



OTUS

ОНЛАЙН-ОБРАЗОВАНИЕ

# Онлайн-образование

Не забыть включить запись!





# Меня хорошо видно && слышно?

Ставьте +, если все хорошо  
Напишите в чат, если есть проблемы

# Правила вебинара



Активно участвуем



Задаем вопрос в чат или голосом



Off-topic обсуждаем в Slack #канал группы или #general



Вопросы вижу в чате, могу ответить не сразу

# Рекламная система



Каморин Михаил

Senior Backend developer

Skyeng

[m.v.kamorin@gmail.com](mailto:m.v.kamorin@gmail.com)

# Цели вебинара | После занятия вы сможете

1

Проектировать архитектуру рекламной системы

2

Планировать highload-архитектуру для типичных бизнес-задач, возникающих для рекламной системы

3

Организовывать резервирование и мониторинг для рекламной системы



# 01. Компоненты приложения



# Состав

- CRM

# Состав

- CRM
- Сервис управления показом рекламы

# Состав

- CRM
- Сервис управления показом рекламы
- Сервис сбора статистики показов

# Состав

- CRM
- Сервис управления показом рекламы
- Сервис сбора статистики показов
- Сервис доставки рекламы

# Состав

- CRM
- Сервис управления показом рекламы
- Сервис сбора статистики показов
- Сервис доставки рекламы
- Таргетинг-сервис

# Состав

- CRM
- Сервис управления показом рекламы
- Сервис сбора статистики показов
- Сервис доставки рекламы
- Таргетинг-сервис
- Сервис анализа пользователей

# Состав

- CRM
- Сервис управления показом рекламы
- Сервис сбора статистики показов
- Сервис доставки рекламы
- Таргетинг-сервис
- Сервис анализа пользователей
- ML-алгоритмы

# Состав

- CRM
- Сервис управления показом рекламы
- Сервис сбора статистики показов
- Сервис доставки рекламы
- Таргетинг-сервис
- Сервис анализа пользователей
- ML-алгоритмы
- Поисковый движок (необязательно)

# Сервис управления показом рекламы

- Объявления показываются сотни тысяч и миллионы раз

# Сервис управления показом рекламы

- Объявления показываются сотни тысяч и миллионы раз
- 100% точность учёта количества показов не важна

# Сервис управления показом рекламы

- Объявления показываются сотни тысяч и миллионы раз
- 100% точность учёта количества показов не важна
- Кратковременная недоступность сервиса не критична

# Сервис доставки рекламы

- Сборка объявлений из статичных компонентов

# Сервис доставки рекламы

- Сборка объявлений из статичных компонентов
- Бинарный формат хранения

# Сервис доставки рекламы

- Сборка объявлений из статичных компонентов
- Бинарный формат хранения
- Проблема с шириной канала

# Сервис доставки рекламы

- Сборка объявлений из статичных компонентов
- Бинарный формат хранения
- Проблема с шириной канала
- Критичен для обновления рекламы

# Сервис доставки рекламы

- Сборка объявлений из статичных компонентов
- Бинарный формат хранения
- Проблема с шириной канала
- Критичен для обновления рекламы
- Не полагаемся на third-party

# Таргетинг-сервис: виды таргетинга

- Проспектинг и ретаргетинг

# Таргетинг-сервис: виды таргетинга

- Проспектинг и ретаргетинг
- Контекстный

# Таргетинг-сервис: виды таргетинга

- Проспектинг и ретаргетинг
- Контекстный
- Социально-демографический

# Таргетинг-сервис: виды таргетинга

- Проспектинг и ретаргетинг
- Контекстный
- Социально-демографический
- Географический

# Таргетинг-сервис: виды таргетинга

- Проспектинг и ретаргетинг
- Контекстный
- Социально-демографический
- Географический
- Временной

# Таргетинг-сервис: виды таргетинга

- Проспектинг и ретаргетинг
- Контекстный
- Социально-демографический
- Географический
- Временной
- Поведенческий

