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

*Pig Latin queries are converted to Map and Reduce jobs and hence they take advantage of parallel processing.*

*It has the ability to perform computations which cannot be done by MapReduce.*

1. What are advantages of pig over MapReduce?

*Pig has the ability to perform computations which can not be done by MapReduce.*

*Pig Latin uses a lot fewer lines of code than the Java MapReduce script.*

*The Pig Latin script is easier to read for someone without a Java background.*

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

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

*Pig Engine converts the pig latin operators or transformations into a series of mapreduce jobs.*

*Pig Engine Parses, compiles Pig Latin scripts into MapReduce jobs run on top of Hadoop.*

1. What are the modes of Pig execution?

*Apache Pig runs in two modes, namely, Local Mode and HDFS mode.*

1. What is grunt shell in Pig?

*The Grunt shell of Apache Pig is mainly used to write Pig Latin scripts. Prior to that, we can invoke any shell commands using sh and fs.*

*Interactive Shell for executing Pig Commands.*

*It is used when script file is not provided.*

*It can execute scripts from Grunt via run or exec commands*

1. What are the features of Pig Latin language?

*Pig Latin is declarative, SQL-like language; the high level language interface for Hadoop.*

1. Is Pig latin commands case sensitive?

*The names (aliases) of relations and fields are case sensitive. The names of Pig Latin functions are case sensitive. The names of parameters (see Parameter Substitution) and all other Pig Latin keywords are case insensitive.*

*The names (aliases) of relations and fields are case sensitive. The names of Pig Latin functions are case sensitive. The names of parameters (see Parameter Substitution) and all other Pig Latin keywords are case insensitive.*

1. What is a data flow language?

*In a dataflow language, we have 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 instructions. Dataflow programming languages are ones that focus on the state of the program and cause operations to occur according to any change in the state. Dataflow programming languages are inherently parallel, because the operations rely on inputs that when met will cause the operation to execute.*