## Installing an Add-on Service

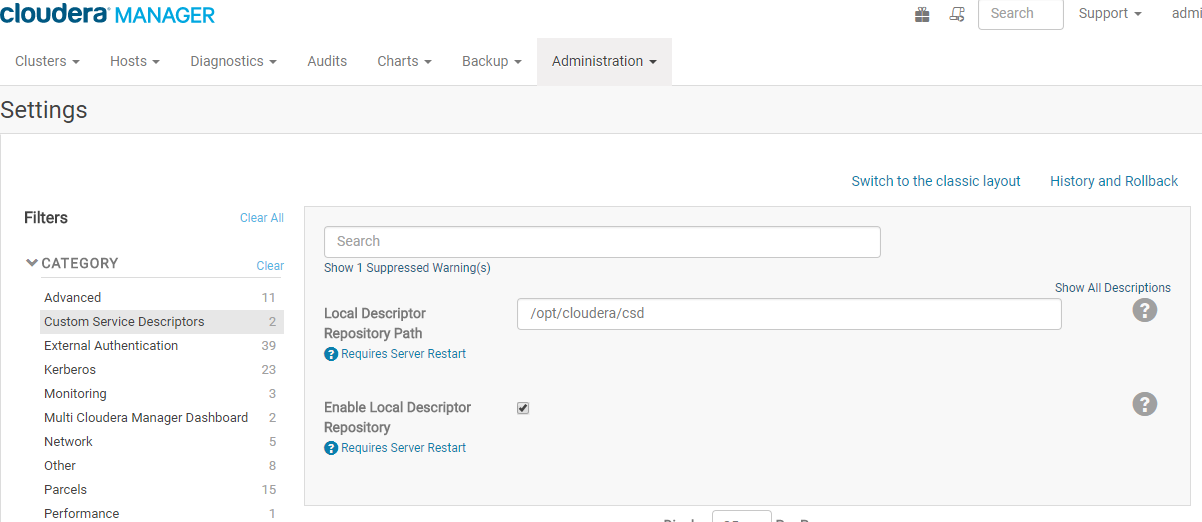
To install any Custom service we need to first download CSD file and placed under Local Descriptor Repository path, here we are trying to install spark2 on Cloudera 5.11

First Download [Spark2 CSD](https://www.cloudera.com/documentation/spark2/latest/topics/spark2_packaging.html#packaging) file

Now you need to place that file under **/opt/Cloudera/csd** folder

The default location for CSD files is /opt/cloudera/csd. You can change the location in the Cloudera Manager Admin Console as follows:

1. Select **Administration** > **Settings**.
2. Click the **Custom Service Descriptors** category.
3. Edit the **Local Descriptor Repository Path** property.
4. Click **Save Changes** to commit the changes.



Set the file ownership to cloudera-scm:cloudera-scm with permission 644.

|  |
| --- |
| [shalaj@mac127 ~]$ cd /opt/cloudera/csd/  [shalaj@mac127 csd]$ ll  total 16  ----r-x--- 1 root root 16109 Jun 30 11:22 SPARK2\_ON\_YARN-2.1.0.cloudera1.jar  [shalaj@mac127 csd]$ exit  logout  You have new mail in /var/spool/mail/root  [root@mac127 ~]# cd /opt/cloudera/csd/  [root@mac127 csd]# chown cloudera-scm:cloudera-scm SPARK2\_ON\_YARN-2.1.0.cloudera1.jar  [root@mac127 csd]# chmod 644 SPARK2\_ON\_YARN-2.1.0.cloudera1.jar |

Restart the Cloudera Manager Server:

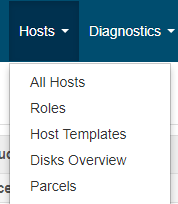
service cloudera-scm-server restart

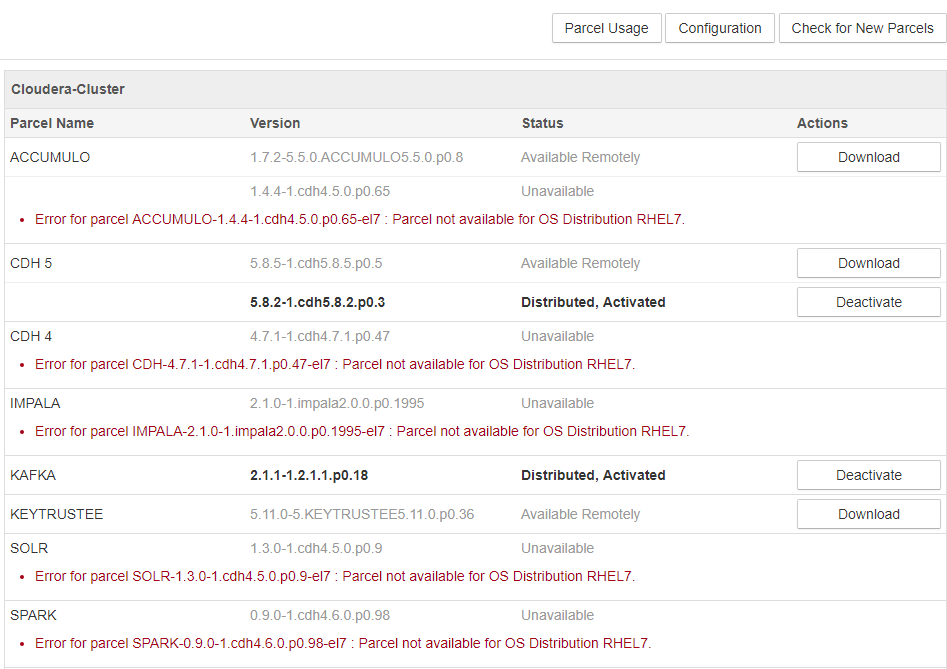
Log into the Cloudera Manager Admin Console and restart the Cloudera Management Service.

* Select **Clusters** > **Cloudera Management Service**.
* Select **Actions** > **Restart**.

