

Instalación del ambiente de desarrollo de programas

IntelliJ IDEA

NOTA: La instalación de IntelliJ IDEA debe ser en una computadora con interfaz gráfica, puede ser asignada a la madre o a un nodo encargado de hacer todo este trabajo como máquina de desarrollo.

La ejecución de Hadoop y otras de las herramientas de Big Data requieren de la máquina virtual de Java «JVM». Para instalar IntelliJ IDEA se descarga del link:

<https://www.jetbrains.com/idea/download/>

Por el momento descargue la versión *community*.

El archivo será descargado en la carpeta descargas, pero se recomienda que se encuentre en el directorio home (“~/”), posteriormente se instalará descomprimiéndolo ahí mismo, la ejecución se hace con:

```
$ tar xvfz ideaIC-2016.1.3.tar.gz
$ cd idea-IC-145.1617.8/bin/
$ ./idea.sh
```

NOTA: Verifica que la instalación es realizada como usuario *hadoop*

En el apartado de **Download feature plugins**, instale las tres opciones (*Scala*, *IdeaVim* e *IDE FeaturesTrainer*)

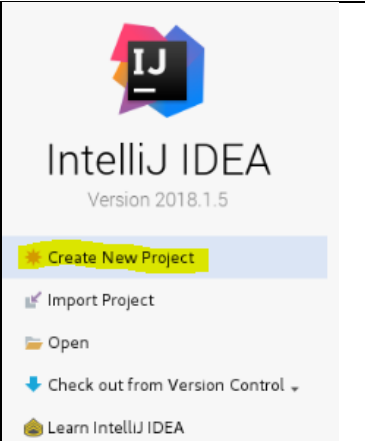
2.1 Programación usando el paradigma Map-Reduce usando IntelliJ IDEA

En esta sección se llevará a cabo el uso del modelo Map-Reduce para la ejecución de varios programas como son los que se mencionaran a continuación

Verificación de la última version actual estable

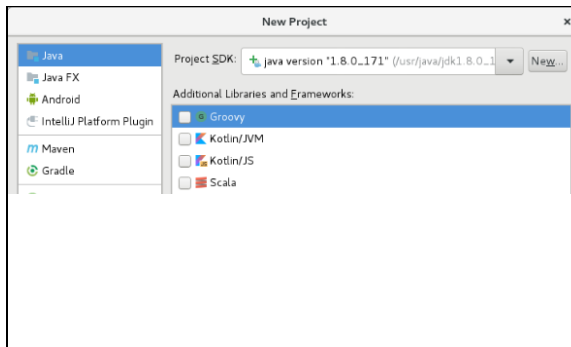
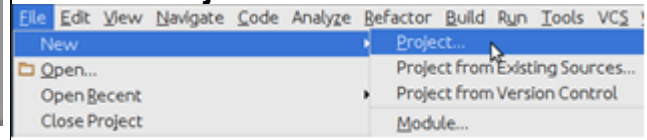
<http://hadoop.apache.org/releases.html>

2.1.2 Creación de un proyecto nuevo

	<p>En algunas ocasiones se presenta la pantalla que se muestra a la derecha. Comience dando «click» en la opción marcada Create New Project.</p>
---	---



En otras ocasiones se presenta una pantalla gris y un menú y opcionalmente una sugerencia en la ventana «Tip of the Day» la cual se cierra presionando el botón «Close». En el menú se seleccionan las opciones: **File/New/Project...**

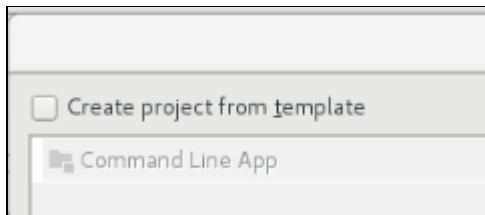


A continuación se presentara una ventana de un nuevo proyecto «New Project» en donde debe elegirse el lenguaje de programación a utilizar, en este caso es Java.

Para continuar presione el botón en la parte inferior.

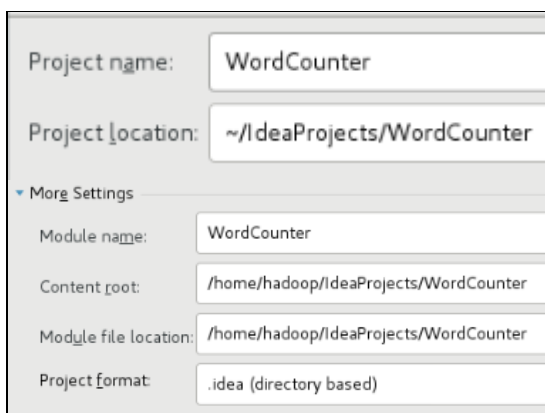


NOTA: Si falta asignar el SDK «Software Development Kit» se debe presionar el botón New... y navegar en los archivos para seleccionar el directorio donde se encuentra instalado el JDK de Java.



Esta ventana se usa si se desea crear un proyecto con una plantilla, en este caso

solamente presionamos el botón en la parte inferior para continuar.



En esta ventana se indica el nombre del proyecto, se sugiere que inicie con mayúscula usando la “notación de camello” de los lenguajes orientados a objetos. En este caso escribimos WordCounter debido a que será nuestro primer programa a realizar.

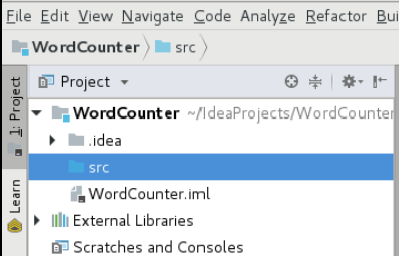
Se deja la ubicación sugerida en «Project location».

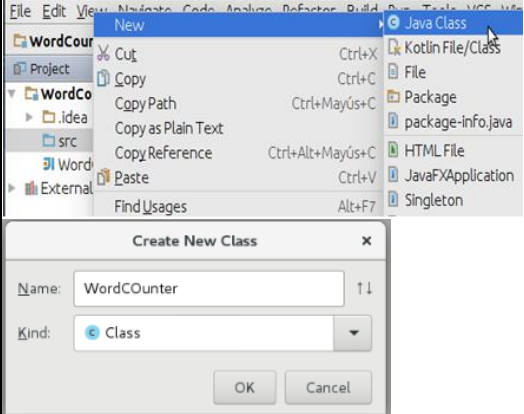
En la parte inferior se indica los valores escritos en la parte superior.

