Java-Success.com

Prepare to fast-track, choose & go places with 800+ Java & Big Data Q&As with lots of code & diagrams.

| search here … | Go |

Home   **Why?** ▼   **300+ Java FAQs** ▼   **300+ Big Data FAQs** ▼   **Courses** ▼

👤 **Membership** ▼   **Your Career** ▼

03. Setting up Scala and Spark with Maven

📅 Posted on October 26, 2016

This assumes that you set up Scala IDE for eclipse as per Setting up Scala IDE & using Maven plugins.

Step 1: Create a Scala project in Maven.

```
1
2  mvn archetype:generate "-DarchetypeGroupId=net.alcl
3
```

Step 2: Import Maven project into eclipse.

File -> Existing Maven Projects -> browse and select "spark-and-scala"

300+ Java Interview FAQs

300+ Java FAQs 🔥   ⌄

16+ Java Key Areas Q&As   ⌄

150+ Java Architect FAQs   ⌄

80+ Java Code Quality Q&As ⌄

150+ Java Coding Q&As   ⌄

300+ Big Data Interview FAQs

300+ Big Data FAQs 🔥   ⌄

Tutorials - Big Data   ›

  TUT - 🔢 Starting Big Data
  TUT - Starting Spark & Scala

Step 3: Open the **pom.xml** and make modification to the "scala-maven-plugin" so that it looks as shown below.

```
1
2  <project xmlns="http://maven.apache.org/POM/4.0.0"
3      xsi:schemaLocation="http://maven.apache.org/P(
4      <modelVersion>4.0.0</modelVersion>
5      <groupId>com.mytutorial</groupId>
6      <artifactId>spark-and-scala</artifactId>
7      <version>1.0-SNAPSHOT</version>
8      <name>${project.artifactId}</name>
9      <description>My wonderfull scala app</descript
10     <inceptionYear>2015</inceptionYear>
11     <licenses>
12         <license>
13             <name>My License</name>
14             <url>http://....</url>
15             <distribution>repo</distribution>
16         </license>
17     </licenses>
18
19     <properties>
20         <maven.compiler.source>1.6</maven.compiler
21         <maven.compiler.target>1.6</maven.compiler
22         <encoding>UTF-8</encoding>
23         <scala.version>2.11.5</scala.version>
24         <scala.compat.version>2.11</scala.compat.v
25     </properties>
26
27     <dependencies>
28         <dependency>
29             <groupId>org.scala-lang</groupId>
30             <artifactId>scala-library</artifactId>
31             <version>${scala.version}</version>
32         </dependency>
33
34         <!-- Test -->
35         <dependency>
36             <groupId>junit</groupId>
37             <artifactId>junit</artifactId>
38             <version>4.11</version>
39             <scope>test</scope>
40         </dependency>
41         <dependency>
42             <groupId>org.specs2</groupId>
43             <artifactId>specs2-core_${scala.compat
44             <version>2.4.16</version>
45             <scope>test</scope>
46         </dependency>
47         <dependency>
48             <groupId>org.scalatest</groupId>
```

800+ Java Interview Q&As

300+ Core Java Q&As ⌄

300+ Enterprise Java Q&As ⌄

150+ Java Frameworks Q&As ⌄

120+ Companion Tech Q&As ⌄

Tutorials - Enterprise Java ⌄

```
49              <artifactId>scalatest_${scala.compat.v
50              <version>2.2.4</version>
51              <scope>test</scope>
52          </dependency>
53      </dependencies>
54
55      <build>
56          <sourceDirectory>src/main/scala</sourceDir
57          <testSourceDirectory>src/test/scala</test:
58          <plugins>
59              <plugin>
60                  <!-- see http://davidb.github.com/
61                  <groupId>net.alchim31.maven</grou{
62                  <artifactId>scala-maven-plugin</ar
63                  <version>3.2.0</version>
64                  <executions>
65                      <execution>
66                          <id>scala-compile-first</
67                          <phase>process-resources</
68                          <goals>
69                              <goal>compile</goal>
70                          </goals>
71                      </execution>
72                  </executions>
73              </plugin>
74              <plugin>
75                  <groupId>org.apache.maven.plugins
76                  <artifactId>maven-surefire-plugin
77                  <version>2.18.1</version>
78                  <configuration>
79                      <useFile>false</useFile>
80                      <disableXmlReport>true</disab
81                      <!-- If you have classpath is
82                      <!-- useManifestOnlyJar>false-
83                      <includes>
84                          <include>**/*Test.*</incl
85                          <include>**/*Suite.*</inc
86                      </includes>
87                  </configuration>
88              </plugin>
89          </plugins>
90      </build>
91 </project>
92
```

