G countries“



- Backward compatible
- Native support in Android (from 9) and iOS (from 13)

# What is EVS?

## Enhanced Voice Services for VoLTE Networks and more

- The next generation 3GPP communication codec (after AMR-WB, 2001)
- Substantially improved
  - Speech quality and compression efficiency
  - Quality for non-speech content (mixed content, music)
  - Audio bandwidth (superwideband, fullband)
- Higher error robustness
- Integrated AMR-WB for seamless switching from/to EVS
- Result of a cooperation of 12 companies:

ERICSSON

HUAWEI

NTT

orange™

QUALCOMM

VoiceAge®

Fraunhofer  
IIS

NOKIA

NTT  
Do Co Mo

Panasonic

SAMSUNG

ZTE中兴

# What is MPEG-H 3D Audio?

## Immersive and Interactive Audio – Universal Delivery

- Switch between languages
- Turn announcer or dialogue up or down
- Listen to your home team or the pit crew



MPEG-H On-Screen Display

- Viewers becomes part of the audience
- Enthusiasts set up speaker system, others use soundbar



3D Soundbar

- Home Theater
- Kitchen TV
- Tablet
- Phone earbuds
- VR Headset



Table 1. The number of participants in each of the four groups, and the number of participants who completed the study

Group	Number of participants	Number of participants who completed the study
Control	10	10
Low-dose	10	10
High-dose	10	10
Low-dose + high-dose	10	10

control group, and the low-dose group. The high-dose group was significantly heavier than the control group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the low-dose group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the control group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the low-dose group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the control group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the low-dose group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the control group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the low-dose group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the control group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the low-dose group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the control group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).

The high-dose group was significantly heavier than the low-dose group ( $p < 0.05$ ), and the low-dose group was significantly heavier than the control group ( $p < 0.05$ ).



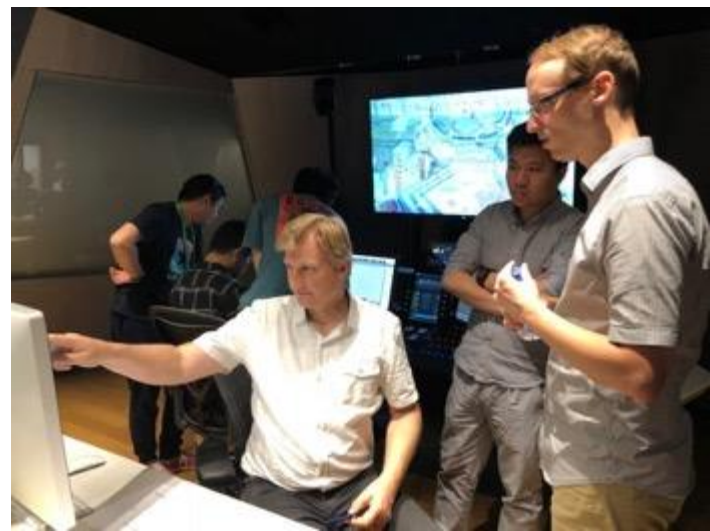
北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

1. Who is Fraunhofer IIS?
2. Fraunhofer IIS' Audio Technologies
3. China 3D Audio vs. MPEG-H 3D Audio
4. Merits of MPEG-H 3D Audio
5. Streaming Signal Chain with MPEG-H 3D Audio
6. Case Studies



