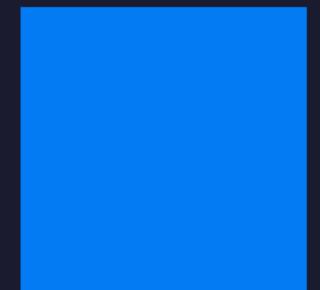


Яндекс Облако



Yandex Database —
эффективная альтернатива
традиционным noSQL-решениям



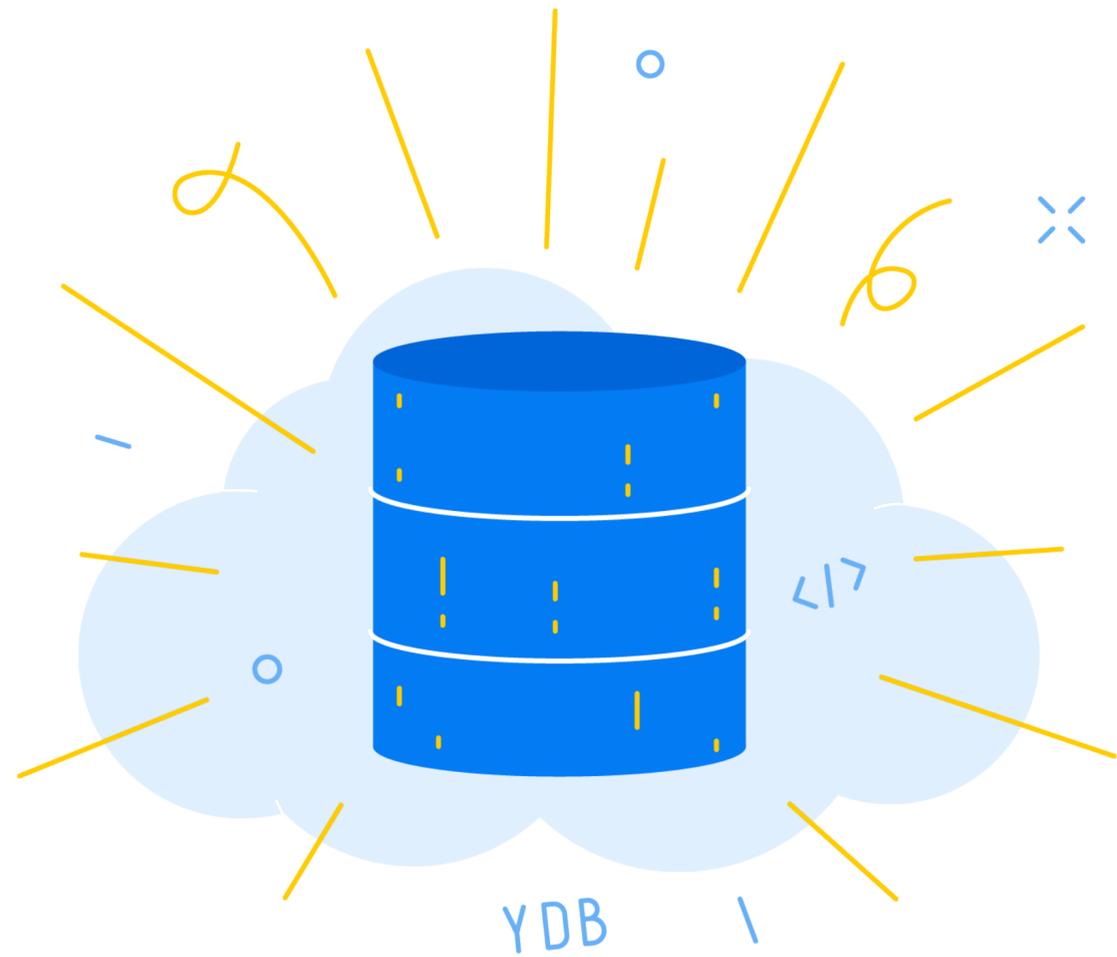
Андрей Фомичев

Руководитель отдела разработки
систем хранения данных

01

Yandex Database: Public Preview

Yandex Database: Public Preview



Доклады про YDB

- Что такое YDB и чем она может быть полезна Вам
- Как YDB используется в Яндексе и за пределами

02

Yandex Database: Rationale

NoSQL Database



Что вы думаете, когда слышите “noSQL Database”?

- Масштабируемость
- Доступность

Вся суть инфраструктуры Яндекса — про эти качества

- Любой сервис Яндекса мечтает вырасти до миллионов пользователей
- Любой сервис Яндекса должен работать 24x7

NoSQL — какой монетой платить?



Согласованность данных

- › Атомарные изменения данных
- › Транзакционные изменения данных

Функциональность

- › Упрощенная модель (Key-Value)
- › Явная спецификация операции вместо декларативного SQL-запроса

