**Explain in detail.**

**●What is meant by FlumeNG?**

Apache Flume is a distributed, reliable, and available system for efficiently collecting, aggregating and moving large amounts of log data from many different sources to a centralized data store. At a high-level, Flume NG uses a single-hop message delivery guarantee semantics to provide end-to-end reliability for the system.

**● Can Flume provides 100 % reliability to the data flow?**

Yes, it provides end-to-end reliability of the flow. By default flume uses a transactional approach in the data flow. Sources and sink encapsulated in a transactional repository provides by the channels. This channels responsible to pass reliably from end to end in the flow. So it provides 100% reliability to the data flow.

**● Can Flume distribute data to multiple destinations?**

Yes flume can distribute the data to multiple destinations. As the flume agent can declare more than one source, channels and sink. Hence they can take inputs from many sources and buffer in multiple channels as well as give them to multiple sinks(Destinations) . For doing this one has to change the configuration of the flume and allocate the respective sinks.

**● Explain about the different channel types in Flume. And which channel type is faster?**

The 3 different built in channel types available in Flume are-

JDBC Channel – JDBC Channel stores the events in an embedded Derby database.

FILE Channel –File Channel writes the contents to a file on the file system after reading the event from a source. The file is deleted only after the contents are successfully delivered to the sink.

MEMORY Channel – Events are read from the source into memory and passed to the sink.

**MEMORY Channel is the fastest channel** among the three however has the risk of data loss. The channel that you choose completely depends on the nature of the big data application and the value of each event.