

Practica de Hadoop MapReduce

Para el desarrollo de la práctica hemos necesitado 3 clases:

- AnalisisLogsDriver.java
- MapperAnalisis.java
- ReducerAnalisis.java

Analizando el ejercicio la técnica que se va usar es la de In mapper combiner esto nos permite realizar las operaciones en el mapper y pasar la mínima información posible al reducer.

Se usaran writables básicos (Text).

AnalisisLogsDriver

Clase principal donde definimos el job que va ser lanzado. En el que se define el mapper y el reducer.

Una vez lanzado el JOB pasaremos a escribir los contadores de las aplicaciones para que sean mostrados por consola.

MapperAnalisis

Es donde se realiza realmente el cálculo del ejercicio usando un in mapper combiner, lo que conseguimos es realizar el conteo de los componentes y agruparlos.

Usamos una clave compuesta para almacenarlo en el HashMap de:

“DD/MM/AAAA-HH /t Componente”

Con esto podemos ir incrementando los valores del componente si la key coincide en la HashMap, ese incremento se almacena como value de la HasMap.

Luego en la salida descomponemos la clave creada y sacamos como valor el componente : nº de veces que se repite.

Definimos en el mapper además los counter para que se vaya incrementando cada vez que aparezca el componente.

Usamos un método setup() para definir la HashMap , y el método cleanup para liberar y reconstruir la info como hemos mencionado anteriormente.

ReducerAnalysis

El reducir aquí solo realiza la función de recorrer los valores e ir juntándolos en un String para luego poder enviarlo a la salida, para conseguir el formato especificado.

La entrada (por fecha)

Key: DD/MM/AAAA-HH

Values: (APP:nº repeticiones,APP:nº Repeticiones.....)

Opciones que se valoraron:

Añadir un writable propio para la parte de la fecha donde asumiera la clave compuesta en vez de usar un writable de TEXT; hice un par de pruebas y vi que realmente no me aportaba nada nuevo o al menos que yo viera.