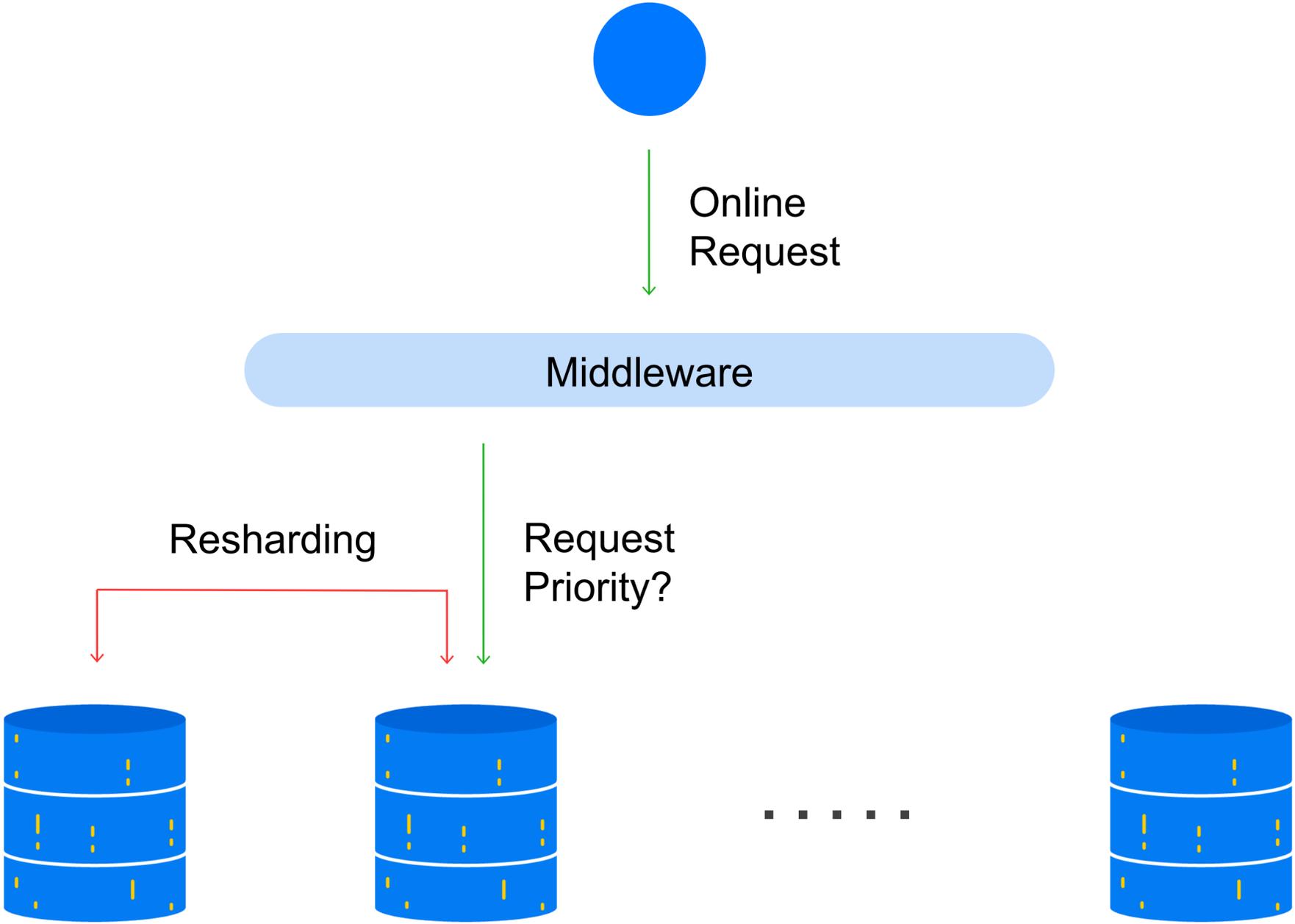
А может шардированная реляционная СУБД?



Шардированная реляционная база — популярный паттерн

- Один шард — нет проблем
- Много шардов — вы делаете работу инфраструктуры
 - Балансировка
 - Решардирование данных, если есть дисбаланс по данным
 - Решардирование данных при добавлении ресурсов

QoS и шардированная реляционная СУБД



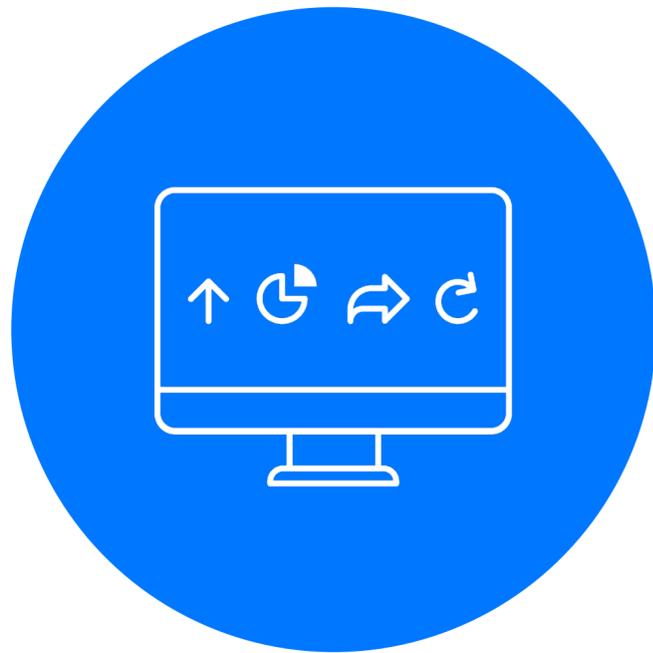
На что тратит время ваша разработка?



Ожидания

- › Подготовиться к «чёрной пятнице»
- › Выкатить подсказки с рекомендациями товаров
- › Подготовить всё для продажи нового продукта
- › Подключить новый регион
- › Оптимизировать стоимость инфраструктуры

На что тратит время ваша разработка?



Реальность

- › Поднять slave
- › Побалансировать
- › Перешардировать
- › Синхронизировать

03

Yandex Database — History

Давным-давно в далеком Яндексе



- Управляем большими данными сами
 - Псевдо реляционная шардированная СУБД
 - Считаем, что машины надежны, ведь там есть RAID
 - Иногда надо обработать все данные сразу (PageRank)
 - Консенсус и распределенные транзакции? Нет, не слышал
 - Решардинг данных — вручную

KiWi: NoSQL KV Database



- › 2011: Интернет уже большой, ДЦ ещё маленькие, перешардирование — боль
- › Key Value — кворум записи
- › Consistent Hashing
- › Отличие: посуточная обработка и репликация встроены в процесс merge
- › Обработка данных: UDF и KiWi Lang
- › Удаление данных — политики хранения

