IPMX PCM Digital Audio Profile Requirements

AIMS
Alliance for IP Media Solutions

Version 1.0 (December 2025)

Published by the Alliance for IP Media Solutions (AIMS)
https://aimsalliance.org
https://ipmx.io

---

1. Introduction (Informative)

This document defines the requirements for the **IPMX PCM Digital Audio Profile**, establishing consistent, interoperable behavior across IPMX devices that implement this media format.

This profile builds upon:

- The IPMX baseline requirements, including NMOS requirements, branding, and certification rules found in the IPMX Product Qualification and Certification Requirements (PQCR) document.
- References the applicable TR-10 specifications
- Defines the specific PCM format combinations required for IPMX compliance
- Specifies sender and receiver requirements particular to PCM audio
- Defines NMOS signaling expectations needed for interoperability

This document does *not* restate TR-10-3 or any of its second-level references.

Where helpful, short informative explanations are included. All uses of shall, should, and may follow [RFC 2119](#) conventions and form the normative portion of the profile.

2. Scope (Normative)

This profile defines:

- The required PCM digital audio formats for IPMX devices implementing this profile.
- Sender-specific requirements.
- Receiver-specific requirements.
- Optional features that vendors may implement.
- NMOS signaling requirements specific to this profile.

Anything not defined here or in TR-10-3 is outside this profile's scope.

3. Normative References (Normative)

Devices conforming to this profile shall comply with:

- **[IPMX PQCR](#)**: IPMX Baseline Requirements, including NMOS requirements, branding, and certification rules.

- **VSF TR-10-1:** System Timing and Definitions
- **VSF TR-10-3:** Internet Protocol Media Experience (IPMX) — PCM Digital Audio.
- **VSF TR-10-8**: NMOS Requirements, referencing **NMOS IS-04** Discovery and Registration and **AMWA NMOS IS-05** Connection Management. TR-10-8 defines the NMOS baseline for IPMX; PQCR extends this with additional mandatory NMOS specifications.
- **AMWA NMOS IS-11:** NMOS Stream Compatibility Management
- **AMWA BCP-004-01:** NMOS Receiver Capabilities (required for devices that implement a Receiver)
- **AMWA BCP-004-02:** NMOS Sender Capabilities (required for devices that implement a Sender)

In addition, where implemented, devices shall comply with the following AMWA specifications, as already referenced by PQCR:

- **AMWA BCP-005-01:** EDID to Receiver Capabilities (where applicable)

Where this profile refers to NMOS parameter names or capability identifiers, it uses the NMOS Parameter Registers:

- **AMWA NMOS Parameter Registers – Formats**
- **AMWA NMOS Parameter Registers – Capabilitie**s

Optionally, this profile may work in conjunction with(see Section 8):

- **IS-08**: Audio Channel Mapping

This profile does not redefine concepts such as packet pacing, network timing, SDP structures, or Info Blocks already defined elsewhere.

4. Terms and Definitions (Informative)

The terms used in this document are as defined in:

- VSF TR-10-1
- VSF TR-10-3

This profile introduces no additional terms.

5. Profile Essence Summary (Informative)

Parameter	Requirement
Essence Type	PCM Digital Audio
Baseline Sample Rate	48 kHz
Required Encodings	L16 and L24
Required Packet Time	1 ms (for at least one supported combination)
Required Channel Counts	1, 2, and 8 channels

6. Profile Format Requirements (Normative)

The media essence for this profile shall be **PCM Digital Audio** as defined in **TR-10-3**.

This profile introduces no essence-level constraints beyond those specified below.

6.1 Required Sender Behavior

A device declaring support as a sender shall implement all requirements in this section.

6.1.1 Required Sender Format Combinations

1. The sender shall encode PCM audio in accordance with TR-10-3.

2. At minimum, the sender shall support 48 kHz operation using at least one of the following TR-10-3-defined encodings:
    ○ L24 (24-bit linear PCM)
    ○ L16 (16-bit linear PCM)
3. For the required 48 kHz format, the sender shall support 1 ms packet time for at least one supported channel configuration.
4. The sender shall support at least one of the following channel counts (48 kHz):
    ○ 1 channel
    ○ 2 channels
    ○ 8 channels

6.1.2 Additional Sender Format Requirements

1. The sender shall meet the Sender Conformance Level requirements defined in TR-10-3, including:
    ○ Producing streams conformant with ST 2110-30 Level A (as required by TR-10-3).
2. Additional channel counts, packet times, or encoding variations defined in TR-10-3 may be implemented but are not required by this profile.
3. The sender shall correctly signal channel order, channel count, encoding, and timing as defined in TR-10-3.

