

## Программа курса

### «GRAD: Администрирование Greenplum / Arenadata DB»

**О курсе:** 4-дневный обучающий курс «Администрирование базы данных Arenadata DB (Greenplum)» предоставляет необходимые знания и навыки по сопровождению и поддержке кластера Arenadata DB при эксплуатации в облачной инфраструктуре и дата центре.

**Аудитория:** Администраторы и специалисты, желающие получить практические навыки по установке, конфигурированию, обслуживанию и управлению кластером Greenplum / Arenadata DB.

**Уровень подготовки:**

- Знание базовых команд Linux (опыт работы с командной строкой, файловой системой, POSIX, текстовыми редакторами vi, nano)
- Основы работы с базами данных
- Опыт работы с PSQL/ SQL

**Продолжительность курса:** 32 академических часа, 4 дня по 8 ак. часов дистанционно

## Содержание программы

### 1. Введение в СУБД Arenadata DB

- Архитектура Arenadata DB (Greenplum)/ MPP
- Greenplum Мастера и Сегменты
- Greenplum Interconnect
- Отказоустойчивость
- Выбор дистрибутива

### 2. Установка и настройка Greenplum / Arenadata DB

- Требования по установке.
- Планирование и масштабирование.
- Установка с использованием ADCM. Ручная установка.
- Настройка Data Layout и параметров окружения.
- Управление сервисами Arenadata DB (Greenplum)
- Клиентское подключение к БД. Клиентские приложения и утилиты.
- Arenadata DB (Greenplum) в облаке
  - **Практическое задание** — установка и настройка Arenadata DB с использованием ADCM.

### 3. Схемы и объекты СУБД Arenadata DB (Greenplum)

- Логическая и физическая структура БД
- Установка по умолчанию: схемы и шаблоны
- Табличные пространства в Arenadata DB
- Внешние таблицы
- Создание и обслуживание индексов
- Создание и управление представлениями (views)
  - **Практическое задание** — предварительная настройка кластера Arenadata DB.

### 4. Операции обслуживания СУБД Arenadata DB

- Управление ресурсами кластера- ресурсные группы
- Управление данными в Arenadata DB
- Вставка, обновление и удаление строк
- Очистка (Vacuum) базы данных
- Масштабирование (Расширение) кластера

- Планы запросов
- Управление файлами логов БД
- Реиндексация
- Задачи операций обслуживания
  - **Практическое задание** — настройка и использование ресурсных групп на кластере Arenadata DB (ADB) Greenplum.
  - **Практическое задание** — операции обслуживания базы данных Arenadata DB.
  - **Практическое задание** — масштабирование кластера Arenadata DB (Greenplum) добавлением узлов сегментных хостов.

## 5. Загрузка и выгрузка данных в Arenadata DB

- Механизмы загрузки и выгрузки данных
- Использование внешних таблиц
- Сценарии загрузки данных в БД
- Интеграции с другими источниками: Kafka, Hadoop, NiFi, S3
  - **Практическое задание** — настройка доступа Arenadata DB (Greenplum) к внешним данным и работа с внешними таблицами.
  - **Практическое задание** — настройка сервиса PXF для кластера Arenadata DB (Greenplum).

## 6. Партиционирование таблиц

- Создание партиционированных таблиц
- Ключ партиционирования
- Разные типы хранения
- Tablespace
- Мульти-партиционирование
- Удаление, разбиение, обмен партиций
- Внешние таблицы (обычные и web-внешние)
  - **Практическое задание** — загрузка в партиционированные таблицы.

## 7. Отказоустойчивость в Arenadata DB (Greenplum)

- Опции отказоустойчивости для Arenadata DB
- Зеркалирование сегментов в Arenadata DB
- Отказоустойчивость Сегментов и восстановление в ADB
- Зеркалирование Мастеров в ADB
- Восстановление сегмента. Восстановление мастера
  - **Практическое задание** — обеспечение отказоустойчивости кластера Arenadata DB с помощью зеркалирования сегментов.
  - **Практическое задание** — обеспечение отказоустойчивости кластера с помощью зеркалирования Мастера ADB.
  - **Практическое задание** — восстановление после сбоя зеркал сегментов и мастера в кластере Arenadata DB.

## 8. Резервное копирование и восстановление БД Arenadata DB

- Основные принципы операций восстановления и резервного копирования
- Backup утилиты Arenadata DB
- Процедура резервного копирования
- Процедура восстановления БД
  - **Практическое задание** — резервное копирование и восстановление кластера Arenadata DB.

## 9. Мониторинг и troubleshooting СУБД Arenadata DB

- 
- Мониторинг и анализ состояния системы (System state)
- Инструменты мониторинга Arenadata DB / ADCM

- Основные причины проблем производительности
- Собираемые метрики и сбор статистики
- Управление нагрузкой и оповещениями
- Мониторинг журналов: анализ сообщений об ошибках и идентификация проблемных запросов
- Оптимизация параметров БД и дизайна БД
- Best Practices
  - **Практическое задание** — настройка мониторинга кластера Arenadata DB.

## 10. Безопасность в Arenadata DB

- Контроль доступа и обеспечение безопасности
- RBAC в СУБД Arenadata DB (Greenplum)
- Поддержка шифрования и аутентификации с использованием SSL
- Интеграция с LDAP
  - **Практическое задание** — настройка LDAP аутентификации для Arenadata DB.

**Примечание: лабораторные работы в курсе проводятся с использованием лабораторного стенда в облаке Яндекс Cloud на community версии дистрибутива Arenadata DB 6.XX**

**! Возможно проведения курса полностью на ванильной версии дистрибутива Greenplum!**