Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Escuela de Ciencias y Sistemas

Seminario de Sistemas 2 Sección A

Primer Semestre 2023



**Práctica 2**

**Apache - Hadoop**

Nombre: William Alejandro Borrayo Alarcón

Carné: 201909103

Contenido

[Procedimiento realizado 2](#_Toc130989749)

[Conclusiones sobre resultados de archivos 2](#_Toc130989750)

[Análisis acerca de los resultados de cada archivo 2](#_Toc130989751)

[Conclusiones acerca de los resultados de cada archivo 3](#_Toc130989752)

[Conclusiones acerca del uso de Hadoop en BigData. 3](#_Toc130989753)

[Bibliografía 3](#_Toc130989754)

# Procedimiento realizado

Para realizar este procedimiento es necesario tener Docker instalado.

## Preparación de Hadoop en contenedor

1. Descarga de imagen de Hadoop: Si vemos el mensaje ‘Already exists’ es porque ya la tenemos descargada.

Texto

Descripción generada automáticamente

1. Creación del contenedor (ejecución de la imagen descargada): En el comando se utiliza un parámetro para darle un nombre específico al contenedor: ‘hadoop’. Si se creó correctamente veremos la consola de esta manera:

Texto

Descripción generada automáticamente

También en localhost:50007 veremos esto:

Captura de pantalla de computadora

Descripción generada automáticamente

En utilities, browser directory se mostrará así:

Captura de pantalla de un celular

Descripción generada automáticamente

1. Luego de esto se crea una carpeta con el nombre ‘201909103\_Practica2’ en el contenedor, aquí se guardarán algunos archivos necesarios (archivos de entrada y el WordCount.java). Para la copia de los archivos se utiliza el comando: docker cp <Ruta en computadora> hadoop:/201909103\_Practica2 en una consola aparte.

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

1. Se exportan las variables HADOOP\_HOME y CLASSPATH de la siguiente forma:

Texto

Descripción generada automáticamente

1. Ejecutamos el archivo .java, si se ejecuta bien deberá crear 3 archivos .class como se muestra en la imagen, la advertencia que se muestra no representa mayor problema:

Texto

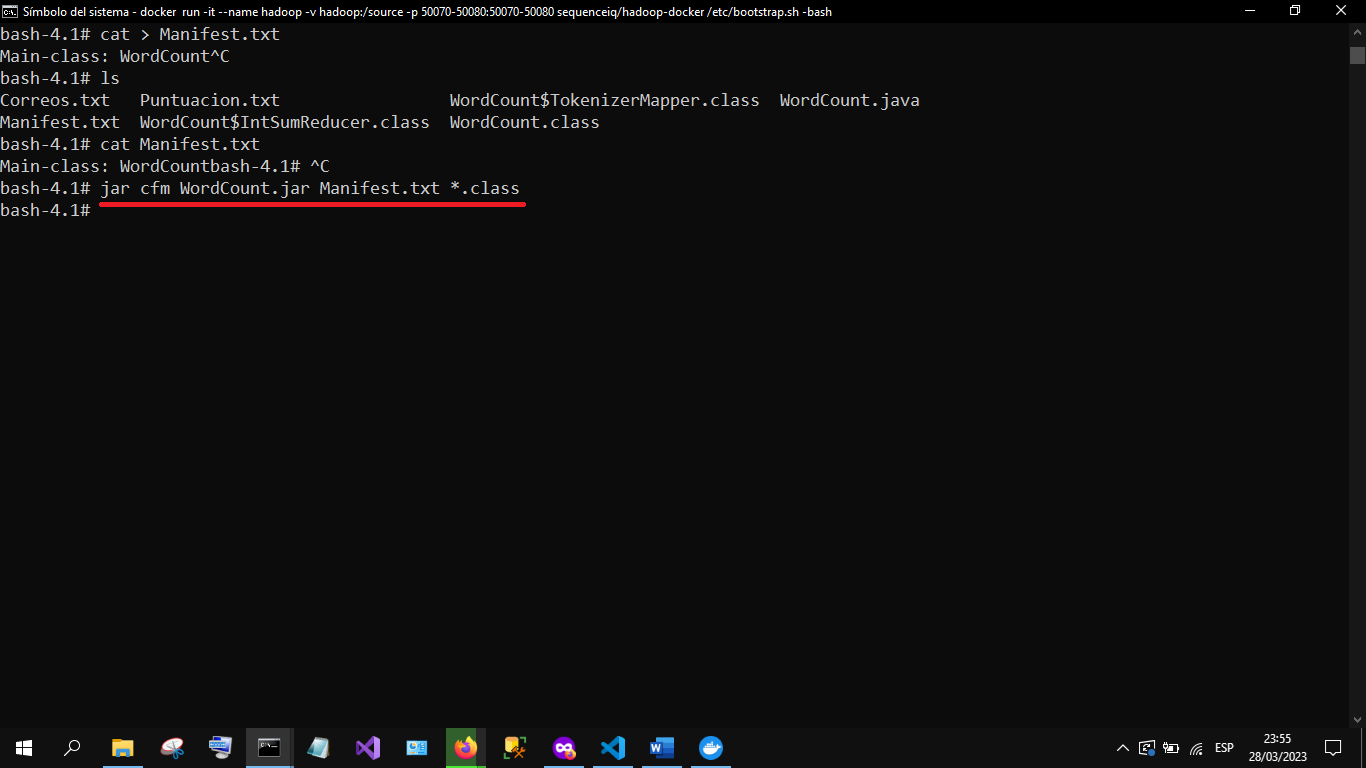
Descripción generada automáticamente

1. Se crea el archivo Manifest.txt con el contenido “Main-class: WordCount”. Se puede utilizar cat y luego de escribir el contenido la combinación Ctrl + D.

