Desfragmentador de archivos para Stratos

Ruth Barragán, Rosa Michel y Cynthia Martínez
Departamento de Sistemas y Computación
Instituto Tecnológico de Ciudad Guzmán
Ciudad Guzmán, Jal.; México
rcbarragan@itcg.edu.mx, [michel91_3, cynthia_amp]@hotmail.com

Abstract— Continued growth of digital information has made applications that help efficiently perform processing, reading and writing large amounts of data in parallel and multiple discs, the discs are located on a different machine. The defragmenter files application handles split a file onto fixed-size blocks to distribute the different divisions of the same between several machines, also storage of large files and has de ability to perform data replication to avoid loss of data in case of failure, this transparently to the user.

Keyword— StratOS, Desfragmenter, HDFS.

Resumen— El continuo crecimiento de la información digital ha hecho que se desarrollen aplicaciones que ayuden a realizar de forma eficiente el procesamiento, lectura y escritura de grandes cantidades de datos en paralelo y en múltiples discos, donde los discos están ubicados en máquinas distintas. Esta aplicación conocida como "desfragmentador de archivos" se encarga de dividir un archivo en bloques de tamaño fijo para distribuir las diferentes divisiones del mismo entre varias máquinas, también del almacenamiento de archivos de gran tamaño y tiene la capacidad de realizar una replicación de datos que evite la pérdida de los mismos en caso de fallo, esto de forma transparente para el usuario.

Palabras claves— StratOS, Desfragmentador, HDFS.

I. Introducción

En la actualidad se trabaja con grandes cantidades de datos, esto ocasiona una gran demanda en el uso de computadoras más veloces y aplicaciones ligeras que ocupen poco espacio en memoria y demanden mayor cantidad de procesamiento el cual llegue a afectar el funcionamiento de la computadora. El sistema de archivos distribuido de Hadoop, está diseñado para guardar grandes cantidades de datos y proveer acceso a esta información. Cuando se habla de grandes cantidades de información se refiere a archivos de cientos de Megabytes, Gigabytes o Terabytes o incluso Pentabytes. Está diseñado para ser tolerante a una alta probabilidad de fallos de la máquina. Es capaz de continuar trabajando sin interrupciones perceptibles por el usuario en caso de que se produzca un fallo.

II. Bigdata y Hadoop

Bigdata permite el almacenamiento de grandes volúmenes de datos y a los procedimientos usados para encontrar patrones repetitivos dentro de esos datos "tal y como lo presenta Barranco [1]". También se le conoce como almacenamiento de datos a gran escala. Además del gran volumen de información también existe gran variedad de datos que se pueden representar de diversas formas de tal forma que las aplicaciones que analizan estos datos requieren que la velocidad de respuesta sea demasiado rápida para obtener los resultados en el momento preciso.

La clasificación de Bigdata se reduce a: Web and social media, que incluye contenido web e información que se obtiene de las redes sociales. Machine to machine, que se refiere a las tecnologías que pueden conectarse a otros dispositivos. Big transaction data, que incluye registros de facturación, registros detallados de las llamadas, etc. Biometrics, en la que se incluyen huellas digitales, escaneo de retina, reconocimiento facial, etc. Human generated, toda la información que se genera a través de

llamadas telefónicas, correos electrónicos, estudios médicos, etc. Toda esa cantidad de información actualmente la está analizando la plataforma de Hadoop.

Hadoop es un proyecto de software abierto "tal y como se presenta en Wikipedia [3]", que permite el procesamiento distribuido de grandes conjuntos de datos en clusters de servidores básicos. Está diseñado para extender un sistema de servidor único a miles de máquinas, con un alto grado de tolerancia a las fallas.

Hadoop es un sistema de código abierto que se utiliza para almacenar, procesar y analizar grandes volúmenes de información de manera eficiente a través de computación distribuida, conectando computadoras y coordinándolas para trabajar juntas en paralelo. Tiene la ventaja de proveer un modelo de programación simple, el cual permite escribir y hacer pruebas en forma rápida. Provee un sistema eficiente de distribución automática de datos.

Los componentes básicos de Hadoop son los siguientes:

1. *HADOOP COMMON.*

Es un conjunto de herramientas que sirven de base para otros componentes Hadoop. Incluye el sistema de archivos básico, manejo de llamadas a procedimientos remotos y bibliotecas de serialización.

