СБЕР

Рабочее место D-people — опыт СБЕР, на примере блока «Розничный Бизнес»

Дмитрий Бугайченко

Немного о себе:

2020 SberRecSys, CDS РБ

2019 Облачная платформа рекомендации

2018 Автоматические A/B тесты

2017 Рекомендации контента для витрин

2016 Рекомендации друзей

2015 Умная лента

2014 Универсальная аналитическая платформа

2013 Рекомендации Групп

2012 Рекомендации музыки

2011 Работа в Mail.ru

2009 Преподаватель СПбГУ

2008 Кандидат физ-мат наук

Зачем бизнесу данные?

Кто такие D-people?

Data Engineer

Data Analyst

Data Scientists

Кто такие D-people?

Data Engineer

Data Analyst

Data Scientists

ML Engineer

D-people РБ в цифрах

Люди

140+ DS

60+ DA

350+ DE

Источники

75+ систем источников данных

200+ витрин

Ресурсы

8000+ ядер

60+ ТБ памяти

1000+ ТБ дисков

Немного истории

Как все начиналось…

Сервер в бою

JDBC/SCP

D-people

E-mail

Внутренние пользователи

Подстольник

Работа с данными на «подстольниках»

+ Быстро организовать
+ Гибкий доступ к данным
+ Свободный выбор инструментов

- Большие риски
- Неуправляемое качество данных
- Отсутствие полной картины
- Ограниченность вычислительных ресурсов
- Много дублирования
- …

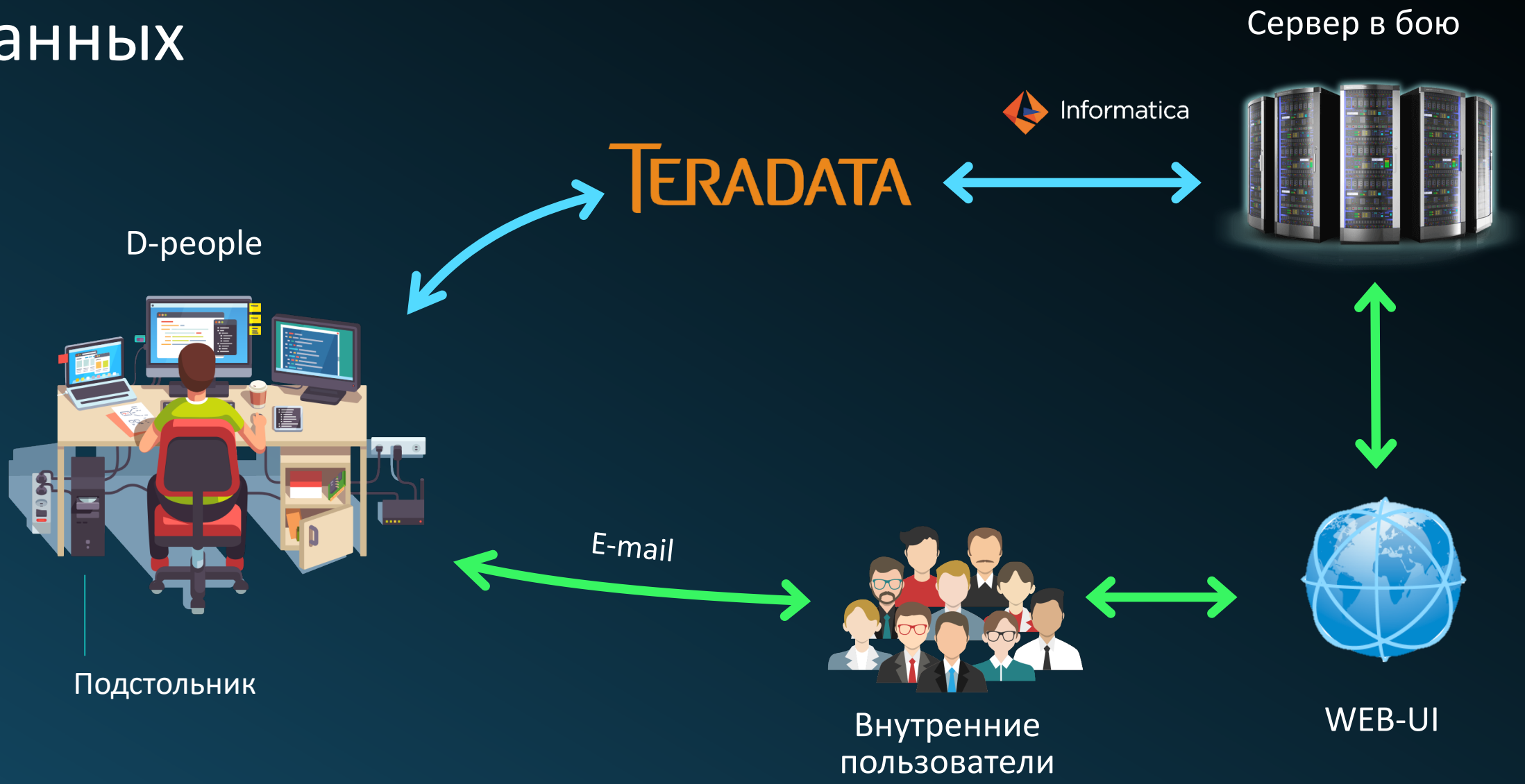
Но потом пришла кибербеза!

