CS286A Metadata Data Mover Installation and Usage Guide

Jialiang Gu, Haoyu Chen

This program will be run on **two separate machines**.

On each machine, clone the files under kafka repo from Github.

**Package Resources and Dependencies:**

* Kafka can be downloaded at: <http://kafka.apache.org/downloads.html> (the version we used is 0.8.2.1.)
* The files we revised from the original kafka package are located at:
  + [cs286A](https://github.com/cs186-spring15/cs286A)/[dataMover](https://github.com/cs186-spring15/cs286A/tree/master/dataMover)/[kafka](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka)/[core](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core)/[src](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src)/[main](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src/main)/[scala](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src/main/scala)/[kafka](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src/main/scala/kafka)/tools/FileConsumer.scala
* In addition, the repo's files should be placed in:
  + [cs286A](https://github.com/cs186-spring15/cs286A)/[dataMover](https://github.com/cs186-spring15/cs286A/tree/master/dataMover)/[kafka](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka)/[core](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core)/[src](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src)/[main](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src/main)/[scala](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src/main/scala)/[kafka](https://github.com/cs186-spring15/cs286A/tree/master/dataMover/kafka/core/src/main/scala/kafka)/tools/

**Runtime Instructions:**

* cd into the kafka directory *(denoted here as $KAFKA>)*
* Run the following commands:
  + $KAFKA> ./gradlew jar
  + $KAFKA> ./gradlew srcJar
  + $KAFKA> bin/zookeeper-server-start.sh config/zookeeper.properties
  + $KAFKA> bin/kafka-server-start.sh config/server.properties
  + $KAFKA> bin/kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic test

*On the client side, locate the file you want to write log into: (denoted here as $LOG\_FILE)*

* cd into the kafka directory *(denoted here as $KAFKA>)*
* $KAFKA> bin/kafka-file-consumer.sh --zookeeper localhost:2181 --topic test --output-file $LOG\_FILE

*On the server side:*

* cd into dataMover directory *(denoted here as $DATAMOVER>)*
* $DATAMOVER> ./inotifywait.sh