What is SparkSession

SparkSession was introduced in version Spark 2.0, It is an entry point to underlying Spark functionality in order to programmatically create Spark RDD, DataFrame, and DataSet. SparkSession's object spark is the default variable available in spark-shell and it can be created programmatically using SparkSession builder pattern.

SparkSession in Spark 2.0

With Spark 2.0 a new class org.apache.spark.sql.SparkSession has been introduced which is a combined class for all different contexts we used to have prior to 2.0 (SQLContext and HiveContext e.t.c) release hence, Spark Session can be used in the place of SQLContext, HiveContext, and other contexts.
As mentioned in the beginning SparkSession is an entry point to Spark and creating a SparkSession instance would be the first statement you would write to program with RDD, DataFrame, and Dataset. SparkSession will be created using SparkSession.builder() builder patterns.
Prior to Spark 2.0, SparkContext used to be an entry point, and it's not been completely replaced with SparkSession, many features of SparkContext are still available and used in Spark 2.0 and later. You should also know that SparkSession internally creates SparkConfig and SparkContext with the configuration provided with SparkSession.
Spark Session also includes all the APIs available in different contexts –

- SparkContext
- SQLContext
- StreamingContext
- HiveContext

How many SparkSessions can you create in an application?

You can create as many SparkSession as you want in a Spark application using either SparkSession.builder() or SparkSession.newSession(). Many Spark session objects are required when you wanted to keep Spark tables (relational entities) logically separated.

SparkSession in spark-shell

Be default Spark shell provides spark object which is an instance of SparkSession class. We can directly use this object when required in spark-shell.

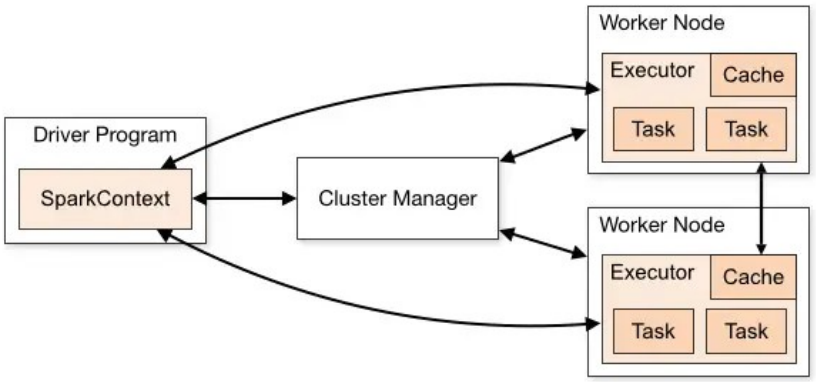
```
// Usage of spark variable
scala> spark.version
```

Similar to the Spark shell, In most of the tools, notebooks, and Azure Databricks, the environment itself creates a default SparkSession object for us to use so you don't have to worry about creating a spark session.

What is SparkContext

Since Spark 1.x, SparkContext is an entry point to Spark and is defined in org.apache.spark package. It is used to programmatically create Spark RDD, accumulators, and broadcast variables on the cluster. Its object sc is default variable available in spark-shell and it can be programmatically created using SparkContext class.

Note that you can create only one active SparkContext per JVM. You should stop() the active SparkContext before creating a new one.



The Spark driver program creates and uses SparkContext to connect to the cluster manager to submit Spark jobs, and know what resource manager (YARN, Mesos or Standalone) to communicate to. It is the heart of the Spark application.

SparkContext in spark-shell

Be default Spark shell provides sc object which is an instance of SparkContext class. We can directly use this object where required.

```
scala>> sc.appName
```

Similar to the Spark shell, In most of the tools, notebooks, and Azure Databricks, the environment itself creates a default SparkContext object for us to use so you don't have to worry about creating a spark context.