

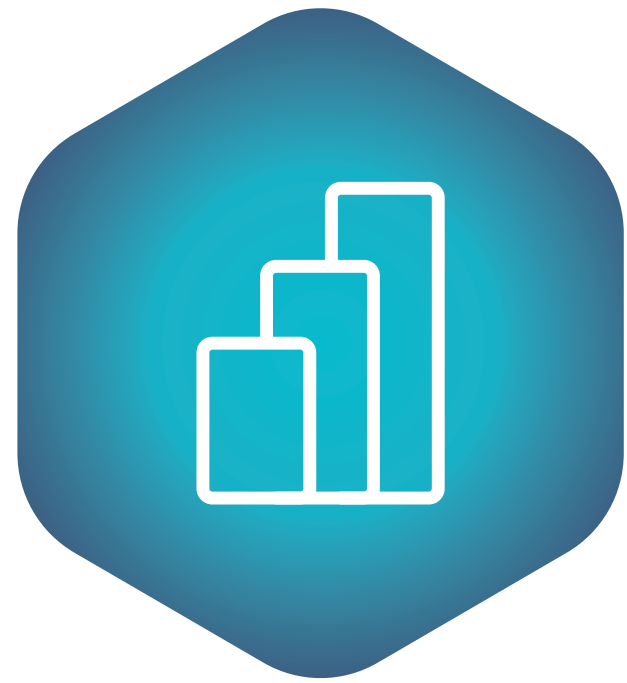


Эластичная платформа больших
данных Arenadata One

Дмитрий Гайворонский
Arenadata One BDM

Обзор

Ключевые вызовы в Big Data



Рост объемов данных

Кратный рост объемов данных год к году в связи с внедрением IoT и AI технологий



Разрыв между бизнесом и данным

Дефицит специалистов для внедрения и использования новейших Big Data-технологий



Баланс «цена-ценность»

Необходимо определить, какие данные представляют наибольшую ценность для бизнеса



Корпоративные требования

Необходимо обеспечить безопасную работу с данными и надежность и отказоустойчивость системы

291 ЗБ

данных будет произведено в мире в 2027 г.

На 10%

ежегодно растет стоимость специалистов

Только 35%

вакансий закрываются

30-40%

признают, что не понимают, как извлекать ценность из своих данных

50-60%

рост атак на Big Data-инфраструктуру ожидают аналитики Gartner в 2025 г.

Arenadata One – ответ рынку Big Data



TCO

Динамическое управление стоимостью хранения и вычислений

T2M

Низкий порог входа для пользователей и администраторов

Корпоративные требования

Сквозные функции безопасности, надежности и отказоустойчивость enterprise-уровня

Ценность и монетизация

Инструменты для извлечения выгоды для бизнеса из данных

Ключевые особенности в Arenadata One

Разделение Compute/Storage

Эффективное использование ресурсов и масштабирование по требованию в эластичной контейнерной среде

Infrastructure Agnostic

On-premise, частное или публичное облако, гибрид. В виде ПО, облачного сервиса или ПАК

Data Governance by Design

Интегрированное средство управления моделью, качеством и связанностью данных

Объектное S3-хранилище

Единое унифицированное внешнее хранилище данных с возможностью аналитики в том числе и над бэкапами

