**Assignment 12 2**

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

Worker nodes are machines that run executors.

Executors runs actual tasks.

Task is a unit of work to execute. Eg: Calculating maximum salary from employee table.

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

* Rich API
* Resilient Distributed Datasets
* DAG based execution
* In memory processing
* Unified platform

**3. What is Spark Driver?**

Driver is the program that runs on the master node of the machine. It connects to cluster manager to allocate resources across applications and sends tasks for executors to run.

Driver in Spark creates SparkContext, connected to a given Spark Master.

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

* Apache spark is very fast compared to MapReduce.
* Coding is very easy in spark (less lines of code).
* Spark has In-memory processing whereas in MapReduce all operations are done in disk, which is a time-consuming process.
* Spark performs iterative computations (computing multiple times on same dataset), whereas Hadoop cannot do iterative computing.

**5. What is Spark Executor?**

Executor runs the tasks and sends result back to the driver. Executors gets tasks from Thread pool.

We can feed multiple tasks to executor in a queue, but One executor can process only one task at a time.