**Installing the Parcel**

Go to host >> Parcel

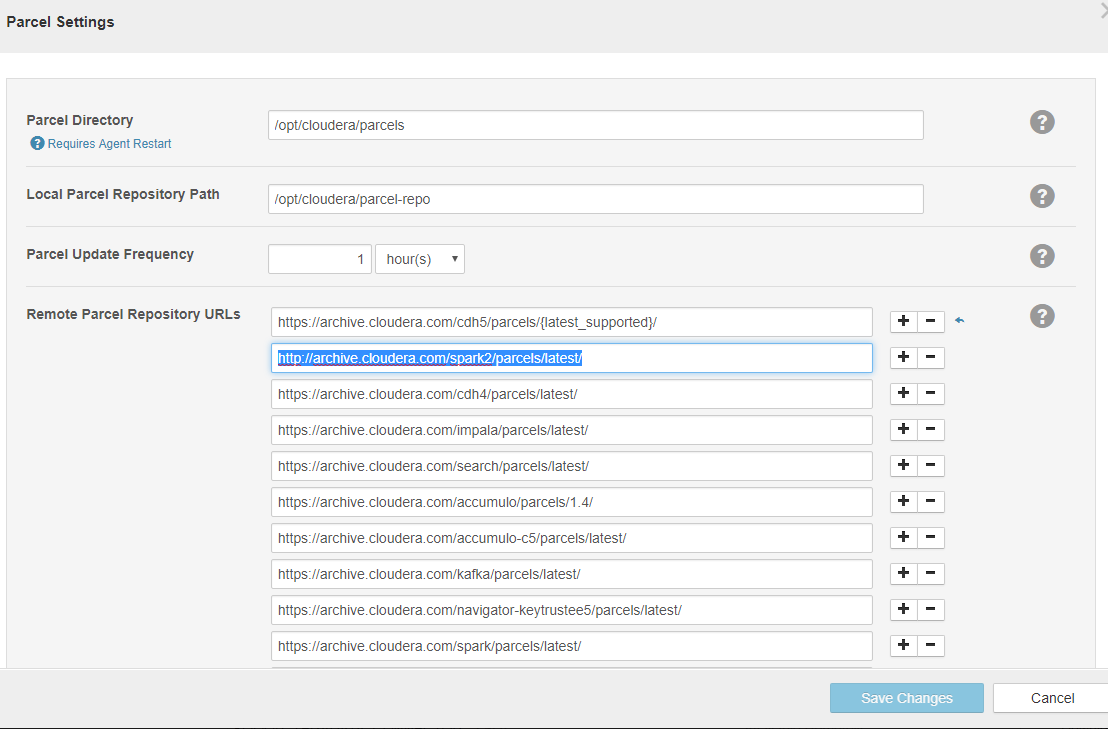




Here I want to add Spark 2.1 version

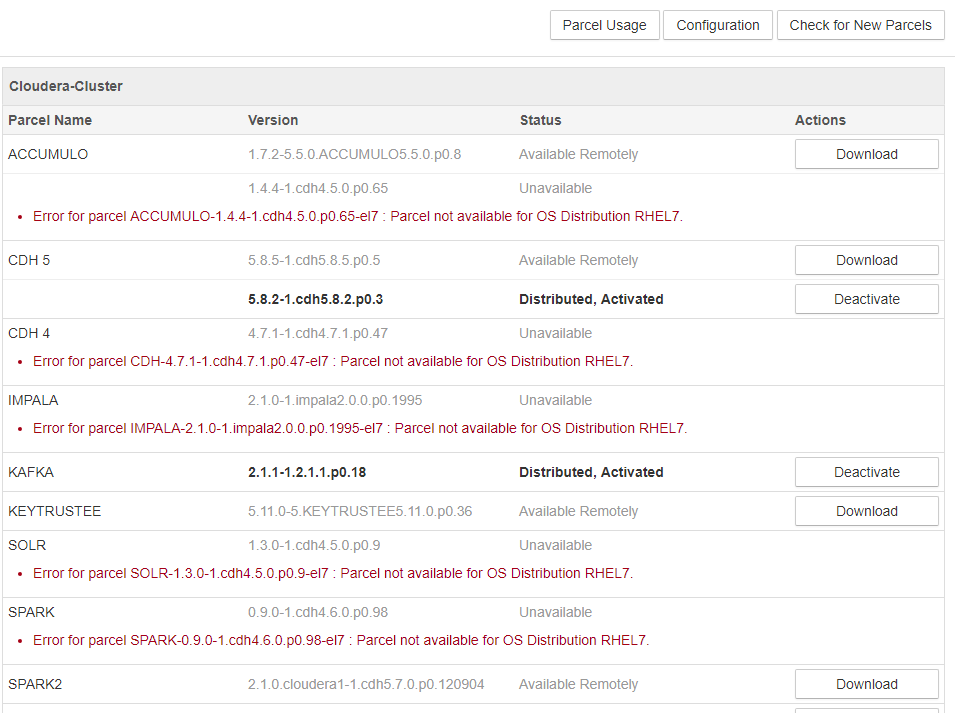
Click Configuration button

Search parcel for spark2 and add it to Remote Parcel Repository URLS



Add <http://archive.cloudera.com/spark2/parcels/latest/> and Click Save changes

Now click **Check for New Parcels** button



Now you will find SPARK2 service

Click Download

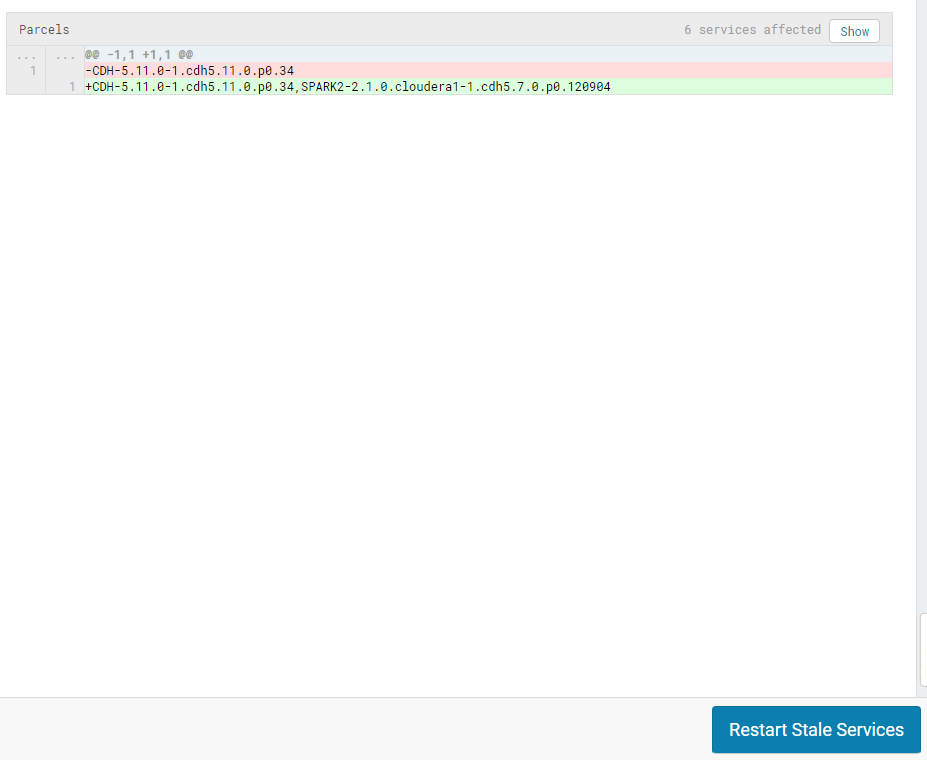


Click Distribute



Click Activate

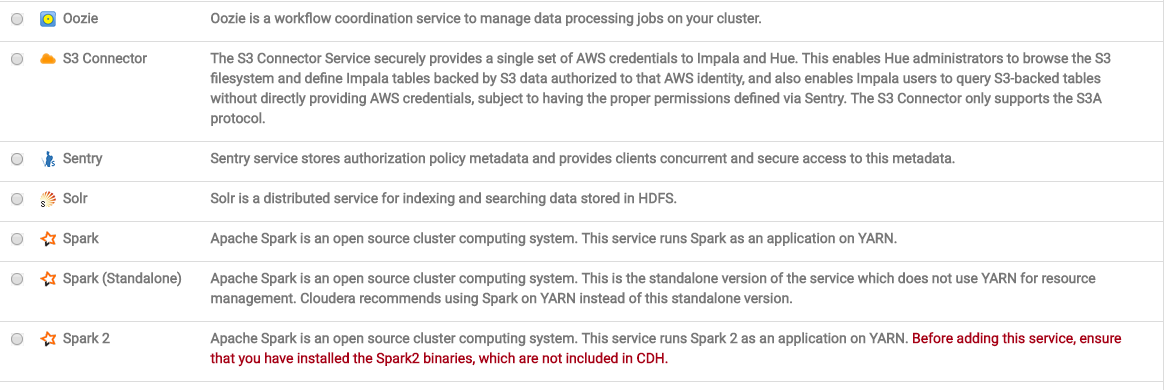
Now restart stale services



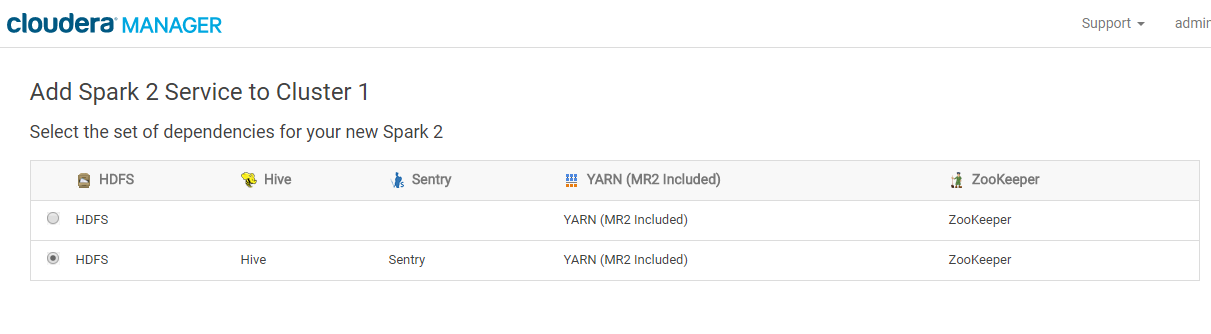
Now you can add new service

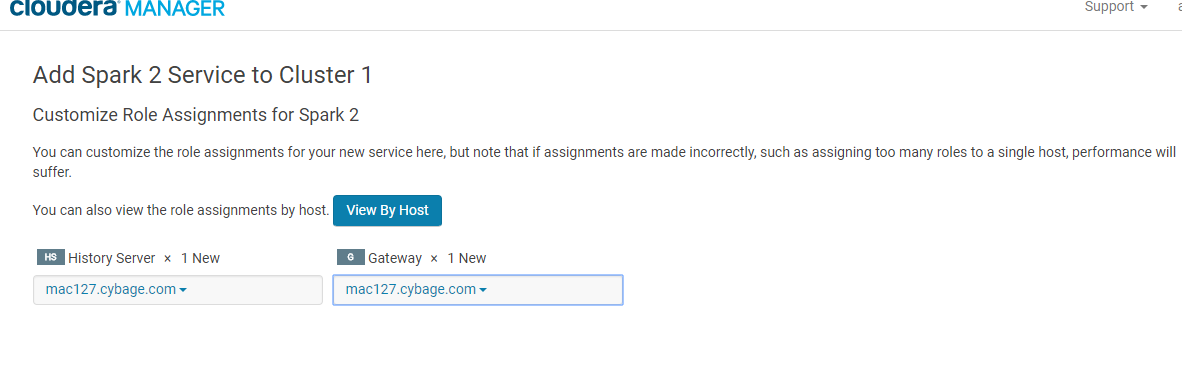
Click Add Service

Now you will find spark2



Now select Spark2 and





Now run spark2-shell from command shell

|  |
| --- |
| [shalaj@mac127 ~]$ spark2-shell  Setting default log level to "WARN".  To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).  17/06/30 12:02:30 WARN spark.SparkContext: Support for Java 7 is deprecated as of Spark 2.0.0  Spark context Web UI available at http://172.27.155.127:4040  Spark context available as 'sc' (master = yarn, app id = application\_1498716529422\_0005).  Spark session available as 'spark'.  