2. *HDFS.*

Consiste en un sistema de archivos distribuido, que permite que el archivo de datos no se guarde en una única máquina sino que sea capaz de distribuir la información a distintos dispositivos. Es esta distribución y redundancia la que permite el acceso rápido y la tolerancia a fallos en los nodos del clúster.

3. *MAPREDUCE.*

Se trata de un conjunto de software de trabajo que hace posible aislar al programador de todas las tareas propias de la programación en paralelo. Es decir, permite que un programa que ha sido escrito en los lenguajes de programación más comunes, se pueda ejecutar en un clúster de Hadoop.

El sistema más común para correr aplicaciones Big Data es Hadoop, desarrollado por Yahoo. Hadoop está basado en el patrón MapReduce. Tiene su propio sistema de archivos que usa replicación de datos para lograr localidad y resistencia a fallos. Si un nodo del sistema o un disco falla, los datos se recuperan de las copias distribuidas en el resto del sistema y la tarea que estaba corriendo en el nodo fallido se reanuda.

La ingestión de datos en los centros de procesamiento de datos es un problema común en Big Data clústers. Esto debido principalmente a la dificultad de crear patrones generales para proyectos que tienen necesidades muy específicas de almacenamiento de datos. Por ejemplo, en las simulaciones de universos paralelos, tal y como lo presentan Nathaniel y Aragón [2], un tipo de análisis requiere leer todos los universos paralelos al mismo tiempo. El almacenamiento más eficiente para I/O se logra cuando cada universo simulado está almacenado en un nodo del clúster, no distribuido a través del clúster, como en el caso default en HDFS. En otro tipo de análisis se requiere leer una sola realización en diferentes tiempos, en este caso el almacenamiento más eficiente se logra cuando los archivos de la realización están distribuidos a través del clúster. Por lo tanto, se propone desarrollar un desfragmentador de archivos para un sistema distribuido, como lo es StratOS. Esta herramienta consiste en un programa que siguiendo instrucciones del usuario sean distribuidos los archivos de acuerdo a su uso más eficiente. El desfragmentador se usará durante el copiado de archivos al clúster o cuando los datos están en el físicamente en un nodo del clúster.

III. HDFS EN HADOOP

Hoy día se vive en la era de los datos. No es fácil medir el volumen total de datos almacenados electrónicamente pero su constante crecimiento exponencial, ha hecho que los avances tecnológicos en áreas de almacenamiento y distribución de grandes cantidades de información estén en constante desarrollo, aunque en algunos casos, las tecnologías de almacenamiento persistente como los discos duros no estén alineados con esta constante, pues éstos presentan un rápido aumento en la capacidad de almacenamiento, pero las velocidades de acceso o transferencia de datos no ha crecido de la misma forma.

Este crecimiento exponencial de información digital y las limitaciones en transferencias de datos en las tecnologías de almacenamiento, ha permitido crear soluciones como Hadoop que permiten realizar de manera eficiente el procesamiento, la lectura y la escritura de grandes cantidades de datos en paralelo y en múltiples discos, donde los discos están ubicados en diferentes máquinas.

Hadoop tiene un componente que gestiona los archivos de gran tamaño, archivos que crecen por encima de la capacidad de almacenamiento de una única máquina física, por lo cual este componerte se encarga de dividir el archivo para distribuir las diferentes divisiones entre varias máquinas, el nombre del componente es HDFS.

HDFS o Hadoop Distributed File System es un sistema de archivos distribuido que se encarga del almacenamiento a través de una red de máquinas “tal y como lo muestra Pérez [5]”, el cual está diseñado para almacenar archivos de gran tamaño con una filosofía de escribir solo una vez y permitir múltiples lecturas, esta filosofía encaja comúnmente con una aplicación Map/Reduce o aplicaciones tipo araña web (web crawler).

El HDFS no requiere de un hardware altamente confiable sino de máquinas comunes del mercado, aunque este tipo de máquinas aumenta la probabilidad de fallo de nodo o máquina, debido a la posibilidad de que una pieza como el disco duro, la memoria o tarjetas de red se averíen, el sistema de archivos tiene la capacidad de realizar una replicación de datos con el fin de que en el caso de fallo de un nodo se utilice una copia disponible de otro nodo o máquina, evitando así la pérdida de datos y poder seguir trabajando sin interrupción perceptible para el usuario.

