


Хроники застывшего кадра, или как победить зависания видео


Алексей Шпагин, ВКонтакте




VideoTech
2024

Обо мне

 Бэкграунд — разработчик C++

 12 лет работы в телекоме,
видеозвонках, стриминге

 В руководстве командами 6 лет

 Руководитель бэкенд-разработки
единой видеоплатформы



Содержание

1

Единая
видеоплатформа
и зависания видео

2

Инструменты
для изучения
зависаний видео

3

Доработки и
эксперименты

4

Результаты

Единая видеоплатформа и зависания видео



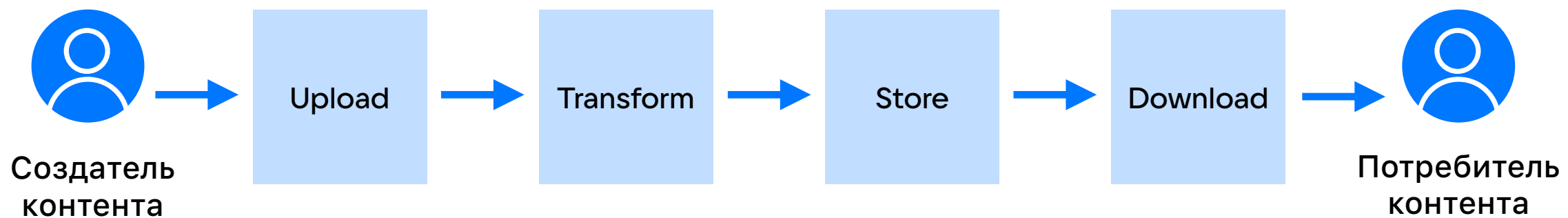
Видеоплатформа

Видео-
ХОСТИНГ

Стриминг

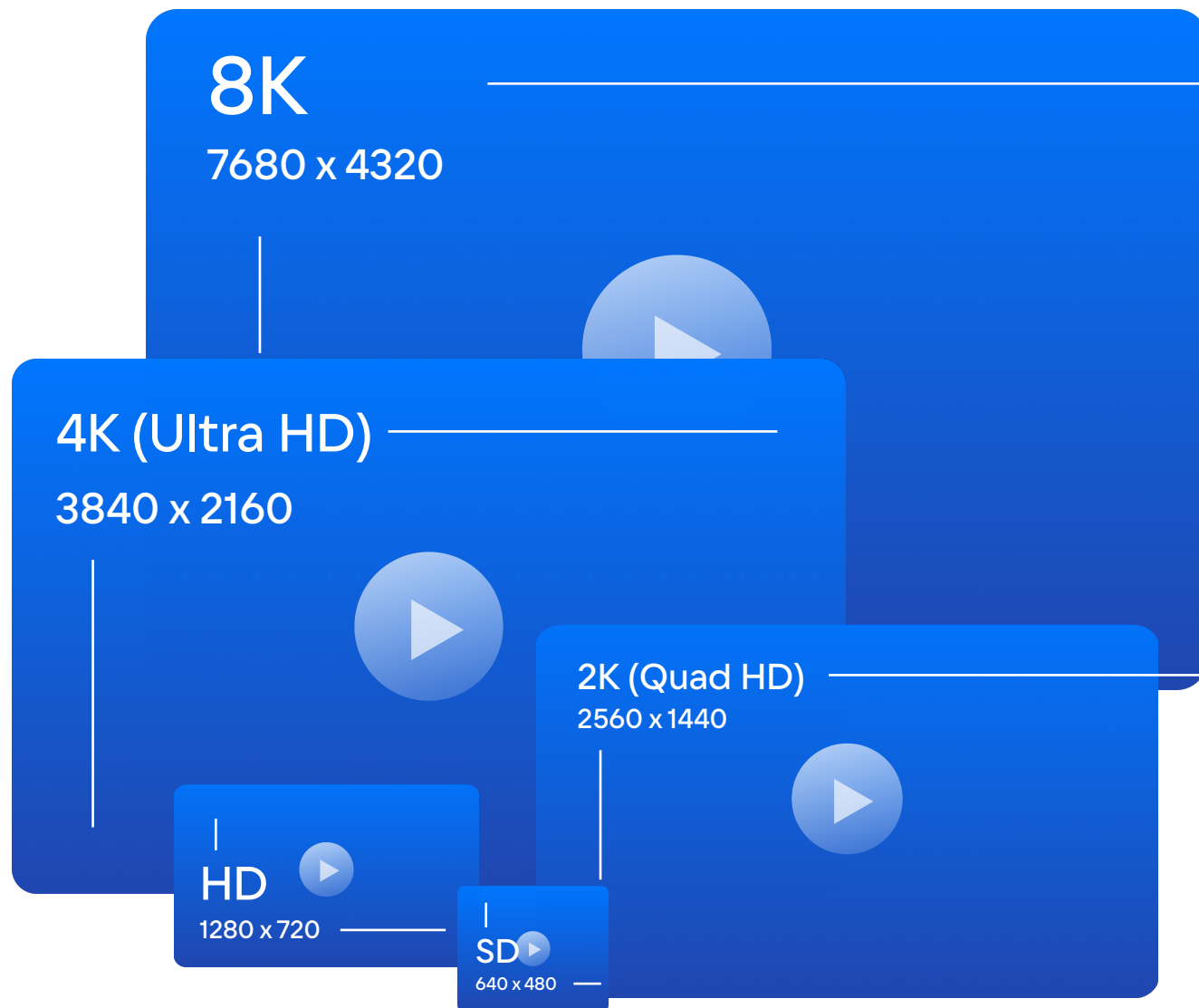