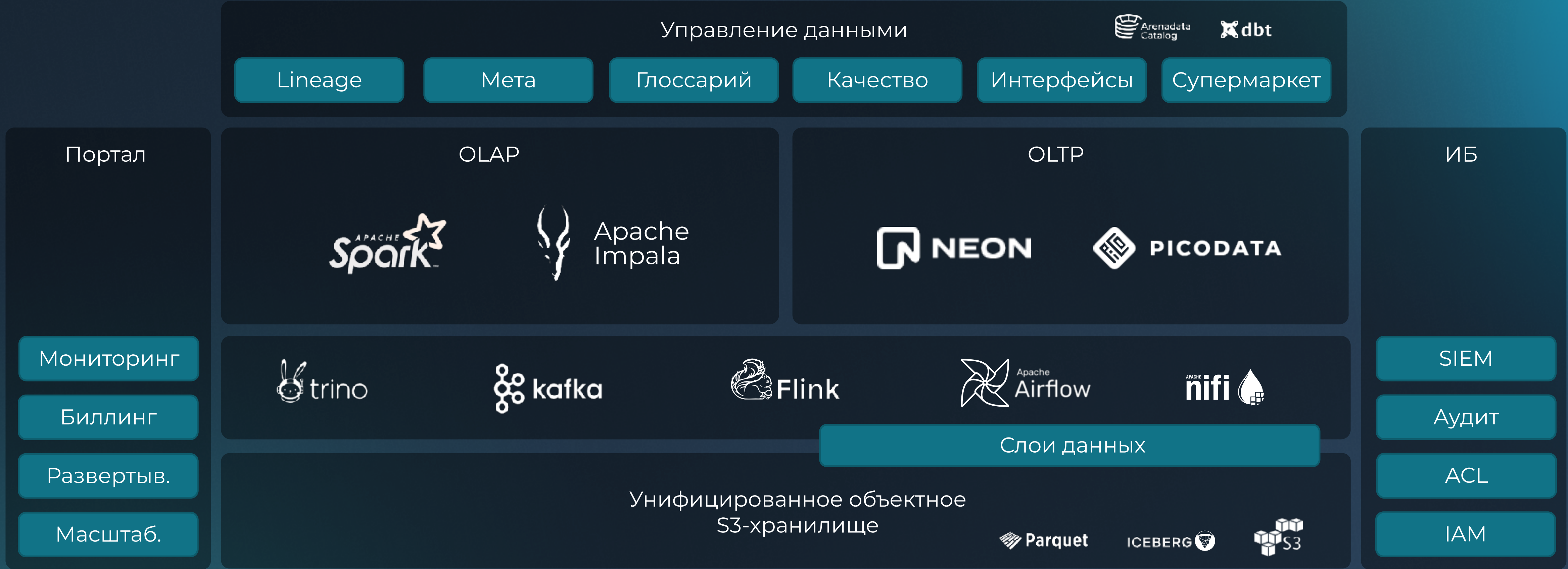
Универсальность

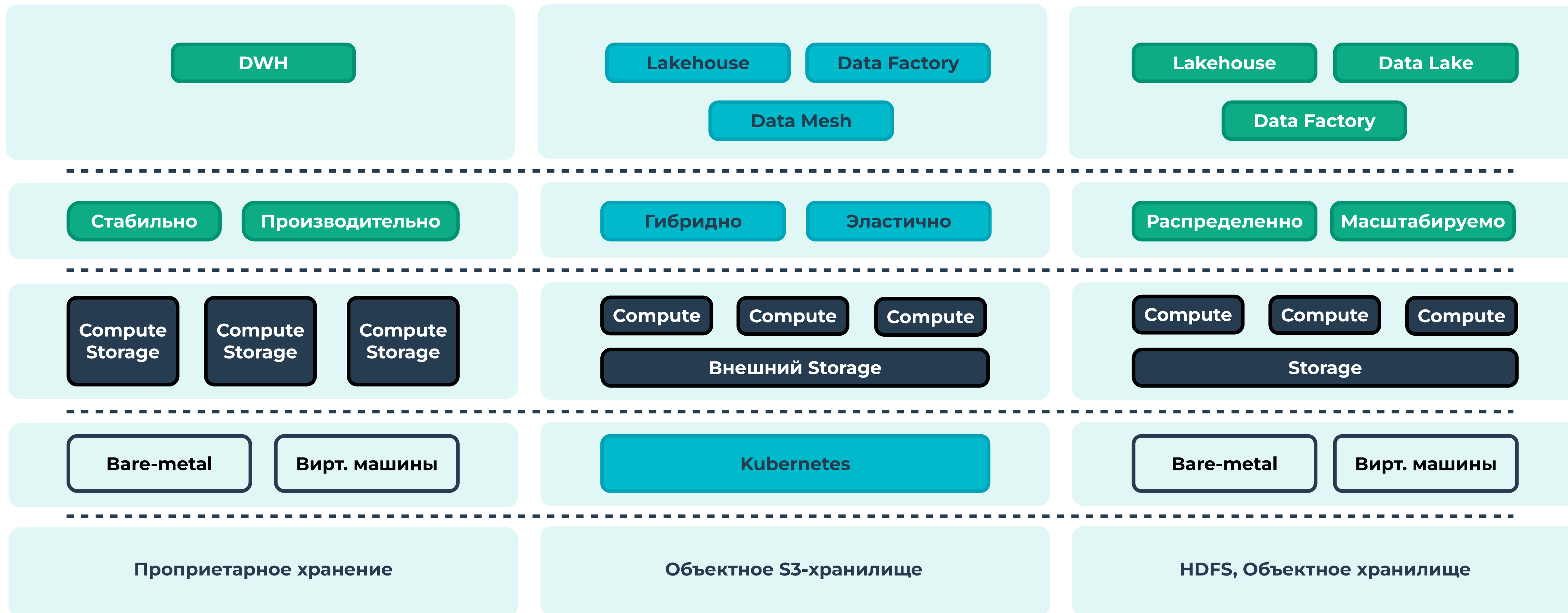
OLTP, OLAP и AI-нагрузки, системы любого класса, включая Lake House, Data Mesh и real-time аналитика

