MPP AVoIP Live Media Standards [20xx-20xx]

601 Platform Standards

1. Contents

2. Version Control

Rev.	Date	Detail	Author(s)
0.1	09/06/22	First draft release of document for review This version of the document is concerned with flow exchange within and between NRK data centres and office locations. It does not currently describe flow exchange standards between NRK cloud and public cloud resources or third party contribution standards.	GH
0.2	13/06/22	Second draft following internal reviews and discussions around multi-channel audio production workflows.	GH / TH
0.3	14/06/22	Further ammendments following further internal review and discussions. First release to NRK.	GH / TH
0.4	22/06/22	Added support for mono flows which may (subject to future design work) be beneficial to Radio	GH

3. Introduction

Fading industry standards such as SDI brought great benefits, not least they were reliable and simple to build with, which brought a high degree of creative freedom. However, the trade-off is that they are also rigid, which, in a rapidly evolving landscape constrains agility within media organisations.

With the transition to AVoIP media organisations have greater freedom to choose the standards and pseudo standards (such as NDI) they wish to support and to embrace new standards as the emerge. However, with this freedom, some of the responsibility for standards documentation, enforcement, and interoperability testing passes from standards bodies and suppliers to media organisations.

This document details NRK's audio, video, and ancillary data standards for live media exchange within the platform. It does not describe Contribution or Distribution standards. It also does not describe NRK file-based media standards.

The aim of this document is to ensure that NRK experiences all the benefits of SDI without its constraints, and with all the benefits AVoIP brings. Benefits include:

1. **Device Interoperability -** High-level standards support does not guarantee interoperability. For example, whilst a device may support ST 2110, ST 2110 offers many options for audio and video transport and support for specific features varies greatly between products.
2. **Improved business agility -** Improved interoperability makes it quicker and easier to create new workflows without the need for heavy research and development processes.
3. **More effective project delivery -** Product failure to comply with these standards can be exposed early. Products can either be rejected or agreed enhancements placed on a roadmap and managed.
4. Reduces system complexity
5. **Less physical equipment** required for signal exchange
6. Significantly reduces control system complexity
7. Minimise format conversions – benefits include:
8. **Cost savings** – from needing less equipment
9. **Maximise quality** – by minimising the need for conversions
10. **Minimise latency** – by minimising the need for conversions
11. Increases system reliability
12. Maximise operational efficiency
13. **Minimise manual interventions** to convert audio and video signals.
14. **Simplifies multi-format working** - when working with multiple video resolutions, framerates, colour formats, and audio formats.
15. Reduces effort spent reinventing basic system constructs

Whilst conformance to these standards is mandatory, this is an active document. Suggestions for updates is encouraged where it brings value to NRK and increases NRK's creative freedom.

4. Normative References

This document makes reference to the standards and quasi-standards detailed below.

Please note the control standards such as the AMWA NMOS suite are specifically exlcuded from this document.

4.1. Standards

- AES67: High-performance streaming audio-over-IP interoperability
- AES-r16-2016: PTP parameters for AES67 and SMPTE ST 2059-2 interoperability
- ITU BT.709: Parameter values for the HDTV standards for production and international programme exchange
- ITU BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange
- ITU BT.2100: Image parameter values for high dynamic range television for use in production and international programme exchange
- SMPTE ST 2110-10: System Timing and Definitions
- SMPTE ST 2110-20: Uncompressed Active Video
- SMPTE ST 2110-21: Traffic Shaping Uncompressed Video
- SMPTE ST 2110-22: Constant Bit-Rate Compressed Video
- SMPTE ST 2110-30: PCM Digital Audio
- SMPTE ST 2110-31: AES3 Transparent Transport
- SMPTE ST 2110-40: SMPTE ST 291-1 Ancillary Data
- SMPTE ST 2110-41: Fast Metadata [once standard is ratified]
- SMPTE ST 2022-1: Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks
- SMPTE ST 2022-2: Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks
- SMPTE ST 2022-7: Seamless Protection Switching of SMPTE ST 2022 IP Datagrams
- VSF TR-08: Transport of JPEG XS Video in ST 2110-22

4.2. Quasi Standards

- Audinate: Dante
- Viz RT: NDI

5. Video

5.1. Formats

Table 1 summarises supported video formats within the platform.

The two main house production formats are highlighted in gold text. Transitioning to and standardising on 1080p50 and 2160p50 WCG HDR HLG will simplify platform design and operations, as well as increase picture quality, by avoiding unnecessary colour space, framerate, and scan conversions. It will also reduce the number of format variations that operations need to manage.

720p50 is included because Playout hands off to some distributors in this format but is not used. It's possible that support for this format may not be required in future.

1080p25 HDR is inlcuded for multiviewer heads. Running these at 1080p50 will consume twice as much WAN bandwidth for minimal or zero production quality gain.

