1. What is meant by FlumeNG ?

FlumeNG is the Flume Next Generation while previous version of Flume is called as FlumeOG i.e. Flume Old Generation. Flume NG uses a single-hop message delivery guarantee semantics to provide end-to-end reliability for the system. To accomplish this, certain new concepts have been incorporated into its design, while certain other existing concepts have been either redefined, reused or dropped completely.

1. Can Flume provides 100 % reliability to the data flow?

Yes, Apache Flume provides end to end reliability because for the data flow it utilizes transactional approach.

1. Can Flume can distributes data to multiple destinations?

Yes, it support flow of data in multiple destinations. The data can flow from one source to multiple channels and many destinations. For this to be achieved flow multiplexer has to be defined.

1. 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-

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.