Scala, Eclipse IDE and Spark application

Environment *Cloudera Quickstart, single node*

Preparing the environment

Download and install sbt

```
wget http://dl.bintray.com/sbt/rpm/sbt-0.13.0.rpm
```

```
sudo yum localinstall sbt-0.13.0.rpm
```

Check sbt version:

```
[root@quickstart new_proj]# sbt --version
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
sbt launcher version 0.13.0
[root@quickstart new_proj]#
```

Project implementation

1) Create folder for working space – In this case, *new_folder* is the root folder of the project:

mkdir -p /home/cloudera/new_folder

Under project's root folder, create *plugins.sbt* file and move it under *project* folder.

Add eclipse plugin as following:

```
[root@quickstart new_proj]#
[root@quickstart new_proj]# pwd
/home/cloudera/new_proj
[root@quickstart new_proj]# more  project/plugins.sbt
addSbtPlugin("com.typesafe.sbteclipse" % "sbteclipse-plugin" % "2.5.0")
[root@quickstart new_proj]#
[root@quickstart new_proj]#
[root@quickstart new_proj]#
```

2)Create folder src/main/scala, where you will add scala program Hello.scala

```
[root@quickstart new_proj]# more src/main/scala/Hello.scala
package foo.bar.baz

object Main extends App {
        println("Hiya, Gump!")
}
[root@quickstart new_proj]#
[root@quickstart new_proj]#
```

3)Under root folder of the project, create *build.sbt*, and add the name (this will be used for the creation of the jar), version, ScalaVersion and spark dependencies:

```
[root@quickstart new_proj]# pwd
/home/cloudera/new_proj
[root@quickstart new_proj]#
[root@quickstart new_proj]# more /home/cloudera/new_proj/build.sbt
name := "simple-spark-scala"

version := "1.0"

scalaVersion := "2.11.8"

mainClass in (Compile, packageBin) := Some("foo.bar.baz")

libraryDependencies += "org.apache.spark" % "spark-core_2.11" % "1.6.2"
[root@quickstart new_proj]#
```

4) Run “sbt” and wait for a few minutes until the project is created

5) After you get at the “>” prompt, type as following: reload, compile, run. If no issues, program should compile successfully:

```
[root@quickstart new_proj]# sbt
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
>
> reload
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
>
> compile
[success] Total time: 0 s, completed Mar 10, 2017 6:30:21 AM
>
> run
[info] Running foo.bar.baz.Main
Hiya, Gump!
[success] Total time: 0 s, completed Mar 10, 2017 6:30:24 AM
>
>
>
```

6) Type “exit” and proceed with jar creation, by using command “sbt package” Under target/scala-2.11 you should be able to notice the newly created jar:

```
[root@quickstart new_proj]#
[root@quickstart new_proj]# sbt package
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
[success] Total time: 1 s, completed Mar 10, 2017 6:31:21 AM
[root@quickstart new_proj]# ls target/
resolution-cache  scala-2.11  streams
[root@quickstart new_proj]# ls target/scala-2.11/
classes  simple-spark-scala_2.11-1.0.jar
[root@quickstart new_proj]#
```

7) Test eclipse, and run “sbt eclipse” - if success, the Eclipse project is created for your project:

```
[root@quickstart new_proj]# sbt eclipse
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
[info] About to create Eclipse project files for your project(s).
[info] Successfully created Eclipse project files for project(s):
[info] simple-spark-scala
[root@quickstart new_proj]#
```

You can also use eclipse from sbt prompt:

```
[root@quickstart new_proj]# sbt
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
> eclipse
[info] About to create Eclipse project files for your project(s).
[info] Successfully created Eclipse project files for project(s):
[info] simple-spark-scala
>
>
```

8) Now, that we have the jar created, we can start the compiling under Scala. This will be implemented with command **sbt "run-main <main class>"**

```
[root@quickstart new_proj]# sbt run
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
[info] Running foo.bar.baz.Main
Hiya, Gump!
[success] Total time: 1 s, completed Mar 10, 2017 5:14:16 AM
[root@quickstart new_proj]# sbt "run-main foo.bar.baz.Main"
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
[info] Running foo.bar.baz.Main
Hiya, Gump!
[success] Total time: 1 s, completed Mar 10, 2017 5:14:47 AM
[root@quickstart new proj]#
```

9) If success, proceed with running spark application:

```
[root@quickstart new_proj]#
[root@quickstart new_proj]# spark-submit  --class "foo.bar.baz.Main"  \
> --master yarn  \
> --executor-memory 512m \
> --total-executor-cores 1  \
> /home/cloudera/new_proj/target/scala-2.11/simple-spark-scala_2.11-1.0.jar
WARNING: User-defined SPARK_HOME (/opt/cloudera/parcels/CDH-5.8.0-1.cdh5.8.0.p0.42/lib/spark) overrides detected (/usr/lib/spark).
WARNING: Running spark-class from user-defined location.
Hiya, Gump!
[root@quickstart new_proj]#
[root@quickstart new_proj]#
```

As a side note, **smc console** will invoke scala interpretor

```
[root@quickstart new_proj]#
[root@quickstart new_proj]# sbt console
Loading /usr/share/sbt/bin/sbt-launch-lib.bash
[info] Loading project definition from /home/cloudera/new_proj/project
[info] Set current project to simple-spark-scala (in build file:/home/cloudera/new_proj/)
[info] Starting scala interpreter...
[info]
Welcome to Scala 2.11.8 (Java HotSpot(TM) 64-Bit Server VM, Java 1.7.0_67).
Type in expressions for evaluation. Or try :help.

scala>
```