为期1年多的标准测试工作  
More than One Year Exhaustive Testing

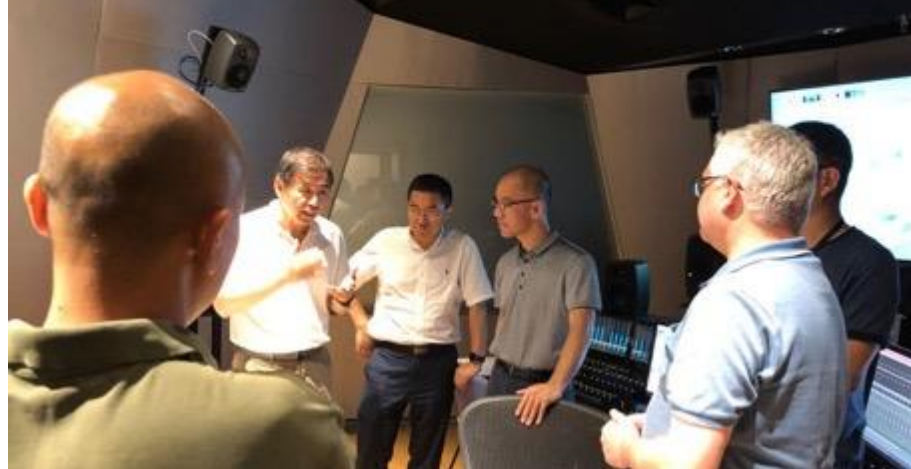


## MPEG-H TV Audio Trials in China, June 2018



向CCTV5领导演示世界杯三维声信号测试成果  
Successful Demonstration to CCTV5 Leadership

## 向工信部领导演示世界杯三维声信号测试成果 Successful Demonstration to MIIT Leadership



向广电总局领导演示世界杯三维声信号测试成果  
Successful Demonstration to NRTA Leadership



# China 3D Audio (C3DA) is Based on MPEG-H LC

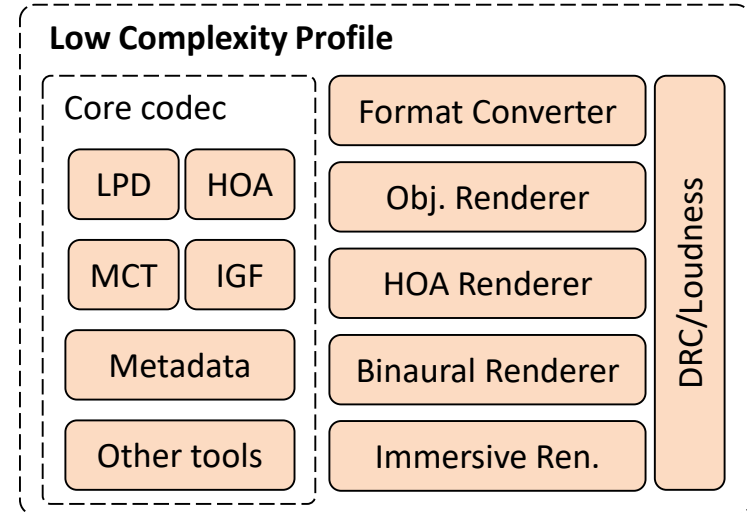
## Development Objectives

- MPEG-H bitstream compatible
- Highest efficiency in terms of bitrate and computational requirements
- Optimized IPR portfolio
- Implementation software to include the full feature set from the beginning
  - Support for 5.1+4 channels plus objects
  - Full interactivity support for language switching and to turn sound elements up and down
  - Software is currently provided for free

# MPEG-H Audio is an International Open Standard (ISO/IEC 23008-3)

## Large set of tools – Subsets (Profiles) for Specific Applications

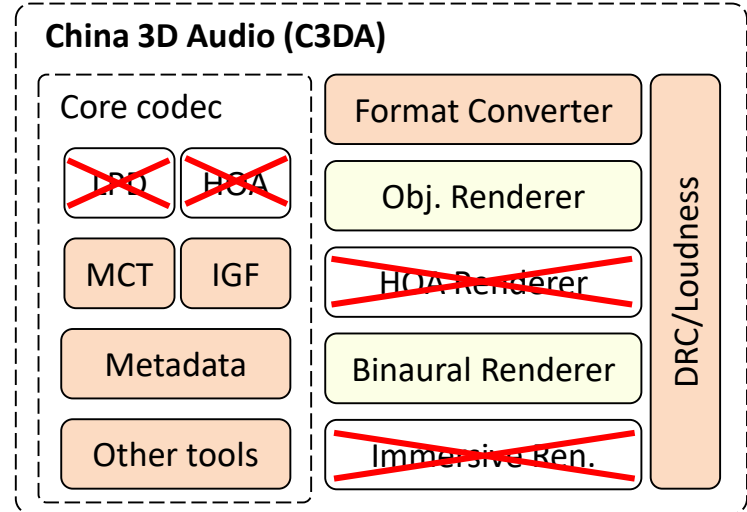
- The LC Profile contains all tools which are needed for Channel, Object and HOA coding and rendering
- The first profile adopted by applications standards (e.g., DVB or ATSC)
- Application standards permitted to apply further restrictions
  - Eg., DVB allows for exchange of rendering engine “if perform at least as well as the reference renderer”



# China 3D Audio (C3DA) is Tailored for Streaming and Broadcasting

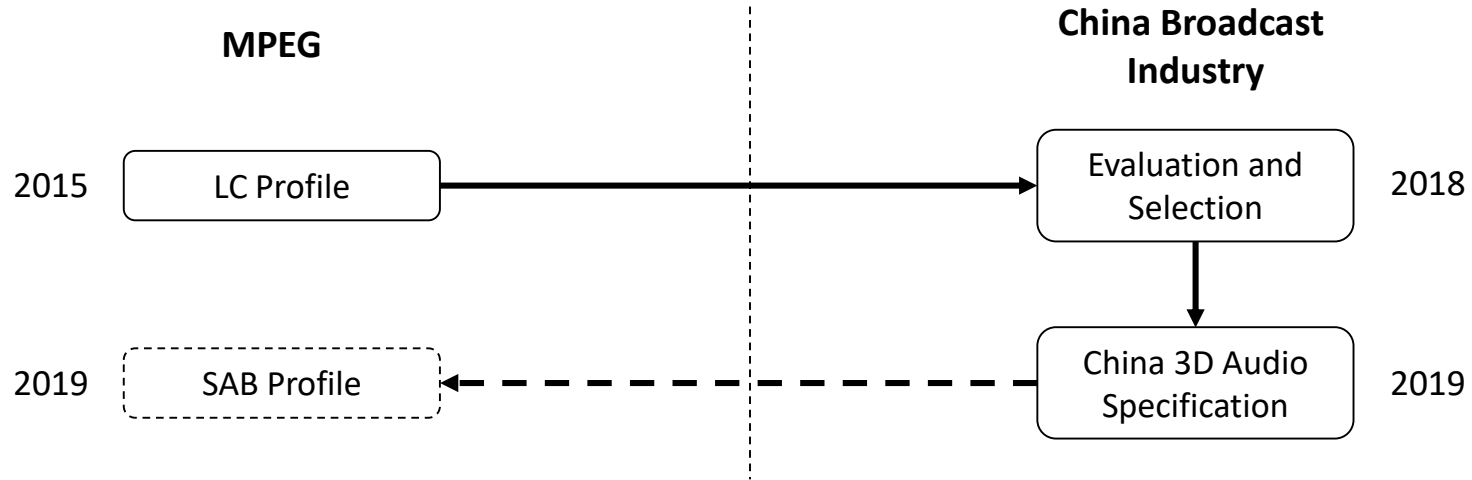
## Removes Unnecessary Tools – Allows for Exchange of Tools

- The current C3DA draft specification contains:
  - Coding tools for Channels and Objects from MPEG-H 3D Audio
  - Auro Object/Binaural Renderer
- The C3DA bitstream is compliant to MPEG-H 3DA LC Profile syntax
- All removed tools (e.g., HOA, LPD) are forbidden in the bitstream.



