



O T U S

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

Почтовый сервис



Каморин Михаил

Senior Backend developer

Skyeng

m.v.kamorin@gmail.com

Цели вебинара | После занятия вы сможете

1

Проектировать архитектуру почтового сервиса

2

Планировать highload-архитектуру для типичных бизнес-задач, возникающих для почтового сервиса

3

Организовывать резервирование и мониторинг для почтового сервиса



01. Компоненты приложения



Состав

- Mail User Agent

Состав

- Mail User Agent
- Mail Transport Agent

Состав

- Mail User Agent
- Mail Transport Agent
- Хранилище данных

Состав

- Mail User Agent
- Mail Transport Agent
- Хранилище данных
- Антивирус

Состав

- Mail User Agent
- Mail Transport Agent
- Хранилище данных
- Антивирус
- Рендерер писем

Состав

- Mail User Agent
- Mail Transport Agent
- Хранилище данных
- Антивирус
- Рендерер писем
- ML-алгоритмы

Mail User Agent

- POP3 устарел, но всё ещё активно используется

Mail User Agent

- POP3 устарел, но всё ещё активно используется
- IMAP создаёт большую нагрузку в десктопных и мобильных клиентах при редких запусках

Mail User Agent

- POP3 устарел, но всё ещё активно используется
- IMAP создаёт большую нагрузку в десктопных и мобильных клиентах при редких запусках
- Проблемы с MFA

Mail Transport Agent

- Массовые рассылки

Mail Transport Agent

- Массовые рассылки
- Разграничение внешнего и внутрисервисного трафика

Mail Transport Agent

- Массовые рассылки
- Разграничение внешнего и внутрисервисного трафика
- Защищённые соединения

Антивирус

- Несколько движков

Антивирус

- Несколько движков
- Декомпрессия архивов

Антивирус

- Несколько движков
- Декомпрессия архивов
- Брутфорс паролей

Антивирус

- Несколько движков
- Декомпрессия архивов
- Брутфорс паролей
- Анализ мультимедиа-файлов

Хранилище данных

- БД для тел писем (сообщений) + движок поверх БД для удобства

Хранилище данных

- БД для тел писем (сообщений) + движок поверх БД для удобства
- Отдельное хранилище для вложений и embedded картинок

Хранилище данных

- БД для тел писем (сообщений) + движок поверх БД для удобства
- Отдельное хранилище для вложений и embedded картинок
- Отдельные пайплайны на обработку сообщений и вложений

Рендерер писем

- Максимальная поддержка HTML и CSS

Рендерер писем

- Максимальная поддержка HTML и CSS
- Необходима безопасность

Рендерер писем

- Максимальная поддержка HTML и CSS
- Необходима безопасность
- Много форматов внутреннего содержимого

Рендерер писем

- Максимальная поддержка HTML и CSS
- Необходима безопасность
- Много форматов внутреннего содержимое
- Ссылки на хранилище

ML-алгоритмы

- Антиспам-фильтр

ML-алгоритмы

- Антиспам-фильтр
- Блокировка спам-аккаунтов

ML-алгоритмы

- Антиспам-фильтр
- Блокировка спам-аккаунтов
- Анализ вложений и ссылок

ML-алгоритмы

- Антиспам-фильтр
- Блокировка спам-аккаунтов
- Анализ вложений и ссылок
- Антифишинг



02. Highload бизнес-процессы



Входящее письмо

- Письмо может содержать вложения + различные виды упаковки

Входящее письмо

- Письмо может содержать вложения + различные виды упаковки
- Препроцессинг заголовков

Входящее письмо

- Письмо может содержать вложения + различные виды упаковки
- Препроцессинг заголовков
- Проверка чёрных и белых списков

Входящее письмо

- Письмо может содержать вложения + различные виды упаковки
- Препроцессинг заголовков
- Проверка чёрных и белых списков
- Делим письмо на непосредственно сообщение и вложения

Входящее письмо

- Письмо может содержать вложения + различные виды упаковки
- Препроцессинг заголовков
- Проверка чёрных и белых списков
- Делим письмо на непосредственно сообщение и вложения
- Сообщение и вложения кладём в разные хранилища

