Twitter
API(Original Source)

Kafka
Producer
(Cleans the tweet, publish to Kafka Topic)

Kafka

Rafka
Consumer
(Takes the data, writes to Hive with Spark)

Spark)

Hive
(Permanent Structured Data Storage)