

PRÁCTICA PIG Local

Buscamos un archivo para analizar. Vamos a utilizar el siguiente:

- `sudo head /var/log/messages`

Lo copiamos a nuestra carpeta local:

- `sudo cp /var/log/messages .`

Comprobamos sus permisos y vemos que no tiene permisos de ejecución:

- `ls -la`
- `ls -la messages`

Le damos todos los permisos:

- `sudo chmod 777 messages`
- `ls -la messages`

Ejecutamos Pig para entrar en el Grunt y trabajar con comandos Pig:

- `pig -x local`

Ahora cargamos el contenido del fichero en la variable messages:

- `messages = LOAD 'messages';`

Filtramos los warnings de ese fichero haciendo uso de expresiones regulares:

- `warns = FILTER messages BY $0 MATCHES '.*WARN+.*';`

Volcamos el resultado a un fichero:

- `DUMP warns;`

Output(s):

Successfully stored records in: "file:/tmp/temp-1781065794/tmp-1068339407"

Salimos de la terminal de Pig:

- `ctrl + c`

Vamos a la carpeta donde está ubicado el fichero de salida:

- `cd /tmp/temp-1781065794/tmp-1068339407`
- `ls`

Mostramos el contenido del fichero de salida (Bag) por pantalla:

- `cat part*`

```
-----  
LOct 23 16:29:11 quickstart console-kit-daemon[23671]: WARNING: no  
sender#012[cloudera@quickstart tmp-1068339407]$  
-----
```