[Using Flume and Kafka](https://www.dezyre.com/hackerday/real-time-decision-using-flume-kafka) to solve the Streaming case of Hadoop small file problem on HBase

## **ARCHITECTURAL OVERVIEW OF THE SOLUTION**

Kafka

Flume Agent

Streaming data

An application that consumes from Kafka and generates a rowKey for each event

The consumer application having processed the records sends it to another kafka topic as a producer

Kafka

Application that consumes from a kafka topic and writes to HBase

HBase on HDFS

Stream processing application

In this architecture with low latency historical data lookup requirement, the stream processing application can also request the database to look up historical records of a shared rowkey.