

WebRTC Con

为开发者赋能
为行业加速

2018. 5. 19 - 20 · 上海光大会展中心

姜健 - Google 软件工程师

吴威麒 - 触宝电话音频技术专家
《触宝电话如何改善通话质量》

和君 - TutorABC 大前端负责人

吴涛 - 陌陌视频流媒体技术负责人
《低延迟音视频传输技术在直播领域的应用》

冯巍 - 爱奇艺技术产品中心研究员
《爱奇艺 WebRTC 在智能音箱领域的应用》

张轲 - 腾讯音视频实验室高级工程师
《腾讯语音通信 QoS 优化实践》

陈域将 - 阿里巴巴高级技术专家
《WebRTC 在阿里巴巴的实践》

刁磊 - 哒哒英语研发工程师
《WebRTC 在直播场景问题和优化》

邱建林 - 英特尔实时通信客户端架构师
《WebRTC 之 H.264/265 硬件编解码优化》

更多分享 敬请期待 ...

来一次纯正的 WebRTC 之旅



即刻享受 8 折优惠

LiveVideoStack
— 音视频技术社区 —

WebRTC Con

为开发者赋能
为行业加速

2018. 5. 19 - 20 · 上海光大会展中心

Niklas Blum - Google PM of WebRTC
《Google 解读 WebRTC 优化与改进》

朱浩齐 - 网易云易盾 CTO

盛骁杰 - 优酷 VR 技术专家

Zoe Liu - Google 软件工程师
《AV1 + WebRTC 助力开发者》

冯巍 - 爱奇艺技术产品中心研究员
《爱奇艺 WebRTC 在智能音箱领域的应用》

吴涛 - 陌陌视频流媒体技术负责人

章琦 - 唐桥科技首席架构师
《基于 WebRTC 的多方视频会议融合架构设计》

周正宁 - Aupera 傲睿智存 CTO

陈功 - 声网首席 WebRTC 架构师
《声网的 WebRTC 服务架构与实践》

更多分享 敬请期待 ...

来一次纯正的 WebRTC 之旅



即刻享受 9 折优惠

LiveVideoStack
— 音视频技术社区 —

LiveVideoStack Meet 多媒体开发新趋势

2018年3月31日
微软大厦 东直门会议室



Audio Codec Trend and 3D audio



Agenda



Introduction

Sharing & Discussion

Q&A

2

.....
General Trend of Audio Industry

3

.....
NGA: Benefit and how

4

.....
3D Spatial Audio/Soundfield Audio

Sharing topics

  Audio Codec/Industry Trend

  New Requirement for Audio codec

  NGA

  3D Spatial Audio



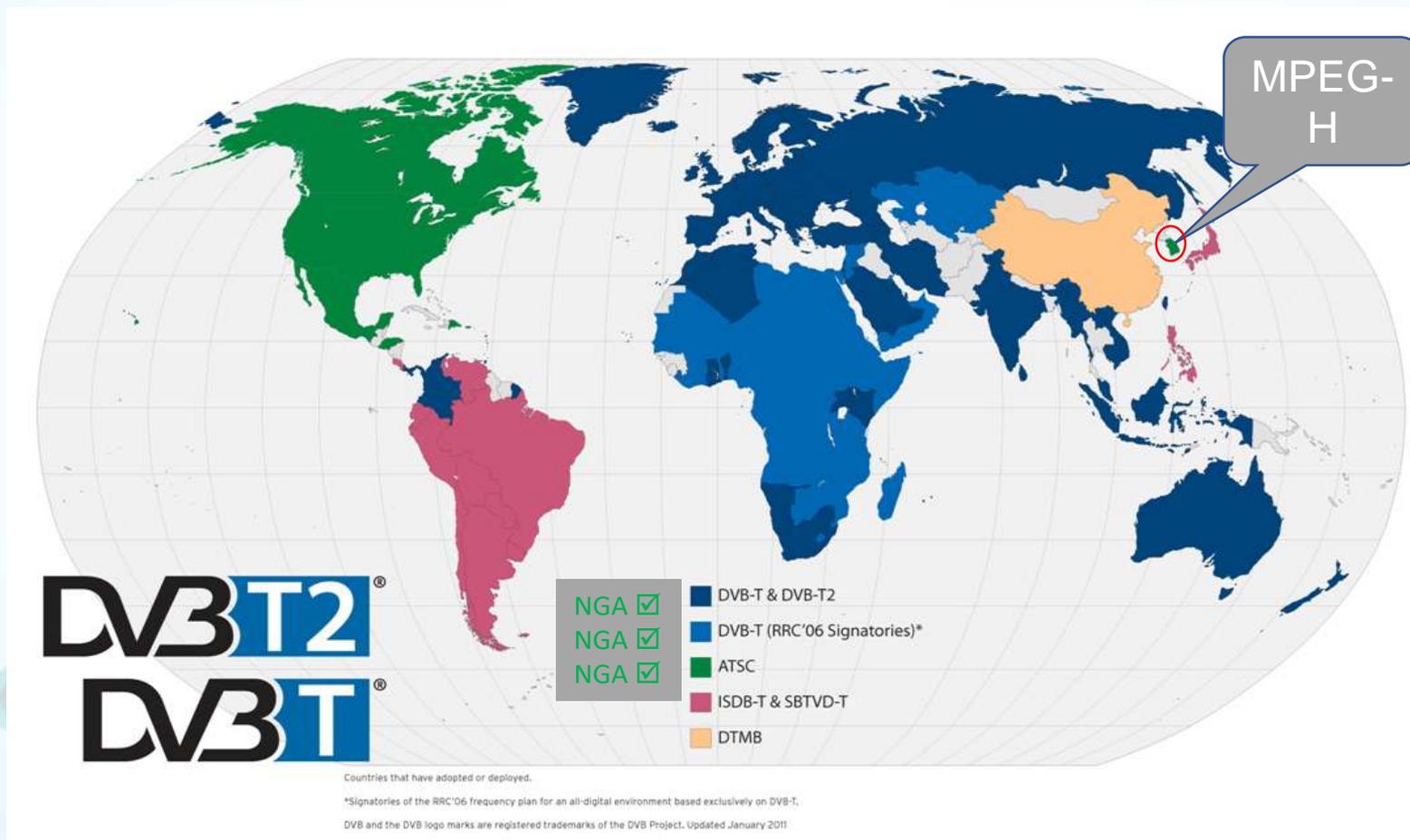
1

Introduction

- **Huang Chuanzeng: Henry & 80s**
- **Master of Acoustics Signal processing from the Institute of Acoustics, CAS**
- **7 year work experiences with audio, mainly with Dolby, just start dating with Bytedance.**