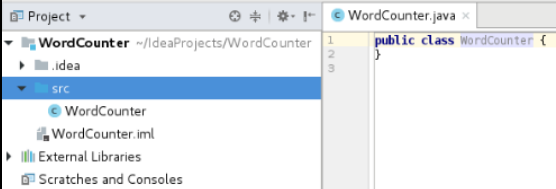
En este caso se presiona el botón para concluir la creación del nombre de un proyecto en lenguaje Java.



2.1.3 Creación de un programa en lenguaje JAVA

	<p>Para añadir archivos en lenguaje Java al proyecto se hace «click» en el triángulo gris en el lado izquierdo del nombre del proyecto para expandirlo. Luego se selecciona la carpeta de archivos de código fuente «src».</p>
---	--

	<p>Presionando el botón derecho se selecciona la opción New, luego Java Class y se hace «click» con el ratón con lo que aparece una ventana pequeña para escribir el nombre de la clase, en este caso la misma del nombre del proyecto WordCounter.</p>
---	---

	<p>En la parte derecha de la pantalla, en el contenido de la pestaña WordCounter.java se completa la clase WordCounter.java con el contenido del listado 1.</p>
---	--

<pre>import java.io.IOException; import java.util.*; import org.apache.hadoop.fs.Path; import org.apache.hadoop.conf.*; import org.apache.hadoop.io.*; import org.apache.hadoop.mapreduce.*; import org.apache.hadoop.mapreduce.lib.input.FileInputFormat; import org.apache.hadoop.mapreduce.lib.input.TextInputFormat; import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat; import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat; public class WordCounter { public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> { private final static IntWritable one = new IntWritable(1); private Text word = new Text(); public void map(LongWritable key, Text value, Context context) throws IOExce</pre>	<p>Al hacer la captura del programa, se puede observar que algunas palabras se colocan de color rojo. Esto es porque falta añadir las bibliotecas de Hadoop y Map-Reduce en el proyecto.</p> <p>Se sugiere concluir la captura del programa y al final añadir estas bibliotecas como se indica más adelante.</p>
---	--

Listado 1:

```
import java.io.IOException;
```

```
// BIBLIOTECAS DE JAVA
```

```

import java.util.*;
import org.apache.hadoop.fs.Path;                // BIBLIOTECAS DE HADOOP Y MAP-REDUCE
import org.apache.hadoop.conf.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.*;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;

public class WordCounter {
    public static class Map extends Mapper <LongWritable, Text, Text, IntWritable> {
        private final static IntWritable one = new IntWritable(1);
        private Text word = new Text();

        public void map(LongWritable key, Text value, Context context)
            throws IOException, InterruptedException {
            String line = value.toString();
            StringTokenizer tokenizer = new StringTokenizer(line);

            while (tokenizer.hasMoreTokens()) {
                word.set(tokenizer.nextToken());
                context.write(word, one);
            }
        }
    }

    public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
        public void reduce(Text key, Iterable<IntWritable> values, Context context)
            throws IOException, InterruptedException {
            int sum = 0;
            for (IntWritable val : values) {
                sum += val.get();
            }
            context.write(key, new IntWritable(sum));
        }
    }

    public static void main(String[] args) throws Exception {
        Configuration conf = new Configuration();
        Job job = new Job(conf, "WordCounter");
        job.setOutputKeyClass(Text.class);
        job.setOutputValueClass(IntWritable.class);

        job.setMapperClass(Map.class);
        job.setReducerClass(Reduce.class);

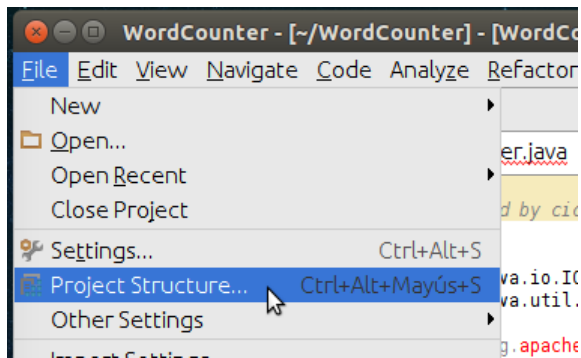
        job.setInputFormatClass(TextInputFormat.class);
        job.setOutputFormatClass(TextOutputFormat.class);

        FileInputFormat.addInputPath(job, new Path(args[0]));
        FileOutputFormat.setOutputPath(job, new Path(args[1]));

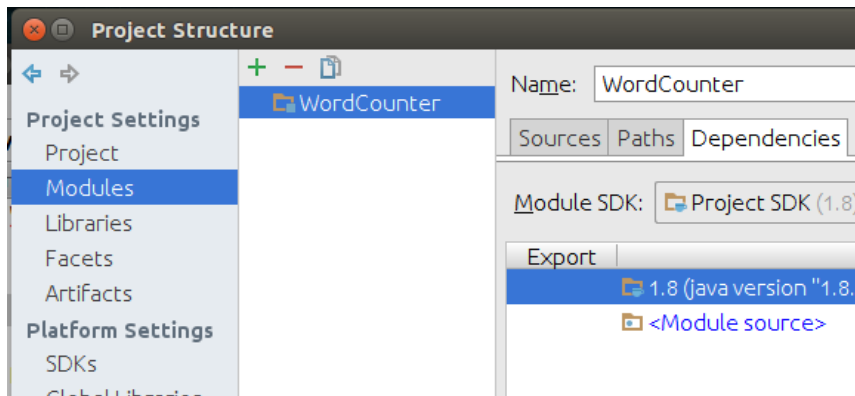
        job.waitForCompletion(true);
    }
}

```

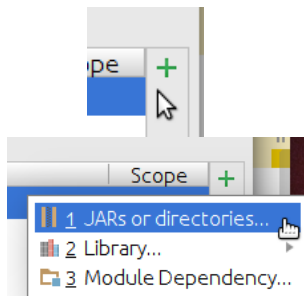
2.1.4 Asignación de las bibliotecas de Hadoop



Para que se reconozcan en el proyecto las bibliotecas de Hadoop y Map-Reduce se selecciona la opción **File** del menú de IntelliJ Idea y de sus opciones se hace «click» en **Project Structure**.

