core components of flume

flume is developed to provide a distributed,reliable and available system for efficiently collecting, aggregating and moving large ampunts of log data from many different sources to a centralized data store. The core components of flume are

event: A sing log of entry or unit of data transported.

Flow: movement of events from the point of origin to their final destination is considered a data flow, or simply flow. This is not rigorous definition and is used only at a high level for description purpose.

Channel:It is the duct between the sink and source.

Agent:any jvm which runs the flume is called agent

Client: An interface implementation that operates at the point of origin of events and delivers them to a flume agent.clients typically operate in the process spapce of the application they are consuming data from.

Sink: An interface implementation that can remove events from a channel and transmit them to the next agent in the flow,or to the event’s final destination.terminal sink are the sink that transmit the event to its’s final destination.the flume avro sinl is an example of a regular sink that can transmit messeges to other agents thet are running an avro source

2:yes, apache flume provides end to end reliability because of its transcational approach in data flow with the help of source and sink,which is responsible for pass reliably from one end to another end in a flow.

3:Consolidation in flume:flume can collect data from different sources even from different flume agents.It collects data from different and flows through channe and link.It finally sends data to hdfs or any other targeted local system.

4:Event in flume:A unit of data with set of string attributes.Each log file is considered as an event to the source.Flume has inbuilt functionality to understand the source format.EEvery event has a header and value sector, which has a header information and appropriate value that is assigned to particular header.

5:Agent in flume: A flume agent is a jvm process that holds the flume core components through which events flow from an external source to the target source such as hdfs.Agent is most important component in flume,(i.e) it is heart of flume.