EBU

OPERATING EUROVISION AND EURORADIO

TECHNICAL REVIEW

Dynamic Metadata in an SMPTE ST 2110 environment

A proof of concept

December 2021

Adrian Hilber	SRF
Andreas Weiss	SRF
Markus Brockmann	SRF

The main purpose of an EBU Technical Review is to critically examine new technologies or developments in media production or distribution. All Technical Reviews are reviewed by 1 (or more) technical experts at the EBU or externally and by the EBU Technical Editions Manager. Responsibility for the views expressed in this article rests solely with the author(s).

To access the full collection of our Technical Reviews, please see: tech.ebu.ch/publications

If you are interested in submitting a topic for an EBU Technical Review, please contact: tech@ebu.ch

1. Introduction

SRF[1] is the national broadcaster in Switzerland, a small multilingual country surrounded by far bigger nations with a multiplicity of TV stations vying for our national audience's attention. It goes without saying that in this environment, SRF needs to continually improve its services to excite and maintain its audience.

Next Generation Audio (NGA) came to surface at the right moment for us. The discussions concerning personalized sound experiences and the ongoing aim of optimizing the use of bandwidth found their solutions in the technologies associated with Object-Based Audio (OBA).

Additionally, SRF stepped into the future of IP-based workflows in broadcast by developing Europe's first all-IP UHD OB Van (EBU Technology & Innovation Award 2019) and building a new Playout and Switching centre completely based on SMPTE ST 2110. One of the lessons learned from this innovation was the need for flexible Metadata-handling in an IP environment.

In this article we will highlight these two main issues and how they got married in a proof of concept in collaboration with Dolby Laboratories Inc.

The choices we made in our proof-of-concept were those available to us and with good support from the vendors and manufacturers concerned. Other solutions are evidently possible, but we are describing our PoC here[2].

2. Personalized Sound Experience

The basic concept of Object-Based Audio is the identification of audio elements using descriptive metadata, and indeed, an audio object is a piece of audio content together with its descriptive metadata. This simple concept at once unlocks almost unimaginable capabilities within a broadcast environment, too many to discuss here, but for our purposes, one of the biggest advantages of OBA is its ability through metadata to separate audio beds and other audio objects such as commentaries in different languages, audio description and so on.

Another part of the wonder of OBA is that once a complete "programme" of audio objects has been made, this one and the same audio environment can be archived for reuse in just about any context without the need for re-authoring.

Yet again, the metadata describing the content may be acted upon by the end-user to personalize the sound experience obtained; choosing the appropriate commentary language, audio description and background audio (audio bed) and varying the

[1] Schweizer Radio und Fernsehen. https://www.srf.ch/ .
[2] This article is also re-purposed as a use-case in the EBU Technical Review article "Practical implementation of new open standards for NGA production and interchange".

relative levels between these elements to suit. All this is fully described in other EBU publications on OBA.

With OBA we can kill two birds with one stone towards the goals of personalizing sound experiences and optimizing the use of bandwidth.

Towards this, the idea of transmitting a blockbuster movie, concert or live sports event with one multichannel audio bed and a variety of three or four different languages plus audio description (AD) seems obvious. Whereas the thought of using static metadata for fixed rendering ratios for these programme elements was not an attractive option for us, the idea of delivering them with dynamic metadata came to mind.

The solution was found by taking the mixing parameters of an audio console (or any other audio processor) and using this data to control an authoring tool generating the desired metadata. Using a standard metadata protocol (in our case, S-ADM[3]) allows the use and routing of these metadata in any standardized broadcast environment.

Open standards and protocols ensure the most interoperability between different products and as technology advances accelerate, they are becoming even more important for us as a broadcaster to embrace.

3. Our test scenario

Our test was designed to prove how dynamic metadata could be used for a sports transmission. Typically, in such a programme we have an international sound object (IS) and several commentary objects.

Each object is transported once only, with accompanying metadata, but at the receiver side we have a different representation for each commentary. The appropriate mix for the IS and the individual commentary is controlled through the metadata.

Representation 1:	IS + Commentary 1
Representation 2:	IS + Commentary 2
Representation 3:	IS + Commentary 3

This means that the IS object needs three different gains in the metadata, one for each representation, whilst each Commentary also needs a gain.

[3] S-ADM. Serial Audio Definition Model, specified in Recommendation ITU-R BS.2125-0.

