Arenadata™ Hadoop

Версия - v1.4

Hadoop Security Guide

2018 Arenadata™

Оглавление

1 Функции безопасности Hadoop **2**
1.1 Функции безопасности Hadoop 2

Глава 1

Функции безопасности Hadoop

В документе приведены функции безопасности Hadoop, такие как безопасность периметра, аутентификация, авторизация, отчетность и защита данных.

Документ может быть полезен администраторам, программистам, разработчикам и сотрудникам подразделений информационных технологий, осуществляющих внедрение кластера.

Important: Контактная информация службы поддержки – e-mail: info@arenadata.io

1.1 Функции безопасности Hadoop

Для организаций, которые хранят конфиденциальные данные в экосистеме **Hadoop**, например, запатентованные или персональные данные, которые подлежат соблюдению нормативных требований (**HIPPA**, **PCI**, **DSS**, **FISAM** и т.д.), безопасность очень важна. Многие организации также должны придерживаться строгой внутренней политики безопасности.

Arenadata Hadoop (далее – **ADH**) обеспечивает комплексный подход к безопасности в следующих ключевых областях: безопасность периметра, аутентификация, авторизация, отчетность, защита данных.

1.1.1 Безопасность периметра

ADH позволяет изолировать кластер **Hadoop** с помощью шлюза и правильно настроенных правил брандмауэра. **ADH** поддерживает следующую защиту периметра:

- Apache Knox Gateway;
- Клиенты шлюза.

1.1.2 Аутентификация

ADH предоставляет единую точку аутентификации для сервисов и пользователей, которая интегрируется с существующими системами идентификации предприятия и доступа. **ADH** поддерживает следующие службы аутентификации:

- Kerberos;
- LDAP;

- Локальная система Unix;
- SSO (по периметру через Apache Knox Gateway).

1.1.3 Авторизация (контроль доступа)

ADH предоставляет функции, позволяющие системным администраторам контролировать доступ к данным **Hadoop** с использованием авторизации на основе ролей. **ADH** поддерживает следующие модели авторизации:

- Высококачественное управление доступом для данных, хранящихся в HDFS;
- Контроль доступа на уровне ресурсов для YARN;
- Контроль доступа на уровне крупнозернистого уровня для операций MapReduce;
- Контроль доступа к семейству на уровне таблицы и столбца для данных HBase;
- Контроль доступа на уровне таблицы для наборов данных Apache Hive.

1.1.4 Отчетность (аудит и мониторинг безопасности)

ADH позволяет отслеживать активность **Hadoop**, используя собственный аудит (журналы аудита), журналы аудита безопасности периметра на шлюзе **Knox** и центральное расположение консоли администрирования **ADH**, включая:

- Запросы доступа;
- Операции обработки данных;
- Изменение данных.

1.1.5 Защита данных

ADH предоставляет механизмы для шифрования данных при их передаче и требует использования партнерских решений для шифрования данных в состоянии покоя, их поиска и маскирования. **ADH** поддерживает следующие методы шифрования:

- SSL для компонентов ADH;
- Шифрование RPC;
- Протокол передачи данных.