# Таргетинг-сервис: задачи

- Сегментация пользователей

# Таргетинг-сервис: задачи

- Сегментация пользователей
- Определение контекста

# Таргетинг-сервис: задачи

- Сегментация пользователей
- Определение контекста
- Связь контекста и сегментов

# Анализ пользователей и ML-алгоритмы

- Сбор открытых данных по социальным сетям

# Анализ пользователей и ML-алгоритмы

- Сбор открытых данных по социальным сетям
- Категоризация поисковых запросов

# Анализ пользователей и ML-алгоритмы

- Сбор открытых данных по социальным сетям
- Категоризация поисковых запросов
- Анализ возможности допродаж

# Анализ пользователей и ML-алгоритмы

- Сбор открытых данных по социальным сетям
- Категоризация поисковых запросов
- Анализ возможности допродаж
- Анализ связанных сегментов



# 02. Highload бизнес-процессы



# Управление показом рекламы, архитектура

- Количество показов может достигать десятков и сотен тысяч запросов в **секунду**

# Управление показом рекламы, архитектура

- Количество показов может достигать десятков и сотен тысяч запросов в **секунду**
- Нужно отдавать объявления и протоколировать это с указанной частотой

# Управление показом рекламы, архитектура

- Количество показов может достигать десятков и сотен тысяч запросов в **секунду**
- Нужно отдавать объявления и протоколировать это с указанной частотой
- На чтение нам необходима «репликация»:

# Управление показом рекламы, архитектура

- Количество показов может достигать десятков и сотен тысяч запросов в **секунду**
- Нужно отдавать объявления и протоколировать это с указанной частотой
- На чтение нам необходима «репликация»:
  - Реплика определяется партнёром (передаём ему ссылку в коде виджета)

# Управление показом рекламы, архитектура

- Количество показов может достигать десятков и сотен тысяч запросов в **секунду**
- Нужно отдавать объявления и протоколировать это с указанной частотой
- На чтение нам необходима «репликация»:
  - Реплика определяется партнёром (передаём ему ссылку в коде виджета)
  - Нужна информация о пользователе и контексте

# Управление показом рекламы, архитектура

- Количество показов может достигать десятков и сотен тысяч запросов в **секунду**
- Нужно отдавать объявления и протоколировать это с указанной частотой
- На чтение нам необходима «репликация»:
  - Реплика определяется партнёром (передаём ему ссылку в коде виджета)
  - Нужна информация о пользователе и контексте
- На запись шардинг не поможет, т.к. нужна онлайн-статистика

# Определение контекста и пользователя

- Не можем ходить в БД за этими данными

# Определение контекста и пользователя

- Не можем ходить в БД за этими данными
- Контекст может быть зафиксирован в ссылке для партнёра

# Определение контекста и пользователя

- Не можем ходить в БД за этими данными
- Контекст может быть зафиксирован в ссылке для партнёра
- Информацию о пользователе берём из куки/локального хранилища, если её нет, то не относим к сегментам, а передаём данные в сервис анализа

# Сбор статистики показов

- Не можем писать поток нужной интенсивности в БД

# Сбор статистики показов

- Не можем писать поток нужной интенсивности в БД
- Возможные инструменты: statsd, redis

# Сбор статистики показов

- Не можем писать поток нужной интенсивности в БД
- Возможные инструменты: statsd, redis
- Логическая «репликация» в БД по одному счётчику раз в  $K$  секунд

# Сбор статистики показов

- Не можем писать поток нужной интенсивности в БД
- Возможные инструменты: statsd, redis
- Логическая «репликация» в БД по одному счётчику раз в  $K$  секунд
- Можем потерять данные:

# Сбор статистики показов

- Не можем писать поток нужной интенсивности в БД
- Возможные инструменты: statsd, redis
- Логическая «репликация» в БД по одному счётчику раз в  $K$  секунд
- Можем потерять данные:
  - Интервал записи на диск

# Сбор статистики показов

- Не можем писать поток нужной интенсивности в БД
- Возможные инструменты: statsd, redis
- Логическая «репликация» в БД по одному счётчику раз в K секунд
- Можем потерять данные:
  - Интервал записи на диск
  - UDP (statsd)