Multi-tier архитектура

Механизм хранения и обработки данных на основе приоритетов и требований бизнеса

Состав платформы

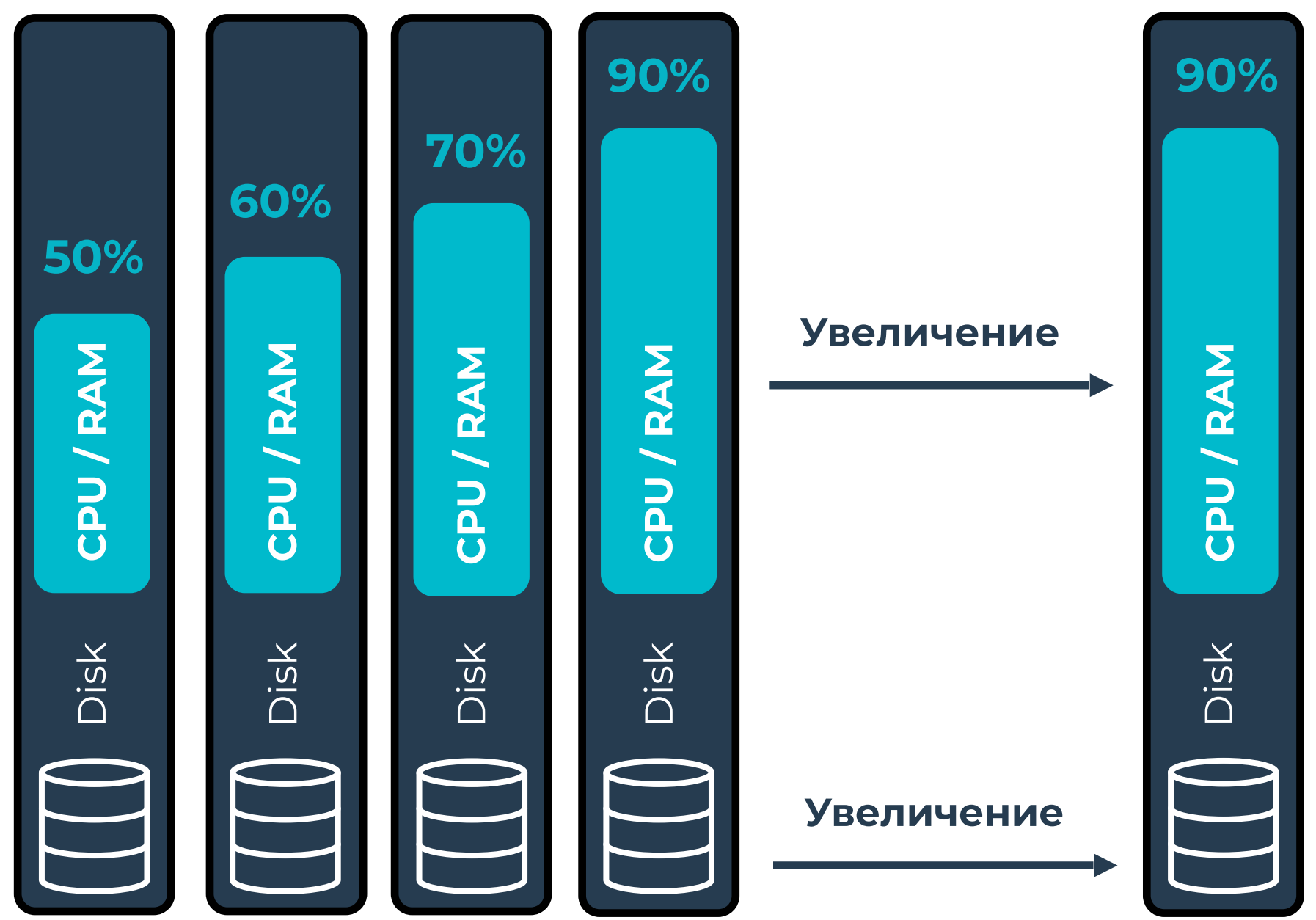




TCO

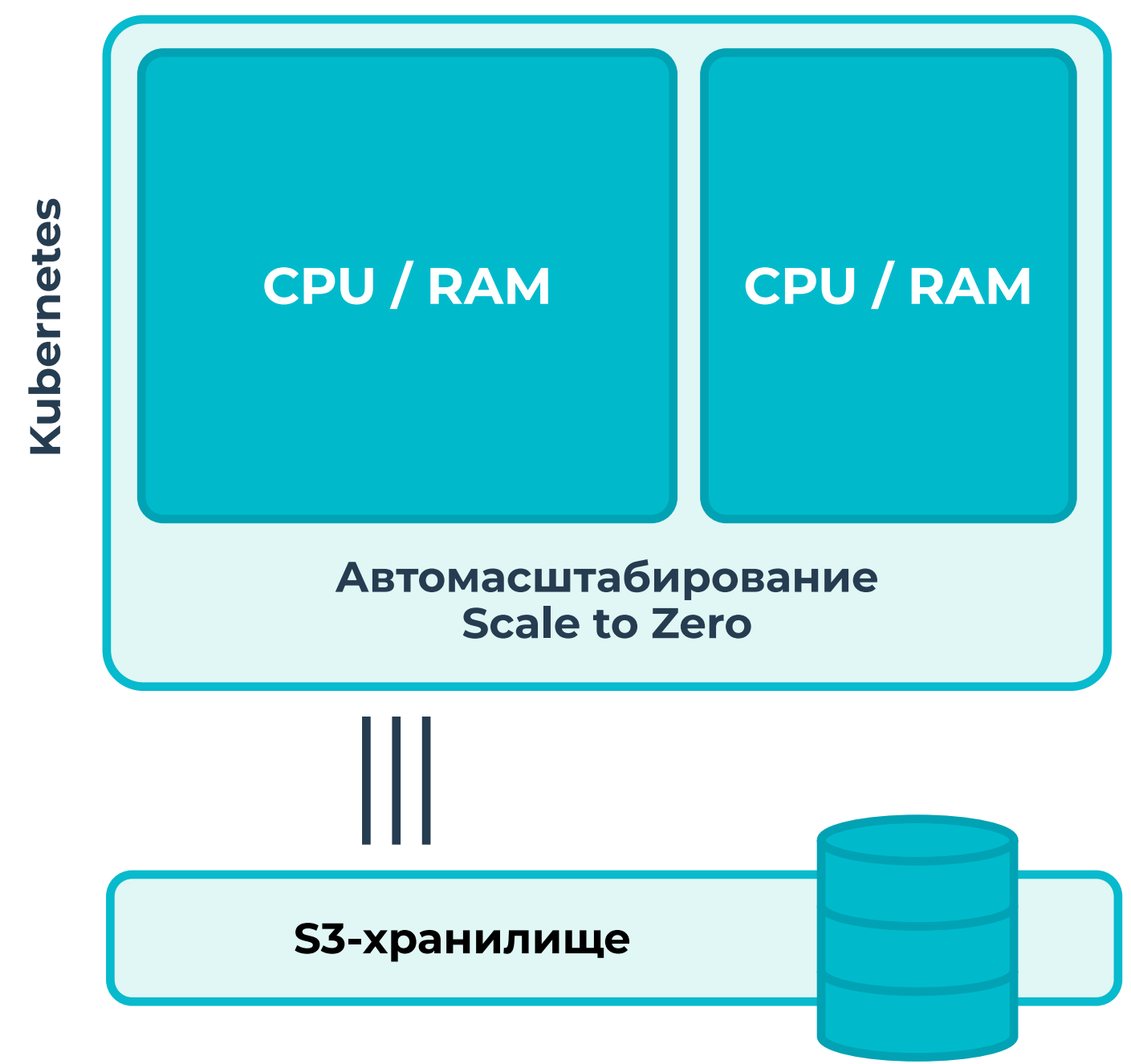
Оптимальное использование ресурсов за счет раздельного масштабирования хранения и вычислений