6.1.3 Transport & Timing Requirements

1. The sender shall use RTP transport and system timing as defined in TR-10-1 and TR-10-3.
2. All packet pacing and network timing behavior shall follow the requirements of TR-10-1 (e.g., RTP timestamp behavior, media clock vs RTP clock relationships).

6.1.4 Required Sender Signaling

The sender shall populate all required PCM Audio Media Info Block fields as defined in TR-10-3, including:

● Sample rate
● Sample size
● Channel count
● Packet time

- Measured sample rate
- Channel order

The sender shall also populate the associated SDP fields required by TR-10-3, ensuring all parameters are consistent between the Info Block, SDP, and actual stream behavior.

6.2 Required Receiver Behavior

A device declaring support as a receiver shall implement all requirements in this section.

6.2.1 Required Receiver Format Combinations

1. The receiver shall decode PCM audio in accordance with **TR-10-3**.
2. At 48 kHz, the receiver shall support both TR-10-3-defined encodings:
   - L16
   - L24
3. The receiver shall support 1 ms packet time for each required encoding and channel combination listed below.

6.2.2 Additional Receiver Format Requirements

1. At 48 kHz, the receiver shall support all of the following channel counts for both L16 and L24:
   - 1 channel
   - 2 channels
   - 8 channels
2. The receiver shall correctly interpret all conveyed TR-10-3 parameters, including:
   - Channel count
   - Channel order
   - Sample size
   - Packet time
   - Media clock vs RTP clock behavior
3. Additional channel counts or packet times may be implemented but must be accurately advertised through NMOS and TR-10-3 signaling.

6.2.3 Interpretation of Transport & Signaling

1. The receiver shall interpret RTP timing, packet structure, and transport behavior as specified in TR-10-3 and TR-10-1.
2. The receiver shall correctly interpret the PCM Audio Media Info Block fields defined in TR-10-3.
3. Receivers shall follow TR-10-3 behavior for reserved fields; this profile does not redefine them.

7. NMOS Representation (Normative)

Devices claiming support for this profile shall model and advertise their media resources through NMOS in a manner consistent with:

- All required NMOS specifications listed in Section 3 (including IS-04, IS-05, IS-11, BCP-004-01/02, and BCP-005-01 where applicable)
- TR-10-8 (the NMOS baseline for IPMX)
- The required format combinations and signaling defined in Section 6 of this document

This section defines only the profile-specific NMOS constraints and expectations.

7.1 Sender NMOS Representation

A device acting as a sender for this profile shall model its NMOS resources as follows:

7.1.1 Flow Format & Media Type

1. The Flow's format attribute shall be:
   - `urn:x-nmos:format:audio`
2. The Flow's media type (encoding) shall correctly indicate the PCM format in use (e.g., `audio/L24` or `audio/L16`) consistent with TR-10-3 signaling.

7.1.2 Transport Requirements

1. Sender resources shall use an RTP-based NMOS transport URN (e.g., `urn:x-nmos:transport:rtp.mcast`), as required by TR-10-8.
2. Transport parameters exposed through NMOS shall match the sender's actual TR-10-3 behavior, including:
   - Sample rate
   - Channel count

- ○ Bit depth
- ○ Packet time

7.1.3 Sender Capabilities (BCP-004-02)

1. The sender shall expose Sender Capabilities accurately describing the PCM format combinations supported for this profile, including at least one of:
    - ○ 48 kHz, L24, with 1, 2, or 8 channels
    - ○ 48 kHz, L16, with 1, 2, or 8 channels (if implemented)
2. Parameter Register identifiers shall be used where applicable (e.g., sample rate, bit depth, channel count).

7.1.4 IS-11 Behavior

1. IS-11 resources for senders shall reflect the actual combinations supported under Section 6.
2. IS-11 signaling shall not contradict TR-10-3 parameters or the combinations required by this profile.

7.2 Receiver NMOS Representation

A device acting as a receiver for this profile shall model its NMOS resources as follows:

7.2.1 Receiver Format Support

1. Receiver resources shall indicate support for urn:x-nmos:format:audio.
2. Receiver media type declarations shall reflect support for **48 kHz L24 and L16** PCM audio.

7.2.2 Receiver Capabilities (BCP-004-01)

1. A receiver shall expose Constraint Sets permitting all required combinations defined in Section 6, specifically:
    - ○ 1-channel L16 and L24
    - ○ 2-channel L16 and L24
    - ○ 8-channel L16 and L24
2. Constraint Sets shall use Parameter Register identifiers for:

- ○ Sample rate
- ○ Bit depth
- ○ Channel count

7.2.3 IS-11 Behavior

1. IS-11 compatibility resources shall describe the receiver's true decoding capabilities.
2. IS-11 data shall not restrict or redefine any combination that is mandatory under Section 6.

