

Live vs VOD vs Near-Live — Decision Card

One page. Three categories. The deadline-driven trade-offs and the protocol families that fit each.

	VOD	Near-live	Live
	<i>Content exists before play</i>	<i>Live with deliberate delay</i>	<i>Every stage races a clock</i>
Latency:	not measured	5-90 s (deliberate)	0.2-30 s (target)
Protocol:	HLS / DASH / CMAF	HLS / DASH / CMAF	LL-HLS / WebRTC / MoQ
Buffer:	20-60 s	6-18 s	0.05-6 s
Cache hit:	~99%	~95%	5-85%
Optimise:	cost, quality	predictable delay	latency, scale
Examples:	Netflix, YouTube, lecture archives	sports betting, live shopping, linear TV	Twitch, sports, telemedicine, auctions
Failure mode:	egress bill, codec reach	drift, moderator overload	stalls during peak moments

Three questions to pick a category

- Q1.** Does the content exist before the viewer presses play?
Yes -> VOD. Otherwise continue to Q2.
- Q2.** Is sub-second interactivity required (bidding, conferencing, telemedicine)?
Yes -> Live, sub-second (WebRTC / MoQ). Otherwise continue to Q3.
- Q3.** Is a deliberate 5-90 s delay acceptable (betting, moderation, broadcast graphics)?
Yes -> Near-live (HLS / DASH / CMAF with 6-18 s buffer). Otherwise: **Standard / LL live (HLS, DASH, LL-HLS, LL-DASH).**

Rule of thumb

Tightening latency by 10x roughly doubles unit cost. Most 'we need low latency' requests are actually near-live in disguise. Pick the loosest deadline your product can live with — and design the architecture around it.

Primary standards (per Fora Soft Learn editorial source-hierarchy)

HLS: IETF RFC 8216 (Aug 2017). LL-HLS: Apple HLS Authoring Specification, rev 2025-09. DASH: ISO/IEC 23009-1:2022. CMAF: ISO/IEC 23000-19:2024. WebRTC: W3C Candidate Recommendation + IETF RFC 8825-8866. QUIC: IETF RFC 9000 (May 2021).



GO DEEPER · VOLUME 2 OF 7

Video Streaming: A Complete Guide to Video Delivery

From Network Protocols and Transport to Ingest, HLS, and MPEG-DASH

By Nikolay Sapunov, CEO at Fora Soft · 1,124 pages · Kindle & paperback on Amazon

Read it on Amazon →

amazon.com/dp/B0HH8F4R82

This card sorts a product into VOD, near-live or live by its deadline and names the protocol families for each. Volume 2 covers that choice in depth: live, VOD and near-live delivery, the glass-to-glass latency budget, TCP versus UDP, QUIC, and HLS, MPEG-DASH and CMAF packaging.

WHAT YOU WILL FIND INSIDE

- Live, VOD and near-live delivery models
- HLS, MPEG-DASH and CMAF segments
- Glass-to-glass latency budgets
- QUIC, HTTP/3 and transport choices

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

Not sure whether your product needs live, near-live or VOD delivery?

We will match your latency target to a protocol and architecture and show what each option costs per viewer. 250+ video products shipped since 2005.

Book a 30-min call →

WhatsApp →

Email us →