2

General Trend of Audio Industry



AV codec and application trend

New video codec, new audio codec opportunity

MPEG2 – Digital

MPEG2 – Digital

HEVC - UHD

New video codec, new audio codec from MPEG

MPEG1 L-II - 2.0 vs
5.1

MPEG1 L-II - 2.0 vs
5.1

MPEG-H - More Open Source nonsense

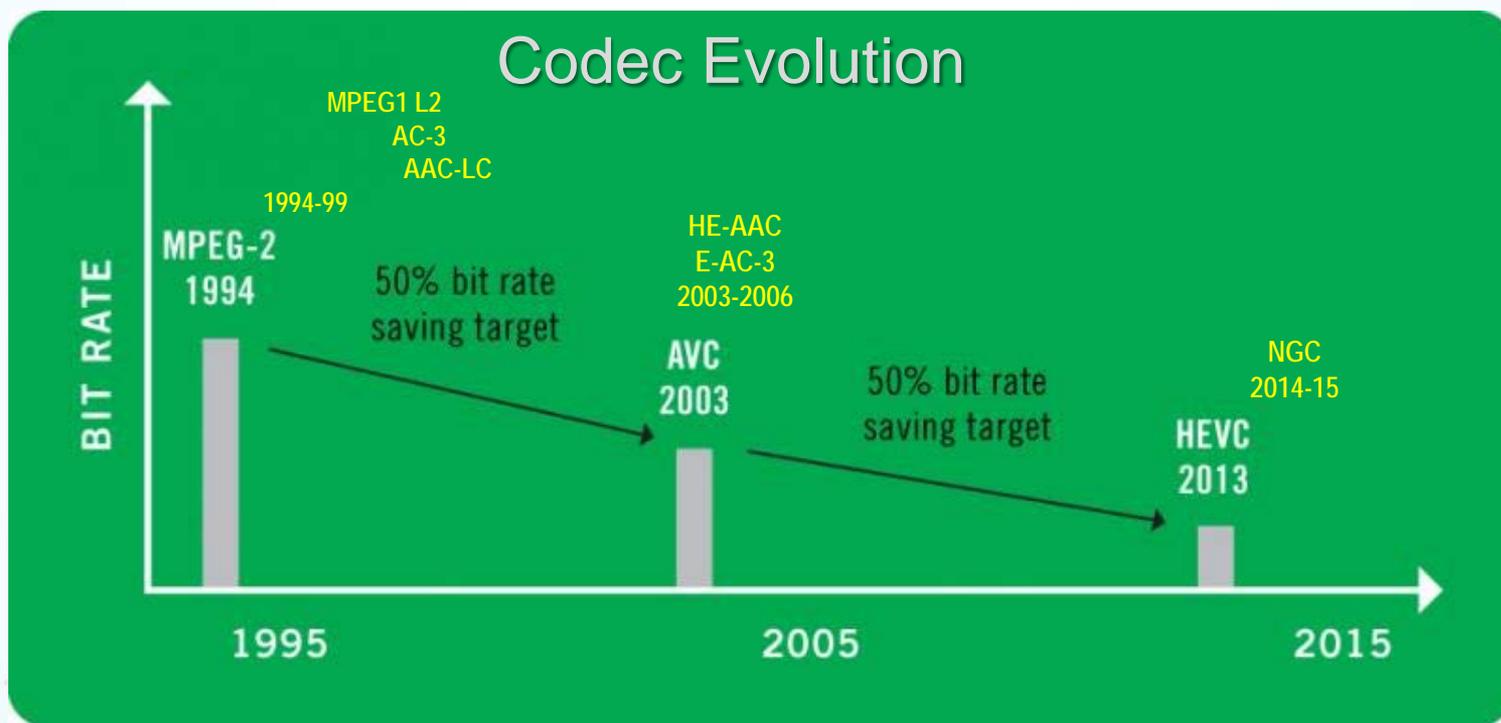
Broadcast and E-Media have merged

DTV - ATSC, DVB

DTV - ATSC, DVB

SmartTV – Netflix, Amazon...