Входящее письмо

- Письмо может содержать вложения + различные виды упаковки
- Препроцессинг заголовков
- Проверка чёрных и белых списков
- Делим письмо на непосредственно сообщение и вложения
- Сообщение и вложения кладём в разные хранилища
- Обрабатываем сообщение и вложения параллельно в разных пайплайнах

Обработка вложения

- Антивирус
 - Декомпрессия архива, брутфорс пароля

Обработка вложения

- Антивирус
 - Декомпрессия архива, брутфорс пароля
 - Анализ исполняемых файлов

Обработка вложения

- Антивирус
 - Декомпрессия архива, брутфорс пароля
 - Анализ исполняемых файлов
 - Анализ мультимедиа-файлов

Обработка вложения

- Антивирус
 - Декомпрессия архива, брутфорс пароля
 - Анализ исполняемых файлов
 - Анализ мультимедиа-файлов
- Дополнительный ML-анализ

Обработка вложения

- Антивирус
 - Декомпрессия архива, брутфорс пароля
 - Анализ исполняемых файлов
 - Анализ мультимедиа-файлов
- Дополнительный ML-анализ
- Если на каком-то этапе находим проблему, то вложение блокируется в хранилище

Обработка сообщения

- Антиспам

Обработка сообщения

- Антиспам
- Антифишинг

Обработка сообщения

- Антиспам
- Антифишинг
- Дополнительный анализ для спецслужб

Обработка сообщения

- Антиспам
- Антифишинг
- Дополнительный анализ для спецслужб
- Анализ текста (Google, Yandex) для сегментации пользователя и получения контекста

Mail User Agent, масштабирование

- Для обслуживания пользователей масштабируем backend горизонтально

Mail User Agent, масштабирование

- Для обслуживания пользователей масштабируем backend горизонтально
- Запрос на получение списка писем лёгкий, но одновременно с ним приходит много пользователей

Mail User Agent, масштабирование

- Для обслуживания пользователей масштабируем backend горизонтально
- Запрос на получение списка писем лёгкий, но одновременно с ним приходит много пользователей
- В результате получаем M серверов * N пользователей на каждом. Backend справляется, но количество соединений к БД становится очень большим

Mail User Agent, масштабирование

- **Решение:** шардируем БД вместе с backend

Mail User Agent, масштабирование

- **Решение:** шардируем БД вместе с backend
- **Проблемы:**
 - Простой алгоритм и ключ не отражают неравномерность заполнения БД

Mail User Agent, масштабирование

- **Решение:** шардируем БД вместе с backend
- **Проблемы:**
 - Простой алгоритм и ключ не отражают неравномерность заполнения БД
 - Не можем управлять активностью (IMAP-клиенты постоянно шлют запросы)

Mail User Agent, масштабирование

- **Решение:** шардируем БД вместе с backend
- **Проблемы:**
 - Простой алгоритм и ключ не отражают неравномерность заполнения БД
 - Не можем управлять активностью (IMAP-клиенты постоянно шлют запросы)
 - Сложности при записи сообщений

Mail User Agent, масштабирование

- **Решение:** шардируем БД вместе с backend
- **Проблемы:**
 - Простой алгоритм и ключ не отражают неравномерность заполнения БД
 - Не можем управлять активностью (IMAP-клиенты постоянно шлют запросы)
 - Сложности при записи сообщений
 - Сложный рещардинг

Mail User Agent, масштабирование

- **Решение:** мультиплексор соединений между БД и backend инстансами

Mail User Agent, масштабирование

- **Решение:** мультиплексор соединений между БД и backend инстансами
- **Преимущества:**
 - Нет проблемы с шардированием

Mail User Agent, масштабирование


- **Решение:** мультиплексор соединений между БД и backend инстансами
- **Преимущества:**
 - Нет проблемы с шардированием
 - Всё ещё можем сделать кластер БД

Mail User Agent, масштабирование

- **Решение:** мультиплексор соединений между БД и backend инстансами
- **Преимущества:**
 - Нет проблемы с шардированием
 - Всё ещё можем сделать кластер БД
 - Легко решардируем