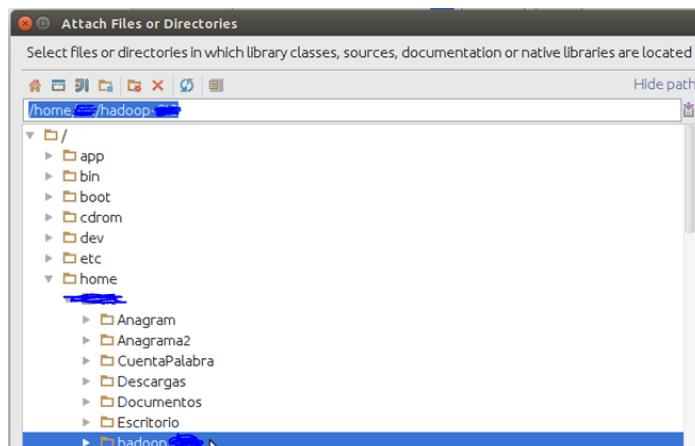


En la pantalla de **Project Structure** se selecciona en el lado derecho la pestaña **Dependencies**.



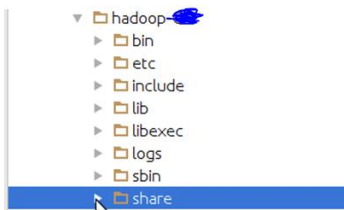
En la ventana de las dependencias se hace «click» en el signo **+** para añadir las bibliotecas requeridas.

En el menú emergente se selecciona la opción **1 JARs or directories...**

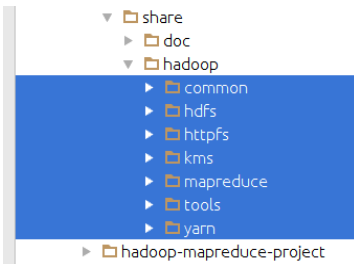


En la ventana de los directorios se elige la ruta donde se ha instalado Hadoop, en este caso:

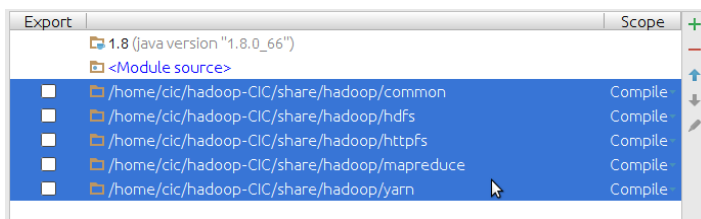
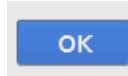
/home/hadoop



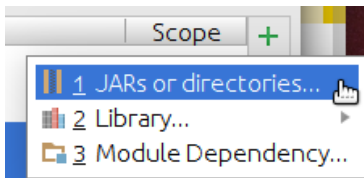
En la carpeta de instalación de Hadoop se accede al directorio **share**, donde se encuentran las bibliotecas de Hadoop.




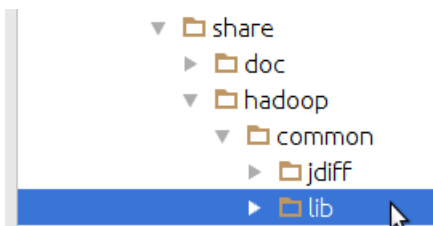
Se abre el directorio hadoop y se seleccionan los directorios common, hdfs, httpfs, kms, mapreduce, tools y yarn. Al concluir se presiona el botón



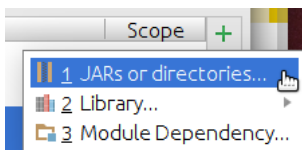
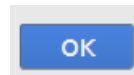
Los directorios seleccionados aparecen en la pantalla de dependencias.




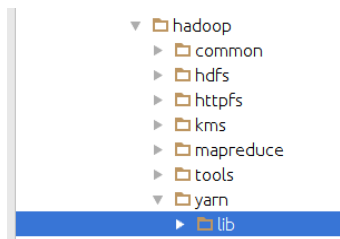
Se vuelve a presionar el botón  y se selecciona la opción **1 JARs or directories...** para añadir otra biblioteca.



Se selecciona el directorio **/home/hadoop/share/common/lib** y se presiona el botón

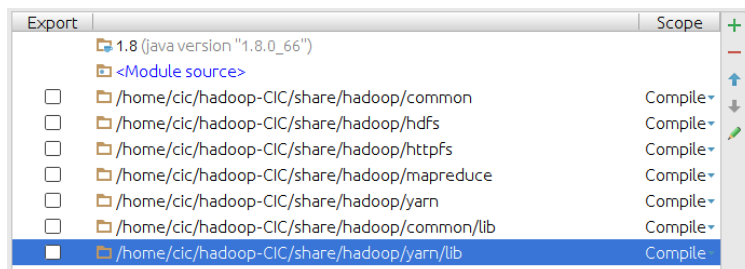
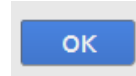


Se vuelve a presionar el botón  y se selecciona la opción **1 JARs or**

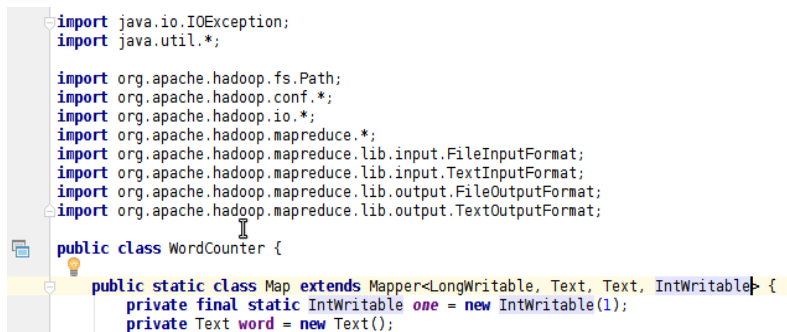


directories... para añadir otra biblioteca.