# Plan: Creating an International Version of C3DA

## MPEG to Embrace C3DA as „Streaming and Broadcast“ (SAB) Profile



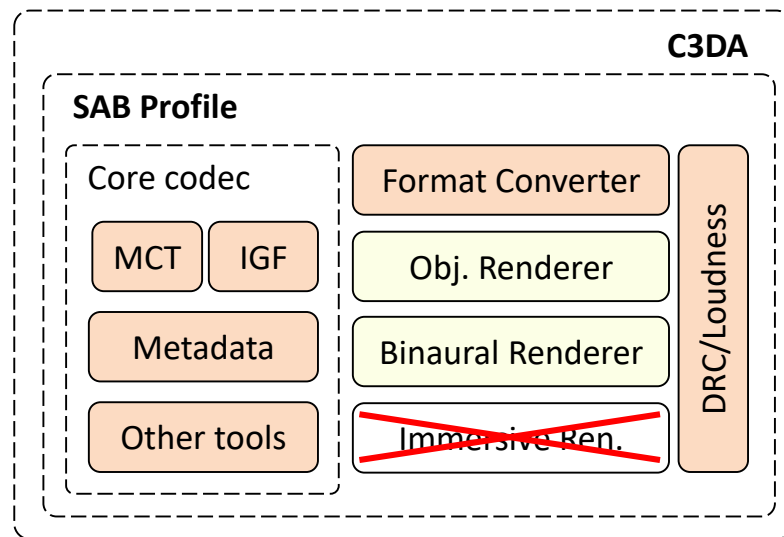
- C3DA represents the best solution for broadcast and streaming applications today
- Contains all essential tools for highest audio quality and best user experience
- This initiative is supported by influential broadcasters and CE companies



# Relation between SAB and C3DA

## National Constraints apply – Format Remains Compatible

- SAB and C3DA use the same bitstream syntax and coding tools
- Audio rendering engine can be exchanged
- Similar to DVB, the C3DA specification can be expressed as an application standard which specifies additional constraints on top of the SAB Profile





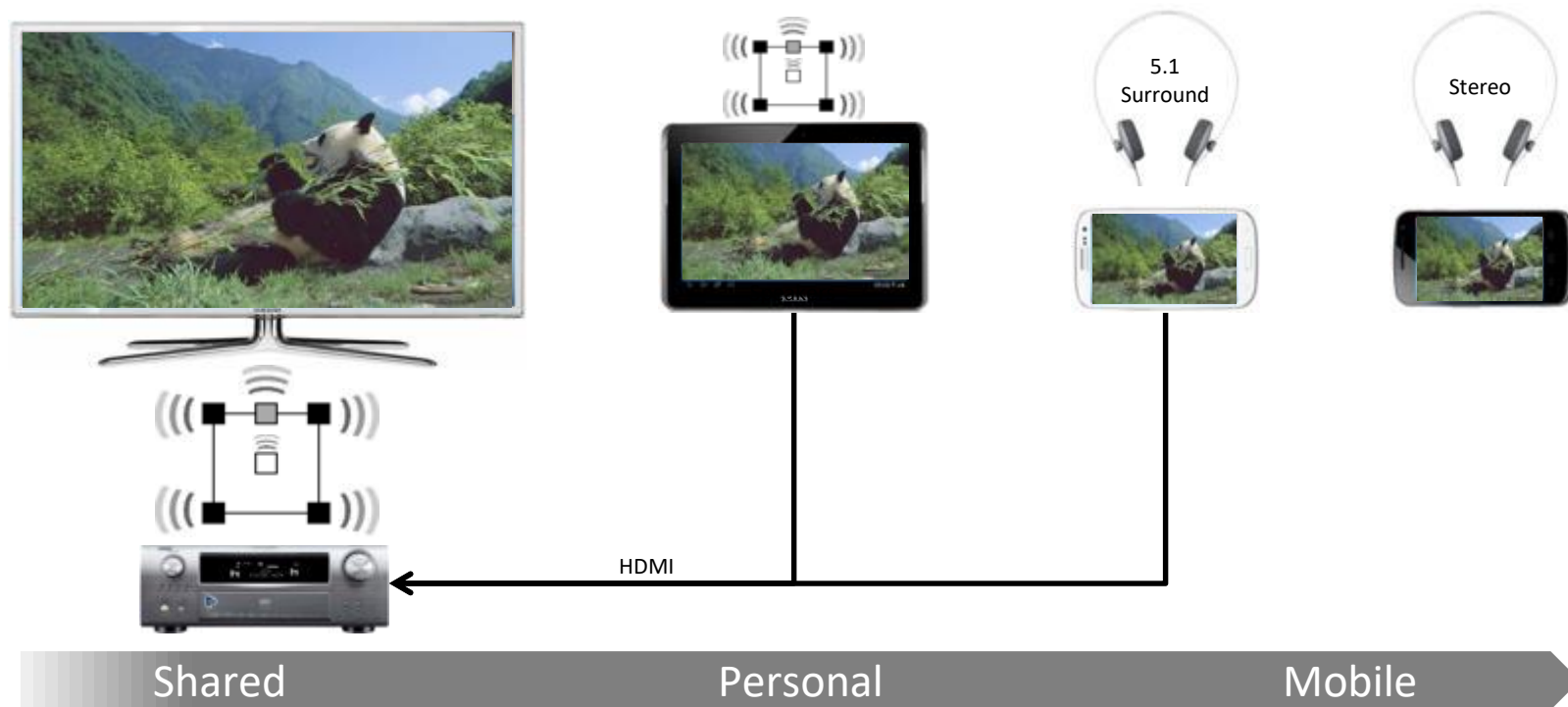
北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

1. Who is Fraunhofer IIS?
2. Fraunhofer IIS' Audio Technologies
3. China 3D Audio vs. MPEG-H 3D Audio
- 4. Merits of MPEG-H 3D Audio**
5. Streaming Signal Chain with MPEG-H 3D Audio
6. Case Studies

