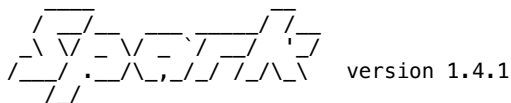


```

[hdfs@sandbox ~]$ spark-shell --packages com.databricks:spark-csv_2.10:1.1.0 --master yarn-
client --driver-memory 512m --executor-memory 512m
Ivy Default Cache set to: /home/hdfs/.ivy2/cache
The jars for the packages stored in: /home/hdfs/.ivy2/jars
:: loading settings :: url = jar:file:/usr/hdp/2.3.2.0-2950/spark/lib/spark-
assembly-1.4.1.2.3.2.0-2950-hadoop2.7.1.2.3.2.0-2950.jar!/org/apache/ivy/core/settings/
ivysettings.xml
com.databricks#spark-csv_2.10 added as a dependency
:: resolving dependencies :: org.apache.spark#spark-submit-parent;1.0
  confs: [default]
  found com.databricks#spark-csv_2.10;1.1.0 in central
  found org.apache.commons#commons-csv;1.1 in central
  found com.univocity#univocity-parsers;1.5.1 in central
downloading https://repo1.maven.org/maven2/com/databricks/spark-csv_2.10/1.1.0/spark-
csv_2.10-1.1.0.jar ...
[SUCCESSFUL ] com.databricks#spark-csv_2.10;1.1.0!spark-csv_2.10.jar (97ms)
downloading https://repo1.maven.org/maven2/org/apache/commons/commons-csv/1.1/commons-
csv-1.1.jar ...
[SUCCESSFUL ] org.apache.commons#commons-csv;1.1!commons-csv.jar (91ms)
downloading https://repo1.maven.org/maven2/com/univocity/univocity-parsers/1.5.1/univocity-
parsers-1.5.1.jar ...
[SUCCESSFUL ] com.univocity#univocity-parsers;1.5.1!univocity-parsers.jar (155ms)
:: resolution report :: resolve 4590ms :: artifacts dl 349ms
  :: modules in use:
  com.databricks#spark-csv_2.10;1.1.0 from central in [default]
  com.univocity#univocity-parsers;1.5.1 from central in [default]
  org.apache.commons#commons-csv;1.1 from central in [default]
-----
|               |               modules               || artifacts |
|               | number| search|dwnlded|evicted|| number|dwnlded|
-----+-----+-----+-----+-----+-----+-----+-----+-----+
|               | 3     | 3     | 3     | 0     || 3     | 3     |
-----+-----+-----+-----+-----+-----+-----+-----+-----+
:: retrieving :: org.apache.spark#spark-submit-parent
  confs: [default]
  3 artifacts copied, 0 already retrieved (265kB/7ms)
15/12/14 03:00:50 WARN SparkConf: The configuration key 'spark.yarn.applicationMaster.waitTries'
has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key
'spark.yarn.am.waitTime' instead.
15/12/14 03:00:50 WARN NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
15/12/14 03:00:50 INFO SecurityManager: Changing view acls to: hdfs
15/12/14 03:00:50 INFO SecurityManager: Changing modify acls to: hdfs
15/12/14 03:00:50 INFO SecurityManager: SecurityManager: authentication disabled; ui acls
disabled; users with view permissions: Set(hdfs); users with modify permissions: Set(hdfs)
15/12/14 03:00:50 INFO HttpServer: Starting HTTP Server
15/12/14 03:00:50 INFO Server: jetty-8.y.z-SNAPSHOT
15/12/14 03:00:50 INFO AbstractConnector: Started SocketConnector@0.0.0.0:41650
15/12/14 03:00:50 INFO Utils: Successfully started service 'HTTP class server' on port 41650.
Welcome to

```



```

Using Scala version 2.10.4 (OpenJDK 64-Bit Server VM, Java 1.7.0_91)
Type in expressions to have them evaluated.
Type :help for more information.
15/12/14 03:00:55 WARN SparkConf: The configuration key 'spark.yarn.applicationMaster.waitTries'
has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key
'spark.yarn.am.waitTime' instead.
15/12/14 03:00:55 INFO SparkContext: Running Spark version 1.4.1
15/12/14 03:00:55 WARN SparkConf: The configuration key 'spark.yarn.applicationMaster.waitTries'
has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key
'spark.yarn.am.waitTime' instead.
