

Como parte de un proceso de transformación, nuestra empresa requiere que definamos los siguientes flujos de datos:

1. Crear un Pipeline donde la fuente sean los ficheros que sean colgados en un directorio del sistema local en la ruta `/home/bigdata/flume/source` (NOTA: Si la ruta no existe crearla), de tal forma que una vez que sea insertado un fichero en dicho directorio se desencadene el proceso (spooldir). Los datos ha de almacenarse en HDFS en la ruta `/home/bigdata/pec2/ex1` y en formato de texto.
2. Crear un Pipeline donde la fuente sean los ficheros que sean colgados en un directorio del sistema local en la ruta `/home/bigdata/flume/source`, de tal forma que una vez que sea insertado un fichero en dicho directorio se desencadene el proceso (fuente spooldir). Los datos ha de almacenarse en HDFS en la ruta `/home/bigdata/pec2/ex2` en formato de texto en una estructura de carpetas de acuerdo a la fecha y hora a la que han llegado por ejemplo `/home/bigdata/pec2/ex2/18-04-15/1020` (significa que el fichero fue colocado en la carpeta el 15/04/2018 a las 10:20 aproximadamente).
3. Crear un pipeline donde la fuente sean los ficheros que sean colgados en un directorio del sistema local en la ruta `/home/bigdata/flume/source`, de tal forma que una vez que sea insertado un fichero en dicho directorio se desencadene el proceso (spooldir). Los ficheros que serán volcados al sistema local tendrán el siguiente formato:
 - 1, José Antonio, Ingeniero
 - 2, María Antonia, Medico
 - 3, María Fernanda, Ingeniero
 - 1, Pedro, Abogado

Los datos ha de almacenarse en HDFS en 3 rutas distintas
`/home/bigdata/pec2/ex3/tipo1`, `/home/bigdata/pec2/ex3/tipo2`,
`/home/bigdata/pec2/ex3/tipo3` De acuerdo al valor entero con el que inicia cada línea dentro de los ficheros el cual no es mas que el código del departamento al que pertenece. Los datos ha de guardarse en HDFS en formato de texto. Es imprescindible que los datos de las personas que son departamento "1" estén únicamente en la ruta `/home/bigdata/pec2/ex3/tipo1` así como las personas del departamento "2" estén en la ruta `/home/bigdata/pec2/ex3/tipo2` y así sucesivamente.

La ponderación de las preguntas es 3, 4, 4 haciendo un total de 10 puntos. Dentro de la puntuación de cada pregunta, valdrá 1 punto el que los ficheros sean almacenados en formato de texto.

Para la solución enviar el fichero el fichero de configuración (agente).