Al igual que en un sistema de archivos de un solo disco, los archivos en HDFS se dividen en porciones del tamaño de un bloque, que se almacenan como unidades independientes, esta abstracción de bloque es la que permite que un archivo pueda ser mayor en capacidad que cualquier unidad de disco de una sola máquina, facilitando el que se pueda almacenar un archivo en múltiples discos de la red al dividirlo en bloques. Además, los bloques encajan bien con la replicación, proporcionando tolerancia a fallos y alta disponibilidad. En el sistema de archivos cada bloque se replica en un pequeño número de máquinas separadas físicamente (normalmente tres). Permitiendo que en casos de que un bloque no está disponible sea porque está corrupto o se averió una máquina o una de sus partes principales, una copia de este bloque se puede leer desde otra ubicación de una manera transparente para el cliente.

El HDFS implementa la replicación utilizando el concepto de bloque de disco, el cual es la cantidad mínima de datos que se pueden leer o escribir en disco. En este caso tiene un bloque por defecto de 64 MB como unidad de tamaño básico para la partición de un archivo, siendo éste muy superior al de los discos. La razón de su gran tamaño es minimizar el costo de búsquedas, ya que este tamaño presenta tiempos de búsqueda de bloque en disco inferior al tiempo de transferencia de bloque desde el disco a la memoria RAM. Para mejorar la velocidad de transferencia de bloque a memoria RAM se debe realizar una disposición de los siguientes bloques del archivo en forma secuencial y no aleatoria en el disco, permitiendo por la secuencia de bloques un flujo continuo de datos hacia la memoria.

HDFS “tal y como lo muestra Cluster Informática [4]”, tiene una característica de los sistemas distribuidos contemporáneos que es la separación de datos de los metadatos, esto es con el fin de simplificar la administración del almacenamiento, ya que en este caso los bloques tienen un tamaño fijo y no almacenan información de los metadatos, lo que facilita el cálculo para determinar la capacidad de bloques por unidad de disco, sin tener que preocuparse por el espacio que genera la información de los metadatos como los permisos de creación, modificación y tiempos de acceso para los archivos, el árbol de directorios, entre otros, el cual se almacena en máquinas (nodos) separadas de los datos.

Para realizar esta separación el sistema HDFS tiene dos tipos de nodos operativos que funcionan con un patrón maestro esclavo, el maestro es el NameNode y el esclavo es el DataNode:

A. Namenode gestiona y almacena la información sobre cada archivo o metadatos, como la ubicación de los bloques que componen el archivo en el datanode, el árbol de directorios, los permisos, el nombre del archivo entre varias funciones más, se debe tener en cuenta que los metadatos son modificables y ocupan poca memoria, por consiguiente, se busca que los metadatos siempre estén en la memoria RAM para un rápido acceso y sincronización.

B. Datanode es el caballo de batalla del sistema de archivos, estos se encargan de almacenar y recuperar bloques, además, periódicamente le informan al namenode las listas de bloques que se están almacenando (sincronización).

IV. METODOLOGÍA

Este proyecto forma parte de la investigación “Analizar imágenes astronómicas en forma paralela y automática utilizando Hadoop [2]”, la cual se encuentra en proceso y consiste en una investigación aplicada, con el objetivo de facilitar el uso de herramientas de Big Data en análisis de datos científicos el doctor Miguel Ángel Aragón Calvo junto con el estudiante de Post Doctorado Nathaniel R. Stickley, han desarrollado una infraestructura de software denominada “StratOS”, basado en Mesos, el cuál resuelve los problemas que se suscitan al analizar imágenes y datos astronómicos con la tecnología común para Big Data, así como también en el manejo de un clúster de la Universidad de California.

Para el desarrollo del proyecto, se llevaron a cabo las etapas de: Análisis, Diseño, Codificación y Pruebas. A continuación se describen estas etapas.

1. Análisis

El Sistema Distribuido de Archivos de Hadoop se basa en el Sistema de Archivos de Google, es muy tolerante a errores y está diseñado para ser instalado en hardware de bajo costo. Proporciona un alto rendimiento en el acceso a datos de aplicaciones y es muy eficiente en aplicaciones con grandes conjuntos de datos. En Hadoop se ejecuta el código a través de un clúster, los datos se dividen en directorios y archivos. Los archivos se dividen en bloques de tamaño uniforme. Estos archivos se distribuyen en los distintos nodos del clúster para su transformación posterior. HDFS supervisa el proceso, los bloques se replican para manejar errores de hardware, comprobar que el código se ejecuta correctamente y enviar los datos ordenados en el nodo correspondiente.