# Ротация объявлений

- Частота появления объявления определяется
  - Приоритетом (константа)

# Ротация объявлений

- Частота появления объявления определяется
  - Приоритетом (константа)
  - Сегментацией пользователей (константа)

# Ротация объявлений

- Частота появления объявления определяется
  - Приоритетом (константа)
  - Сегментацией пользователей (константа)
  - Оставшимся количеством показов (переменная величина)

# Ротация объявлений

- Частота появления объявления определяется
  - Приоритетом (константа)
  - Сегментацией пользователей (константа)
  - Оставшимся количеством показов (переменная величина)
- Статистика показов может запаздывать + она собирается не здесь

# Ротация объявлений

- Частота появления объявления определяется
  - Приоритетом (константа)
  - Сегментацией пользователей (константа)
  - Оставшимся количеством показов (переменная величина)
- Статистика показов может запаздывать + она собирается не здесь
- Нужен либо быстрый алгоритм пересчёта частоты, либо больше инстансов сервиса управления показом

# Ротация объявлений, очередь показов

- Для каждой пары (партнёр, сегмент пользователей) строится очередь объявлений

# Ротация объявлений, очередь показов

- Для каждой пары (партнёр, сегмент пользователей) строится очередь объявлений
- Новые объявления забираются с головы нужной очереди

# Ротация объявлений, очередь показов

- Для каждой пары (партнёр, сегмент пользователей) строится очередь объявлений
- Новые объявления забираются с головы нужной очереди
- В конец добавляются объявления на базе рассчитанной частоты

# Ротация объявлений, очередь показов

- Для каждой пары (партнёр, сегмент пользователей) строится очередь объявлений
- Новые объявления забираются с головы нужной очереди
- В конец добавляются объявления на базе рассчитанной частоты
- Очередь может разбираться быстрее, чем добавляться

# Ротация объявлений, очередь показов

- Для каждой пары (партнёр, сегмент пользователей) строится очередь объявлений
- Новые объявления забираются с головы нужной очереди
- В конец добавляются объявления на базе рассчитанной частоты
- Очередь может разбираться быстрее, чем добавляться
  - Нужно уметь забирать объявление напрямую, без очереди

# Ротация объявлений, очередь показов

- Для каждой пары (партнёр, сегмент пользователей) строится очередь объявлений
- Новые объявления забираются с головы нужной очереди
- В конец добавляются объявления на базе рассчитанной частоты
- Очередь может разбираться быстрее, чем добавляться
  - Нужно уметь забирать объявление напрямую, без очереди
- Очередь может заполняться слишком быстро (нужно ограничение)

# Доставка объявлений

- Хотим максимально быстро пускать объявление в ротацию, поэтому обычное хранилище не подходит

# Доставка объявлений


- Хотим максимально быстро пускать объявление в ротацию, поэтому обычное хранилище не подходит
- Суммарный объём бинарных данных со всеми объявлениями – десятки или единицы сотен мегабайт

# Доставка объявлений

- Хотим максимально быстро пускать объявление в ротацию, поэтому обычное хранилище не подходит
- Суммарный объём бинарных данных со всеми объявлениями – десятки или единицы сотен мегабайт
- В гигабитный канал проходят единицы отправок в секунду

# Доставка объявлений

- Хотим максимально быстро пускать объявление в ротацию, поэтому обычное хранилище не подходит
- Суммарный объём бинарных данных со всеми объявлениями – десятки или единицы сотен мегабайт
- В гигабитный канал проходят единицы отправок в секунду
- Нужен механизм логической репликации бинарных данных, чтобы отправлять инкрементальные дампы

The image features a blue-tinted aerial view of a dense city skyline, likely New York City, with numerous skyscrapers. A semi-transparent blue band with a white network pattern of lines and dots runs horizontally across the middle of the image. The text is centered within this band.