Звонки

Backend видеоплатформы



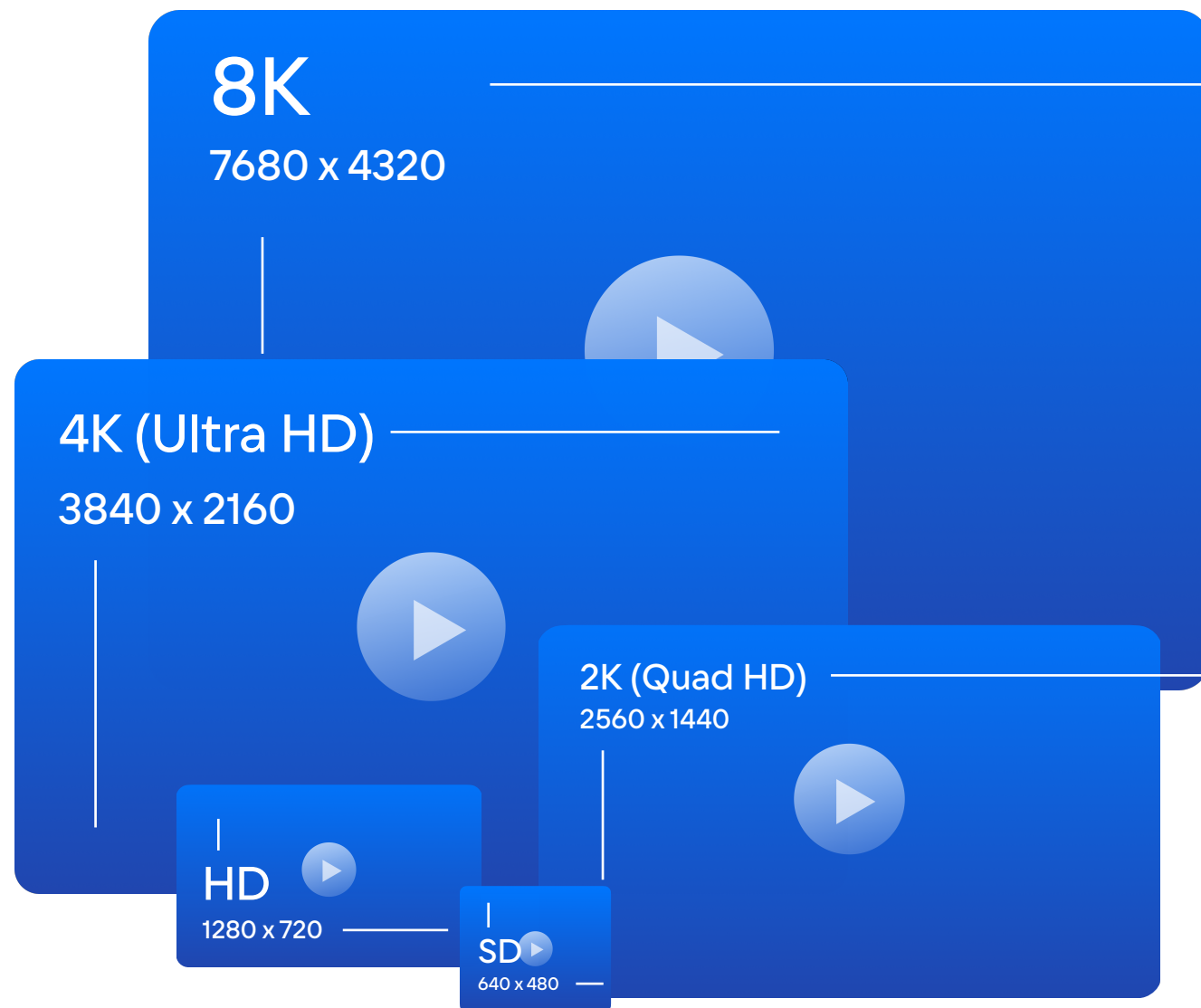
SDK

- Плеер
- MP4, HLS, DASH
- Приложения VK
- Партнеры

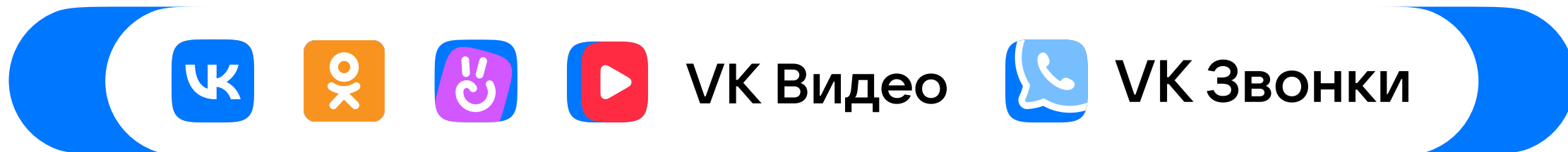



Клиентские платформы

- Web
- iOS
- Android Mobile
- Android TV



Единая видеоплатформа



 **4,1 млрд**
просмотров VK Видео и
VK Клипов в сутки

 **> 1 Эбайт**
хранилище

 **7+ Тбит/с**
трафик

 **1 Пбайт**
новых видео ежедневно



В поддержку
поступают жалобы
на зависания
(стóллы*) видео
при просмотре

* Стóлл, от англ. «to stall» - задержать, застопорить, буксовать, глохнуть

В чем может быть причина СТОЛЛОВ?

