Workshop: Кибер Безопасность в кластере Hadoop (платформа Cloudera Distributed Hadoop)

Ближайшая дата курса: 1-3 ноября

3 дня практического обучения по установке и конфигурированию эшелонированной защиты кластера **Hadoop** на базе дистрибутива **Cloudera Distributed Hadoop**, с использованием протоколов безопасности **Kerberos**, настройки аутентификации **Active Directory** с поддержкой механизмов авторизации и аудита событий безопасности **Apache Sentry** и **Cloudera Navigator**. Политики резервного копирования, репликации и восстановления. Защита данных для файловой системы HDFS посредством списков управления доступа ACL и Posix, шифрования данных с помощью **Cloudera Navigator** и встроенного маскирования персональной информации. Шифрование и аутентификация трафика с TLS между компонентами экосистемы **Hadoop**: **Spark**, **YARN, Hive**, **HBase, Kafka, HDFS**, **MapReduce**. **Best Practices** по защите данных и периметра озера данных Hadoop под управлением **Cloudera Manager**.

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

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 системы безопасности компонент экосистемы **Hadoop** для унифицированного входа с использованием **Single-Sign-On**, шлюза безопасности **httpFS** и политик разграничения доступа **Apache Sentry**. Практические занятия выполняются на локальных рабочих станциях и в кластерной среде **Amazone Web Services** с использованием дистрибутива **Cloudera Distributed Hadoop.**

Соотношение теории к практике 30/70

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

Программа курса

1. Cybersecurity для больших данных в Hadoop

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

- Особенности дистрибутива **CDH** и базовая безопасность(по умолчанию)
- Настройка **Cloudera Manager** для аутентификации с **Kerberos**
- Настройка протокола **Kerberos** для аутентификации с **Active Directory** (**FreeIPA**)
- Настройка **httpFS** для защиты периметра кластера **Cloudera Hadoop**
- **Best Practices** для аутентификации данных и защиты периметра

4. Настройка авторизации в озере данных Cloudera Hadoop

- Установка и настройка политик **Apache Sentry**
- Настройка мапирования групп **Ldap** для **Hadoop** аутенитификации
- Настройка **Apache Sentry и Cloudera Navigator** для авторизации компонент экосистемы **Hadoop** с использованием протокола **Kerberos**
- Настройка политик **Apache Sentry** для разграничения полномочий доступа
  - Управление пользователями в **Cloudera Manager**
  - **RBAC** - ролевые политики для разграничения доступа
  - Строковая фильтрация для разграничения доступа

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

- Фильтр на колонки для разграничения доступа
- **Best Practices** для политик разграничения полномочий

5. Защита данных HDFS

- Шифрование данных при передаче (**Data @ Wire encryption**)
  - Настройка **TLS** шифрование для подключения к **Web UI** компонентам экосистемы **Hadoop** в **Cloudera Manager**
  - Поддержка протокола **SPNEGO**
  - **Best Practices** для шифрования трафика
- Шифрование данных на хранении (**DARE**)
  - Настройка **Cloudera Navigator Encrypt** для шифрования данных и метаданных
  - Настройка **Cloudera Navigator Key Trustee** для управления ключами
  - **HDFS** шифрование
  - **Best Practices** для шифрования данных файловой системы
- Управление доступом к **HDFS**
  - **Posix** и **ACL** для **HDFS**
  - Расширенные списки управления доступом в **HDFS**
  - Маскирование персональных данных
  - Шифрование паролей
  - **Best Practices** для управления списками доступа для файловой системы
- Антивирусная защита в озере данных

6. Настройка политик аудита с Cloudera Navigator

- Использование **Cloudera Navigator** для аудита событий в кластере
- Включение аудита для **Cloudera** кластера
- **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)

• Практические занятия с меткой (**опционально**) выполняются по желанию и при наличии свободного времени у слушателей

Скачать программу курса **«Безопасность кластера Cloudera Hadoop»** в формате pdf

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