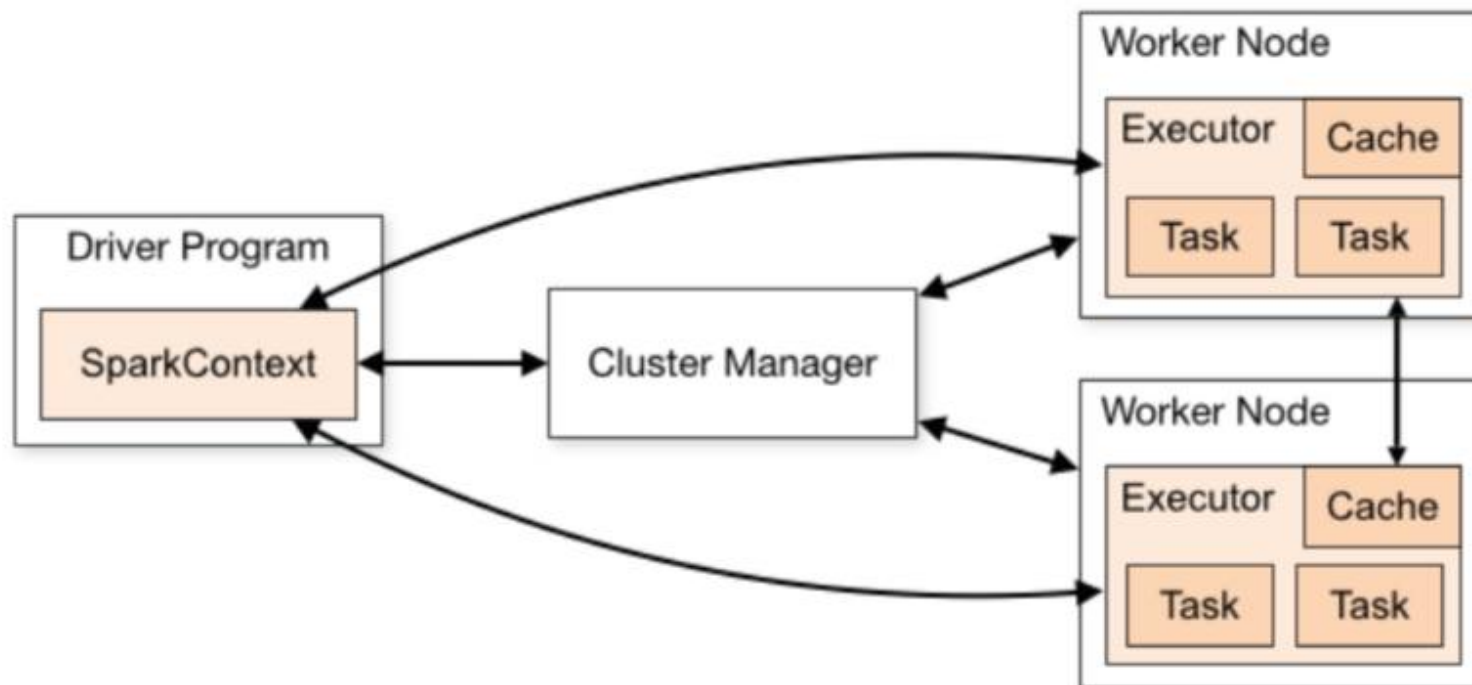
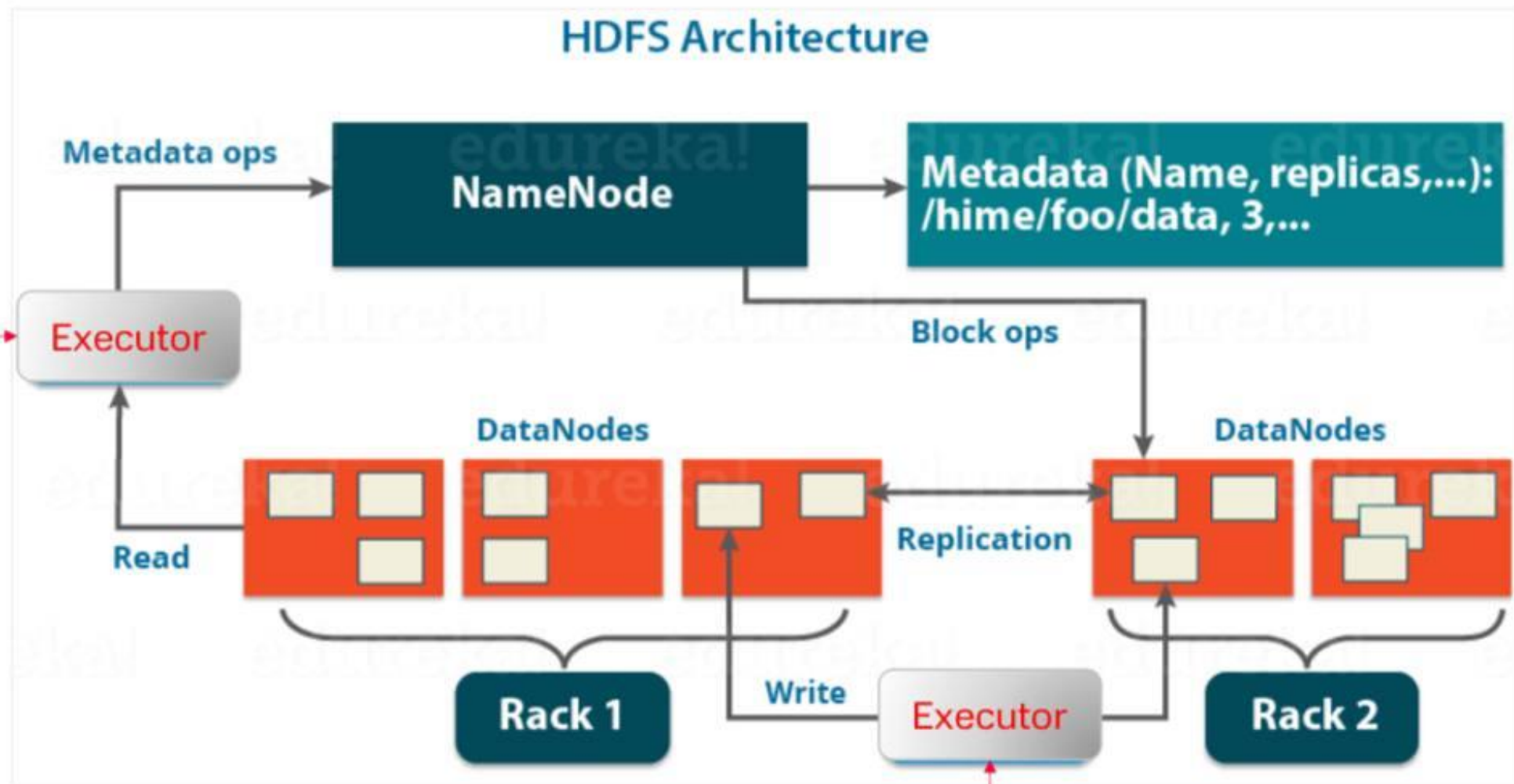


# Архитектура Spark



# Место Spark в экосистеме Hadoop



# Выделение ресурсов

- spark.app.name – **имя** приложения (видно в UI YARN)
- spark.driver.memory – объем **оперативной памяти** для **драйвера**
- spark.driver.maxResultSize – максимальный размер данных, которые могут быть получены на драйвер (например, с помощью команды **collect**)
- spark.executor.memory – объем **оперативной памяти** для **эксекютора**
- spark.local.dir – директория для хранения **временных файлов**
- spark.log.level – управление **логирование** (DEBUG, WARNING, INFO, ERROR)
- spark.driver.extraJavaOptions – параметры **JVM для драйвера**
- spark.executor.extraJavaOptions – параметры **JVM для эксекютора**
- spark.files – **файлы**, которые нужно **отправить на эксекюторы**
- spark.jars.packages – **java-зависимости** для их установки (например, jdbc-коннектор к БД)