Yandex Cloud

Платформа Yandex.Cloud

Григорий Атрепьев
Главный архитектор облачных решений

Содержание

1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Мы постоянно улучшаем нашу платформу

Расширяем ключевые сценарии использования

9 сервисов

+ Базовая инфраструктура

Размещение веб-сервисов

Автоматизация работы колл-центров

>50 клиентов

2018

27 сервисов

+ Платформа данных

Архитектура микросервисов на базе Kubernetes®

Бизнес-аналитика

>5 тыс. клиентов

2019

35 сервисов

+ ERP-системы в облаке

Приложения cloud-native

Рабочее место для data science

>10 тыс. клиентов

2020

50 сервисов

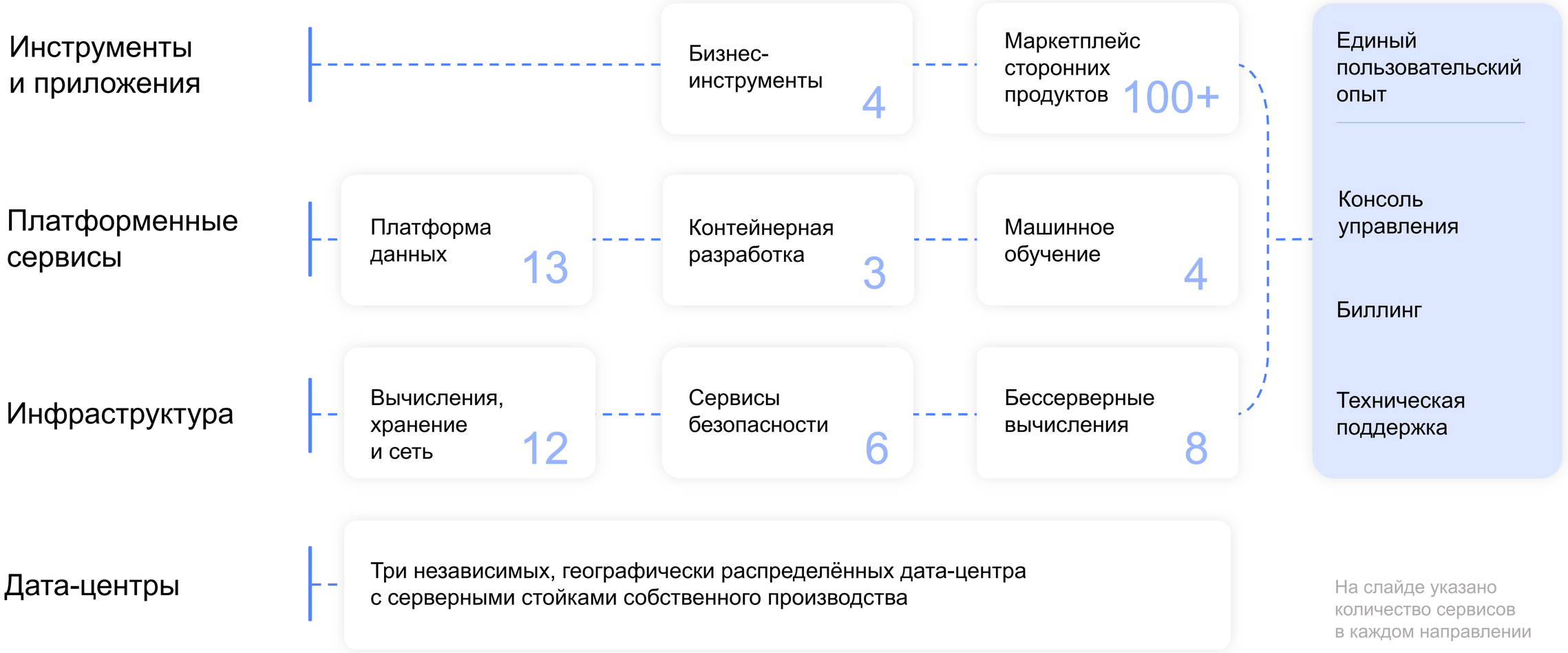
+ Потоковая аналитика

Удалённые рабочие станции

Ускоренная доставка контента

2021

Платформа Yandex.Cloud — единый хаб новых технологий



На слайде указано количество сервисов в каждом направлении

Сценарии использования

- Сайт в облаке
- Интернет-магазин
- Хранение и обработка персональных данных
- 1С в облаке
- Сервисы Microsoft в облаке
- Автоматизация колл-центров
- Рекомендательная система для ритейла и e-commerce
- Корпоративное хранилище данных
- Бизнес-аналитика и визуализация данных
- Serverless
- Чат-боты на Serverless
- Distributed Cloud
- Миграция в облако
- Security Solution Library
- Solution Library for AWS

Конкурентоспособные цены

			aws		Azure
Compute Конфигурация 2-4-80 Linux	$ 28,75	$ 28,82	$ 79,36	$ 68,36	$ 68,64
Object Storage Горячее хранилище объёмом 1 ТБ в месяц	$ 16,15	$ 19,55	$ 24,50	$ 23,00	$ 19,60
Managed PostgreSQL Ресурсы vCPU — RAM — SSD: 8 — 32 — 500 в месяц	$ 208,39	$ 217,69	$ 703,60	$ 587,60	$ 677,03
Managed MySQL Ресурсы vCPU — RAM — SSD: 8 — 32 — 500 в месяц	$ 214,30	$ 217,69	$ 677,32	$ 587,60	$ 677,03
Speech-to-Text	$ 7,69	—	$ 24,00	$ 22,56	$ 16,67

Действующие цены в долларах США на ноябрь 2021 г. с учётом НДС

Конкурентоспособные цены

			aws		Azure
Compute Конфигурация 2-4-80 Linux	2 242 руб.	2 580 руб.	5 920 руб.	5 100 руб.	5 121 руб.
Object Storage Горячее хранилище объёмом 1 ТБ в месяц	1 260 руб.	1 750 руб.	1 828 руб.	1 716 руб.	1 462 руб.
Managed PostgreSQL Ресурсы vCPU — RAM — SSD: 8 — 32 — 500 в месяц	17 719 руб.	19 488 руб.	52 489 руб.	43 835 руб.	50 506 руб.
Managed MySQL Ресурсы vCPU — RAM — SSD: 8 — 32 — 500 в месяц	18 180 руб.	18 288 руб.	50 528 руб.	43 835 руб.	50 506 руб.
Speech-to-Text	600 руб.	-	1 790 руб.	1 683 руб.	1 243 руб.

Действующие цены в рублях на ноябрь 2021 г. с учётом НДС

Простые инструменты для IT-специалистов

- Полнофункциональная консоль управления
- Доступ к сервисам через API, CLI и Terraform
- SDK для Python, Go, Java, C#
- Документация на русском и английском языках
- Многоуровневая техподдержка

1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Инфраструктурные сервисы

Compute Cloud
BM, диски и GPU

Object Storage
Масштабируемое хранилище данных

Cloud Interconnect
Выделенные сетевые подключения

API Gateway
Интеграция с сервисами Yandex.Cloud

Virtual Private Cloud
Управление облачной сетью

Load Balancer
Сетевые балансировщики нагрузки

DDoS Protection
Защита от DDoS-атак

Application Load Balancer
L7-балансировщики нагрузки

Load Testing Preview
Тестирование и анализ производительности

Cloud DNS
Управление DNS

Cloud CDN
Организация сетей доставки контента (CDN)

Monitoring
Сбор и анализ метрик ресурсов

Resource Manager
Управление облаками, каталогами и другими облачными ресурсами

Cloud Desktop Preview
Удалённые рабочие места в облаке

Managed Service for Kubernetes®
Управление кластерами Kubernetes

Container Registry
Управление Docker-образами

Serverless Containers
Запуск контейнеров без Kubernetes

Cloud Organization
Управление сервисами организации

Identity and Access Management
Управление облачными ресурсами

Cloud Logging Preview
Формирование логов для сервисов Yandex.Cloud

Compute Cloud

Платформы

- Intel Broadwell / Cascade Lake / Ice Lake
- Макс. 96 vCPU / 640 (1024) ГБ RAM
- Burst (совместно используемые ядра vCPU) и прерываемые ВМ
- Выделенные хосты
- Варианты хранения без репликации (высокая скорость ввода и вывода)

GPU

- GPU Tesla V100/A100 (до 8 GPU на 1 ВМ)
- vGPU (мин. ¼ Tesla V100)

Yandex Object Storage

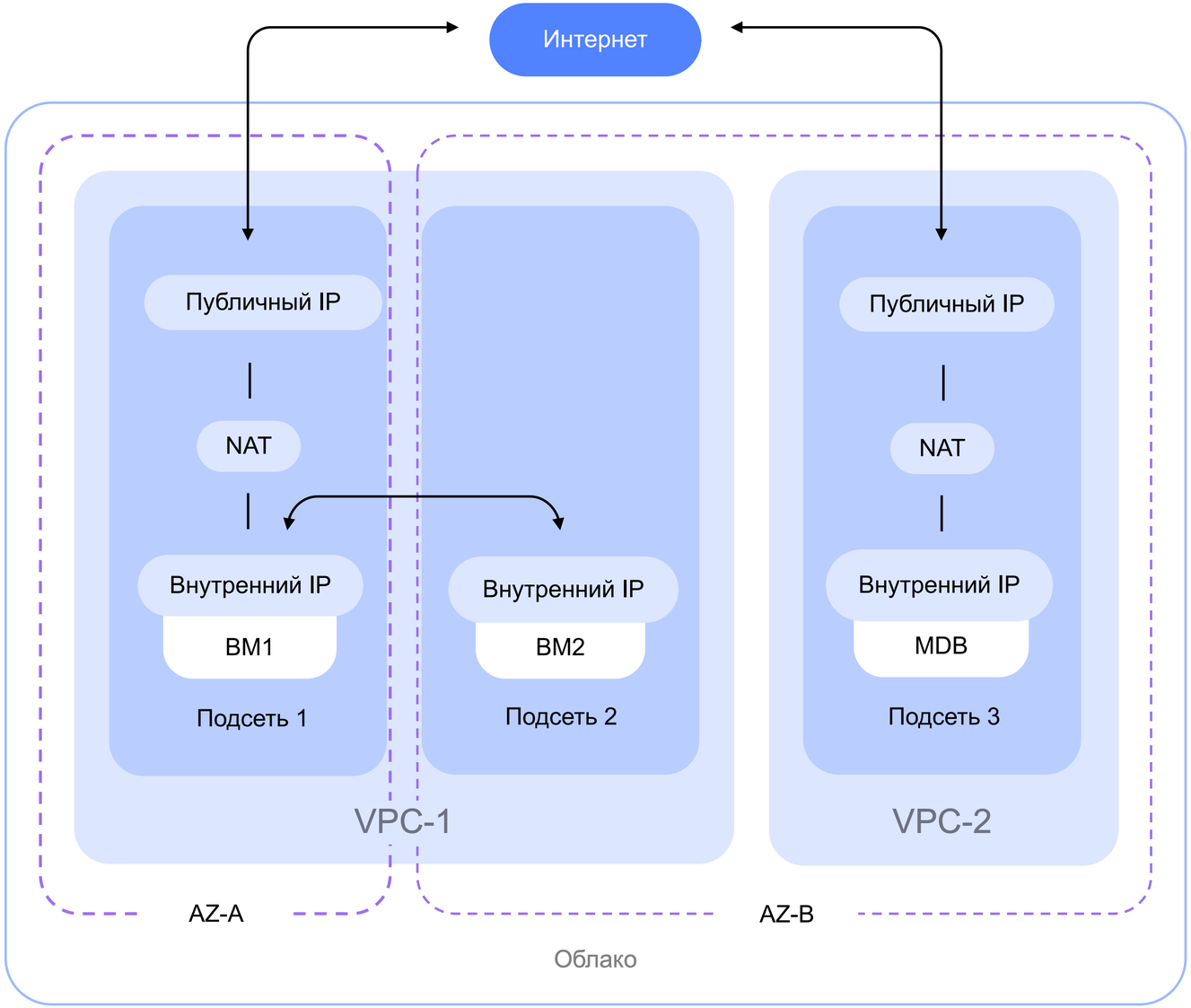
- S3-совместимый API-интерфейс
- Репликация данных по трём зонам доступности
- Горячее и холодное хранение данных для разных сценариев
- Тесная интеграция со множеством сервисов Yandex.Cloud
- Технология Cloud SDN доступна как опция

VPC с точки зрения клиента

Yandex VPC

Функциональные возможности

- Охват по регионам
- Подключения по приватным IP-адресам
- Балансировщик нагрузки (L4 + L7)
- Interconnect (выделенное соединение)
- Анти-DDoS
- Выходной шлюз
- Статическая маршрутизация и поддержка сетевых устройств
- Группы безопасности

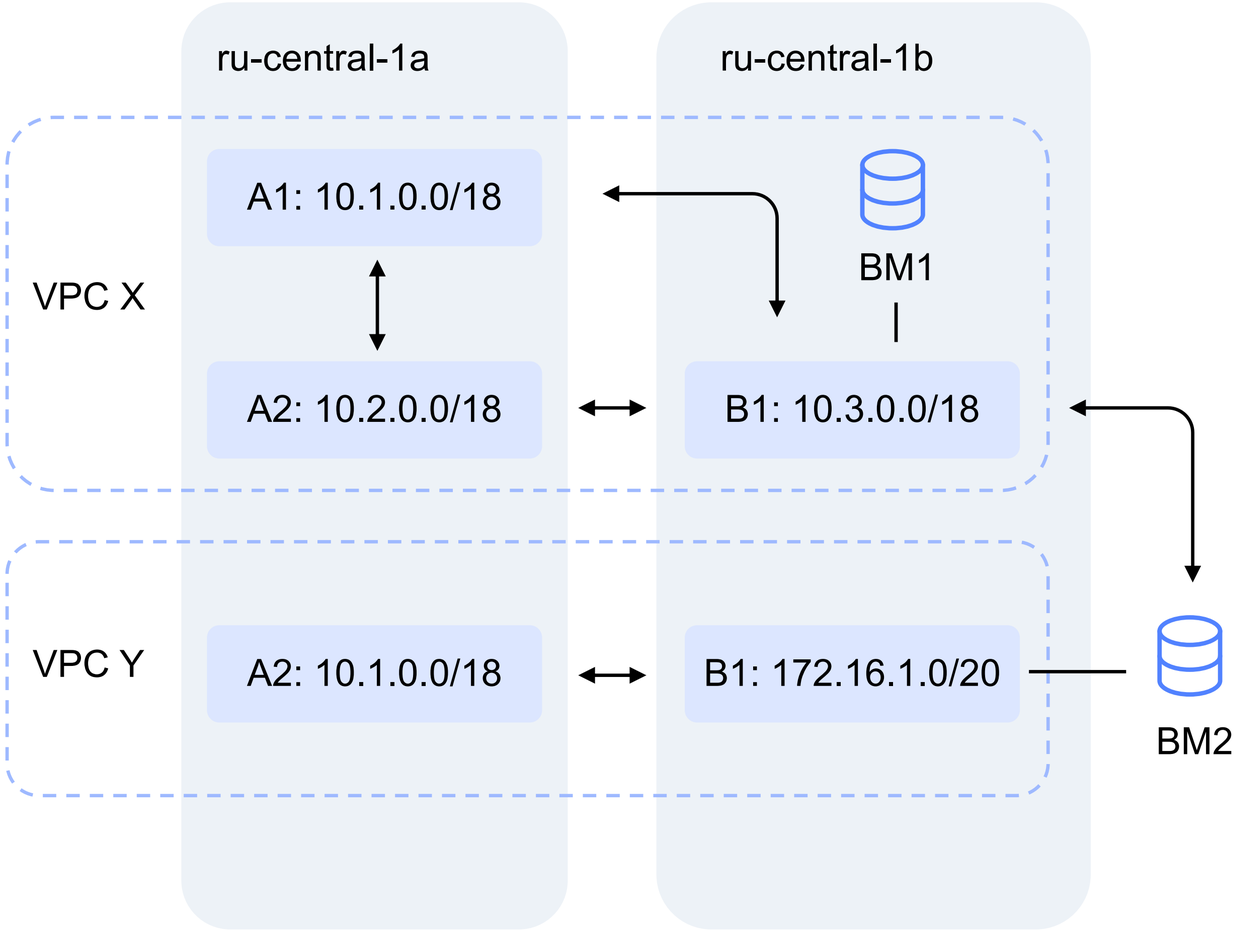


Пример организации виртуального частного облака (VPC) в Yandex.Cloud

Обзор решения VPC с позиций клиента

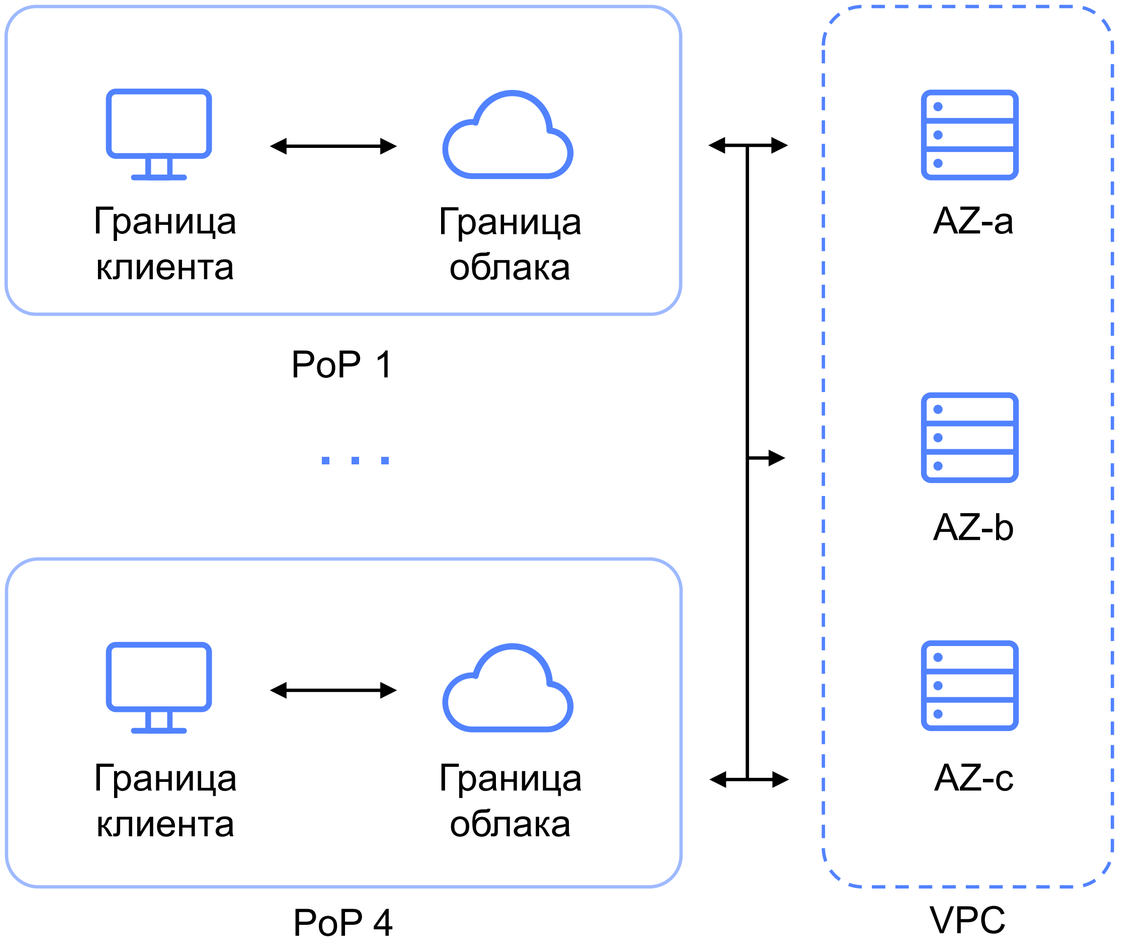
Основные абстракции

- **Сеть VPC** — контейнер подсетей (или VRF) — покрытие по регионам
- **Подсеть** — диапазон IP-адресов в облачной сети. Адреса из этого диапазона могут быть выделены облачным ресурсам, например виртуальным машинам и БД-кластерам. Сервис AZ-scope
- **Статический маршрут** — маршрут из таблицы маршрутизации, применяемый на уровне подсети для передачи трафика на другой интерфейс ВМ (например, файрвол или VPN-устройство)
- **Группа безопасности** — сетевое правило, которое может быть применено к интерфейсу виртуальной машины или другого сервиса



Cloud Interconnect

- Создание приватных подключений между вашей инфраструктурой и Yandex Virtual Private Cloud
- Четыре точки присутствия в Москве (M9, STOREDATA, DATALINE OST, DATALINE NORD)
- Сервис может предоставляться через сертифицированного партнёра Cloud Interconnect или любого интернет-провайдера, работающего в Москве
- Можно использовать мультиоблачные топологии на базе AWS/GCP/Azure
- Можно использовать для объединения виртуальных частных облаков (VPC peering)



Важные моменты о поддержке VPC

Поддерживается только маршрутизация L3

Мы производим ARP-спуфинг внутри подсети, при этом в ответ на все ARP-запросы выдаётся MAC-адрес виртуального интерфейса маршрутизатора. Поэтому маршрутизируются даже пакеты, передаваемые между двумя ВМ, находящимися в одной и той же подсети

Используются публичные IP-адреса NAT 1 : 1

Доступ к общедоступной сети интернет управляется через IP-адреса Elastic, обеспечивающие механизм NAT 1 : 1

Запрещены плавающие и VRRP/HSRP IP-адреса

Мы не поддерживаем L2-протоколы, которые создают плавающие (или общие) IP-адреса в подсетях. Для обеспечения избыточности клиенты могут использовать сетевые балансировщики нагрузки

Поддержка файрвола

Можно реализовать с использованием групп безопасности либо сетевых устройств сторонних производителей (например, Cisco, Palo Alto). VPC поддерживает статическую маршрутизацию трафика через такие устройства