KiWi: NoSQL KV Database



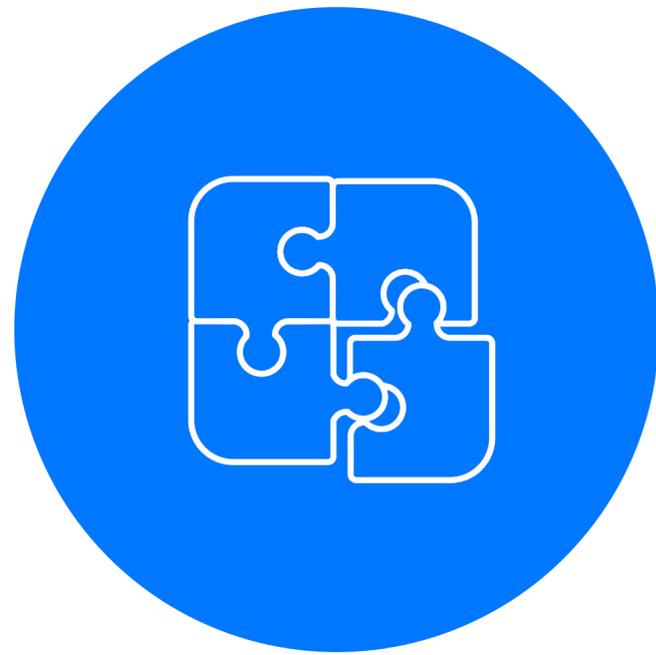
- Стали замечать аномалии
- Оказалось, что согласованность данных между записями важна
- Возникла потребность в консистентных срезах данных для анализа
- Каждую проблему можно было разобрать и исправить силами прикладных разработчиков, но это было неудобно
- Другие большие компании прошли похожий путь

Yandex Database Platform



- **One size fits all — платформа для наших нужд**
- Решили сделать платформу для наших нужд
- Time series Monitoring
- Persistent Queues
- Обработка данных в реальном времени (RTMR)
- Key Value Store
- Relational Database

Yandex Database Platform



Давайте раз и навсегда решим задачи

- Распределённый консенсус
- Репликация данных для отказоустойчивости
- Поддержка геораспределённости
- Распределённые транзакции

YDB — интерактивные приложения



Мечта: быстро делать интерактивные отказоустойчивые приложения

- > Отзывы
- > Online-магазин
- > Биллинг
- > Gaming
- > Собственно, само Облако 😊