… а после пришло осознание ответственности перед клиентом

а потом пришли осознание ответственности перед клиентом и кибербеза!

И появилась…

Аналитическое хранилище данных



Централизованное хранилище данных на базе Teradata

TERADATA

+ Общая картина в одном месте
+ Контроль качества данных
+ Возможность строить промышленные процессы

– Большой T2M подключение новых источников
– Ограниченность функционала для ДС
– Плохая масштабируемость
– Vendor lock

Централизованное хранилище данных на базе Teradata

TERADATA

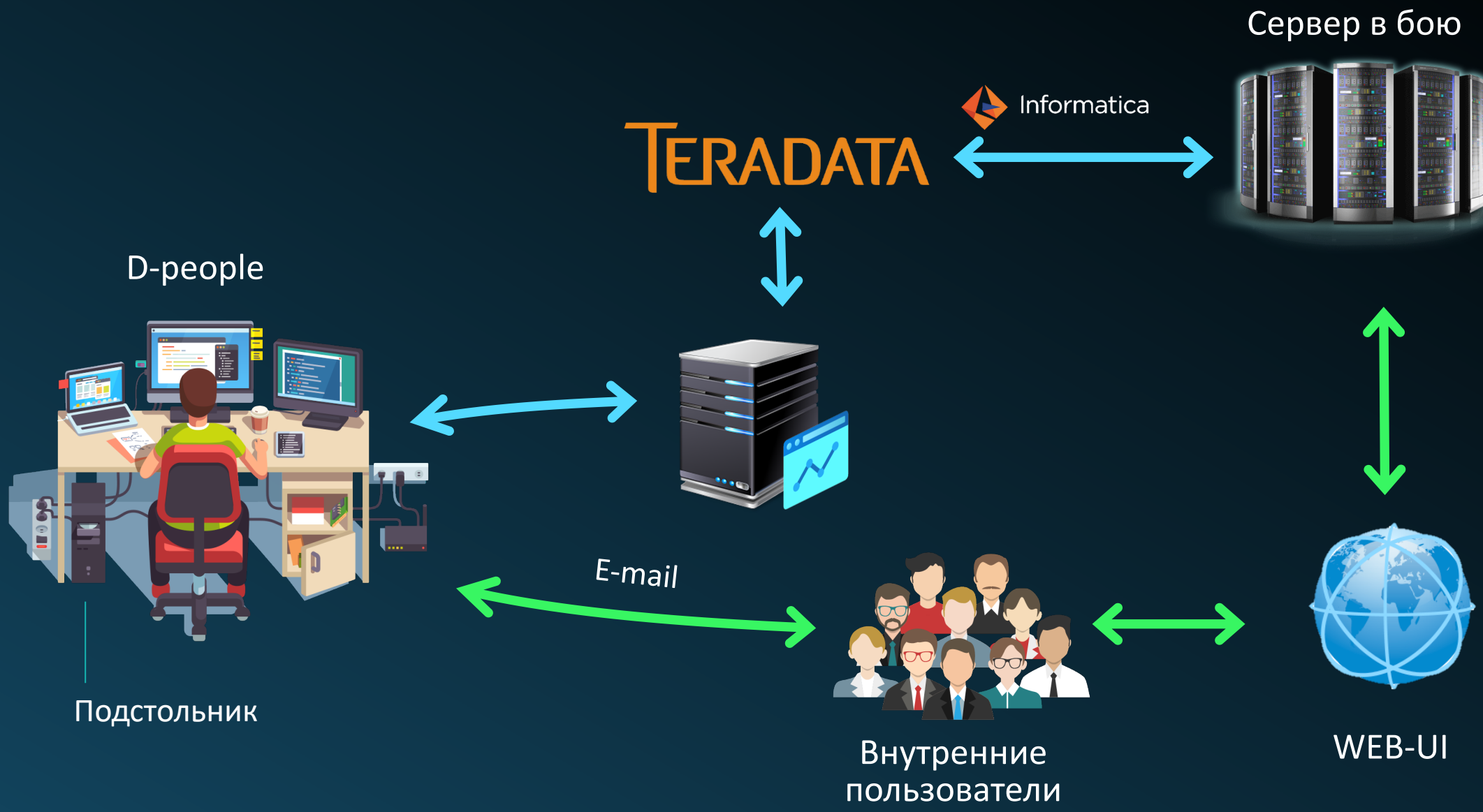
+ Общая картина в одном месте
+ Контроль качества данных
+ Возможность строить промышленные процессы

