[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

## **ARCHITECTURAL OVERVIEW OF THE SOLUTION**

Flume Agent

(hdp\_sfp\_flume\_hdfs\_kafka\_agent.properties)

**HDFS**

Kafka Server

Streaming data

## **TECHNICAL DETAILS FOR THE LAYER 2 FLUME AGENT (**hdp\_sfp\_flume\_hdfs\_kafka\_agent.properties**)**

Kafka Consumer (stream data processing app)

Custom Interceptor to generate a primary key for each record

**Event**

**Source**

**HDFS Sink**

Kafka Messaging Broker Running on Zookeeper

Kafka Channel

Memory Channel