MPP



Только увеличение серверов
Compute + Storage

Lakehouse

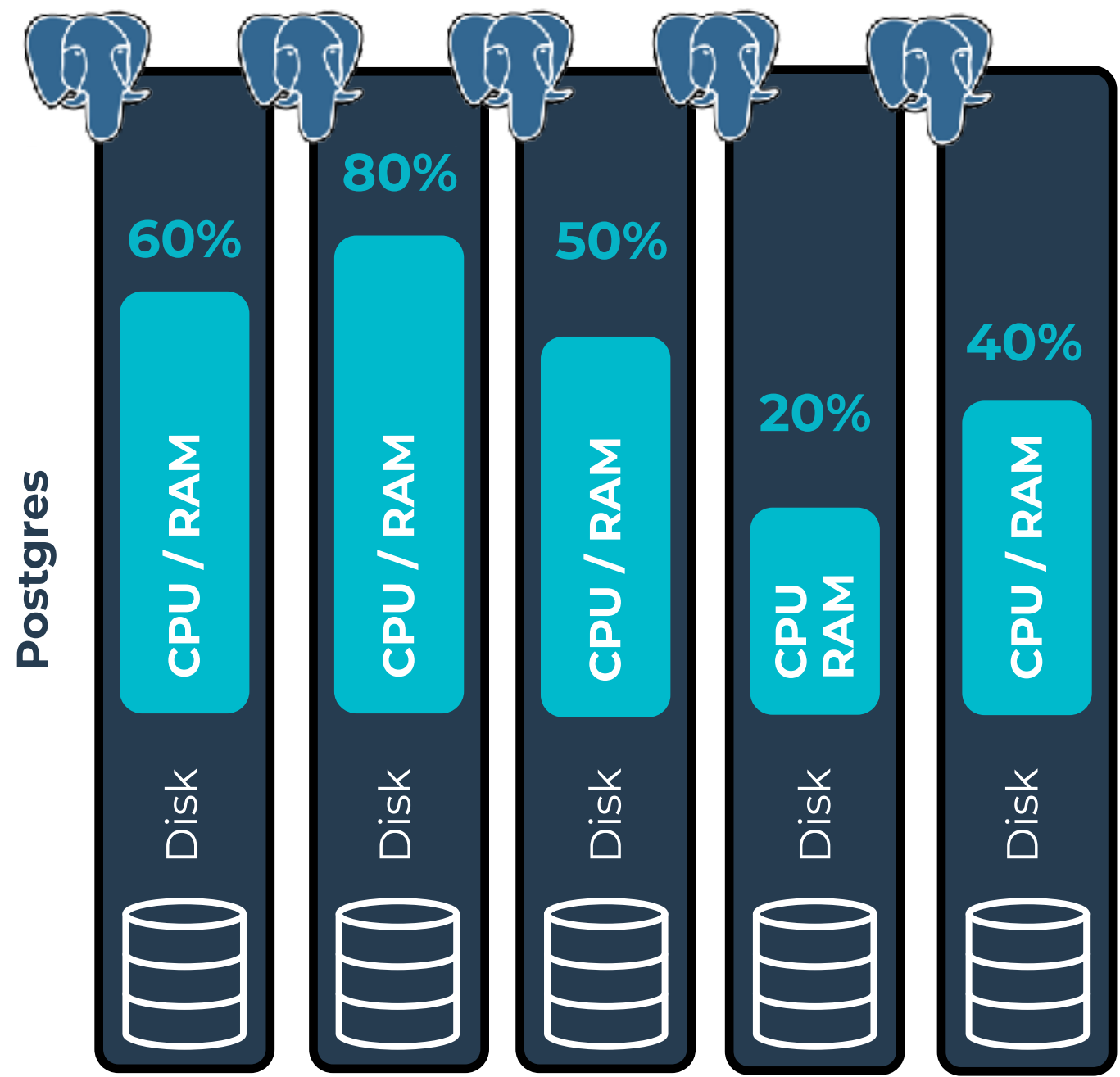


Storage – статичный
Compute – динамический

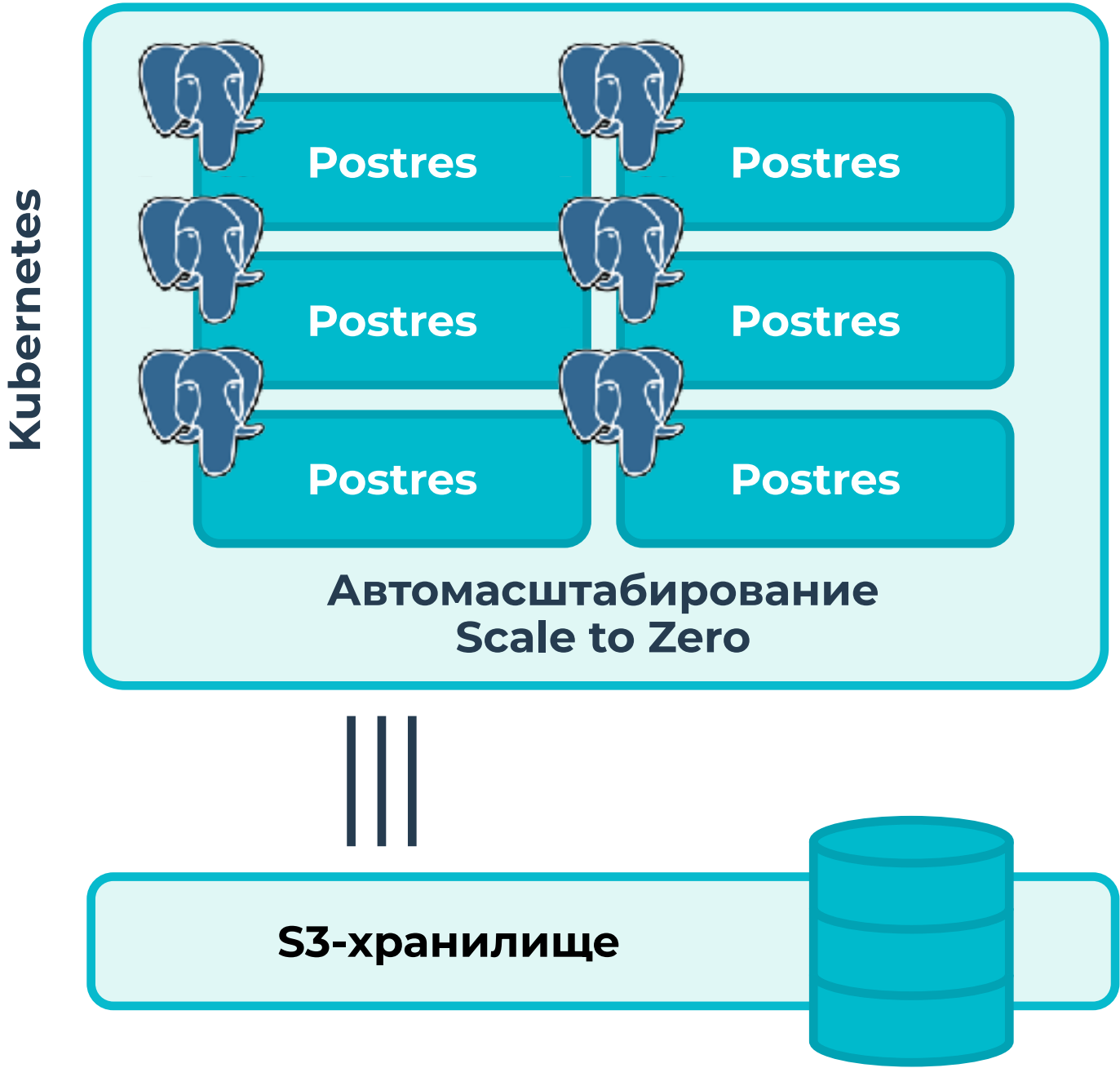
Экономия до 25-50% ресурсов на Postgres за счет единого пула ресурсов и масштабирования



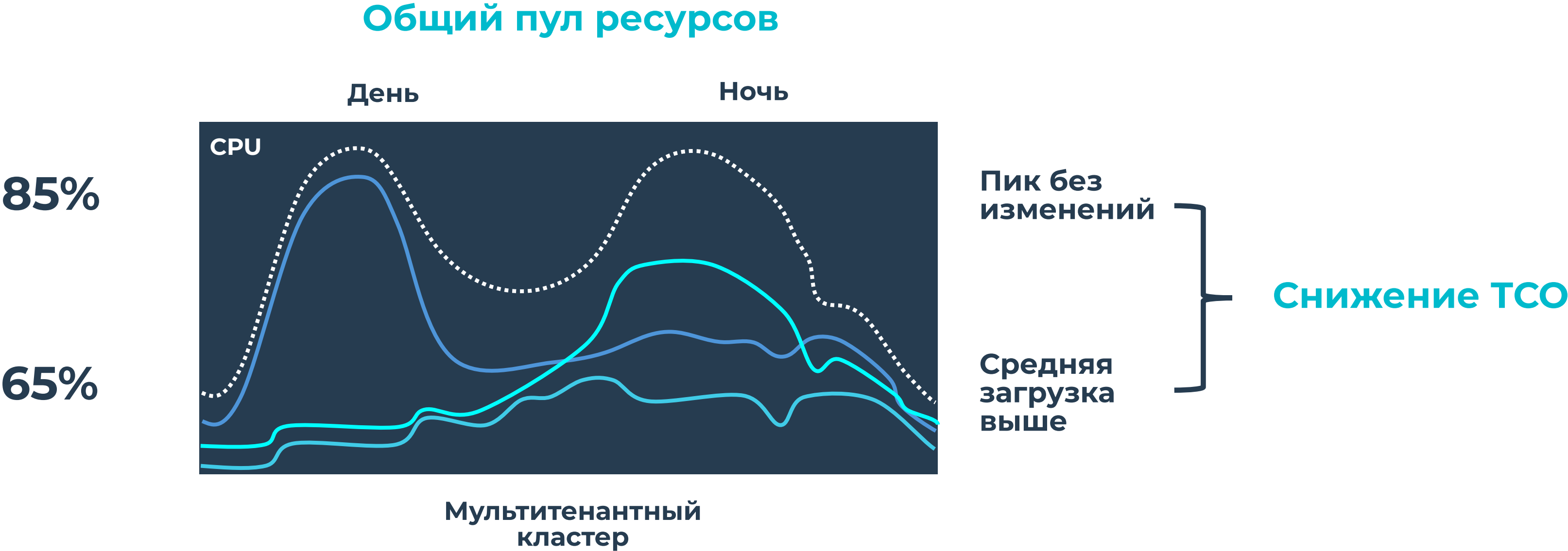
Множественные кластеры PostgreSQL



Единый кластер Neon
UI CLI API управление



Мультитенантность



Биллинг

Пример биллинга в облаке

Детализация Расходы на сервисы Ограничения Бонусы

01.06.2025 — 30.06.2025

Детализация за день

При отображении детализации все пополнения и списания рассчитываются на основе часового пояса: GMT+3 (Москва, Россия)