Mail User Agent, масштабирование

- **Решение:** мультиплексор соединений между БД и backend инстансами
- **Преимущества:**
 - Нет проблемы с шардированием
 - Всё ещё можем сделать кластер БД
 - Легко решардируем
- **Недостатки:**
 - Задержка при обработке запроса на ожидание канала в мультиплексоре

The image features a blue-tinted aerial view of a 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

- Хранилище сообщений
 - Потеря данных критична – тройное резервирование

Критичные сервисы

- Хранилище сообщений
 - Потеря данных критична - тройное резервирование
 - Дубликаты необработанных сообщений во внешнее файловое хранилище

Критичные сервисы

- Хранилище сообщений
 - Потеря данных критична - тройное резервирование
 - Дубликаты необработанных сообщений во внешнее файловое хранилище
 - Локальная репликация

Критичные сервисы

- Хранилище сообщений
 - Потеря данных критична - тройное резервирование
 - Дубликаты необработанных сообщений во внешнее файловое хранилище
 - Локальная репликация
 - Асинхронная репликация реплики в другой ЦОД

Критичные сервисы

- Хранилище сообщений
 - Потеря данных критична - тройное резервирование
 - Дубликаты необработанных сообщений во внешнее файловое хранилище
 - Локальная репликация
 - Асинхронная репликация реплики в другой ЦОД
- Хранилище вложений
 - Облачное хранилище (обсудим в следующий раз)

Критичные сервисы

- Хранилище сообщений
 - Потеря данных критична - тройное резервирование
 - Дубликаты необработанных сообщений во внешнее файловое хранилище
 - Локальная репликация
 - Асинхронная репликация реплики в другой ЦОД
- Хранилище вложений
 - Облачное хранилище (обсудим в следующий раз)
- Заблокированный контент не удаляем в течение срока по соглашению

Интерфейсы Mail User Agent

- Самый важный – POP3, должен работать всегда
 - Балансировщик перенаправляет POP3-трафик при недоступности шарда на другой

Интерфейсы Mail User Agent

- Самый важный – POP3, должен работать всегда
 - Балансировщик перенаправляет POP3-трафик при недоступности шарда на другой
- Следующий – web-интерфейс

Интерфейсы Mail User Agent

- Самый важный – POP3, должен работать всегда
 - Балансировщик перенаправляет POP3-трафик при недоступности шарда на другой
- Следующий – web-интерфейс
- Затем IMAP
 - При любых проблемах можем игнорировать

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб
 - Белые/чёрные списки

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб
 - Белые/чёрные списки
 - Препроцессинг заголовков

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб
 - Белые/чёрные списки
 - Препроцессинг заголовков
 - Антивирус

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб
 - Белые/чёрные списки
 - Препроцессинг заголовков
 - Антивирус
 - Антиспам

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб
 - Белые/чёрные списки
 - Препроцессинг заголовков
 - Антивирус
 - Антиспам
 - Антифишинг

Обработка входящих

- В порядке убывания важности:
 - Анализ для спецслужб
 - Белые/чёрные списки
 - Препроцессинг заголовков
 - Антивирус
 - Антиспам
 - Антифишинг
 - Всё остальное

Служебный мониторинг

- Нагрузка на backend-инстансах

Служебный мониторинг

- Нагрузка на backend-инстансах
- Отдельно % ошибок по каждому из протоколов

Служебный мониторинг

- Нагрузка на backend-инстансах
- Отдельно % ошибок по каждому из протоколов
- Health-check всех обрабатывающих сервисов

Служебный мониторинг

- Нагрузка на backend-инстансах
- Отдельно % ошибок по каждому из протоколов
- Health-check всех обрабатывающих сервисов
- Нагрузка на каждом из этапов обработки

Служебный мониторинг

- Нагрузка на backend-инстансах
- Отдельно % ошибок по каждому из протоколов
- Health-check всех обрабатывающих сервисов
- Нагрузка на каждом из этапов обработки
- Проблемы с загрузкой вложений