Больше информации о VPC

Контейнерные решения

Комплекс контейнерных решений Yandex.Cloud

Container Solution

Managed Service for Kubernetes®

Serverless Containers

Container Registry

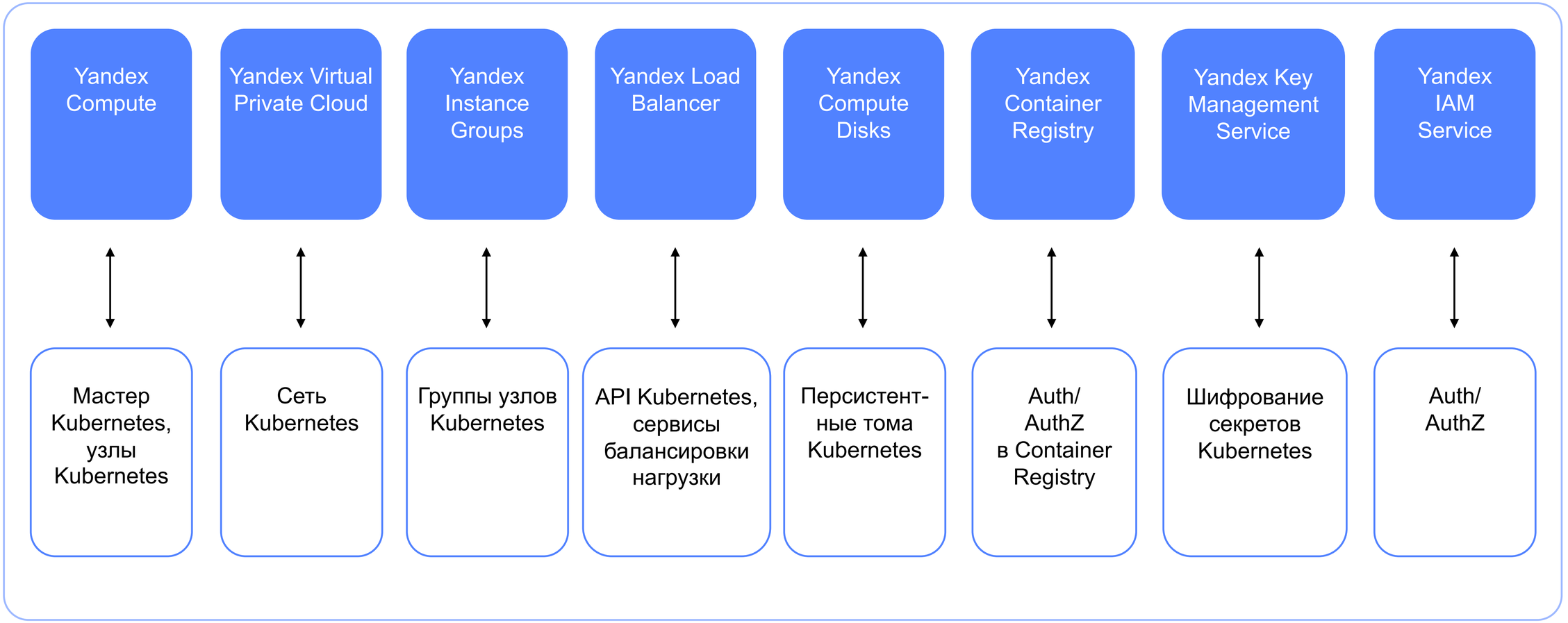
Инфраструктура

Yandex Managed Service for Kubernetes®

- Полномасштабное управление контейнеризованными приложениями на базе Kubernetes®
- Полное резервирование при использовании сервиса Cross AZ
- В качестве worker-узлов можно использовать любые конфигурации BM Compute Cloud
- Поддержка до 56 PVC на один узел
- Поддержка GPU-узлов для задач машинного обучения и искусственного интеллекта
- Поддержка Horizontal Cluster Autoscaler, Split DNS, Calico Network Policy, Cilium dataplane Preview
- Поддержка версий 1.20, 1.19, 1.18
- Мониторинг и управление кластером через облачный пользовательский интерфейс

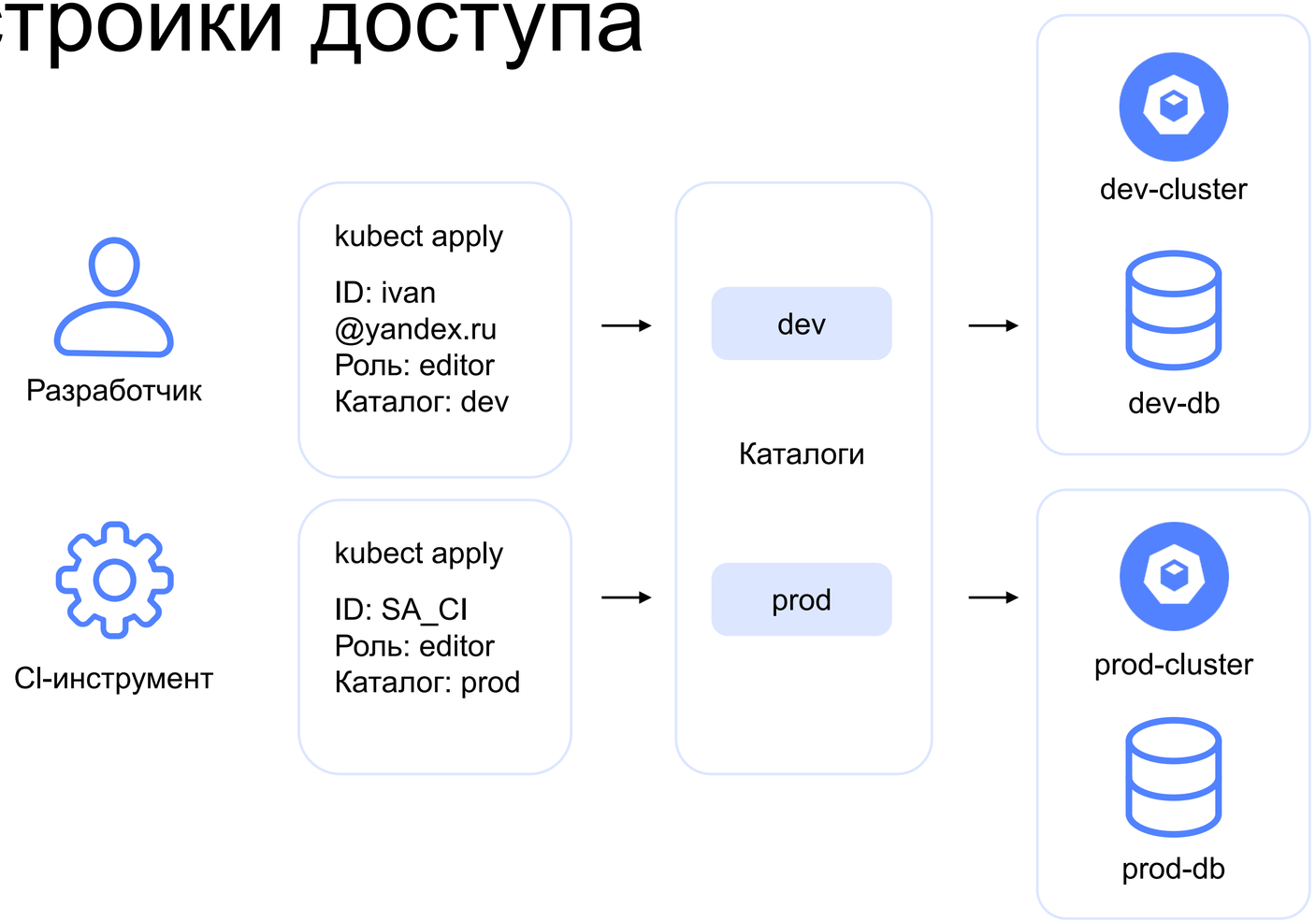
Интеграции для Kubernetes

Подробности интеграции рассмотрены
с позиций пользователей Kubernetes



Каталоги можно использовать для детальной настройки доступа

- IAM можно использовать для предоставления доступа сразу к нескольким кластерам и БД Kubernetes
- Можно использовать пользовательские и сервисные аккаунты
- IAM интегрируется в kubectl
- Пользовательские аккаунты могут быть предоставлены внешними поставщиками удостоверений через SAML-федерацию



Управление доступом и ресурсами

Yandex Cloud Organization

Сервис предназначен для управления оргструктурой, настройки интеграции с каталогом сотрудников, а также разделения доступа пользователей к облачным ресурсам организации

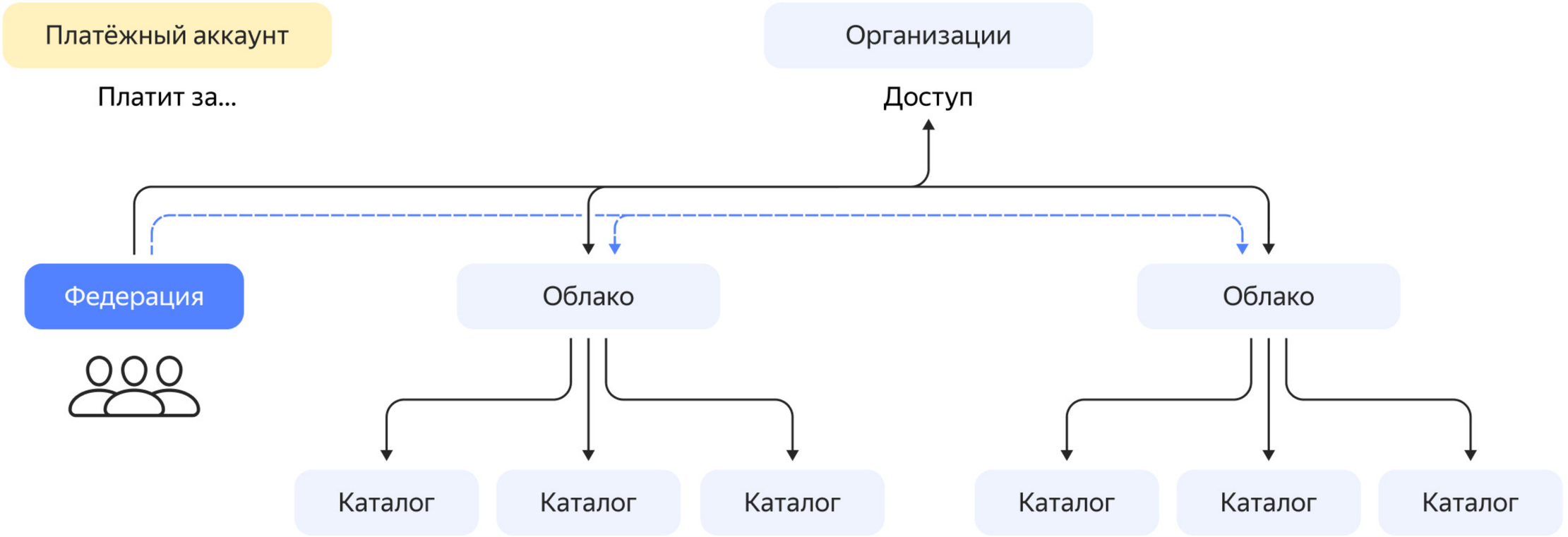
Корневой контейнер для всех ресурсов в организации:

- Пользователей и организаций
- Облаков
- Дополнительных сервисов

Область видимости для назначения доступов:

- Пользователям на всю организацию
- Сервисным аккаунтам в облаках организации

Управление доступом и ресурсами



Федерация удостоверений

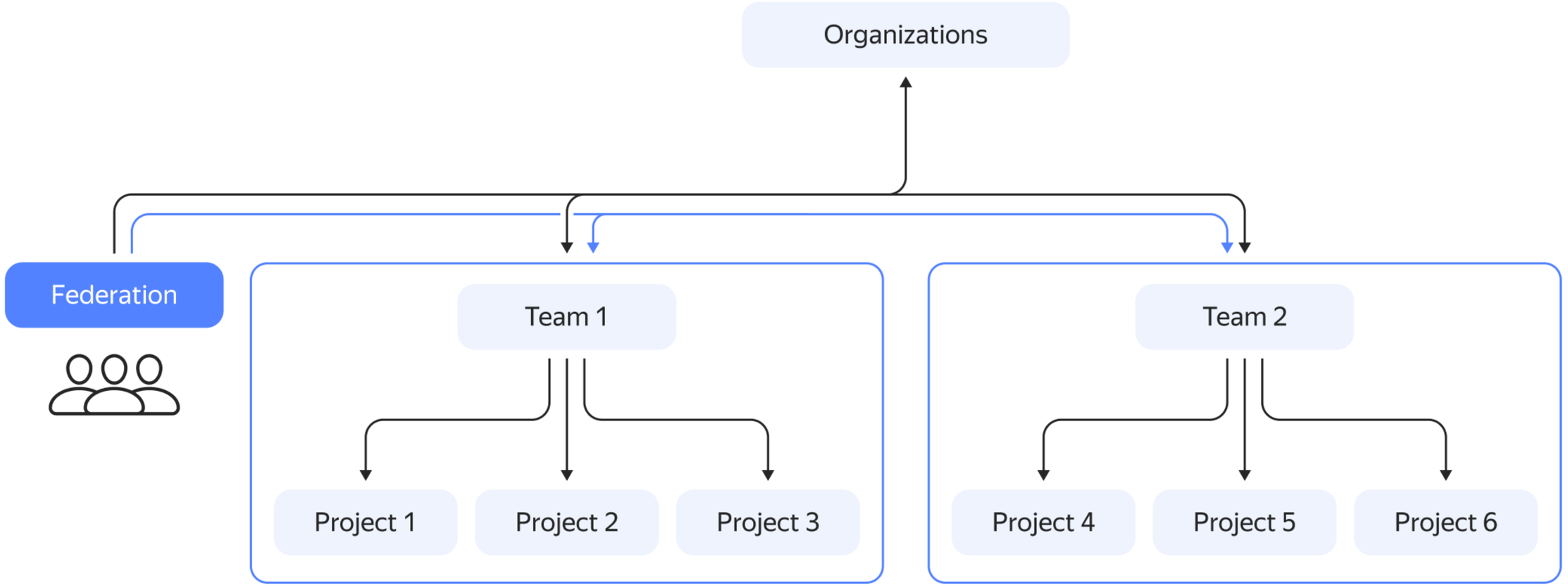
Контроль аутентификации на стороне клиента

Любой SAML-совместимый поставщик удостоверений

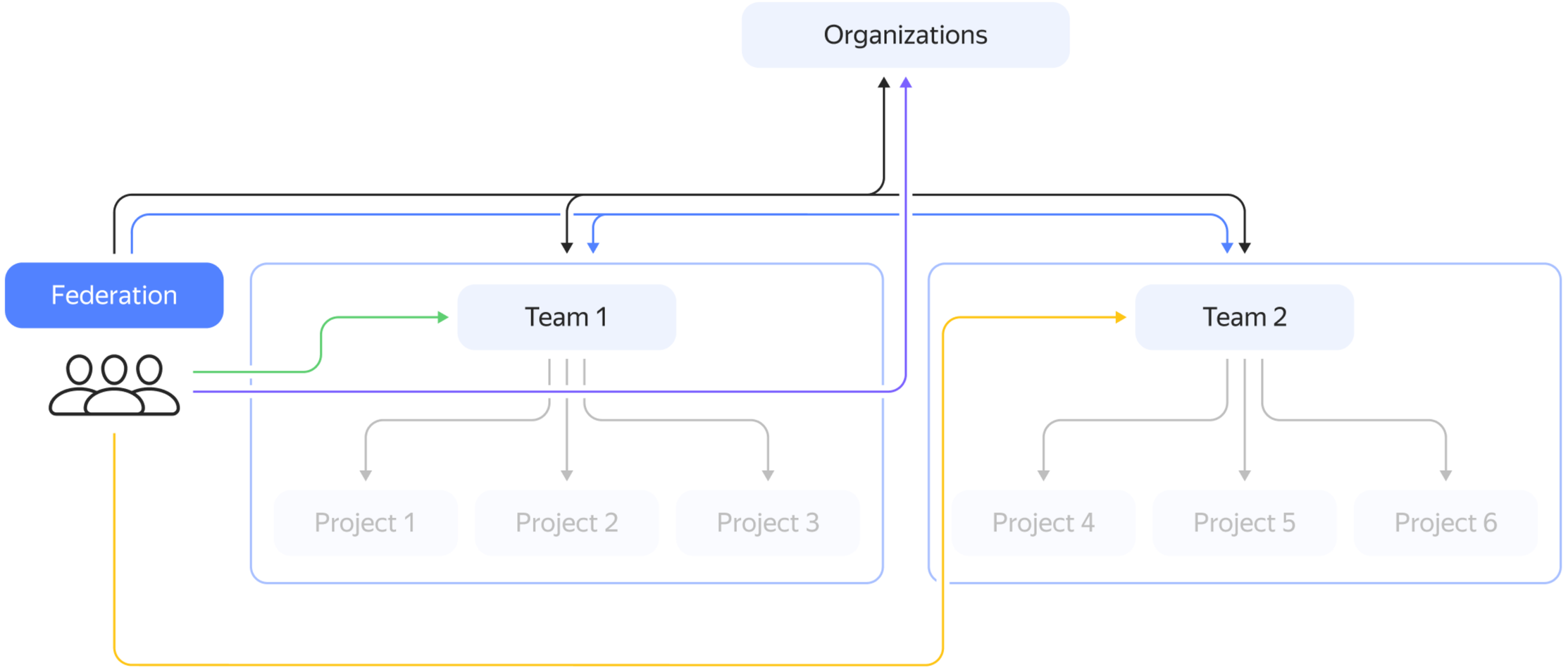
G Suite, OneLogin

Поддерживается в Kubernetes RBAC

Пример структуры



Как работает SAML-федерация



1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Сервисы платформы данных

Managed Service for PostgreSQL
Управление БД PostgreSQL

Managed Service for ClickHouse
Управление БД ClickHouse

Managed Service for Apache Kafka®
Управление кластерами Apache Kafka®

Managed Service for Redis™
Управление БД Redis™

Managed Service for MongoDB
Управление БД MongoDB

Managed Service for MySQL®
Управление БД MySQL®

Data Transfer
Инструмент для миграции БД

Managed Service for Elasticsearch
Управление кластерами Elasticsearch

Managed Service for SQL Server™
Управление базами данных SQL Server™

DataLens
Визуализация и анализ данных

Managed Service for Greenplum® Preview
Управление БД Greenplum®

Yandex Database
Распределённая отказоустойчивая СУБД

Message Queue
Очереди для организации обмена сообщениями между приложениями

Monitoring
Сбор и визуализация метрик

Data Streams
Управление потоками данных

Data Proc
Управление кластерами Apache Hadoop®

Object Storage
Масштабируемое хранилище данных

Yandex Data Platform

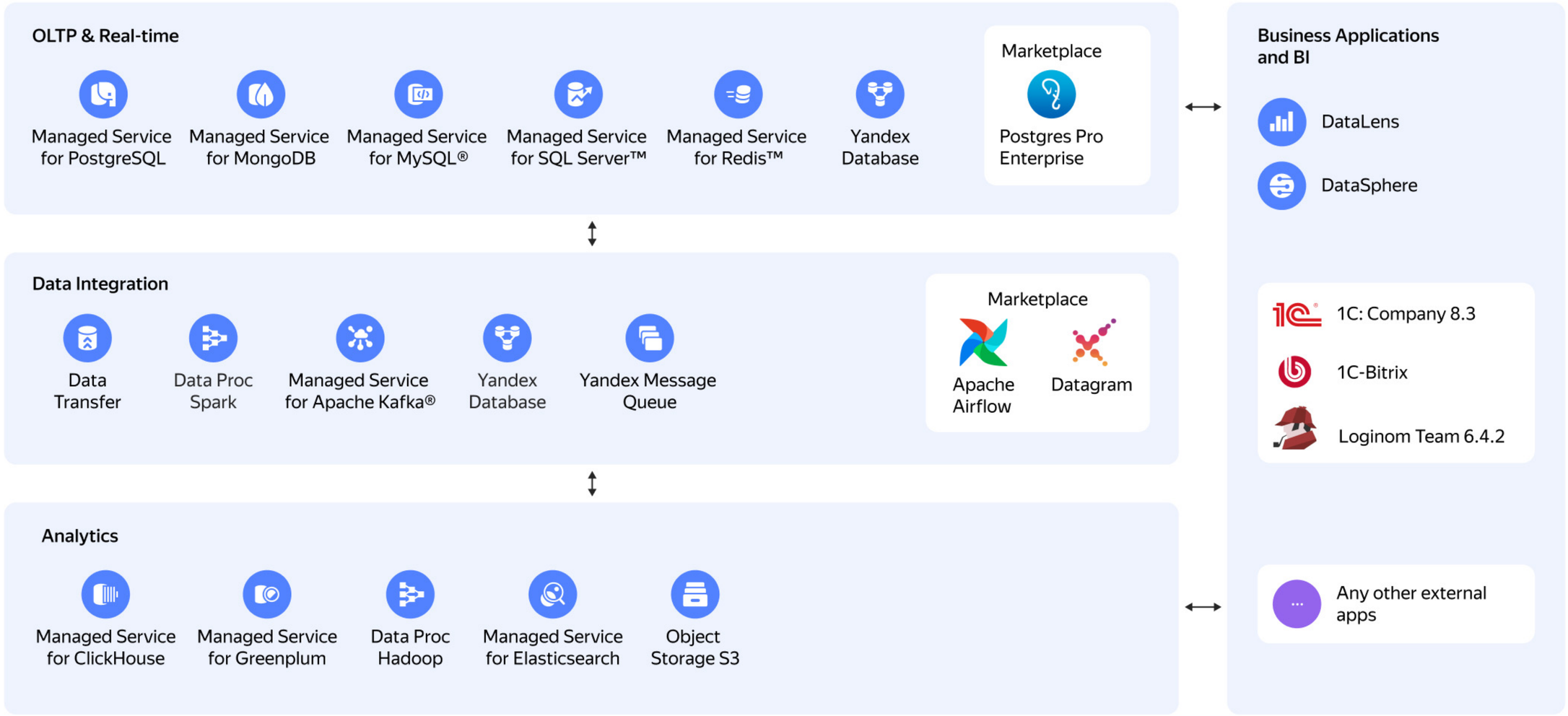
Экосистема сервисов для работы с данными

- Лучшие сервисы для хранения и обработки данных, используемые ведущими мировыми компаниями*
- Перекрёстно интегрированные сервисы — дополнительная разработка не нужна
- Простая в использовании консоль снижает технологический барьер
- Возможности кастомизации: применение конкретных интеграций и компонентов в соответствии с вашей проектной архитектурой
- Защита данных соответствует самым строгим требованиям федерального закона ФЗ № 152, а также индустриальных стандартов GDPR, ISO и PCI DSS

* Примеры: Uber, Tesla, Wallmart, Bloomberg, CERN и др.

