



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

Облачное хранилище



Каморин Михаил

Senior Backend developer

Skyeng

m.v.kamorin@gmail.com

Цели вебинара | После занятия вы сможете

1

Проектировать архитектуру
облачного хранилища

2

Планировать highload-архитектуру
для типичных бизнес-задач,
возникающих для облачного хранилища

3

Организовывать резервирование
и мониторинг для облачного хранилища



01. Компоненты приложения



Состав

- Биллинг

Состав

- Биллинг
- Сервис хранения данных

Состав

- Биллинг
- Сервис хранения данных
- Сервис доставки данных

Состав

- Биллинг
- Сервис хранения данных
- Сервис доставки данных
- Пользовательский клиент

Состав

- Биллинг
- Сервис хранения данных
- Сервис доставки данных
- Пользовательский клиент
- Управление доступом

Состав

- Биллинг
- Сервис хранения данных
- Сервис доставки данных
- Пользовательский клиент
- Управление доступом
- ML-алгоритмы

Сервис хранения данных

- Разные типы данных требуют разную организацию хранилища

Сервис хранения данных

- Разные типы данных требуют разную организацию хранилища
- Данные и метаданные храним отдельно

Сервис хранения данных

- Разные типы данных требуют разную организацию хранилища
- Данные и метаданные храним отдельно
- Хотим управлять числом копий идентичных данных сами

Сервис хранения данных

- Разные типы данных требуют разную организацию хранилища
- Данные и метаданные храним отдельно
- Хотим управлять числом копий идентичных данных сами
- Отвечает за контроль доступности данных

Сервис хранения данных

- Разные типы данных требуют разную организацию хранилища
- Данные и метаданные храним отдельно
- Хотим управлять числом копий идентичных данных сами
- Отвечает за контроль доступности данных
- Оптимизация чтения – основной приоритет

Сервис хранения данных

- Разные типы данных требуют разную организацию хранилища
- Данные и метаданные храним отдельно
- Хотим управлять числом копий идентичных данных сами
- Отвечает за контроль доступности данных
- Оптимизация чтения - основной приоритет
- Разные уровни доступа

Сервис доставки данных

- Разные типы данных требуют разные протоколы доставки

Сервис доставки данных

- Разные типы данных требуют разные протоколы доставки
- Требуется кэш с интеллектуальной инвалидацией

Сервис доставки данных

- Разные типы данных требуют разные протоколы доставки
- Требуется кэш с интеллектуальной инвалидацией
- Географическая балансировка

Сервис доставки данных

- Разные типы данных требуют разные протоколы доставки
- Требуется кэш с интеллектуальной инвалидацией
- Географическая балансировка
- Не контролирует доступ к данным

Управление доступом

- Разные типы доступа
 - Общие файлы

Управление доступом

- Разные типы доступа
 - Общие файлы
 - Доступ по ссылке

Управление доступом

- Разные типы доступа
 - Общие файлы
 - Доступ по ссылке
 - Экспирация по времени

Управление доступом

- Разные типы доступа
 - Общие файлы
 - Доступ по ссылке
 - Экспирация по времени
 - Экспирация по количеству загрузок

Управление доступом

- Разные типы доступа
 - Общие файлы
 - Доступ по ссылке
 - Экспирация по времени
 - Экспирация по количеству загрузок
 - Комбинированные варианты

Управление доступом

- Разные типы доступа
 - Общие файлы
 - Доступ по ссылке
 - Экспирация по времени
 - Экспирация по количеству загрузок
 - Комбинированные варианты
- Отзыв доступа

Управление доступом

- Разные типы доступа
 - Общие файлы
 - Доступ по ссылке
 - Экспирация по времени
 - Экспирация по количеству загрузок
 - Комбинированные варианты
- Отзыв доступа
- Доступ к хранилищу немгновенного доступа

Пользовательский клиент

- Виртуальная файловая система

Пользовательский клиент

- Виртуальная файловая система
- Проблема частичной синхронизации

Пользовательский клиент

- Виртуальная файловая система
- Проблема частичной синхронизации
- Конфликт версий

ML-алгоритмы

- Незашифрованный контент используется как источник данных для обучения
 - Распознавание лиц

ML-алгоритмы

- Незашифрованный контент используется как источник данных для обучения
 - Распознавание лиц
 - Задачи кластеризации