Audio Codecs evolution



Improving Today



Improve Reach

Efficient delivery
to all devices

Simplify Operations

Intelligent loudness,
A/V frame alignment

Enhance Accessibility

Descriptive audio,
dialogue enhancement

Bridge to the Future

Deliver today's audio
and add new features
later

Enabling Tomorrow



Scalable



Personalized



Immersive

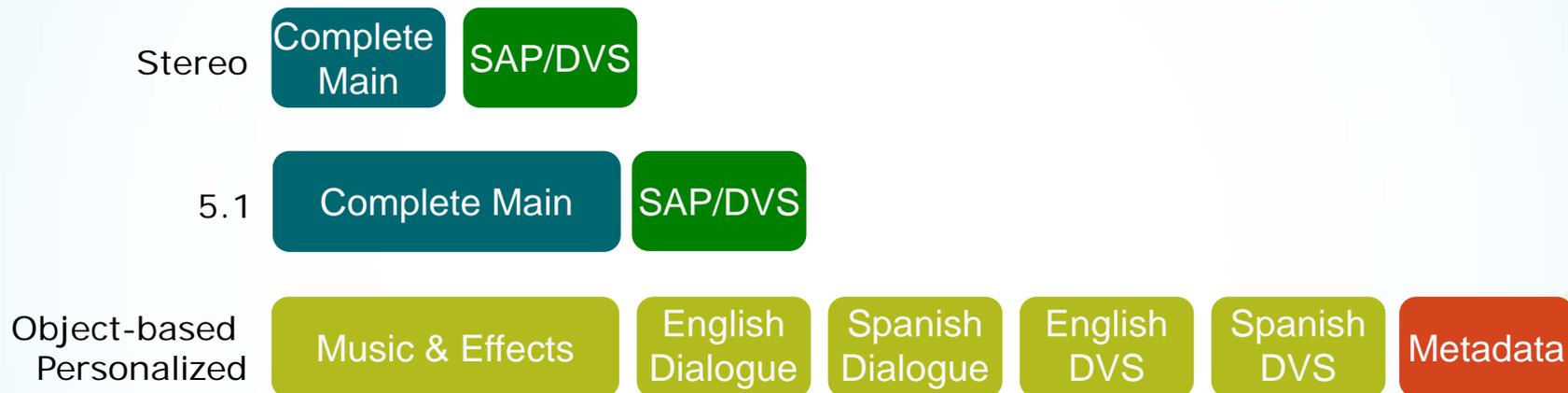


End to End

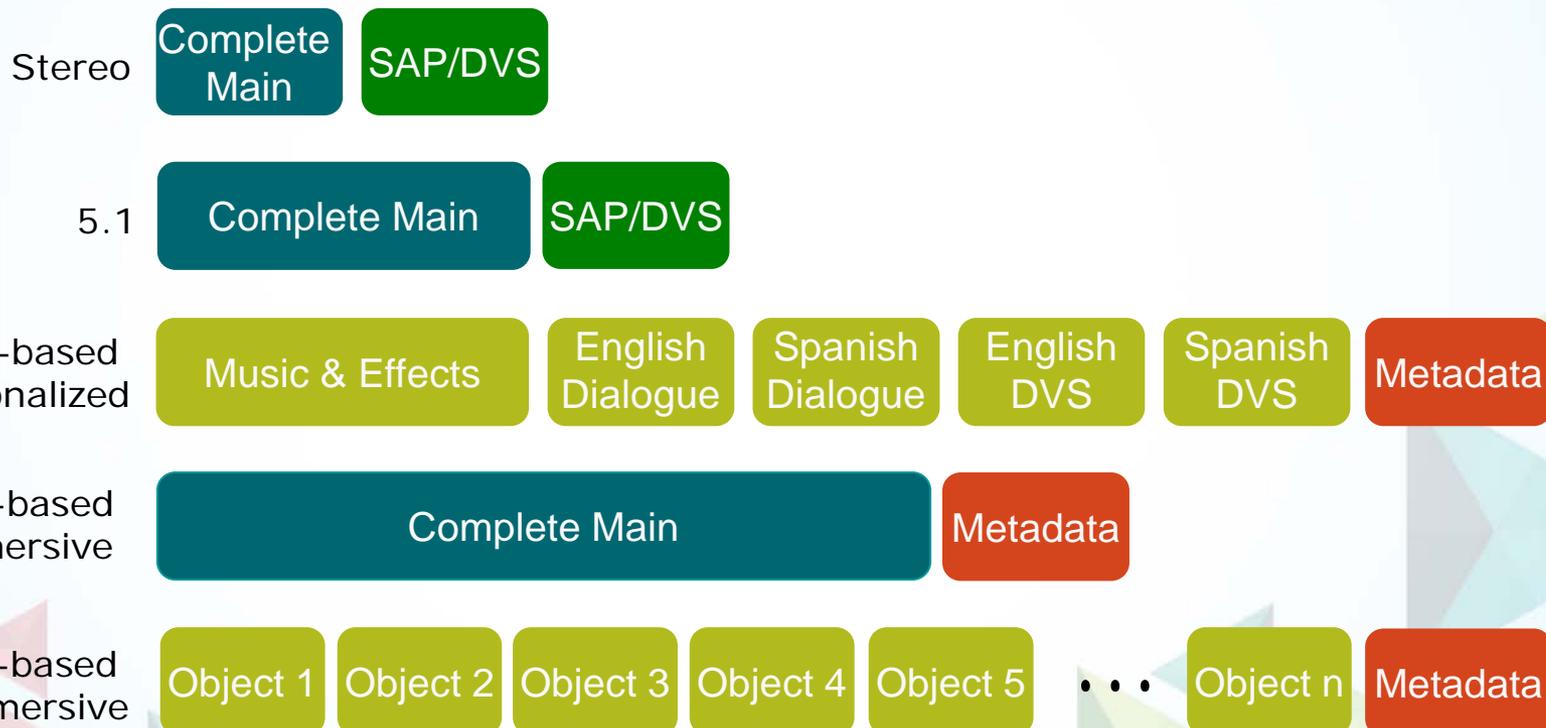
Today's Broadcast Audio