Бизнес-метрики

- Объём данных по пользователям

Бизнес-метрики

- Объём данных по пользователям
- % блокировок писем

Бизнес-метрики

- Объём данных по пользователям
- % блокировок писем
- % блокировок аккаунтов

Следующий вебинар

Тема: Облачное хранилище



Пятница 28.02.19 в 20.00



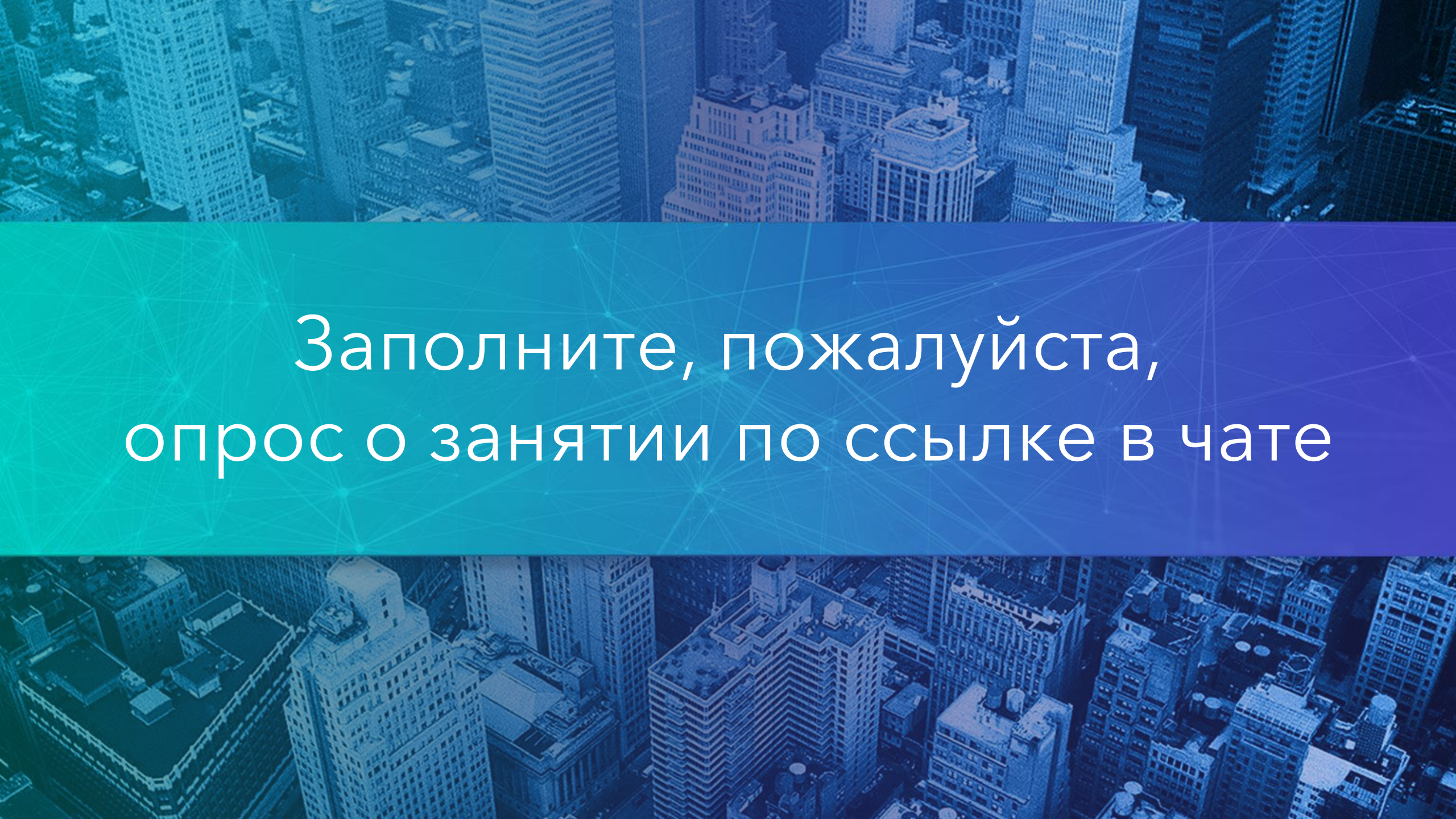
Ссылка на вебинар будет в ЛК за 15 минут