Step 4: You can now run the following maven
command to build a package.

```
1
2 C:\users\xx\projects\spark-and-scala>mvn clean pacl
3
```

Step 5: Lets add the Spark library dependencies.

```xml
1
2   <project xmlns="http://maven.apache.org/POM/4.0.(
3       xsi:schemaLocation="http://maven.apache.org/l
4       <modelVersion>4.0.0</modelVersion>
5       <groupId>com.mytutorial</groupId>
6       <artifactId>spark-and-scala</artifactId>
7       <version>1.0-SNAPSHOT</version>
8       <name>${project.artifactId}</name>
9       <description>My wonderfull scala app</descri[
10      <inceptionYear>2015</inceptionYear>
11      <licenses>
12          <license>
13              <name>My License</name>
14              <url>http://....</url>
15              <distribution>repo</distribution>
16          </license>
17      </licenses>
18
19      <properties>
20          <maven.compiler.source>1.6</maven.compil
21          <maven.compiler.target>1.6</maven.compil
22          <encoding>UTF-8</encoding>
23          <scala.version>2.11.6</scala.version>
24          <scala.compat.version>2.11</scala.compat
25          <spark.version>1.3.0</spark.version>
26      </properties>
27
28      <dependencies>
29          <!-- Scala -->
30          <dependency>
31              <groupId>org.scala-lang</groupId>
32              <artifactId>scala-library</artifactI
33              <version>${scala.version}</version>
34          </dependency>
35
36          <!-- Spark -->
37          <dependency>
38              <groupId>org.apache.spark</groupId>
39              <artifactId>spark-core_${scala.compa
40              <version>${spark.version}</version>
41              <scope>provided</scope>
42          </dependency>
43
44          <!-- Test -->
45          <dependency>
46              <groupId>junit</groupId>
47              <artifactId>junit</artifactId>
48              <version>4.11</version>
49              <scope>test</scope>
50          </dependency>
51          <dependency>
```

```
52              <groupId>org.specs2</groupId>
53              <artifactId>specs2-core_${scala.comp
54              <version>2.4.16</version>
55              <scope>test</scope>
56          </dependency>
57          <dependency>
58              <groupId>org.scalatest</groupId>
59              <artifactId>scalatest_${scala.compat
60              <version>2.2.4</version>
61              <scope>test</scope>
62          </dependency>
63      </dependencies>
64
65      <build>
66          <sourceDirectory>src/main/scala</sourceD
67          <testSourceDirectory>src/test/scala</tes
68          <plugins>
69              <plugin>
70                  <!-- see http://davidb.github.co
71                  <groupId>net.alchim31.maven</grou
72                  <artifactId>scala-maven-plugin</
73                  <version>3.2.0</version>
74                  <executions>
75                      <execution>
76                          <id>scala-compile-first<
77                          <phase>process-resources-
78                          <goals>
79                              <goal>compile</goal>
80                          </goals>
81                      </execution>
82                  </executions>
83              </plugin>
84              <plugin>
85                  <groupId>org.apache.maven.plugin
86                  <artifactId>maven-surefire-plugi
87                  <version>2.18.1</version>
88                  <configuration>
89                      <useFile>false</useFile>
90                      <disableXmlReport>true</disa
91                      <!-- If you have classpath i
92                      <!-- useManifestOnlyJar>fals
93                      <includes>
94                          <include>**/*Test.*</inc
95                          <include>**/*Suite.*</in
96                      </includes>
97                  </configuration>
98              </plugin>
99          </plugins>
100     </build>
101 </project>
102
```

Step 6: Create a simple Spark job.

```scala
1
2  package com.mytutorial
3
4  import org.apache.spark.SparkConf
5  import org.apache.spark.SparkContext
6
7  object SimpleSparkInScala {
8
9
10   def main(args: Array[String]) : Unit = {
11     val conf = new SparkConf()
12               .setAppName("Simple does nothing")
13               .setMaster("local")
14
15     val sc = new SparkContext(conf)
16
17     val list = List(1,2)
18
19     val simpleRDD = sc.parallelize(list)
20
21     simpleRDD.foreach { x => println(x) }
22
23   }
24 }
25
```