Ahora se selecciona el directorio **/home/hadoop/share/yarn/lib** y se presiona el botón

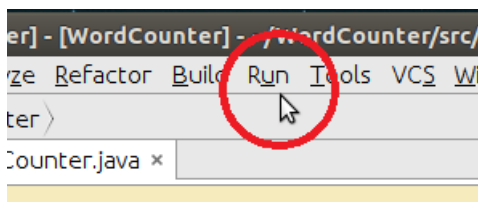


Una vez que se han añadido las bibliotecas requeridas se presionan los botones **Apply** y **Accept**.



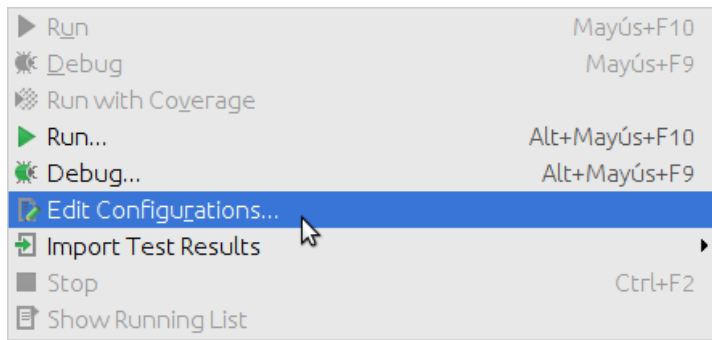
En el código se puede observar que las palabras marcadas con color rojo ahora se tornan de color negro indicando que ya son reconocidas. Si algunas palabras persisten de color rojo significa que hay errores cometidos por el usuario y se deben corregir.

2.1.5 Configuración de la ejecución del proyecto

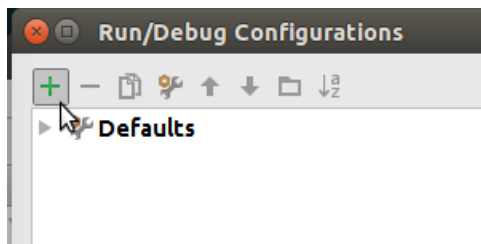


La ejecución de este proyecto requiere que se indiquen los archivos de entrada y salida que se deben indicar en la línea de comando porque en el programa se especificó con las líneas:

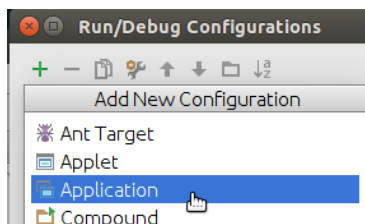
```
FileInputFormat.addInputPath(job, new
Path(args[0]));
FileOutputFormat.setOutputPath(job, new
Path(args[1]));
```



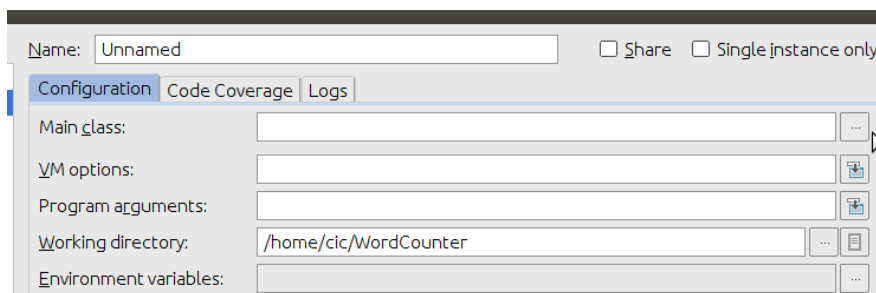
Al seleccionar **Run** se despliega un menú y de estas se elige **Edit Configurations...**



Se presenta entonces una pantalla en la que se debe presionar el ícono **+** para crear una configuración de ejecución.

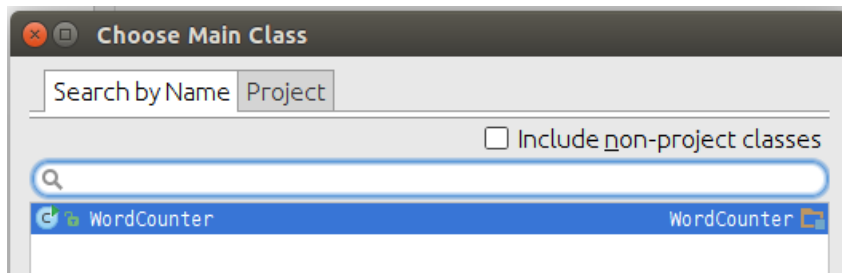



De las opciones que se presentan se elige **Application**.

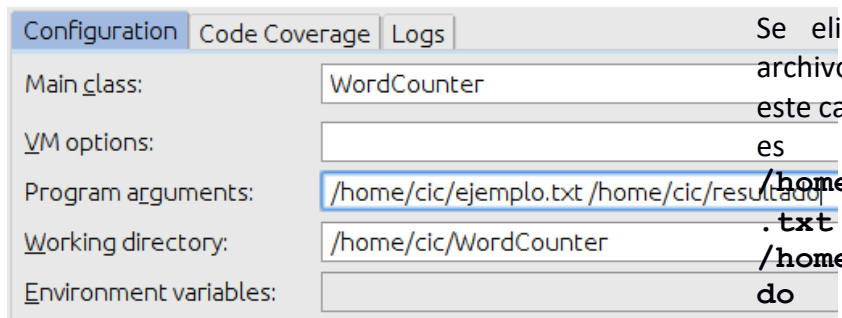


En la pantalla de configuración se deben especificar los valores de **Main class:** y **Program arguments:**

El **Name**: puede dejarse como Unnamed o colocar otro que haga referencia al proyecto.