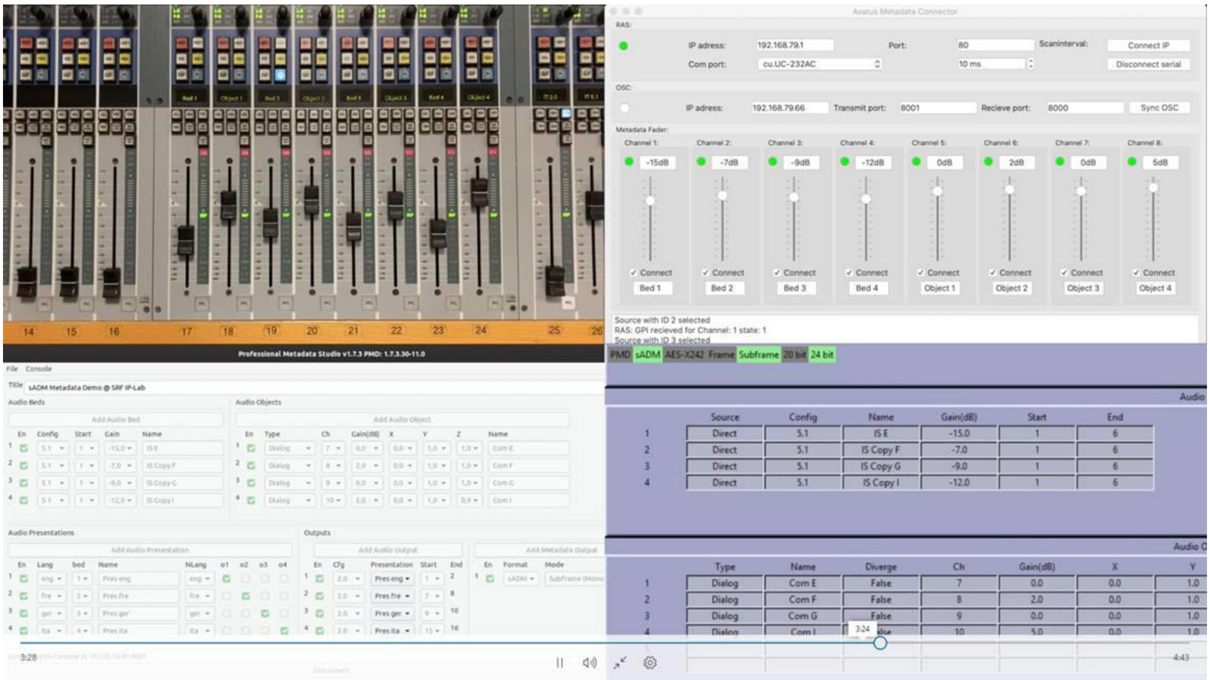

Figure 1: Top left: Faders on the Desk controlling metadata only. Top right: protocol convertor between desk and PMD Studio. Bottom left: PMD Studio with different gains for the same IS input. Bottom right: AM Viewer on the receiver side

4. Transport of dynamic metadata in a SMPTE ST 2110 environment

S-ADM is an appropriate choice for such object-based audio metadata. It is clear to us that even in a modern facility there is always the need of backward compatibility as legacy SDI infrastructure and links are often encountered. A full IP–based workflow is of course the goal.

To us as audio engineers, it is important to be able to generate metadata directly from a mixing desk. As there are nowadays so many controllers and web interfaces in a control room, having yet another UI for the metadata control is not an option to contemplate.

As is shown in Figure 2, things rapidly get complicated, even in a very simple setup. The first challenge was to get the fader positions from our console to the metadata authoring tool (Dolby PMD Studio, in our case). There are different protocols possible, but no common standard.

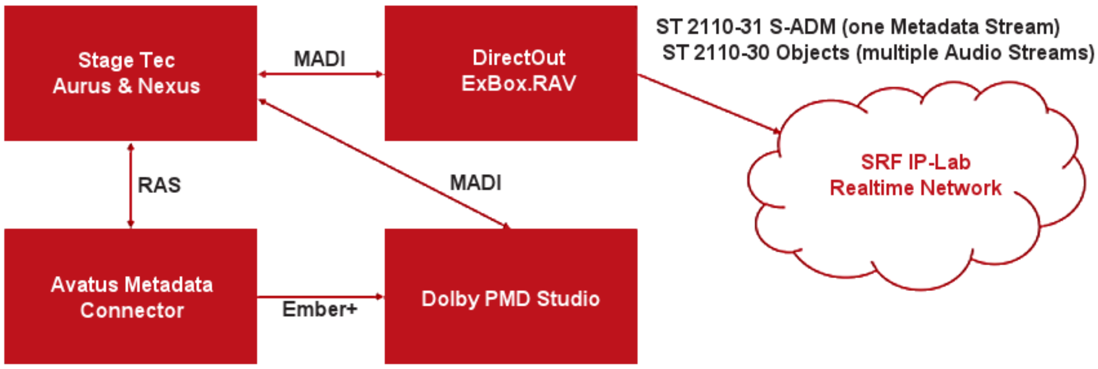

Figure 2: S-ADM setup in the audio control room

Even with the chosen Ember+ protocol, parameters vary substantially between different products and vendors, as was the case here. As a result, no direct connection was available, and we had to develop our own protocol converter to interface our console with the PMD Studio.