ML-алгоритмы

- Незашифрованный контент используется как источник данных для обучения
 - Распознавание лиц
 - Задачи кластеризации
 - Машинное зрение



02. Highload бизнес-процессы



Блочное хранилище

- Хранение данных, не нуждающихся в файловой системе:
 - Образы разделов

Блочное хранилище

- Хранение данных, не нуждающихся в файловой системе:
 - Образы разделов
 - Образы дисков виртуальных машин

Блочное хранилище

- Хранение данных, не нуждающихся в файловой системе:
 - Образы разделов
 - Образы дисков виртуальных машин
 - Тома для контейнеров

Блочное хранилище

- Хранение данных, не нуждающихся в файловой системе:
 - Образы разделов
 - Образы дисков виртуальных машин
 - Тома для контейнеров
- Запись-чтение довольно сбалансированы

Блочное хранилище

- Хранение данных, не нуждающихся в файловой системе:
 - Образы разделов
 - Образы дисков виртуальных машин
 - Тома для контейнеров
- Запись-чтение довольно сбалансированы
- Сам образ как данные - не нужен, нужна возможность его использовать в соответствующем решении

Блочное хранилище

- Хранилище метаданных просто содержит внутренний путь до сервера с блоком

Блочное хранилище

- Хранилище метаданных просто содержит внутренний путь до сервера с блоком
- Протокол доступа предоставляет не данные, а интерфейс, например, для монтирования раздела в виртуальную машину/контейнер

Блочное хранилище

- Хранилище метаданных просто содержит внутренний путь до сервера с блоком
- Протокол доступа предоставляет не данные, а интерфейс, например, для монтирования раздела в виртуальную машину/контейнер
- Управление доступом на уровне хранилища метаданных без особых проблем

Хранилище картинок

- Выделяется отдельно, т.к. обычно с картинками работают ML-алгоритмы

Хранилище картинок

- Выделяется отдельно, т.к. обычно с картинками работают ML-алгоритмы
- После записи практически не изменяются, большой дисбаланс на чтение

Хранилище картинок

- Выделяется отдельно, т.к. обычно с картинками работают ML-алгоритмы
- После записи практически не изменяются, большой дисбаланс на чтение
- Требуется кэш

Хранилище картинок

- Выделяется отдельно, т.к. обычно с картинками работают ML-алгоритмы
- После записи практически не изменяются, большой дисбаланс на чтение
- Требуется кэш
- Данные могут дублироваться для повышения производительности

Хранилище картинок

- Выделяется отдельно, т.к. обычно с картинками работают ML-алгоритмы
- После записи практически не изменяются, большой дисбаланс на чтение
- Требуется кэш
- Данные могут дублироваться для повышения производительности
- Данные могут быть общими для нескольких пользователей (выясняется на базе хэша + служебной метаданных)

Хранилище мультимедиа

- Пишется один раз, читается много раз, часто параллельно

Хранилище мультимедиа

- Пишется один раз, читается много раз, часто параллельно
- Потокоевое видео может сильно нагружать канал чтения

Хранилище мультимедиа

- Пишется один раз, читается много раз, часто параллельно
- Поток видео может сильно нагружать канал чтения
- Постепенная деградация + масштабирование «на лету»

Хранилище мультимедиа

- Пишется один раз, читается много раз, часто параллельно
- Поток видео может сильно нагружать канал чтения
- Постепенная деградация + масштабирование «на лету»
- Проблема с определением возможности удаления «лишних» копий

Объектное хранилище

- Для пользователя выглядит как файловая система, на самом деле – хранилище объектов (noSQL-термин «документ»)

Объектное хранилище

- Для пользователя выглядит как файловая система, на самом деле – хранилище объектов (noSQL-термин «документ»)
- Часто используется для бэкапа и синхронизации, т.е. пишется редко, читается ещё реже

Объектное хранилище

- Для пользователя выглядит как файловая система, на самом деле – хранилище объектов (noSQL-термин «документ»)
- Часто используется для бэкапа и синхронизации, т.е. пишется редко, читается ещё реже
- Редко есть возможность разделить один документ между пользователями, редко нужно масштабирование чтения