# MPEG-H 3D Audio Facilitates Universal Delivery

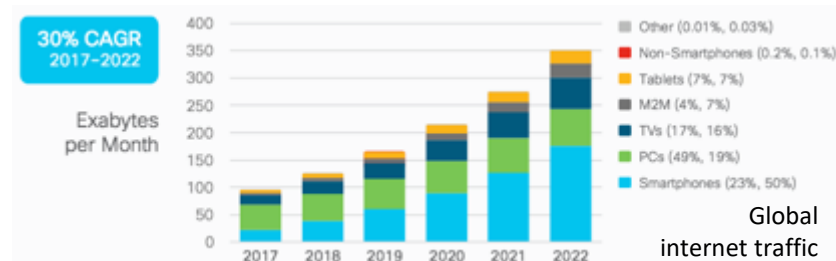
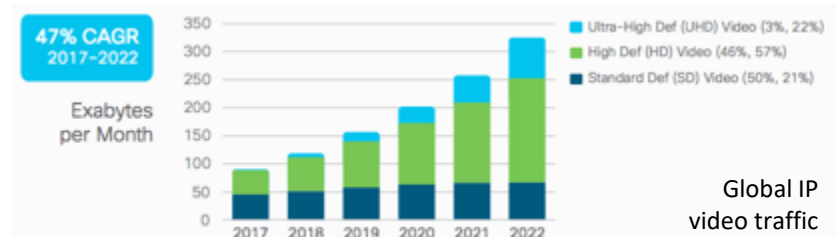
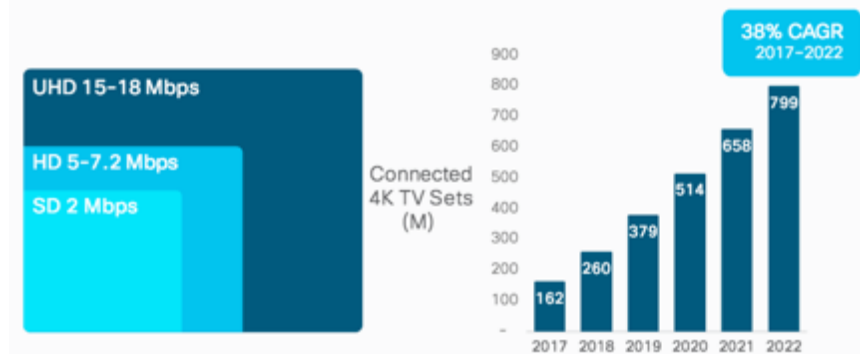
Providing High Quality Interactive and Immersive Audio to All Device Categories



# MPEG-H 3D Audio Targets High Quality Entertainment Consumption

## Connected TV Sets to Become the Second Biggest Streaming Receiver Population

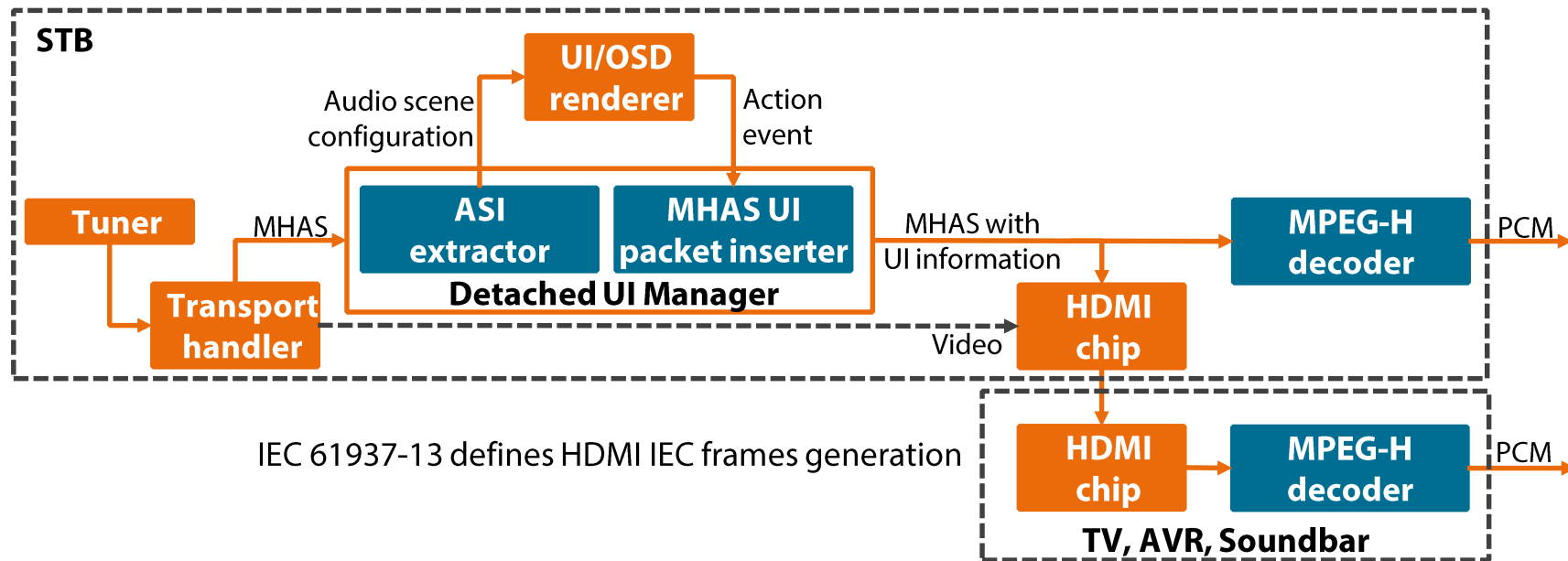
- Rapid adoption of 4k flat-panel TV
- 62% (800m) internet connected by 2022



Source: Cisco global IP traffic forecast 2019

# MPEG-H 3D Audio System Comes with Built-in Interactivity Possibilities

## The User Interface (UI) and On-Screen Display (OSD) Manager



# With MPEG-H the User Can Adjust the Sound to Individual Preferences

## – Within the Limits the Content Producer Defines



- Manually select a different language of a program
- Change balance between dialogue and background
- Select additional tracks (On-Off)
- Change the position of sound events
- Gain Control of each audio element
- ...