El HDFS trabaja con una cantidad mínima de información llamada bloque que normalmente tiene un tamaño de 64 – 128 Mb debido a que se trabaja con archivos muy grandes. Su objetivo es dividir el archivo en bloques de tamaño fijo y distribuirlo en los distintos nodos del clúster.

En el sistema de archivos distribuidos de Hadoop se tiene el concepto de bloque, el cual permite dividir un archivo que se desea subir al clúster en fragmentos de tamaño de bloque. El tamaño de bloque

de un sistema de archivos es la unidad más pequeña que el sistema de archivos puede manejar y siempre se tiene ese tamaño de bloque.

2. Diseño

Para HDFS se necesita configurar el archivo hdfs-site.xml. Cuando se modifica el tamaño de bloque dentro de este archivo, todos los archivos subsecuentes que ingresen tendrán ese tamaño de bloque. Los archivos que ya se encontraban dentro permanecen con su tamaño de bloque anterior.

Se abre la terminal para poder ingresar al archivo ~/.bashrc y poder anexar las variables de entorno.

```
$cd /
$sudo su
password:
#cd home/
# gedit <nombredeusuario>/.bashrc
```

Dentro del archivo .bashrc se agrega la siguiente línea de código al final. Sin modificar nada del código superior, y se guarda.

```
export HADOOP_HOME=/usr/local/hadoop
```

Se guarda y actualiza el código con el siguiente comando en la terminal.

```
# source ~/.bashrc
Ahora se hace el mismo proceso pero ahora con él .bashrc de root.
 # gedit /root/.bashrc
Al final del archivo se ingresan los siguientes comandos igual que en el archivo anterior.
export HADOOP_HOME=/usr/local/hadoop
export PATH=$PATH:$HADOOP_HOME/bin
export PATH=$PATH:$HADOOP_HOME/sbin
export HADOOP_MAPRED_HOME=${HADOOP_HOME}
export HADOOP_COMMON_HOME=${HADOOP_HOME}
export HADOOP_HDFS_HOME=${HADOOP_HOME}
export YARN_HOME=${HADOOP_HOME}
```

Se guarda y actualiza el código con el siguiente comando en la terminal.

```
#source ~/.bashrc
```

Antes de seguir adelante, se necesita comprobar que Hadoop está trabajando bien. Sólo hay que utilizar el comando siguiente:

```
 $ hadoop version
```

Hadoop está configurado para que se ejecute en un modo distribuido en una sola máquina.

3. Codificación

Para llevar a cabo esta tarea se creó un software denominado “desfragStratOs.sh” el cual contiene las librerías necesarias para que puedan modificar los tamaños de bloque de archivos para almacenarlos en el clúster. Este software contiene un menú principal:

- Fragmentar
- Desfragmentar
- Visualizar carpetas HDFS
- Crear carpeta de HDFS
- Eliminar carpeta HDFS
- Salir

En la fragmentación se solicita la carpeta para realizar la fragmentación del archivo, se ingresa la ruta del archivo a fragmentar, el nombre del archivo con extensión y el tipo de fragmentación, ya sea en Mb, Gb, Tb, Pb ó Exb.

En la desfragmentación se solicita la ruta donde se encuentra el archivo fragmentado, además se solicita la ruta donde va a descargarse el archivo.

Lo mismo sucede para visualizar carpetas HDFS, solicita el nombre de la carpeta a visualizar.

Para crear carpeta HDFS se solicita el nombre de la carpeta a crear y la ruta si se desea crear dentro de otra carpeta.

Para eliminar carpeta, de igual manera se solicita el nombre de la carpeta a eliminar y la ruta específica de la carpeta.

Para la opción de salir solo se elige la opción 6 para que el software termine su ejecución.

4. Pruebas y Resultados

Para iniciar el script primeramente se deben iniciar los servicios del clúster con el siguiente comando:

```
$ sudo /usr/local/hadoop/sbin/start-all.sh
```

Después se tienen que asignar permisos especiales para su ejecución con el comando siguiente:

```
$ chmod +x ./desfragStratOS.sh
```

Después de esto se debe iniciar el software “desfragStratOs.sh”

El menú principal muestra 6 opciones, como se visualiza en la figura 1.

