

Министерство науки и высшего образования Российской Федерации

Национальный исследовательский университет ИТМО

Инфраструктура больших данных

Весна

2025

Лабораторная работа №8

МИГРАЦИЯ НА KUBERNETES

Цель работы:

Получить навыки оркестрации контейнеров с использованием Kubernetes путём миграции сервиса модели на PySpark, сервиса витрины на Spark и сервиса источника данных.

Ход работы:

1. Создать инфраструктуру для Spark вычислений, настроить репликацию:
<https://spark.apache.org/docs/latest/running-on-kubernetes.html>
<https://habr.com/ru/companies/neoflex/articles/511734/>
2. Запустить сервис модели (версии лабораторной №5) в кластере k8s. Проверить работоспособность.
3. Запустить сервис источника данных (версии лабораторной №6) и накатить обновление сервиса модели в кластере k8s. Проверить работоспособность.
4. Запустить сервис витрины данных (версии лабораторной №7) и накатить обновления остальных двух сервисов в кластере k8s. Проверить работоспособность.
5. Оптимизировать утилизацию ресурсов. Возможно упрощение инфраструктуры.
6. Допускается сборка в helm chart, использование OpenShift, а также размещение секретов в k8s services.

Результаты работы:

1. Отчёт о проделанной работе;
2. Ссылка на репозиторий GitHub;
3. Актуальный дистрибутив модели в zip архиве.

Полезные материалы:

1. <https://kubernetes.io/ru/docs/tutorials/kubernetes-basics/>
2. https://youtube.com/playlist?list=PLg5SS_4L6LYvN1RqaVesof8KAf-02fJSi
3. <https://mcs.mail.ru/blog/razvorachivanie-prilozhenie-na-apache-spark-v-kubernetes>