

Drift Correction Cheat Sheet

Two clocks, two fixes: continuous resampling vs frame drop / insert - on one page.

The root cause: two clocks, never the same

Two clocks	Capture and playback run on separate crystals; rates differ by parts per million (ppm).
Consumer spec	Typical crystals +/-20 to +/-50 ppm; two devices can differ by up to 100 ppm.
The squeeze	Faster producer overflows the buffer; slower producer empties it into silence.

How drift accumulates

50 ppm	$= 50 / 1,000,000 = 0.00005$ fractional clock difference.
per second	$\times 1 \text{ s} = 0.05 \text{ ms}$ gained or lost; at 48 kHz that is 2.4 samples/s.
per hour	$\times 3,600 \text{ s} = 180 \text{ ms}$ - at the ITU-R BT.1359-1 limit (+90 / -185 ms).

Soft vs hard correction

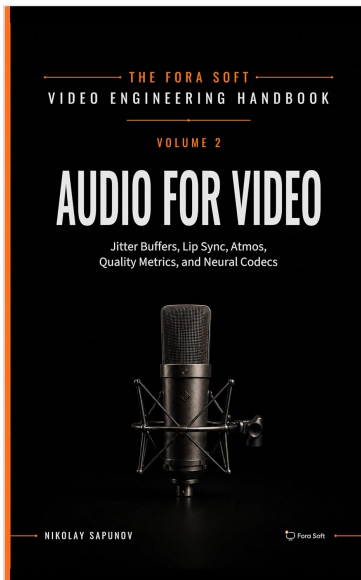
Soft / resample	Nudge a fraction of every sample. Inaudible. Higher CPU. The audio default.
Hard / drop+insert	Remove or repeat a whole 10-20 ms frame. Cheap. Audible click, gap, stutter.
WebRTC NetEQ	Accelerate shortens, PreemptiveExpand lengthens - pitch-preserving time stretch.
Video	Frames are discrete: drop or repeat only. The eye forgives it; the ear does not.

The audio-master rule for video

Audio master	Audio is the reference timeline; video drops or repeats frames to follow it.
Why	A 10 ms audio gap is more noticeable than a duplicated video frame.

Four common mistakes

- ☐ Blaming the network: jitter is random; drift is a steady one-directional slide.
- ☐ Reactive-only buffer: correct continuously, do not wait for empty or full.
- ☐ Pitch-shifting resampler: use pitch-preserving time stretch, not a raw rate change.
- ☐ Inverted master: drop video frames to follow audio, never the other way round.
- ☐ Tell-tale: fine at the start, worse over time = clock drift, never a one-off bug.



GO DEEPER · VOLUME 5 OF 7

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

From Jitter Buffers and Lip Sync to Atmos, Quality Metrics, and Neural Codecs

By Nikolay Sapunov, CEO at Fora Soft · 1,288 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0HH6XPY46

This sheet explains why two clocks drift apart and when to resample or drop and insert frames. Volume 5 covers this in depth: timestamps and clock domains, jitter buffers and NetEq, lip sync and the audio-master rule, and incident analysis by symptom.

WHAT YOU WILL FIND INSIDE

- Clock domains and timestamps
- Jitter buffers and NetEq
- Lip sync and A/V alignment
- Incident analysis by symptom

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 slowly drifts out of sync during long calls or streams?

We will measure clock drift on your devices, add continuous resampling and keep video following the audio clock. 250+ video products shipped since 2005.

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