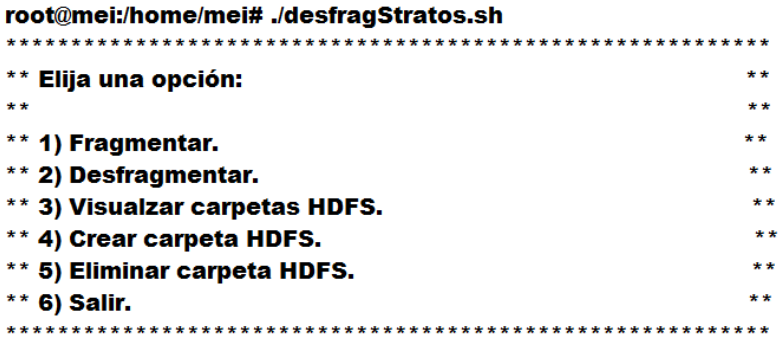

Fig. 1. Menú principal del software desfragStratOs en ejecución.

Dentro de la opción 1, se puede hacer uso de la herramienta fragmentación en base a HDFS lo cual solo se tiene que seguir el menú guiado completando las opciones indicadas por el menú de dicha opción, como se muestra en la figura 2.

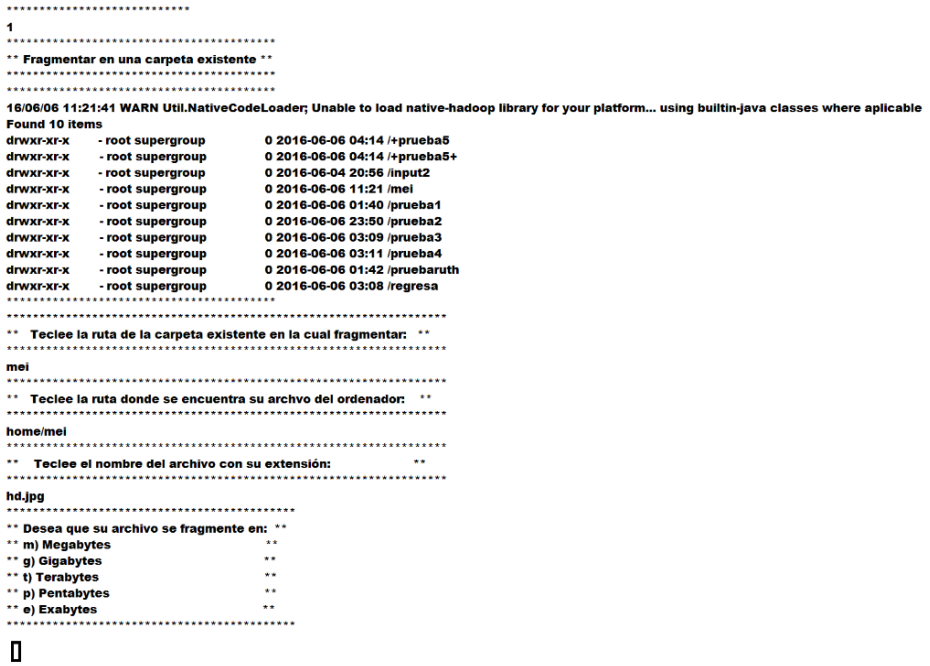

Fig. 2. Fragmentación de archivos.

Dentro de la opción 2, se puede hacer uso de la herramienta desfragmentación en base a HDFS lo cual solo se tiene que seguir el menú guiado completando las opciones indicadas por el menú de dicha opción como se muestra en la figura 3.