Платформа данных

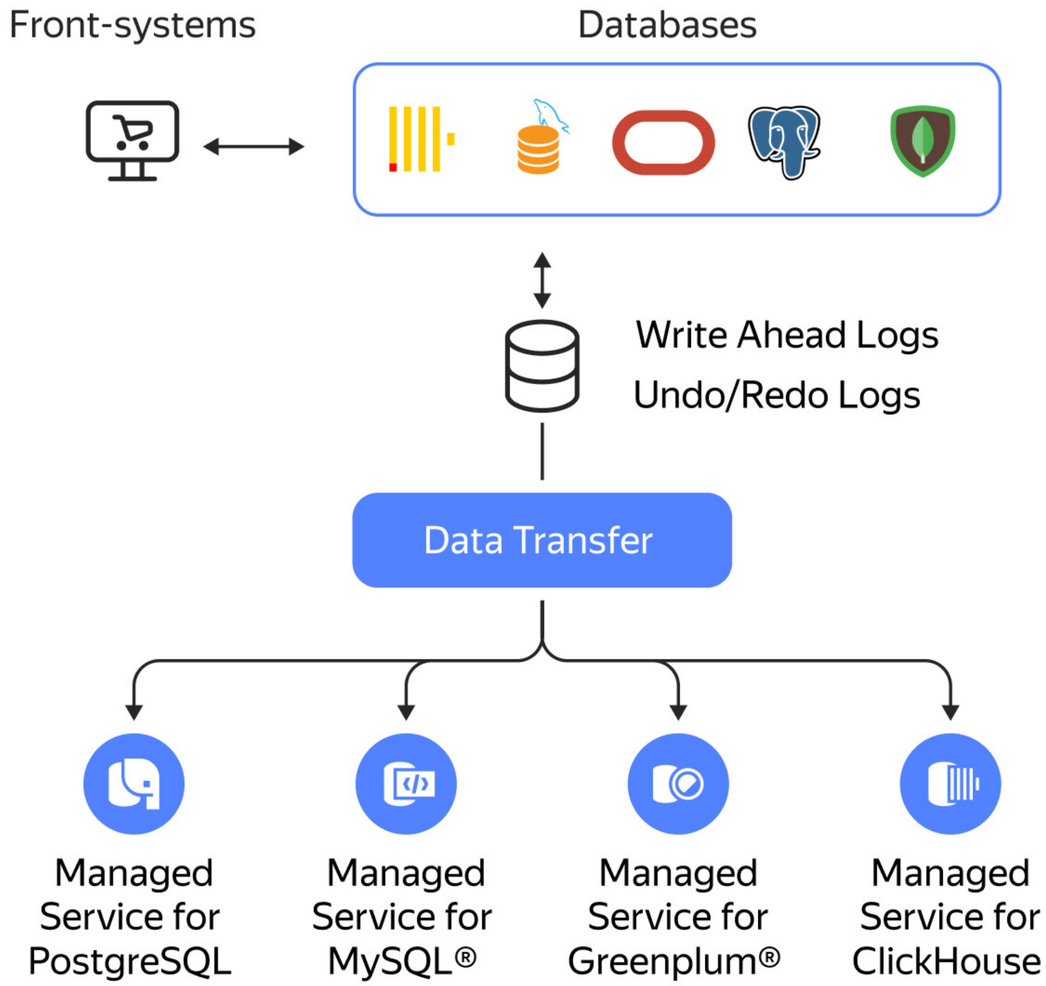
Yandex.Cloud Data Platform



Изначальная поддержка Change Data Capture (CDC)

Сбор данных из СУБД

- Нет снижения производительности из-за доступа к БД-источникам
- Нет необходимости в прямом доступе к таблицам (отсутствие полного сканирования)
- Репликация происходит в режиме реального времени

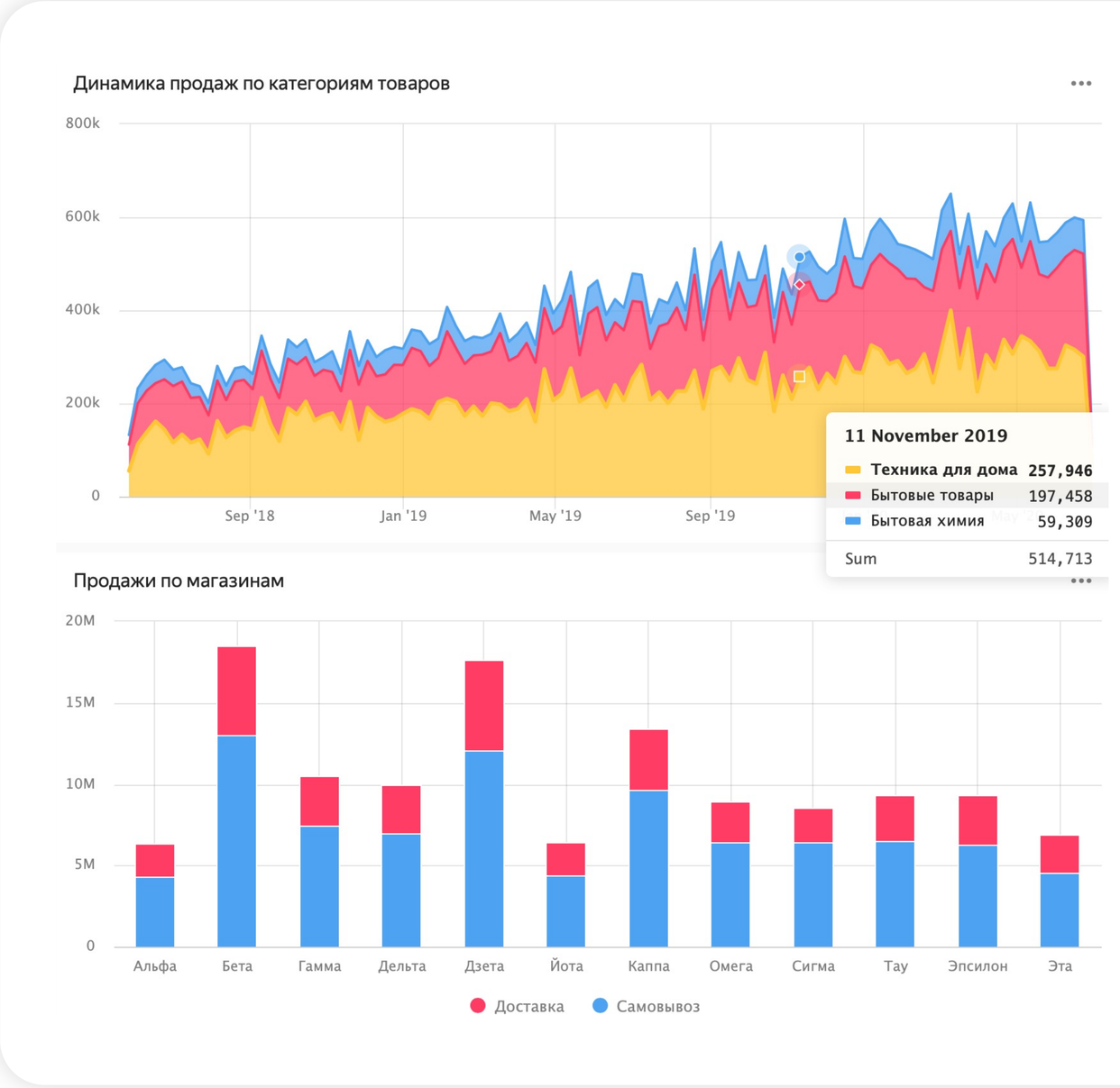


Платформа данных

Yandex DataLens

Бесплатный облачный инструмент корпоративной бизнес-аналитики с возможностью визуализации данных

- Быстрая проверка гипотез на реальных данных
- Сбор ключевых бизнес-метрик из различных источников на едином дашборде
- Обсуждение результатов анализа внутри команды, с партнёрами и клиентами — по ссылке



Наш опыт работы на Yandex Data Platform

Одни только PostgreSQL БД нашей команды занимают более 2 ПБ данных, осуществляя более 1 млн транзакций в секунду

Elasticsearch

Jupyter

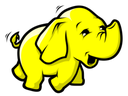
Redis

ClickHouse

Greenplum Database

Postgre SQL

Mongo DB



Hadoop

ClickHouse

Мы разработали высоко-производительную аналитическую СУБД, обрабатывающую миллиарды строк в секунду и применяемую более чем в 2 тыс. компаний по всему миру

Партнёрства

Мы сотрудничаем с ведущими мировыми вендорами, включая MongoDB, Elasticsearch, Microsoft и др.

Проекты с открытым исходным кодом

Мы участвуем во многих других опенсорсных проектах. Например, мы создали Odyssey — пулер подключений, а также разрабатываем систему резервного копирования WAL-G

Внутренние проекты

Сервисы Yandex Data Platform используются во многих продуктах Яндекса, включая Почту, Такси, Драйв, Диск и др. Мы понимаем, чего хотят пользователи, и даём им это

Примеры использования Data Platform

Розничная торговля

- Оптимальное предложение: системы рекомендаций для сайтов и мобильных приложений
- Бизнес-отчётность
- Программы лояльности
- Снижение розничных остатков
- Автоматизация склада
- Снижение расходов на мерчандайзинг
- Снижение доли возвратов
- Локализация цепочек поставок

Банки

- Скоринг
- Customer 360
- Регуляторная отчётность
- Управленческая отчётность
- Продуктовая аналитика
- Борьба с мошенничеством
- Обогащение внешних данных
- Умное страхование

Телекоммуникации

- Безопасность
- Системы рекомендаций
- Продуктовая аналитика
- Скоринг
- Планирование сетевых нагрузок
- Борьба с мошенничеством
- Снижение оттока клиентов
- Геотаргетинг
- Прогноз нагрузки на поддержку

Производство

- Прогноз среднего времени между значениями
- Продуктовая аналитика
- Таргетинг БД с устройств
- Контроль доступа
- Снижение пробега транспортных средств
- Борьба с мошенничеством
- Оптимизация логистики
- Оптимизация производства

Общая стоимость владения решением on-premise по сравнению с Yandex.Cloud

Пример: БД PostgreSQL

On-premise

Сервер 24 ядра, 96 ГБ RAM,
500 ГБ SSD

Размещение серверов
в дата-центре

Затраты на администратора БД

Затраты на системного
администратора

Не включены: поддержка
и лицензии для ОС и сетей

4 252 812 руб.

Yandex.Cloud

ВМ PostgreSQL:
24 ядра, 96 ГБ RAM,
500 ГБ SSD

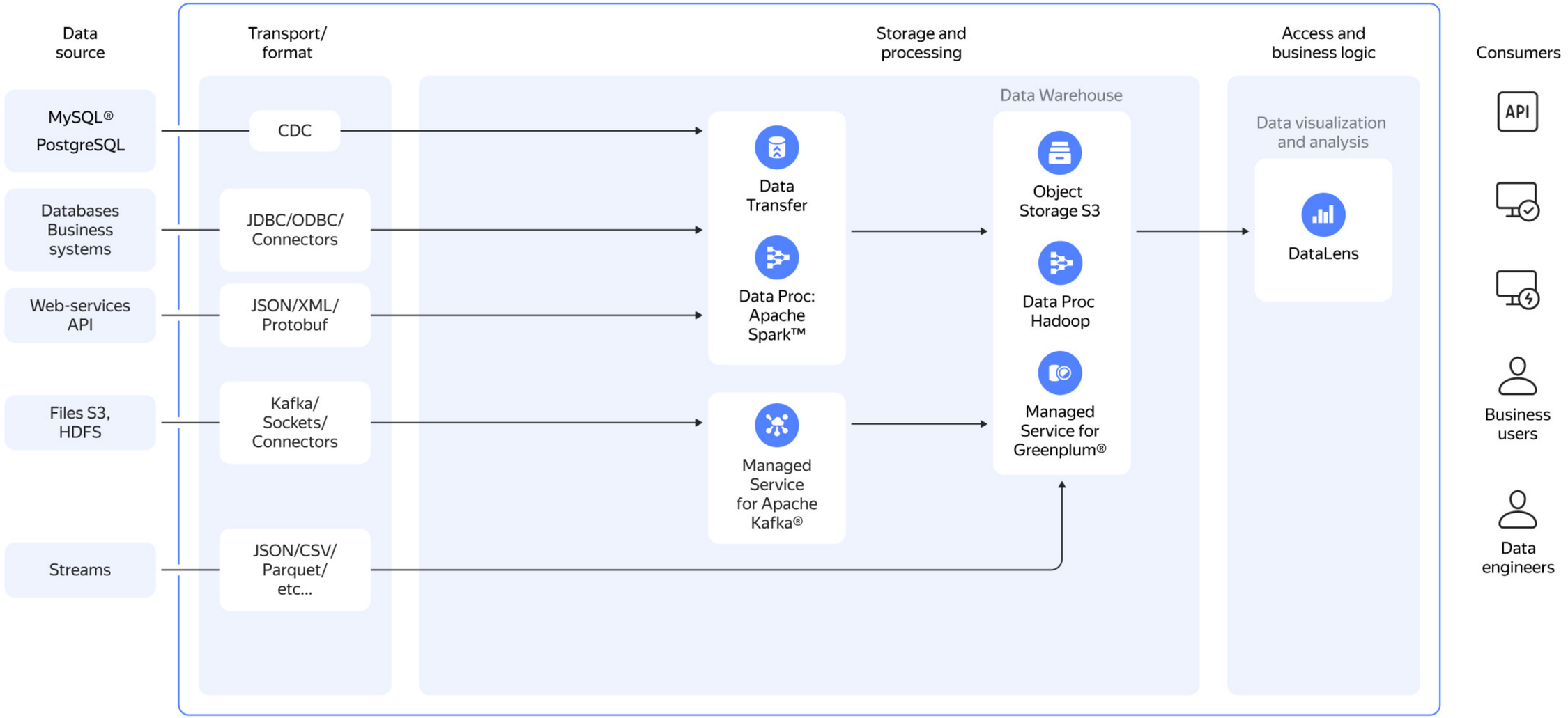
1 489 965 руб.

