## What is Hive?

* A data warehousing solution for Hadoop
* It uses HDFS as underlying data store
* It translate the SQL queries to MapReduce . Creates java process to distribute the data
* The meta data of file goes to the meta storex
* HiveQery language is used to analyze data in Hive
* Used to analyze the data in HDFS
* It cannot be used form real-time queries because it will take some time for the query to return. So its mainly used for batch oriented applications
* Supports basic indexes
* Mainly used for DBA or Data Analysts
* Mainly used for structured data
* Recommended for people familiar with SQL
* Has a provision for partition which can be used to process a sub set of data by date or alphabetical order

## What is HBase?

* It’s a column oriented storage
* It’s a key value store
* Used for quick look up
* Used for real-time query
* Used to store clicks on a website by id. Easy to retrieve
* Just like HDFS, HBase has a Master Node (It manages the cluster) and many Region Nodes (Stores portions of the table and perform the work on the data)

## What is Pig?

* Apache Pig is a high-level platform for creating programs that run-on Apache Hadoop.
* Tools to process the data stored in HDFS
* Used in data analysis
* Mainly used for Data Scientist and Programmers
* Used in semi-structured or unstructured data (hence the name Pig which eats anything)
* Used for programing purposes
* Recommended for people with scripting knowledge like Python
* Does not have provision for portion, but this might be achieved by filter

## What is Impala?

* It’s a Real time query engine. Its not based on MapReduce
* Process SQL queries directly on the distributed file system
* It does not use Java, so no JVM
* It’s a native query engine in the HDFS

## What is Elasticsearch ?

* Open source search engine written on top of Apache Lucene search library
* Designed to take data from any source, analyze it and search through it

## What is Solr?

* Open source search engine written on top of Apache Lucene search library
* Mainly used for full text searching
* Most adapted indexing engine
* It is integrated with Zookeeper
* It is the second popular search engine after Elasticsearch

## What is Hue?

* It’s a GUI for the Hadoop echo systems