Статистика

Виртуальные серверы

Диски виртуальных серверов

Виртуальные серверы Kubernetes

Диски виртуальных серверов Kubernetes

Виртуальные машины Баз Данных

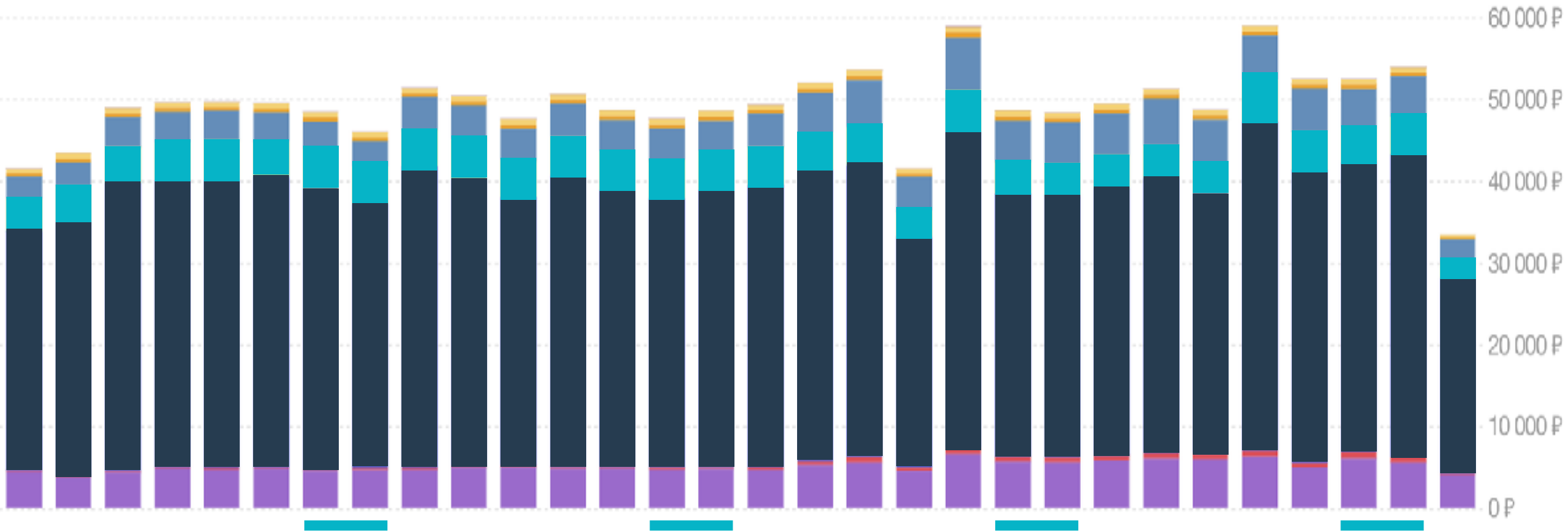
Диски виртуальных машин Баз Данных

Объектное хранилище Hotbox

Объектное хранилище Icebox

Объектное хранилище резервного копирования

Снимки дисков



Измерение

Управляешь только тем, что можешь измерить

Ответственность

Предупреждение о потенциальных затратах и согласие перед запуском нагрузки

Контроль

Дисциплина и контроль потребления ресурсов, все лишнее отключается

Бюджет

Защита бюджетов подразделений на ресурсы на основе биллинга

Дата-контракты

Основа для сопоставления с ценностью от дата-продуктов

T2M

Спрос на сокращение времени вывода на рынок продуктов

DBaaS

Спрос на ускорение доступа к инфраструктуре, (DBaaS подход) как пользователям, так и ИИ-агентам

No-code/Low-code

Спрос на no-code и low-code решения для проектирования и подготовки пайп-лайнов

Аналитик vs Инженер

Тренд упрощение построения конвейеров обработки данных силами аналитика, а не разработчика

Real-time

Спрос на инструменты работы с поведением пользователя в реальном времени для гиперперсонализации

Единая контрольная панель

Консоль провайдера

Для развертывания и мониторинга ADB
через ADCM и интерфейс консоли
облачного провайдера

Кластер менеджер ADOne

Развертывание и управление движками
платформы Arenadata One из одного окна

Единая платформа данных

Возможность развертывания, мониторинга
и управления ресурсами всех продуктов
Арены через единую контрольную панель



Your current organization
auth.arenadata.io ^

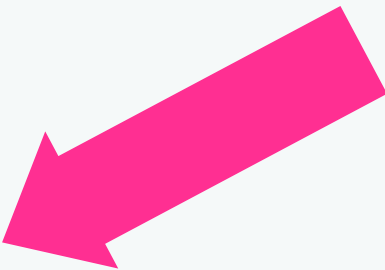
Common info
Users
Groups
Permissions

Projects ^
Projects list

Products ^
Spark
SparkSQL
Trino
Neon
ADB

Components ^
Hive Metastore
Jupiter
HUE

IaaS ^
PostgreSQL



[← Back to list](#)