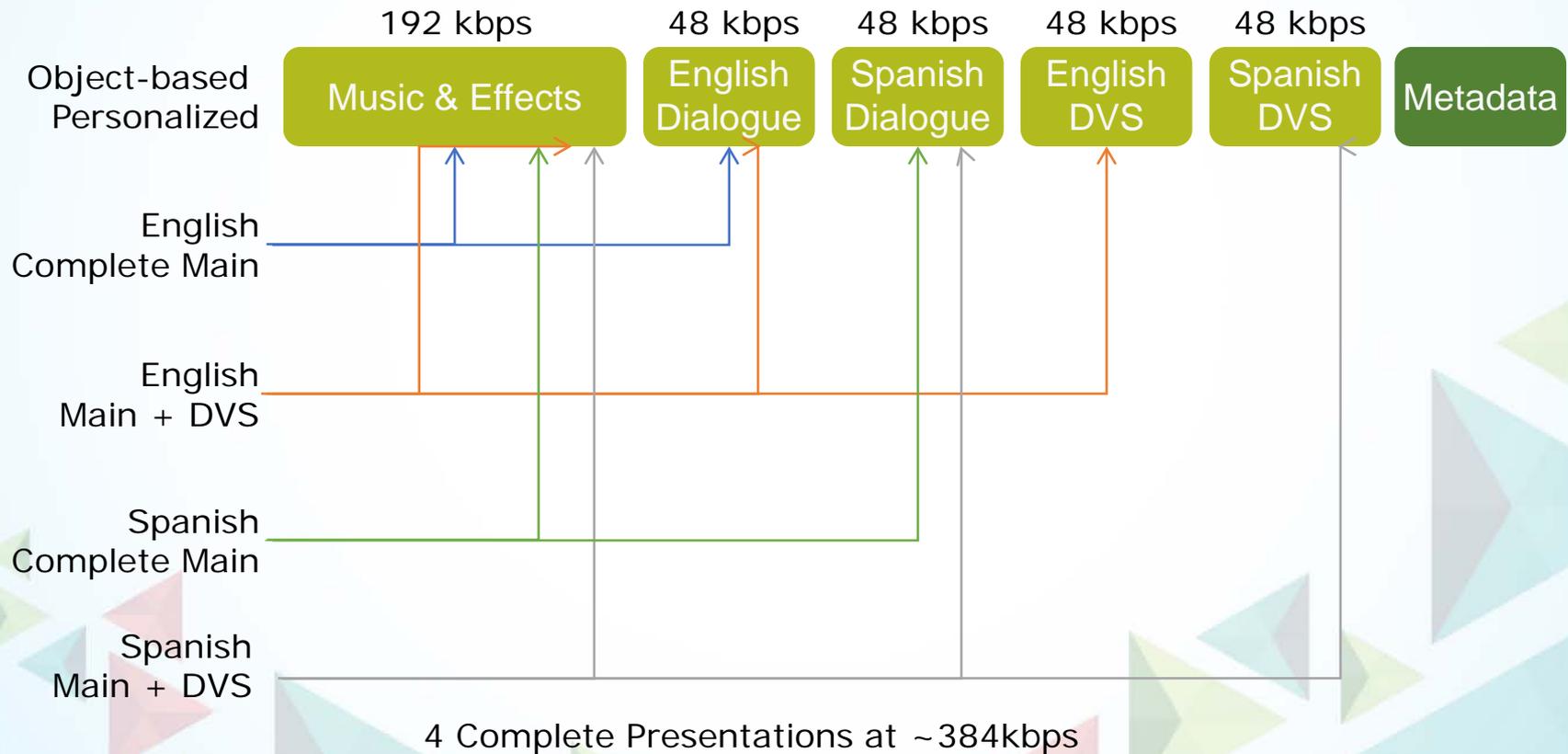
Improving Upon Today



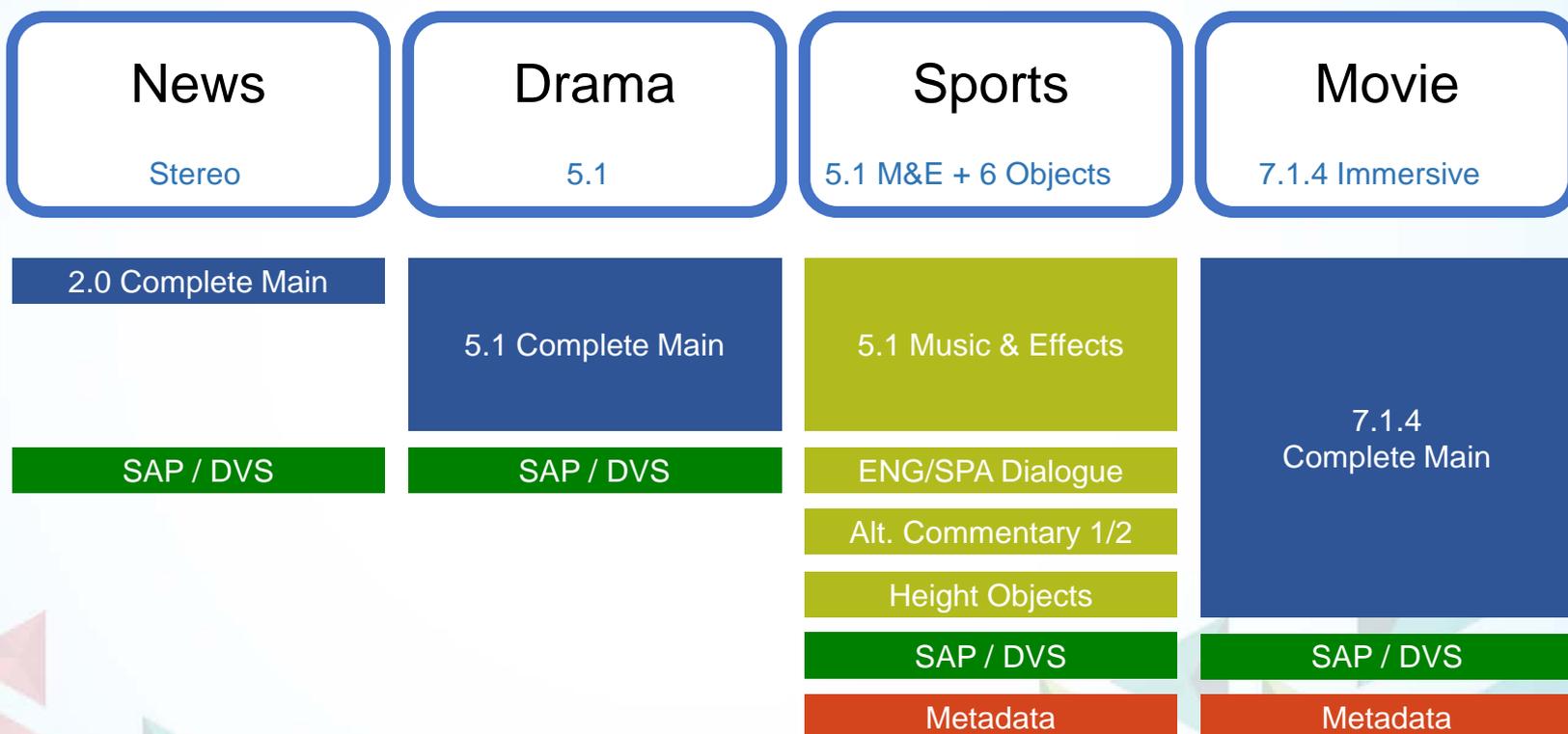
Tomorrow's Broadcast Audio



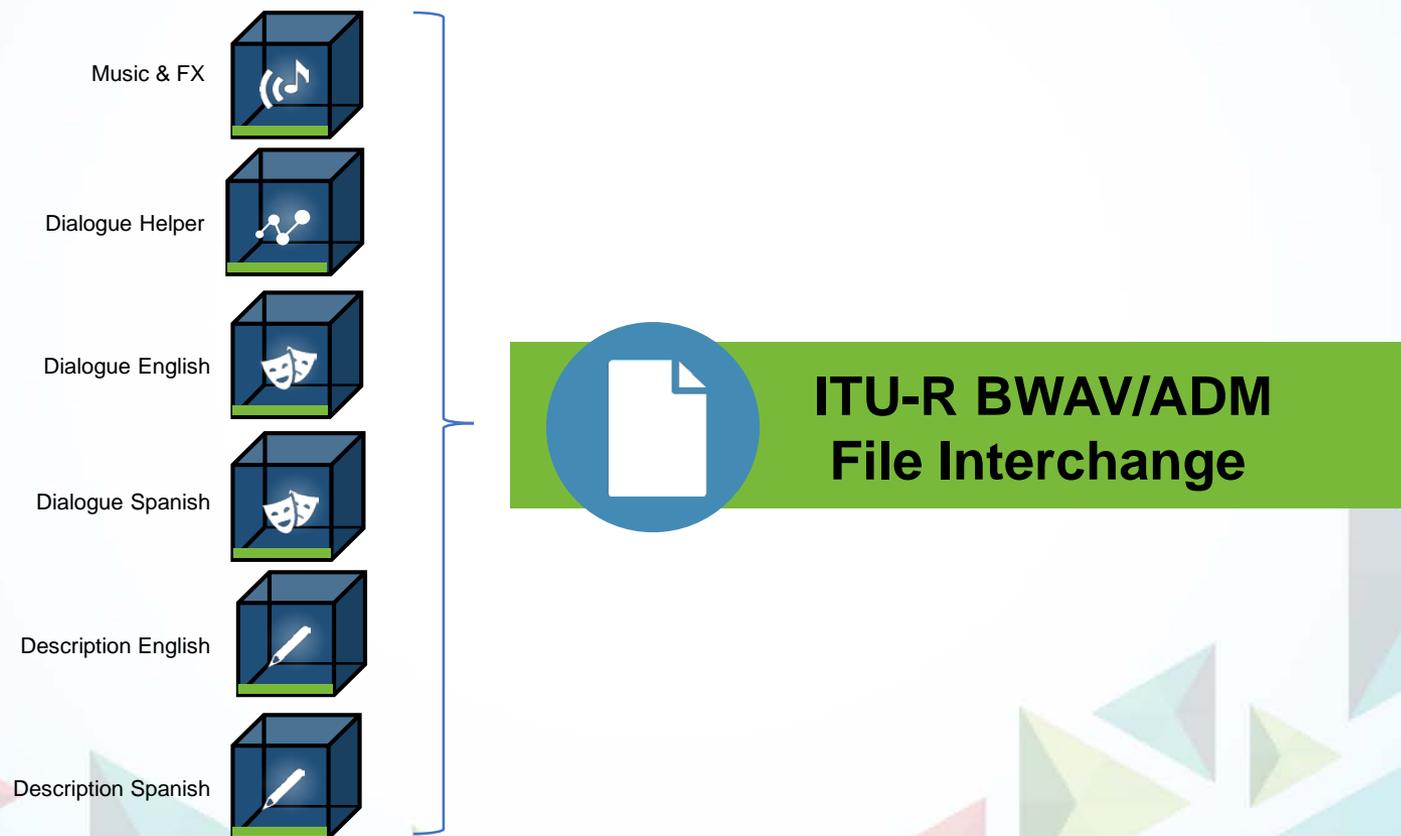
Efficient Personalized Audio



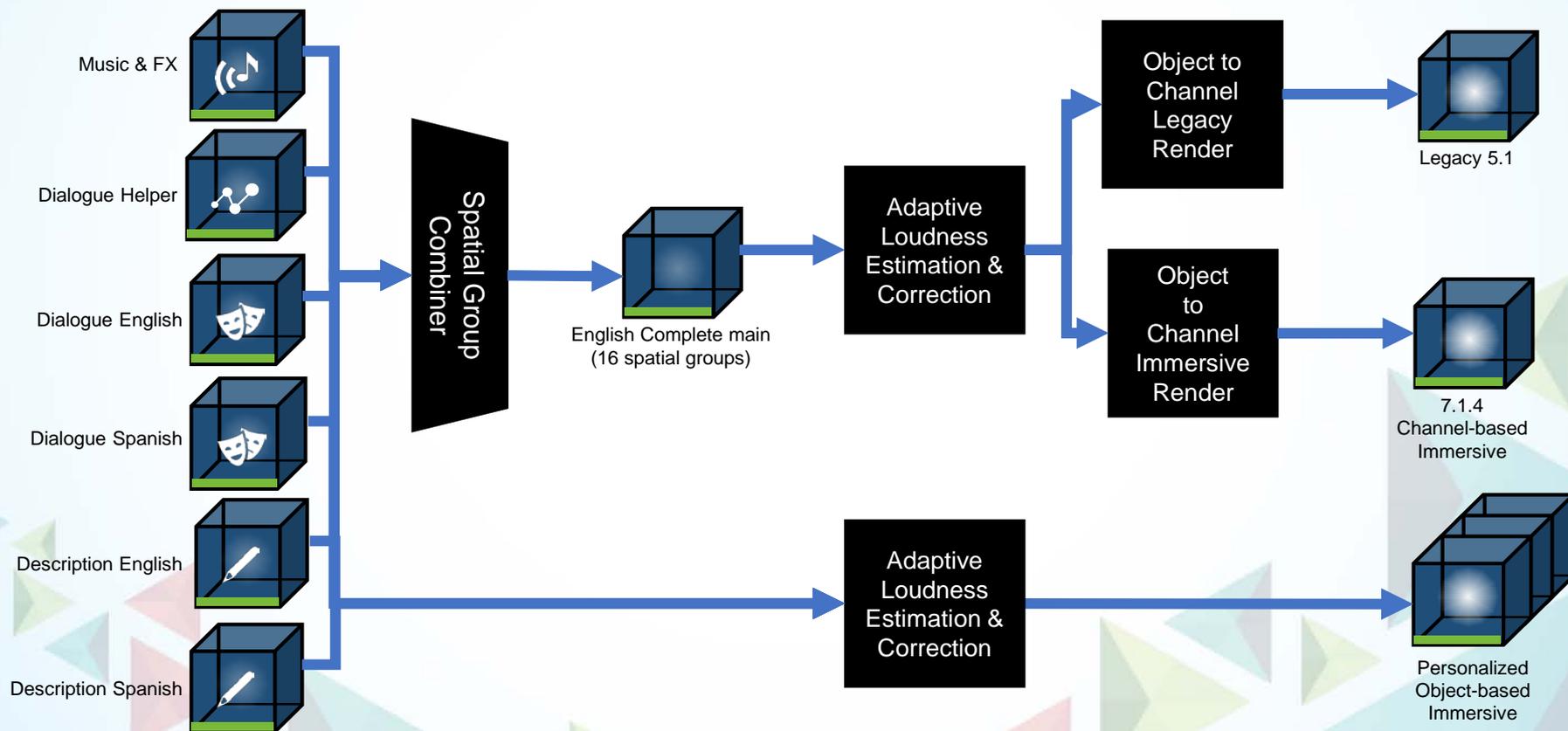
A Future Night of Broadcast Content



Dynamic Object-based Delivery



File-based Object Delivery



Broadcast and E-Media(OTT)

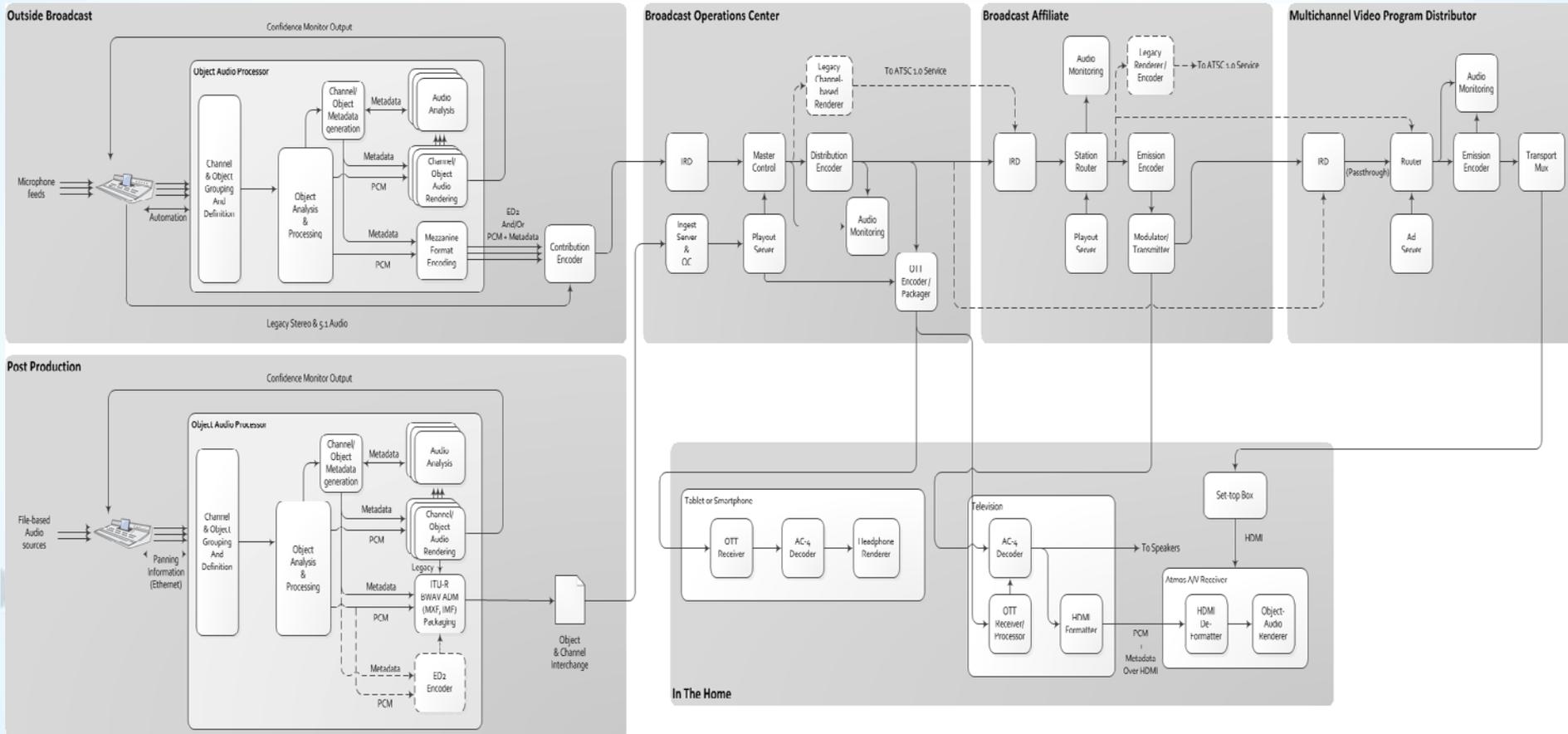
The domain of Broadcast and E-media permeates and merges somehow.