– Большой T2M подключение новых источников
– Ограниченность функционала для ДС
– Плохая масштабируемость
– Vendor lock

Клиентская аналитика РБ

- Несколько больших серверов (1ТБ памяти, GPU) с подключением к Teradata
- Небольшой кластер Hadoop для ETL
- Airflow для запуска регулярных процессов
- Инструменты SAS для DS и DA
- Классический Python

Клиентская аналитика РБ



Клиентская аналитика РБ

+ Привычные ДС-ам инструменты
+ Большей мощности, чем «подстольники»

– Ресурс ограничен одним сервером
– Людей больше чем серверов
– Ограничены возможности по опромышливанию

Клиентская аналитика РБ

+ Привычные ДС-ам инструменты
+ Большей мощности, чем «подстольники»

– Ресурс ограничен одним сервером
– Людей больше чем серверов
– Ограничены возможности по опромышливанию

Первый опыт Hadoop – Лаборатория Данных 2.0

- Один большой кластер Hadoop для всех D-people
- Hue, Spark, Jupyter
- «Условно-бесплатный»

Первый опыт Hadoop – Лаборатория Данных 2.0

- Один большой кластер Hadoop для всех D-people
- Hue, Spark, Jupyter
- «Условно-бесплатный»
- Очень быстро пришел в «негодность»
- При 100+ пользователей частые ошибки с сайдэффектами на другие приложения
- Встроенные механизмы управления ресурсами (очереди, preemption, динамическая аллокация Spark) не помогли
- **Во многом из-за недостатка опыта у пользователей**

Текущий статус

При поддержке SberData, SberInfra, Блока КИБ и многих хороших людей ☺

Извлекаем уроки - каким должно быть рабочее место D-people?

- Быстрый поиск и получение доступа к данным
- Достаточно вычислительной мощности для работы с большими данными и ДС задачами
- Одновременная работа большого числа пользователей
- Безопасность и контроль качества данных
- Открытые технологии без vendor lock
- Понятные пути по опромышливанию

Каждому D-people - свое рабочее место!

DE

- Распределенный MapRed-like движок
- Язык программирования
- IDE, unit tests
- Схемы данных для отладки
- Реальные данные для НТ

DA

- Массивно-параллельный движок, кубы
- Диалект SQL
- Ad-hock-запросы, визуализация
- Только реальные качественные данные

DS

- Суперкомпьютер
- Python с ML инструментами
- Итеративные вычисления, подбор гиперпараметров
- Только реальные качественные данные

Шаги D-процесса

Найти данные  Получить ресурсы  Получить данные  Получить результат

Шаги D-процесса

Найти данные

Получить ресурсы

Получить данные

Получить результат

Ищем данные

Карта данных

- Поиск по витринам
- Техническая
  и административная информация

МЛ360

- Поиск признаков
- Детальная статистика
- Информация
  по использованию
- Кодогенерация