04

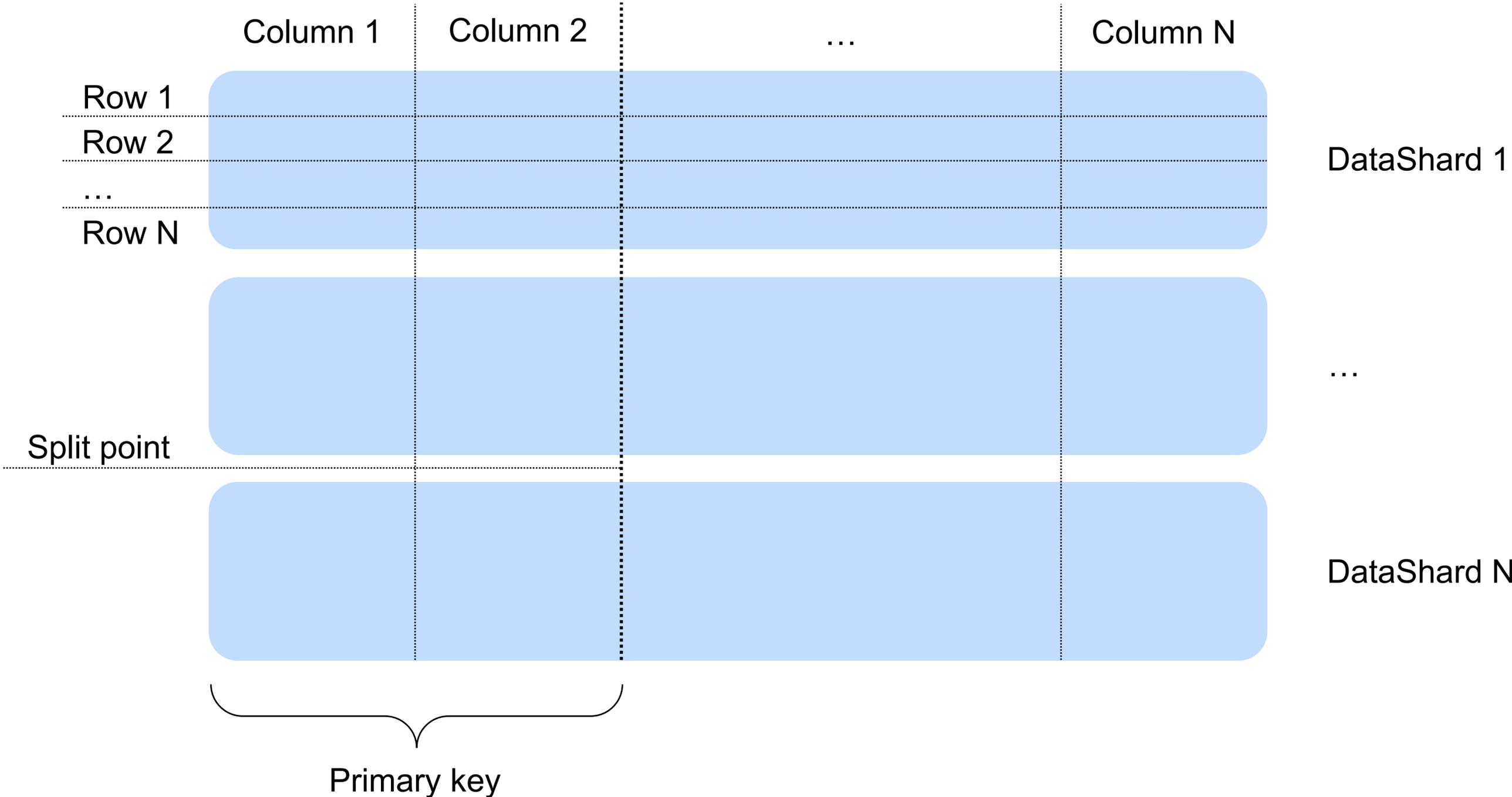
Yandex Database — NewSQL Database

Yandex Database

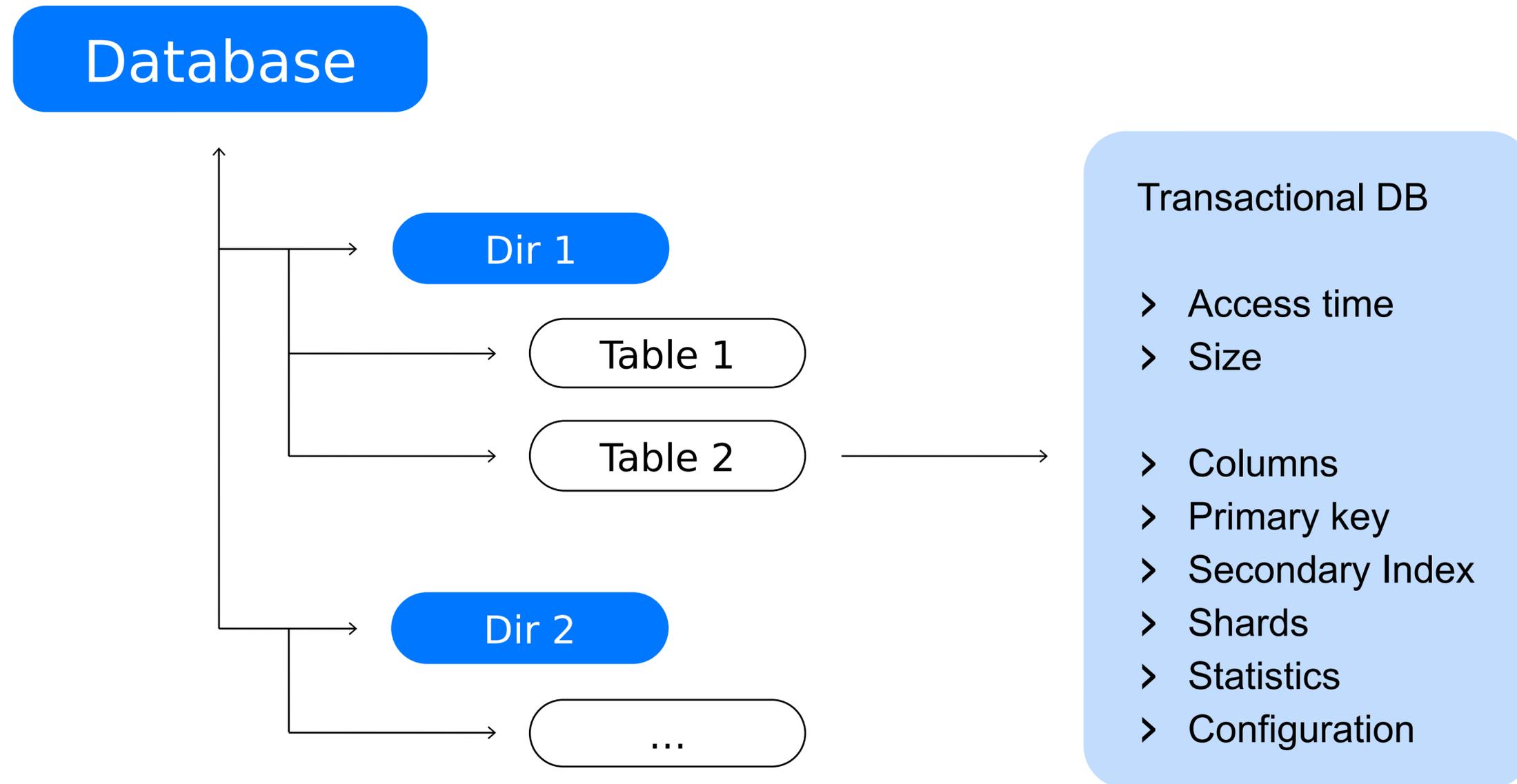
Yandex Database — это геораспределённая база данных, предоставляющая

- Надежное хранение данных с автоматической репликацией
- Механизм распределенных ACID-транзакций со строгой консистентностью
- Высокую пропускную способность при малом времени отклика
- Автоматическое восстановление после сбоев
- Горизонтальную масштабируемость на десятки тысяч нод
- Диалект SQL (YQL)

YDB: таблица или реляционное отношение



YDB: таблицы организованы в иерархию



Свойства Yandex Database

- Автоматический split / merge шардов
 - По размеру шарда
 - По нагрузке (в разработке)
- Автоматическое восстановление после сбоев за время обнаружения отказа
- Синхронная репликация данных, в том числе в геораспределённой конфигурации данных
- Механизм serializable-транзакций между записями базы данных

Свойства Yandex Database

- Предназначена для OLTP-нагрузки
 - Интерактивные приложения
 - Аналитика (в разработке)
- Вторичные индексы (онлайн-построение — в разработке)
- Механизм резервной копии (backup)

Декларативный язык запросов YQL



- > Диалект SQL
- > Строгая типизация
- > Именованные подзапросы
- > Явная параметризация
- > Богатый набор встроенных функций
 - DateTime
 - Regexp
 - Math
 - String
 - Unicode
 - Etc

Декларативный язык запросов YQL

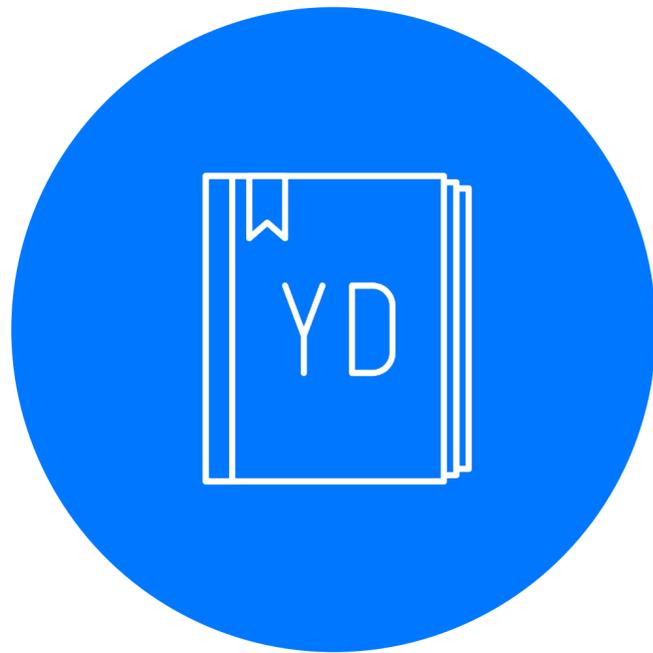


- Стандартные DML-конструкции
 - SELECT, INSERT, UPDATE, DELETE
 - JOIN, GROUP BY, ORDER BY

