



Группа компаний
Программный
Продукт

Российские решения по обработке больших данных и мониторингу инфраструктуры

Импортозамещение на примере решения задач в сфере
больших данных и мониторинга инфраструктуры для
государственных заказчиков и крупных компаний



ОСНОВНОЙ ВЫЗОВ

к российскому ИТ-рынку

Рост объемов данных и инфраструктуры.

Импортозамещение – невозможность использования иностранных решений.

!Владимир Путин запретил использовать иностранное программное обеспечение госорганам и госзаказчикам на критической инфраструктуре. Соответствующий указ опубликован 30 марта.

Запрет вводится с 1 января 2025 года. До этой даты владельцы критической инфраструктуры должны перейти на отечественный софт. Закупки иностранного программного обеспечения без согласования запрещаются уже с 31 марта этого года.

27,1

миллиарда устройств к 2025 году*

175

зеттабайт данных ежегодно к 2025 году**



*По данным IoT Analytics к 2025 году, будет более 27 миллиардов подключений к Интернету устройств.

** По данным Data Age Report человечество сформировало порядка 51 зеттабайта информации в 2020 году. К 2025 году объем этих данных вырастет до 175 зеттабайтов ежегодно.



Текущие вызовы рынка

в области решений обработки
больших данных



Компоненты в open-source репозиториях

Компоненты для работы с большими данными находятся в open-source репозиториях и не имеют общего стабильного дистрибутива Hadoop в открытом доступе.



Отсутствие enterprise-решений

Стоимость enterprise-решений, представленных Cloudera, значительно выросла и не доступна для покупки. Другие Community-версии не предлагают полного функционала для работы.



Невозможность восстановления системы

Большинство заказчиков не имеют репозитория для восстановления компонентов системы. При полном разрушении компонента или инфраструктуры восстановление системы может стать невозможным, что приведет к полной потере данных.



Негибкий подход

Использование большинства имеющихся иностранных систем не предполагает гибкого подхода к требованиям заказчика.



Переход на российские системы

Невозможность приобретения иностранных enterprise решений и требование по переходу на российские системы.



Группа компаний
Программный
Продукт

Российские решения

по обработке больших данных
и мониторингу инфраструктуры



СОСТАВ РЕШЕНИЙ

ПЛАТФОРМА ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

GlobalData

ядро Hadoop и ETL-компоненты

Integration

взаимодействие компонентов
и интеграцию, рекомендацию данных

Admin

администрирование, управление, бэкап,
логирование и инвентаризация
распределенной вычислительной среды

Visualization

инструменты, позволяющие
пользователям исследовать
и визуализировать данные

Платформа обработки больших данных

Особенности подхода



Индивидуализированная сборка под нужды каждого заказчика



Каждый экземпляр платформы должен быть собран под задачи заказчика. В полной мере это можно обеспечить только наличием продукта собственной сборки



Минимальное отставание версий компонентов от open-source версий



Разработка собственных модулей, полностью не зависящих от внешних продуктов



Собственная сборка Hadoop на базе open-source решений и собственных разработок



Регулярное обновление системы и добавление в нее обновленных open-source компонентов с предварительным проведением подготовительных работ по стабилизации работы компонента с основной системой.

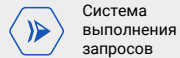
Платформа обработки больших данных

Состав платформы

P³-GlobalData

— основной компонент, включающий в себя ядро Hadoop и ETL-компоненты.

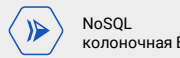
Подразумевает наличие системы установки.



Система выполнения запросов



Система планирования задач



NoSQL колоночная БД



Система обработки больших данных

ЯДРО HADOOP



Файловая система для хранения больших данных



Фреймворк распределённых вычислений



Планировщик ресурсов

P³-Integration

— система, обеспечивающая взаимодействия компонентов и интеграцию/рекомендацию данных



Распределенная система обмена сообщениями



SFTP-сервер



APP сервер



Система управления обработкой информации

P³-Admin

— инструменты администрирования, управления, мониторинга, бэкапа, логирования и инвентаризации распределенной вычислительной среды



Система управления конфигурациями



Инфраструктура для безопасности данных



Система бэкапа и восстановления



Система инвентаризации



Централизованное ведение журналов



Визуализация мониторинга ресурсов



Система мониторинга



Менеджер управления конфигурациями

P³-Visualization

— инструменты, позволяющие пользователям исследовать и визуализировать данные



Инструмент для создания ноутбуков



Приложение для исследования данных



Визуальные представления для данных



Каталог данных



Группа компаний
Программный
Продукт

СИСТЕМА МОНИТОРИНГА И УПРАВЛЕНИЯ ИНФРАСТРУКТУРОЙ

Почему важен централизованный мониторинг?

- Выявление потенциальных аварийных ситуаций.
- Сокращение времени выявления корневых причин аварий.
- Сокращение времени назначения инцидентов на исполнителей.
- Контроль SLA сервисов.
- Контроль выполнения плановых работ.
- Планирование вычислительных ресурсов инфраструктуры.

Система мониторинга и управления инфраструктурой

Одно решение вместо
множества

01. Система управления
обработкой информации



ORACLE®



Sonic ESB

02. Система инвентаризации



comptel



03. Система мониторинга

COMARCH



ГОРОДСКОЙ ТРАНСПОРТ

КЕЙС

Big Data для решения актуальных задач в сфере транспорта для администраций городов и для крупного бизнеса:



Транспортная загруженность

Анализ загруженности уличной сети, общественного транспорта, парковочных мест как на основании ситуации «реального времени» так и прогнозных моделей.



Тарификация и управление

Ведение гибкой тарификации для оптимизации нагрузки на транспортную сеть (например, сниженные тарифы для разгрузки сети в час пик).



Прогнозная модель

Прогнозы прибытия транспорта и построение маршрутов с учетом загруженности дорог, типов груза и персональных предпочтений водителя.



Взаимодействие с пользователем

Предоставление уведомления о дорожной ситуации, взаимодействия на базе прогнозных моделей и обработки сырых данных.



Выявление трендов и аномалий

Отслеживание и анализ трендов и аномалий, прогнозирование аварий и пробок, синхронизация с службами метеорологии и МЧС.



Проектирование транспортной среды

Анализ транспортных потоков и матриц корреспонденций для оптимального проектирования схем развития транспортной среды.

КЕЙСЫ

ВНЕДРЕНИЕ

Продуктовая аналитика
– как инструмент
быстрого внедрения



Построение ML-моделей

Экспертиза в построение быстрых и качественных ML-моделей позволит спрогнозировать, предсказывать результаты и оптимизировать различные процессы



Интеграционный модуль

Наличие готовых алгоритмов и модуля интеграции позволит обеспечить быстрое подключение источников данных по разным протоколам



Data Quality

Разработанный инструмент с уже готовыми метриками качества данных позволит поднять показатели на новый уровень



Автоматизация процессов обработки данных

Экспертиза в автоматизации/оптимизации различных процессов обработки данных позволит увеличить скорость работы и бесшовно интегрировать платформу в экосистему компании



Хранилище – как живой организм

Экспертиза в построении хранилищ данных дает возможность оперативного прогнозирования потребностей заказчика



Бизнес-аналитика

Подход к аналитике с точки зрения бизнес анализа позволит посмотреть на данные с другой стороны и извлечь максимальную выгоду



Группа компаний
Программный
Продукт

Спасибо за
внимание!