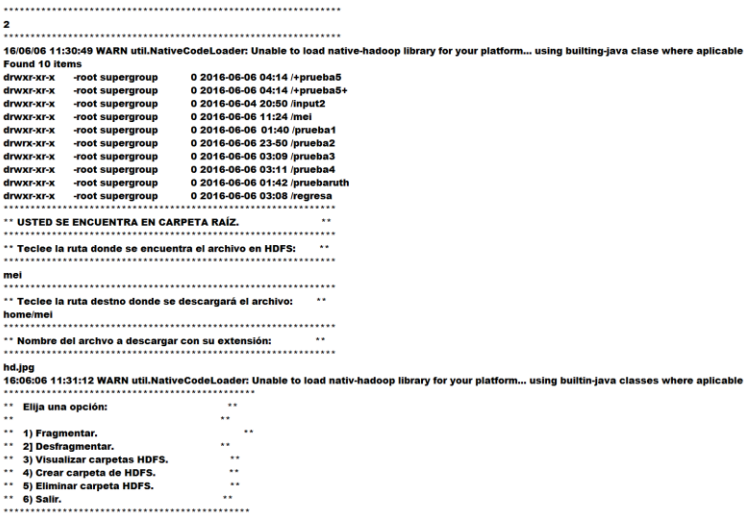

Fig. 3. Desfragmentación de archivos.

Dentro de la opción 3, puede hacer uso de la herramienta Visualizar Carpetas en base a HDFS lo cual solo se tiene que seguir el menú guiado completando las opciones indicadas por el menú de dicha opción como se muestra en la figura 4.

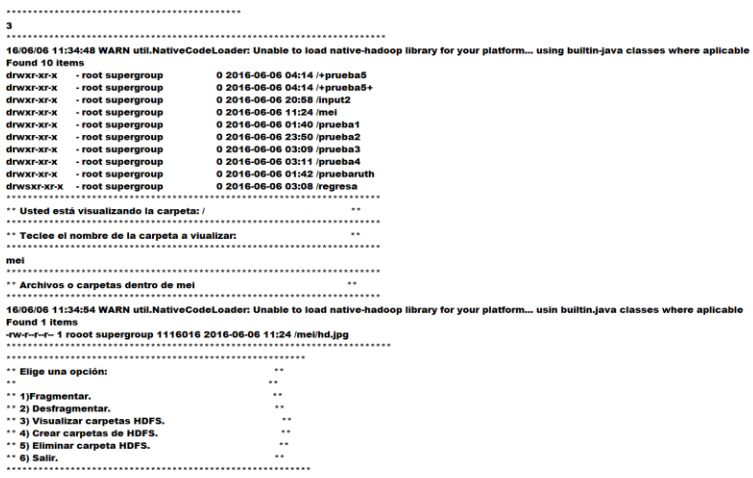

Fig. 4. Visualizar Carpetas.

Dentro de la opción 4, se puede hacer uso de la herramienta Crear Carpetas en base a HDFS lo cual solo se tiene que seguir el menú guiado completando las opciones indicadas por el menú de dicha opción como se muestra en la figura 5.

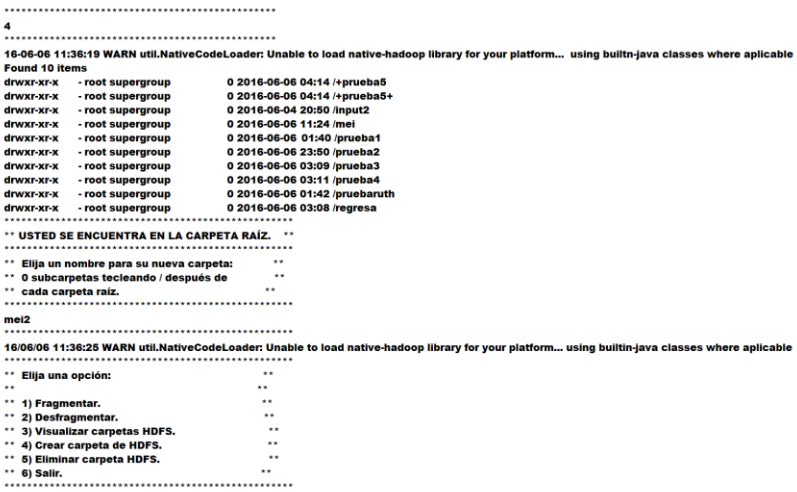

Fig. 5. Crear carpetas.

Dentro de la opción 5, se puede hacer uso de la herramienta Eliminar Carpetas en base a HDFS lo cual solo se tiene que seguir el menú guiado completando las opciones indicadas por el menú de dicha opción como se muestra en la figura 6.