7.2.4 Alignment With TR-10-3 Signaling

1. All NMOS properties including sample rate, packet time, bit depth, and channel count shall be consistent with:
   - ○ the PCM Digital Audio Media Info Block, and
   - ○ the SDP signaling required by TR-10-3.

8. Optional Features (Normative)

The following feature is optional for devices implementing this profile. Where implemented, the requirements stated in this section shall apply.

8.1 IS-08: Audio Channel Mapping

8.1.1 Normative Requirements (If Implemented)

1. A device that implements **AMWA IS-08** shall expose an IS-08 mapping resource consistent with:
   - ○ its actual channel counts and PCM format combinations defined in Section 6
   - ○ the channel order signaled through the PCM Digital Audio Media Info Block
   - ○ the active NMOS Connection API (IS-05) transport parameters
2. IS-08 mappings shall not contradict any required channel-count combinations of this profile (1, 2, and 8 channels at 48 kHz).
3. IS-08 shall use the NMOS Parameter Register identifiers when expressing channel mapping constraints or capabilities.

4. Devices implementing both IS-08 and EDID-derived behavior (via BCP-005-01) shall ensure that all IS-08 mappings remain consistent with the EDID-derived Receiver Capabilities.

8.1.2 IS-08 Overview & Usage (Informative)

What IS-08 Is

AMWA IS-08: Audio Channel Mapping provides a common, interoperable API that lets a receiver *reorder, shuffle, combine, or select* audio channels from one or more incoming streams.

It defines:

● A standard JSON-based "mapping" resource
● How mappings relate to receivers, senders, and IS-05 connection resources
● How control systems can modify the mapping at runtime

IS-08 does not perform audio processing itself. Instead, it exposes a *control surface* for describing how channels are wired inside the device.

Why IS-08 Matters in IPMX

IPMX PCM audio can arrive with:

● 1, 2, or 8 channels
● Either L16 or L24
● Arbitrary channel orders, depending on the source device
● Multichannel program audio used in:
    ○ lecture capture
    ○ live production
    ○ conferencing
    ○ distributed audio
    ○ signage
    ○ displays with downmix constraints

Without IS-08, a receiver has no standardized way to:

● pick which channels to monitor
● map 8 channels into a 2-channel output
● reorder channels for compatibility with a display or DSP

- construct stereo pairs from multichannel content
- select a subset of channels from multiple streams

IS-08 turns those ad-hoc proprietary behaviors into interoperable, NMOS-based behaviors.

Typical Use Cases for IS-08 in IPMX

1. Display / Monitor Devices (EDID-Driven)

A common scenario:

- A display advertises **2-channel audio** via EDID.
- An IPMX sender produces **8-channel** PCM audio.
- IS-05 connects the stream, but the receiver must reduce 8→2 channels.

Using IS-08, the device can:

1. Receive all 8 channels from the IPMX stream
2. Select the desired pair (e.g., Ch1/Ch2 or Ch3/Ch4)
3. Reorder or downmix (if the device supports internal DSP)
4. Expose the final 2-channel mapping in NMOS

This allows the display to operate correctly without breaking network-level interoperability.

2. Multichannel Audio Monitoring Panels

Many studio-grade or Pro AV monitor devices need to:

- solo or mute individual channels
- swap L/R
- switch between several stereo pairs
- fold down multichannel feeds

IS-08 provides the controller with a standard way to request these internal routing behaviors.

3. DSP / Matrix Processors

Devices with internal DSP (AEC processors, matrix engines, etc.) often:

- accept multiple incoming IPMX audio streams
- build internal audio graphs

- route subsets into processing blocks

IS-08 provides a control-plane mapping that aligns NMOS routing with the internal DSP graph.

4. Live Production & Events

Audio consoles and zone routing systems may:

- map different channel groups to different zones
- maintain consistent channel order across multiple senders
- reorder channels when ingesting from different vendors' encoders

IS-08 ensures that control systems can do this in a vendor-neutral way.

5. Emergency / Paging Systems

When paging overlays need specific channels (e.g., Ch1 must be "Paging", Ch2 must be "Program Right"), IS-08 ensures the receiver consistently extracts the correct channel.