Ищем данные

Карта данных – для DE

- Поиск по витринам
- Техническая
  и административная информация

МЛ360 – для DS

- Поиск признаков
- Детальная статистика
- Информация
  по использованию
- Кодогенерация

Шаги D-процесса

Найти данные

Получить ресурсы

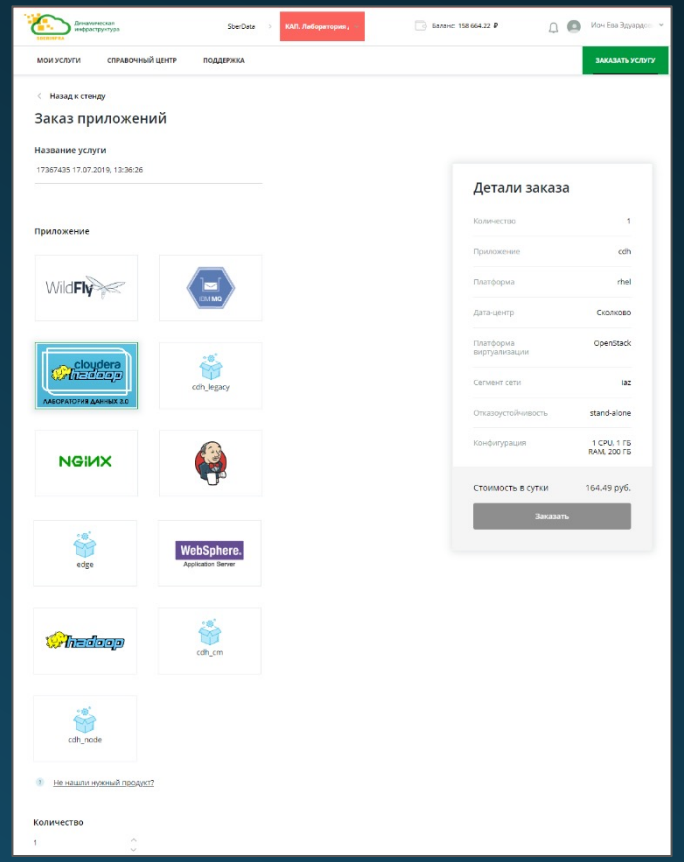
Получить данные

Получить результат

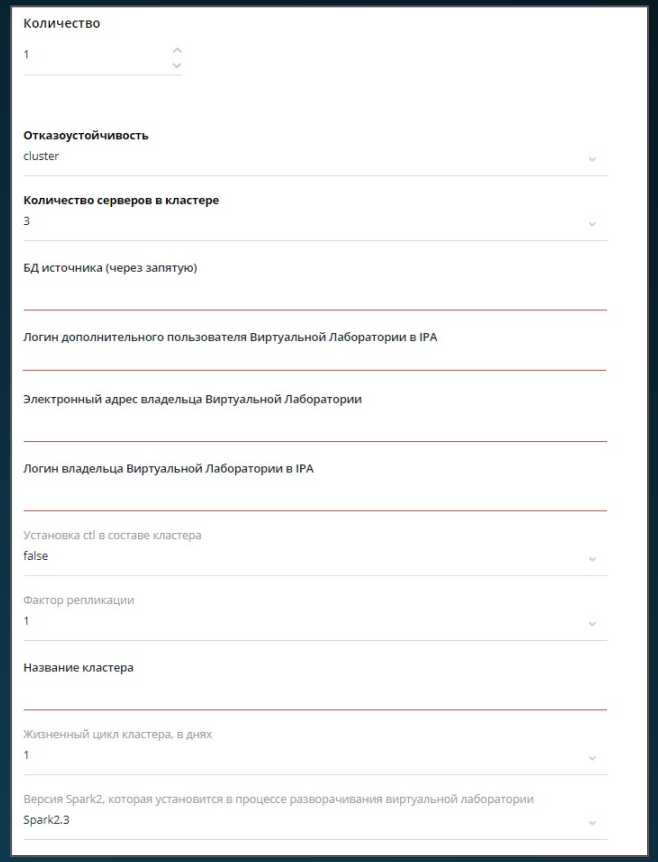
Получение ресурсов – лаборатория данных 3.0