- Uptime платформы видео 99.99%
- Нет корреляции жалоб с часом пик по нагрузке
- Есть запас по капасити бэкенда
- «Детские» болезни давно вылечены
- Нет очевидных узких мест

Единая видеоплатформа

1

Видеоплатформа

2

SDK, продукты на базе платформы

3

Жалобы на зависания

4

Причины столлов не очевидны

Инструменты для изучения и отладки



Причины зависания видео

Сеть

Производительность

- ✓ Недостаточная скорость интернет соединения
- ✓ Недостаточно совершенный алгоритм выбора качества видео на клиенте
- ✓ Проблемы на раздаче видео

Причины зависания видео

Сеть

Производительность

- ✓ Слабое устройство
- ✓ Videopоток слишком «жирный» для текущего устройства
- ✓ Плеер конкурирует за ресурсы с другими запущенными приложениями

Верхнеуровневое решение

1

Доработать метрику
для оценки
количества столов

2

Разработать
инструменты
для отладки

3

Провести ревью
всех ключевых
узлов

Метрика оценки количества столов



Данные, необходимые для метрики столов

TVT — Total View Time

Количество секунд, которые суммарно все пользователи платформы провели за просмотром видео

Empty Buffer

Время в мс, когда буфер с видеоданными был пуст.
Суммируем по всем сессиям просмотра, исключая Empty Buffer перед началом просмотра

Метрика количества
СТОЛЛОВ

Empty Buffer/TVT —
сколько было
миллисекунд столлов
на секунду просмотра



Свойства данной метрики

Относительная, не зависит от
длины видео и периода оценки

Различные платформы можно
сравнивать друг с другом

Подходит для всех платформ

Позволяет увидеть только
столлы, связанные с сетевым
взаимодействием

Целевые показатели

KPI

столлы по всем платформам
< 5 мс/с (суммарно 18 секунд
столлов за час просмотра)

OKR

столлы по всем платформам
< 3 мс/с (суммарно 11 секунд
столлов за час просмотра)

Стартовая
позиция по
метрике
Empty Buffer/TVT

Платорма	Сейчас, мс/с	Цель, мс/с
Мобильный Веб	12.94	3
Веб	6.54	3
Android ВКонтакте	4.65	3
Android VK Видео	3.86	3

Где будем чинить

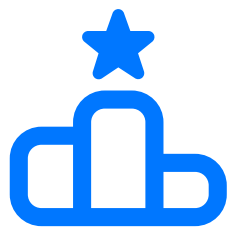
- ✓ Android приложение ВКонтакте и отдельное приложение VK Видео
- ✓ Мобильный Веб
- ✓ Веб для десктопа

Остальные платформы показывают столлы ниже порога, которым задались сейчас (< 3 мс/с), поэтому их на данном этапе не рассматриваем

Телеметрия



Телеметрия: предпосылки



Подозрение, что есть конкуренция за ресурсы между приложениями или даже модулями приложения ВКонтакте.



Нужно более детально видеть, где происходит «затык» при скачивании видео.

Телеметрия: реализация

- ✓ Вдохновлялись OpenTelemetry
- ✓ Воспользовались наработками Одноклассников, продуктом Tracer
- ✓ Внедрили телеметрию на бэкенде