# 03. Резервирование и мониторинг

# Критичные сервисы

- Сервис управления показом рекламы
  - Кратковременная недоступность некритична

# Критичные сервисы

- Сервис управления показом рекламы
  - Кратковременная недоступность не критична
  - Поскольку сервисы взаимозаменяемые, можем легко подменять один сервис другим через балансировщик

# Критичные сервисы

- Сервис управления показом рекламы
  - Кратковременная недоступность не критична
  - Поскольку сервисы взаимозаменяемые, можем легко подменять один сервис другим через балансировщик
  - Штатный режим - **без** балансировщика

# Критичные сервисы

- Сервис управления показом рекламы
  - Кратковременная недоступность не критична
  - Поскольку сервисы взаимозаменяемые, можем легко подменять один сервис другим через балансировщик
  - Штатный режим - **без** балансировщика
- Сервис доставки рекламы
  - Критичен для обновления набора объявлений

# Критичные сервисы

- Сервис управления показом рекламы
  - Кратковременная недоступность не критична
  - Поскольку сервисы взаимозаменяемые, можем легко подменять один сервис другим через балансировщик
  - Штатный режим - **без** балансировщика
- Сервис доставки рекламы
  - Критичен для обновления набора объявлений
  - Недоступность в течение нескольких десятков минут скажется только на оперативности ввода объявления в ротацию

# Критичные сервисы

- Сервис анализа пользователей
  - При недоступности всем неизвестным пользователям будем показывать общую рекламу (более дешёвую)

# Критичные сервисы

- Сервис анализа пользователей
  - При недоступности всем неизвестным пользователям будем показывать общую рекламу (более дешёвую)
  - Имеет смысл резервировать, при условии:
    - Много неизвестных пользователей (вошли на новые ресурсы или рынки)

# Критичные сервисы

- Сервис анализа пользователей
  - При недоступности всем неизвестным пользователям будем показывать общую рекламу (более дешёвую)
  - Имеет смысл резервировать, при условии:
    - Много неизвестных пользователей (вошли на новые ресурсы или рынки)
    - Скорость сегментирования выше, чем средний time on site (успеем показать хотя бы одно таргетированное объявление)

# Критичные сервисы

- Сервис сбора статистики показов
  - Недоступность ведёт к неточному учёту

# Критичные сервисы

- Сервис сбора статистики показов
  - Недоступность ведёт к неточному учёту
  - Потери пропорциональны времени простоя, легко посчитать экономику

# Критичные сервисы

- Сервис сбора статистики показов
  - Недоступность ведёт к неточному учёту
  - Потери пропорциональны времени простоя, легко посчитать экономику
  - Проблемы горизонтального масштабирования:
    - Обязательно отдельные входящие запросы

# Критичные сервисы

- Сервис сбора статистики показов
  - Недоступность ведёт к неточному учёту
  - Потери пропорциональны времени простоя, легко посчитать экономику
  - Проблемы горизонтального масштабирования:
    - Обязательно отдельные входящие запросы
    - Расхождение данных

# Служебный мониторинг

- Время обновления данных об объявлениях

# Служебный мониторинг

- Время обновления данных об объявлениях
- Пропускная способность сервиса сбора статистики о показах

# Служебный мониторинг

- Время обновления данных об объявлениях
- Пропускная способность сервиса сбора статистики о показах
- Суммарное время недоступности сервиса сбора статистики о показах

# Служебный мониторинг

- Время обновления данных об объявлениях
- Пропускная способность сервиса сбора статистики о показах
- Суммарное время недоступности сервиса сбора статистики о показах
- Длина очередей объявлений для показа

