



Hewlett Packard
Enterprise

УНИФИЦИРОВАННАЯ ПЛАТФОРМА ХРАНЕНИЯ И ОБРАБОТКИ ДАННЫХ

Сергей Карташев, архитектор решений
8 февраля 2022

ВАШИ ДАННЫЕ – АКТИВ?



Ресурс для получения прибыли



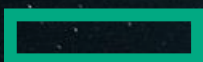
Безопасность и защита



Управление жизненным циклом



Ответственность за ввод,
изменение, использование



ВАША СИСТЕМА ХРАНЕНИЯ И АНАЛИЗА ДАННЫХ ДЕЙСТВИТЕЛЬНО BIG DATA? ИЛИ ОНА ПРОСТО БОЛЬШАЯ?



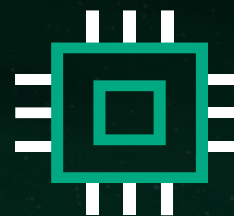
Пределы масштабирования объема



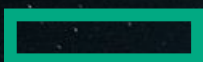
Масштабирование
производительности



Стоимость апгрейдов



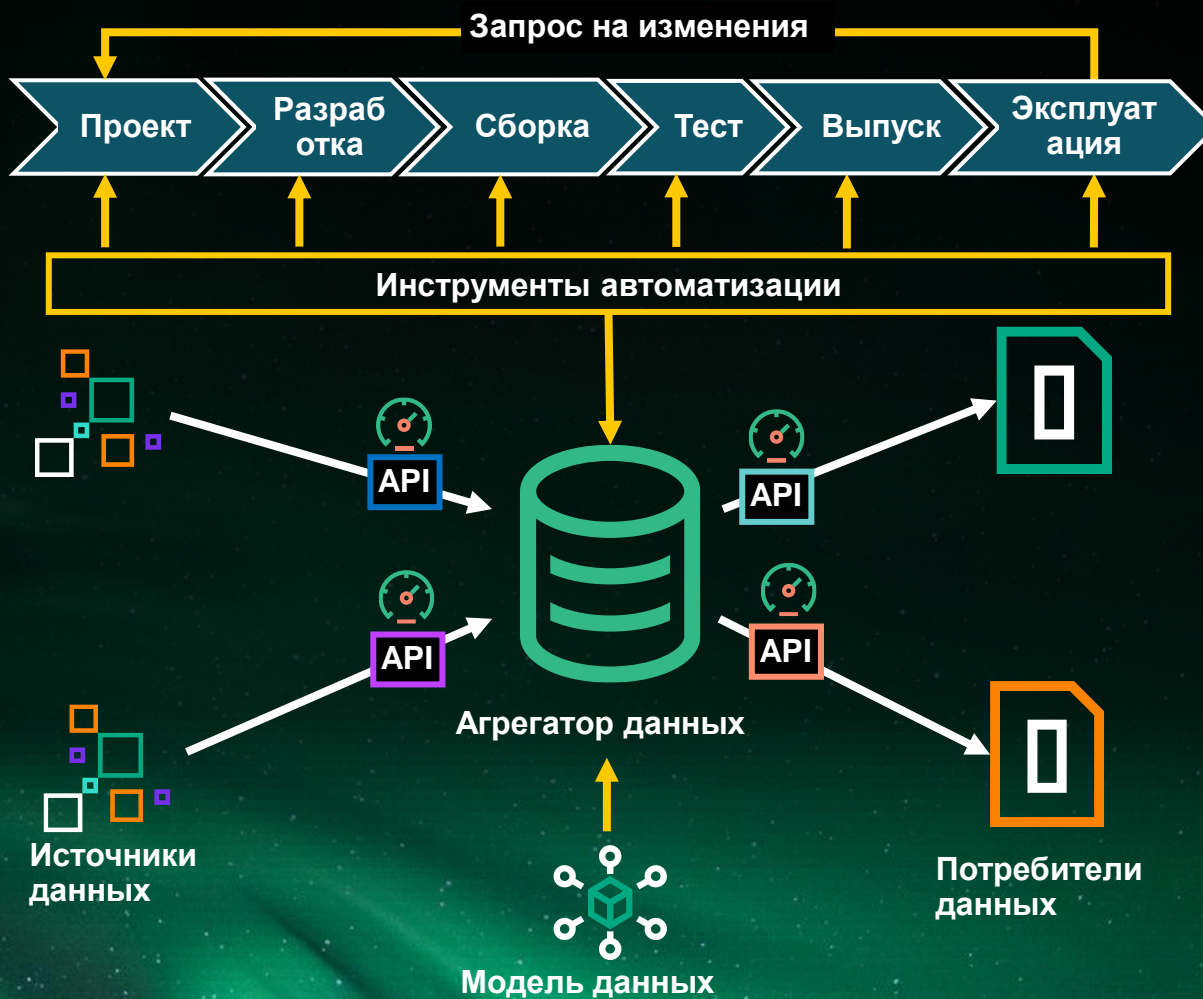
Архитектура и управление
(просто/сложно)