65%

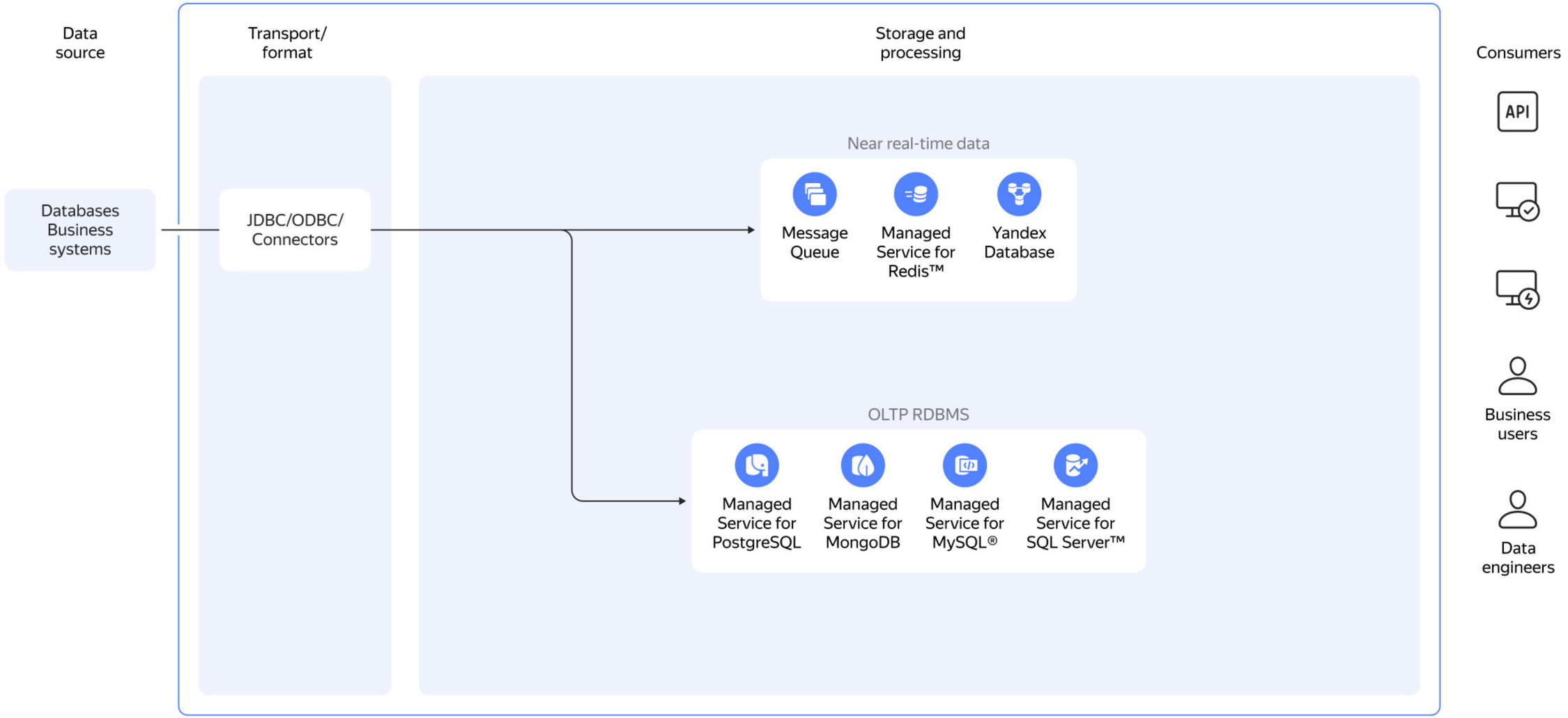
экономия за три года

Примеры использования Data Platform

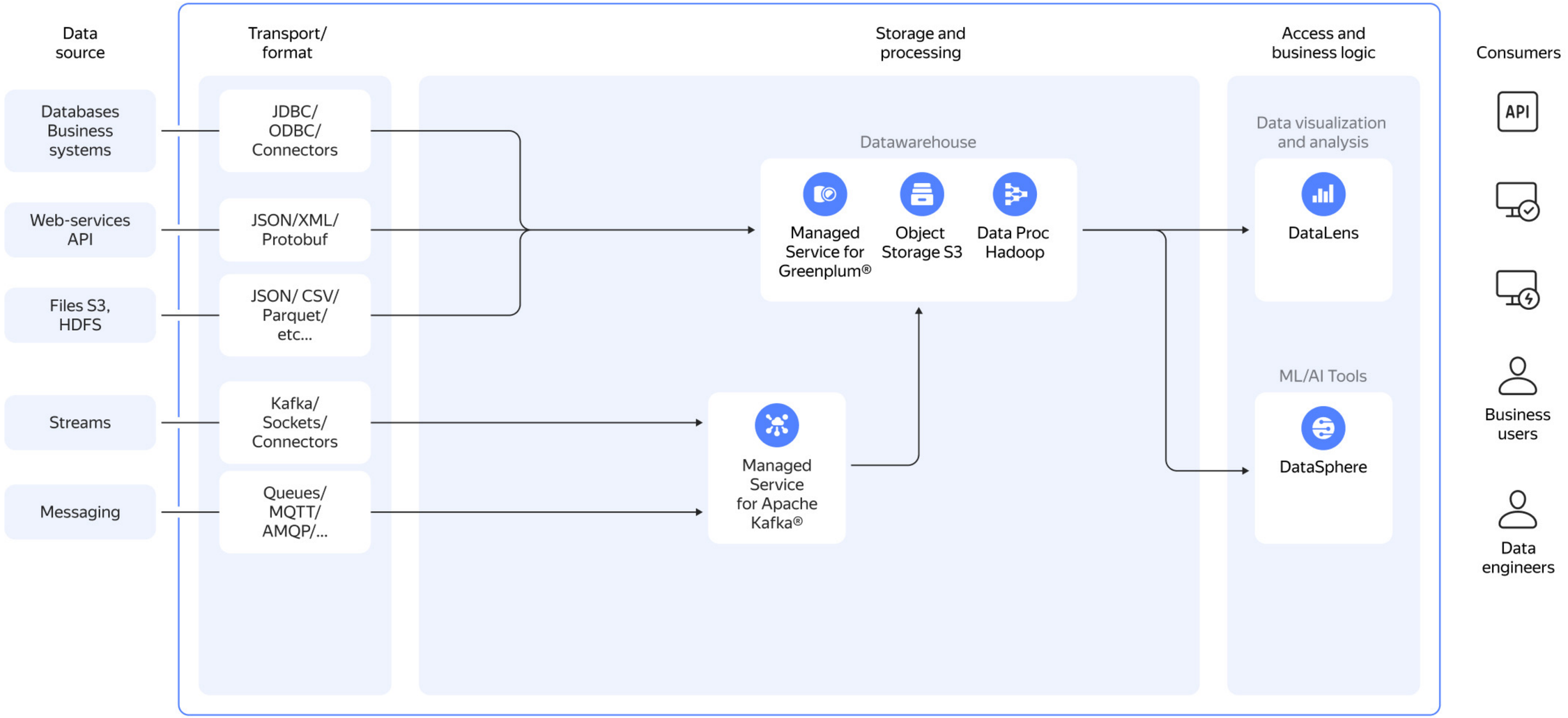
Хранилище данных



Система администрирования БД для прикладного бэкенда

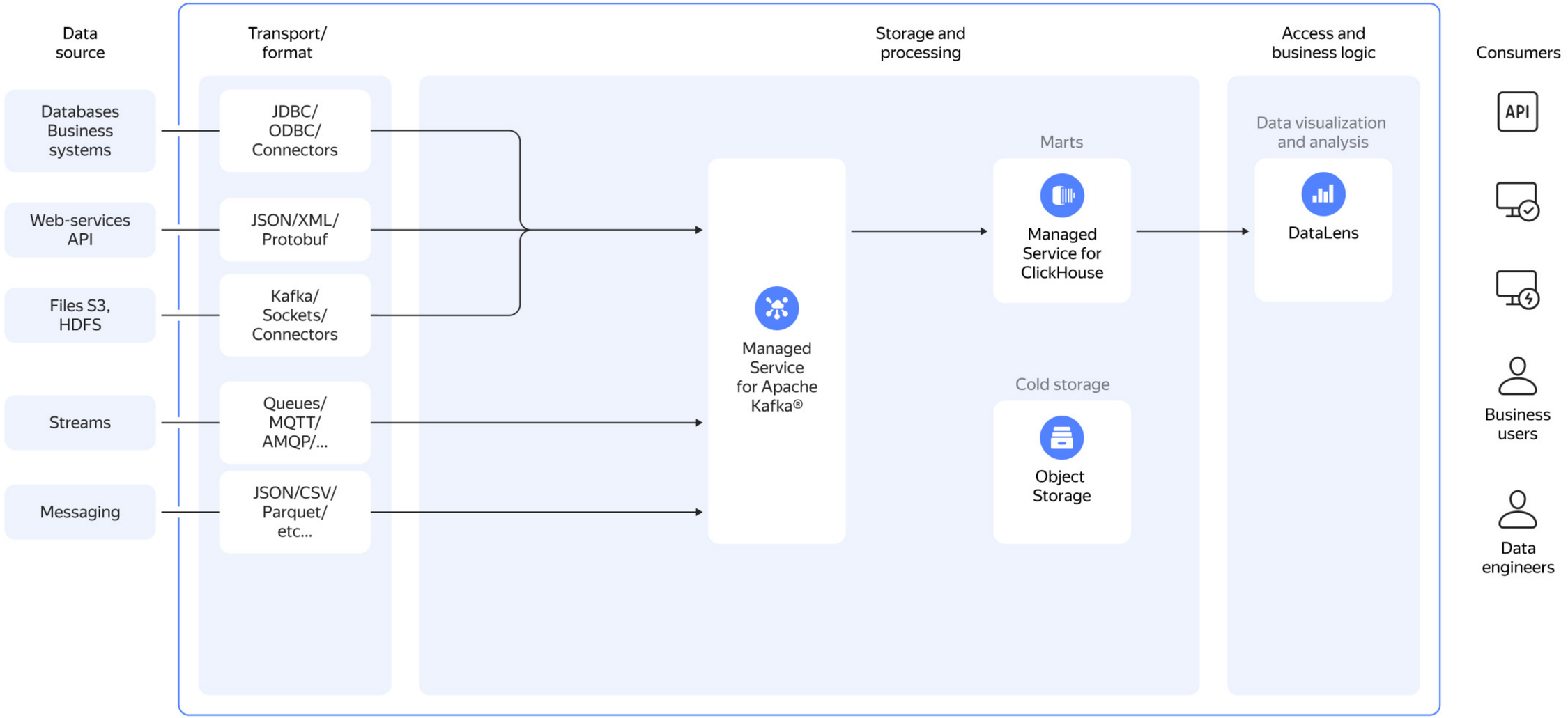


Data science и машинное обучение



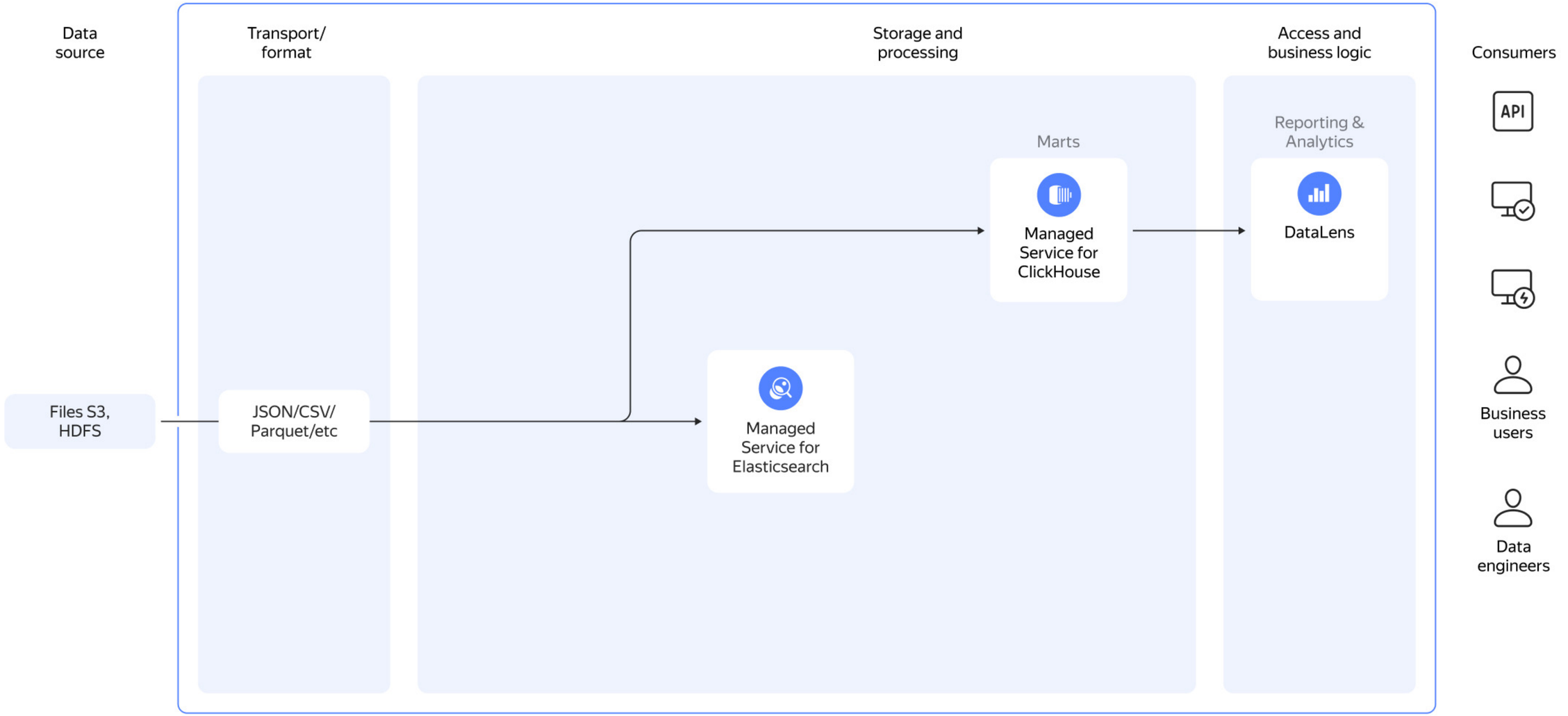
Платформа данных

Аналитика в режиме реального времени



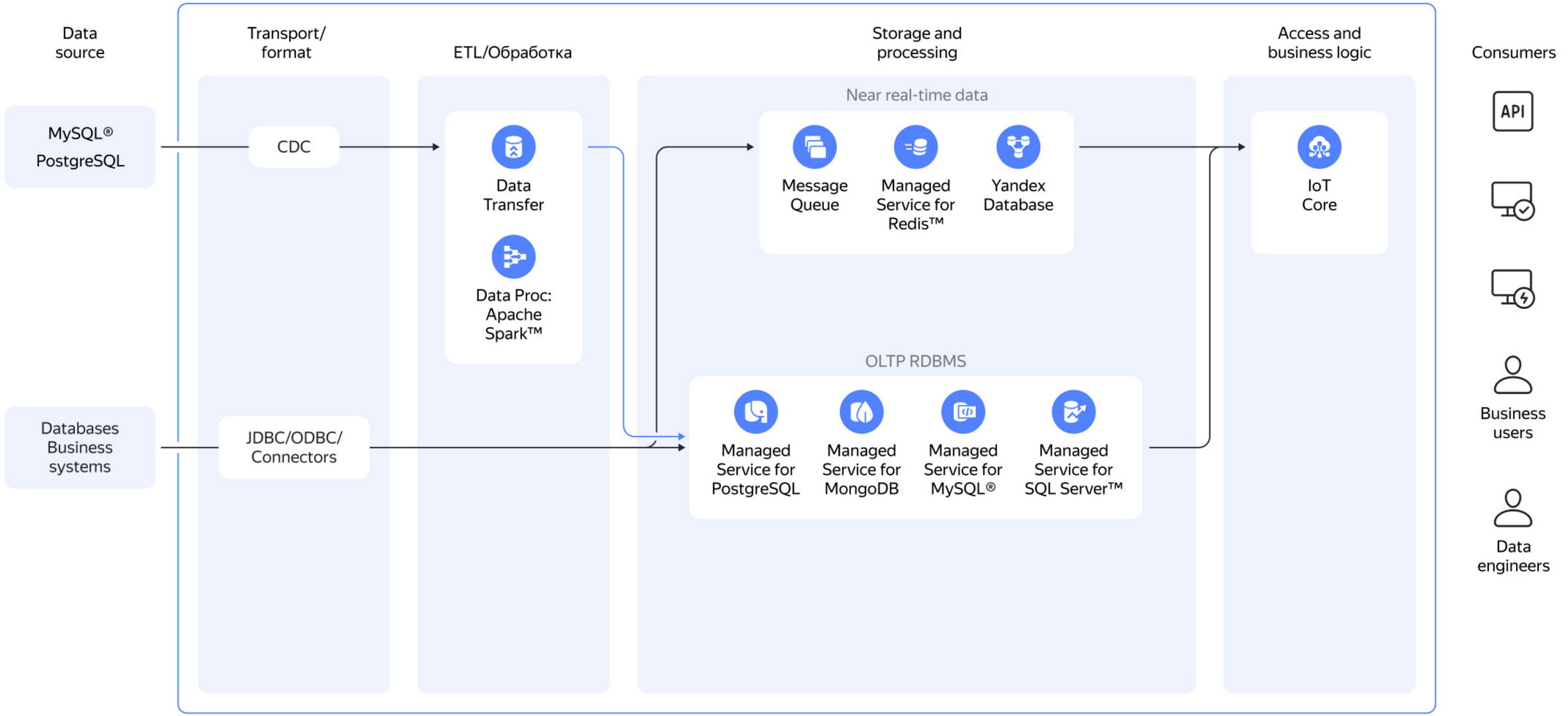
Платформа данных

Анализ логов

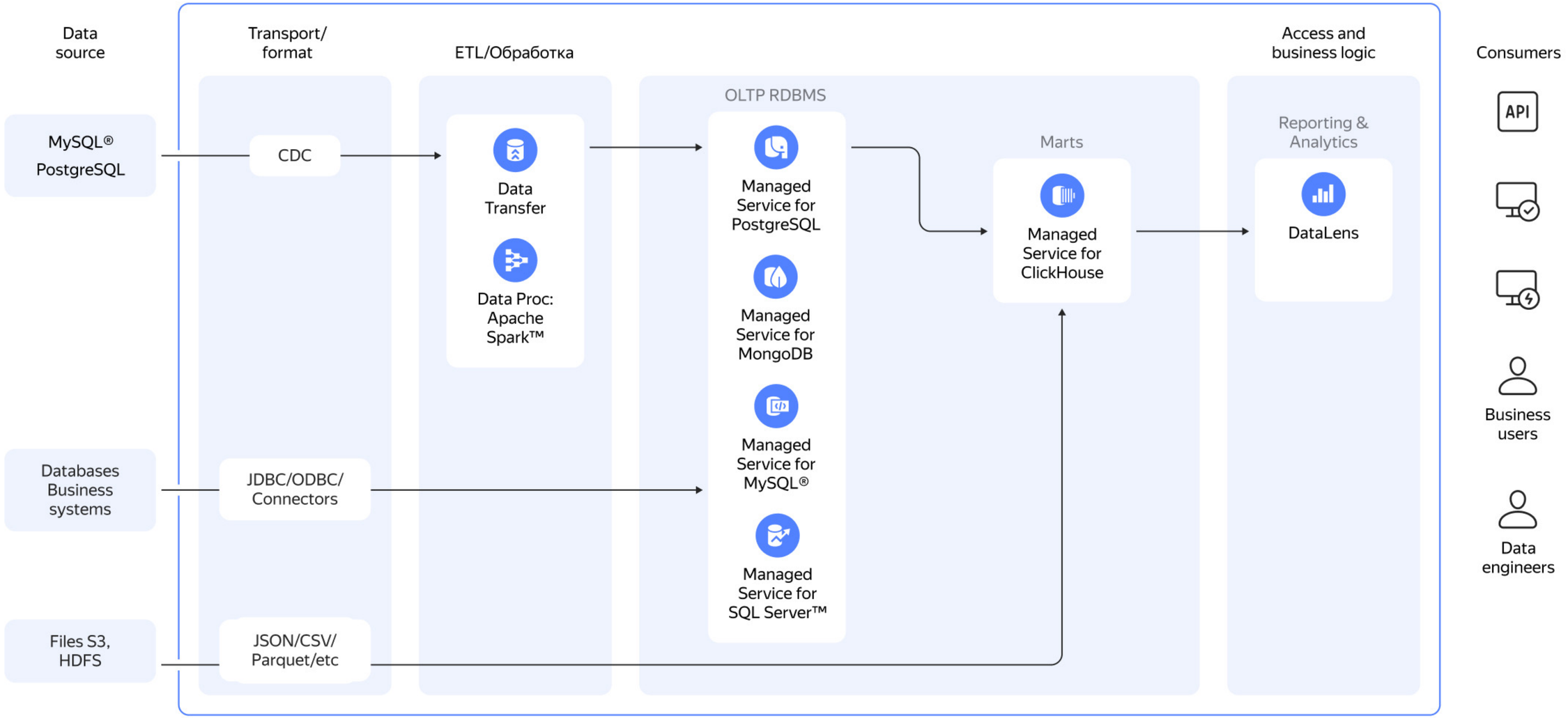


Платформа данных

IoT-хаб



Аналитическое хранилище данных



1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Инструменты для бессерверных вычислений

Cloud Functions
Выполнение программного кода как функции

IoT Core
Решения для интернета вещей

Object Storage
Масштабируемое хранилище данных

Message Queue
Очереди для организации обмена сообщениями между приложениями

Yandex Database
Распределённая отказоустойчивая СУБД

API Gateway
Интеграция с сервисами Yandex.Cloud

Serverless Containers
Запуск контейнеров без Kubernetes

Data Streams
Управление потоками данных

Serverless

Разработка приложений, хранение данных и настройка интеграций с другими платформами без создания виртуальных машин и обслуживания инфраструктуры

- Способ запуска вычислений
- Полностью управляемая среда
- Платформа отвечает за конфигурацию, масштабирование, отказоустойчивость и безопасность
- Запуск вычислений по требованию
- Оплата только за потреблённые ресурсы

Cloud Functions

Сервис Cloud Functions позволяет запускать приложения в защищённой, отказоустойчивой и масштабируемой среде без создания или обслуживания виртуальных машин

На чём можно писать функции:


JavaScrypt


R


PHP


Python


Bash


Java


C#


Go

Сценарии использования Serverless

Чат-боты

Микросервисные приложения

Бэкенд для мобильных приложений

Одностраничные приложения

Поддержка IoT-приложений

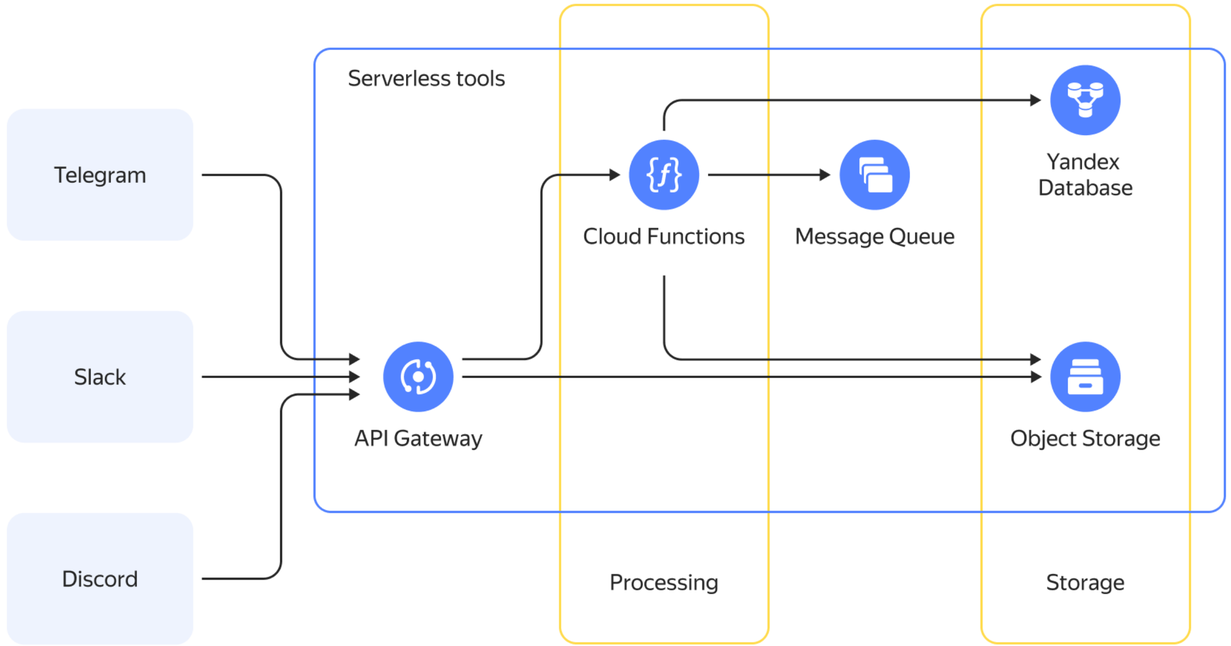
Ввод данных в системы хранения

Трассировка сервисов при помощи Jaeger

Chat bots

Методы ChatOps для поддержки ваших DevOps-процессов

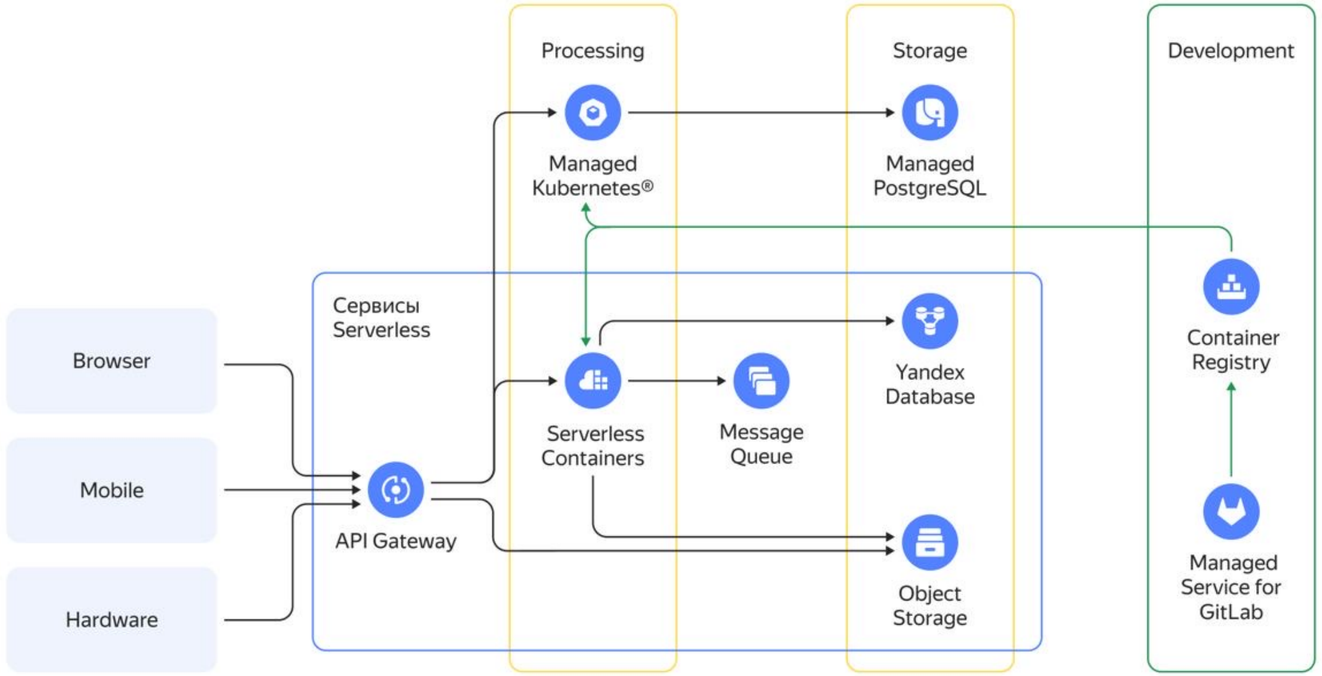
- Событийную модель обработки событий легко реализовать с помощью набора отдельных функций
- Реализация скрыта от внешних наблюдателей благодаря API Gateway
- Хранение данных можно реализовать с помощью Object Storage и Yandex DataBase, запущенной в serverless-режиме



Application backend

Экосистема сервисов для надёжных, простых
в управлении, масштабируемых serverless-проектов

- Бессерверные решения,
которые можно развернуть рядом
с существующей инфраструктурой
- Возможность перераспределять
внешние вызовы к созданным
ресурсам за счёт API Gateway

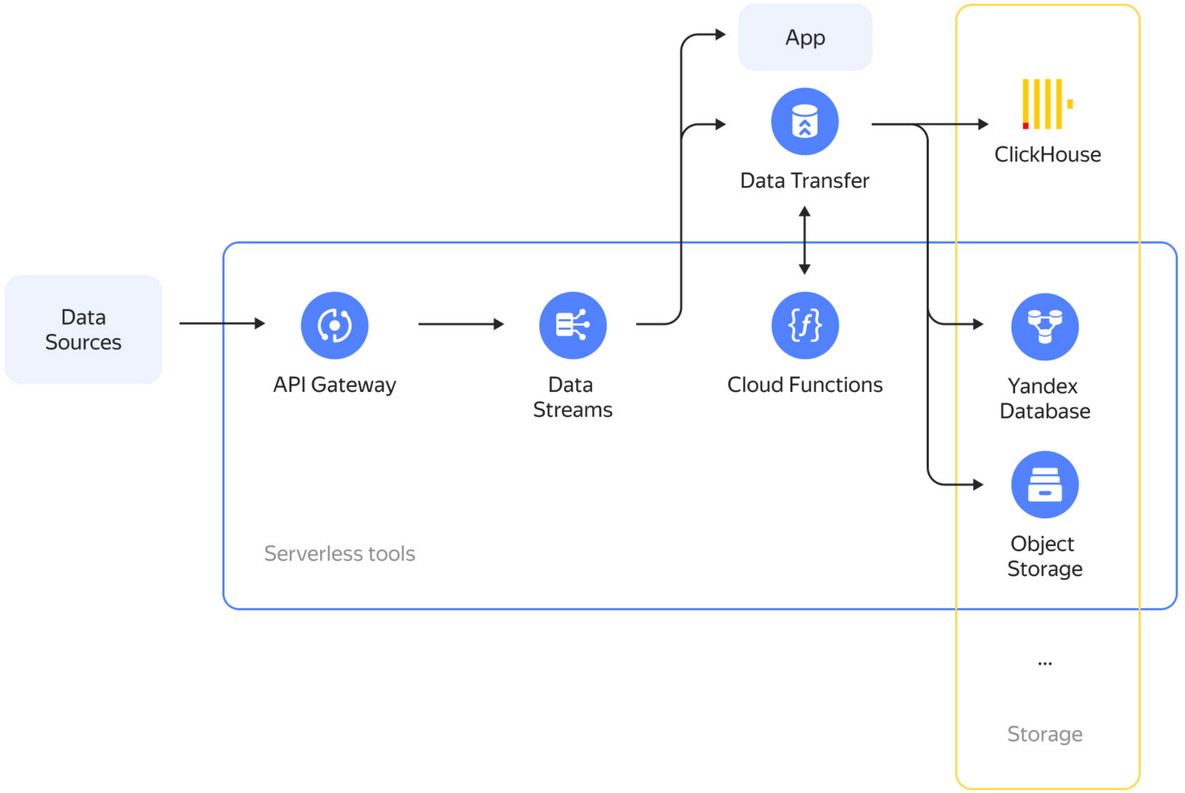


Ввод данных в системы хранения

Yandex Data Streams как шина потоков данных обеспечивает оптимальные режимы работы источников и приёмников

Data Streams:

- принимает входящие данные с высокой частотой и скоростью, не блокируя источники
- сохраняет принятые данные у себя
- формирует пакеты данных и отправляет их в принимающие системы, снижая нагрузку на них



1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Инструменты для машинного обучения и искусственного интеллекта

DataSphere
Разработка модели машинного обучения, обучение модели и аналитика данных

Translate
Машинный перевод с поддержкой более 90 языков

SpeechKit
Распознавание и синтез речи

Vision
Анализ изображений с использованием моделей машинного обучения

SpeechKit

Распознавание и синтез речи на нескольких языках. Голосовой помощник Алиса использует голосовые технологии, адаптированные к реальным бизнес-решениям

- Контекстное распознавание
- Синтез в реальном времени
- Разработка уникального голоса для вашего бренда
- Гибридные технологии для конфиденциальных данных (развёртывание на собственной инфраструктуре)
- Поддержка русского, английского, турецкого и казахского языков (планируется поддержка других европейских языков)
- Обширная партнёрская сеть

SpeechKit Hybrid

Полномасштабный контроль конфиденциальности информации в сочетании с облачными возможностями масштабирования

Вариант 1
Для оптимизации начислений и ресурсов, наиболее конфиденциальные данные обрабатываются на вашей стороне, а все остальные – в облаке

Вариант 2
Все данные обрабатываются внутри вашей инфраструктуры. Фиксированные платежи в расчете на канал или облачный биллинг на основе телеметрических данных

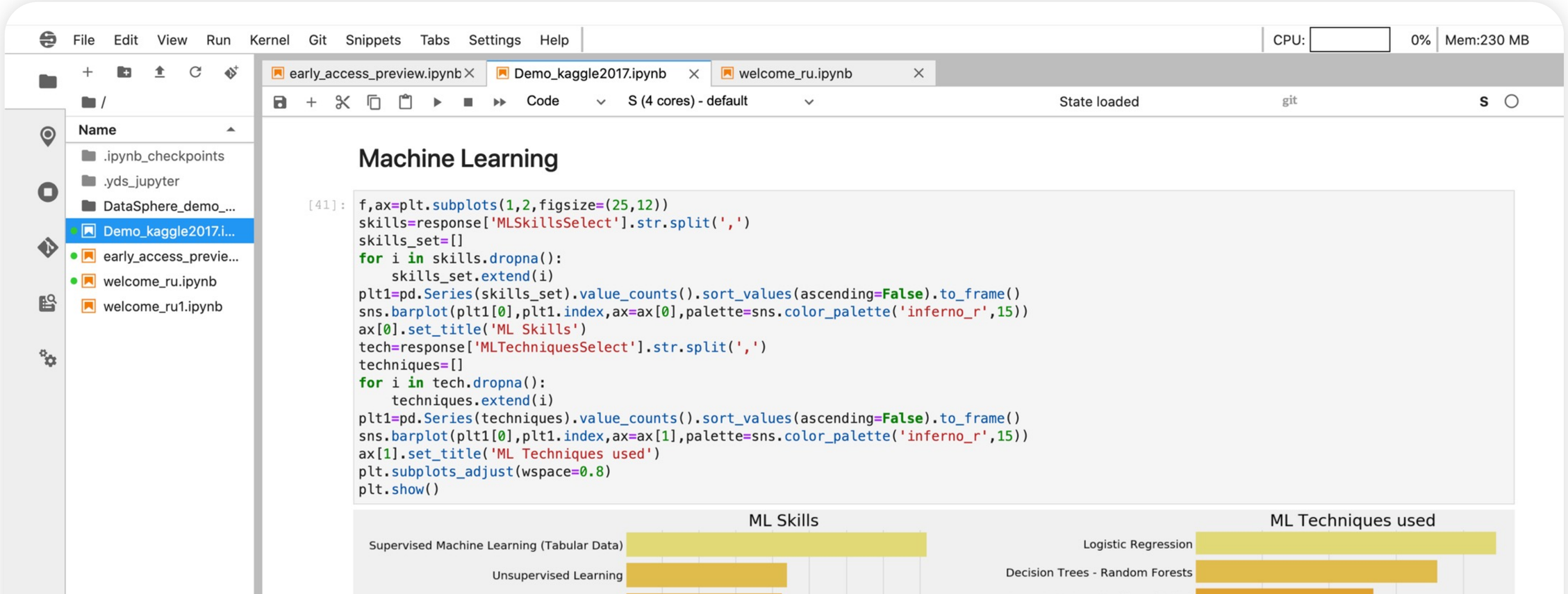
Машинное обучение и искусственный интеллект