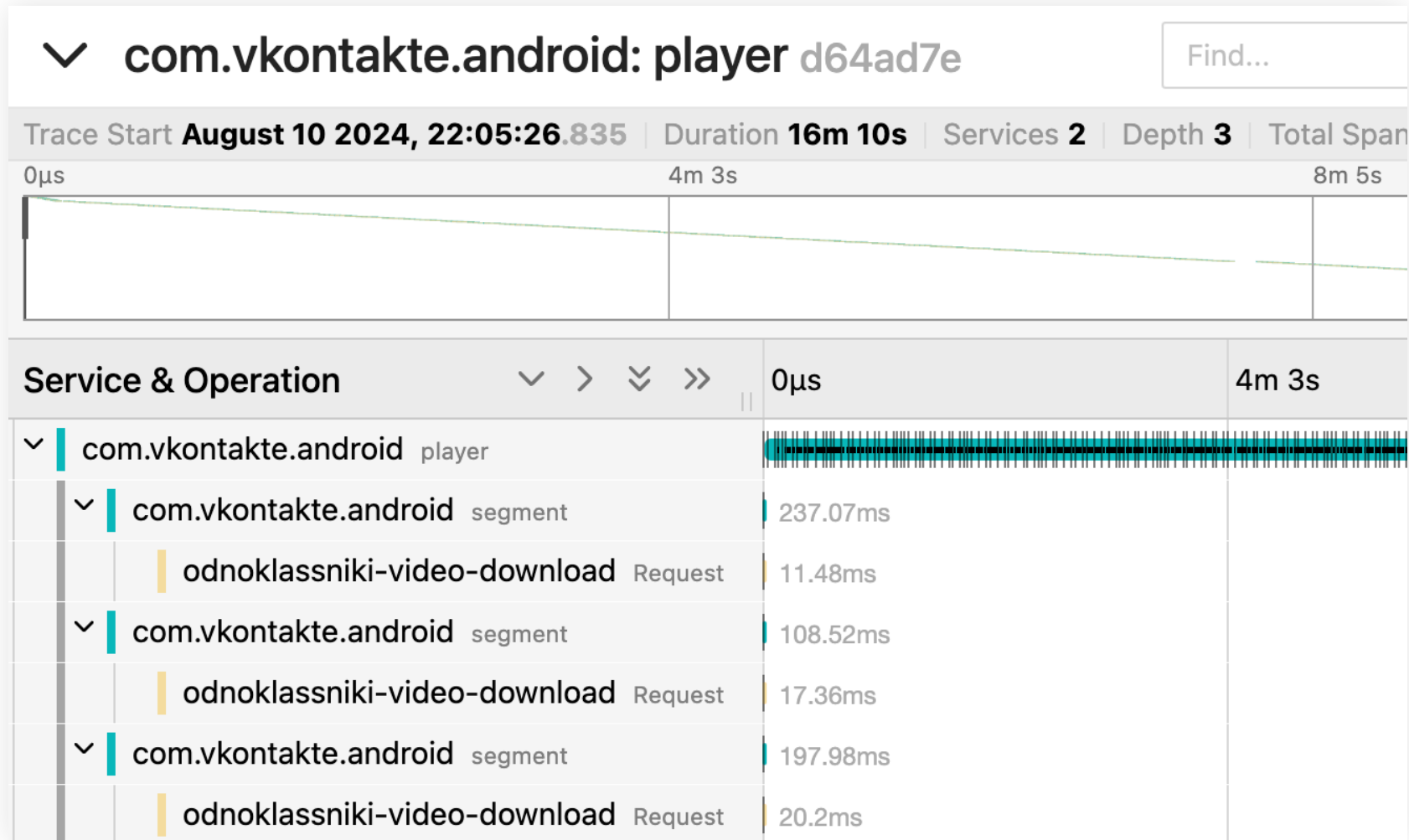


КИРИЛЛ ПОПОВ. MOBIUS
2023



APPTRACER.RU

Телеметрия: распределенный трейсинг



player

Service: **com.vkontakte.android** | Duration: **16m 9s** | Start Time: **580.91ms**

> **Tags:** content.id = 6244984818398 | user.id = [REDACTED]

> **Process:** service.name = com.vkontakte.android | telemetry.sdk.language = java | telemetry.sdk.name = ru.ok.tracer | telemetry.sdk.version = 0...

Logs (313)

> **185.42ms:** event = onPlayerInit

> **185.45ms:** event = NEW_STATE: STATE:INIT

> **185.58ms:** event = onPlayerBuffering

> **185.58ms:** event = NEW_STATE: STATE:BUFFERING

> **405.64ms:** event = onFirstBytes

> **535.31ms:** event = bandwidthEstimate | buf.durationMs = 0 | bw.bitrateEstimate = 17435331 | bw.totalBytesLoaded = 16256 | bw.totalLoad...

> **698.76ms:** event = onVideoSizeChanged | height = 1080 | pixelWidthHeightRatio = 1 | unappliedRotationDegrees = 0 | width = 1920

> **698.87ms:** event = onCurrentVideoTrackChanged | videoTrack.bitrate = 1136470 | videoTrack.formatId = 5 | videoTrack.frameRate = 25 | vi...

> **699.12ms:** event = onFirstFrameDecoded

> **718.25ms:** event = onFirstFrameRendered

> **2.65s:** event = bandwidthEstimate | buf.durationMs = 5379 | bw.bitrateEstimate = 13336679 | bw.totalBytesLoaded = 245137 | bw.totalLoa...

> **5.2s:** event = bandwidthEstimate | buf.durationMs = 17545 | bw.bitrateEstimate = 8470306 | bw.totalBytesLoaded = 477814 | bw.totalLoadT...

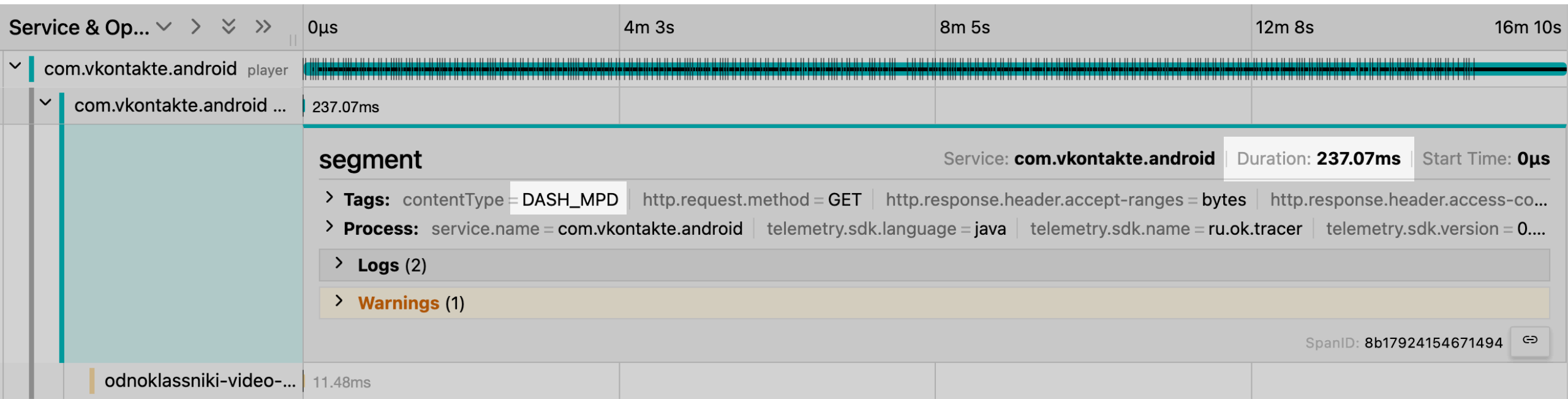
> **7.37s:** event = bandwidthEstimate | buf.durationMs = 39173 | bw.bitrateEstimate = 12946609 | bw.totalBytesLoaded = 581381 | bw.totalLoa...