Broadcast: Cable->Satellite->on-demand

E-media: streaming podcast, Samsung Galaxy Beam, etc.



Broadcast and Streaming Delivery



Agenda

3

.....
NGA: Benefit and how

Tomorrow's needs (NGA commercial requirements)

	Requirement	NGA	NGA detail
1	Enable immersive audio	✓	Enables broadcast delivery of the industry leading immersive experiences such as movie content, and immersive audio including height and objects
2	Support for delivery of channel-based audio	✓	NGA through the use of NGA part-1 has already been included as a channel-based codec, Part2 now brings further possibilities for delivering channel based audio
3	Operating capabilities	✓	The NGA system is capable of delivering much more than the capabilities given in the CM requirements. Through informed use of the tools of NGA experiences and use cases can be delivered with fewer elements in the bitstream
4	Creation of stereo output from channel based, or immersive audio	✓	The NGA standard provides a specification for predictable stereo down-mix capability
5	Efficient support for supplementary audio	✓	NGA efficiently supports supplementary streams having the ability to seamlessly re-combine these at the decoder
6	Improved Accessibility (e.g. clean audio, audio description)	✓	NGA has been designed with efficient and comprehensive accessibility features
7	Loudness Management	✓	NGA's dual-ended loudness management enables robust compliance with all broadcast loudness regulation
8	Personalisation	✓	Advanced personalisation allows control of aspects such as volume or position of elements to be personalised
9	Binaural audio	✓	NGA can carry both pre-rendered binaural audio and in addition can enable the rendering of any NGA stream to a binaural headphone output
10	Hybrid delivery and companion device	✓	NGA contains the elements needed to allow audio frame accurate synchronisation and allows for audio presentations to be built from elements that are delivered through multiple paths, if needed NGA also contains the ability to carry the identification and sync metadata as part of the audio stream
11	Preservation and interoperability of production metadata	✓	NGA was built with Metadata in mind and, with the other elements of the audio system allows for the production metadata to flow from capture to playback
12	Codec Performance Evaluation	✓	Dolby are happy to submit NGA for codec evaluation as specified in this CR. NGA strikes the perfect balance of practicality ease of integration and performance
13	Applicability to Broadband and Mobile Delivery	✓	NGA is part of the DVB-DASH specification, and prototype tablets and mobiles can already decode NGA, the seamless splicing and sync features also make it ideal for IP
14	Timing	✓	NGA is commercially available NOW, and major manufacturers such as Sony and TP Vision have already committed to integrating NGA. Encoders are already available

NGA – Delivering on NGA requirements



Accessible

Descriptive
Audio, Dialogue
Enhancement,
Multiple
Language



Personalized

Modify the
presentation
to the
listener's
preference



Immersive

Put the
consumer in
the action
and the
venue

Lifelike
experience



Adaptable

Optimal
playback on
every device

NGA: Next Generation Audio

Efficiency



Spectrum Loss



Reach/Opex



Compliance



Accessibility



Loudness



NGA

Content goodness VS overhead to make it

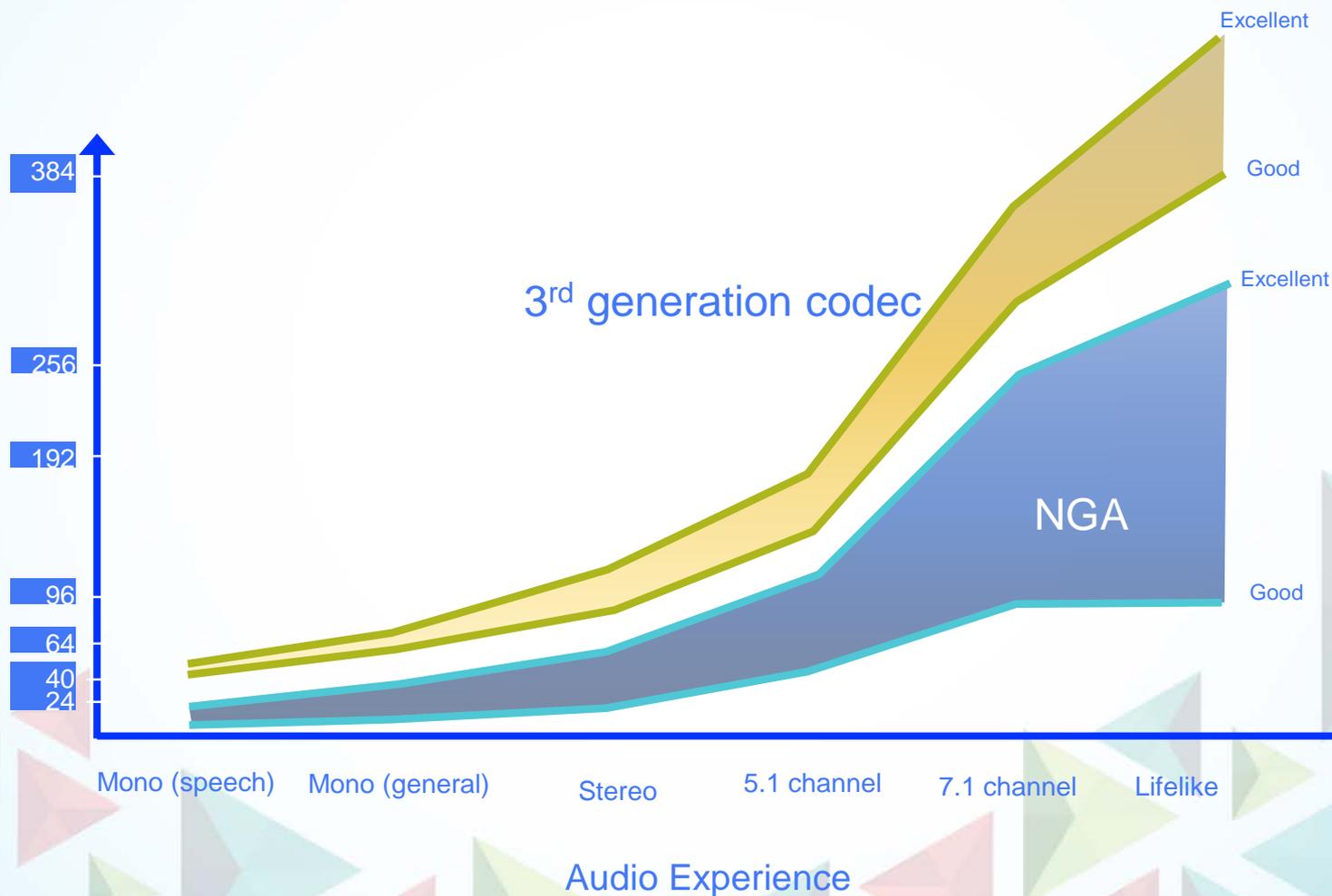


MORE MEDIA GOODNESS



HALF THE BIT-RATE

Current Audio Codec VS NGA



Legacy audio codec

Codec	Age	Efficiency	Seamless	Accessibility	Dialogue	Loudness	Metadata
MPEG1 L2	1993	-	-	*	-	-	-
AC3 (DD)	1997	-	-	-	-	*	**
AAC-LC	1999	-	-	*	-	-	*
HE-AAC	2003	**	-	*	-	-	*
E-AC3	2006	**	*	**	-	*	**