15/12/14 03:00:55 INFO SecurityManager: Changing view acls to: hdfs
15/12/14 03:00:55 INFO SecurityManager: Changing modify acls to: hdfs
15/12/14 03:00:55 INFO SecurityManager: SecurityManager: authentication disabled; ui acls
disabled; users with view permissions: Set(hdfs); users with modify permissions: Set(hdfs)
15/12/14 03:00:55 INFO Slf4jLogger: Slf4jLogger started
15/12/14 03:00:55 INFO Remoting: Starting remoting
15/12/14 03:00:55 INFO Remoting: Remoting started; listening on addresses :[akka.tcp://

```



has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key 'spark.yarn.am.waitTime' instead.  
spark.yarn.driver.memoryOverhead is set but does not apply in client mode.  
15/12/14 03:00:56 INFO TimelineClientImpl: Timeline service address: http://sandbox.hortonworks.com:8188/ws/v1/timeline/  
15/12/14 03:00:56 INFO RMPProxy: Connecting to ResourceManager at sandbox.hortonworks.com/10.0.2.15:8050  
15/12/14 03:00:56 INFO Client: Requesting a new application from cluster with 1 NodeManagers  
15/12/14 03:00:56 INFO Client: Verifying our application has not requested more than the maximum memory capability of the cluster (2250 MB per container)  
15/12/14 03:00:56 INFO Client: Will allocate AM container, with 896 MB memory including 384 MB overhead  
15/12/14 03:00:56 INFO Client: Setting up container launch context for our AM  
15/12/14 03:00:56 INFO Client: Preparing resources for our AM container  
15/12/14 03:00:57 WARN DomainSocketFactory: The short-circuit local reads feature cannot be used because libhadoop cannot be loaded.  
15/12/14 03:00:57 INFO Client: Uploading resource file:/usr/hdp/2.3.2.0-2950/spark/lib/spark-assembly-1.4.1.2.3.2.0-2950-hadoop2.7.1.2.3.2.0-2950.jar -> hdfs://sandbox.hortonworks.com:8020/user/hdfs/.sparkStaging/application\_1450062003577\_0002/spark-assembly-1.4.1.2.3.2.0-2950-hadoop2.7.1.2.3.2.0-2950.jar  
15/12/14 03:00:59 INFO Client: Uploading resource file:/tmp/spark-431bf7fe-1863-4a32-9d88-b970d08bb794/\_\_\_hadoop\_conf\_\_7144304655670299658.zip -> hdfs://sandbox.hortonworks.com:8020/user/hdfs/.sparkStaging/application\_1450062003577\_0002/\_\_\_hadoop\_conf\_\_7144304655670299658.zip  
15/12/14 03:00:59 INFO Client: Setting up the launch environment for our AM container  
15/12/14 03:00:59 INFO SecurityManager: Changing view acls to: hdfs  
15/12/14 03:00:59 INFO SecurityManager: Changing modify acls to: hdfs  
15/12/14 03:00:59 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(hdfs); users with modify permissions: Set(hdfs)  
15/12/14 03:00:59 INFO Client: Submitting application 2 to ResourceManager  
15/12/14 03:00:59 INFO YarnClientImpl: Submitted application application\_1450062003577\_0002  
15/12/14 03:00:59 WARN SparkConf: The configuration key 'spark.yarn.applicationMaster.waitTries' has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key 'spark.yarn.am.waitTime' instead.  
15/12/14 03:00:59 INFO YarnExtensionServices: Spark YARN services:  
org.apache.spark.deploy.yarn.history.YarnHistoryService  
15/12/14 03:00:59 INFO TimelineClientImpl: Timeline service address: http://sandbox.hortonworks.com:8188/ws/v1/timeline/  
15/12/14 03:00:59 WARN SparkConf: The configuration key 'spark.yarn.applicationMaster.waitTries' has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key 'spark.yarn.am.waitTime' instead.  