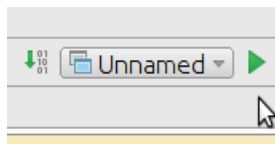
La clase principal se selecciona presionando el botón . Del menú de opciones se elige la clase principal, en este caso **WordCounter**.



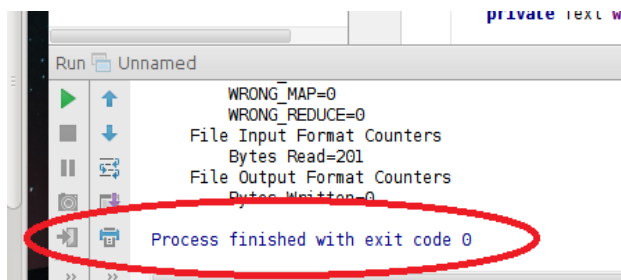
Se eligen las rutas de los archivos de entrada y salida. En este caso el archivo de entrada es **/home/upiita/ejemplo.txt** y el de salida es **/home/upiita/resultado.txt**.

Para terminar se debe presionar el botón Apply y

luego el botón .

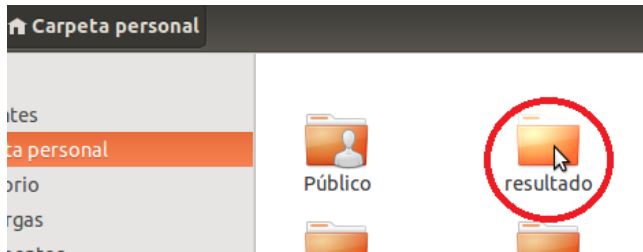


El programa se ejecuta ahora con el ícono de triángulo verde ubicado en la parte superior derecha de la pantalla.



La ejecución tarda de algunos segundos a varios minutos dependiendo del tamaño de los archivos de entrada. Si la terminación es correcta se despliega el mensaje:

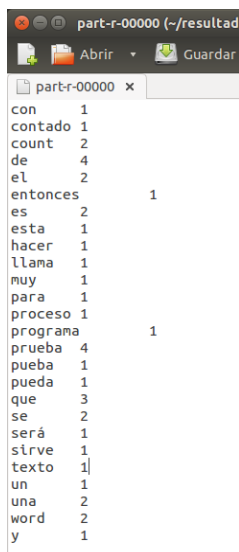
Process finished with exit code 0



La salida crea una carpeta **resultado** en el directorio **/home/upiita** que en el visualizador gráfico de archivos corresponde a **Carpeta personal**.



Al abrir la carpeta de salida se pueden observar dos archivos **_SUCCESS** que se crea cuando la ejecución es correcta y **part-r-00000** que contiene los datos.



Al abrir el archivo **part-r-00000** con algún editor de texto se pueden ver los resultados de los cuales se muestra un fragmento.

Para analizar un libro, se puede descargar los libros que desee del siguiente link.
<http://www.gutenberg.org/browse/languages/es>

2.1.6 Ejecución almacenando el resultado en Hadoop

```

cic@maestra:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/cic/hadoop-CIC/logs/hadoop-cic-namenode-maestra.out
localhost: starting datanode, logging to /home/cic/hadoop-CIC/logs/hadoop-cic-datanode-maestra.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/cic/hadoop-CIC/logs/hadoop-cic-secondarynamenode-maestra.out
starting yarn daemons
starting resourcemanager, logging to /home/cic/hadoop-CIC/logs/yarn-cic-resourcemanager-maestra.out
localhost: starting nodemanager, logging to /home/cic/hadoop-CIC/logs/yarn-cic-nodemanager-maestra.out
cic@maestra:~$

```

En algunas ocasiones se requiere almacenar o leer datos en Hadoop para lo cual es necesario ejecutar el servidor de Hadoop en una terminal con el script:

start-all.sh

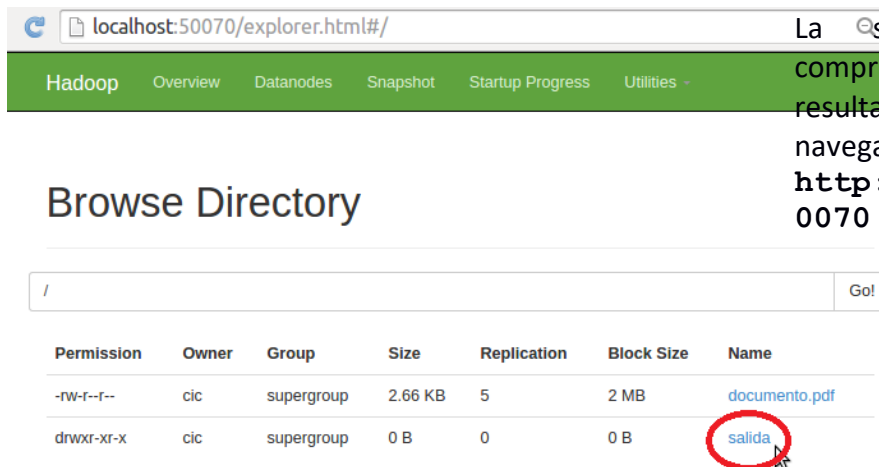
```

rn-cic-nodemanager-maestra
cic@maestra:~$ jps
11793 Launcher
20215 Main
13195 NameNode
13867 NodeManager
13339 DataNode
13742 ResourceManager
14142 Jps
13519 SecondaryNameNode
cic@maestra:~$

```

El comando **jps** permite verificar que el servidor se activó correctamente.

