**ASSIGNMENT-6.2**

1. Why MapReduce program is needed in Pig Programming?

* Pig's infrastructure layer consists of a compiler that produces sequences of Map-Reduce programs, for which large-scale parallel implementations already exist.
* Users start with a Pig script or the Pig command line (called Grunt).
* Pig parses, compiles, optimizes, and fires MapReduce statements.
* MapReduce accesses HDFS and returns the results.
* This is done because ***Hadoop understands only mapreduce programming***.

1. What are advantages of Pig over MapReduce?

* Hadoop MapReduce is a compiled language whereas Apache Pig is a ***scripting language.***
* Pig provides ***higher level of abstraction*** whereas Hadoop MapReduce provides low level of abstraction.
* Hadoop MapReduce requires more lines of code when compared to Pig.
* [Hadoop](https://www.dezyre.com/Hadoop-Training-online/19)MapReduce requires more development effort than Pig.
* Performing a Join operation in Apache Pig is pretty simple.
* There is no need for compilation. On execution, every Apache Pig operator is converted internally into a MapReduce job.

1. What is Pig engine and what is its importance?

***Pig engine:***

 Apache Pig has a component known as Pig Engine that accepts the Pig Latin scripts as input and converts those scripts into MapReduce jobs.

***Importance:***

* It acts as interpreter between Pig Latin script and MapReduce Jobs.
* It provides an environment to execute Pig scripts into series of mapreduce jobs.

1. What are the modes of Pig execution?

Pig has two execution modes:

* ***Local Mode*** - To run Pig in local mode, you need access to a single machine; all files are installed and run using your local host and file system. Specify local mode using pig -x local.
* ***Mapreduce Mode*** - To run Pig in mapreduce mode, you need access to a Hadoop cluster and HDFS installation. Mapreduce mode is the default mode; you can, but don't need to, specify it using pig OR pig -x mapreduce.

1. What is Grunt Shell in Pig?

* The Grunt shell of Apache Pig is mainly used to write Pig Latin scripts.
* We can invoke the Grunt shell in a desired mode (local/MapReduce).
* After invoking the Grunt shell, we can run your Pig scripts in the shell.
* Grunt is pig's interactive shell.
* It is started when no file is specified for pig to run.

1. What are the features of Pig Latin language?

* It providesRich set of operators
* Pig Latin is similar to SQL and it is a Ease of programming
* The tasks in Apache Pig optimize their execution automatically
* Using the existing operators, users can develop their own functions
* Pig provides the facility to create User-defined Functions
* Apache Pig analyzes all kinds of data, both structured as well as unstructured.

1. Is Pig Latin commands case sensitive?

***Pig Latin cannot decide whether it is case-sensitive or not.***

* Keywords in Pig Latin are not case-sensitive; for example, LOAD is equivalent to load.
* Relation and field names are case -sensitive. So A = load 'foo'; is not equivalent to a = load 'foo’.
* User-defined Functions names are also case-sensitive, thus COUNT is not the same User-defined functions as count.

1. What is a data flow language?

* In a dataflow language, a stream of data which is passed from instruction to instruction to be processed.
* Conditional execution jumps and procedure calls route the data to different instruction.
* Apache Pig is a tool/platform which is used to analyze larger sets of data representing them as data flows.
* Pig can efficiently perform those jumps, loops and process in an efficient manner.