**1. How are worker, executor and task related to each other?**

Executor

Task

Task

Executors are **worker** nodes processes in charge of running individual **tasks** in a given Spark job. They are launched at the beginning of a Spark application and typically run for the entire lifetime of an application. Once they have run the task they send the results to the driver.They also provide in-memory storage for RDDs that are cached by user programs through Block Manager.

Each application gets its own executor processes, which stay up for the duration of the whole application and run tasks in multiple threads. This has the benefit of isolating applications from each other, on both the scheduling side (each driver schedules its own tasks) and executor side (tasks from different applications run in different JVMs).

**2. What are the key features of Spark?**

* The main feature of Spark is its in-memory cluster computing that increases the processing speed of an application.
* Spark is designed to cover a wide range of workloads such as batch applications, iterative algorithms, interactive queries and streaming.
* It reduces the management burden of maintaining separate tools.
* Spark is designed for fast computation.
* Spark supports multi languages.
* Runs Everywhere.

**3. What is Spark Driver?**

Spark Driver is the program that runs on the master node of the machine and declares transformations and actions on data RDDs. The driver also delivers RDD graphs to the “Master”, where the standalone cluster manager runs.

**4. What are the benefits of Spark over MapReduce?**

* 100x faster than MapReduce.
* Real –time processing.
* Storing Data in memory.
* Apache Spark is simpler to program and does not require any abstractions.
* Spark provides an interactive mode.
* lower latency computations by caching the partial results across its memory of distributed workers.

**5. What is Spark Executor?**

Executors are Spark processes that run computations and store the data on the worker node. When Spark Context connects to a cluster manager, it acquires an Executor on the cluster nodes. The final tasks by SparkContextare transferred to executors.