Test ADB

UUID: 491177f1-1f14-430d-a9df-bdf87a2bc3d8

Installation in progress

Current installation state: creating

Common info

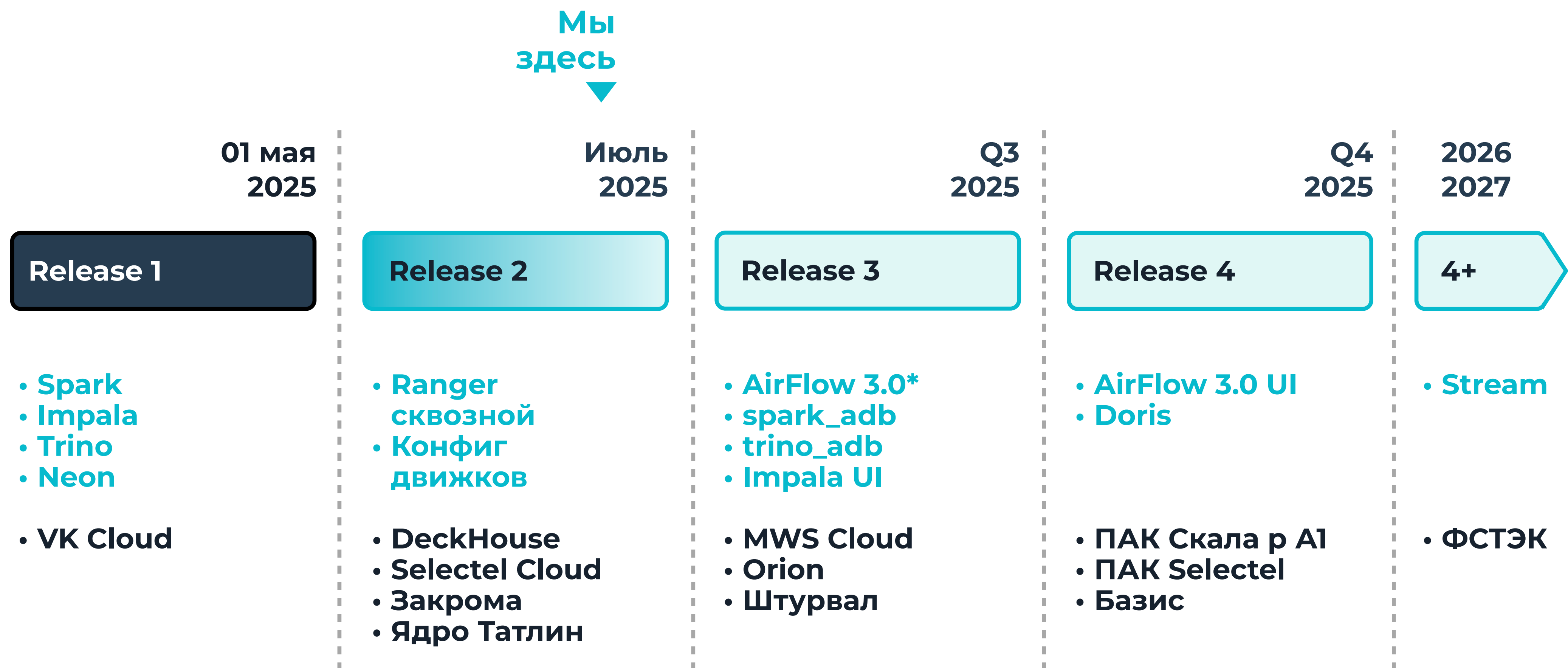
Created date	09.06.2025 15:53:02 +03:00
State	creating
Kind	ADB
Owner	5a1ef9ed-8f62-4725-9055-0984ecbbad26
Project ID	a89b0ff6-d515-4e69-a0f9-bac00163f686

Resource options

Enterprise	☐
Product Options Type	ADB ▾
Size	10

Enterprise Ready

Дорожная карта



Функциональный состав Release 1.0.0 и планы по развитию

Вычисления	Транспорт	Мониторинг	Каталог
Apache Spark3.5.2	Trino451	Grafana11.0.0	Data Catalog0.8.1
Apache Impala4.4.1	Apache Airflow*3.0	Victoria Metrics1.93.5	
Neon7381	Apache Flink	OpenSearch2.17.0	Ресурсы
Picodata	Apache Kafka		Cluster Manager1.0.0
Apache DorisResearch	Apache NiFiResearch	Интерфейсы	Безопасность
StarRocksResearch	DBT	JupyterHub5.1.0	Apache Knox2.5.0
DuckDBResearch		Apache Hue4.11.0	Apache Ranger2.0.0
		Apache Kyuubi1.9.0	