Возможность быстро (~1 день) развернуть выделенный (но небольшой) кластер в облаке

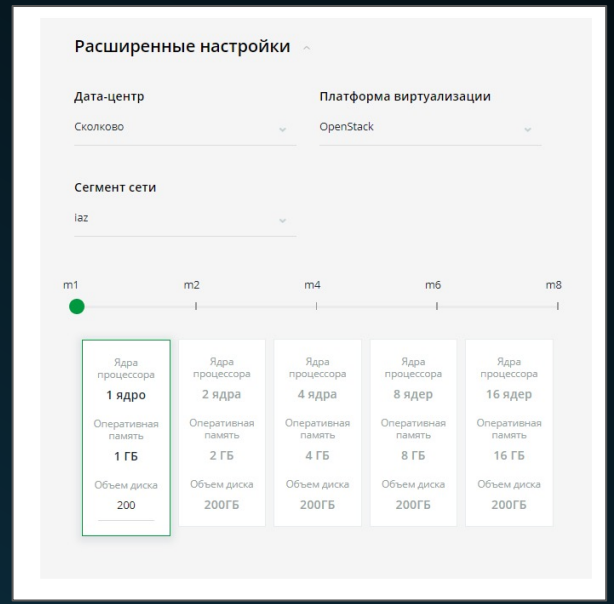
1. Выбор услуги CDH



2. Заполнение данных по кластеру



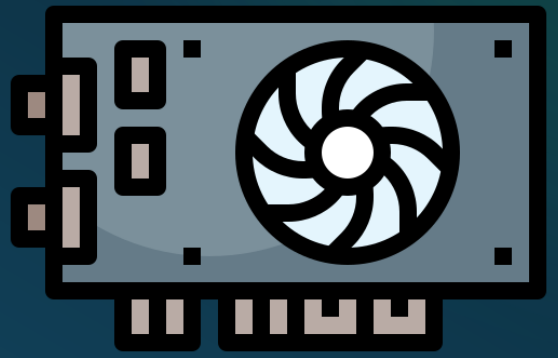
3. Выбор конфигурации



Получение ресурсов – лаборатория данных 3.0

- **Появилась изоляция, но с ней и новые проблемы**
  - Нужно планировать свой бюджет
  - Размер кластера ограничен, не реализовать преимущества платформы Hadoop
  - Утилизация неравномерная
  - Утилизация в целом по системе очень низкая
- **Удобно для DE, но мало возможностей для DS**

Получение ресурсов – хочу GPU, что делать?!



- **Идти на Кристофари**
  - Но без большей части данных
- **Припрятать DGX-2**
  - А если найдут?
- **Подключить «машину дата сайентиста» к ЛД**
  - Закончились ☹
- **Использовать сервис Datalab!**

Получение ресурсов – сервис Datalab

- Возможность получить контейнер с GPU в облаке
- До 16 GPU на контейнер и Jupyter ноутбук
- Подключение к HDFS своей лаборатории данных

Шаги D-процесса

Найти данные

Получить ресурсы

Получить данные

Получить результат

Супермаркет данных – универсальный инструмент поставки данных

Копирование

- Данные копируются из кластера-источника в кластер потребитель
- Минимум накладных расходов после
- – Очень дорого в процессе

Проксирование

- Данные доступны в кластере источнике через прокси
- Не накладных расходов на распространение
- – Плохо предсказуемая нагрузка на источник

Супермаркет данных – универсальный инструмент поставки данных

Копирование

- Небольшие справочники
- Витрины с:
  - Партиционированием по дате
  - Среднего размера
- Востребованные во многих процессах на потребителе

Проксирование

- Крупные витрины без партиционирования по дате
- Редко используемые витрины

Проблема транзитивных доступов

Вывод – доступ к данным надо оформлять
для лабораторий, а для D-people к лабораториям

Шаги D-процесса

Найти данные  Получить ресурсы  Получить данные  Получить результат

Инструменты DE

- **Spark submit is all you need ☺**
- **Но можно использовать и**
  - IDE
  - Hue, Jupiter + Toree
  - HBase, Flink
  - …
- **Bitbucket**

Инструменты DS

Теория

- Добавляем к ЛД узел с Jupiter и PySpark
- Через парсели ставим нужные пакеты Python

Практика

- Узел с Jupiter на порядок меньше чем «большие тачки» прошлого
- Привычные пакеты Python не могут утилизировать кластер Hadoop

Куда податься DS-у?