> **9.71s:** event = bandwidthEstimate | buf.durationMs = 57431 | bw.bitrateEstimate = 13165366 | bw.totalBytesLoaded = 339035 | bw.totalLoa...

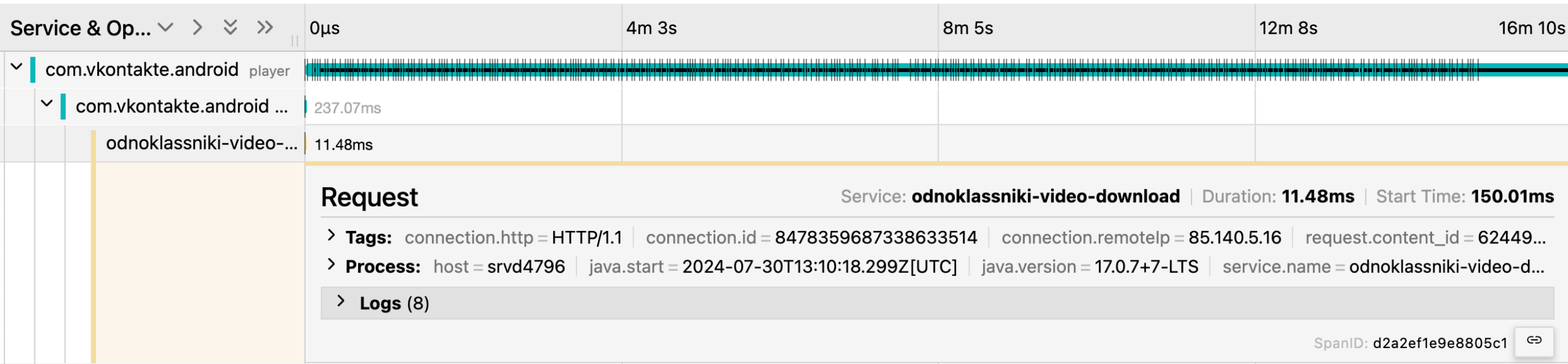
> **12.44s:** event = bandwidthEstimate | buf.durationMs = 69775 | bw.bitrateEstimate = 12845749 | bw.totalBytesLoaded = 112613 | bw.totalLo...

> **15.63s:** event = bandwidthEstimate | buf.durationMs = 69586 | bw.bitrateEstimate = 12350623 | bw.totalBytesLoaded = 399960 | bw.totalL...

Телеметрия: спан запроса манифеста



Телеметрия: спан сервера раздачи



A/B тестирование



Цели A/B тестирования

1

Увидеть эффект от доработок

2

Уменьшить негативный эффект на пользователей, если эксперимент окажется неудачным

Проведение и интерпретация A/B

Проводим так же эксперименты A/B/C, A/B/C/D и т.д.

Смотрим на рост целевой метрики столов

Смотрим на дополнительные метрики:
first bytes, количество завершений просмотров во время столов и др.