Configuration	Code Coverage	Logs	
Main class:	WordCounter		La ejecución accedendo datos en Hadoop se requiere que la ruta de los archivos inicie con:
VM options:			hdfs://localhost:54310/salida.txt
Program arguments:	/home/cic/ejemplo.txt hdfs://localhost:54310/salida.txt		hdfs://localhost:54310/
Working directory:	/home/cic/WordCounter		En este caso el archivo es salida.
Environment variables:			



La salida se puede comprobar visualizando el resultado mediante un navegador en la dirección: **http://localhost:50070**

Size	Replication	Block Size	Name
0 B	3	128 MB	_SUCCESS
189 B	3	128 MB	part-r-00000

Al dar «click» en salida se puede notar que tiene dos archivos: **_SUCCESS** creado cuando la ejecución es correcta y **part-r-00000** que contiene los datos.

```
que      3
se       2
será     1
sirve    1
texto    1
un       1
una      2
word     2
y        1
cic@maestra:~$ hdfs dfs -text /salida/part-r-00000
```

El contenido de **part-r-00000** es posible visualizarlo en la terminal con el comando de Hadoop:
hdfs dfs -text /salida/part-r-00000

2.1.7 CREACION DE ARCHIVOS .JAR

PASOS PARA COMPILAR JAVA EN CLUSTER

PASOS PARA COMPLILAR EL .JAR USANDO EL HADOOOP DEL CLUSTER

Para compilar archivos .jar en el cluster modo minimal debe agregar:

nano .bashrc

```
export HADOOP_CLASSPATH:$JAVA_HOME/lib/tools.jar
```

3.Creación de clusters multinodo

Como primer paso respalde los siguientes archivos

```
core-site.xml
hadoop-env.sh
hdfs-site.xml
yarn-site.xml
mapred-site.xml
masters
slaves
```

NOTA: El nombre de este equipo de un solo nodo que se sugiere llamar maestra se registra en el archivo del sistema operativo /etc/hosts

Editar el archivo hosts

```
sudo nano /etc/hosts
```

El contenido de este archivo es una lista de direcciones IP, un tabulador y el nombre del equipo, por ejemplo:

```
192.168.1.2 maestra
192.168.1.10 nodo01
192.168.1.11 nodo02
```

Nota: Usted debe agregar esta información después de las dos líneas que ya existen en el archivo, además de poder agregar el dns si es que así usted lo desee por cada ip.

```
maestra con permisos de super usuario
ip addr add 192.168.1.2/24 broadcast 192.168.1.255 dev enp2s0

nodo01
ip addr add 192.168.1.10/24 broadcast 192.168.1.255 dev enp4s0
```

NOTA: Error Port 22, problemas con conexión de ssh, debido a que no están ingresadas las direcciones IP0 correspondientes a las configuraciones.

NOTA: No se recomienda que se utilice wifi para conectarse entre nodos, la manera correcta será conectando la computadora maestra y los nodos por medio de un switch o hub.

- Checar nombre del usuario

```
hostname
```

Asignar los nombres correspondientes (maestra, nodo01), esto se realiza para tener un mayor control al momento de inicializar los servicios. **PUEDE OMITIRLO**

```
sudo nano /etc/hostname
```

```
maestra
```

Este segundo comando es para el secondary namenodes

```
nano /etc/sysconfig/network
```

```
HOSTNAME=maestra
```

Reiniciar la tarjeta de internet

```
/etc/init.d/network restart
```

```
Hostnamectl
```

Repita este mismo paso en todas las computadoras, puede realizarlo desde la maestra con el protocolo ssh.

```
su hadoop
```

```
ssh-keygen -t rsa
```

```
ssh-copy-id -i ~/.ssh/id_rsa.pub hadoop@maestra
```

```
ssh-copy-id -i ~/.ssh/id_rsa.pub hadoop@nodo01
```

Comprueba conectándote a los demás usuarios, o haciendo ping al usuario.

Para copiar el archivo a todos los usuarios, solo copie este archivo a cada nodo, de la máquina maestro, cópielo a cada uno de los esclavos

```
scp
```

Editar Core-site.xml

```
<configuration>
```

```
<property>
```

```
<name> fs.defaultFS</name>
```

```
<value>hdfs://maestra:54310</value>
```

```
</property>
```

```
<property>
```

```
<name> fs.default.name</name>
```

```
<value>hdfs://maestra:54310</value>
```

```
</property>
```

```
</configuration>
```

Editar hdfs-site.xml

```
<configuration>

  <property>
    <name>dfs.replication</name>
    <value>3</value>
  </property>
  <property>
    <name>dfs.blocksize</name>
    <value>67108864</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>/home/hadoop/hadoopdata/hdfs/namenode </value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>/home/hadoop/hadoopdata/hdfs/datanode </value>
  </property>

</configuration>
```

NOTA: El número de replicaciones es 3 cuando tenemos más de 3 nodos en nuestra configuración.

Editar Mapred-site.xml

```
<property>
  <name>mapred.job.tracker</name>
  <value>NODO MAESTRA:54311</value>
</property>
```

NOTA: la versión 3.0.3 cambio `mapred.job.tracker` a `mapred.jobtracker.address`

Editar Yarn-site.xml

```
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
<property>
  <name>yarn.nodemanager.aux-services.mapreduce_shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>

<property>
```

```

    <name>yarn.resourcemanager.resource-tracker.address</name>
    <value>maestra:8031</value>
  </property>
</property>
    <name>yarn.resourcemanager.scheduler.address</name>
    <value>maestra:8030</value>
  </property>
</property>
    <name>yarn.resourcemanager.address</name>
    <value>maestra:8032</value>
  </property>
</configuration>

```

nano masters, quitar localhost

```
maestra
```

nano slaves

```
maestra
nodo01
nodo02
```

Copia las configuraciones a los demás nodos, verifica que los parámetros del bloque sean los requeridos por los equipos

```

scp core-site.xml mapred-site.xml hdfs-site.xml yarn-site.xml hadoop-env.sh
  masters slaves nodo01:/home/hadoop/etc/hadoop/
scp /.bashrc nodo01: <Ruta origen>

```

Nota: Borre la anterior configuración de la ruta creada en la configuracions de hdfs-site.xml

```

hdfs namenode -format
start-dfs.sh
start-yarn.sh

```

Panorama de nuestros nodos vivos en consola

```
hdfs dfsadmin -report|more
```

Puede comprobar el funcionamiento de sus nodos

```
ssh nodo01 jps
```

Nota: Cuando tuve problemas de levantar los dos nodos, detuve los servicios, escribi: **sudo iptables -F**

Desabilita el firewall de Centos 7

```

systemctl disable firewalld
systemctl stop firewalld
systemctl status firewalld

```


Otra manera

```
sestatus  
sudo setenforce 0
```

Vuelva a inicializar los servicios.

