

RTP / RTCP Sync Reference Card

Timestamp clock rates, the sender-report fields, the wall-clock math, and the failure modes — on one page.

The two pieces

RTP ts	Counts media sampling ticks from a random start. Not packets, not wall time.
RTCP SR	Pairs one RTP reading with the NTP wall-clock time at the same instant.

RTP media clock rates

Opus audio	48,000 Hz — 960 ticks per 20 ms packet (RFC 7587).
G.711 audio	8,000 Hz — 160 ticks per 20 ms packet (RFC 3551).
Video	90,000 Hz — 3,000 ticks per frame at 30 fps (RFC 3551).

Media clock to wall clock

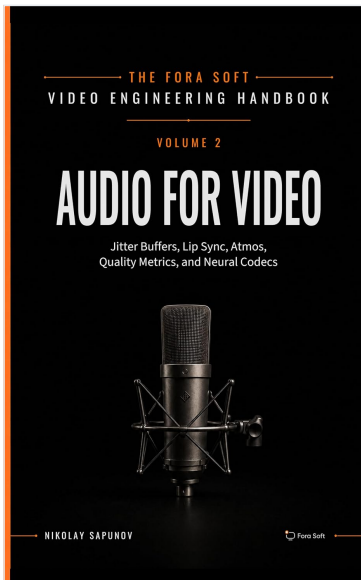
NTP format	64-bit fixed point: seconds since 0h UTC 1 Jan 1900; 32 whole + 32 fraction.
SR pair	RTP ts R0 maps to NTP time T0 (same instant).
Convert	$\text{wall}(R) = T0 + (R - R0) / \text{clock rate}$. Do it per stream.
Lip-sync	Both streams on one wall clock; render matching audio + frame together.

When sync breaks

Failure mode	Cause	Symptom
Clock drift	ppm crystal mismatch	Lip-sync drifts steadily; worse over a long call.
Network jitter	uneven arrival	Sync wobbles, unstable, not one-directional.
Missing / wrong SR	no valid anchor	Streams fine alone, permanently out of sync.
Sync lock delay	RTCP rate limit	First ~1-5 s off, then snaps in (RFC 6051 fixes).

Remember

- ☐ Audio and video are separate RTP streams with separate random-start clocks.
- ☐ The RTP timestamp counts media sampling ticks, not packets and not wall time.
- ☐ Audio clocks at 48,000 Hz (Opus) or 8,000 Hz; video at 90,000 Hz.
- ☐ The RTCP sender report pairs one RTP reading with an NTP wall-clock time.
- ☐ One pair per stream puts both on one timeline — that is lip-sync.
- ☐ Sender reports must keep arriving; each fresh one re-anchors against drift.
- ☐ RFC 6051 and WebRTC abs-capture-time carry the mapping in-band for fast sync.
- ☐ A middlebox that regenerates SRs on its own clock silently breaks sync.



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 card gives the RTP clock rates, sender-report fields, wall-clock math and sync failure modes on one page. Volume 5 covers synchronization in depth: lip sync, timestamps and clock domains, jitter buffers and NetEq, and how loss and packet timing show up as symptoms.

WHAT YOU WILL FIND INSIDE

- Lip sync, timestamps and clock domains
- Jitter buffers and NetEq
- Packet loss concealment, FEC and RED
- Diagnosing sync incidents 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

Chasing lip-sync drift in your calls or live streams?

We will trace RTP timestamps and sender reports end to end, find where the clock mapping breaks, and fix it in your media pipeline. 250+ video products shipped since 2005.

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