

Audio Containers — Cheat Sheet

MP4, fMP4, Matroska/WebM and MPEG-TS — where each one stores its audio.

The one rule

Audio lives in self-describing tracks, so one file can carry many languages and mixes.

The codec compresses the sound (Opus, AAC, AC-3, FLAC). The container labels and bundles it (MP4, MKV, TS).

The four containers

Container	Structure	Audio named by	Best at
MP4 / fMP4	Box tree (moov/mdat)	stsd entry: mp4a + esds	Universal playback; CMAF streaming
Matroska / MKV	EBML element tree	CodecID: A_OPUS, A_AAC	Flexibility, archival, many tracks
WebM	EBML (Matroska subset)	CodecID SHOULD be A_OPUS	Open web <video>, royalty-free
MPEG-TS	188-byte packets + PIDs	stream_type in the PMT	Broadcast, loss recovery, legacy HLS

MPEG-TS audio stream_type (in the PMT)

0x0F	AAC (ADTS framing)
0x11	AAC (LATM framing)
0x81	AC-3 (ATSC convention)
0x03 / 0x04	MPEG-1 / MPEG-2 audio

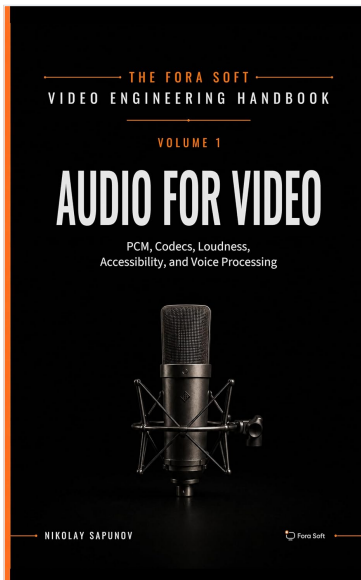
How an MP4 stores audio (ISO/IEC 14496-12 / -14)

moov = table of contents | mdat = cargo (raw samples) | path: moov > trak > mdia > minf > stbl > stsd

AAC codec entry = mp4a; setup bytes = esds. fast-start moves moov to the front for progressive play.

Silent-track checklist (audio plays in one player, fails in another)

- ☐ MP4: the esds box (AAC AudioSpecificConfig) is present and correct.
- ☐ Matroska: the CodecPrivate setup bytes are present for the audio track.
- ☐ MPEG-TS: the PMT lists the audio PID with the right stream_type.
- ☐ Language tag set on each audio track so players show the menu.
- ☐ fast-start: moov relocated to the front for progressive download.



GO DEEPER · VOLUME 4 OF 7

Audio for Video: A Complete Guide to Sound in Video Products

From PCM and Codecs to Loudness, Accessibility, and Real-Time Voice Processing

By Nikolay Sapunov, CEO at Fora Soft · 1,229 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0HH6PRZH1

This cheat sheet shows where MP4, Matroska, WebM and MPEG-TS store audio and what to check when a track plays silent. Volume 4 covers the audio itself in depth: codecs such as Opus and AAC, channel layouts and multiple language tracks, loudness and accessibility.

WHAT YOU WILL FIND INSIDE

- Opus and AAC configuration details
- Accessibility tracks and audio description
- Channel layouts and multi-language tracks
- Loudness normalisation with LUFS

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI

Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF

Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery

Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing

Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs

Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video

Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Audio playing in one player and silent in another?

We will inspect your files and packaging, find the missing setup bytes or track flags and fix the pipeline that writes them. 250+ video products shipped since 2005.

[Book a 30-min call →](#)
[WhatsApp →](#)
[Email us →](#)