- **На Datalab с GPU**
- **Использовать пакеты распределенного МЛ**
  - LightGBM, XGBoost, (CatBoost)
- **Учить Scala!**

Подключаем инструменты - DA

Лаборатория Greenplum

- Большой коммунальный кластер
- Быстрый переход с Teradata

Но есть проблемы

- Нет нормального разделения доступов
- Грабли ЛД2.0 на лицо

Поделиться результатами – Persistent Storage

- **Кластер ЛД3.0 – эфемерная сущность в облаке**
- **Для хранения важных результатов – отдельный кластер HDFS**
- **Можно использовать для обмена результатами**
- **Проблемы**
  - Транзитивные доступы
  - Durability vs. reproducibility

Рабочее место D-people – helicopter view

Найти данные

Получить ресурсы

Получить данные

Получить результат

АРМ Alpha

Контур Фабрики Данных

Карта данных

ML360

BitBucket

Кластер источник 1

Кластер источник N

Супермаркет данных

АРМ Фабрики данных

Кластер ЛД 3.0

Jupiter Node

Spark Node 1

Spark Node N

GPU container

Datalab

Следующие шаги

Нехватка ресурсов при низкой утилизации в лаборатории данных 3.0

Проблема:

Часть кластеров ЛД3.0 простаивают, тогда как на других остро не хватает ресурсов в моменте

Что будем делать?

- Инструменты для быстрого временного расширения кластера под задачу
- Построение командных кластеров
- Обучение DS/DA эффективному использованию платформы

Долгое оформление доступов

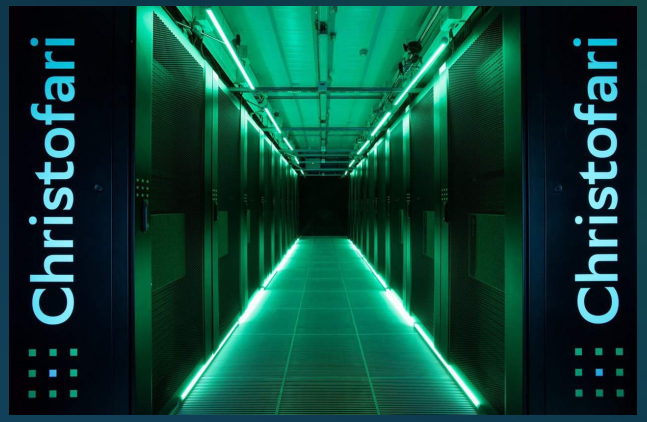
Проблема:

Доступы приходится оформлять часто, это отнимает время, но не дает защиты из-за транзитивности

Что будем делать?

- И витрины, и D-people помечаются тегами
- Кластер ЛД наследует теги от всех витрин
- Если тэги кластера есть подмножество тэгов D-people, он имеет доступ к нему и всем витринам на нем

Сложно получить контейнер на много GPU



Проблема:

В Datalab есть много GPU,
но получить их на одном контейнере
сложно

Что будем делать?

- Для тех задач, где не нужны быстрые каналы между GPU, внедряем инструменты оркестрации на базе Ray/Horovod
- Подходит для подбора гиперпараметров GBDT

Нет удалёнки для DS/DA

Проблема:

Если у D-people есть доступ
к персональным данным, у него закрывают
удаленный доступ

Что будем делать?

- Выделенный контур для кластер ЛД с удаленным доступом
- Анонимизация данных и синтетика
- ML Ops для поставки кода
из удаленного сегмента
во внутренний

Вывод результатов в ПРОМ

Проблема:

Выводить результаты
в промышленную эксплуатацию долго и мучительно

Что будем делать?

- TO BE CONTINUED

СБЕР

Создавайте будущее вместе с нами!

https://sberbank-talents.ru