Output:

```
1
2  Using Spark's default log4j profile: org/apache/sp
3  16/10/25 22:59:40 INFO SparkContext: Running Spar
4  16/10/25 22:59:45 WARN NativeCodeLoader: Unable t
5  16/10/25 22:59:45 INFO SecurityManager: Changing
6  16/10/25 22:59:45 INFO SecurityManager: Changing
7  16/10/25 22:59:46 INFO SecurityManager: SecurityM
8  16/10/25 22:59:46 INFO Slf4jLogger: Slf4jLogger s
9  16/10/25 22:59:46 INFO Remoting: Starting remoting
10 16/10/25 22:59:48 INFO Remoting: Remoting started
11 16/10/25 22:59:48 INFO Utils: Successfully starte
12 16/10/25 22:59:48 INFO SparkEnv: Registering MapOu
13 16/10/25 22:59:48 INFO SparkEnv: Registering Bloc
14 16/10/25 22:59:48 INFO DiskBlockManager: Created
15 16/10/25 22:59:48 INFO MemoryStore: MemoryStore s
16 16/10/25 22:59:48 INFO HttpFileServer: HTTP File
17 16/10/25 22:59:48 INFO HttpServer: Starting HTTP
18 16/10/25 22:59:48 INFO Server: jetty-8.y.z-SNAPSH
19 16/10/25 22:59:48 INFO AbstractConnector: Started
20 16/10/25 22:59:48 INFO Utils: Successfully starte
21 16/10/25 22:59:48 INFO SparkEnv: Registering Outpu
22 16/10/25 22:59:48 INFO Server: jetty-8.y.z-SNAPSH
```

```
23  16/10/25 22:59:48 INFO AbstractConnector: Started
24  16/10/25 22:59:48 INFO Utils: Successfully started
25  16/10/25 22:59:48 INFO SparkUI: Started SparkUI at
26  16/10/25 22:59:48 INFO Executor: Starting executor
27  16/10/25 22:59:49 INFO AkkaUtils: Connecting to He
28  16/10/25 22:59:49 INFO NettyBlockTransferService:
29  16/10/25 22:59:49 INFO BlockManagerMaster: Trying
30  16/10/25 22:59:49 INFO BlockManagerMasterActor: Re
31  16/10/25 22:59:49 INFO BlockManagerMaster: Registe
32  16/10/25 22:59:50 INFO SparkContext: Starting job
33  16/10/25 22:59:50 INFO DAGScheduler: Got job 0 (fc
34  16/10/25 22:59:50 INFO DAGScheduler: Final stage:
35  16/10/25 22:59:50 INFO DAGScheduler: Parents of f-
36  16/10/25 22:59:50 INFO DAGScheduler: Missing parer
37  16/10/25 22:59:50 INFO DAGScheduler: Submitting S
38  16/10/25 22:59:53 INFO MemoryStore: ensureFreeSpa
39  16/10/25 22:59:53 INFO MemoryStore: Block broadca
40  16/10/25 22:59:53 INFO MemoryStore: ensureFreeSpa
41  16/10/25 22:59:53 INFO MemoryStore: Block broadca
42  16/10/25 22:59:53 INFO BlockManagerInfo: Added br
43  16/10/25 22:59:53 INFO BlockManagerMaster: Updated
44  16/10/25 22:59:53 INFO SparkContext: Created broad
45  16/10/25 22:59:53 INFO DAGScheduler: Submitting 1
46  16/10/25 22:59:53 INFO TaskSchedulerImpl: Adding
47  16/10/25 22:59:53 INFO TaskSetManager: Starting ta
48  16/10/25 22:59:53 INFO Executor: Running task 0.0
49  1
50  2
51  16/10/25 22:59:53 INFO Executor: Finished task 0.(
52  16/10/25 22:59:53 INFO TaskSetManager: Finished ta
53  16/10/25 22:59:53 INFO TaskSchedulerImpl: Removed
54  16/10/25 22:59:53 INFO DAGScheduler: Stage 0 (fore
55  16/10/25 22:59:53 INFO DAGScheduler: Job 0 finishe
56
```

Disclaimer

© 2022 java-success.com

Top