**Descriptions**

**Spark**

Apache Spark is a general-purpose & lightning fast cluster computing system. It provides high-level API

**Scala**

Scala is a pure object-oriented language, in which conceptually every value is an object and every operation is a method-call

**Kafka**

Apache Kafka is a fast, scalable, durable, and fault-tolerant publish-subscribe messaging system.

**Hadoop**

Haddop is an open-source software framework for storing data and running applications on clusters of commodity hardware. It provides massive storage for any kind of data, enormous processing power and the ability to handle virtually limitless concurrent tasks or jobs.

**Hadoop Admin**

 The role of a Hadoop Admin is mainly associated with tasks that involve installing and monitoring hadoop clusters