# Бизнес-метрики

- Неоплаченные показы

# Бизнес-метрики

- Неоплаченные показы
- Качество сегментации

# Бизнес-метрики

- Неоплаченные показы
- Качество сегментации
- Качество таргетинга

# Следующий вебинар

**Тема:** Почтовый сервис



Среда 26.02.19 в 20.00



Ссылка на вебинар будет в ЛК за 15 минут



Материалы к занятию  
в ЛК – можно изучать

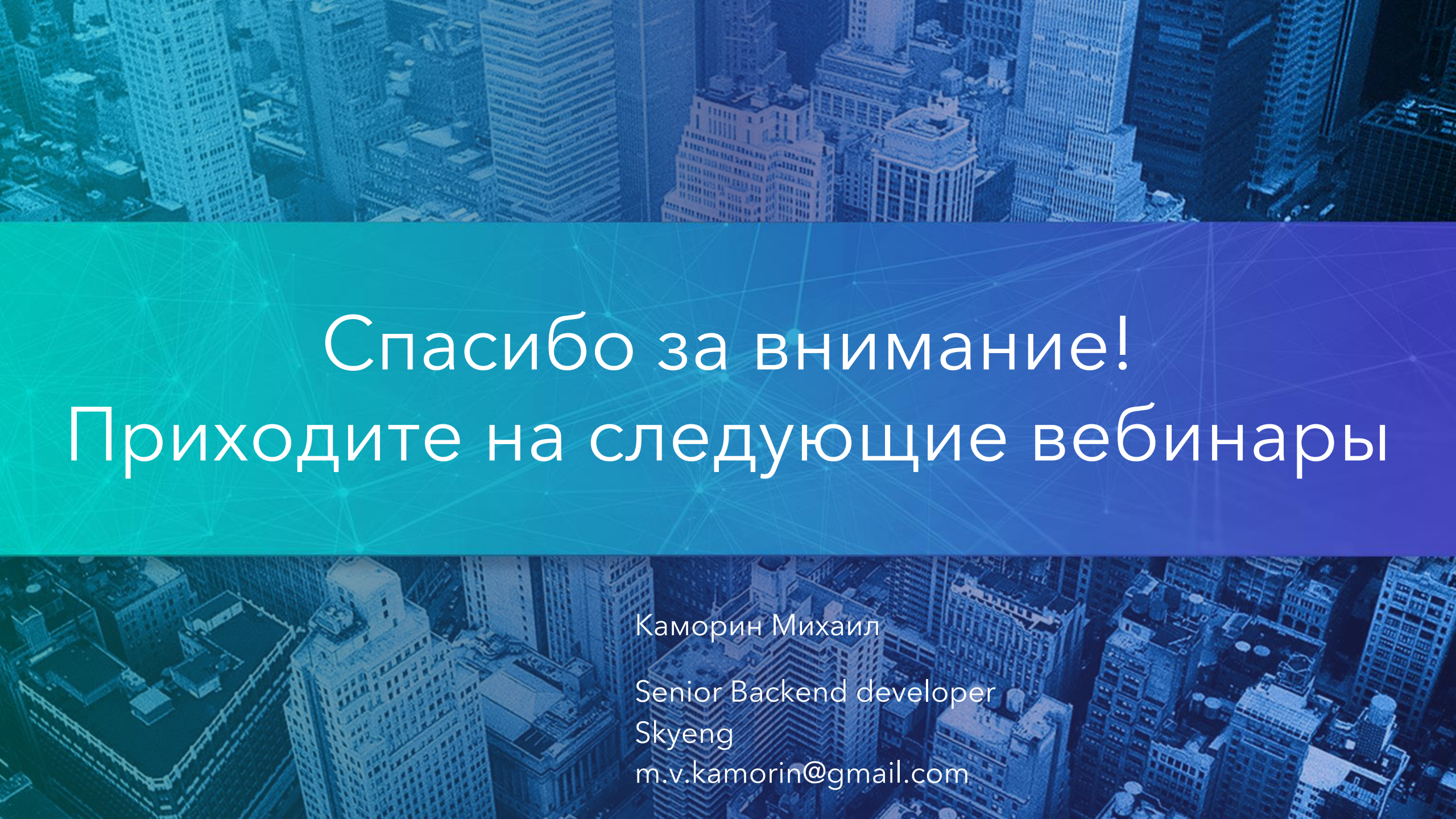


Обязательный  
материал обозначен  
красной лентой



Заполните, пожалуйста,  
опрос о занятии по ссылке в чате





Спасибо за внимание!  
Приходите на следующие вебинары

Каморин Михаил

Senior Backend developer  
Skyeng

[m.v.kamorin@gmail.com](mailto:m.v.kamorin@gmail.com)