Modules:

1.mapper module

2.aggeregator module

3.reducer module

Mapping module:

In mapper module, map tasks are launched in parallel to convert the original input splits into intermediate data in a form of key/value pairs. These key/value pairs are stored on local machine and organized into multiple data partitions. Map is a user-defined function, which takes a series of key-value pairs and processes each one of them to generate zero or more key-value pairs.

Aggrigator module:

In aggregator module,each aggregator can reduce merged traffic from multiple map tasks

Reducing module:

In reducer module, each reduce task fetches its own share of data partitions from all map tasks to generate the final result. The Reducer takes the grouped key-value paired data as input and runs a Reducer function on each one of them. Here, the data can be aggregated, filtered, and combined in a number of ways, and it requires a wide range of processing. Once the execution is over, it gives zero or more key-value pairs to the final step.