Spark adalah sub-proyek Hadoop ini. Oleh karena itu,

lebih baik untuk menginstal Spark ke dalam sistem berbasis Linux

Langkah-langkah berikut menunjukkan cara menginstal Apache Spark.

Langkah 1: Memeriksa Instalasi Java

instalasi Java adalah salah satu hal yang wajib dalam memasang Spark. Cobalah perintah berikut untuk memverifikasi versi JAVA.

$java -version

Jika Java sudah, diinstal pada sistem Anda, Anda bisa melihat respon berikut -

java version "1.7.0\_71"

Java(TM) SE Runtime Environment (build 1.7.0\_71-b13)

Java HotSpot(TM) Client VM (build 25.0-b02, mixed mode)

Dalam kasus Anda tidak memiliki Java diinstal pada sistem Anda, kemudian Instal Java sebelum melanjutkan ke langkah berikutnya.

Langkah 2: Memeriksa instalasi Scala

Anda harus bahasa Scala untuk melaksanakan Spark. Jadi mari kita memeriksa instalasi Scala menggunakan perintah berikut.

$scala -version

Jika Scala sudah diinstal pada sistem Anda, Anda bisa melihat respon berikut -

Scala code runner version 2.11.6 -- Copyright 2002-2013, LAMP/EPFL

Dalam kasus Anda tidak memiliki Scala diinstal pada sistem Anda, kemudian lanjutkan ke langkah berikutnya untuk instalasi Scala.

Langkah 3: Men-download Scala

Download versi terbaru dari Scala oleh kunjungan link berikut Unduh Scala .

Untuk tutorial ini, kita menggunakan versi scala-2.11.6. Setelah men-download, Anda akan menemukan Scala file tar di folder download.

Langkah 4: Instalasi Scala

Ikuti langkah-langkah di bawah ini diberikan untuk menginstal Scala.

Ekstrak Scala file tar

Ketik perintah berikut untuk mengekstraksi Scala file tar.

$ tar xvf scala-2.11.6.tgz

Memindahkan file software Scala

Gunakan perintah berikut untuk memindahkan Scala file perangkat lunak, direktori masing-masing (/ usr / local / scala).

$ su –

Password:

# cd /home/Hadoop/Downloads/

# mv scala-2.11.6 /usr/local/scala

# exit

Set PATH untuk Scala

Gunakan perintah berikut untuk pengaturan PATH untuk Scala.

$ export PATH = $PATH:/usr/local/scala/bin

Instalasi Memverifikasi Scala

Setelah instalasi, lebih baik untuk memverifikasinya. Gunakan perintah berikut untuk memverifikasi instalasi Scala.

$scala -version

Jika Scala sudah diinstal pada sistem Anda, Anda bisa melihat respon berikut -

Scala code runner version 2.11.6 -- Copyright 2002-2013, LAMP/EPFL

Langkah 5: Mendownload Apache Spark

Download versi terbaru dari Spark dengan mengunjungi berikut link di download Spark

Untuk tutorial ini, kita menggunakan versi spark-1.3.1-bin-hadoop2.6.

Setelah men-download, Anda akan menemukan file Spark tar di folder download.

Langkah 6: Instalasi Spark

Ikuti langkah-langkah berikut untuk menginstal Spark.

Penggalian Spark tar

perintah berikut untuk mengekstrak file percikan tar.

$ tar xvf spark-1.3.1-bin-hadoop2.6.tgz

Memindahkan file software Spark

Berikut perintah untuk memindahkan file software Spark ke direktori masing-masing (/ usr / local / spark).

$ su –

Password:

# cd /home/Hadoop/Downloads/

# mv spark-1.3.1-bin-hadoop2.6 /usr/local/spark

# exit

Menyiapkan lingkungan untuk Spark

Tambahkan baris berikut ke ~ /.bashrc berkas. Ini berarti menambah lokasi, di mana file software percikan terletak pada variabel PATH.

export PATH = $PATH:/usr/local/spark/bin

Gunakan perintah berikut untuk sumber file ~ / .bashrc.

$ source ~/.bashrc

Langkah 7: Memeriksa Instalasi Spark

Menulis perintah berikut untuk membuka shell Spark.

$spark-shell

Jika percikan berhasil diinstal maka Anda akan menemukan output berikut.

Spark assembly has been built with Hive, including Datanucleus jars on classpath

Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties

15/06/04 15:25:22 INFO SecurityManager: Changing view acls to: hadoop

15/06/04 15:25:22 INFO SecurityManager: Changing modify acls to: hadoop

15/06/04 15:25:22 INFO SecurityManager: SecurityManager: authentication disabled;

ui acls disabled; users with view permissions: Set(hadoop); users with modify permissions: Set(hadoop)

15/06/04 15:25:22 INFO HttpServer: Starting HTTP Server

15/06/04 15:25:23 INFO Utils: Successfully started service 'HTTP class server' on port 43292.

Welcome to

\_\_\_\_ \_\_

/ \_\_/\_\_ \_\_\_ \_\_\_\_\_/ /\_\_

\_\ \/ \_ \/ \_ `/ \_\_/ '\_/

/\_\_\_/ .\_\_/\\_,\_/\_/ /\_/\\_\ version 1.4.0

/\_/

Using Scala version 2.10.4 (Java HotSpot(TM) 64-Bit Server VM, Java 1.7.0\_71)

Type in expressions to have them evaluated.

Spark context available as sc

scala>