- Стандартные DML-конструкции
 - UPSERT / REPLACE
 - UPDATE ON
 - DELETE ON

- Стандартные DDL-конструкции
 - CREATE TABLE
 - DROP TABLE
 - ALTER TABLE

Клиентские SDK



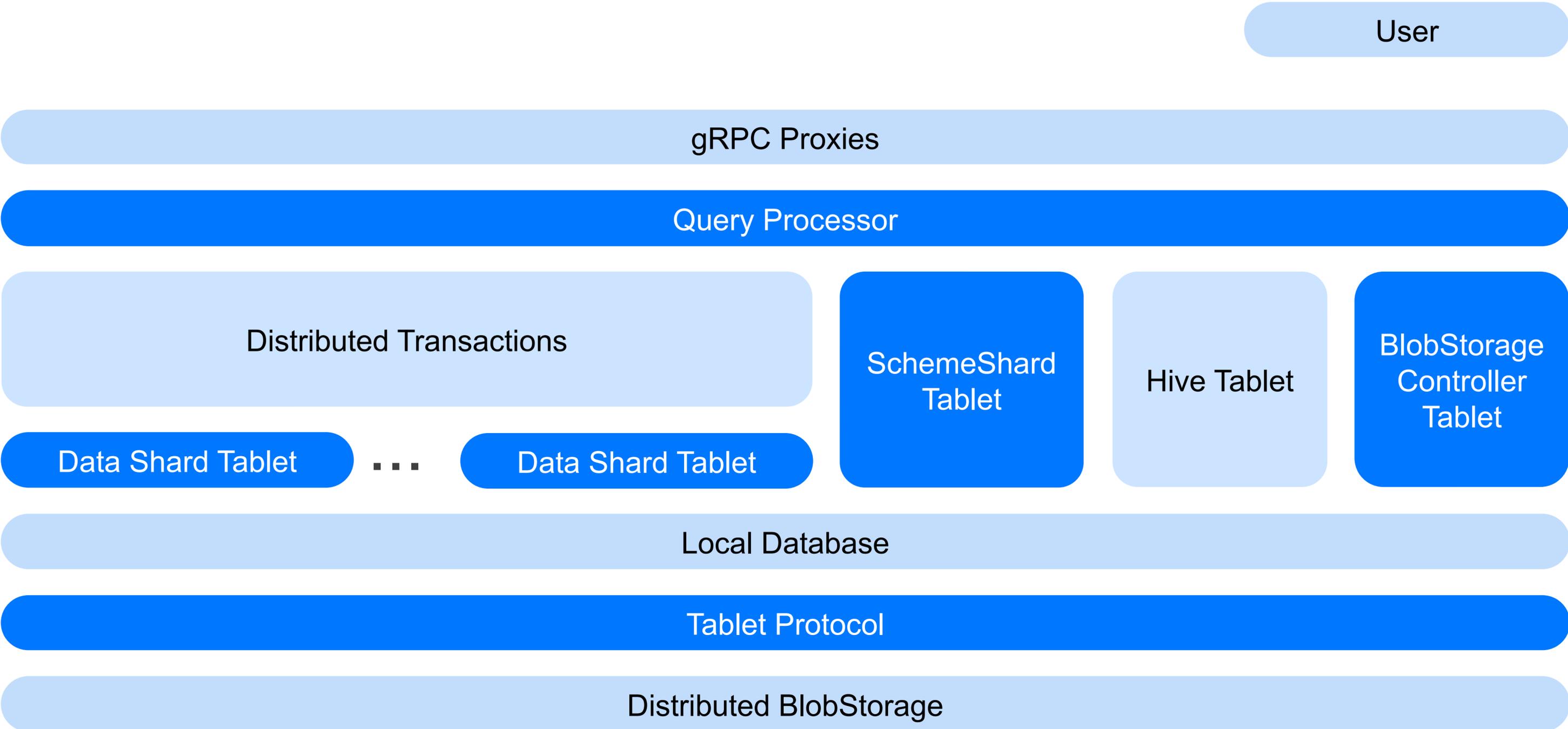
- Поддерживаемые языки программирования
 - Java
 - Python
 - Go

- Клиентские библиотеки реализуют
 - Пул подключений
 - Кеш подготовленных запросов
 - Балансировку соединений на стороне клиента
 - Автоматическое обнаружение серверной топологии
 - Обработку ошибок

05

Yandex Database — Internals

YDB — технологический стек



YDB: Distributed Blob Storage



- › Сетевое хранилище
- › Синхронная репликация
- › Работа с дисковыми устройствами напрямую
- › Коммит транзакции — запись на устройство
- › Собственный user-space RT-планировщик дисковых операций — изоляция пользователей и QoS
- › Борьба с длинным хвостом latency
- › Различные схемы избыточности данных: mirror-3, erasure-4-2

YDB: Network Interconnect



- Прозрачный интерконнект поверх TCP
- Мультиплексирование и приоритезация трафика (QoS)

YDB: Tablets and Local Database



- Tablet — отказоустойчивый шард, решает проблему распределенного консенсуса
- Tablet используется для системных нужд и для хранения пользовательских данных (шарда таблицы)
- Данные хранятся в Local Database — Log Structured Merge Tree
- Логически данные представлены в виде реляционных таблиц с обязательным Primary Key

YDB: Distributed Transactions



- ACID, serializable, но можно ослаблять
- Распределённые транзакции дороги, но иногда необходимы
- Традиционно распределённые транзакции реализуются через 2PC, в YDB deterministic transactions

YDB: Query Processing



- YQL — диалект SQL
- Сейчас — OLTP-нагрузка
- YQL-запрос выполняется как набор deterministic transactions
- Между распределёнными транзакциями берутся optimistic locks
- Уровни изоляции
 - Serializable
 - Read committed
 - ...