ПРИНЦИПЫ ПРОЕКТИРОВАНИЯ АРХИТЕКТУРЫ ПЛАТФОРМ, ОРИЕНТИРОВАННЫХ НА ОБРАБОТКУ БОЛЬШИХ ДАННЫХ

- Все данные передаются в виде потока событий через агрегатор данных
- Все источники публикуют метаданные
- Поддержка модели данных с помощью каталога данных
- Каждый интерфейс агрегатора данных – управляемый микросервис
- Элементы каталога данных доступны через эти интерфейсы с помощью API
- Интеграция с конвейером разработки (CI/CD Pipeline)

[От анархии данных к их монетизации | Открытые системы. СУБД | Издательство «Открытые системы» \(osp.ru\)](#)



ПЛАТФОРМЫ ХРАНЕНИЯ ДАННЫХ



Разрозненные хранилища
(Silos)

- Нет полного представления о данных
- Нет возможности аналитики по всему набору данных



Хранилище данных
(Data Warehouse)

- Централизованное хранение структурированных данных
- Данные преобразованы и отсортированы
- Нет возможности или очень дорого хранить и анализировать неструктурированных и потоковые данные



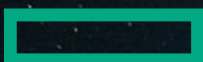
Озеро данных
(Data Lake)

- Централизованное хранение без преобразования
- Лучшее масштабирование
- Иногда требуется значительное время на внедрение
- Ниже производительность при работе со структурированными данными



Дом у озера
(Data Lakehouse)

- Централизованное хранение без преобразования
- Данные (полу-) структурированы
- Может использоваться для BI и ML
- Эффективная стоимость хранения

