Session7 Assignment 1- answers

1. Why Map-reduce program is needed in Pig Programming?

Ans : Pig progrmming internally creates map-reduce programs for execution.

2. What are advantages of pig over MapReduce?

Ans : 1. Do not require programming interms of map and reduce tasks, especially Java programming.

2. Provides High level scipring language that can be used by other than programmers.

3. Less number of code lines compared to MapReduce.

3. What is pig engine and what is its importance?

Ans: Pig engine takes the Pig Latin scripts written by users, parses them, optimizes them and then executes them as a series of MapReduce jobs on a Hadoop Cluster.

4. What are the modes of Pig execution?

Ans : MapReduce/Hadoop Mode and Local mode

5. What is grunt shell in Pig?

Ans : Grunt is an interactive shell for executing Pig Commands

6. What are the features of Pig Latin language?

Ans: 1. It is a data flow programming language, consisting of collection od operational/transformation statements.

2. Unlike RDBMS, Pig is more relaxed about the data processesing. For example user can define optional schema at run-time.

3. However pig do not support low-latency random read and write/update queries.

4. Keyword are not case sensitive, but all other identifiers are case sensitive.

7. Is Pig latin commands case sensitive?

And: No

8. What is a data flow language?

Ans: Is a programming language where the program/application can be represented as directed graph similar to dataflow diagram, with set of nodes(blocks),each node representing a source /processing block/function.