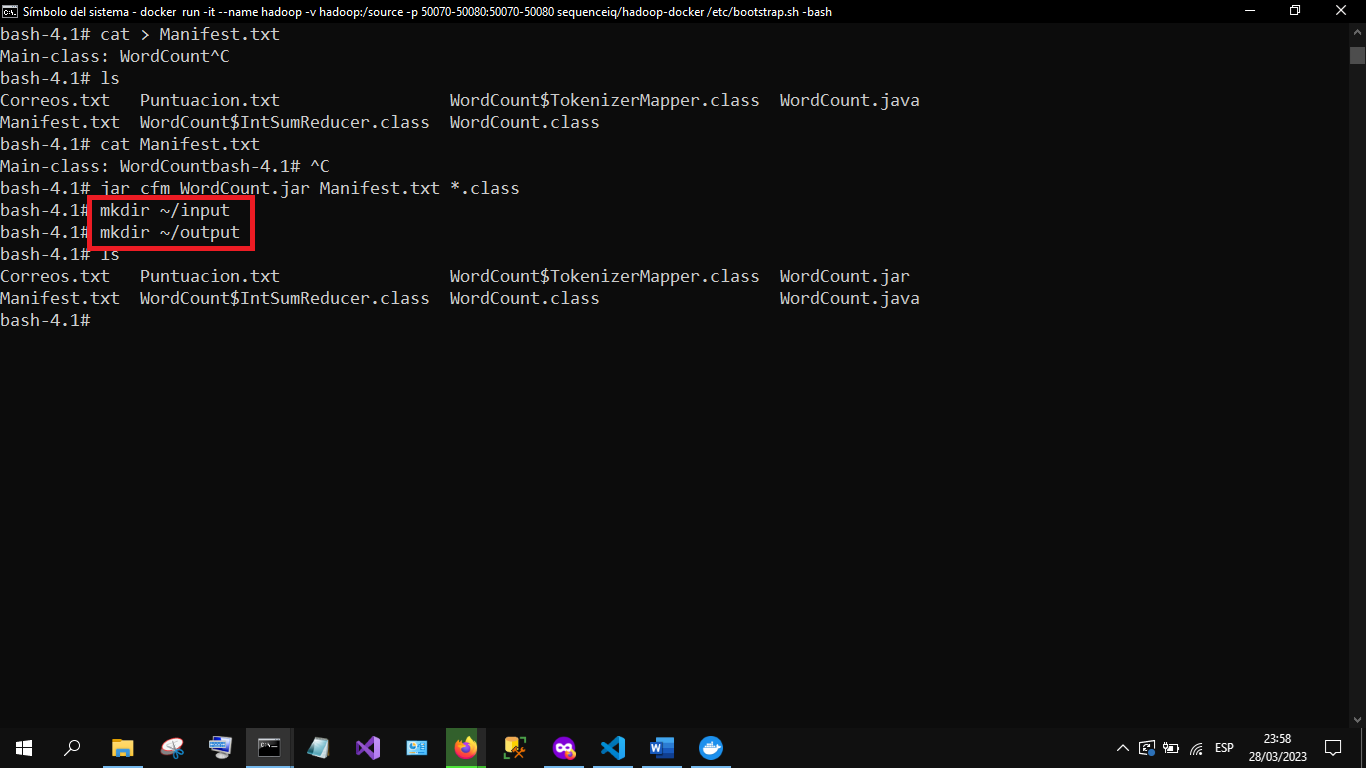
Texto

Descripción generada automáticamente

1. Se debe crear un archivo .jar utilizando el archivo Manifest.txt previamente creado:



1. Se crean las carpetas ‘input’ y ‘output’ en el home del usuario root:



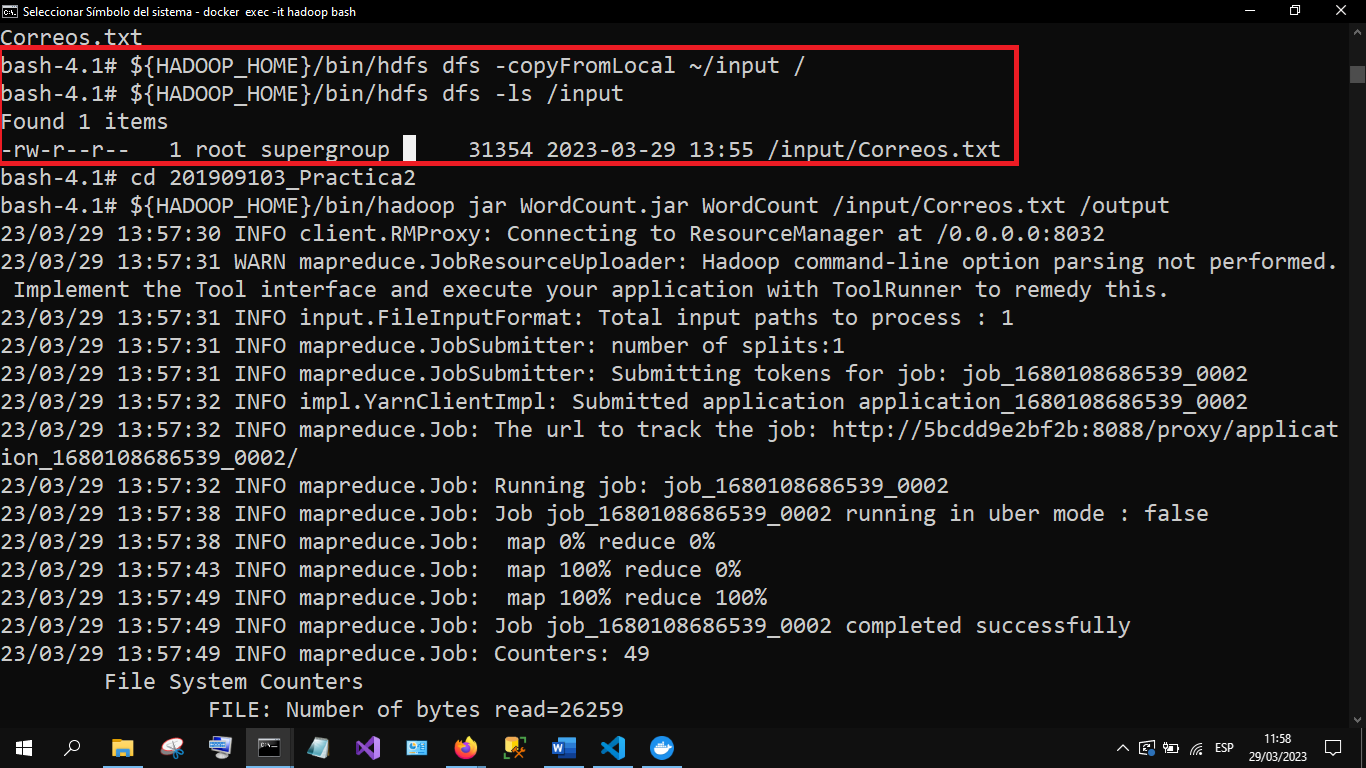
## Procedimiento para el conteo de palabras en el archivo Correos.txt

