

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

I. OBJETIVO

Conocer cómo implementar los componentes de Hadoop.

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. Cargar los archivos de ejemplo.
3. Crear el script en Pig.
4. Definir la relación.
5. Guardar y ejecutar el script
6. Definir la relación con un esquema.
7. Defina una relación de una relación existente.
8. Ver los datos
9. Seleccionar las columnas específicas de una relación.
10. Seleccionar las columnas específicas de una relación.
11. Realizar una relación entre 2 relaciones
12. Realizar una relación entre 2 relaciones
13. Filtre y agrupe los datos usando "Group By"