При снижении продуктовых метрик не пускаем доработку в прод

Решение и инструменты

1 Направление движения

2 Ключевая метрика

3 Целевые показатели

4 Инструменты отладки

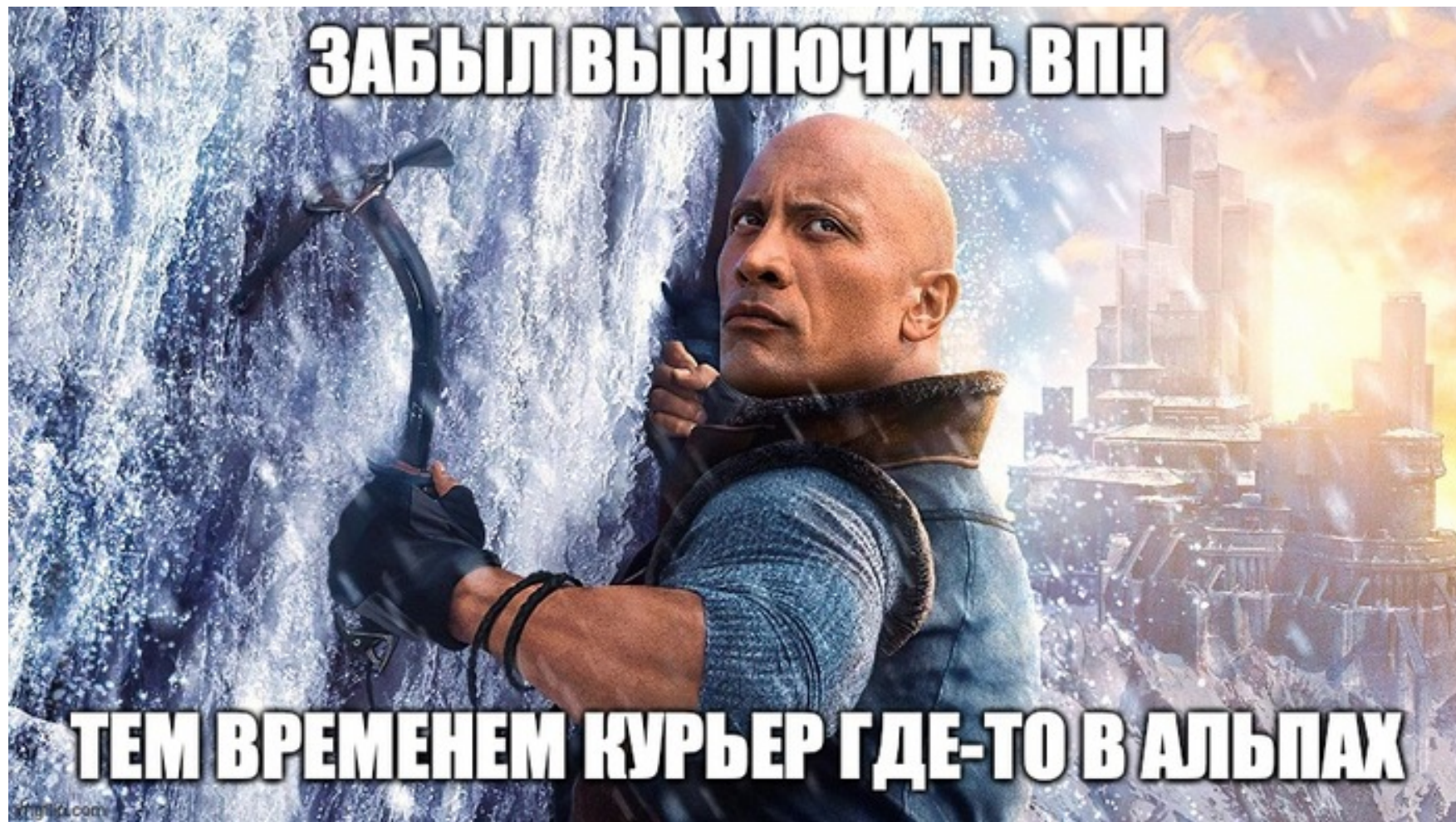
Доработки и эксперименты



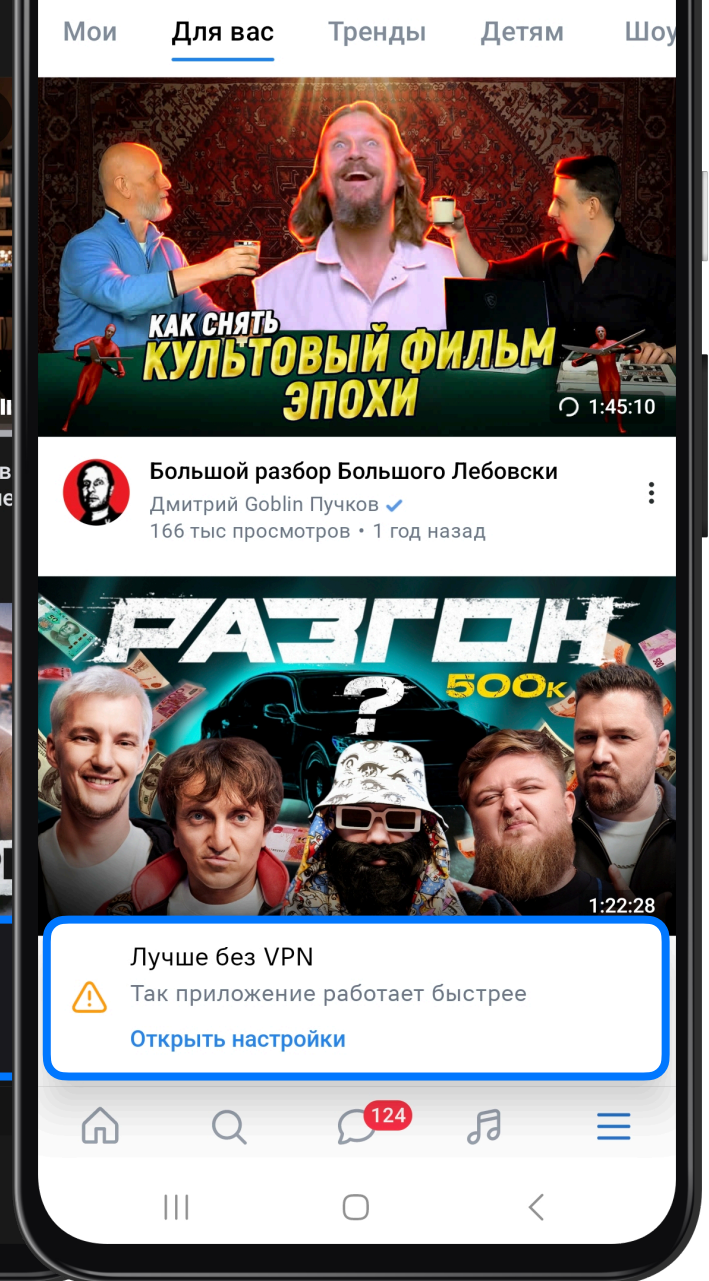
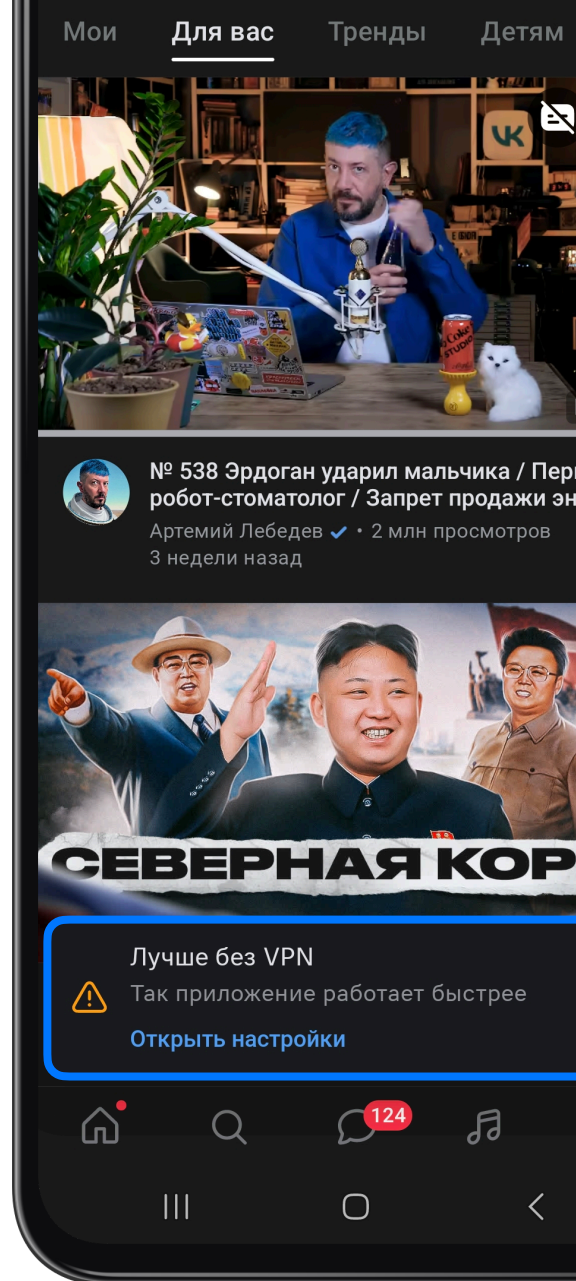
VPN