#	Wide	High	Scan	FR	Ratio	Gamut	Range	Notes
1	1280	720	p	50	16:9	Rec 709	SDR	Playout distribution handoff
2	1920	1080	i	25	16:9	Rec 709	SDR	
3	1920	1080	p	50	16:9	Rec 709	SDR	
4	1920	1080	p	25	16:9	Rec 2100	HDR_HLG	Used for multiviewer heads
5	1920	1080	p	50	16:9	Rec 2100	HDR_HLG	House production format
6	3840	2160	p	50	16:9	Rec 2100	HDR_HLG	House production format

Table 1 - Video Formats

5.2. Transports

The following transports are supported within the platform:

1. SMPTE ST 2110-20 (uncompressed)
2. SMPTE ST 2110-22 (JPEG XS compressed)
3. NDI (SHQ)
4. SMPTE 2022-2 Transport Streams

5.2.1. SMPTE ST 2110-20 & -22

All flows shall be compressed using the JPEG-XS codec with the codestream encapsulated with an ST 2110-22 transport.
Two quality levels shall be supported:

1. **Full** – 6:1 compression - Used for content production and quality monitoring
2. **High** – 8:1 compression - Used for monitoring, where pristine picture quality isn't required, e.g. the majority of multiviewers.

JPEG-XS profiles for different resolution and frames rate combinations shall adhere to the interoperability points defined within VSF TR-08:2022:

https://vsf.tv/download/technical_recommendations/VSF_TR-08_2022-04-20.pdf

***Note 8/6/2022 -** Platform adoption of intoPIX's "Flawless Imaging Profile" for JPEG-XS is currently under investigation.

The rationale behind adopting compressed can be summarised as follows:

Whilst WAN costs have continued to fall, the bandwidth requirements necessary to transport uncompressed live media flows media flows for datacentres to offices is so great that compression is necessary. JPEG-XS is the ideal codec for this. It's fast (just a few lines of latency), high-quality, low power, and simple to implement. Lower bandwidth flows can also be routed across un-orchestrated networks, allowing data only carriers to bid on connectivity and reducing connectivity costs. It does mean however that a pair of codecs is required for every multiviewer and quality monitor in a facility.

Furthermore, supporting uncompressed 2160p50 production and playout requires provisioning a lot of 100Gb network ports, at great cost. Using JPEG-XS compression is a practical way to reduce the network bandwidth requirements, allowing 10Gb and 25Gb ports to be used for 2160p50 production.

Once the decision to compress 2160p50 flows is made along with the necessary production selections and licence investments, it's reasonable to apply compression to all flows. This allows 1Gb and 10Gb network ports to be used for transport of these flows, which reduces networking costs.

In addition to network cost savings within datacentres, once all flows are JPEG-XS encoded, various benefits can be realised:

1. The need for JPEG-XS codecs between the DCs and offices goes away, reducing platform cost and complexity, and increasing platform reliability.

2. Confidence – What operators see is the same codestream that is in the datacentres, not a compressed representation of datacentre side uncompressed flows.
3. Hybrid cloud – Having a single codestream for private and public cloud becomes a reality and the need for transcode at datacentre boundaries disappears. Reducing platform complexity, latency and cost, and increasing picture quality.
4. Remote production – Opens the possibility of not having to recode flows from remote events. Reducing platform complexity, latency and cost, and increasing picture quality.

5.2.2. NDI

A dedicated NDI VRF is configured on NRK AVoIP networks for NDI traffic.

To minimise coding latency and maximise picture quality, full bandwidth NDI (~100Mbps) shall be used for on premises production and, where adequate network bandwidth exists, remote production scenarios.

NDI uses a proprietary codec call Speed HQ (SHQ), which offers various profiles:

- SHQ0/SHQ1 are 4:2:0 with/without alpha
- SHQ2/SHQ3 are 4:2:2 with/without alpha
- SHQ4/SHQ5 are 4:4:4 with/without alpha
- SHQ7 is 422 with alpha coded the same way as luma
- SHQ9 is 444 with alpha coded the same way as luma

A combination of SHQ 2 and SHQ 7 profiles will be used such that alpha is encoded in the same way as luma. This configuration has become the industry norm for production workflows.

For bandwidth constrained workflows, NDI HX2 shall be used. HX2 offers more coding options than HX as well as equivalent functionality to full bandwidth NDI.

- 1080p50 workflows shall use the h.264 codec configured as follows: details TBD [profile, bitrate, CBR/ABR, GOP size]
- 2160p50 workflows shall use the h.265 codec configured as follows: details TBD [profile, bitrate, CBR/ABR, GOP size]

5.2.3. SMPTE ST 2022-2 Transport Streams

Two dedicated ST 2022-2 VRFs are configured on NRK AVoIP networks to accommodate 2022-2 Transport Stream traffic which is 2022-7 protected.

The VRF is not orchestrated. Uplink bandwidth is assigned on a case-by-case basis. For example, leaves hosting playout and distribution devices will have a 2022-2 reservation, whereas leaves hosting vision mixers will not.