Yandex DataSphere

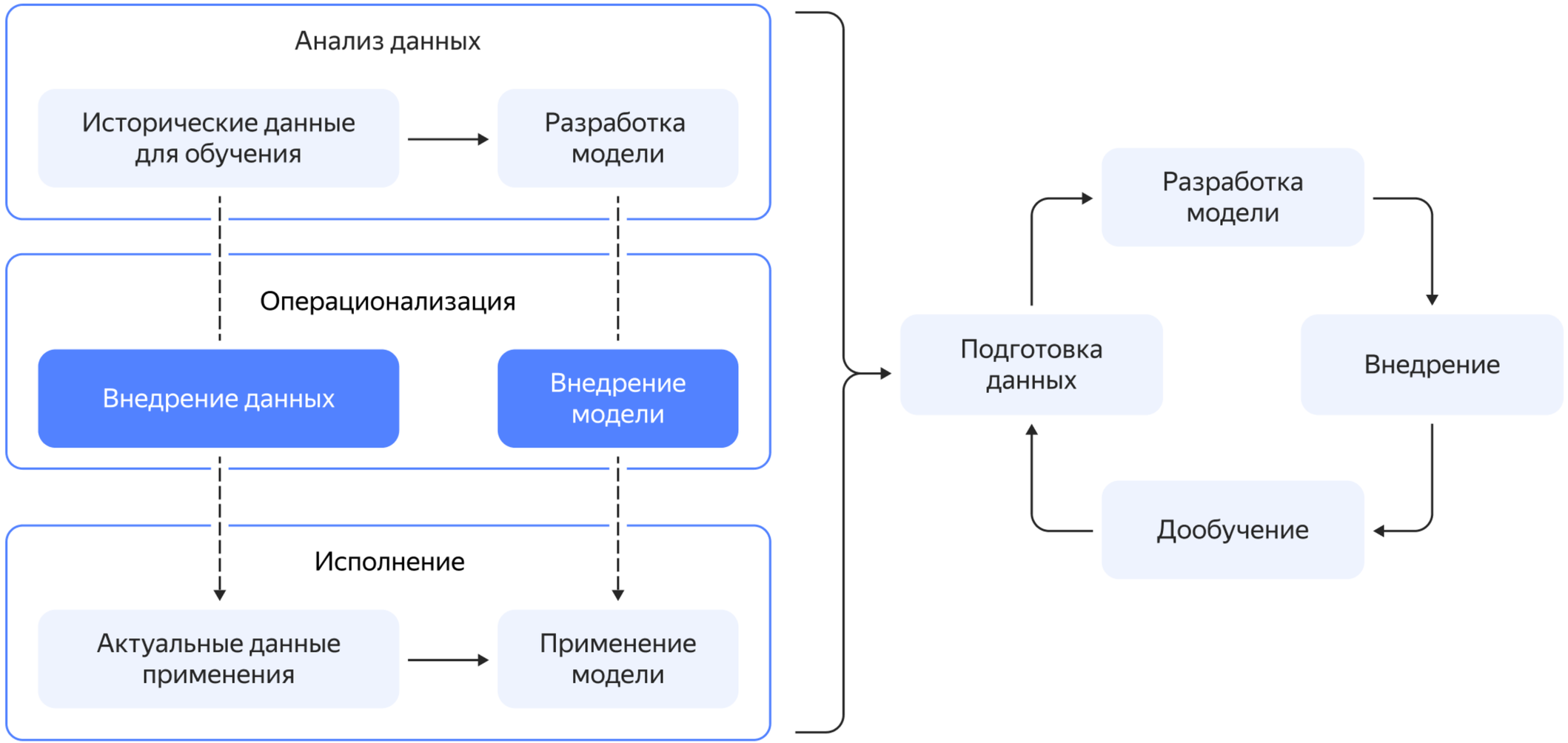
Полный цикл машинного обучения

Сервис для ML-разработки, предоставляющий самые необходимые инструменты и масштабируемые ресурсы для реализации полного цикла машинного обучения — от эксперимента до построения окончательной модели

- Дружественный интерфейс
- Развёртывание в один клик
- Обучение модели как услуга
- Бессерверные вычисления



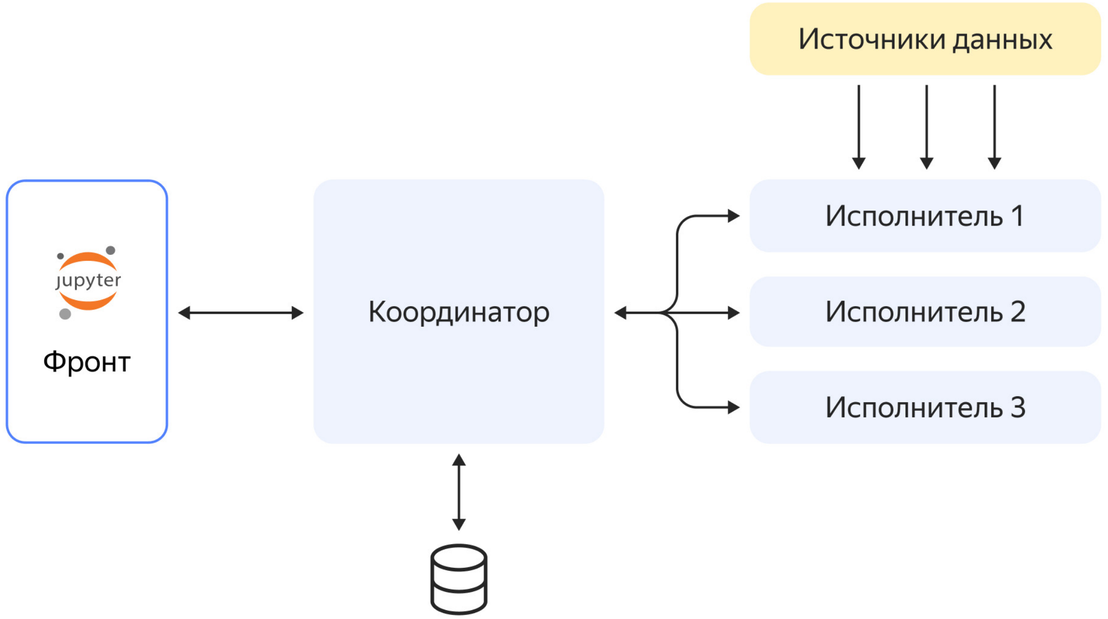
Полный цикл разработки машинного обучения



Контур разработки в DataSphere

Привычный интерфейс, новая реализация

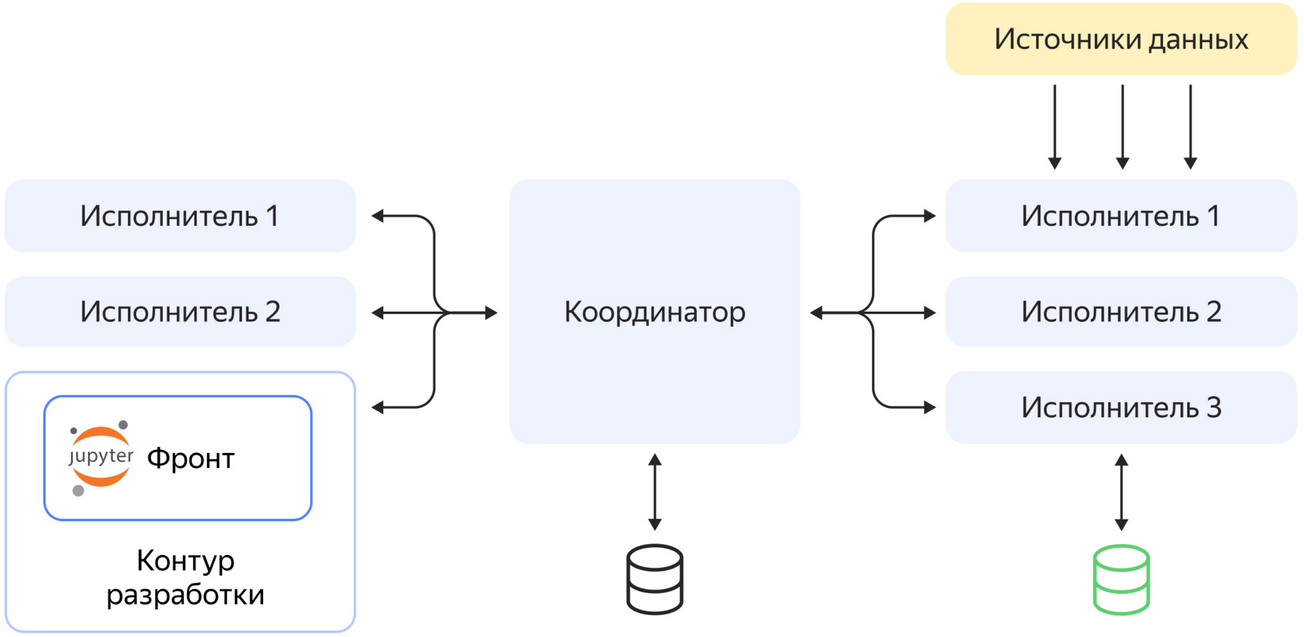
- В качестве интерфейса взяли привычный Jupyter Notebook
- Делегируем исполнение ячеек распределённой системе
- Координатор делегирует исполнение ячеек
- Координатор синхронизирует исполнение и доступ к данным



Контур эксплуатации в DataSphere

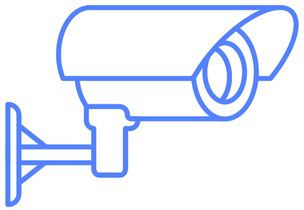
Ячейка определяет микросервис

- Используем те же исполнители, что в контуре разработки
- Необходимые для работы ресурсы доставляются тем же механизмом
- Координатор осуществляет балансировку и поставляет данные
- Граф исполнения определяется специальной ячейкой



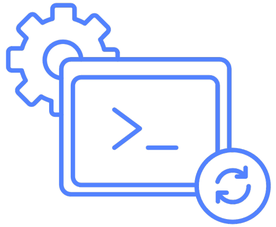
1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Безопасность при разработке и функционировании Yandex.Cloud



Физическая безопасность

- Техническое обслуживание серверов строго регламентируется
- Объекты находятся под постоянным видеонаблюдением
- При доступе к носителям данных, а также при хранении и уничтожении носителей применяются дополнительные меры безопасности



Безопасность разработки

- Сотрудники, участвующие в разработке облачных сервисов, регулярно проходят обучение в области безопасности
- Аудит безопасности кода приложений
- Политика управления обновлениями, задающая максимальное время установки для каждого типа ПО

Шифрование данных

- На всех облачных сервисах данные хранятся в зашифрованном виде
- Данные, передаваемые через интернет, защищены протоколом TLS

Сервисы в области безопасности

Identity and Access Management
Идентификация пользователей и контроль доступа к облачным ресурсам

Lockbox Preview
Создание и хранение секретов

Key Management Service
Управление криптографическими ключами

Аппаратный модуль шифрования (HSM) как услуга Pre-preview
Управление криптографическими ключами

DDoS Protection
Защита от DDoS--атак

Audit trails Preview
Сервис сбора и выгрузки аудитных логов

Менеджер сертификатов
Управление TLS-сертификатами

Безопасность

Функции безопасности для решения любой проблемы

Защита инфраструктуры

Выделенные хосты Preview

Антивирусы Marketplace

Интеграция Audit Trails с SIEM

Управление доступом

Сервисные роли

Использование Active Directory через федерацию удостоверений

Управление политикой доступа (bucket policy) — ACL и сетевые политики

Организации

Автоматизация

Библиотека решений в области безопасности

Аварийное восстановление Marketplace

Защита приложений

WAF Marketplace

Сканер уязвимостей в Container Registry Preview

Сетевая безопасность

Группы безопасности Preview

Interconnect

NGFW Marketplace

NAT

GOST VPN

AntiDDoS L7

Сетевые политики Cillium в Kubernetes

Шифрование

Шифрование в Object Storage при помощи KMS-ключей

HSM-модуль для KMS-ключей Preview

HashiCorp Vault с поддержкой KMS-ключей Preview

Интеграция секретов Kubernetes с KMS и Lockbox Marketplace

Соответствие российским регуляторным требованиям и индустриальным стандартам

✅ **Стандарты ISO**
ISO 27001, ISO 27017, ISO 27018

✅ **PCI DSS**
Для облачных провайдеров и дата-центров

✅ **Cloud Security Alliance**
Платформа прошла первый этап программы Security, Trust, Assurance and Risk (STAR)

✅ **GDPR**
На данный момент Yandex.Cloud удовлетворяет ключевым требованиям ЕС

✅ **Российские нормативные акты**
ФЗ № 152 «О персональных данных»
Постановление Правительства РФ № 1119
Приказ ФСТЭК № 21

✅ **ГОСТ Р 57580**
Безопасность финансовых (банковских) операций

✅ **Реестр программного обеспечения**
Запись в реестре № 9286 от 20.02.2021

Audit trails

Сервис сбора событий безопасности в Yandex.Cloud

- Сбор событий безопасности от сервисов облака
- Обогащение событий
- Отгрузка событий в сторонние системы
  (в том числе SIEM*)

*SIEM (Security information and event management) — система для анализа событий безопасности приложений в реальном времени и реакции на них до наступления существенного ущерба

Платформа предоставляет несколько сервисов, работающих с секретами. Каждый сервис работает с отдельным типом секретов: ключи шифрования, секреты в виде пары ключ — значение и сертификаты

Key Management Service

Сервис генерирует и хранит ключ. Пользователь управляет доступом к ключу и версионностью

Payment HSM as a service

Pre-preview

HSM Thales PayShield 10K предоставляется по модели аренды и может быть использован для хранения PIN-кодов, платёжных и других чувствительных данных в рамках финансовых транзакций

Lockbox

Preview

Чтобы не хранить различные секреты (например, логин и пароль) в коде, для доступа к БД целесообразно воспользоваться сервисом LockBox, который защищает секретные значения, шифруя их на KMS-ключе

Certificate Manager

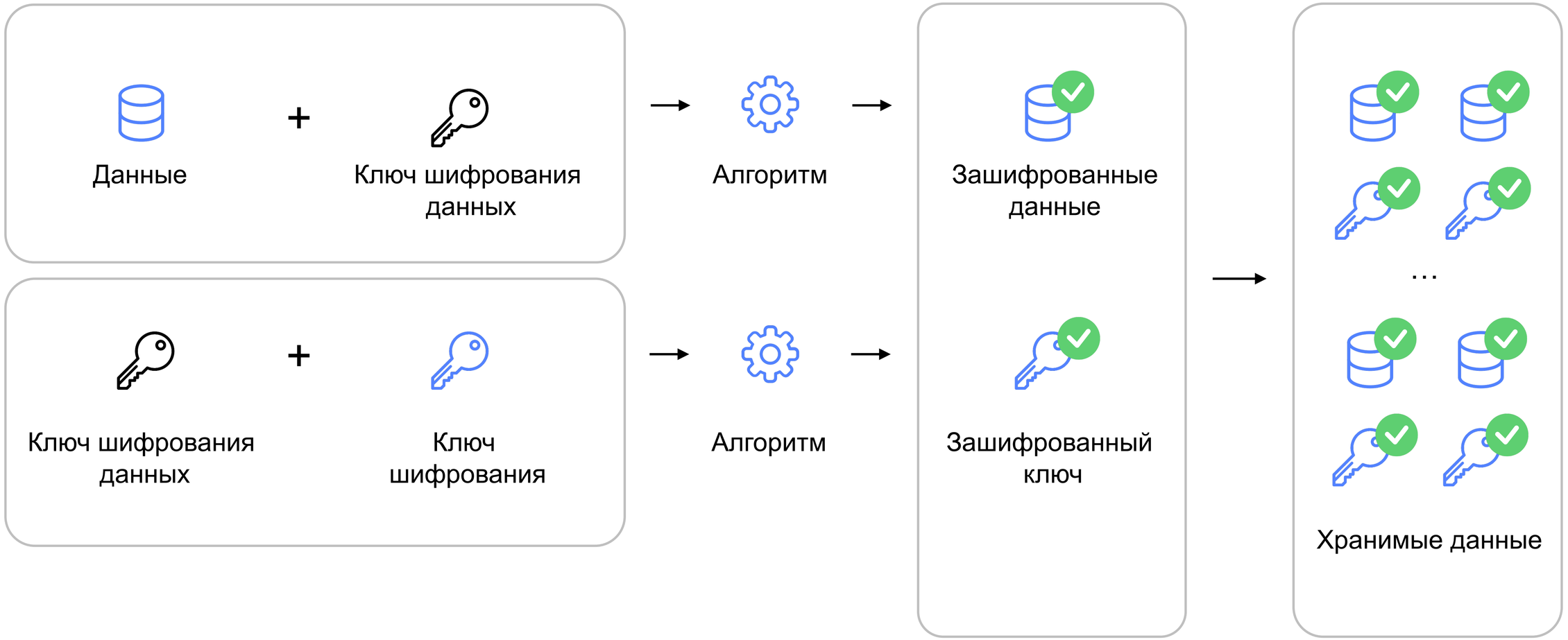
Certificate Manager управляет TLS-сертификатами. Можно получать и обновлять сертификаты Let’s Encrypt, а также добавлять собственные сертификаты

Key Management Service

Ключи шифрования, encrypt/decrypt API

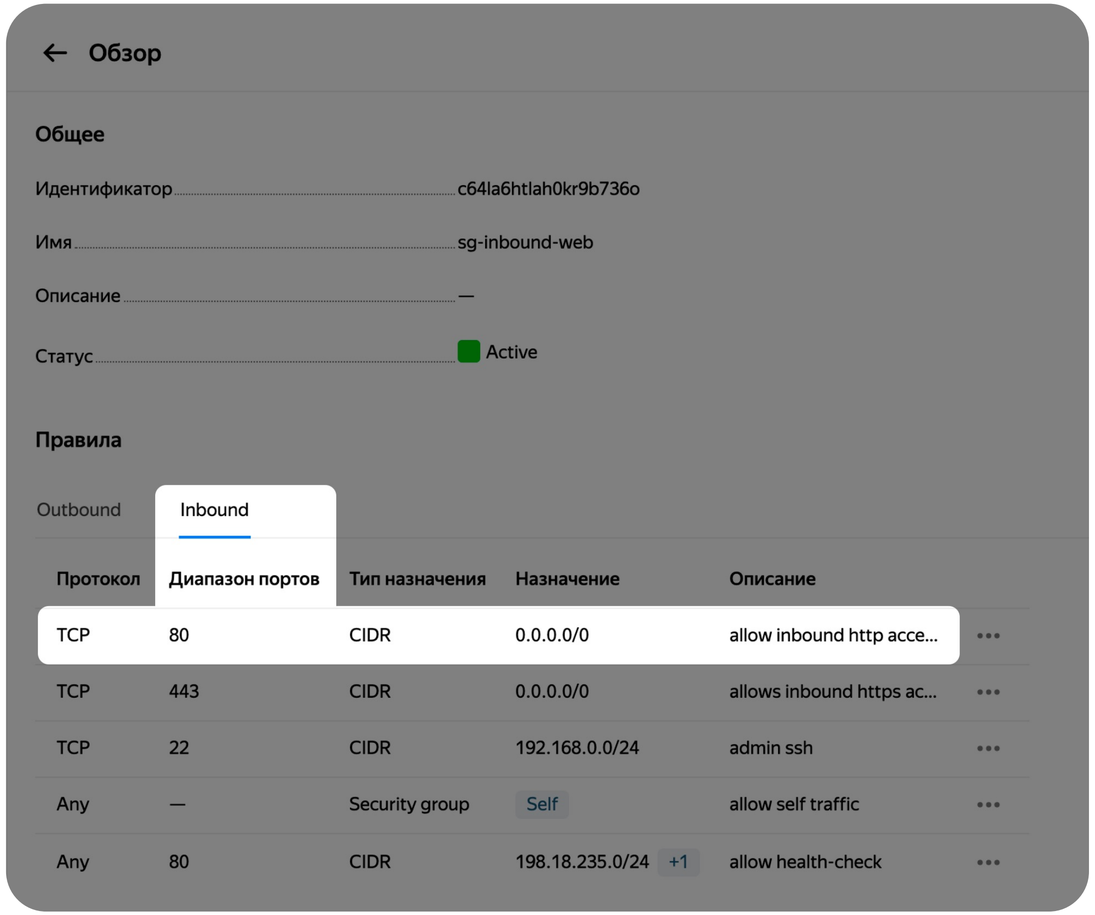
- Шифрование данных на ключах клиента
- История операций с ключами доступна пользователю
- Ключ не выходит за рамки сервиса KMS
- Возможность дополнительно защитить ключ с помощью аппаратного модуля HSM
- Доступен через API и в виде интеграций с другими сервисами Yandex.Cloud

Шифрование по схеме envelope encryption



Security Groups

- Правила для входящего трафика. Определяют диапазоны адресов и портов или другие группы безопасности, откуда VM смогут принимать трафик
- Правила для исходящего трафика. Определяют диапазоны адресов и портов или другие группы безопасности, куда VM смогут отправлять трафик
- Если в группе безопасности существует правило для исходящего трафика, ответный трафик всё равно сможет поступать на VM



* Выделенная строка разрешает входящие подключения по 80-му порту TCP с любого адреса

1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Миграция в облако состоит из двух основных компонентов

Инструменты

Виртуальные машины Compute Cloud

- Основное решение — Hystax
- VMware и Hyper-V
- Win Srv и агенты Linux
- Миграция и аварийное восстановление

Kubernetes

- Независимость от облаков и вендоров
- Версии upstream vanilla
- Уровень абстракций

БД и данные

- Основной инструмент — Data Transfer
- Простота подхода dump — restore
- Также поддерживается логическая репликация для PG и MS SQL
- Flexify.io для крупных томов S3

Люди

Архитекторы решений Yandex.Cloud

Консалтинг:

- Инструменты
- Общедоступные отработанные сценарии применения
- Плейбуки

Технологические партнёры

Реализация миграции:

- Планирование
- Оценка масштабов и сроков
- Выполнение

Профессиональные услуги Yandex.Cloud

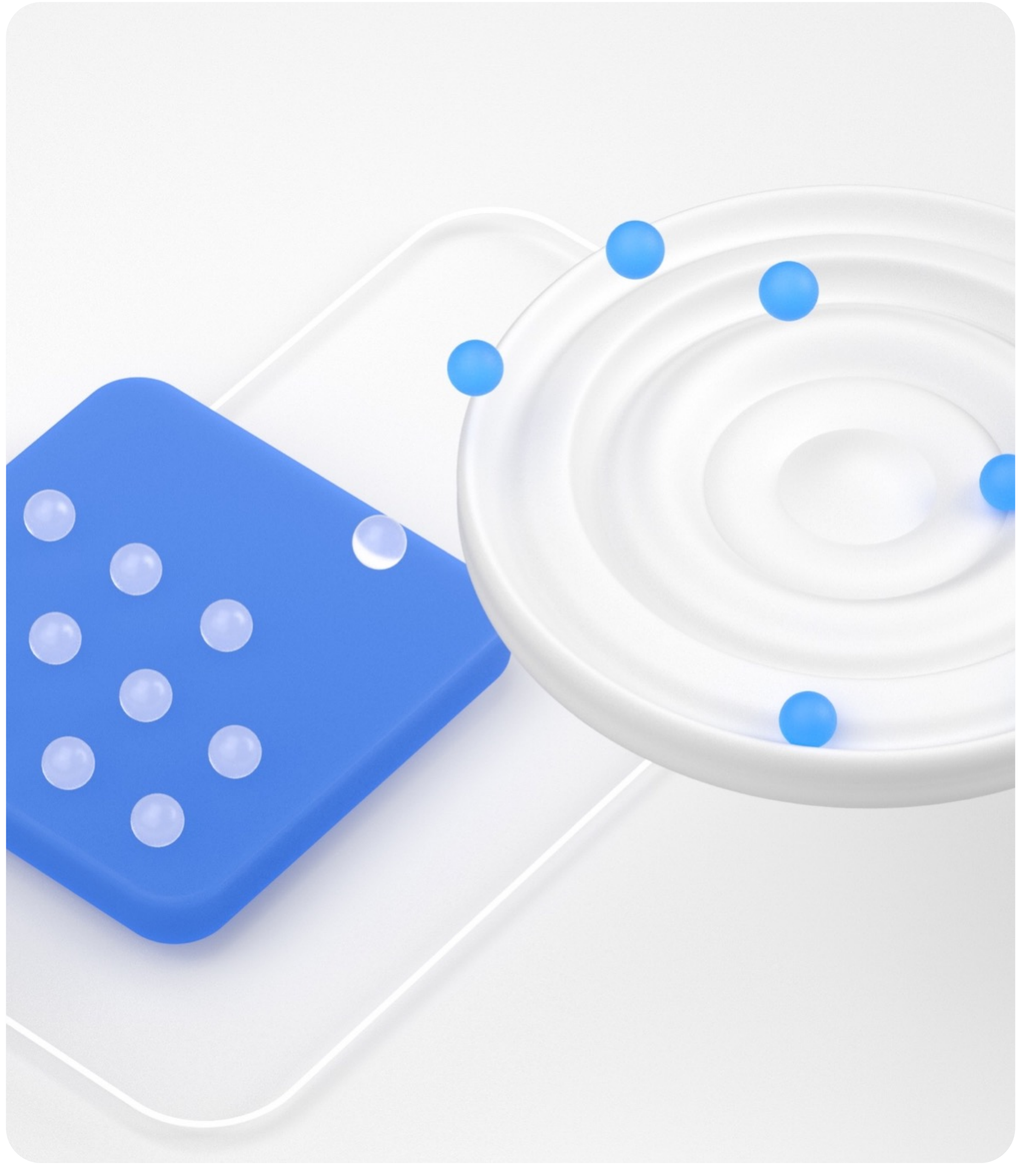
Технический контроль и обеспечение качества

- Партнёрское обучение
- Партнёрский контроль и оценка дизайна
- Эксперты включены в состав проектной команды

Миграция — это не перепроектирование

Мы больше сосредоточены на простом перемещении данных

- Перенесли → Быстро запустились
- Перепроектирование — гораздо более сложная задача



Hystax: разные виды снимков

Windows

Снимки VSS snapshot
Модули VSS writer

SQL Server,
Exchange Server

Linux

Снимок блочного
устройства
с собственным
драйвером

ESXI

Создание снимков ВМ
через VMWare API

Hystax: разные платформы

Охват платформ:

- Amazon Web Services
- Google Cloud Platform
- Microsoft Azure
- Oracle Cloud
- Alibaba Cloud
- VMware
- Hyper-V
- OpenStack
- Bare Metal

1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Альфа Банк

Голосовой помощник с уникальным характером

Задача: разработать голосового робота, который будет помогать клиентам банка с финансовыми операциями в приложении

Решение: использование технологии Yandex SpeechKit Brand Voice, позволяющей создать полную цифровую копию голоса пользователя

Результаты: помощник «Альф» с голосом актёра озвучки Всеволода Кузнецова может отвечать на вопросы клиентов банка, продолжает «обучаться» и увеличивать качество взаимодействия с клиентами

≈100%

не могут отличить «Альфа» от человека

50

часов

записи в студии для создания цифровой копии голоса актёра

SpeechKit

RENAULT

Ускорение разработки продуктов

Задача: стандартизировать подходы к разработке и администрированию

Решение: выстроили несколько уровней инфраструктуры с использованием управляемых сервисов Yandex.Cloud и cloud-native практик

Результаты: значительно ускорились циклы разработки — получение новых ресурсов в dev-окружении сократилось с часов и дней до нескольких минут. Оптимизировали вложения в развитие инфраструктуры

3-5 минут

уходит на получение новых ресурсов

100%

готовность к аудиту регуляторов

Compute Cloud · Object Storage · Message Queue · Resource Manager

М.ВидеоЭльдорадо

Обновления быстрее на 80%

Задача: сократить время внедрения изменений при работе с данными, не увеличивая стоимость

Решение: переход на облачную модель. Использование управляемых сервисов платформы данных Yandex.Cloud, ML-сервисов и BI-инструмента Yandex DataLens

Результаты: отказались от CapEx, перераспределили 50% стоимости бизнес-проекта в пользу актуальных для компании продуктов, получили возможность проводить быстрые эксперименты с данными

50%

стоимости бизнес-проекта идёт на развитие

15%

экономия на инфраструктурном стеке

3 месяца

на новый сервис ценообразования

DataLens Cloud Interconnect Managed Service for ClickHouse

Леруа Мерлен

Платформа хранения данных

Задача: построить платформу для управления данными, которые накапливались децентрализовано

Решение: использование облачных сервисов Yandex Compute для создания парка виртуальных машин и Yandex Object Storage для масштабируемого хранилища

Результаты: облачный кластер, способный принять терабайты данных, запуск тестовых стендов Hadoop, S3 и Spark

10 минут

занимает создание песочницы Greenplum, Spark или Hadoop

1 клик

нужен для увеличения количества нод в Greenplum

Compute Cloud Object Storage

Пятёрочка

Премодерация контента для «Тролли. Караоке»

Задача: быстро разработать систему премодерации, которая будет маркировать потенциально опасный аудио и видеоконтент

Решение: система интеллектуальной премодерации контента на облачной платформе, которая распознаёт элементы, которые не должны присутствовать в кадре

Результаты: более 15 тысяч видеозаписей, посланных для участия в конкурс, было обработано системой и только 6 тысяч выступлений попало в голосование на сайте

15K
роликов
обработано системой премодерации

152-ФЗ
соблюдение закона о персональных данных

Object Storage · Vision · SpeechKit · Cloud Functions

АльфаСтрахование

Ускорение бизнеса

Задача: построить гибкую и безопасную инфраструктуру и ускорить вывод новых продуктов на рынок

Решение: сделали облако надёжным и безопасным продолжением собственной инфраструктуры. Активно используются платформенные сервисы

Результаты: сократили срок вывода продуктов на рынок, научились быстро проверять бизнес-гипотезы, расширили экспертизу сотрудников и сэкономили до 30% на инфраструктуре отдельных проектов

x4

ускорение выхода новых продуктов на рынок

30%

сокращение затрат на инфраструктуру

100%

тестовых сред планируется перенести в облако

Managed Service for MySQL® · Managed Service for Kubernetes® · Compute Cloud · Object Storage

Банк Санкт-Петербург

Интернет-банкинг в облаке

Задача: повысить эффективность инвестиций и ускорить создание новых продуктов

Решение: объединение внутреннего контура с облачной платформой, которая соответствует требованиям российского законодательства и международных стандартов

Результаты: сократили срок вывода продуктов и научились быстро проверять бизнес-гипотезы. Расширили экспертизу сотрудников и сэкономили до 30% на инфраструктуре отдельных проектов

аудит пройден

благодаря соответствию PCI DSS

на 3

года растягиваются инвестиции в инфраструктуру

на 5

месяцев быстрее стартуют новые проекты

Cloud Interconnect · Managed Service for Kafka® · Managed Service for PostgreSQL · Managed Service for Kubernetes®

МГУ имени Ломоносова

Мониторинг промышленных садов

Задача: создать систему прогнозирования урожайности, учитывающую погодные условия и сортовые особенности плодовых культур

Решение: использовали сервисы облачной платформы для хранения и обработки большого количества данных с помощью нейросетей и ML-алгоритмов

Результаты: распределённая система мониторинга насаждений и урожая успешно протестирована на площадках в ботаническом саду МГУ и в экспериментальном саду ФНЦ имени Мичурина

9

облачных сервисов

используются для создания системы прогнозирования

первый опыт

создания систем прогнозирования урожая в России

DataLens Object Storage DataSphere Cloud Functions

skyeng

Тестовые среды в облаке

Задача: перенести тестовые среды Skyeng и снизить пинг

Решение: развернули в облаке серверы с необходимыми техническими характеристиками без ограничения на дисковые операции и с требуемым пингом

Результаты: снизили пинг в два раза, сняли ограничения на дисковые операции и уменьшили стоимость содержания в полтора раза. Появилась возможность изменять отдельные параметры сервера для специфических задач

200+

разработчиков используют тестовые серверы

1 год

потребовался, чтобы окупить затраты на проект

x10

сокращение пинга до Москвы

Compute Cloud

Genotek

Сервис генеалогических деревьев

Задача: разработать сервис, который позволит клиенту бесплатно создать свое генеалогическое древо и объединить его с деревьями-родственниками

Решение: использование возможностей бессерверных технологий и serverless-сервиса Yandex Cloud Functions, а также управляемых баз данных Yandex.Cloud

Результаты: новый сервис был разработан силами команды биоинформатиков. Бессерверные технологии позволили не задумываться о производительности, обслуживании и масштабировании баз данных

>100к
человек
добавлено в сервис генеалогических деревьев

152-ФЗ
соблюдение федерального закона

Managed Service for ClickHouse · Cloud Functions · Managed Service for ElasticSearch · Managed Service for MongoDB

МОЭК

Автоматический приём показаний счётчиков по телефону

Задача: автоматизация приёма показания счётчиков по телефону и обзвона клиентов с напоминанием о задолженностях

Решение: использование сервиса SpeechKit от Yandex.Cloud для распознавания и генерации речи

Результаты: разработана систему для автоматического распознавания и голосового приема показаний счётчиков и модуль автоматического информирования граждан о дебиторской задолженности за потребление ГВС и тепла

80%

сокращение нагрузки на операторов в период передачи показаний счетчиков

1

месяц

заняла реализация проекта

SpeechKit

Bauman Racing Team

Нейронные сети для беспилотного болида

Задача: выбрать для гоночного болида архитектуру нейросети, оптимальную с точки зрения скорости и качества распознавания объектов на трассе с учетом того, что вычислительные ресурсы на борту ограничены

Решение: с помощью сервисов Yandex DataSphere и Compute Cloud студенты обучили и протестировали широкий класс нейросетей YOLOv5, в котором архитектуры значительно отличаются друг от друга по числу параметров распознавания

Результаты: участники команды выбрали оптимальный вариант архитектуры, получив в сжатые сроки и без посторонней помощи точные и эффективные инструменты для решения распознавания объектов на трассе

20K
изображений
для обучения каждой нейросети

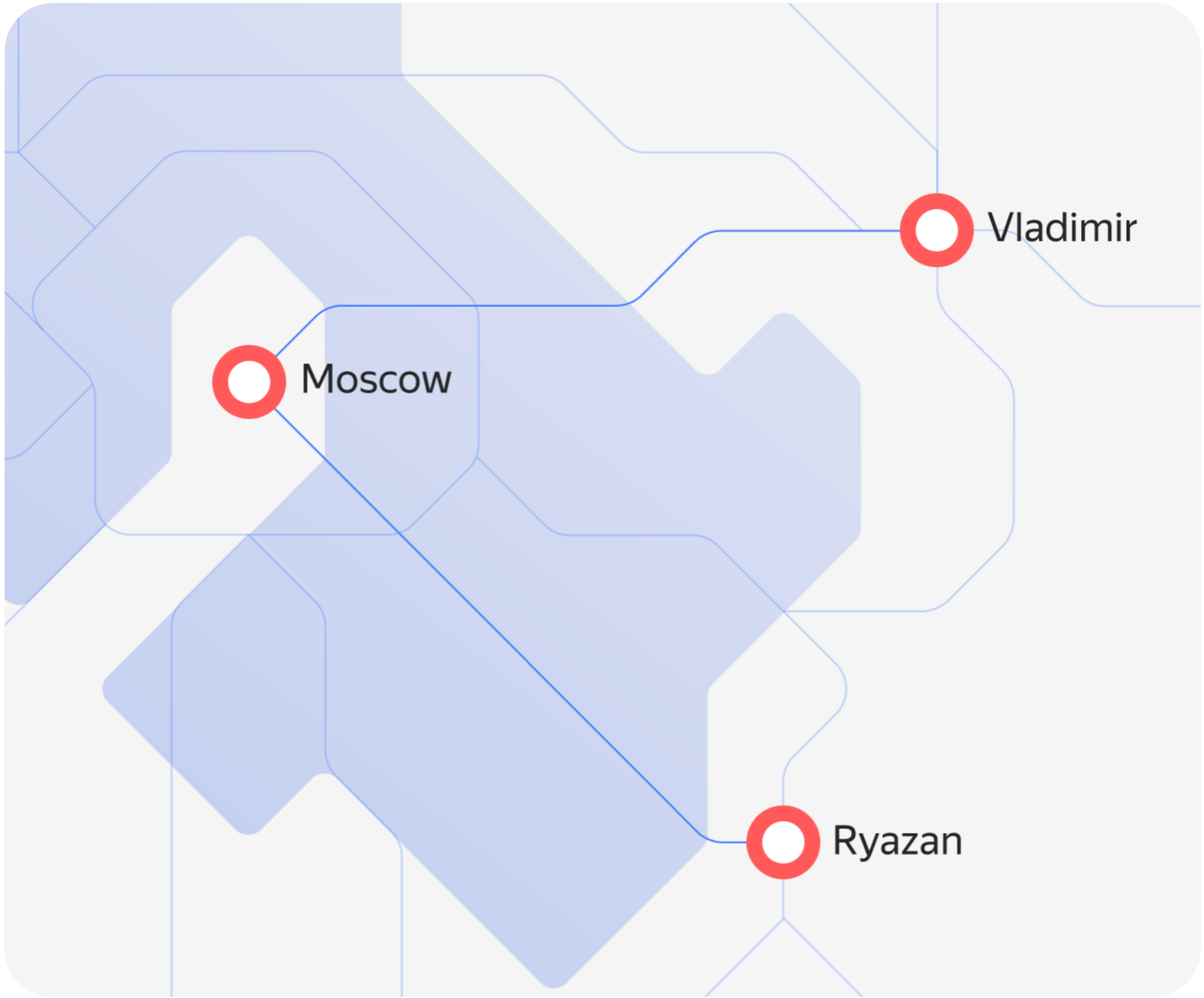
66
часов
обучали четыре нейросети

DataSphere Compute Cloud

1. Обзор платформы
2. Инфраструктура
3. Платформа данных
4. Serverless
5. Машинное обучение и искусственный интеллект
6. Безопасность
7. Опыт миграции
8. Кейсы
9. Yandex.Cloud: что стоит за облаком?

Собственная физическая инфраструктура

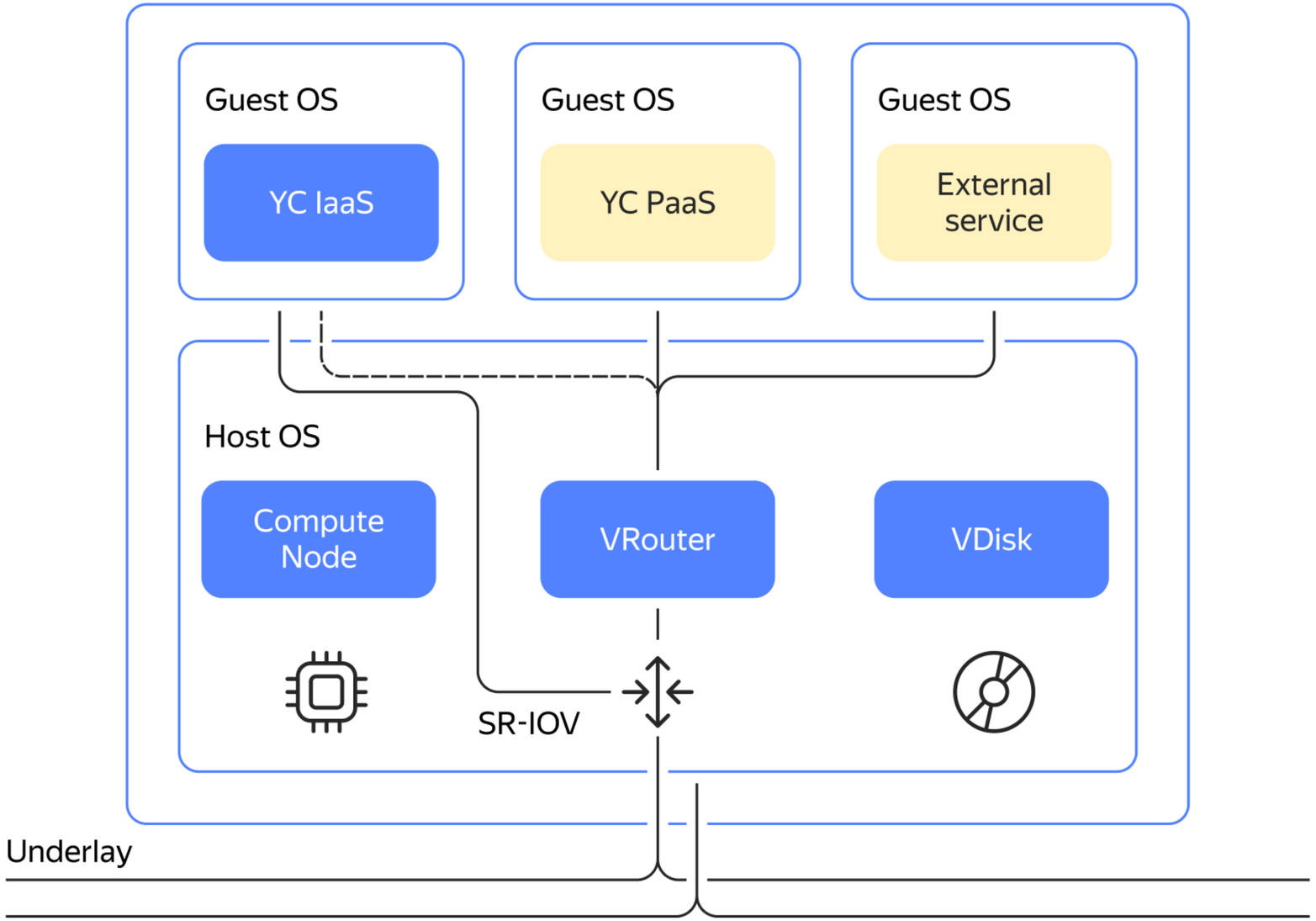
- У Яндекса пять дата-центров в России и один в Финляндии
- Географически распределённые дата-центры расположены на расстоянии более 300 км друг от друга
- Yandex.Cloud размещается в трёх зонах доступности
- Независимое энергоснабжение в каждом дата-центре
- Терабитная полоса пропускания обеспечивается собственной оптоволоконной DWDM-сетью



Серверное оборудование собственной разработки

- Серверные стойки разработаны под дата-центры, дата-центры — под серверные стойки
- Единообразие аппаратного обеспечения позволяет нам работать с разными вендорами
- Наше внутреннее аппаратное обеспечение работает в нужном интервале температур
- Нагрузка до 500 Вт на сервер
- Дисковые накопители поддерживают режим горячей замены

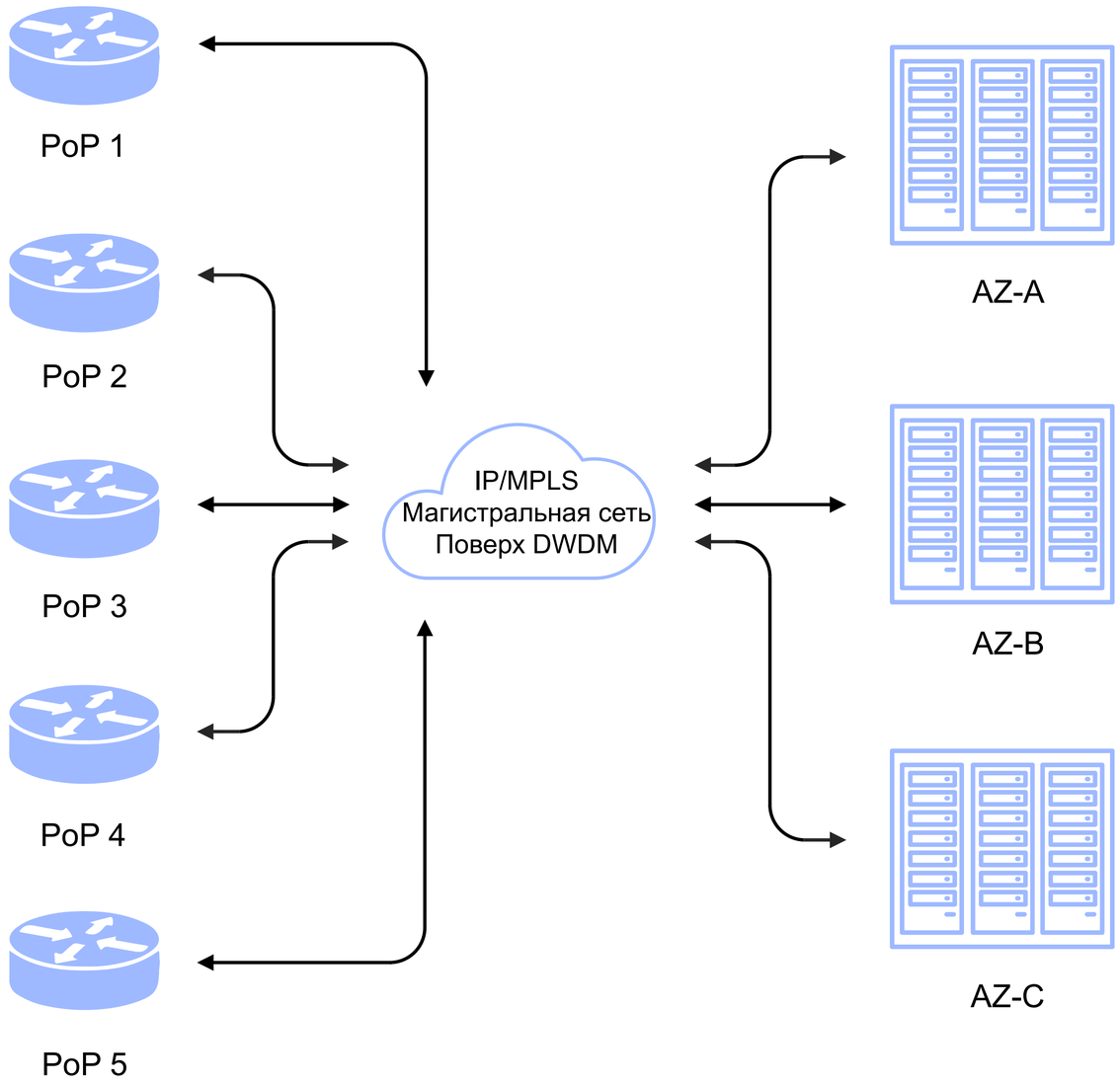
Конфигурация хоста Yandex.Cloud



Сетевая инфраструктура

Сеть Yandex.Cloud

- 3 AZ
- 5 PoP
- Тысячи серверов
- Исходящий канал до 800 Гбит/с
  от каждой стойки
- 5-каскадная сеть Клоза / 2-каскадная
  сложенная сеть Клоза
- 10+ Тбит/с внутри дата-центров
- ~1 Тбит/с от дата-центров к внешнему миру
- Международные проекты и установка
  распределённых облачных инфраструктур