15/12/14 03:00:59 INFO YarnHistoryService: Starting Dequeue service for AppId application\_1450062003577\_0002  
15/12/14 03:00:59 INFO YarnHistoryService: Service History Service in state History Service: STARTED endpoint=http://sandbox.hortonworks.com:8188/ws/v1/timeline/; bonded to ATS=true; listening=true; batchSize=3; flush count=0; current queue size=0; total number queued=0, processed=0; post failures=0;  
15/12/14 03:00:59 INFO YarnExtensionServices: Service  
org.apache.spark.deploy.yarn.history.YarnHistoryService started  
15/12/14 03:01:00 INFO Client: Application report for application\_1450062003577\_0002 (state: ACCEPTED)  
15/12/14 03:01:00 INFO Client:  
client token: N/A  
diagnostics: N/A  
ApplicationMaster host: N/A  
ApplicationMaster RPC port: -1  
queue: default  
start time: 1450062059498  
final status: UNDEFINED  
tracking URL: http://sandbox.hortonworks.com:8088/proxy/application\_1450062003577\_0002/  
user: hdfs  
15/12/14 03:01:01 INFO Client: Application report for application\_1450062003577\_0002 (state: ACCEPTED)  
15/12/14 03:01:02 INFO Client: Application report for application\_1450062003577\_0002 (state: ACCEPTED)  
15/12/14 03:01:03 INFO Client: Application report for application\_1450062003577\_0002 (state: ACCEPTED)  
15/12/14 03:01:04 INFO YarnSchedulerBackends\$YarnSchedulerEndpoint: ApplicationMaster registered as AkkaRpcEndpointRef(Actor[akka.tcp://sparkYarnAM@10.0.2.15:43680/user/YarnAM#1994571821])  
15/12/14 03:01:04 INFO YarnClientSchedulerBackend: Add WebUI Filter.  
org.apache.hadoop.yarn.server.webproxy.amfilter.AmIpFilter, Map(PROXY\_HOSTS -> sandbox.hortonworks.com, PROXY\_URI\_BASES -> http://sandbox.hortonworks.com:8088/proxy/application\_1450062003577\_0002), /proxy/application\_1450062003577\_0002  
15/12/14 03:01:04 INFO JettyUtils: Adding filter:

```
org.apache.hadoop.yarn.server.webproxy.amfilter.AmIpFilter
15/12/14 03:01:04 INFO Client: Application report for application_1450062003577_0002 (state:
RUNNING)
15/12/14 03:01:04 INFO Client:
  client token: N/A
  diagnostics: N/A
  ApplicationMaster host: 10.0.2.15
  ApplicationMaster RPC port: 0
  queue: default
  start time: 1450062059498
  final status: UNDEFINED
  tracking URL: http://sandbox.hortonworks.com:8088/proxy/application_1450062003577_0002/
  user: hdfs
15/12/14 03:01:04 INFO YarnClientSchedulerBackend: Application application_1450062003577_0002 has
started running.
15/12/14 03:01:05 INFO Utils: Successfully started service
'org.apache.spark.network.netty.NettyBlockTransferService' on port 49489.
15/12/14 03:01:05 INFO NettyBlockTransferService: Server created on 49489
15/12/14 03:01:05 INFO BlockManagerMaster: Trying to register BlockManager
15/12/14 03:01:05 INFO BlockManagerMasterEndpoint: Registering block manager 10.0.2.15:49489 with
265.4 MB RAM, BlockManagerId(driver, 10.0.2.15, 49489)
15/12/14 03:01:05 INFO BlockManagerMaster: Registered BlockManager
15/12/14 03:01:12 INFO YarnClientSchedulerBackend: Registered executor:
AkkaRpcEndpointRef(Actor[akka.tcp://sparkExecutor@sandbox.hortonworks.com:49909/user/
Executor#1069864297]) with ID 1
15/12/14 03:01:13 INFO BlockManagerMasterEndpoint: Registering block manager
sandbox.hortonworks.com:60064 with 265.4 MB RAM, BlockManagerId(1, sandbox.hortonworks.com,
60064)
15/12/14 03:01:26 INFO YarnClientSchedulerBackend: SchedulerBackend is ready for scheduling
beginning after waiting maxRegisteredResourcesWaitingTime: 30000(ms)
15/12/14 03:01:26 INFO SparkILoop: Created spark context..
Spark context available as sc.
15/12/14 03:01:26 WARN SparkConf: The configuration key 'spark.yarn.applicationMaster.waitTries'
has been deprecated as of Spark 1.3 and and may be removed in the future. Please use the new key
'spark.yarn.am.waitTime' instead.
15/12/14 03:01:26 INFO HiveContext: Initializing execution hive, version 0.13.1
15/12/14 03:01:26 INFO metastore: Trying to connect to metastore with URI thrift://
sandbox.hortonworks.com:9083
15/12/14 03:01:26 INFO metastore: Connected to metastore.
15/12/14 03:01:26 INFO SessionState: No Tez session required at this point.
hive.execution.engine=mr.
15/12/14 03:01:26 INFO SparkILoop: Created sql context (with Hive support)..
SQL context available as sqlContext.
```

```
scala>
```