Инвалидация кэша

- **Проблема:** пользователь сделал доступ по ссылке, а затем удалил файл. Из-за наличия кэша файл может быть доступен по ссылке ещё довольно долго (до 24 часов)

Инвалидация кэша

- **Проблема:** пользователь сделал доступ по ссылке, а затем удалил файл. Из-за наличия кэша файл может быть доступен по ссылке ещё довольно долго (до 24 часов)
- **Решение:** при удалении объекта асинхронно выполнять
 - Сброс кэша по полному «пути» на backend, работавшем с пользователем

Инвалидация кэша


- **Проблема:** пользователь сделал доступ по ссылке, а затем удалил файл. Из-за наличия кэша файл может быть доступен по ссылке ещё довольно долго (до 24 часов)
- **Решение:** при удалении объекта асинхронно выполнять
 - Сброс кэша по полному «пути» на backend, работавшем с пользователем
 - Сброс кэша листинга для всех промежуточных узлов «пути» на backend, работавшем с пользователем

Инвалидация кэша

- **Проблема:** пользователь сделал доступ по ссылке, а затем удалил файл. Из-за наличия кэша файл может быть доступен по ссылке ещё довольно долго (до 24 часов)
- **Решение:** при удалении объекта асинхронно выполнять
 - Сброс кэша по полному «пути» на backend, работавшем с пользователем
 - Сброс кэша листинга для всех промежуточных узлов «пути» на backend, работавшем с пользователем
 - Аналогичные запросы на все другие backend

Инвалидация кэша

- **Проблема:** пользователь сделал доступ по ссылке, а затем удалил файл. Из-за наличия кэша файл может быть доступен по ссылке ещё довольно долго (до 24 часов)
- **Решение:** при удалении объекта асинхронно выполнять
 - Сброс кэша по полному «пути» на backend, работавшем с пользователем
 - Сброс кэша листинга для всех промежуточных узлов «пути» на backend, работавшем с пользователем
 - Аналогичные запросы на все другие backend
- **Результат:** Мгновенного удаления не будет, но время доступности сильно сократится



03. Резервирование и мониторинг

Критичные сервисы

- Сервис доставки данных
 - Географический шардинг + возможность перенаправления трафика между регионами

Критичные сервисы

- Сервис доставки данных
 - Географический шардинг + возможность перенаправления трафика между регионами
 - Решардинг внутри региона

Критичные сервисы

- Сервис доставки данных
 - Географический шардинг + возможность перенаправления трафика между регионами
 - Решардинг внутри региона
- Сервис хранения данных
 - Резервирование на уровне «железа»

Критичные сервисы

- Сервис доставки данных
 - Географический шардинг + возможность перенаправления трафика между регионами
 - Решардинг внутри региона
- Сервис хранения данных
 - Резервирование на уровне «железа»
 - Кросс-резервирование между регионами

Критичные сервисы

- Управление доступом
 - Критический сервис для корпоративных клиентов

Критичные сервисы

- Управление доступом
 - Критический сервис для корпоративных клиентов
 - Полная реплика БД в каждом регионе

Критичные сервисы

- Управление доступом
 - Критический сервис для корпоративных клиентов
 - Полная реплика БД в каждом регионе
- Пользовательский клиент (backend)
 - Поточная загрузка с возобновлением и кэшированием

Служебный мониторинг

- % избыточности хранилища

Служебный мониторинг

- % избыточности хранилища
- Нагрузка на сервис доставки потокового видео

Служебный мониторинг

- % избыточности хранилища
- Нагрузка на сервис доставки потокового видео
- Нагрузка на региональные кластеры

Служебный мониторинг

- % избыточности хранилища
- Нагрузка на сервис доставки потокового видео
- Нагрузка на региональные кластеры
- Нагрузка на ноды хранилища

Служебный мониторинг

- % избыточности хранилища
- Нагрузка на сервис доставки потокового видео
- Нагрузка на региональные кластеры
- Нагрузка на ноды хранилища
- Доступность нод хранилища

Служебный мониторинг

- % избыточности хранилища
- Нагрузка на сервис доставки потокового видео
- Нагрузка на региональные кластеры
- Нагрузка на ноды хранилища
- Доступность нод хранилища
- % данных, доступных в единственном экземпляре (логически)

