

Tipo : Enunciado de laboratorio
Capítulo : Componentes de Hadoop
Duración : 30 minutos

I. OBJETIVO

Conocer cómo implementar Flume, realizando un ejemplo de envío de archivo de un directorio a HDFS.

II. REQUISITOS

Los siguientes elementos de software son necesarios para la realización del laboratorio:

- Virtual Box
- Hortonworks Sandbox for HDP
- Putty
- Winscp

III. EJECUCIÓN DEL LABORATORIO

1. Ingresar al servidor Hortonworks, para esto deberá ingresar a la siguiente dirección:
<http://127.0.0.1:8080>
2. Ir al menú de Canvas
3. Seleccionar la opción File View
4. Luego de esto navegar hasta seleccionar la carpeta /user/maria_dev, luego de esto dar click en el botón New Folder, crear el directorio flume, clic al botón +Add.
5. Luego de esto ingresar por putty al servidor,
6. Creamos una carpeta con el nombre spool
7. Creamos el archivo flumelogs.conf, con la configuración
8. Ir al directorio de configuración de Flume
9. Ejecutar el agente de flume, indicando el archivo de configuración flumelogs.conf, esperar que finalice la inicialización (Mensaje: Starting accept handler)
10. Ingresar en una segunda ventana de putty al servidor, volver a realizar el paso 5.
11. Ejecutar el comando
12. Revisar el contenido de la carpeta HDFS /user/maria_dev/flume. ¿Qué ha pasado?
13. Realice la copia de otro, utilice el comando del paso 11.
14. Para salir de flume, recuerde presionar CTRL+C, en la pantalla que se ejecuta el agente.