Las salidas a obtener de jps

maestra	namenode, datanode secondarynamenode
Nodo01	Datanode, resourcemanager

Para comprobar que tu cluster funciona correctamente es necesario realizar una prueba de verificación, para ello usted debe crear un archivo más grande que el tamaño de uno de sus nodos, poderlo subir y descargar de hadoop.

NOTA: El cluster puede ser ejecutado desde cualquier nodo.

CHECKlist BIG DATA

- /etc/hosts
- masters, slaves
- Permisos ssh a cada canal
- Comunicación bidireccional maestro esclavo
- Core-site.xml
- Hdfs-site.xml
- Mapred-site.xml
- Yarn-site.xml
- Hadoop-env.sh
- Hdfs namenode –format
- Start-dfs.sh

Nota: Si no aparece datanode ni namenode con jps no continuar. Eliminar archivos de namenode

- Start-yarn.sh

NOTA: Para verificar posibles errores generados al levantar varios nodos, entre a las carpetas de namenode y datanode para ver el archivo que se genera y los posibles errores que se puedan escribir en esas carpetas. Verifique la carpeta logs.

//*****

INFORMACIÓN ADICIONAL

*****/

CONEXIÓN DE INTERNET PARA PC SIN TARJETA DE RED PARA IPS ESTATICAS PARA CENTOS MODO MINIMAL EN EL UPIITA

Uno de los comandos para saber mis interfaces de red es con

```
ip link
```

De esta forma puedo saber cuáles son mis adaptadores de red wifi.

Para poder configurar las interfaces de red en CentOS, se procederá a ingresar a la siguiente dirección, la cual fue modificada la tarjeta de red ETHERNET para este manual.

```
yum vi /etc/sysconfig/network-scripts/ifcfg-enp0s25
```

Estas son las configuraciones que se realizaron para ingresar una IP estática (solicítala con el Dr. que cuenta con una) en el UPIITA debido a que en este lugar se maneja esta configuración.

```
DEVICE=enp0s25
BOOTPROTO=static
IPADDR=192.168.1.2
NETWORKING=YES
NETMASK=255.255.255.0
GATEWAY=10.0.1.1
DNS1=
ONBOOT=yes
```

NOTA: La información mencionada anteriormente vari.

Ahora reinicie el demonio encargado de activar la interfaz de red.

```
service network restart
```

NOTA: Puede probar utilizando el comando **yum update** para comprobar que ya cuenta con acceso de internet. De no funcionar este comando, reinicie la computadora.

Para iniciar la tarjeta de red alámbrica

```
ifup enp4s0
```

Cibergrafía: <http://xmodulo.com/how-to-configure-network-interfaces-in.html>

CONFIGURACIÓN DE WIFI EN CENTOS 7 MINIMAL

```
chkconfig NetworkManager on
service NetworkManager on
nmtui
```

NOTA: Si no puedes activar el wifi, ingresa el siguiente comando:

```
/etc/init.d/network restart
```

Reinicie la computadora.

CREAR UN DIRECTORIO PARA MONTAR UNA MEMORIA USB

NOTA: TODAS LAS CONFIGURACIONES A CONTINUACIÓN SON PARA EL SUPER USUARIO **root**, o **permisos de super usuario**.

PASO 1. Crear el directorio para montar la memoria usb

```
mkdir /media/usb
```

NOTA: Este comando crea la carpeta donde estará montada nuestra memoria usb. En caso de crear montar 2 memorias USB de forma simultánea, usted debe crear una ubicación, por ejemplo /media/usb2.

PASO 2. IDENTIFICAR EL NOMBRE DE LA UNIDAD QUE QUEREMOS MONTAR

Utilice el siguiente comando, conectando y desconectando la memoria usb para identificar cual es el nombre del dispositivo correspondiente.

```
ls -l /dev/sd*
```

Mensajes desplegados sin la pendrive conectada

```
brw-rw-r-- T 1 root disc 8, 0 mar 16 11:21 /dev/sda
brw-rw-r-- T 1 root disc 8, 1 mar 16 11:21 /dev/sda1
brw-rw-r-- T 1 root disc 8, 2 mar 16 11:21 /dev/sda2
brw-rw-r-- T 1 root disc 8, 5 mar 16 11:21 /dev/sda5
```

Mensaje desplegado con pendrive conectado

```
brw-rw-r-- T 1 root disc 8, 0 mar 16 11:21 /dev/sda
brw-rw-r-- T 1 root disc 8, 1 mar 16 11:21 /dev/sda1
brw-rw-r-- T 1 root disc 8, 2 mar 16 11:21 /dev/sda2
brw-rw-r-- T 1 root disc 8, 5 mar 16 11:21 /dev/sda5
brw-rw-r-- T 1 root floppy 8, 16 mar 16 11:24 /dev/sdb
brw-rw-r-- T 1 root floppy 8, 17 mar 16 11:24 /dev/sdb1
```

Se ha identificado que nuestro dispositivo cuenta con el nombre de la partición **sdb1**.

PASO 3: MONTAR LA MEMORIA USB CON LA TERMINAL

En los apartados donde se menciona la palabra **INSTALE**. Puede seguir los siguientes pasos.

El siguiente comando está hecho para una memoria USB de formato FAT o FAT32.

```
mount -t vfat /dev/sdb1 /media/usb
```

creo

Umount /media/usb

Cibergrafía

<https://geekland.eu/montar-la-memoria-usb-en-la-terminal/>

CREACIÓN DE UNA COPIA DE ARCHIVOS DESDE LA MEMORIA USB A LA DIRECCIÓN DESTINO

Después de haber creado la carpeta **/media/usb**, entre como usuario **hadoop** y copie los archivos necesario a su computadora, en este ejemplo se copiara el archivo de java a la dirección raíz.

```
sudo cp /media/usb/jdk-8u171-linux-x64.rpm ~/.
```

+

NOTA: Verifique los archivos copiados tengan los permisos de **hadoop**, si tienen permisos de **superusuario** usted no podrá inicializar **hadoop**.