Бизнес-метрики

- Объём данных по пользователям

Бизнес-метрики

- Объём данных по пользователям
- Объём скачиваемых данных по пользователям

Бизнес-метрики

- Объём данных по пользователям
- Объём скачиваемых данных по пользователям
- % редко запрашиваемых данных

Бизнес-метрики

- Объём данных по пользователям
- Объём скачиваемых данных по пользователям
- % редко запрашиваемых данных
- % данных, запрашиваемых из других регионов

Следующий вебинар

Тема: Сайт знакомств



Пятница 06.03.20 в 20.00



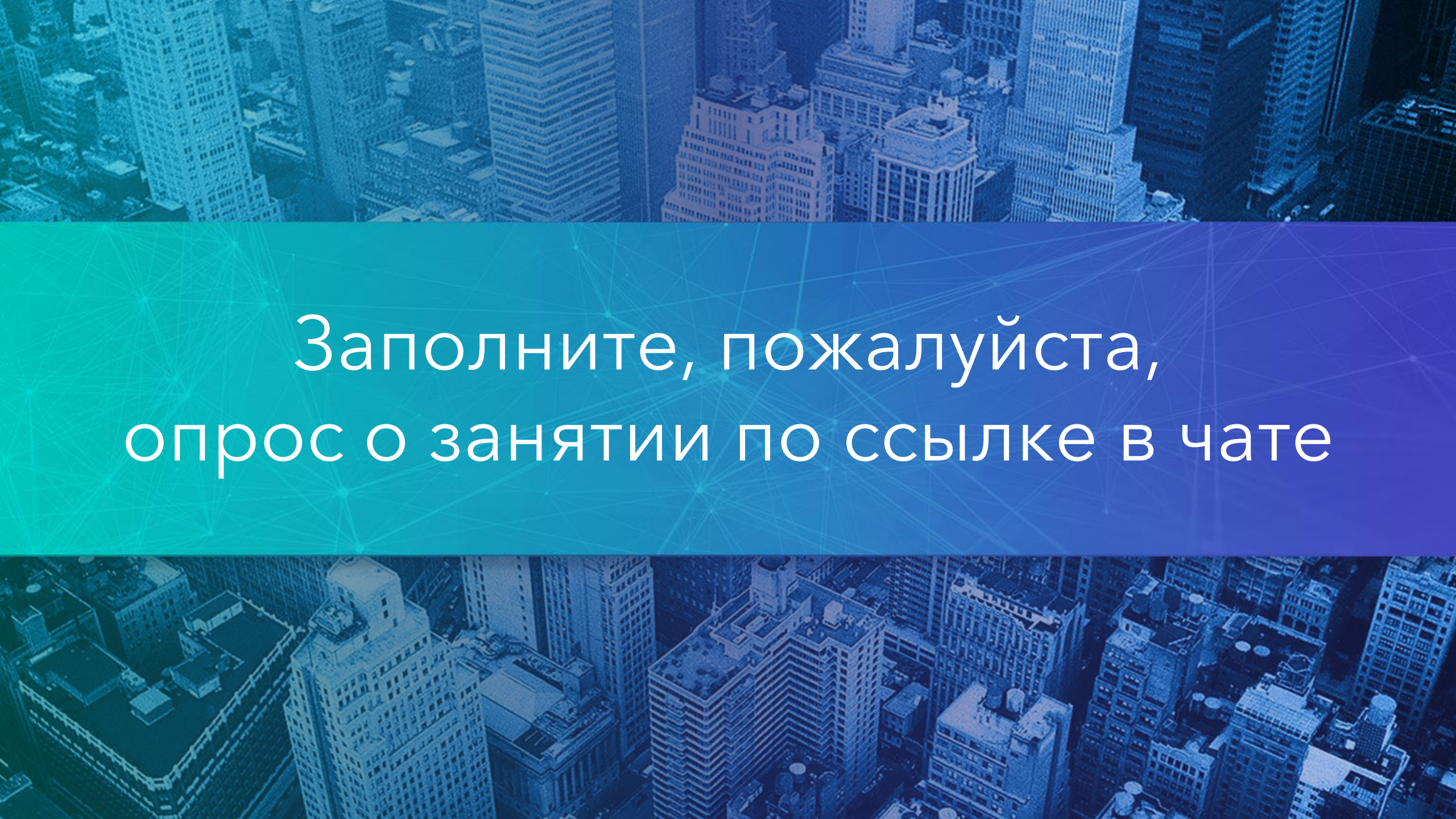
Ссылка на вебинар будет в ЛК за 15 минут