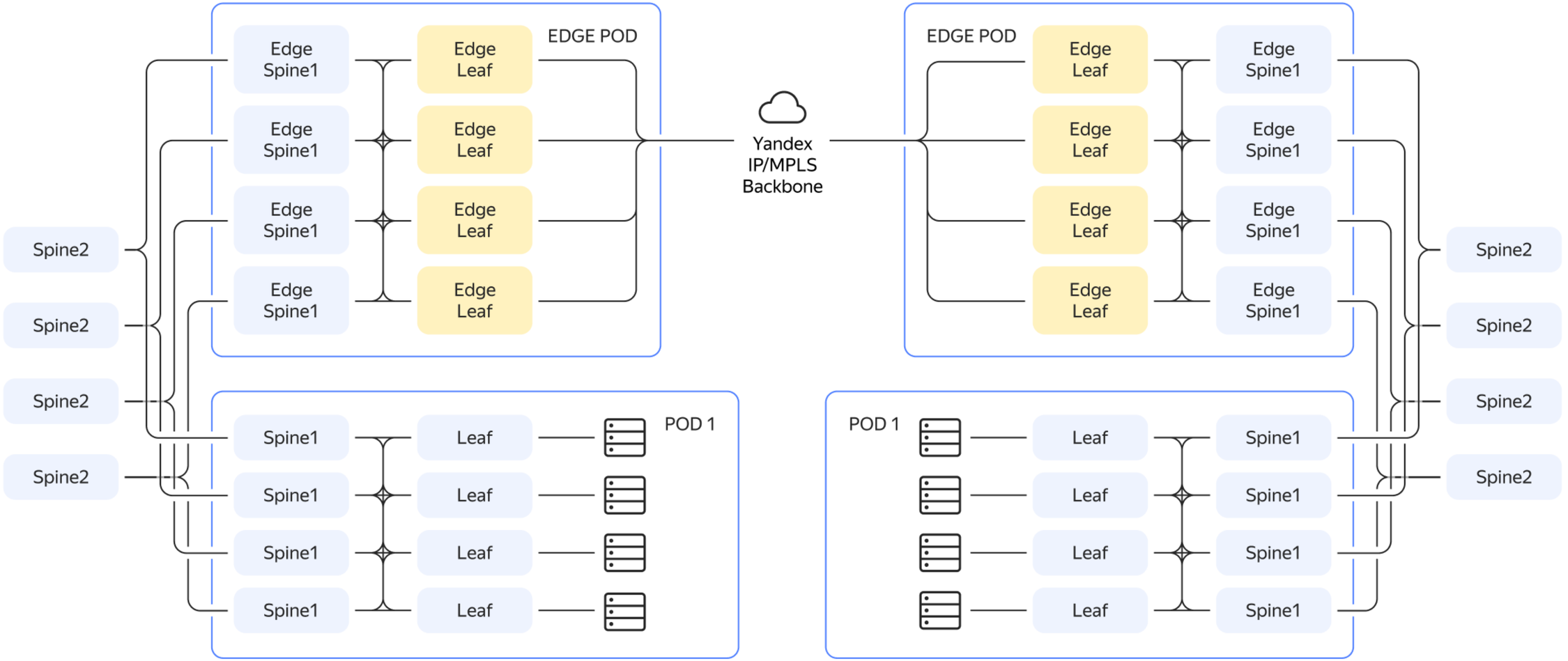


Сетевая инфраструктура

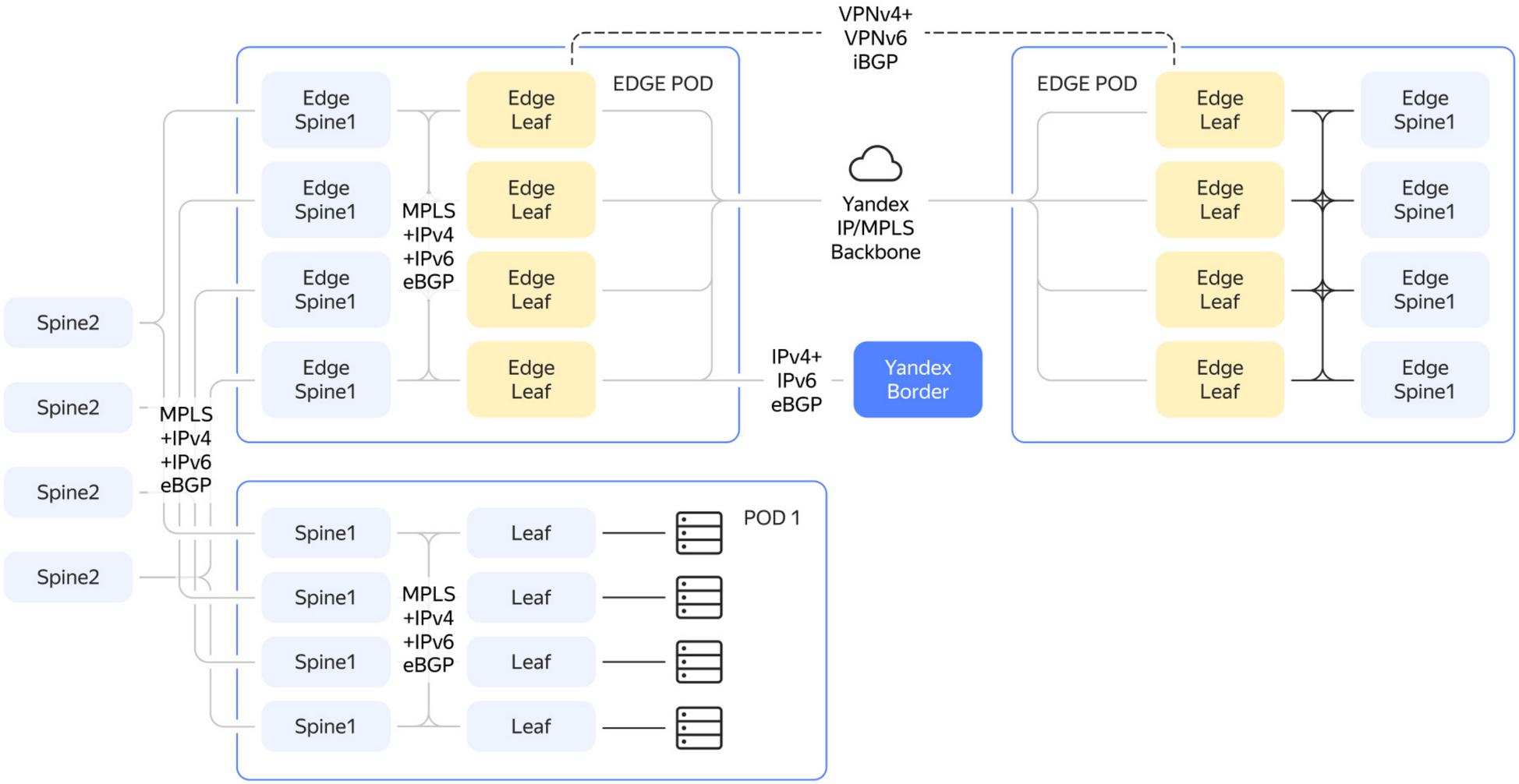
- В Yandex.Cloud используется классическая архитектура CLOS Fabric с 25/100G портами в направлении к серверам и 100/200G исходящими каналами от ToR к Spine
- Мы используем магистральную IPv6-сеть в качестве транспорта и выстраиваем физическую сеть (андерлей) IPv4-поверх-MPLS между хостами
- Наложенная сеть (оверлей) использует технологию Tungsten Fabric (Juniper Contrail) для выстраивания тоннелей MPLS-over-UDP между хостами
- Особый демон (под названием vrouter), запущенный на каждом хосте, отвечает за подключение хоста к другим хостам.
- Особые шлюзы (так называемые облачные шлюзы) маршрутизируют трафик между дата-центрами, а также между внешними сетями (например, интернетом или Cloud Interconnect) посредством MPLS-over-GRE-тоннелей к демону vrouter

Физическая сеть

Сетевая инфраструктура

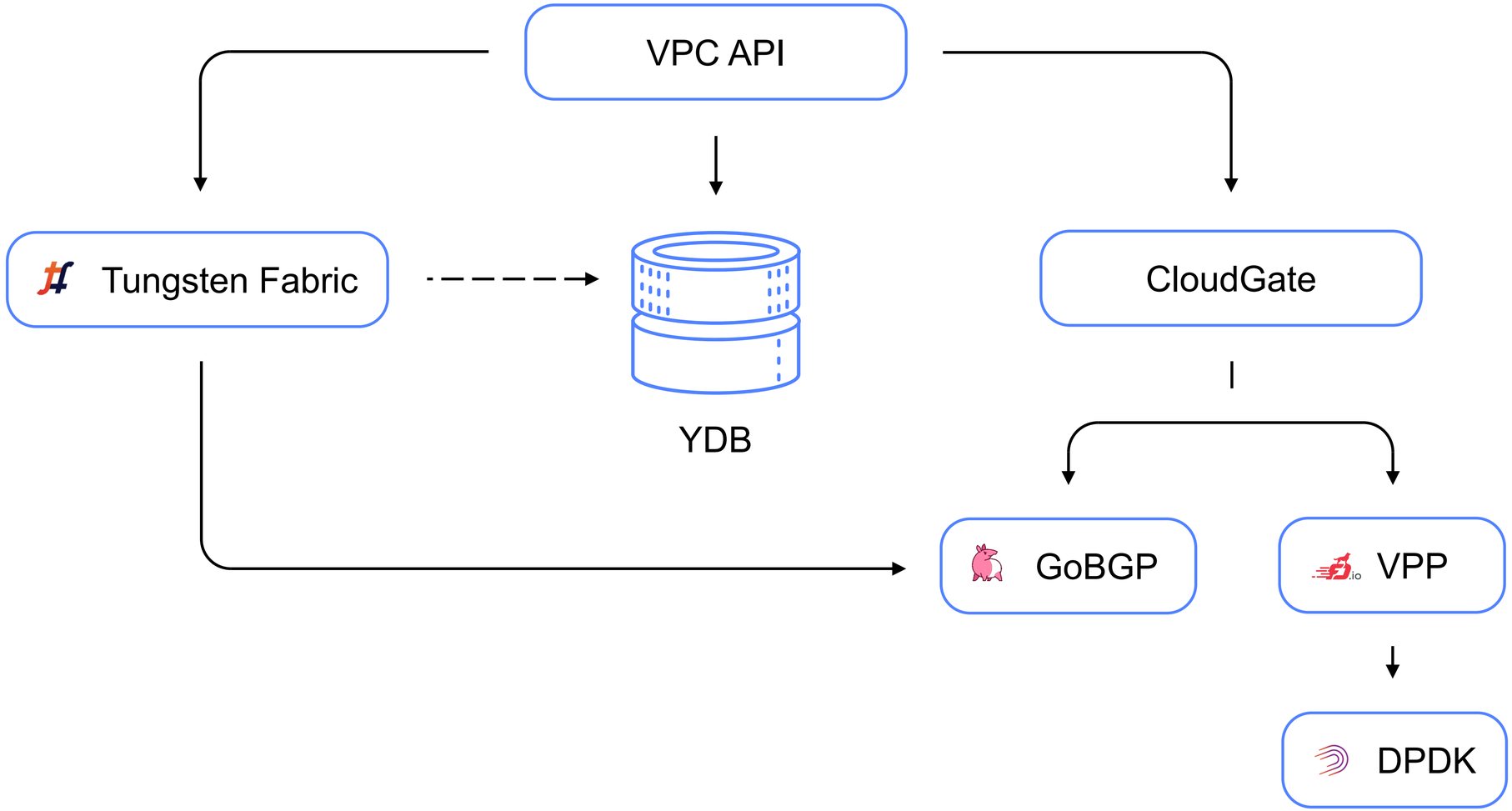


Протоколы сетевой инфраструктуры



VPC-архитектура

Высокоуровневая VPC-архитектура



VPC-компоненты

Описание компонентов

VPC-API — интерфейс для выполнения действий над VPC-ресурсами

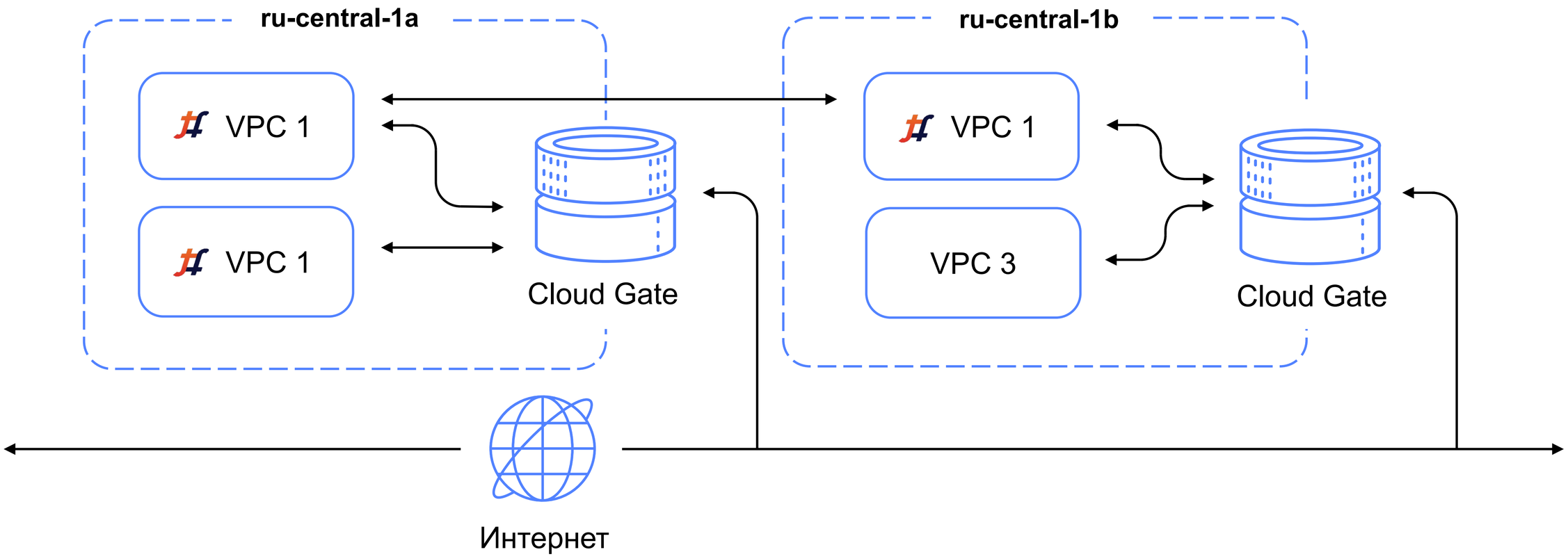
Tungsten Fabric (Juniper Contrail) — SDN-сеть, которая обеспечивает оверлейное сетевое соединение между хостами и уровень данных для передачи пакетов (vrouter). Она использует BGP для обмена информацией о доступности с другими компонентами (например, узлами Load Balancer или облачными шлюзами Cloud Gateway)

Cloud Gateway — виртуальная машина, которая предоставляет сетевые сервисы, в том числе NAT, Load Balancing, Cloud Interconnect (DCI). При этом используется инфраструктура передачи пакетов Cisco VPP и уровень управления goBGP. Обменивается BGP-маршрутами с Tungsten Fabric, предоставляя виртуальным машинам, запущенным на хостах (т. е. за vrouter), доступ к сетевым сервисам.

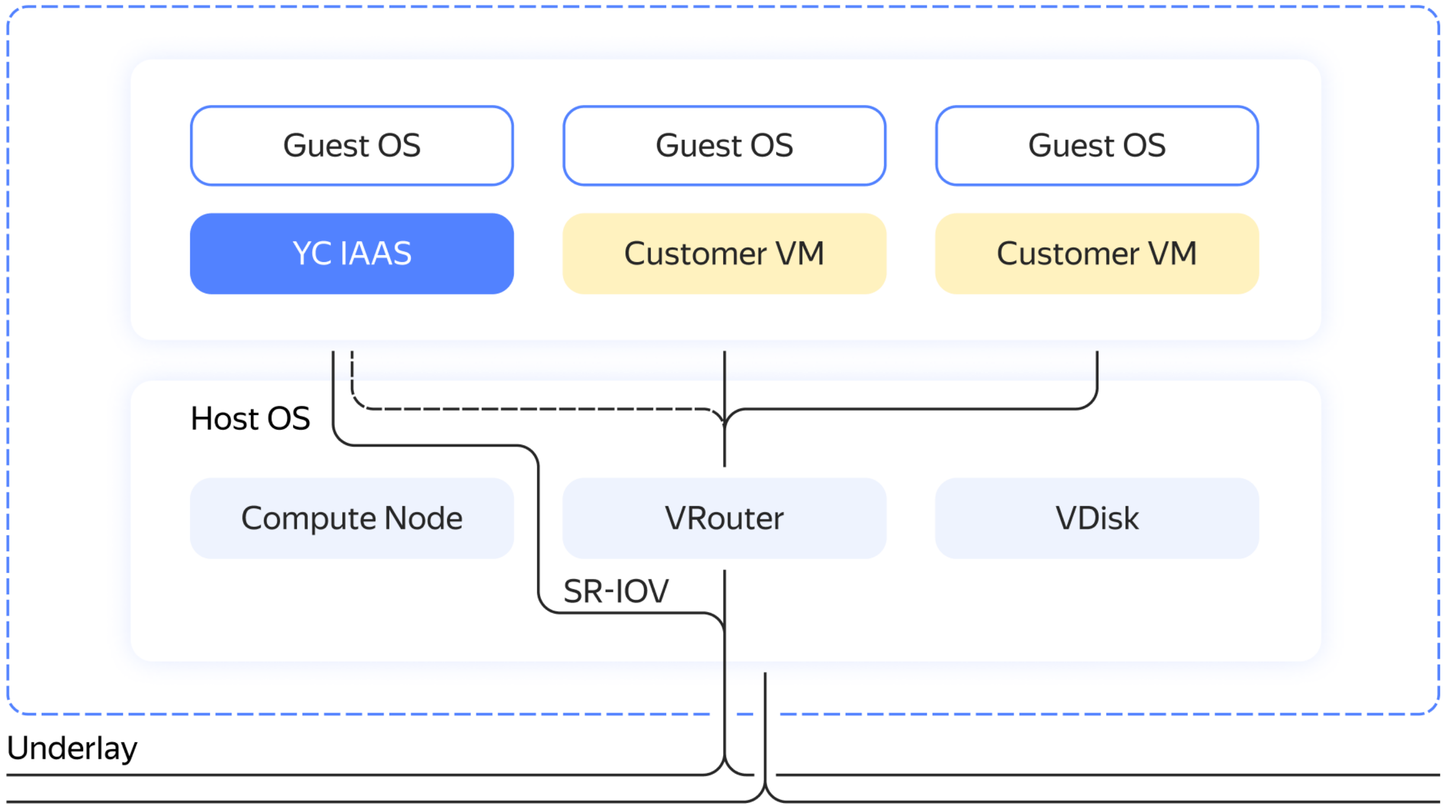
Каждая зона доступности (AZ) имеет несколько облачных шлюзов, используемых несколькими клиентами. Например, может быть 20 облачных шлюзов Cloud Gateway, осуществляющих трансляции NAT-адресов для всех виртуальных машин конкретной AZ. Их число масштабируется в соответствии с потребностями трафика.

VPC-компоненты: Cloud Gate

Cloud Gate предоставляет интернет-доступ (NAT и маршрутизацию)
для ресурсов, подключённых к подсетям внутри VPC-облаков



VPC-компоненты: Vrouter

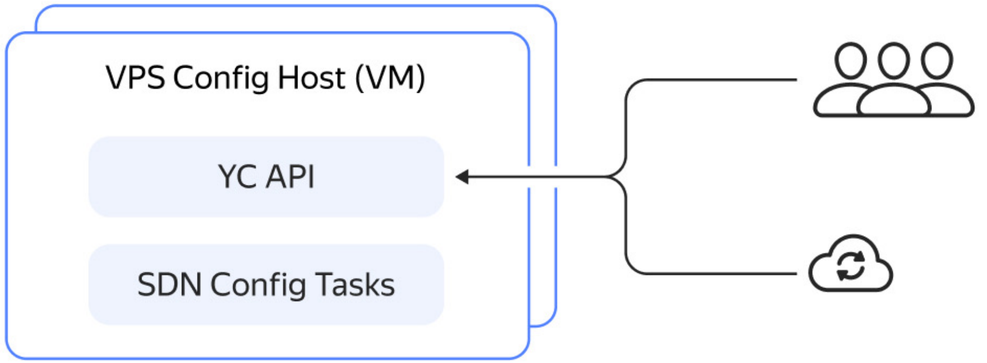


- Процесс Tungsten Fabric Vrouter обеспечивает маршрутизацию и передачу пакетов для рабочих нагрузок на хосте
- Для управляемых машин Yandex.Cloud (например, Cloudgate) может обеспечиваться доступ SR-IOV для NIC

Поток операций VPC

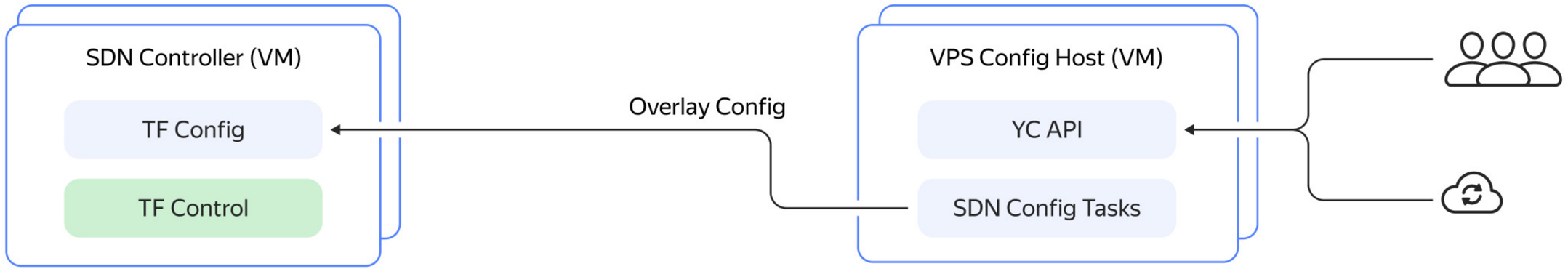
VPC состоит из следующих элементов:

- Уровень управления, отвечающий за желаемое состояние системы. Настраивается через VPC API (консоль, Terraform и т. д.)
- Уровень управления Control Plane настраивает уровень данных для достижения необходимого состояния
- Уровень данных Data Plane, отвечающий за передачу пакетов



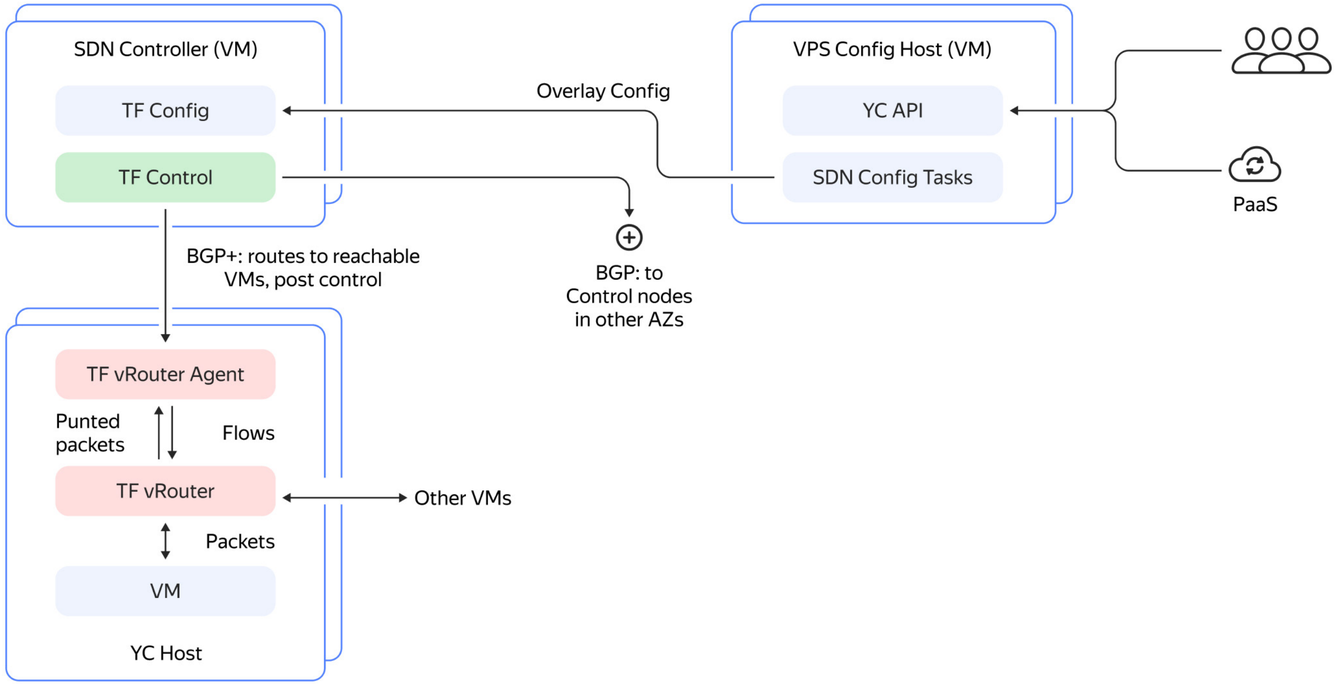
Поток операций VPC

После получения API-вызова запускается задача.
Она резервирует необходимые ресурсы в SDN,
например интерфейсы, адреса, подсети, сети