Why IS-08 Is Optional (But Highly Recommended)

IS-08 is optional in IPMX because many devices:

- only support a fixed channel configuration
- only output a single format (e.g., 2-ch audio for a display)
- do not reorder audio internally

However, any device that:

- presents HDMI or SDI outputs
- performs monitoring
- implements EDID behavior (BCP-005-01)
- performs downmixing
- acts as a multichannel bridge
- integrates with DSP pipelines

will almost certainly benefit from IS-08.

9. Interoperability Notes (Informative)

This section provides context regarding how the IPMX PCM Digital Audio Profile relates to AES67 and ST 2110-30. These are not additional requirements; they clarify how interoperability behaves in practice.

9.1 Relationship to AES67 and SMPTE ST 2110-30

9.1.1 Clocking and Synchronization Behavior

AES67 and ST 2110-30 (hereinafter AES67, if applicable to both) require that all streams be time-synchronous and locked to PTP.

This means:

- AES67 senders always produce media timestamps aligned to the system PTP clock.
- AES67 receivers assume that incoming streams are synchronized to that same system clock.

By contrast, some IPMX use cases may involve media clocks not inherently tied to the PTP epoch (for example, embedded processors with fixed-rate local clocks).

Interoperability implication

If an IPMX device produces audio driven by an internal media clock that is *not* synchronized to PTP:

- AES67 / ST 2110-30 receivers may reject the stream or drift against it.
- Interoperability requires the IPMX sender to resample or otherwise reconcile the internal media clock to PTP prior to transmission.

Implementers targeting AES67 / ST 2110-30 environments should explicitly account for this requirement.

9.1.2 Supported Channel Counts

AES67 and ST 2110-30 require that receivers accept any valid channel count from 1 to 8 channels, not just fixed sets such as 1, 2, 4, or 8.

The IPMX PCM Digital Audio Profile defines mandatory support for the following channel-count combinations:

- 1 channel
- 2 channels
- 8 channels

These combinations cover the most common IPMX and Pro AV use cases; however:

Interoperability implication

IPMX receiver implementers who expect to operate in AES67 / ST 2110-30 environments should strongly consider supporting *all* channel counts between 1 and 8 (inclusive).

This ensures full interoperability with AES67 transmitters, which may send (for example) 3-, 5-, 6-, or 7-channel streams.

9.1.3 RTP Timing and the PTP Epoch

AES67 and ST 2110-30 use the same PTP epoch, but differ in how they define the RTP timestamp offset:

- AES67 (following RFC 3550) allows and recommends using random RTP timestamp offsets for individual streams.
- ST 2110-10 requires the RTP timestamp offset to be zero for all media streams.
- IPMX, through TR-10-1 and TR-10-3, follows the ST 2110 rule and therefore also requires an offset of 0.

Interoperability implication:

IPMX receivers must accept AES67 streams with non-zero RTP timestamp offsets, while IPMX senders will always produce zero-offset RTP timestamps.

9.2 RTP Timing, RTCP Sender Reports, and AES67 Behavior

AES67 / ST 2110-30 systems typically do not send RTCP Sender Reports, because:

- These standards rely fully on PTP for synchronization.
- Senders are required to derive and align their media clocks from PTP.
- Receivers are expected to use the PTP clock as the reference for stream alignment.

TR-10-3 notes that IPMX receivers must accept such streams and operate correctly in the absence of RTCP Sender Reports.

Interoperability implication

When an IPMX receiver ingests an AES67 or ST 2110-30 stream without RTCP:

- It may rely on the fact that AES67 / 2110-30 streams are always PTP-synchronous.
- The receiver should use PTP as the master reference for alignment and drift management.

This behavior is already addressed by TR-10-3, but is highlighted here to clarify expectations for implementers targeting mixed AES67 / IPMX environments.

9.3 Summary of Key Differences

Topic	AES67 / ST 2110-30	IPMX PCM Digital Audio	Interoperability Notes
Clocking	Always PTP-locked	May originate from internal clock	IPMX sender may have to resample non-PTP locked input sources to interoperate with AES67 receivers
Channel Counts	Any 1–8	Required: 1, 2, 8	IPMX receivers should consider supporting full 1–8 range
RTP timestamp offset	random	Always 0	Handled automatically by receivers; implementers should be aware
RTCP Sender Reports	Typically absent	TR-10-3 requires receivers to accept this	IPMX receivers rely on PTP-only timing when RTCP is absent

Appendix A: Revision History (Informative)

Version	Date	Summary of Changes
1.0	2025-12-01	Initial Release
0.3	2025-11	Corrections to AES67 interoperability.
0.2	2025-11	Refactored draft to reduce redundancy. Added complete interoperability notes for AES67 and ST 2110-30
0.1	2025-11	Initial Draft