1. Why MapReduce program is needed in Pig Programming?

Ans –

Pig is not a separate technique by which one can process the large data it’s just the toll by which one can produce mapreduce programs. It’s just a higher abstraction of the map reduce. Pig is tool which has a scripting language called as the pig latin which is given to the pig engine with the help of the shell called as the grunt. The pig engine produces an optimized code according the scripting language code.

Hence MapReduce program is needed in Pig Programming.

It is also true that pig cannot always used for getting the map reduce programs. Sometimes writing the map reduce programs is better than that of the Pig. Pig is basically used by the people who are no programming background.

1. What are advantages of pig over MapReduce?

Ans—

PIg just a higher abstraction of the map reduce

1. Less complex codes to type
2. Easy to learn pig latin than that of Java.
3. Many built in functions.
4. People of non-programming background an also use pig easily.
5. Map reduce takes more time construct the code.
6. 1/20 lines of code are needed than that map reduce.
7. What is pig engine and what is its importance?

Ans—

Pig Engine is where the pig latin scripted code is passed with the help of the grunt shell.

Once the scripted code is passed to engine the parser checks the syntax and generates the DAG(Direct Acyclic Graph). In the DAG, the logical operators of the script are represented as the nodes and the data flows are represented as edges.

The optimizer in the engine optimizes the code and then sends it to the complier where we get the map reduce program which can be sent to Hadoop for processing.

1. What are the modes of Pig execution?

Ans—

There are two modes of pig execution

1. Local Mode
2. HDFS Mode

Local Mode - In this there is no need of HDFS, only the files form local file system are accessible. This mode is used for the testing purpose only.

HDFS Mode -

HDFS mode is where we load or process the data that exists in the Hadoop File System (HDFS) using Apache Pig. In this mode, whenever we execute the Pig Latin statements to process the data, a MapReduce job is invoked in the back-end to perform a particular operation on the data that exists in the HDFS.

1. What is grunt shell in Pig?

Ans –

Grunt shell is by which one can communicate with the pig engine. This is place one can write the pig Latin script language and grunt shell communicates with system kernel to give these coded lies to the pig engine in the language it understands.

1. What are the features of Pig Latin language?

Ans—

1. Easy to learn
2. Considered to map reduce programs less complex.
3. Lesser lines of code
4. More built in functions
5. Is Pig latin commands case sensitive?

Ans—

Pig Latin is not case sensitive language.

Only the relation names and the file name should be case sensitive.

1. What is a data flow language?

Ans—

This type of programming language models a program as a [directed graph](https://en.wikipedia.org/wiki/Directed_graph) of the data flowing between operations, thus implementing [dataflow](https://en.wikipedia.org/wiki/Dataflow) principles and architecture. This different than that of the sequential execution the data. In contrast, dataflow programming emphasizes the movement of data and models programs as a series of connections.