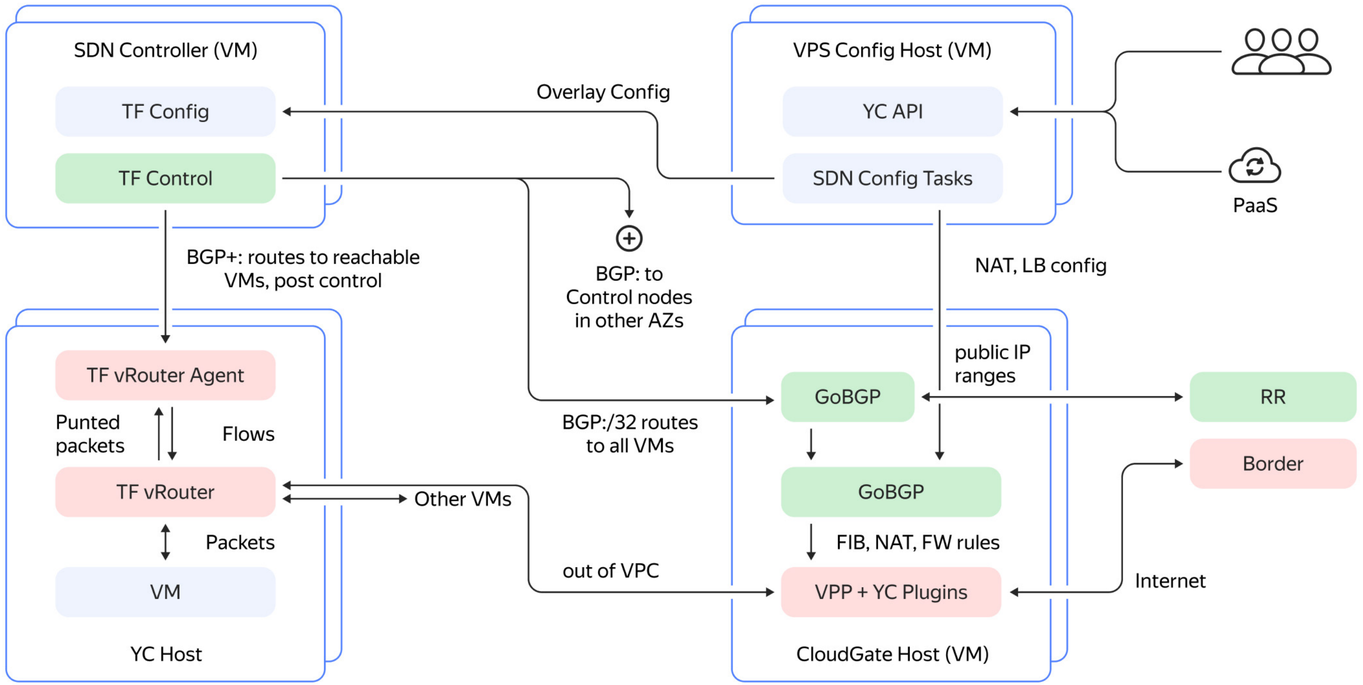
Поток операций VPC

- Tungsten Fabric использует BGP + XMPP для предоставления на уровне хоста сервиса Vrouter Agent с необходимой информацией, например о доступности виртуальной машины и о её текущей конфигурации. В целом отвечает за программирование уровня данных на хостах
- Также обменивается BGP-данными с другими узлами SDN, расположенными в других AZ

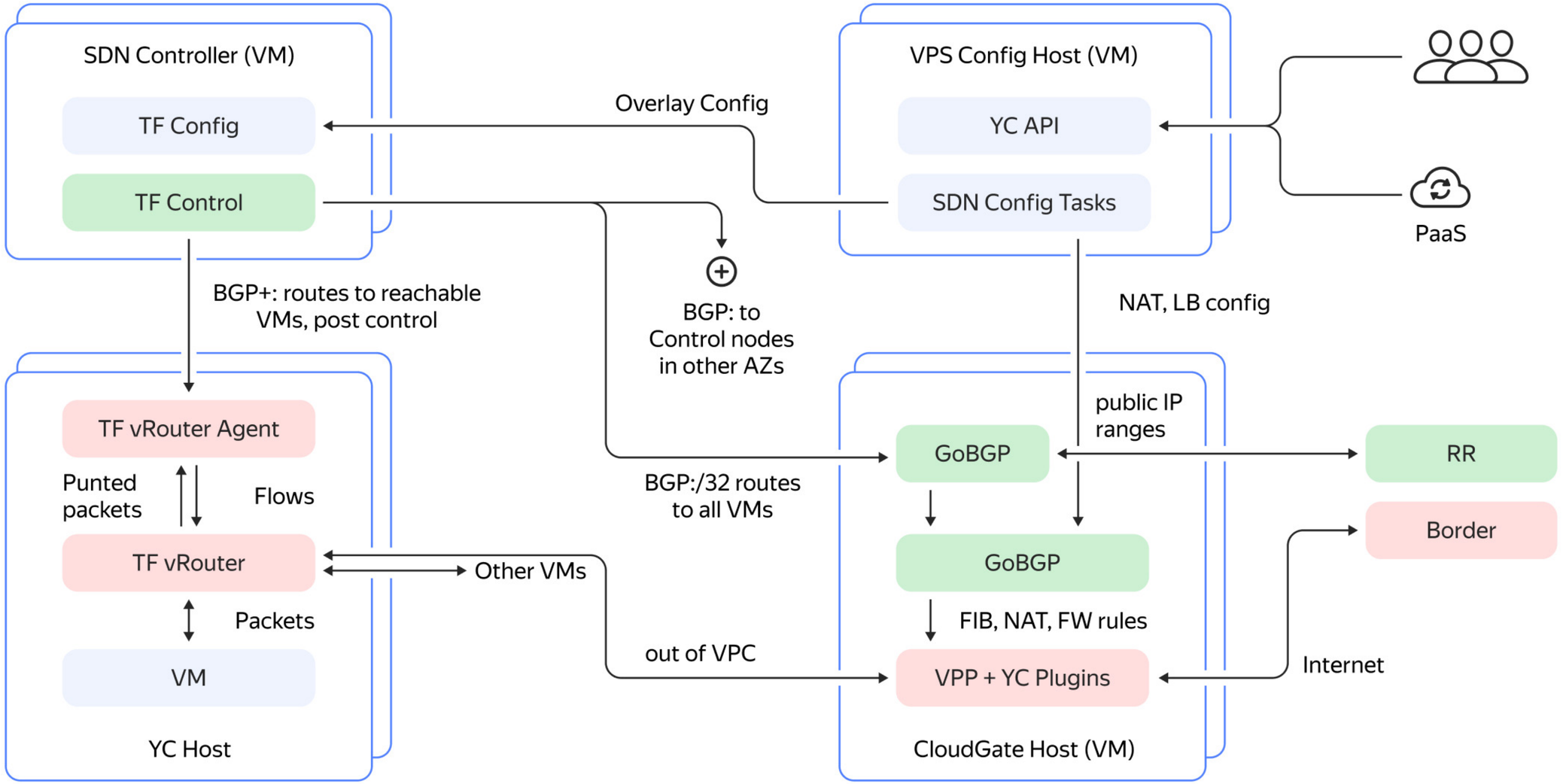


Поток операций VPC

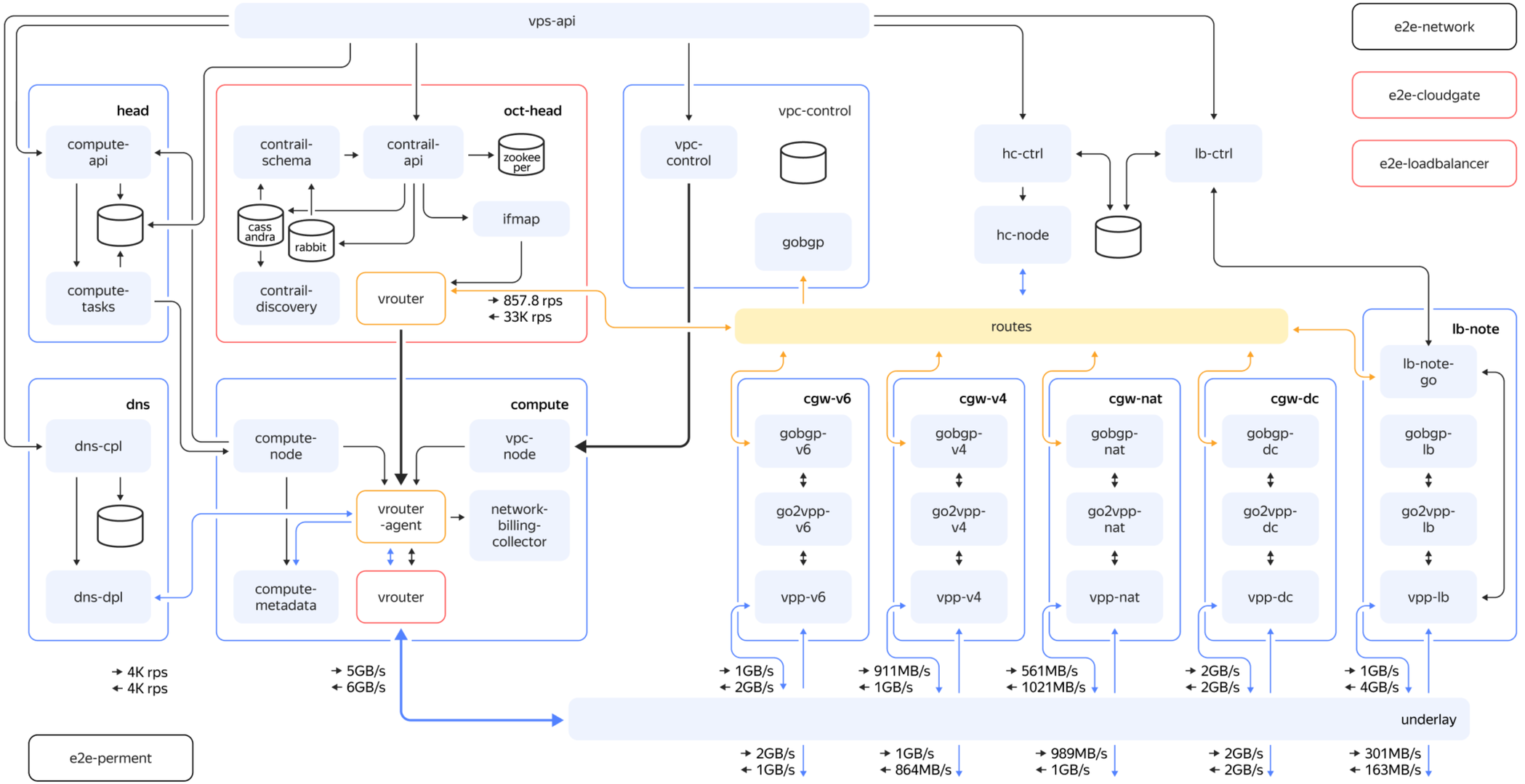
- Tungsten Fabric также взаимодействует с Cloud Gates (балансировщики нагрузки, NAT-шлюзы) для предоставления сервисов шлюзов CGW виртуальным машинам в VPC-облаках
- CloudGate использует VPP для более быстрой обработки пакетов
- Сервисы, предоставляемые шлюзом CGW, настраиваются напрямую через VPC-API



Поток операций VPC

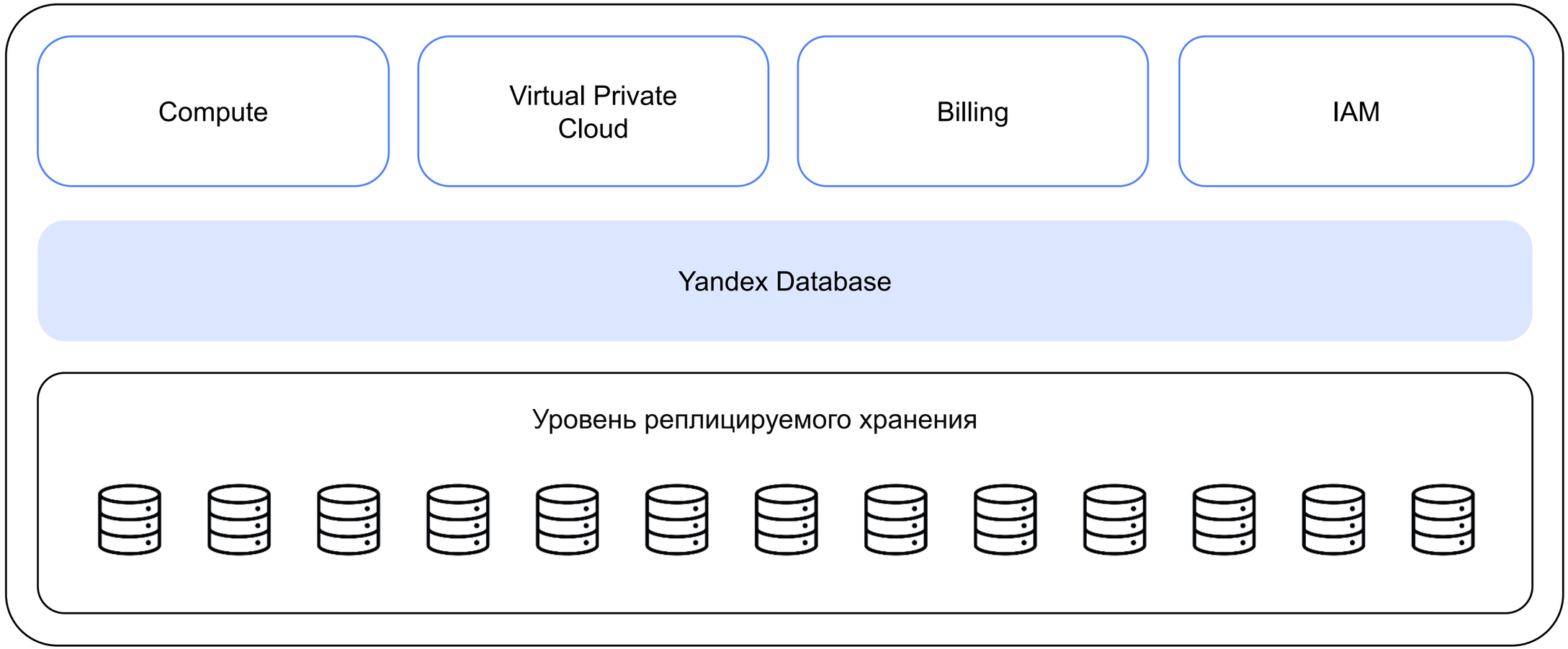


Детальный обзор VPC

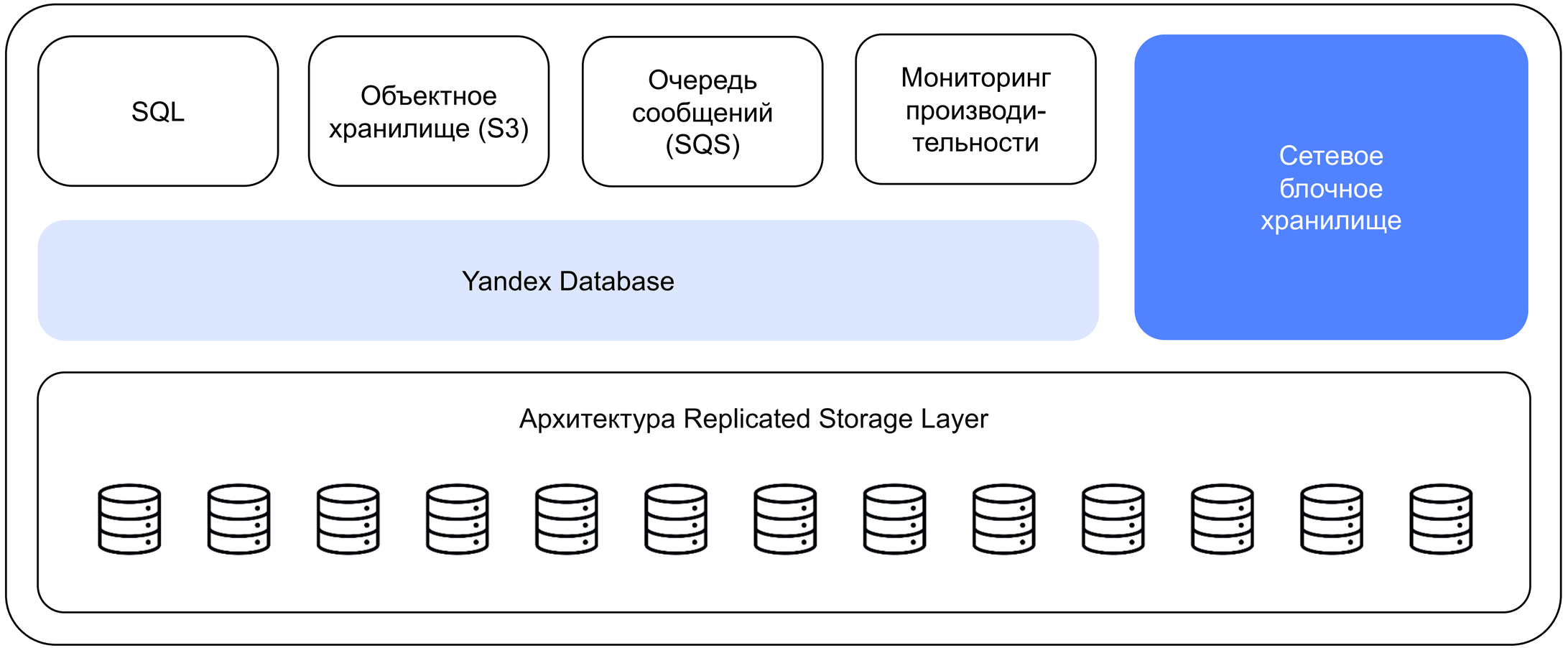


Архитектура хранения данных

Унифицированное хранилище метаданных



Программно-определяемое хранилище



Yandex Database — что это такое?

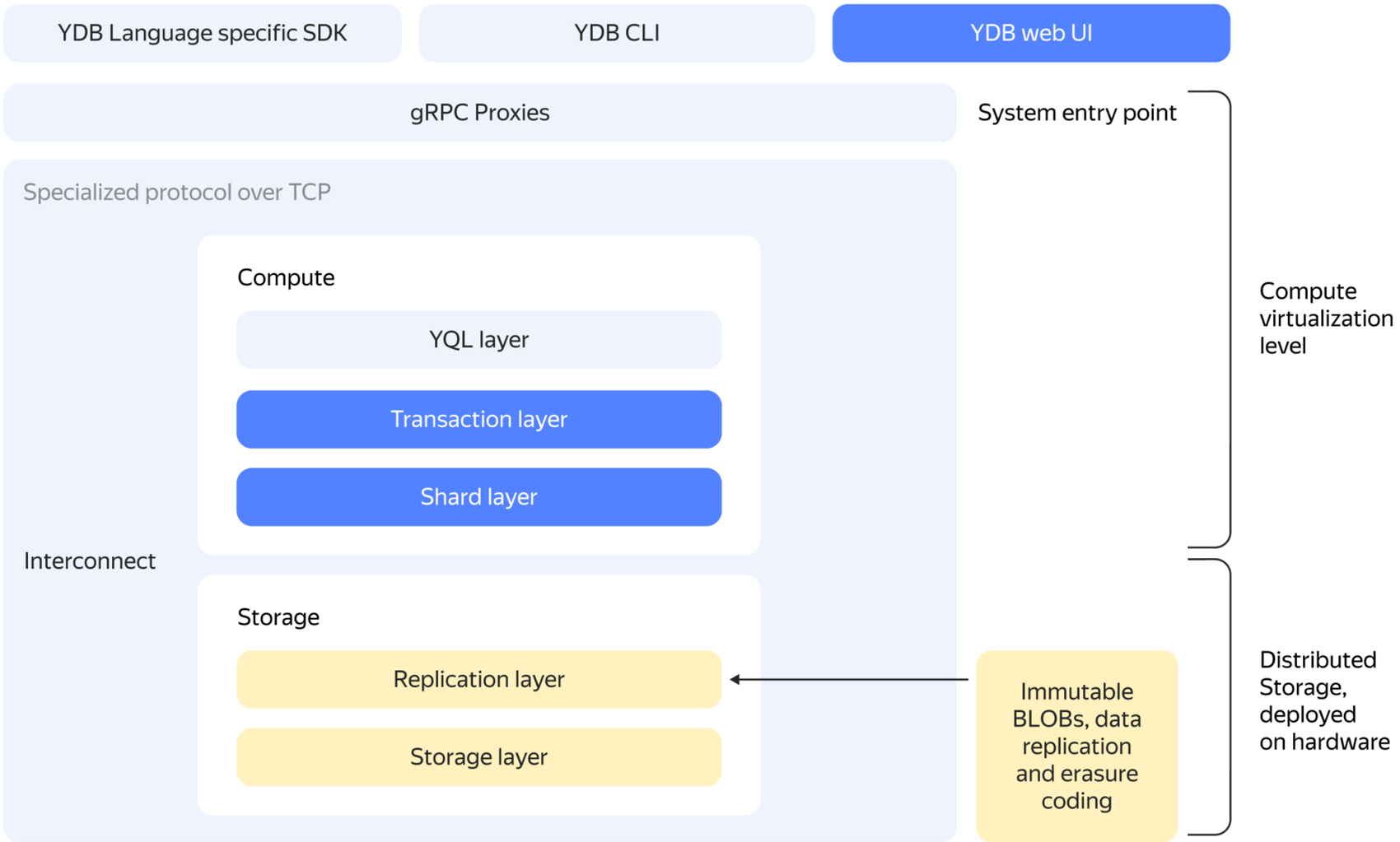
Платформа для специализированной СУБД-разработки

- Запись подтверждения, после того как все необходимые копии были физически зафиксированы
- Уникальный распределённый алгоритм консенсуса

Распределённая отказоустойчивая СУБД типа NewSQL

- Строгая консистентность
- Поддержка схем для таблиц
- Распределённые ACID-транзакции, охватывающие несколько таблиц
- YQL (SQL-диалект от Яндекса)
- Ориентированность на OLTP-нагрузку

Многоуровневая архитектура Yandex Database



Спасибо!

Григорий Атрепьев

Главный архитектор облачных решений

gatrepyev@yandex-team.ru

Полезные ссылки

Документация

Цены

Сценарии использования

YouTube-канал

Бесплатный курс на Яндекс.Практикуме