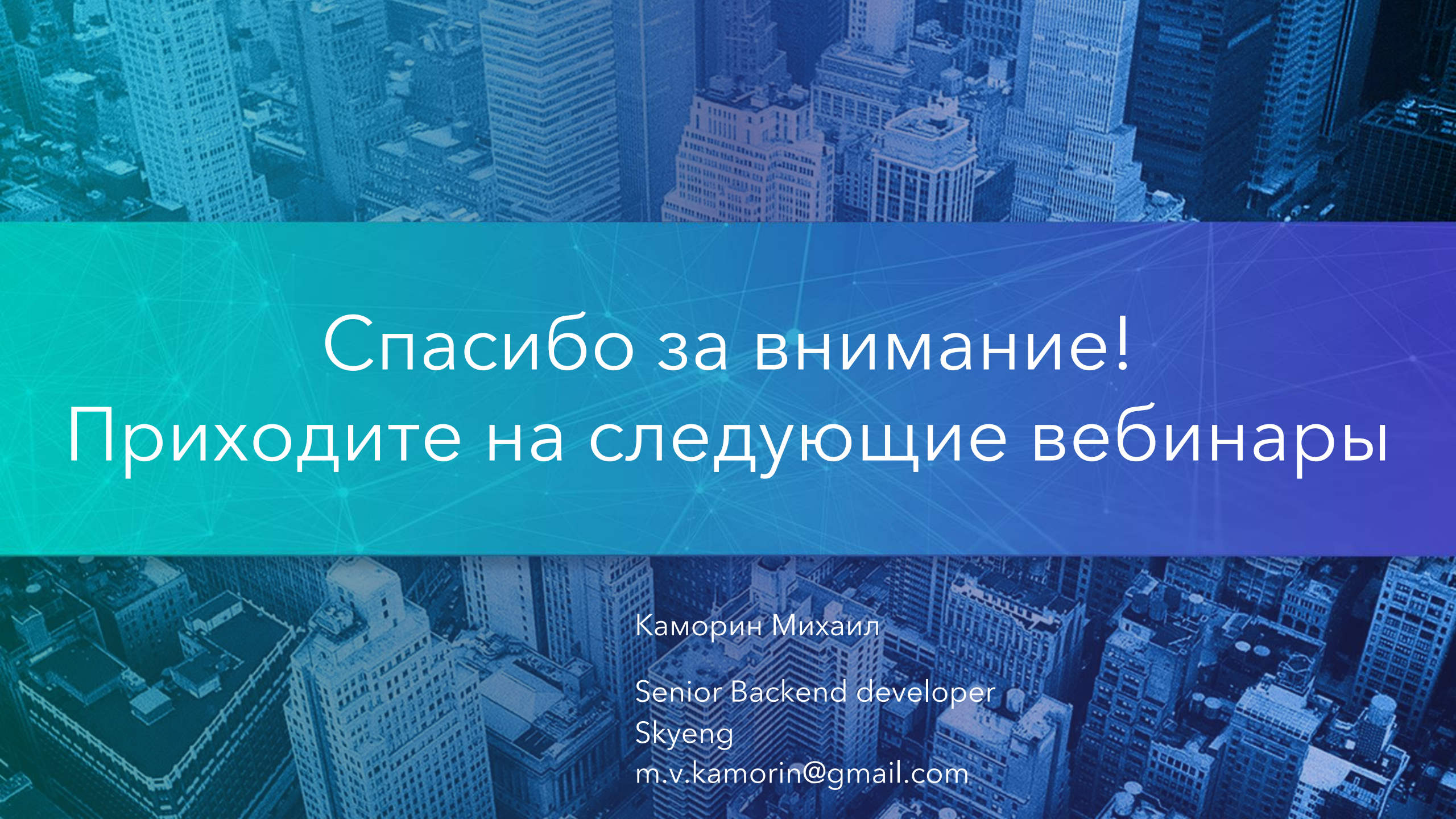
Материалы к занятию
в ЛК – можно изучать



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