VPN



Плашка-напоминалка, что включен VPN



**Доработки
на клиентах**



Эксп. 1/8: подбор качества под viewport

Было

Качество видео
выбиралось только
исходя из измеренной
ширины канала

Стало

Помимо ширины
канала принимается
во внимание размер
области отображения
(viewport), чтобы
не качать излишне
большое видео

Результат по A/B

-5%

Столлы снизились

Эксп. 2/8: Количество сегментов за один запрос

Было

В EchoPlayer есть настройка сколько сегментов скачивать за один HTTP запрос. По умолчанию == 1

Стало

По А/В выбрали скачивать около 5 с контента за один запрос

В зависимости от нарезки получается 1 или 2 сегмента

Результат по А/В

-15%

Столлы снизились

Эксп. 3/8: выбор эстиматора канала

Было

Для оценки скорости соединения использовался дефолтный эстиматор EchoPlayer, который не учитывает RTT

Запустили A/B/C/D эксперимент с

- exponential weighted average
- sliding window weighted percentile
- sliding window weighted average

Эксп. 3/8: выбор эстиматора канала

Результаты

Победил sliding
window weighted
percentile

Параметры:
3 семпла
и 128К данных

-15%
Столлы снизились

Эксп. 4/8: адаптивность по аудио

Было

В dash-манифесте указано от трех до пяти аудиодорожек, но выбиралась всегда дорожка с максимальным качеством 256/192 кбит/с.

Могла быть ситуация, что видео 80 кбит/с, а аудио 256 кбит/с.

Стало

Выбираем дорожку так, чтобы битрейт аудио был ниже, чем видео

Результат по A/V

-12%

Столлы снизились

Эксп. 5/8: предзагрузка контента в ленте видео

Было

Два плеера, один показывает текущее видео, второй готовит (подгружает) следующее видео из ленты

Стало

Реализовали предзагрузку с помощью дискового кеша

Результат по A/B

-5%

Столлы снизились, но отрицательно прокрасились продуктовые метрики. Изучаем.

Эксп. 6/8: мобильный Web: включение dash

Было

По историческим
причинам на m.vk.com
игрался не
сегментированный
mp4

Стало

Перевели
m.vk.com на dash

Результат по A/B

-30%

Столлы снизились

-20%

Время первого кадра
снизилось

Эксп. 7/8: кодирование H.264. CRF vs ABR

Было

Для кодирования видео при нарезке использовался подход ABR (Average Bitrate)

Задавали некоторый target bitrate, константный по всему видео

Стало

Перешли на подход CRF (Constant Rate Factor) с верхней границей (Capped CRF)

Результат по A/V

-25-30%

В среднем битрейт снизился при сохранении визуального качества

Эксп. 8/8: лимит на ретрансмиты

odnoklassniki-video-download Request	144µs
com.vkontakte.android DASH_WEBM_AUDIO:LOW	10.64s
odnoklassniki-video-download Request	134µs
odnoklassniki-video-download Request	151µs
com.vkontakte.android DASH_WEBM_VIDEO:SD	10.68s
odnoklassniki-video-download Request	4.05ms
odnoklassniki-video-download Request	302µs
com.vkontakte.android DASH_WEBM_VIDEO:SD	275.69ms

odnoklassniki-video-download Requ...	180µs
com.vkontakte.android DASH_WEBM_VIDE...	23.64s
odnoklassniki-video-download Requ...	200µs
odnoklassniki-video-download Requ...	199µs
com.vkontakte.android DASH_WEBM_AUDI...	23.6s
odnoklassniki-video-download Requ...	240µs
odnoklassniki-video-download Requ...	139µs
com.vkontakte.android DASH_MPD:MOBILE	59.36ms
odnoklassniki-video-download Requ...	995µs

Эксп. 8/8: лимит на ретрансмиты

- ✓ Увеличили количество ретрансмитов до 6
- ✓ Отключили сужение окна между запросами сегментов

```
# /etc/sysctl.conf
```