YDB: gRPC proxy



- › База доступна как набор endpoints
- › Общение по протоколу gRPC
- › Протокол открыт, можно разработать SDK для разных языков программирования
- › SDK реализует клиентскую балансировку — хорошо масштабируемое и отказоустойчивое решение

06

Как используют Yandex Database в Яндексе

Yandex Database — сценарии использования

auto.ru



07

Yandex Database — ключевой компонент Облака

Облако и Yandex Database



Гиперконвергентная архитектура

- › Storage и Compute живут на одном железе
- › В качестве Storage Облака используется YDB Storage
- › Сетевые диски живут на YDB Storage
- › Control Plane Облака живет в YDB

Облако и Yandex Database



Реляционная БД по умолчанию — сервисы Яндекс.Облака используют базы YDB для хранения данных

- ⚙ Compute Cloud
- ⚡ Load Balancer
- 🏠 Instance Groups
- 🏠 Managed Service for Kubernetes®

....

Пользовательские сервисы Облака



-  Monitoring — сервис для сбора и визуализации метрик приложений
-  Object Storage — масштабируемое хранилище данных
-  Message Queue — очереди для обмена сообщениями между приложениями

08

Yandex Database — сервис в Яндекс.Облаке



PREVIEW

Yandex Database

Yandex Database (YDB) — это распределённая отказоустойчивая NewSQL СУБД.

YDB обеспечивает высокую доступность и масштабируемость, и, в то же время, строгую консистентность и поддержку ACID-транзакций. Для запросов используется декларативный SQL-подобный язык запросов YQL.

[Запросить доступ](#)[Документация](#)

Имя ?

Вычислительные ресурсы ?

s1.large	s2.large
vCPU 16 ядер	vCPU 16 ядер
Память 64 ГБ	Память 64 ГБ

Количество 1 6

Группы хранения ?

Тип диска ?

Количество ? 1 6 x 100 GB

Сеть

Облачная сеть или

Подсети

ru-central1-a или

ru-central1-b или

ru-central1-c или



- Yandex Database
- ydb680
- Обзор
- Навигация
- Мониторинг

Корневая папка

Фильтр по имени Создать

Имя	Владелец
episodes	ajegq762fp6sr4lpaaha0@as
seasons	ajegq762fp6sr4lpaaha0@as
series	ajegq762fp6sr4lpaaha0@as

- Скопировать путь
- Вставить путь в запрос
- Использовать как шаблон
- Создать таблицу
- Изменить таблицу
- Запросить данные
- Добавить данные
- Удалить

seasons Таблица

Схема

#	Имя	Ключ	Тип
0	series_id	P	UInt64
1	season_id	P	UInt64
2	first_aired		UInt64
3	last_aired		UInt64
4	title		Utf8

Запрос SQL-запрос

```

1 SELECT
2     series_id,           -- имена колонок (series_id, release_date, title)
3     title AS series_title, -- перечислены с запятой в качестве разделителя
4     CAST(release_date AS Date) AS release_date -- используя AS в качестве ключевого слова можно
5     -- переименовать столбцы или дать имя произвольному
6     -- выражению
7     FROM series;
8 COMMIT;

```

Выполнить Completed

Результат #1

#	series_id	series_title	release_date
0	1	"IT Crowd"	2006-02-03
1	2	"Silicon Valley"	2014-04-06

- Документация
- Начать работу с YDB
- Концепции YDB
- Рекомендации по работе с YDB
- Инструкции для работы с YQL
- Справочник YQL

Yandex Database

ydb241

Обзор

Навигация

Мониторинг

Документация

[Начать работу с YDB](#)