Welcome to  \_\_\_\_ \_\_  / \_\_/\_\_ \_\_\_ \_\_\_\_\_/ /\_\_  \_\ \/ \_ \/ \_ `/ \_\_/ '\_/  /\_\_\_/ .\_\_/\\_,\_/\_/ /\_/\\_\ version 2.1.0.cloudera1  /\_/  Using Scala version 2.11.8 (Java HotSpot(TM) 64-Bit Server VM, Java 1.7.0\_67)  Type in expressions to have them evaluated.  Type :help for more information. |

Spark2 used scala version 2.11 so we need to use the same version while creating spark program

Refer below link

<https://www.cloudera.com/documentation/spark2/latest/topics/spark2_installing.html>

pom file for spark2

|  |
| --- |
| <project xmlns=*"http://maven.apache.org/POM/4.0.0"* xmlns:xsi=*"http://www.w3.org/2001/XMLSchema-instance"* xsi:schemaLocation=*"http://maven.apache.org/POM/4.0.0 http://maven.apache.org/xsd/maven-4.0.0.xsd"*>  <modelVersion>4.0.0</modelVersion>  <groupId>org.cyb.spark</groupId>  <artifactId>Spark211</artifactId>  <version>0.0.1-SNAPSHOT</version>  <dependencies>  <dependency>  <!-- Spark dependency -->  <groupId>org.apache.spark</groupId>  <artifactId>spark-core\_2.11</artifactId>  <version>2.1.0</version>  </dependency>  <dependency>  <groupId>org.apache.spark</groupId>  <artifactId>spark-sql\_2.11</artifactId>  <version>2.1.0</version>  </dependency>  <dependency>  <groupId>com.databricks</groupId>  <artifactId>spark-csv\_2.11</artifactId>  <version>1.2.0</version>  </dependency>    <dependency>  <groupId>org.apache.spark</groupId>  <artifactId>spark-streaming-kafka\_2.11</artifactId>  <version>1.6.0</version>  </dependency>  <dependency>  <groupId>org.apache.spark</groupId>  <artifactId>spark-streaming\_2.11</artifactId>  <version>2.1.0</version>  <scope>provided</scope>  </dependency>    <dependency>  <groupId>org.apache.spark</groupId>  <artifactId>spark-hive\_2.11</artifactId>  <version>2.1.0</version>  </dependency>  <!--  <dependency>  <groupId>org.apache.spark</groupId>  <artifactId>spark-streaming\_2.11</artifactId>  <version>1.6.0</version>  <scope>provided</scope>  </dependency>  <dependency>  <groupId>org.apache.spark</groupId>  <artifactId>spark-streaming-twitter\_2.11</artifactId>  <version>1.6.0</version>  </dependency>  <dependency>  <groupId>org.twitter4j</groupId>  <artifactId>twitter4j-stream</artifactId>  <version>3.0.3</version>  </dependency>  <dependency>  <groupId>mysql</groupId>  <artifactId>mysql-connector-java</artifactId>  <version>5.1.31</version>  </dependency>  <dependency>  <groupId>au.com.bytecode</groupId>  <artifactId>opencsv</artifactId>  <version>2.4</version>  </dependency>  -->  </dependencies>  </project> |

Also we need to set scala compiler to 2.11