# Virtually magic.

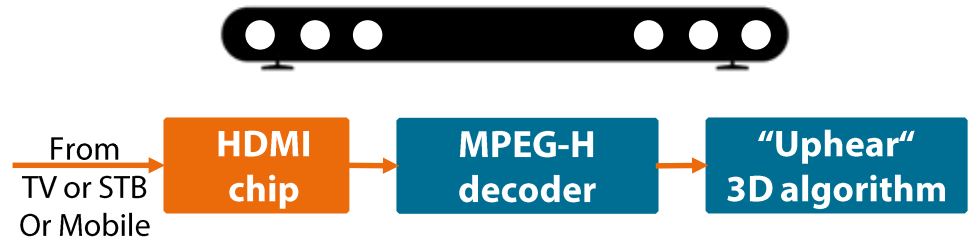
Using the latest virtualization technology jointly developed with Fraunhofer, the AMBEO Soundbar captures knowledge of your room size and its reflective surfaces, adapting the acoustics to fit your individual environment. Soundbar's 13 independent, high end drivers work as virtual speakers around the room, reproducing a 5.1.4 speaker-like installation and placing you in the center of the action.



 **SENNHEISER**

## Immersive Listening Experience Possible with Soundbars

- 3D audio content requires 3D audio listening equipment in the users' homes
  - 5.1 surround plus 4 height speaker
  - 3D audio soundbar





北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

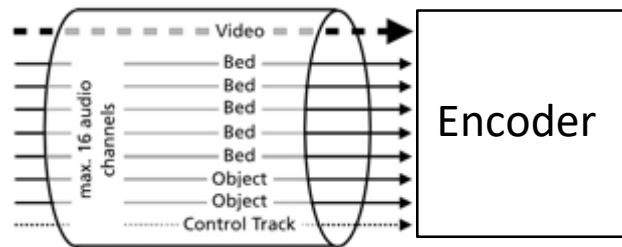
1. Who is Fraunhofer IIS?
2. Fraunhofer IIS' Audio Technologies
3. China 3D Audio vs. MPEG-H 3D Audio
4. Merits of MPEG-H 3D Audio
5. Streaming Signal Chain with MPEG-H 3D Audio
6. Case Studies

# MPEG-H Interactive and Immersive Audio Requires Metadata

## How to Transmit Metadata to the End-User?

Metadata contain information about

- Channel configuration
- Object description (position, volume, on/off)
- Labels of objects for on screen display
- Allowed range of interactivity
- Loudness profiles and downmix parameters
- Descriptive data



Solution: Modulate and transmit metadata as PCM audio track = control track (CT)

- Live broadcast and IPTV: Conveys CT in SDI audio channel 16
- File based playback (VoD): Store CT in WAV/BWF/CMAF

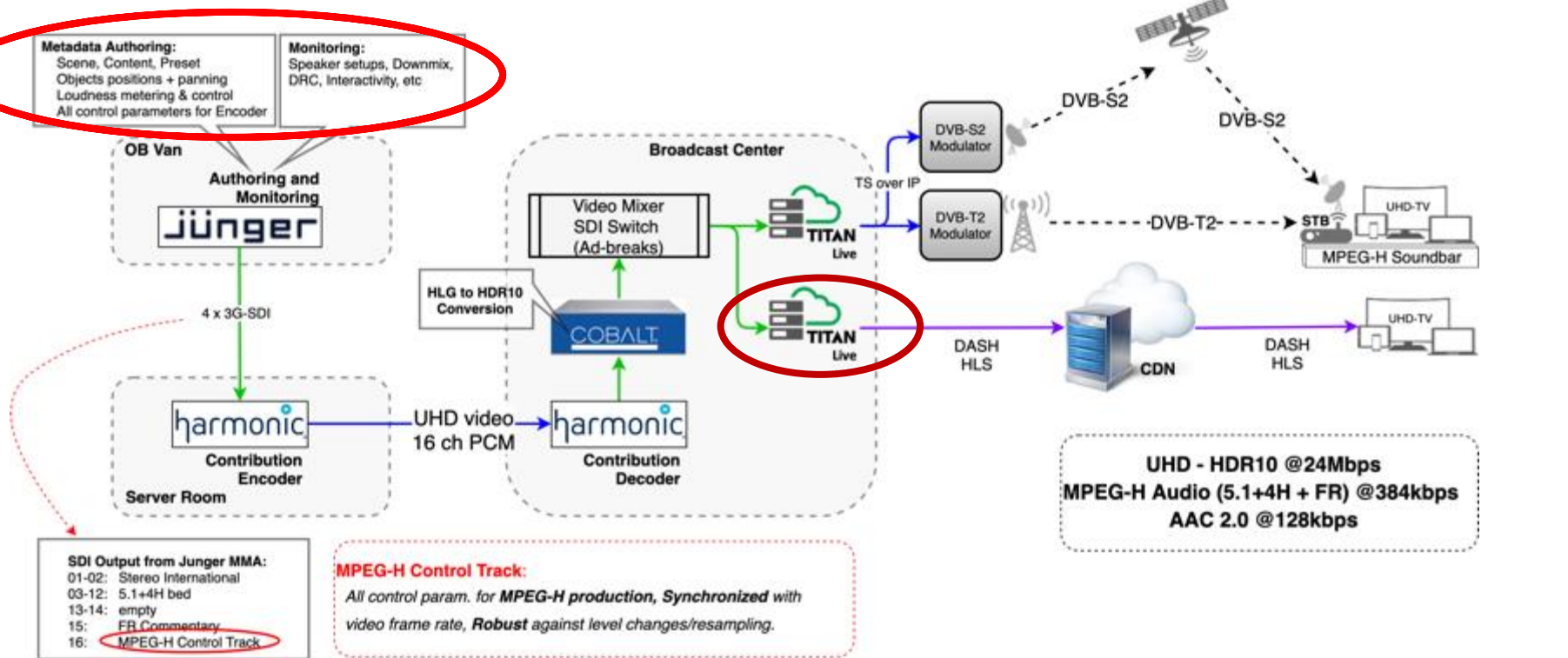
# MPEG-H Audio during the French Tennis Open 2019