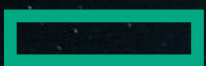


РАЗРОЗНЕННЫЕ ХРАНИЛИЩА ДАННЫХ

Инженеры по данным

Бизнес-аналитики

Ученые по данным



УНИФИЦИРОВАННАЯ ПЛАТФОРМА ХРАНЕНИЯ ДАННЫХ

Инженеры по данным



Бизнес-аналитики



Ученые по данным



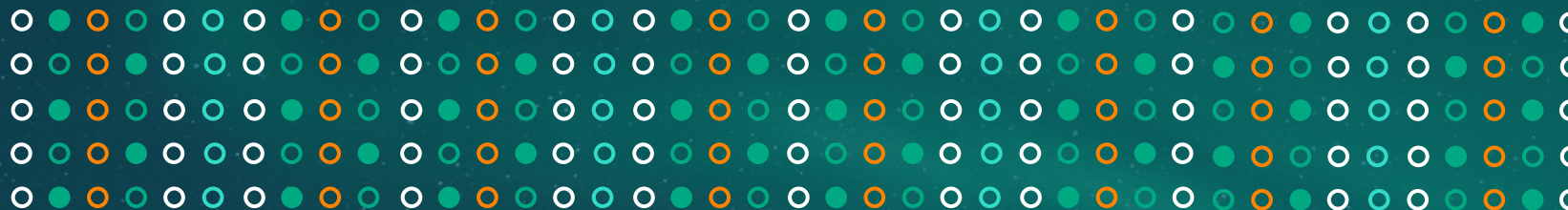
HPE EMERALD

Озеро данных и хранилище данных в одной унифицированной платформе

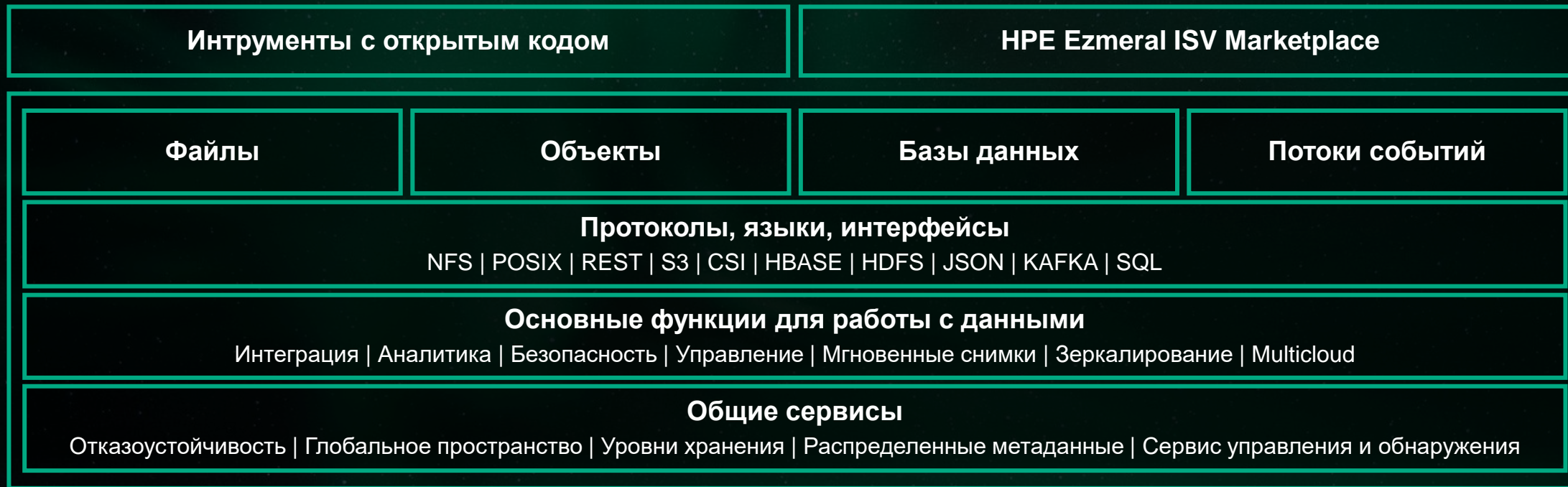
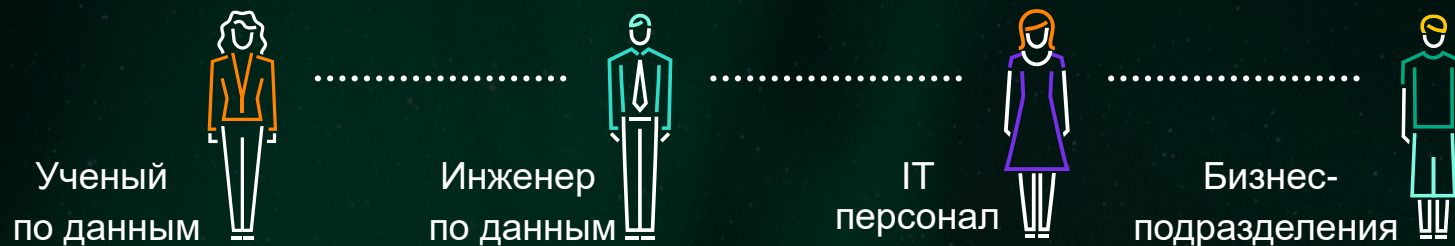
Безопасное
использование
данных

