

Live vs VOD vs near-live — карточка решения

Одна страница. Три категории. Компромиссы по дедлайну и семейства протоколов под каждый.

VOD

Контент готов до play

Задержка не измеряется

Протокол: HLS / DASH / CMAF

Буфер: 20–60 с

Cache hit: ~99%

Оптимизирует:
стоимость, качество

Примеры:
Netflix, YouTube,
архив лекций

Режим отказа: счёт egress, охват ко

Near-live

Live с намеренной задержкой

Задержка 5–90 с (намеренно)

Протокол: HLS / DASH / CMAF

Буфер: 6–18 с

Cache hit: ~95%

Оптимизирует:
предсказуемую задержку

Примеры:
ставки, live-шоппинг,
linear TV

Режим отказа: дрейф, перегруз мод

Live

Каждая стадия гонится за часами

Задержка 0,2–30 с (таргет)

Протокол: LL-HLS / WebRTC / MoQ

Буфер: 0,05–6 с

Cache hit: 5–85%

Оптимизирует:
задержку, масштаб

Примеры:
Twitch, спорт,
телемедицина, аукционы

Режим отказа: подвисания в пик

Три вопроса, чтобы выбрать категорию

- B1.** Контент существует до того, как зритель нажмёт play? **Да — VOD.** Иначе — к B2.
- B2.** Нужна sub-second-интерактивность (ставки, конференции, телемедицина)?
Да — Live, sub-second (WebRTC / MoQ). Иначе — к B3.
- B3.** Допустима намеренная задержка 5–90 с (ставки, модерация, графика)?
Да — Near-live (HLS / DASH / CMAF, буфер 6–18 с).
Иначе: **стандартный / LL live (HLS, DASH, LL-HLS, LL-DASH).**

Правило большого пальца

Ужесточение задержки в 10 раз примерно удваивает юнит-стоимость. Большинство запросов «нам нужен low-latency» на деле — near-live в маске. Выбирайте самый щадящий дедлайн, который продукт может себе позволить, и проектируйте архитектуру под него.

Первичные стандарты (по редакционной иерархии источников Fora Soft Learn)

HLS: IETF RFC 8216 (август 2017). LL-HLS: Apple HLS Authoring Specification, ревизия 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 (май 2021).