## Successful Field Test with Broadcast and IP Live Transmission

france.tv

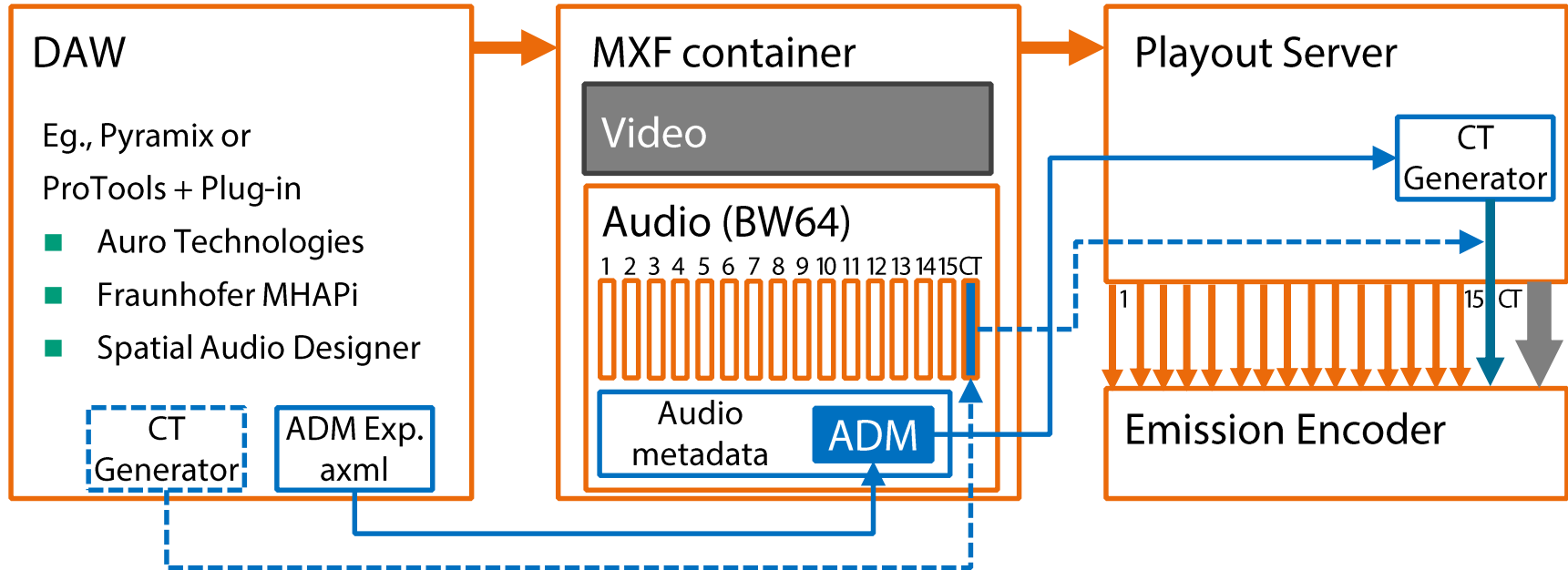


DIFFUSEUR OFFICIEL DEPUIS 30 ANS



# Playout: Archived Content

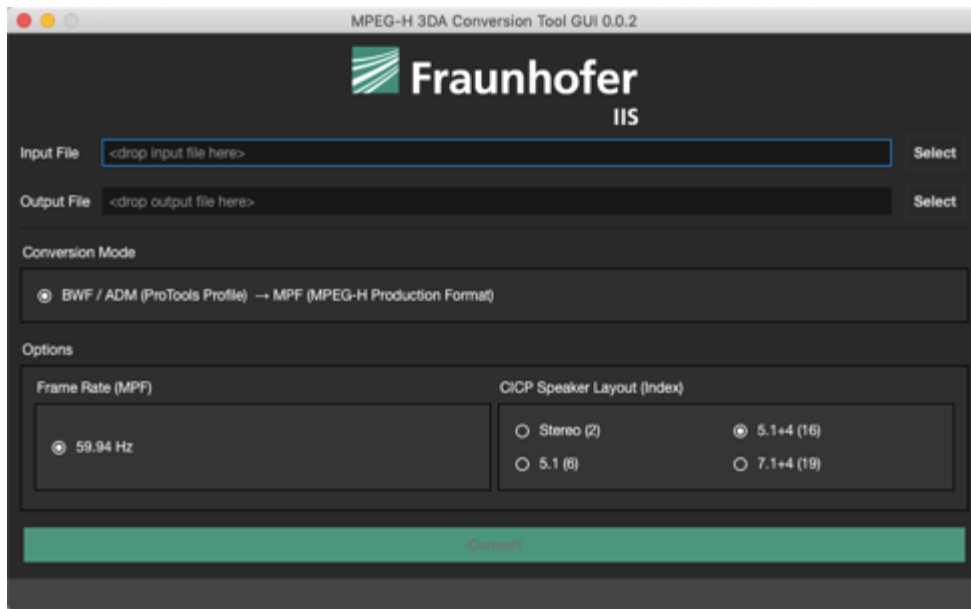
## Storing the Metadata Control Track



# ITU Audio Description Model (ADM)

## Makes Conversion Between Different Production Formats Possible

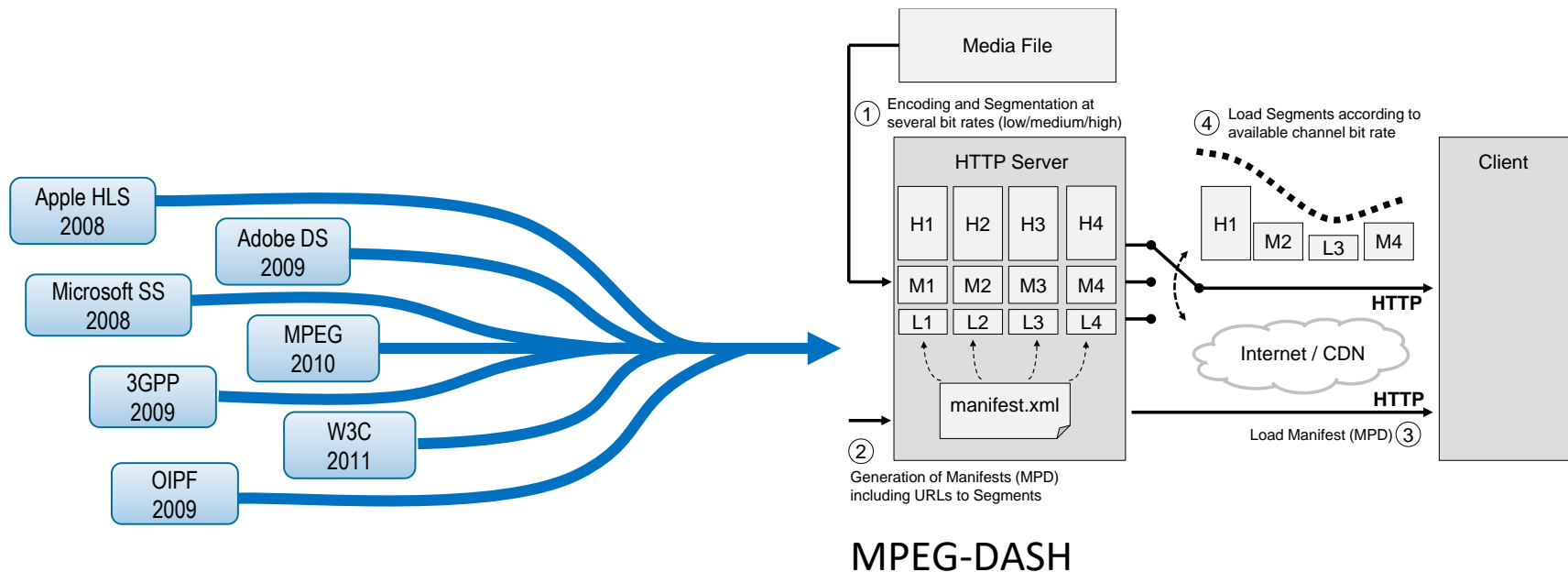
- ITU-R BS.2076
- New: ITU-R BS.2127 (6/2019): ADM renderer for advanced sound systems
- Conversion tool based on ADM available
