Workshop: Кибер Безопасность в кластере Hadoop (платформа HortonWorks HDP)

3 дня практического обучения по установке и конфигурированию интегрированной безопасности кластера **Hadoop** с использованием протоколов безопасности **Kerberos**, настройки аутентификации **Active Directory** с поддержкой механизмов авторизации и аудита событий безопасности **Apache Ranger**. Рассматривается вопрос настройки защищенного периметра сети с поддержкой **Single-Sign-On** средствами **Apache Knox Gateway.** Политики ограничения доступа **Apache Ranger** на уровне строк, колонок и значений с использованием **Apache Atlas**. Поддержка резервного копирования, репликации и восстановления. Создание и сопровождение **защищенного озера данных** (**Data Lake**) с использованием Apache Metron. Защищенное взаимодействие с компонентами экосистемы **Hadoop**: **Spark**, **Hive**, **HBase, Kafka, HDFS**, **MapReduce**.

Аудитория: Системные администраторы, системные архитекторы, разработчики **Hadoop** желающие получить практические навыки по установке, конфигурированию, обслуживанию и управлению защищенной средой кластера **Hadoop** с использованием дистрибутива **HortonWorks Hadoop Data Platform**.

Предварительный уровень подготовки:

- Начальный опыт работы в **Unix**
- Знания **Hadoop** в рамках курса «**Администрирование кластера Hadoop**»
- Понимание основ информационной безопасности

Продолжительность: **3** дня, **24** академических часа

Курс построен на сквозных практических примерах развертывания и администрирования защищенной архитектуры кластера **Hadoop** вместе с протоколами безопасности **Kerberos**, **SSL**; настройки интегрированной безопасности компонент экосистемы **Hadoop** для унифицированного входа с использованием **Single-Sign-On**, шлюза безопасности **Apache Knox Gateway** и политик разграничения доступа **Apache Ranger**. Практические занятия выполняются на локальных рабочих станциях и в кластерной среде **Amazone Web Services** с использованием дистрибутива **HortonWorks Hadoop Data Platform.**

Соотношение теории к практике 30/70

Программа курса

1. Cybersecurity для больших данных в Hadoop

- Особенности реализации информационной безопасности (далее ИБ) в озере данных Hadoop

mailto:sales@bigdataschool.ru
+7 (985) 162-29-63

- Специфические угрозы ИБ существующие в озере данных
- Организационные меры по ИБ для озера данных

2. Обзор подсистем безопасности озера данных

- Автоматизация
- Аутентификация и защита периметра
- Авторизация
- Аудит
- Защита данных
  - шифрование данных
  - антивирусная защита данных
  - **snapshots**
  - репликация данных
  - резервное копирование и восстановление данных
- **Hardening security** для базовых компонент
  - операционные системы
  - базы данных
  - веб сервисы

3. Построение безопасности озера данных на базе компонентов HortonWorks Hadoop Data Platform

- Особенности дистрибутива **HortonWorks HDP** и базовая безопасность(по умолчанию)
- Администрирование **Apache Ambari** для аутентификации с **Kerberos**
- Настройка протокола **Kerberos** для аутентификации с **Active Directory** (**FreeIPA**)
- Настройка безопасности периметра с **Apache Knox Gateway**
- Настройка **Apache Knox Single-Sign-On**
- **Best Practices** для аутентификации данных и защиты периметра

4. Настройка авторизации в озере данных Hadoop

- Установка **Apache Ranger** с помощью **Apache Ambari**
- Настройка мапирования групп **Ldap** для **Hadoop** аутенитификации
- Настройка **Ranger** плагинов для авторизации компонент экосистемы **Hadoop** с использованием протокола **Kerberos**
- Настройка политик **Rangers** для разграниения полномочий доступа
  - **RBAC** - ролевые политики для разграничения доступа
  - **ResourceBAC** - ресурсные политики разграничения доступа
  - Строковая фильтрация для разграничения доступа
  - Фильтр на колонки для разграничения доступа
  - Политки разграничения на основании меток (**tags**)
  - **Best Practices** для политик разграничения полномочий

5. Защита данных HDFS

- Шифрование данных при передаче (**Data @ Wire encryption**)

    - **SSL** шифрование для подключения к **Web UI** компонент экосистемы **Hadoop**
    - Протокол **SPNEGO**
    - **Best Practices** для шифрования трафика
- Шифрование данных на хранении (**DARE**)
    - Настройка **Ranger KMS**
    - **HDFS** шифрование
    - **Best Practices** для шифрования данных файловой системы
- Управление доступом к **HDFS**
    - **Posix** и **ACL** для **HDFS**
    - **Best Practices** для управления списками доступа для файловой системы
- Антивирусная защита в озере данных

6. Настройка политик аудита в Hadoop

- Использование **Apache Solr** для аудита событий
- Включение аудита для **Ambari** кластера
- Использование аудита для управления в **Ranger**
- **Best Practices** для политик аудита

7. Hardening Security для узлов кластера

- Конфигурация узлов для **non-root** установки
- **Endpoint security** подход
- **Best Practices** для защиты конечных узлов

8. Организационные меры информационной безопасности данных в озере данных

- **Best Practices** для построения защищенного озера данных
- Рекомендации по использованию **ПО** верхнего уровня для защиты озера данных
- Технологии **Machine Learning** для построения защищенного озера данных
- Использование **Apache Metron** для создания защищенной инфраструктуры озера данных

Примечание:
• Доступ к лабораторному стенду на **Amazon Web Services** предоставляется на время учебных курсов с **8:30** до **18:30** (возможно продление времени по запросу)