

WebRTC Lip-Sync Debugging Checklist

Two timestamps, one SFU problem, three engines — how to find a group-call sync bug.

The mechanism: two timestamps + a CNAME

RTP timestamp	On every packet. 48 kHz for Opus audio, 90 kHz for video. Measures intervals in one stream only.
RTCP Sender Report	Every few seconds. Pairs an NTP wall-clock time with the RTP timestamp — the cross-stream anchor.
CNAME	Same value for one speaker's audio and video, so the receiver knows they belong together.

Why the SFU breaks it

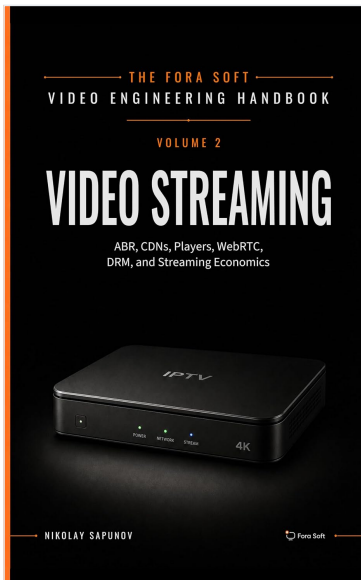
1:1 call	Sender's own Sender Reports reach the receiver untouched. Sync is solved for free.
SFU group call	SFU terminates RTCP and regenerates Sender Reports on its own clock. Danger point.
Absolute Capture Time	Header extension stamping original capture NTP time on packets — survives the SFU if forwarded.

mediasoup vs Pion vs LiveKit

mediasoup	Rewrites RTP timestamps from the SR NTP relationship on stream switch. App must wire A/V identically.
Pion	Correct SR primitives + report interceptor. You own all the sync logic; nothing is automatic.
LiveKit	Server-side sync from track SR NTP timestamps out of the box. Edge cases (RED, egress) drop an SR.

Five checks when sync drifts

- ☐ **Split media paths:** audio on one server, video on another = no shared anchor. Go all-in SFU or all-in MCU.
- ☐ **Missing SR for one stream:** check <chrome://webrtc-internals> shows Sender Reports for BOTH audio and video.
- ☐ **Capture-time extension dropped:** verify Absolute Capture Time is negotiated AND forwarded on every hop.
- ☐ **Blaming the network:** steady one-directional drift = clock or timestamp problem, never jitter.
- ☐ **Tell-tale:** fine 1:1, broken in group calls = the SFU's timing, not the codec or the connection.



GO DEEPER · VOLUME 3 OF 7

Video Streaming: A Complete Guide to Video Delivery

From ABR and CDNs to Players, WebRTC, DRM, and the Economics of Streaming

By Nikolay Sapunov, CEO at Fora Soft · 1,182 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0HH76ZYWC

This checklist explains the RTP and RTCP timestamps behind lip sync and where an SFU breaks them. Volume 3 covers WebRTC and real-time delivery in depth, together with players, stall recovery and QoE monitoring that catch drift before users do.

WHAT YOU WILL FIND INSIDE

- WebRTC and real-time delivery
- QoE monitoring
- Players and stall recovery
- CDN and caching

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 and video drift apart in your group calls?

We will check Sender Reports and capture-time extensions on every hop of your SFU and fix the sync logic in mediasoup, Pion or LiveKit. 250+ video products shipped since 2005.

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