1. Se copia el archivo de entrada hacia la carpeta ‘input’ creada en el paso anterior:

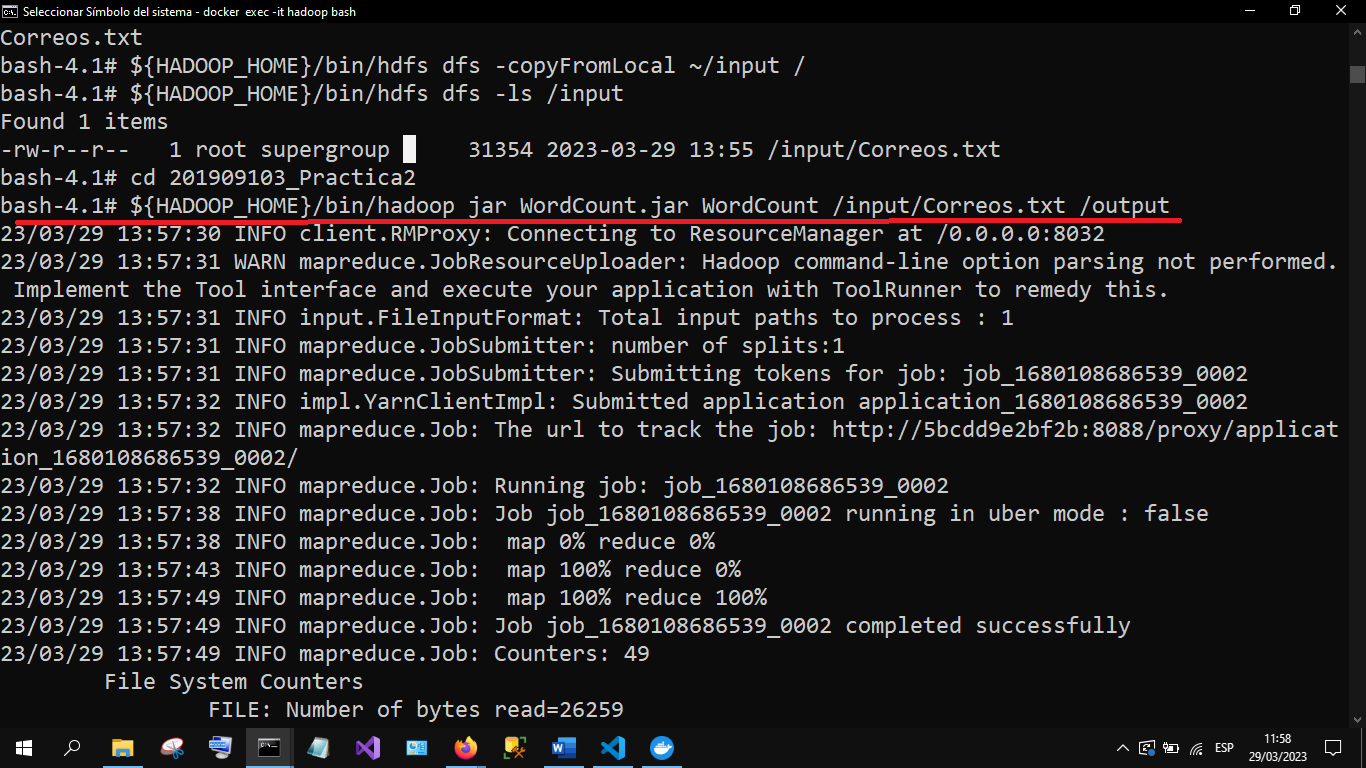
Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Se copia el archivo de entrada hacia el sistema de archivos de hadoop. Y se verifica la correcta copia del archivo:



1. Para iniciar con el conteo de palabras es necesario estar en la carpeta que tiene el archivo WordCount.java (201909103\_Practica2). Desde ahí se ejecuta el comando: ${HADOOP\_HOME}/bin/hadoop jar WordCount.jar WordCount /input/Correos.txt /output:



Se iniciará el proceso de conteo, al final veremos detalles como la cantidad de bytes que fueron escritos, leídos, si hubieron errores u otras cosas:

Texto

Descripción generada automáticamente

1. Podremos ver los archivos que fueron generados con el primer comando, el segundo hará que se muestre el contenido del archivo de resultado:

Texto

Descripción generada automáticamente

1. Renombramos el archivo a ‘Resultado\_Correos.txt’ y verificamos que tenga el mismo contenido:

Texto

Descripción generada automáticamente

1. Copiamos el archivo de resultado a la carpeta output del home del usuario root, luego hacia la carpeta Salidas (es necesario crearla en el directorio inicial del contenedor), para su fácil extracción del contenedor.

Texto

Descripción generada automáticamente

1. Finalmente pasamos el archivo de resultado desde el contenedor hacia nuestra computadora, esto desde una nueva consola y hacia el directorio de nuestra preferencia.

Captura de pantalla de un celular

Descripción generada automáticamente

Podremos comprobar que el archivo ya está en nuestra computadora y con el mismo contenido:

Captura de pantalla de computadora

Descripción generada automáticamente

Captura de pantalla de computadora

Descripción generada automáticamente

# Conclusiones sobre resultados de archivos

* El sistema experto es capaz de tomar la decisión de indicar qué hoteles pueden ser la mejor opción para el hospedaje de los clientes según sus criterios y preferencias.
* Se ofrecen diferentes criterios para el inicio de búsquedas con las preguntas más adecuadas.
* Al finalizar cada camino de preguntas se muestran las mejores opciones de hoteles para el hospedaje, ayudando a los clientes a tomar buenas decisiones.
* Los reportes muestran presentablemente información obtenida desde la base de conocimientos.

# Análisis acerca de los resultados de cada archivo

# Conclusiones acerca de los resultados de cada archivo

# Conclusiones acerca del uso de Hadoop en BigData.