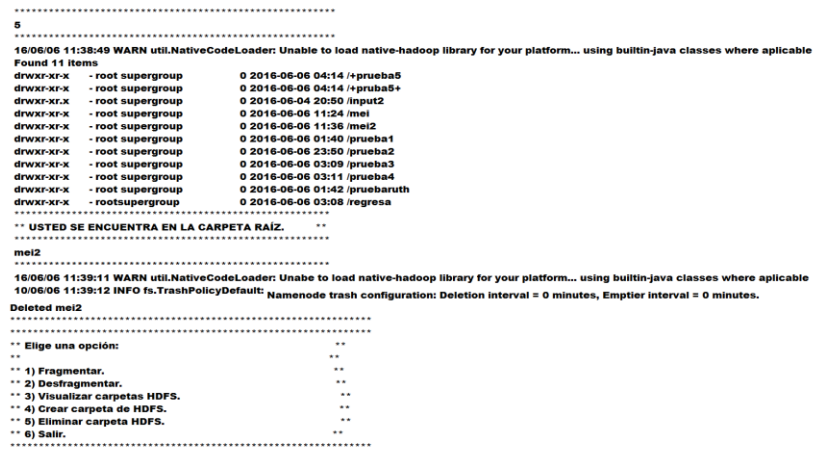

Fig. 6. Eliminar carpetas.

Dentro de la opción 6, puede hacer uso de la herramienta Salir en base a HDFS lo cual sólo termina la ejecución de dicho software, como se muestra en la figura 7.

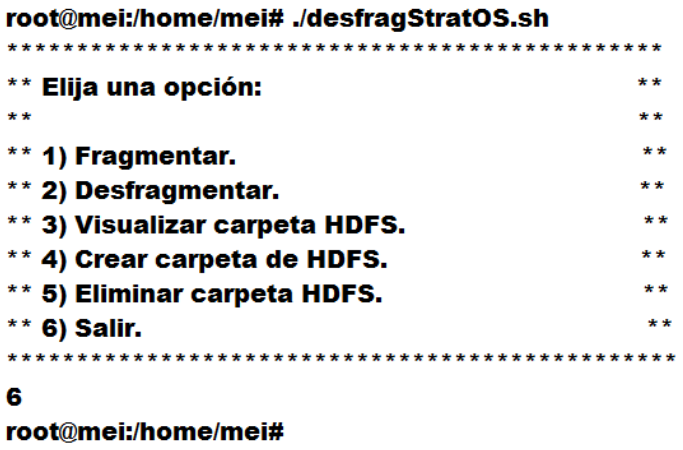

Fig. 7. Salir.

CONCLUSIONES

Hablar de Hadoop es hablar de un gran ecosistema de archivos distribuidos, donde se procesan grandes cantidades de datos que siguen creciendo de manera exponencial. Los archivos de entrada de esos datos se dividen en bloques de tamaño fijo que se almacenan en forma distribuida en un clúster. Un archivo está compuesto de uno o varios bloques, HDFS trata de colocar cada bloque en nodos de datos separados, distribuyendo la información a lo largo del clúster.

El software desarrollado está diseñado para la fragmentación de archivos del clúster de la Universidad de California Riverside; en el cual consiste en cambiar de forma manual la configuración que tiene por default el sistema de archivos de bloques de HDFS, adecuándose a las necesidades como cliente. El software cuenta con un menú desplegable modo consola en el cual con una serie de instrucciones será de gran facilidad hacer uso de este.

REFERENCIAS

1. Barranco, R. IT Specialist for Information Management, IBM Software Group México, ¿QUE ES BIG DATA?, Articulo Web, Obtenido el 5 de Diciembre de 2016 de: https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/
2. Nathaniel R, Aragón M. StratOS: A Big Data Framework for Scientific Computing, Articulo Web Obtenido el 5 de Diciembre de 2016 de: https://arxiv.org/pdf/1503.02233.pdf
3. Wikipedia, Hadoop, Articulo web https://es.wikipedia.org/wiki/Hadoop

4. Wikipedia, Cluster Informática, Artículo Web, Obtenido el 5 de Diciembre de 2016 de: https://es.wikipedia.org/wiki/Cl%C3%BAster_(inform%C3%A1tica).
5. Pérez M. BIG DATA Técnicas, herramientas y aplicaciones, Artículo web Obtenido el 5 de Diciembre de 2016 de: http://mx.casadellibro.com/libro-big-data-tecnicas-herramientas-yaplicaciones/9788494305559/2533124