**12.3.1 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

**12.3.2 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 which ensures the reliability of data flow. Sources and sinks are encapsulated in a transactional repository provided by the channels.

This ensures that the set of events are reliably passed from point to point in the flow. In the case of a multi-hop flow, the sink from the previous hop and the source from the next hop both have their transactions running to ensure that the data is safely stored in the channel of the next hop.

So it provides 100% reliability in the data flow.

**12.3.3 Can Flume can distributes data to multiple destinations?**

Yes, It supports Multiplexing flow. The Event flows from one source to multiple channels and multiple destinations. It is achieved by defining a flow multiplexer that can replicate or selectively route an event to one or more channels.

In Multiplexing Flow, an event is delivered to a subset of available channels when an event’s attribute matches a preconfigured value

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

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