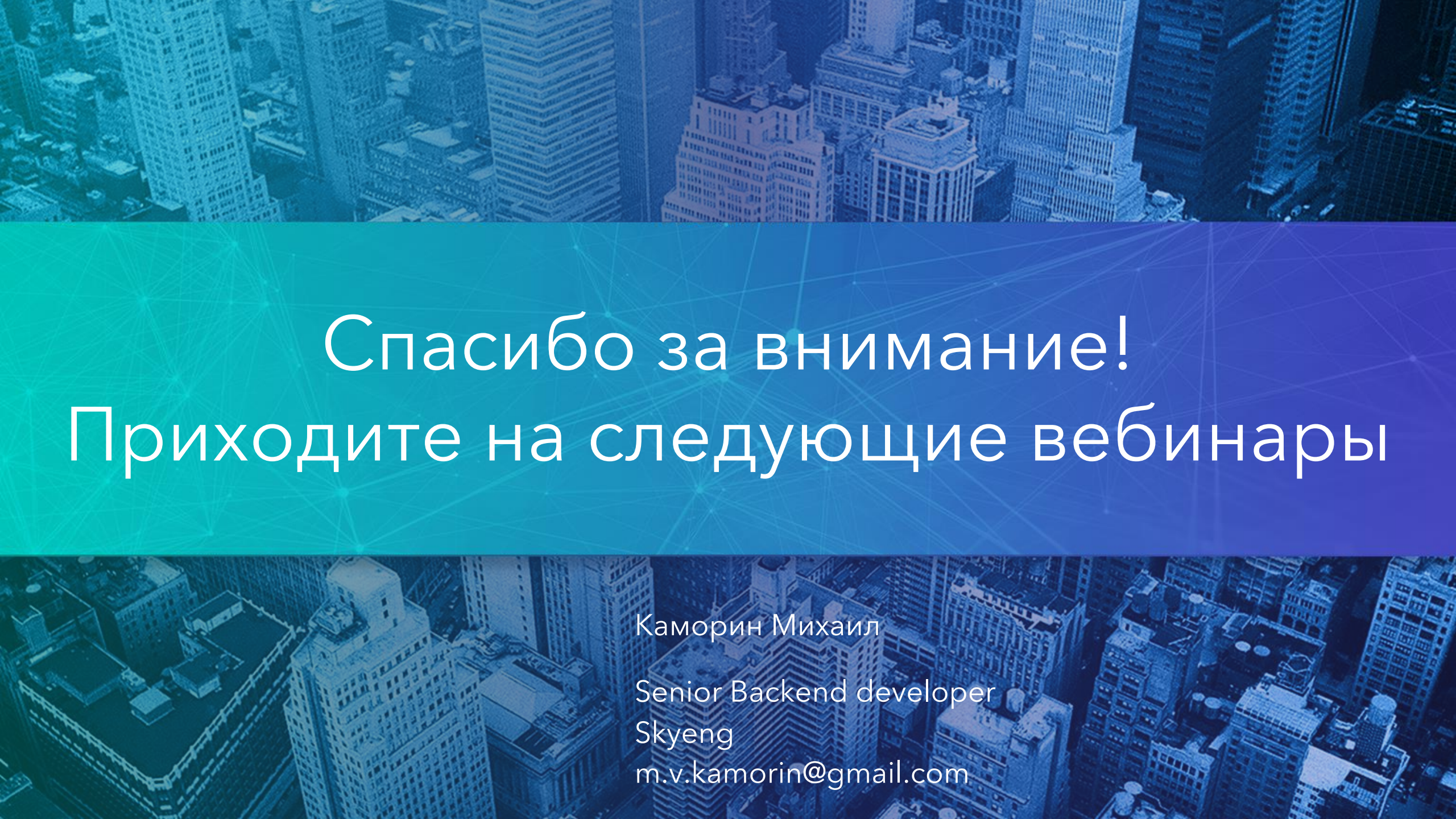
Материалы к занятию
в ЛК – можно изучать



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