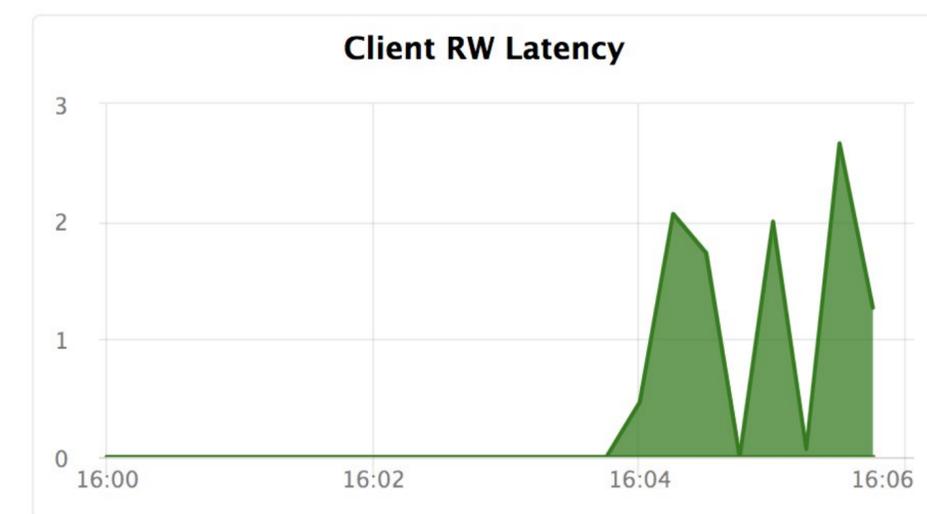
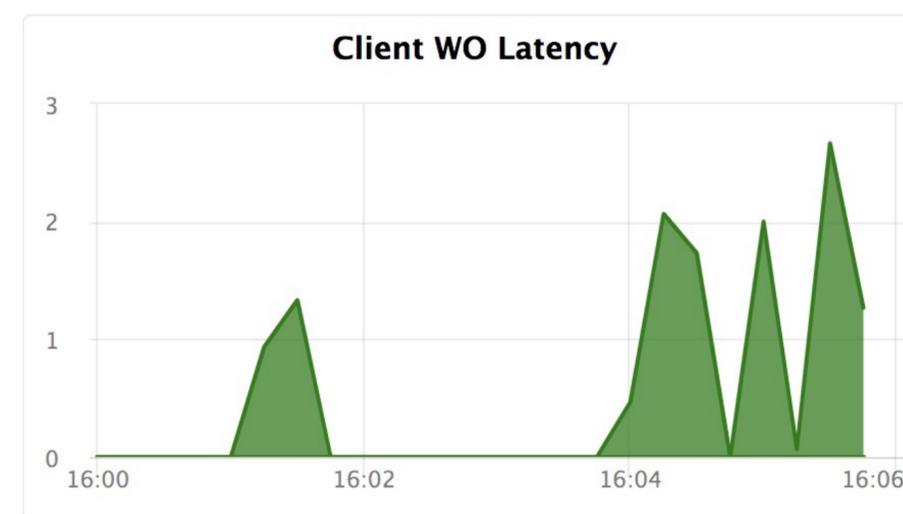
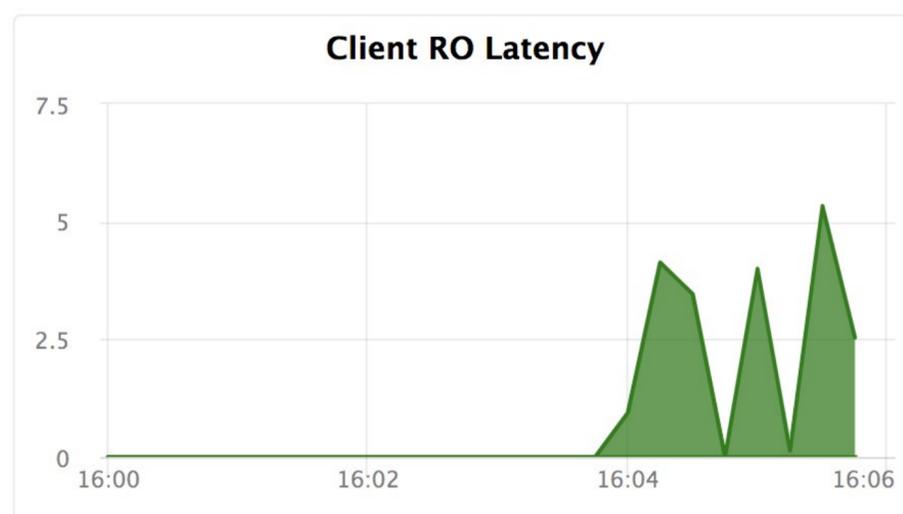
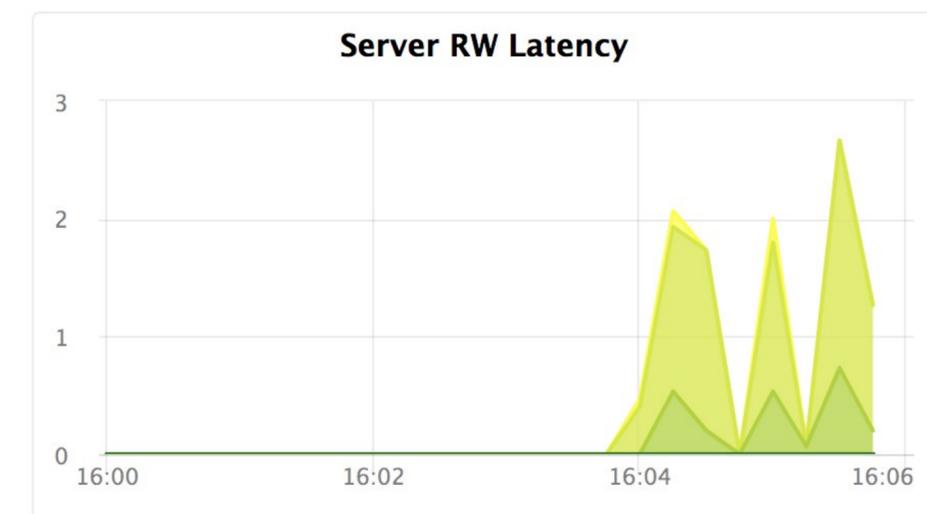
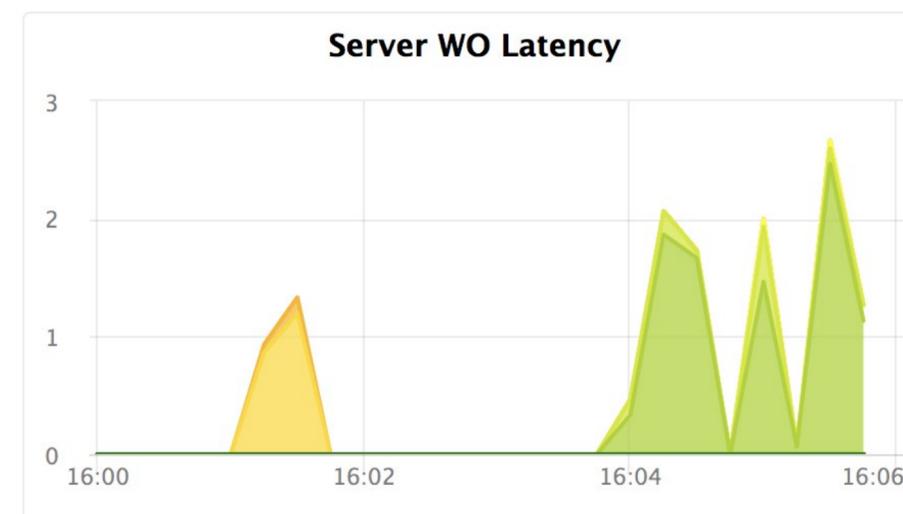
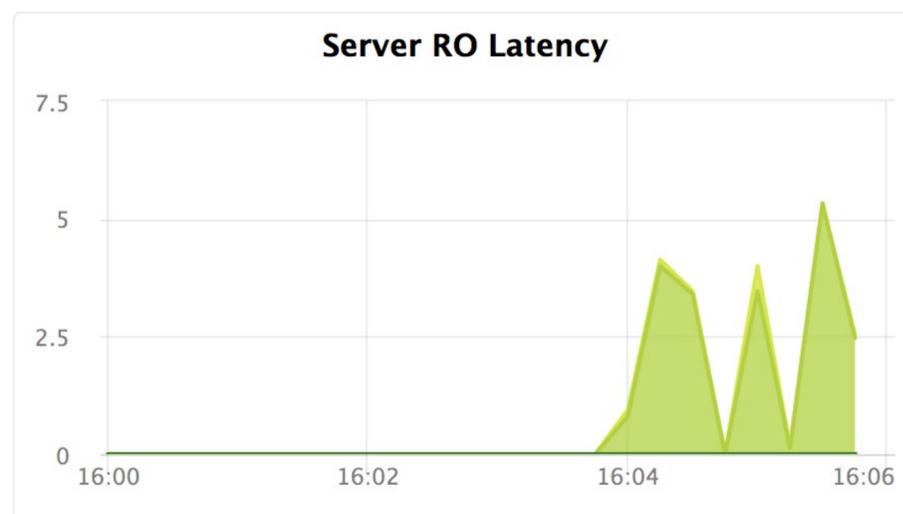
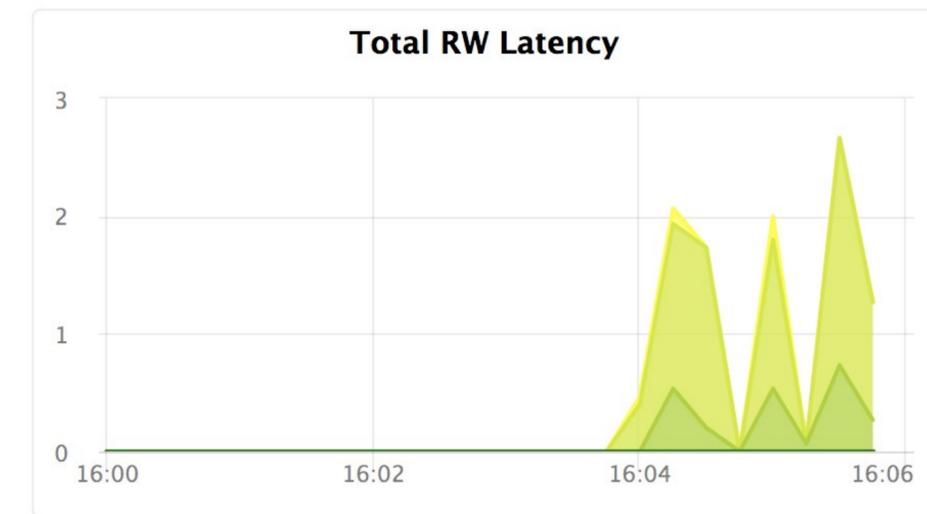
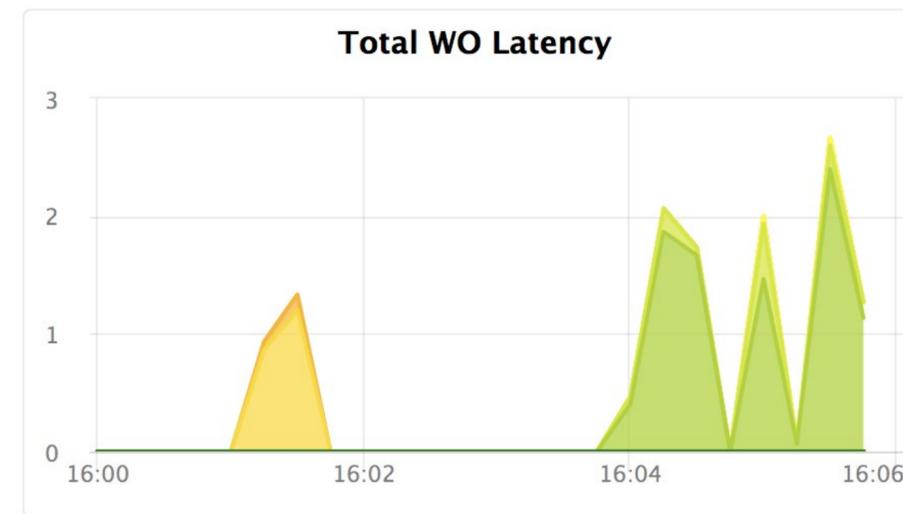
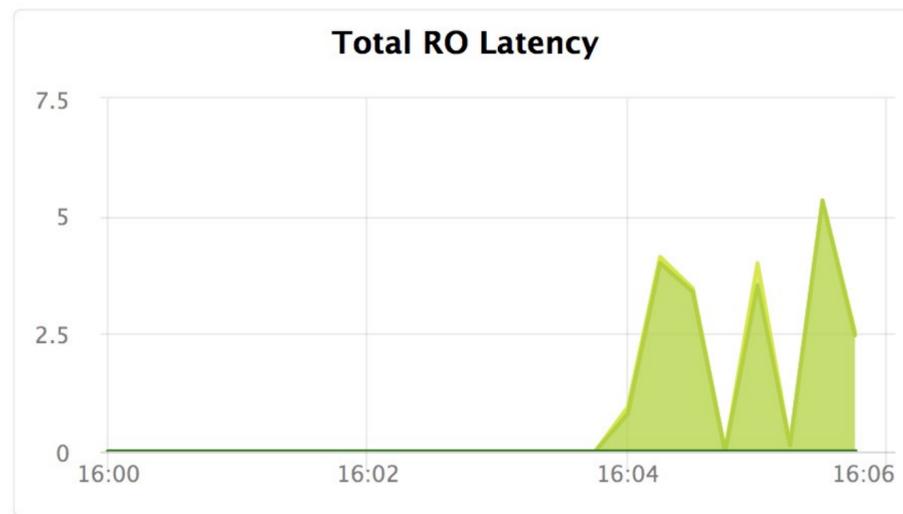
[Концепции YDB](#)

[Рекомендации по работе с YDB](#)

[Инструкции для работы с YQL](#)

[Справочник YQL](#)

28 сентября 2019 16:01 — 28 сентября 2019 16:06 Час День Неделя Месяц





cloud.yandex.ru/services/ydb

Yandex Database — Grow With Us

Присоединяйтесь к Public Preview

Андрей Фомичев

Руководитель отдела разработки
систем хранения и обработки данных

fomichev@yandex-team.ru