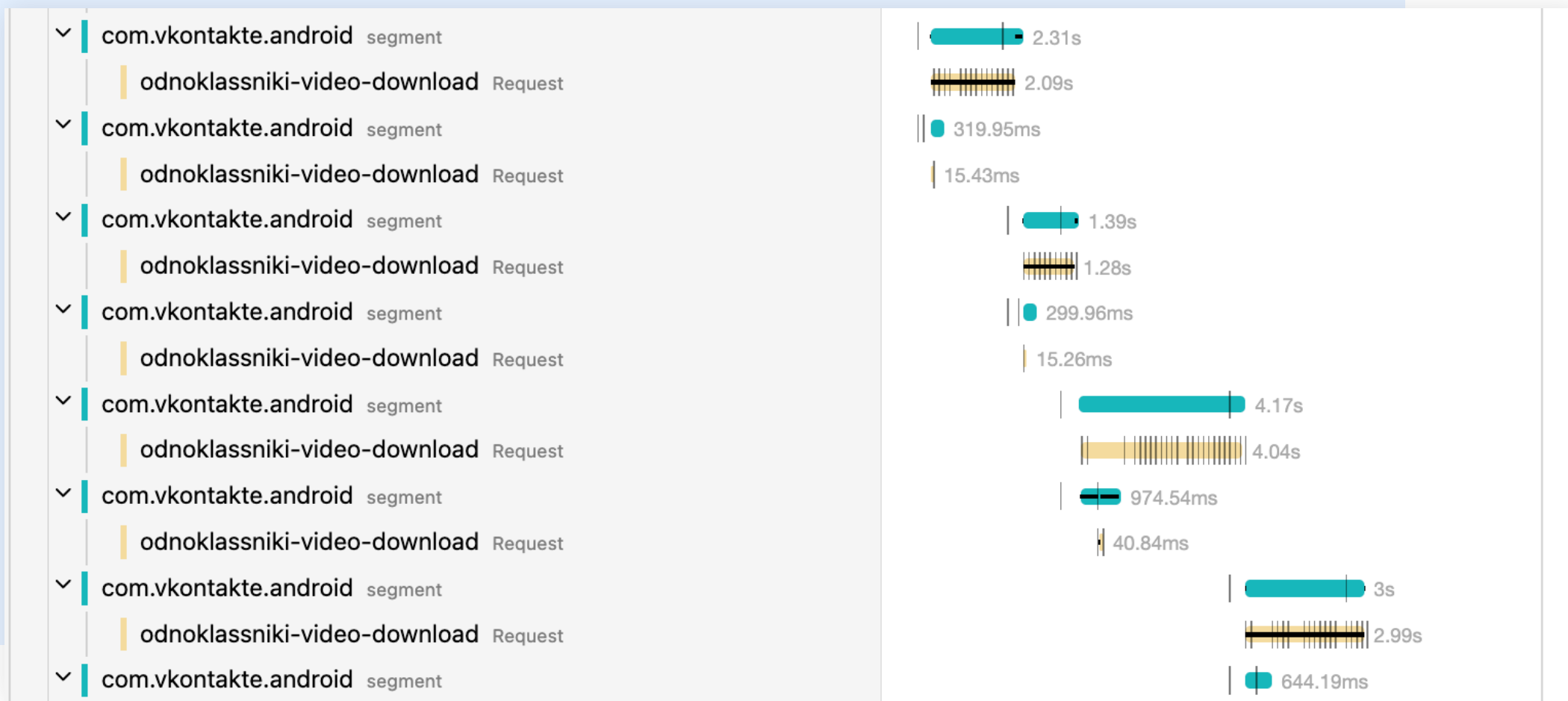
```
net.ipv4.tcp_keepalive_time=60  
net.ipv4.tcp_keepalive_intvl=10  
net.ipv4.tcp_keepalive_probes=6
```

```
# Поменяли в рамках решения проблемы  
net.ipv4.tcp_retries2=6  
net.ipv4.tcp_slow_start_after_idle=0
```

Запланированные эксперименты



Долгие запросы



Перейти на mp4 вместо dash для коротких видео

Сейчас

- Dash используется для всех видео, вне зависимости от их длительности

Планируем

- У нас больше всего столлов именно по коротким видео. Гипотеза, что эстиматор канала не успевает прийти в устойчивое состояние.
- Для коротких видео скачивается mp4 вместо dash

Отмена длинных запросов

Сейчас

- При резком падении скорости сети, эстиматор может существенно запаздывать с оценкой скорости и, как следствие, переключение на более низкое качество происходит не своевременно

Планируем

- Мы можем понять, что скорость упала, если время выполнения запросов стала приближаться к длительности скачиваемого контента. В этом случае мы можем переключиться на качество ниже.

Не сразу переходить на качество выше

Сейчас

- Если на основе оценки канала было принято решение, что можно перейти на качество выше, то это делается сразу

Планируем

- Переходить на качество выше, только если в буфере есть достаточное количество предзагруженных данных текущего качества
- Например, если буфер всего 50 с, то переходить на более высокое качество, если буфер заполнен до 30 с

Доработки и эксперименты

1 Напоминалка о VPN

2 Эксперименты на клиентах

3 Оптимизация битрейта

4 Запланированные экспы

Результаты



Снижение столлов по метрике Empty Buffer/TVT

Платформа	Было, мс/с	Стало, мс/с	Цель, мс/с
Мобильный Веб	12.94	5.61	3
Веб	6.54	3.16	3
Android ВКонтакте	4.65	2.77	3
Android VK Видео	3.86	2.88	3

ИТОГИ

Метрики — чиним то,
что можем измерить

Инструменты для
трассировки и отладки

Эксперименты на группах
пользователей, контролируя
метрики

Решения «в лоб» не всегда
работают, нужна адаптивность

Будем
ВКонтакте!



VideoTech

2024