Ускоряет продуктивность
команд аналитиков при
работе с данными

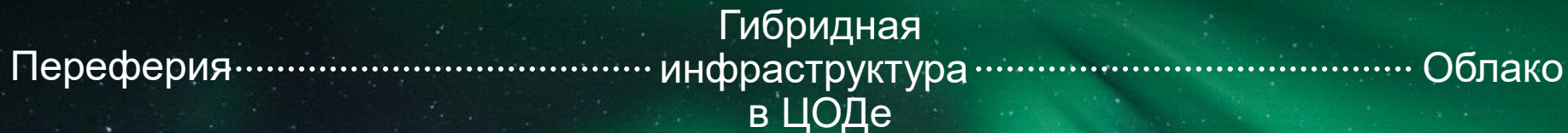
Нет припятствия
для инноваций на
основе данных



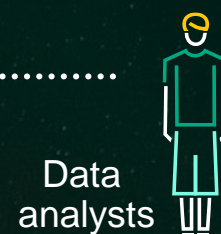
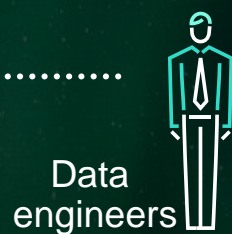
HPE EZMERAL DATA FABRIC ANALYTICS



Использование данных независимо от их местоположения



HPE EZMERAL UNIFIED ANALYTICS



Инструменты анализа данных

Notebooks | Workflow (AirFlow) | App store | ISV tools

Движки анализа и управления данными

Apache Spark + Delta Lake | ISV SQL engines (Presto, Dremio, and Starburst)

Метаданные

Общее и безопасное хранилище метаданных (hive metastore)

Интерфейсы подключения приложений и данных

Доступ к внешним данным с DataTap и FS Mount | Thrift Server | Сервисы доступа к данным (POSIX, DFS, Object Store)

Хранилище данных

HPE Ezmeral Data Fabric file and object | Apache HDFS | Сторонние решения (ISV)

Runtime orchestration

Кластеры Kubernetes (100% CNCF) | Импортирование кластеров | Работа с GPU

Периферия..... Гибридная инфраструктура в ЦОДе Облако

Инженеры по данным



Бизнес-аналитики



Ученые по данным



Инструменты автоматизации

Портал/магазин приложений



S3

NFS

HDFS

Data Fabric File and Object

ЦОД



Файлер



Озеро



Инфраструктура

Облако



Инфраструктура

Периферия



Инфраструктура



ingenico
GROUP

5+ млрд.
транзакций

300+
ТИПОВ
платежей

Работа в
170
странах

“This data simplification has allowed us to shift the focus from platform capacity to business and functional goals. In business today, as they say, vision without execution is just an illusion.”

– Frédéric Delhalle, Retail Director of Strategy and Architecture, Ingenico Group



~65 ТЫС
корпоративных
заказчиков

300 МЛН
потребителей
сервиса

85 ТЫС
потоков данных

“MapR gives us the agility to meet demands of the business and the ability to shift when those demands change. The flexibility we have now is earth-shattering given the scale we’re running.”

– Kevin McClowry, Director of Analytic Solution Development, TransUnion

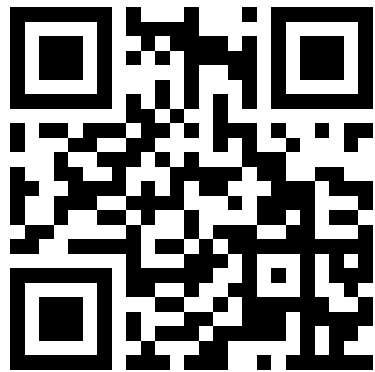
ПРИСОЕДИНЯЙТЕСЬ К КАНАЛАМ НРЕ В СОЦИАЛЬНЫХ СЕТЯХ



TELEGRAM

@hpedigitize

<https://t.me/hpedigitize>



ВКОНТАКТЕ

@hperussia

<https://vk.com/hperussia>

- Топ-новости компании
- Информация о новых продуктах и решениях
- Расписание вебинаров и мероприятий

СПАСИБО