Поддерживаемые платформы S3 и K8S

Протестировано

План

VK Cloud S3

Apache Ozone®

Minio

Selectel S3

Ceph

Скала р МХД.О

Yadro Tatlin.Object

MWS S3

Закрома.Хранение*

Протестировано

План

VK Cloud Containers

Deckhouse K8S

Selectel Managed K8S

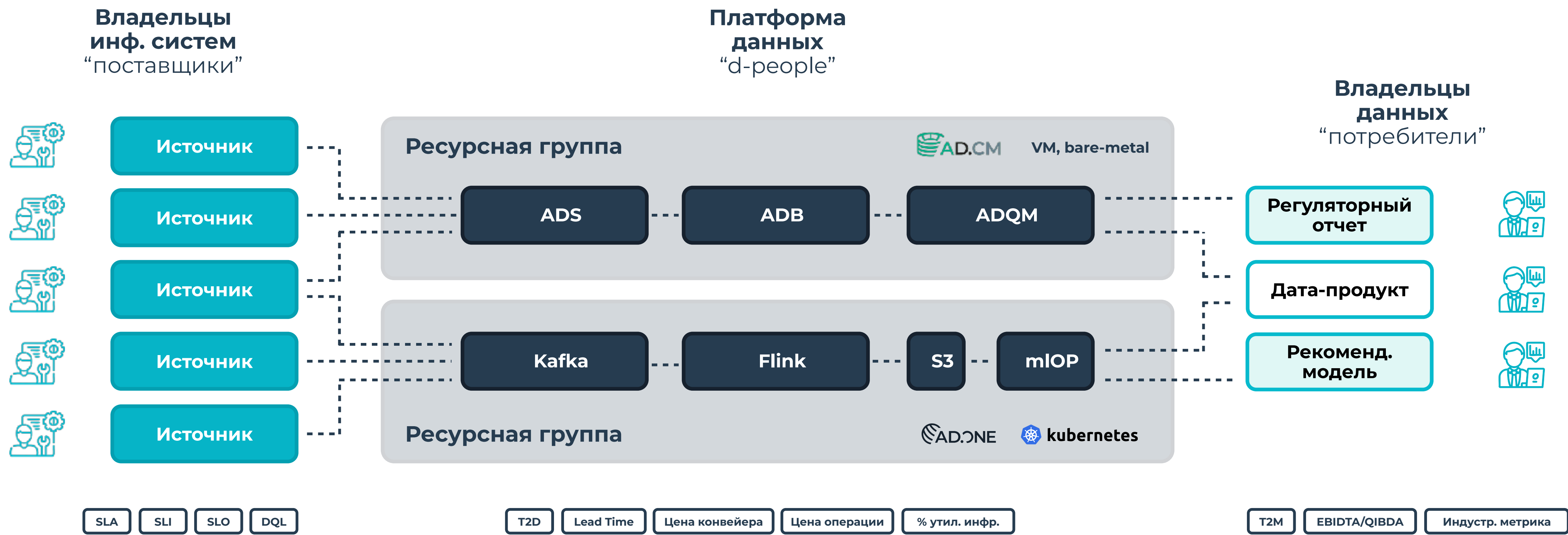
Selectel ПАК

Скала р МВ.К

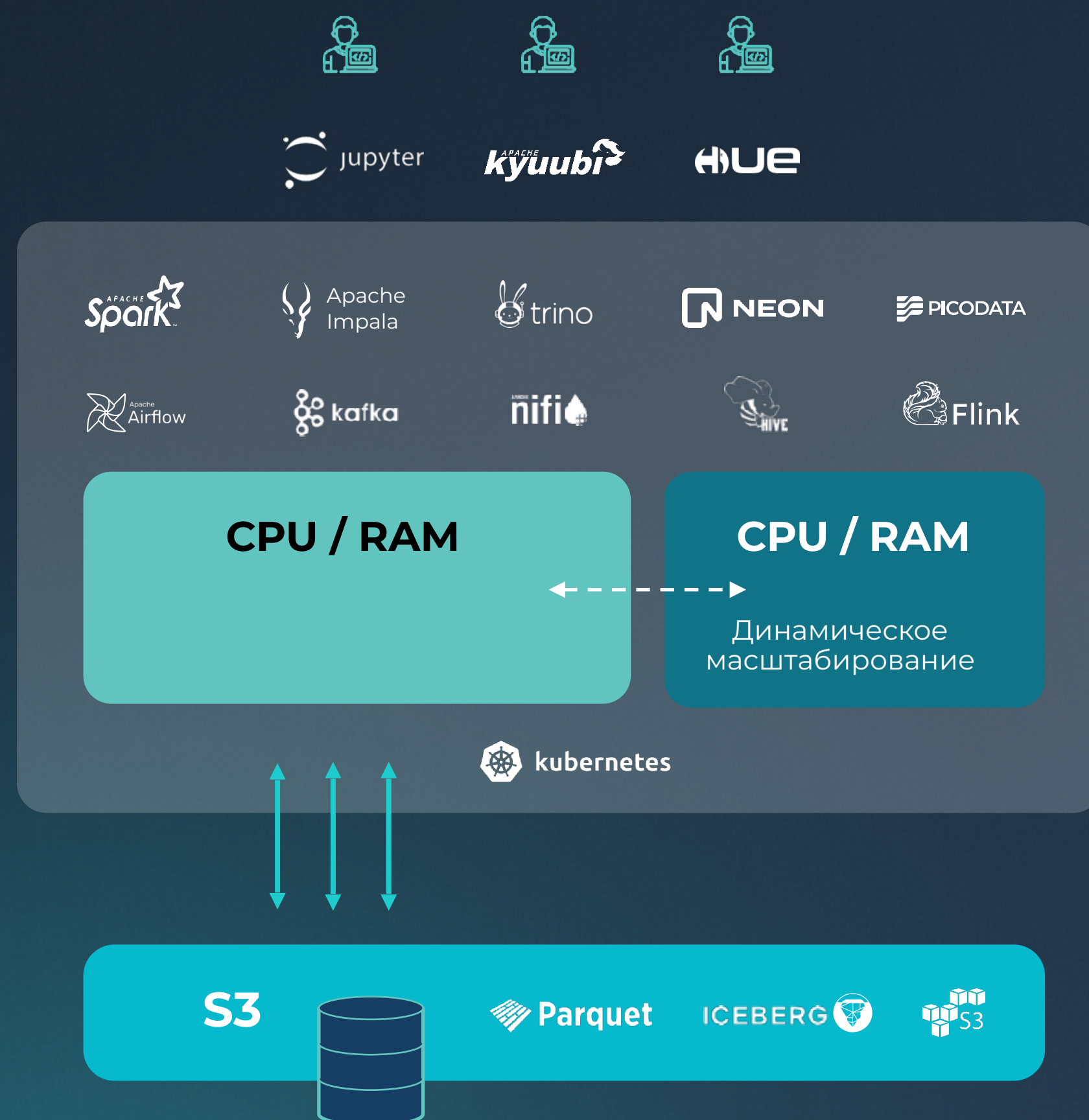
Orion Nova

Баланс цена/ценность

Дата-контракт - модель взаимодействия поставщиков и потребителей на основе баланса стоимости и ценности данных



Платформа Arenadata One адаптирована для хранения и обработки больших данных



ТСО

- Контроль и оптимизация затрат на хранение и обработку данных на основе S3-хранилища и эластичности K8S

T2M

- Инструменты управления кластером, типовые конвейеры данных, доступ пользователей к данным через каталог, типовые инструменты работы с данными

ЦЕННОСТЬ ДАННЫХ

- Фокус извлечение ценности из данных, на сценарии использования, партнерские кейсы и контракты данных для оптимального баланса стоимость/ценность

ENTERPRISE READY

- Исполнение корпоративных требований в области безопасности, надежности, производительности, а также финансовой устойчивости и дорожной карты развития



Дмитрий Гайворонский

Arenadata One BDM
d.gaivoronsky@arenadata.io