**Platform :** CentOS with the latest version of Java

**Software Packages needed :**

Kafka - <http://kafka.apache.org/downloads.html>

Storm - <https://storm.incubator.apache.org/downloads.html>

Cassandra for CentOS- <http://www.datastax.com/docs/1.0/install/install_rpm>

DataStax Java libraries for cassandra - <http://www.datastax.com/download>

**Problem Definition**

**Process Trade Information(simulated) coming from all over the world to see how many times people have traded a particular Symbol during the day**

**Dataflow**

Kafka producers log information into Kafka Cluster

Storm spouts read from Kafka cluster, gives the information to Bolts to filter unwanted information, count the occurrences and update cassandra database.

There is also a utility project that constantly polls TradeFrequencies table in Cassandra to see real time updates

Complete demo is in youtube,

<https://www.youtube.com/watch?v=TrsIHE-cVjc>

I have shown the important lines of code and a demo to show how these two technologies come together to solve a the problem defined above

**Projects**

Projects are organized into folders. You could easily port them into Eclipse

KafkaAndStorm -> Storm project - Has Kafka spout, bolts and topologies

KafkaDemo -> Kafka Producer -> simulates real time trade information

QueryTradeFrequencies -> Uses DataStax adapters to poll database to show real time updates