Audio Matrix comparison Codes for Video Stack

HEVC & UHD

— 音视频技术社区 —

Codec	Age	Efficiency	Seamless	Accessibility	Dialogue	Loudness	Metadata
MPEG1 L2	1993	-	-	*	-	-	-
AC3 (DD)	1997	-	-	-	-	*	**
AAC-LC	1999	-	-	*	-	-	*
HE-AAC	2003	**	-	*	-	-	*
E-AC3	2006	**	*	**	-	*	**
NGA	2015	***	***	***	***	***	***

On February 26th 2015 the DVB added a new Audio codec NGA to the specification TS 101 154, and the DVB signalling spec EN 300 468

Turns into this

Codec	Age	Efficiency	Seamless	Accessibility	Dialogue	Loudness	Metadata	Personal	Immersive	Hybrid	Adaptable
MPEG1 L2	1993	-	-	*	-	-	-	-	-	-	-
AC3 (DD)	1997	-	-	-	-	*	**	-	-	-	-
AAC-LC	1999	-	-	*	-	-	*	-	-	-	-
HE-AAC	2003	**	-	*	-	-	*	-	-	-	-
E-AC3	2006	**	*	**	-	*	**	-	**	-	**
NGA pt1 <small>(TS101 154 V2.2.1)</small>	2014	***	***	***	***	***	***	-	-	-	-
NGA pt2 MPEG-H	2015	****	***	****	****	***	****	***	***	**	***

Enable immersive audio

- Enables broadcast delivery of the industry leading immersive experiences such as movie content, and immersive audio including height and objects
 - TS 103 190-2 [2], clause 4.3, 5.2ff: Immersive CBA (7.1+4, 9.1+4, 22.2 and subsets)
 - TS 103 190-2 [2], clause 5.7, 5.9: Immersive OBA
- Dolby ATMOS as a differentiator
- MPEG-H claim HOA (Scene based)

Capabilities: MPEG-H 3D Audio LC Profile vs. NGA

NGA Level 3

Level	Sampling rate	Total number of elements (channels or objects)	Channel-Based Non-Immersive				Channel-Based Immersive	Object-Based Immersive (A-JOC coded)	Objects (discretely codec)
			Total	M&E	D	AA			
3	192kHz <i>(Max core decode 48kHz)</i>	17.1 elements <i>(Max core decode elements: 10.1 channels or 10 objects/clusters)</i>	10.1	5.1	3	2	7.1.4 <i>(Max core decode elements: 5.1.2)</i>	17.1 <i>(Max core decode elements: 5.1 static or 10 adaptive)</i>	10

MPEG-H 3D Audio LC Profile Level 3

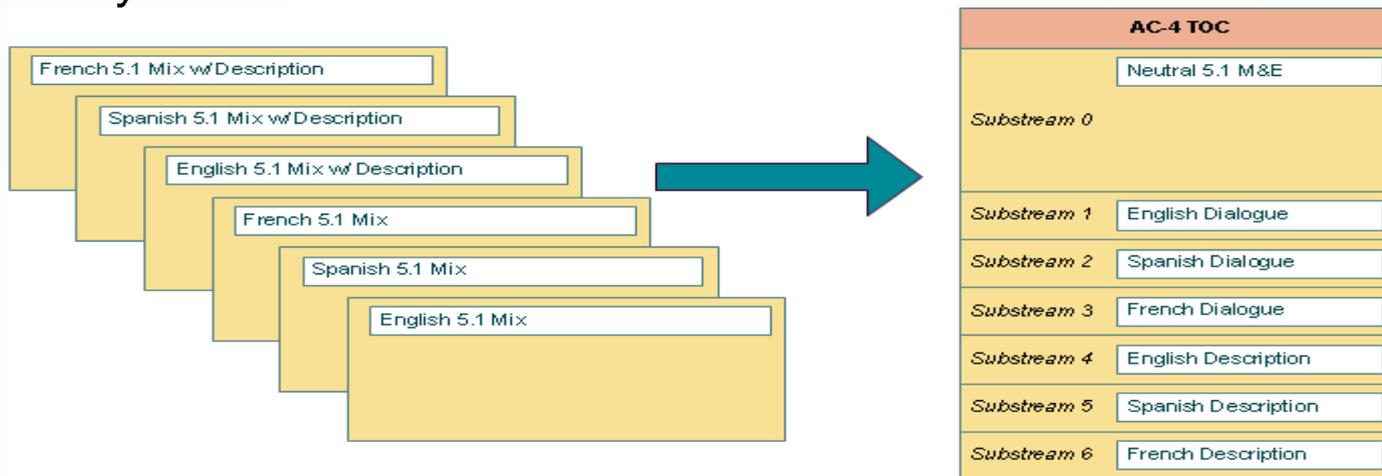
Level	Sampling rate	No of Core channels	No output channels	Max. Config C+O	Worst case dmx config	Max. HOA order+Objects	Max Objects only
3	48kHz	16	12	7.1.4 + 4 obj	12->6	6th order + 4 obj	16

Creation of stereo output from channel based, or immersive audio

- The NGA standard provides a specification for predictable and custom stereo down-mix capability
 - [1], clause 4.3.12.2, [2], clause 6.3.8 & clause 6.3.9
- Both NGA part1 and part2 and MPEG-H can do this.
- The Renderer does creep in here – More later

Efficient support for supplementary audio

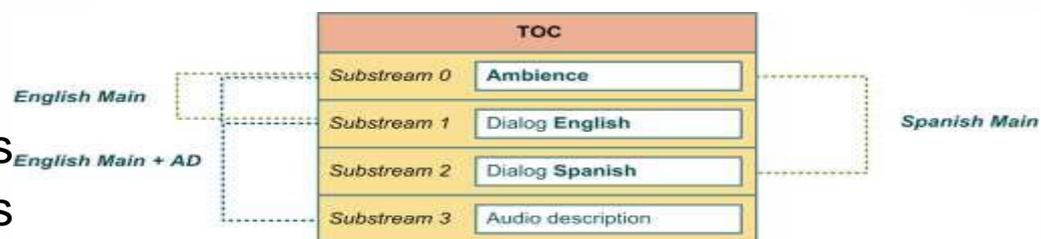
- NGA efficiently supports supplementary streams having the ability to seamlessly recombine these at the decoder



- [2], clause 4.5, [1], clause 4.3.3
- TS 101 154, clause 6.7.4

Improved accessibility

- NGA has been designed with efficient and comprehensive accessibility features

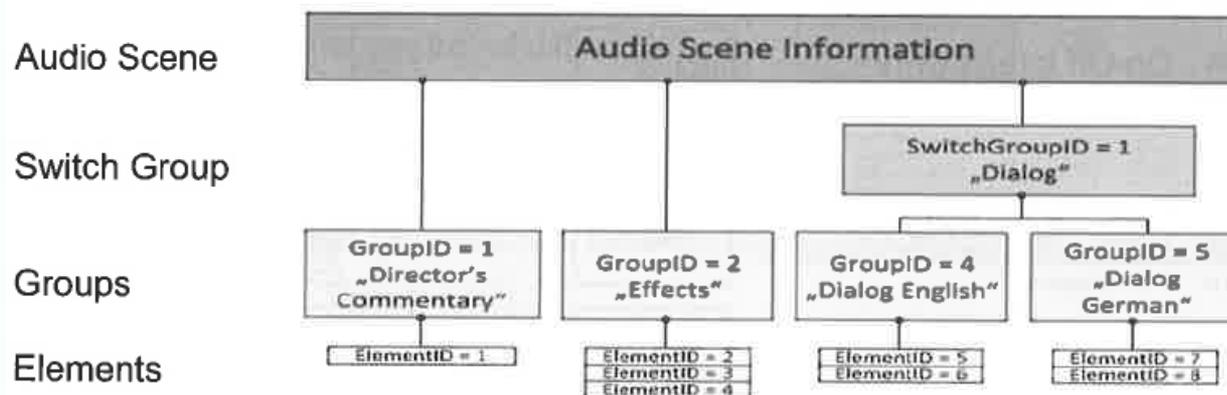


- [1], clause
- [2], clause
- TS 101 154, clause 6.7.4

- We are engaged with accessibility groups and encourage you to do so too, we have improvements to offer, and they can help us build the case for NGA in specs

MPEG-H can do it too

MPEG-H Object-based Audio Audio Scene Example



Loudness Management

- NGA's dual-ended loudness management enables robust compliance with all broadcast loudness regulation
 - [1], clauses 4.3.13, 4.3.12 and [2], clause 4.8.5, 6.3.8
 - TS 101154 clause 6.7.5
- NGA's loudness management is pretty unique.
- MPEG-H seems to be weaker although adequate here
- But no one has worked out how to do loudness for NGA yet!