All ST 2022-2 transport streams must pass EBU/ETSI ETR-290 health checks.

https://www.etsi.org/deliver/etsi_etr/200_299/290/01_60/etr_290e01p.pdf

Multiple TS payload profiles are allowable within the platform. Any new profiles must be captured within this document.
Currently only one profile has been defined.

5.2.3.1. Playout and Distribution TS Format

22/6 Note: This is a template definition which is subject to review confirmation by Playout and Distribution projects

Type	Attribute	Value
Video	Resolution	1920x1080
	Encoding	h.264
	Profile/Level	High@4.0
	Bitrate	35Mbps
	Type	CBR
	Frame Rate	50
	Scan Type	Progressive
	Mode	TFF
	GOP	50
	Colour	4:2:2
	Pixel	1:1
Audio	Encoding	AAC-LC
	Bitrate	256Kbps
	Mode	CBR
	Sample	48KHz
Data	Type 1	SCTE-35 splice_insert & time_signal Examples 1. COM - commercials - splice in/out 2. LIVE - Live - start/stop 3. PROGRAM - Program - start/ID/stop.
	Type 2	CEA 708 Captions
	Type 3	etc.
Transport	Type	ST 2022-2 MPEG-2 TS
	Protection	ST 2022-7
	FEC	If required

Table 2 - Playout & Distribution TS Format

6. Audio

6.1. Formats and Quantities of Channels per Flow

Audio shall be captured exchanged with a sample frequency of 48kHz and sample accuracy of 24bits.
All production audio within the platform shall be exchanged as uncompressed PCM.
Compressed audio shall be decompressed on ingress to the facility.

Audio shall be exchanged within the platform as 2, 8 or 16 channels flows.

Audio channel assignments shall adhere to the mappings defined in Table 3.

13/6 Note: Surround configurations are proposed and based on notional audio workflows which have not yet been designed. Left and right stereo channels are placed in the same flow as the surround channels to ensure phase coherence, which is unlikely to be possible to achieve using discreete flows. For example, if there's a problem with the decoded Dolby audio of an incoming feed, a sound desk operator can switch cleanly to the stereo mix. Contributions arriving with multichannel audio may need to pass through processing to generate a 16 channel composite flow and two 2 channel stereo to service non-surround/Atmos operations.

Channels per flow	Description	Order of channels in group	Notes
1	Mono	Mono	To support radio workflows
2	mono	Mono, Silent	
2	dual mono	M1, M2	
2	stereo	Left, Right	
8	5.1 surround	LSt, RSt, Lf, Rf, C, LFE, Ls, Rs	Proposed
16	Mixed & Effects	Mix: LSt, RSt, Effects: LSt, RSt, Mix: Lf, Rf, C, LFE, Ls, Rs Effects: Lf, Rf, C, LFE, Ls, Rs	Proposed
16	Atmos e.g. 7.1.4	L, R, C, LFE, Lss, Rss, Lrs, Rrs Ltf, Rtf, Ltr, Rtr, silent, silent, silent, silent	Proposed. Silent channels can be used for commentary objects.

Table 3 - Audio flow channel assignments

6.2. Transports

The following transports are supported within the platform:

1. SMPTE ST 2110-30
2. SMPTE ST 2110-31

3. Dante
4. NDI

6.2.1. SMPTE ST 2110-30 & 31 Conformance Levels

Table 3 summarises supported ST 2110-30 and -31 profiles used to carry audio data.

Whilst Conformance Level A, 1ms packet timing offers the widest compatibility with AES67 devices, use of 125uS packet timing has been mandated across all channel counts to avoid integration issues. Specifically, not all Receivers are able to concurrently receive multiple audio flows with different packet timings.

#	Level	channels	Timing	Freq	Bits	Notes
ST2110-30 Conformance Levels						
1	C	2	125us	48kHz	24	
2	C	8	125us	48kHz	24	To support 5.1 surround
3	C	16	125us	48kHz	24	To support 16ch hand-off & future Atmos workflows
ST2110-31 Conformance Levels						
4	C	2xAES	125us	48kHz	24	To support coded audio (Dolby ED2 & Dolby Atmos)

Table 4 - Permitted ST 2110-30 & -31 audio formats

6.2.2. NDI

To minimise coding latency and maximise audio quality, full bandwidth NDI (~100Mbps) shall be used for on premises production and, where adequate network bandwidth exists, remote production scenarios. Audio shall be uncompressed.

For bandwidth constrained workflows, NDI HX2 shall be used. NDI HX2 encoders shall be configured to encode audio using the AAC codec, configured at = 128Kbps for stereo pair

7. Ancillary Data

7.1. Contribution

Currently no requirements have been identified for the carriage of ancillary data.

7.2. Playout

Currently no requirements have been identified for the carriage of ancillary data. Within playout, no data is carried in SDI, it's all inserted at the distribution compression stage.

14/6 – Note: Live multi-channel audio workflows may yet require SMPTE ST 2020 metadata, for example to signal the presence/absence of ST 2020 to indicate that a production is in stereo/ 5.1 /Atmos.