Spark is a unified analytics engine for large-scale data processing including built-in modules for SQL, streaming, machine learning and graph processing.

Apache Spark is an open-source cluster computing framework. Its primary purpose is to handle the real-time generated data.

Spark was built on the top of the Hadoop MapReduce. It was optimized to run in memory whereas alternative approaches like Hadoop's MapReduce writes data to and from computer hard drives. So, Spark process the data much quicker than other alternatives.