- Advanced personalisation allows control of aspects such as volume or position of elements to be personalised
 - [2], clauses 4.8.3.17/18, 6.3.8.5, 6.3.9
 - TS 101154, clause 6.7.4, 6.7.6
- Personalisation is a paper battle at the moment
- NGA Level3 has limitations on the individual items that can be personalised, but is very flexible in how to carry elements and groups
- MPEG-H can in theory carry more individually personalised objects but is much more rigid in how it works
- The DVB is not expecting all elements to be personalisable

Binaural audio

- NGA can carry and signal pre-rendered binaural audio and in addition can enable the rendering of any NGA stream to a binaural headphone output
 - [1], clause 4.3.3.5
- Binaural is becoming important to broadcasters
- BBC and Orange for example are trialling services
- 2 kinds of Binaural are being discussed –
 - pre-rendered Binaural or Binaural recordings and
 - Binauralised Rendering at the device
- Not much competitive difference here, although MPEG-H makes noise about HOA

Hybrid delivery and companion device

- NGA contains the elements needed to allow audio frame accurate synchronisation and allows for audio presentations to be built from elements that are delivered through multiple paths, if needed NGA also contains the ability to carry the identification and sync metadata as part of the audio stream
 - [1] clause 6.2.15 and [2] clause 5.1.3, 5.1.2
 - The DVB have not defined a mechanism nor signalling for this yet
 - So neither system will really support it on day1
 - This however is in development at DVB SSS, but is a future feature
 - Tread carefully with any Hybrid requests and consult a product expert

Preservation and interoperability of production metadata

- NGA was built with Metadata in mind and, with the other elements of the audio system allows for the production metadata to flow from capture to playback
 - [2] clause 6.3.8, 6.3.9
 - This requirement pertains to ADM metadata
 - NGA is compatible with ADM, but does not carry it natively
 - It needs to translate it to its own format (OAMD-NGA).
 - Engineering has a mapping from one format to the other for the relevant parameters
 - MPEG-H does a similar conversion although we would argue that ours is truer to the artistic intent

Applicability to Broadband and Mobile Delivery

- NGA is part of the DVB-DASH specification, and prototype tablets and mobiles can already decode NGA, the seamless splicing and sync features also make it ideal for IP
 - [2] clause 5.11
 - TS 101154, clause 6.7.7
- Inserted to try to minimise the gap between what's used in Broadcast and Mobile
- No distinct advantage for either solution here but audio application which proves that NGA is mobile-friendly