CONFIGURACIÓN CONEXIÓN DE UN CLUSTER EN VIRTUAL BOX

Clone las maquinas completamente y reinicie la MAC

Configure la red como red interna, asigne el mismo nombre a todas las maquinas virtuales, asigne direcciones ip, ping, hdfs dfsadmin report

LA PRUEBA FINAL ES subir a hadoop un archivo mayor al tamaño de un nodo, subir y bajar el archivo.

EDITOR VI

Movimiento del cursor:

Comando (teclas)	Acción
Flechas	Mover en la dirección de la flecha
h	Mover hacia la izquierda
l	Mover hacia la derecha
k	Mover hacia arriba
j	Mover hacia abajo
1G	Lleva el cursor hasta el comienzo del archivo
G	Lleva el cursor hasta el final del archivo

Cambio de modo comando a texto:

Comando	Acción
i	Inserta texto a la izquierda del cursor
a	Inserta texto a la derecha del cursor
A	Inserta texto al final de la línea donde se encuentra el cursor
I	Inserta texto al comienzo de la línea donde se encuentra el cursor
o	Abre una línea debajo de la actual
O	Abre una línea encima de la actual

Borrar texto:

Comando	Acción
x	Borra el carácter bajo el cursor
dd	Borra la línea donde se encuentra el cursor
ndd	Borra las próximas n líneas
D	Borra desde donde se encuentra el cursor hasta el final de la línea
dw	Borra desde donde se encuentra el cursor hasta el final de una palabra

Cortar y pegar:

Esto implica mover partes del archivo de un lugar a otro del mismo. Para esto se debe:

- Cortar el texto que se desea mover utilizando alguno de los comandos usados para borrar texto.
- Mover el cursor (con alguno de los comandos utilizados para desplazar el cursor en el texto) hasta el lugar donde se desee pegar el texto.
- Pegar el texto con el comando **p**.

Copiar y pegar:

Esta operación difiere de la anterior. En este caso lo que se hace es repetir partes del texto en otro lugar del archivo. Para esto se debe:

- Utilizar el comando **yy**, cuya función es copiar la línea donde se encuentra situado el cursor.
- Mover el cursor (con alguno de los comandos utilizados para desplazar el cursor en el texto) hasta el lugar donde se desee pegar el texto.
- Pegar el texto con el comando **p**.

Deshacer cambios:

Se puede deshacer el último cambio realizado, utilizando el comando **u**.

Buscar texto:

Comando	Acción
/texto	Busca hacia adelante la cadena de caracteres "texto"
?texto	Busca hacia atrás la cadena de caracteres "texto"

Salir de vi, salvar, no salvar cambios, etc.:

Comando	Acción
:q	Salir si no hubo cambios
:q!	Salir sin guardar cambios
:w	Guardar cambios
:w archivo1	Guardar cambios en archivo1
:wq	Guardar cambios y salir

EXTRAS

MOSTRAR EL NÚMERO DE LÍNEA EN LOS EDITORES DE TEXTO NANO, VI, VIM Y NANO.

Muestrá el número de renglón del archivo
nano -c <RUTA>

vi/vim
:set number

Dejar de tener el número del renglón
:set nonumber

AUMENTAR NUMERO DE PANTALLAS EN CENTOS

Diríjase a la ruta **Tweak Tools -> Workspaces** y aumente en número de pantallas que desea tener.

//PONER EL ARCHIVO Y FOTO

VINCULAR DROPBOX A LINUX

Descargar la actualización más actual del 2015

```
cd ~ && wget -O - "https://www.dropbox.com/download?plat=lnx.x86_64" | tar xzf -
```

Ten una terminal solo para ejecutar este comando, ya que será el que sincronice Dropbox
~/ .dropbox-dist/dropboxd

La carpeta de Dropbox aparecerá en la ruta de origen

Cibergrafía

<https://www.dropbox.com/es/install-linux>

CTRL+R	Buscar comandos ya utilizados
ALT + F1,F2,F3,F4,F5,F6	Cambiar de terminales en modo mínimo
pwd	Saber la ruta en la cual nos encontramos

EDITORES DE SCREENSHOOT

```
sudo yum install kdegraphics
```

RASPBERRY PI CENTOS

CAMBIAR EL IDIOMA DEL TECLADO A ESPAÑOL

```
sudo yum install kbd  
localectl set-keymap es
```

Cibergrafía:

<https://fututel.com/es/numero-virtual-did/10690-tutoriales-videotutoriales/2712-cambiar-el-idioma-del-teclado-en-centos-7-linux>

Para la instalación de JAVA, primero ingrese a la página oficial y descargue la versión más reciente para arm de 32 bits.

<http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html?ssSourceSiteId=otnes>

Después de descargar el paquete, descomprimelo en la siguiente ruta, utilizando estos comandos.

```
sudo mkdir /usr/java  
sudo cp jdk-8u181-linux-arm32-vfp-hflt.tar.gz /usr/java  
cd /usr/java  
sudo tar xf jdk-8u181-linux-arm32-vfp-hflt.tar.gz
```

Puedes crear un link simbolico en la ruta /usr/bin (para poder usar el comando java donde sea) agrega las siguientes líneas:

```
sudo update-alternatives --install /usr/bin/java java  
/usr/java/jdk1.8.0_181/bin/java 1000
```

```
sudo update-alternatives --install /usr/bin/javac javac  
/usr/java/jdk1.8.0_181/bin/javac 1000
```

//Y HASTA AHI NO ME DEJO HACER JAVA --VERSION

<https://blog.livthomas.net/installing-java-8-on-raspberry-pi-3/>

Reinicie

instale hadoop en Raspberry pi

la configuración de .bashrc es en .bash_profile en raspberry pi

df -h tamaño de espacio de la raspberry

tamaño de toda la micro sd

lsblk

<https://geekytheory.com/tutorial-raspberry-pi-11-como-instalar-java/>

sudo yum update && sudo yum install oracle-java7-jdk

CIBERGRAFIA

<https://medium.com/@khandelwal12nidhi/install-java-on-linux-dbbc67eb4cba>

<http://hadoop.apache.org/releases.html#Download>

https://www.tutorialspoint.com/big_data_tutorials.htm