**Assignment-9.5**

1. **Explain the core components of flume**.

The core components include:

**Event**– a singular unit of data that is transported by Flume (typically a single log entry

**Source**– the entity through which data enters into Flume. Sources either actively poll for data or passively wait for data to be delivered to them. A variety of sources allow data to be collected, such as log4j logs and syslogs.

**Sink**– the entity that delivers the data to the destination. A variety of sinks allow data to be streamed to a range of destinations. One example is the HDFS sink that writes events to HDFS.

**Channel**– the conduit between the Source and the Sink. Sources ingest events into the channel and the sinks drain the channel.

**Agent**– any physical Java virtual machine running Flume. It is a collection of sources, sinks and channels.

* **Client**– produces and transmits the Event to the Source operating within the Agent

1. **Can flume provide 100% reliability while transfering the data**.

Flume can guarantee that all data received by an agent node will eventually make it to the collector at the end of its flow as long as the agent node keeps running.

That is, data can be **reliably** delivered to its eventual destination.The end-to-end reliability level guarantees that once Flume accepts an event, that event will make it to the endpoint - as long as the agent that accepted the event remains live long enough.

Even when using the end-to-end reliability level, a Flume flow can fail to make progress if there aren’t any nodes available to process events.

**3. Explain the consolidation in flume**.

Consolidation collect data from different sources even it's different flume agents.

Flume source can collect all data flow from different sources and flows through channel and sink.

Finally send this data to HDFS or target destination.

1. **Explain what is an event in flume.**

Aneventis the basic unit of the data transported insideFlume.

It contains a payload of byte array that is to be transported from the source to the destination accompanied by optional headers.

A Flume agent is a (JVM) process that hosts the components through which events flow from an external source to the next destination (hop).

**5. Explain what is an agent in flume**.

Anagentis an independent daemon process (JVM) in Flume.

It receives the data (events) from clients or other agents and forwards it to its next destination (sink or agent).

Flume may have more than one agent.