Agenda

4



3D Spatial Audio/Soundfield Audio



3D Audio

- **MPEG-H part3: HOA**
- **Object-based ATMOS**
- **ISF**



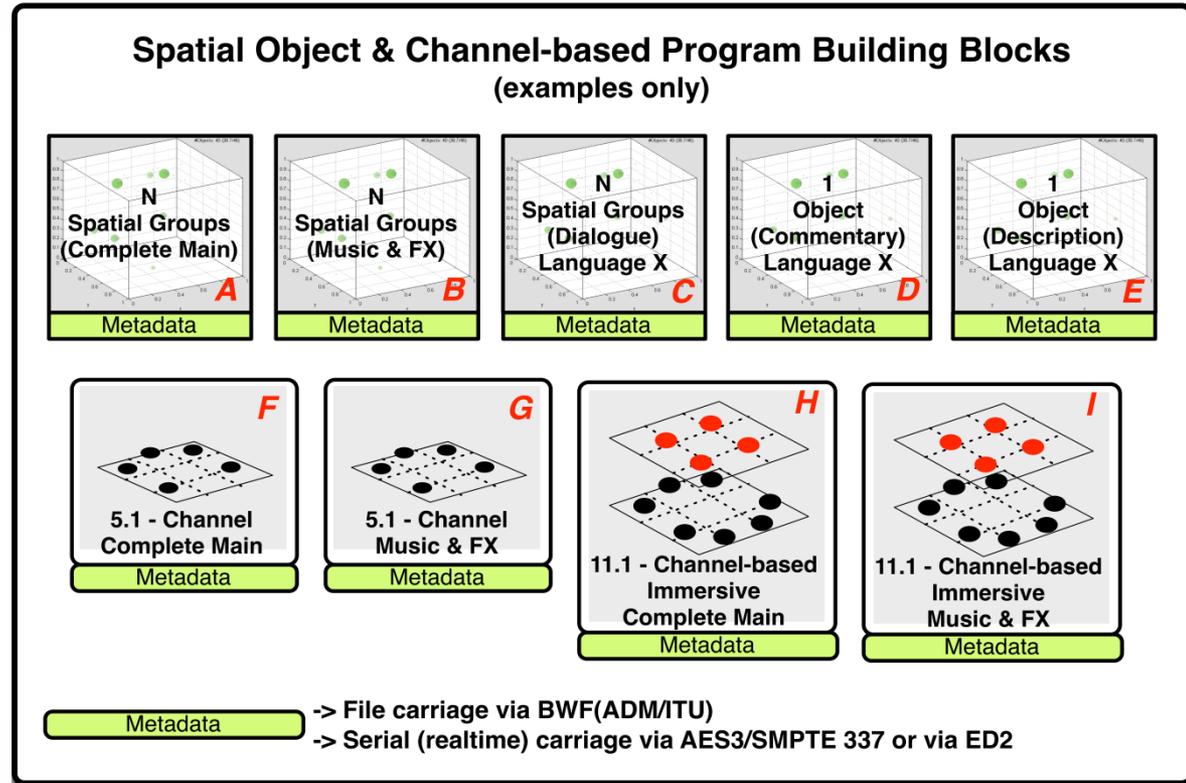
THE CONSTRAINTS OF CHANNELS



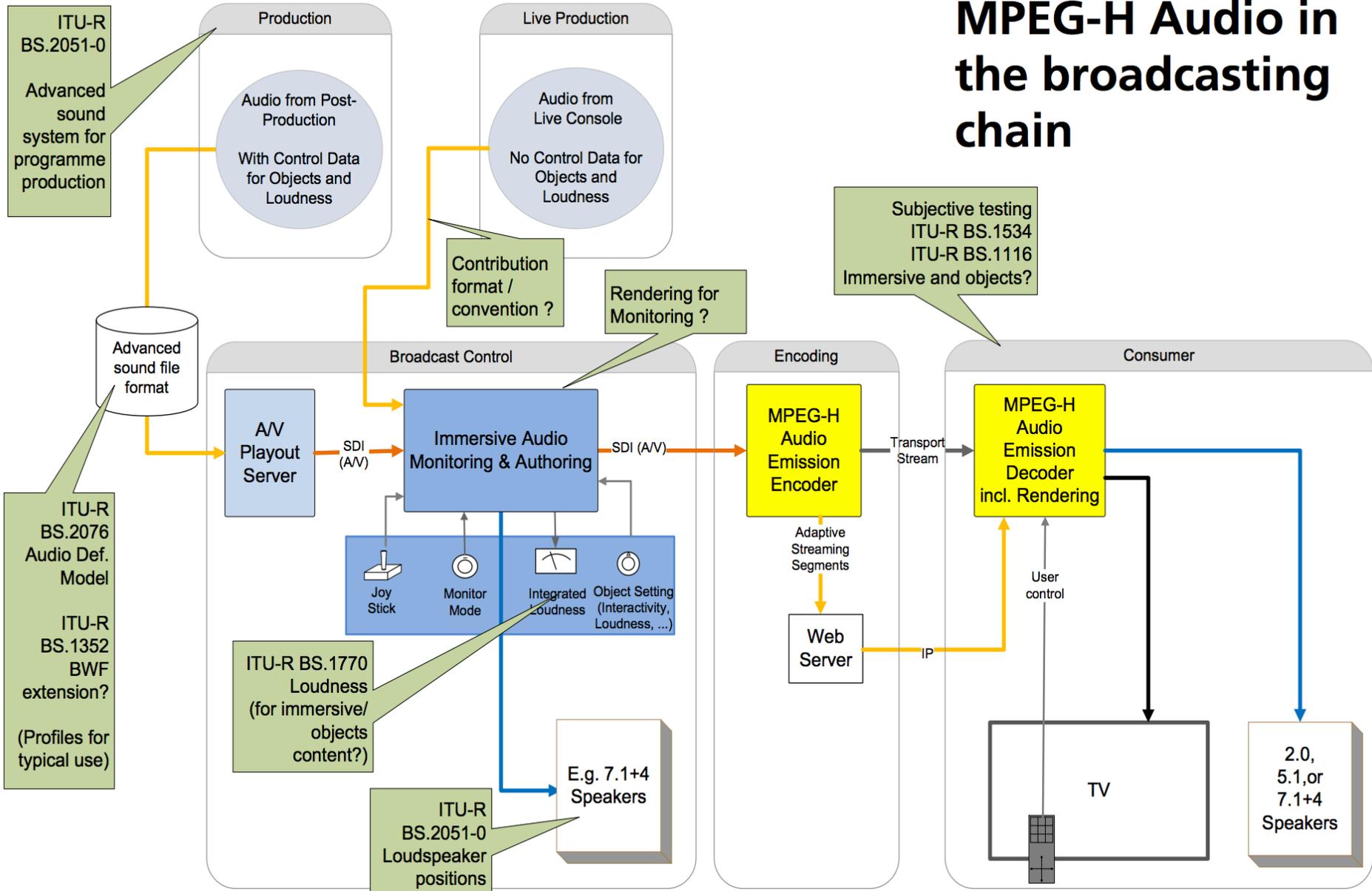


Example of Spatial Object Groups

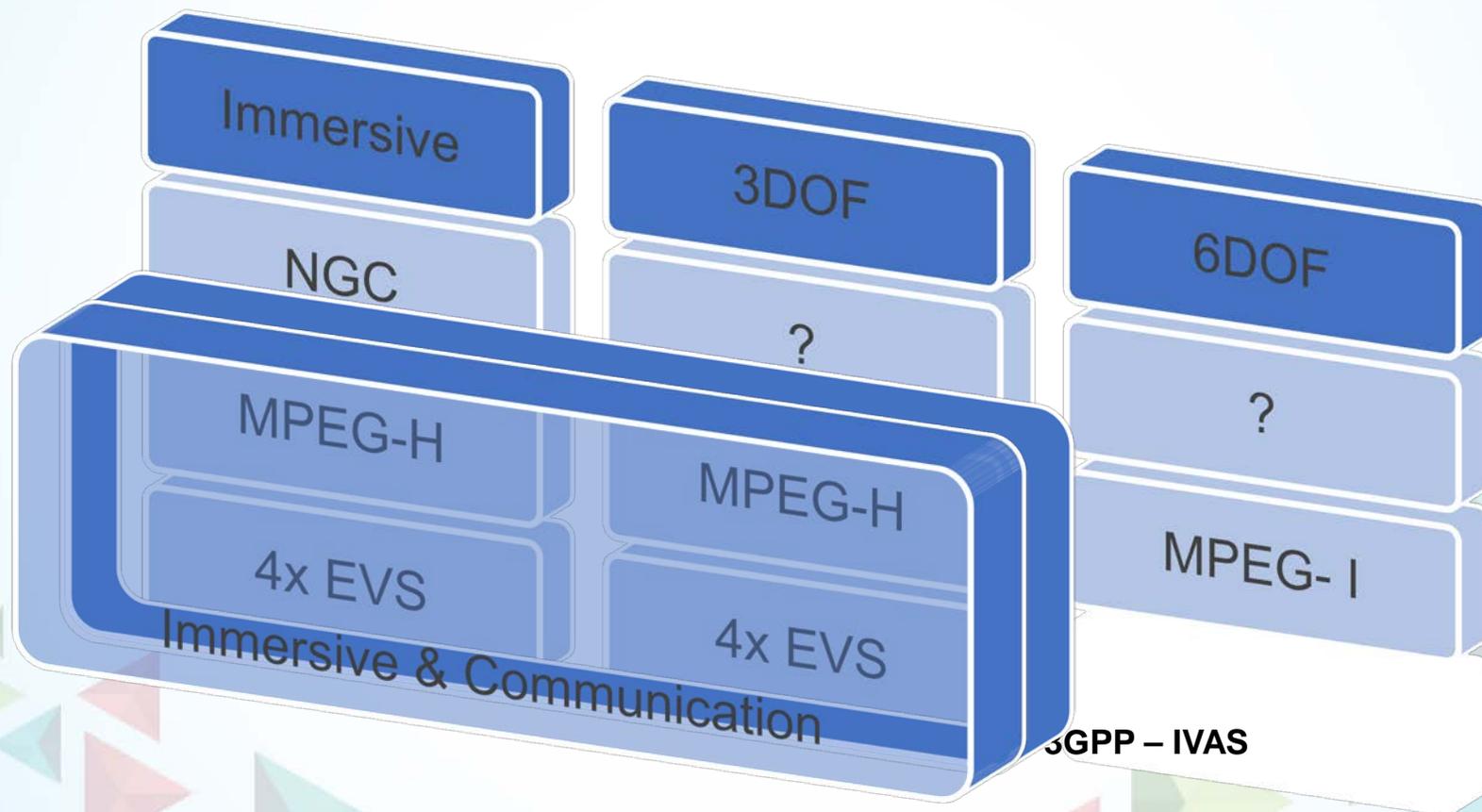
- Building blocks for program creation
- Taken from the Dolby ATSC3.0 proposal



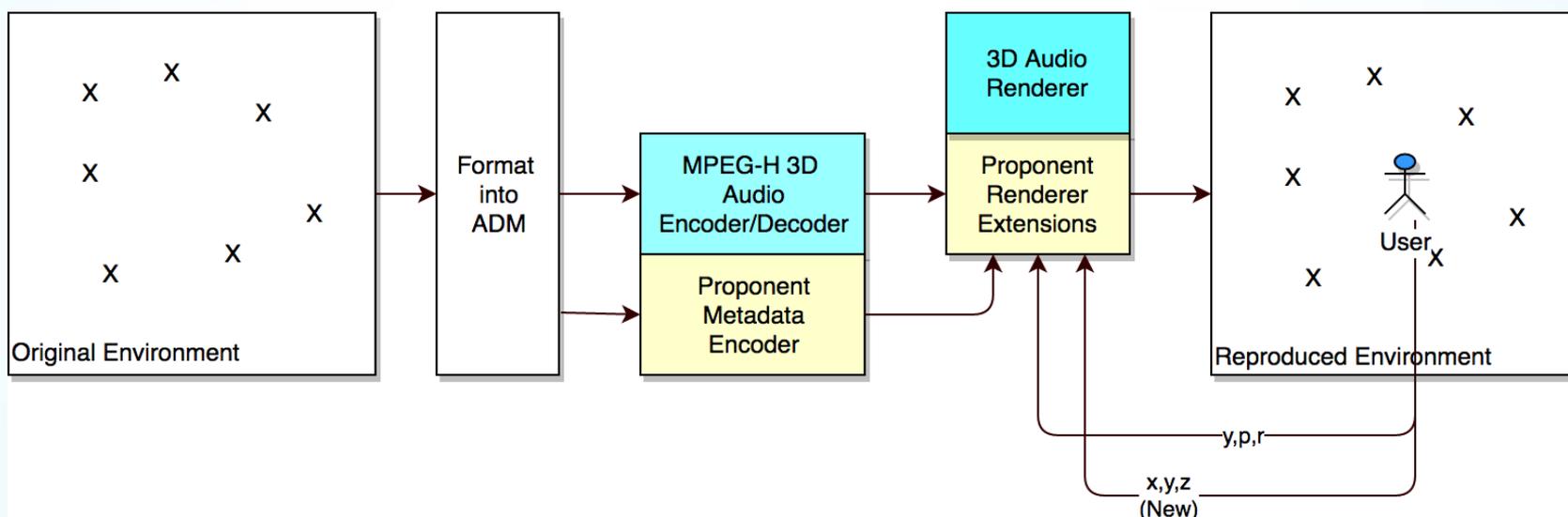
MPEG-H Audio in the broadcasting chain



Audio codec in VR



MPEG-I answers 6DOF



Thank You