The generated S-ADM is carried from the PMD Studio as an AES-subframe via a transparent MADI link to an AES67 Gateway (DirectOut ExBox.RAV) and is transported as a ST 2110-31 stream through our IP test environment.

Some research is still needed in this field to standardize parameters so as to reduce complexity. In our view a good way to go is OSC[4] with an S-ADM profile. As OSC is widely used in audio gear and with a defined S-ADM profile, the vocabulary of each device would be compatible. Perhaps in local setups, a pure OSC workflow could be a viable solution as many renderers for multichannel PA setups already use OSC.

5. Signal path overview

Referring to Figure 3, this works just as expected, even if there is conversion to and from an embedded SDI link through gateways. The open-source Dolby AM-Viewer decoded the metadata stream without any problems.

The downsides of this method may be summed up as scalability and bandwidth consumption. In a modern IP infrastructure this is not an issue though, but we think it is not a smart solution to transport dynamic metadata of a few kbit/s in a ST 2110-31 stream of approximately 3 Mbit/s. Working towards a solution with a ST 2110-41 metadata stream seems an obvious next step to us.

[4] Open Sound Control, http://opensoundcontrol.org/ .

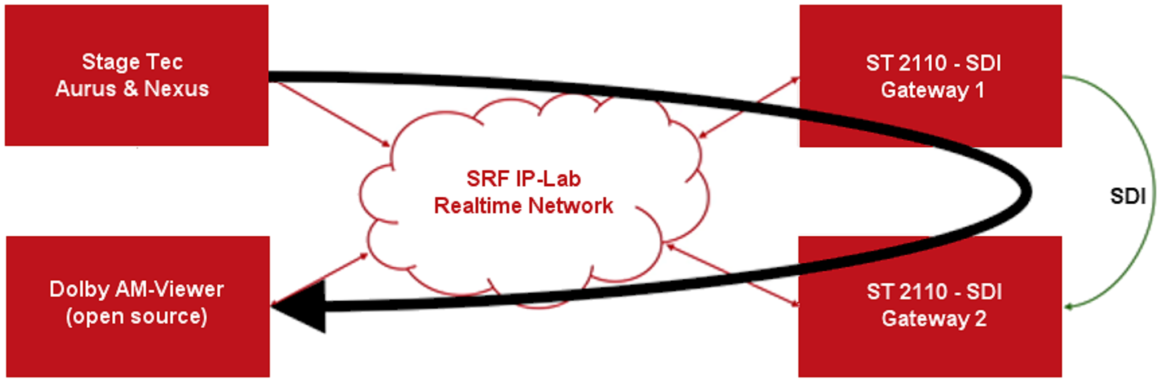

Figure 3: Signal path overview

A possible pitfall is also the fact that within the S-ADM, audio objects are only referenced by index. Whilst this is not an issue if you have a single stream containing both the audio and the metadata, in practice there will often be separate streams for metadata and the audio. In our concept we even have different streams for the audio objects, as we decided to combine what belongs together and have a separate stream per object.

This means that for a sport production we have one stream, for example in 5.1, for international sound and a mono stream for each commentary. When adding metadata to this scenario we must be very carefully when it comes to channel mapping because in the renderer the discrete audio channels must match the index within the S-ADM stream and the chances are very high that on a journey through a broadcast facility, the channels get rearranged at some point to meet (for example) a local track concept.

6. Bibliography

SMPTE ST 2110-30	*Audio transport, based on AES67*
SMPTE ST 2110-31	*Transport of AES3 formatted audio*
SMPTE ST 2110-41	*Transport of extended metadata, with definition of the manner of transporting dynamic or extended metadata in ST-2110 context.*
MADI (AES10)	*Multichannel Audio Digital Interface*
AES67	*Standard for audio applications of networks - High-performance streaming audio-over-IP interoperability*

7. Author(s) biographies

	Adrian Hilber, SRF Adrian started his career in telematics before diving into the broadcast world. He is responsible for many projects like the development of Europe’s first all-IP UHD OB-Van. He leads the IP-Lab at SRF.
	Andreas Weiss, SRF Andreas is sound supervisor for SRF. He is responsible for the broadcast mix of many outstanding concerts and live events as well as TV-Shows such as ‘Musicstar’, ‘The voice of Switzerland’ or ‘Switzerland’s got talent’.
	Markus Brockmann, SRF Markus is head of audio outside broadcast at SRF where he leads the implementation of Next Generation Audio. He is responsible for the sound of numerous productions in different genres e.g., Alpine Skiing at the winter Olympics games.

Published by the European Broadcasting Union, Geneva, Switzerland

ISSN: 1609-1469

Editor-in-Chief: Patrick Wauthier
E-mail: wauthier@ebu.ch