SOLUTION THAT ARE PROVIDED TO HANDLE BIG & COMPLEX DATA

Some essential Hadoop Tools :

The Hadoop Distributed File System (HDFS) is designed to store very
large data sets reliably, and to stream those data sets at high bandwidth to user applications

HBase is a column-oriented database management system that runs on top of HDFS. It is well suited for sparse data sets, which are common in many big data use cases.

The **Apache Hive** data warehouse software facilitates querying and managing large datasets residing in distributed storage.

Sqoop is a tool designed to transfer data between Hadoop and relational databases.

Pig is a platform for analyzing large data sets that consists of a
high-level language for expressing data analysis programs, coupled with infrastructure for evaluating these programs.

Apache Spark fits into the Hadoop open-source community, building on top of the Hadoop Distributed File System (HDFS). **Spark** is not tied to the two-stage MapReduce paradigm, and promises
performance up to 100 times faster than Hadoop MapReduce for certain.