- Similar activity in SMPTE: „Immersive Audio Bitstream“ (IAB) (SMPTE ST 2098-2:2018)





# MPEG-H 3D Audio Supports Adaptive Streaming

## HLS Still Widely Used – MPEG-DASH is the Future



## Third Party Infrastructure Products to Embrace MPEG-H 3D Audio

### Packetizer – Streaming Server – Cloud Solutions

- MPEG-H 3D audio is a completely new designed system
- Standards documentation just published
- Promoting the adoption in related products

Products available:

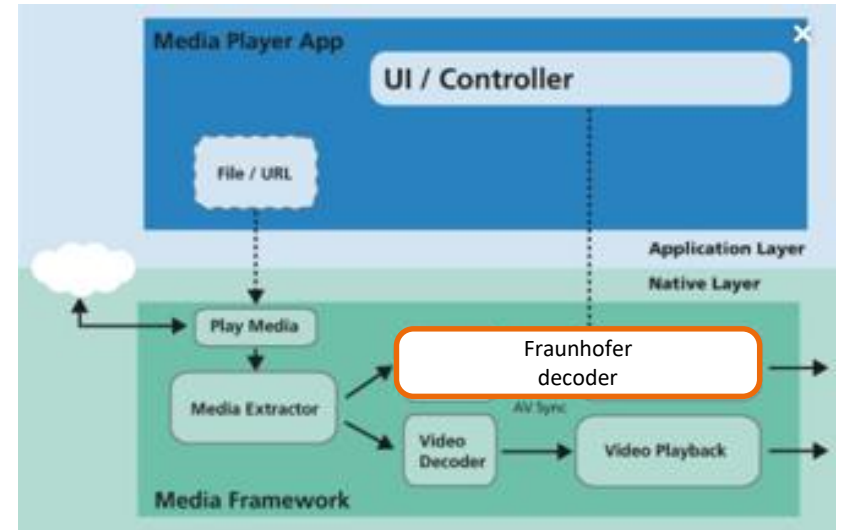
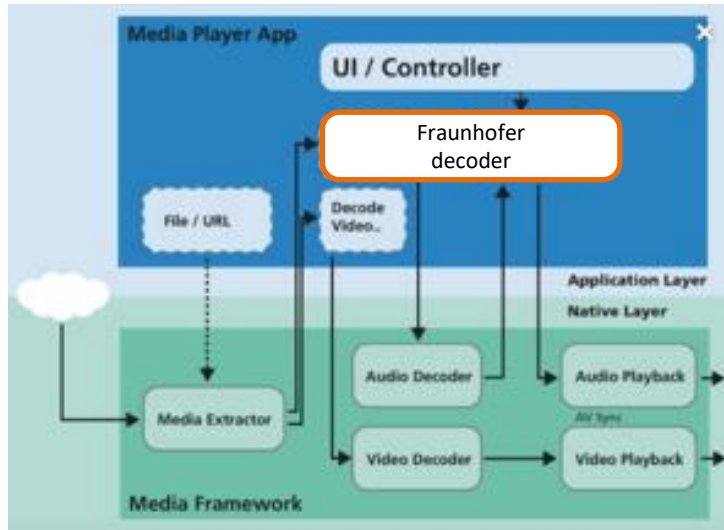


# Media Player Applications

## Native Integration in OS is the Objective



App implementation (eg., ffmpeg) → native OS implementation





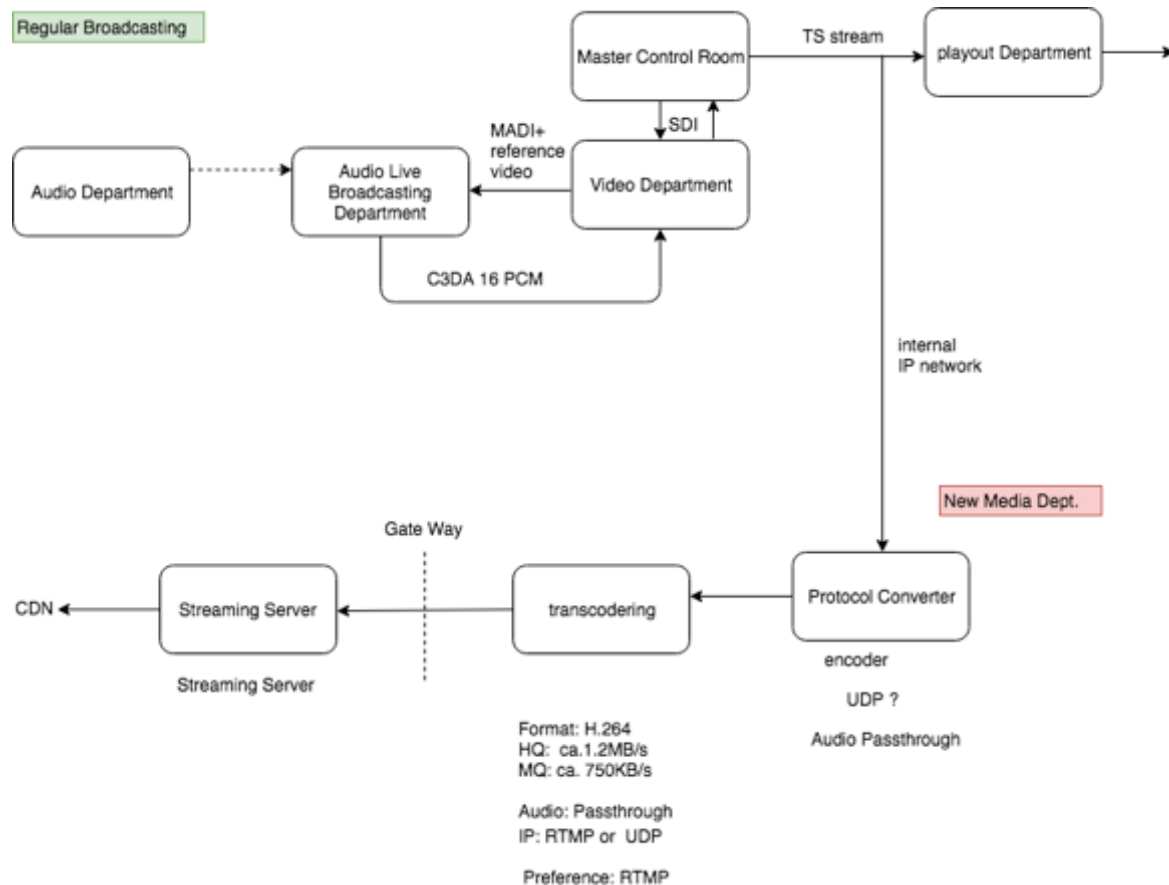
北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

1. Who is Fraunhofer IIS?
2. Fraunhofer IIS' Audio Technologies
3. China 3D Audio vs. MPEG-H 3D Audio
4. Merits of MPEG-H 3D Audio
5. Streaming Signal Chain with MPEG-H 3D Audio
6. Case Studies

# Individual Infrastructure

## Fact Finding



# Identifying User Benefits

## Field Test

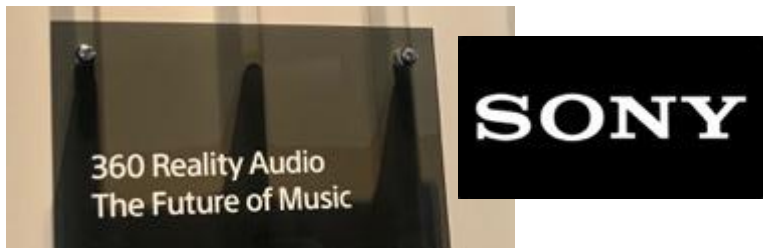
- Additional buttons
  - Add. language / commentary
  - Celebrity speaker
  - Behind the scenes / Fan corner
- Combine with other improvements for premium content eg., HDR
- Links to mobile payment or order services
- New device eg., soundbar along with subscription
- True 360 degree audio for XR



# SONY 360 Reality Audio Music Experience based on MPEG-H

## A New Ecosystem for Music Content Creation, Distribution and Playback

**360**  
REALITY  
AUDIO



- Announced at CES 2019
- An object-based spatial audio platform
- Independent from the loudspeaker configuration at the receiver side.
  - a complete ecosystem from music creation, distribution to playback.
- Supported by Sony Music, Warner Music, UMG, Tidal, Deezer

<https://www.sony.net/SonyInfo/News/Press/201901/19-002E/index.html>

<https://www.sony.de/electronics/360-reality-audio>









北京  
2019

遨游“视”界 做你所想  
Explore World, Do What You Want

# 谢谢大家!

**WE ARE HIRING!!**

Toni Fiedler  
中国区负责人  
[toni.fiedler@iis-extern.fraunhofer.de](mailto:toni.fiedler@iis-extern.fraunhofer.de)  
<http://www.iis.fraunhofer.de/amm/>

