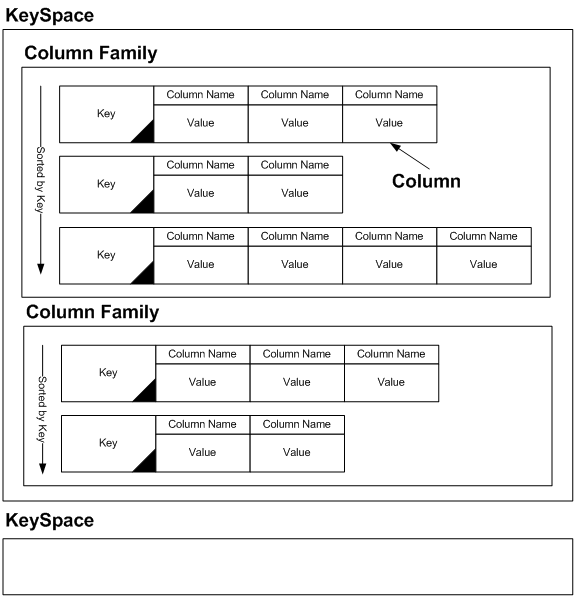
**Apache-Cassandra Database Set Up:**

**Introduction :**

Cassandra is distributed database system. It is donated to Apache open source group by Facebook at 2008.The Cassandra is based on Google Big Table data model and Facebook Dynamo distributed architecture. It doesn’t use SQL and optimized to high scale size of data & transaction handling. Even though Cassandra is implemented with Java language, other language can use the Cassandra as a client. (It supports Ruby,Perl,Python,Scala,PHP etc).

It is used to High Scale Size SNS like Face book,Digg,Twitter etc. It doesn’t support complex relationship like Foreign Key. It just provides Key & Value relationship like Java Hashmap. It is very easy to install and use.

Let’s look at data model of Cassandra



Getting Cassandra to run on Mac OSX is very easy just download the latest binary from<http://cassandra.apache.org/>. At the time of this writing the version was 0.6.3, but I updated the important parts to 0.7rc3 so it should work with that version as well.

Create the directories Cassandra needs to work:

sudo mkdir -p /opt/cassandra

sudo chown -R `whoami` /opt/cassandra

sudo mkdir -p /var/log/cassandra

sudo chown -R `whoami` /var/log/cassandra

sudo touch /var/log/cassandra/system.log

sudo mkdir -p /var/lib/cassandra

sudo chown -R `whoami` /var/lib/cassandra

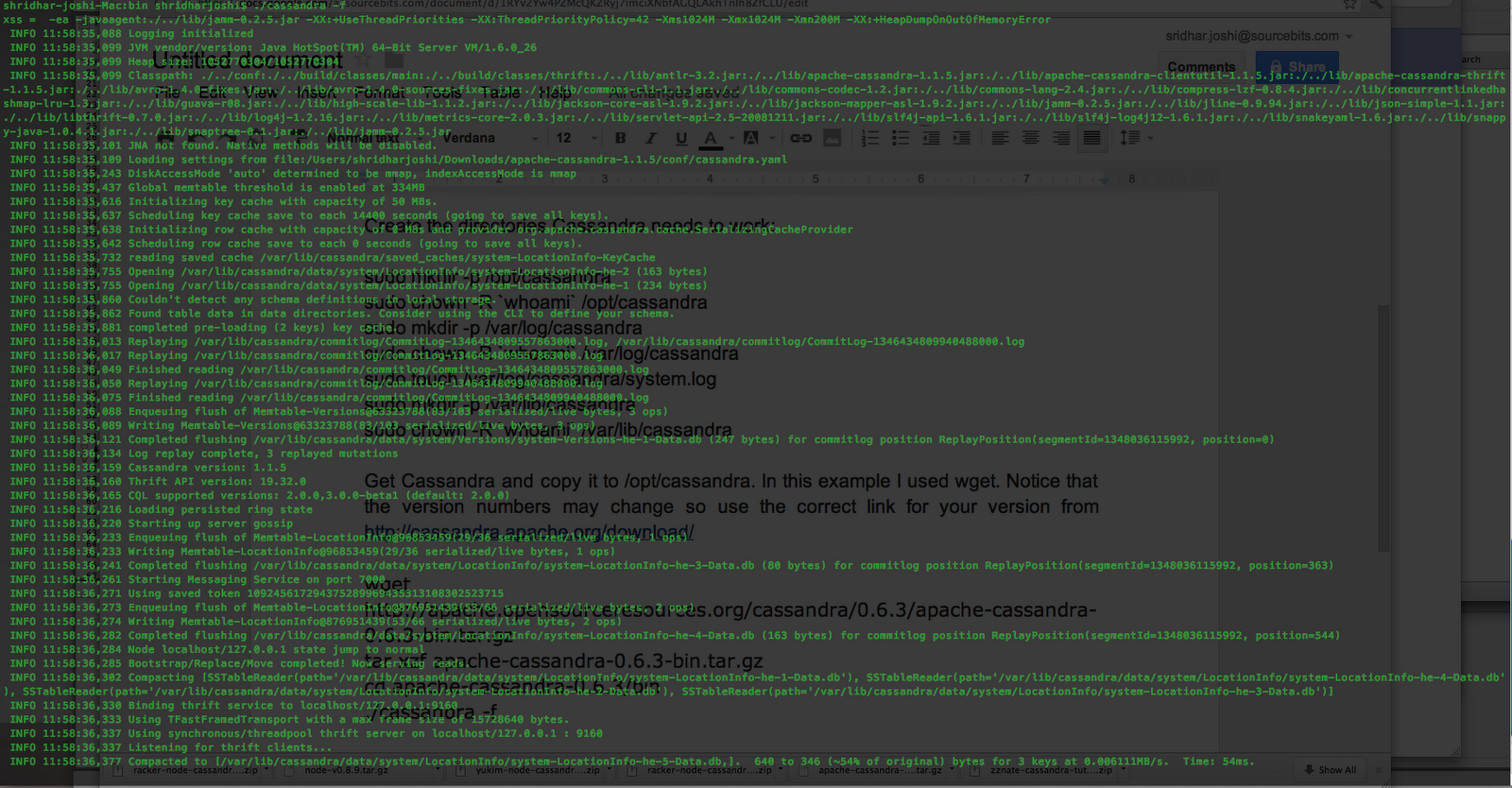
Get Cassandra and copy it to /opt/cassandra. In this example I used wget. Notice that the version numbers may change so use the correct link for your version from <http://cassandra.apache.org/download/>

wget http://apache.opensourceresources.org/cassandra/0.6.3/apache-cassandra-0.6.3-bin.tar.gz

tar xzf apache-cassandra-0.6.3-bin.tar.gz

cd apache-cassandra-0.6.3/bin

./cassandra -f



If all is OK u should see the above like screen......

Now lets test Cassandra. Open a new Terminal window, or tab, and:

cd /opt/cassandra/bin

./cassandra-cli --host localhost --port 9160

Cassandra 0.7rc3 output and testing is a bit different so check with the comments bellow.

At the prompt (for 0.6.x)

cassandra> set Keyspace1.Standard2['jsmith']['first'] = 'Shridhar'

Value inserted.

cassandra> set Keyspace1.Standard2['jsmith']['last'] = 'Joshi'

Value inserted.

cassandra> set Keyspace1.Standard2['jsmith']['age'] = '23'

Value inserted.

cassandra> get Keyspace1.Standard2['shrijosh']

(column=age, value=23; timestamp=1249930062801)

(column=first, value=Shridhar; timestamp=1249930053103)

(column=last, value=Joshi; timestamp=1249930058345)

Returned 3 rows.

cassandra> quit

If you’re able to get the values out of Keyspace1 all is good and Cassandra is up and running.

If you are installing 0.7 or higher the test Keyspace1 is not loaded automatically and you need to run jconsole to load the “schema”. And go to MBeans -> org.apache.cassandra.service -> StorageService -> Operations -> loadSchemaFromYAML as documented on [LiveSchemaUpdates](http://wiki.apache.org/cassandra/LiveSchemaUpdates)

Please note that jconsole needs to connect to localhost on port 8080, to talk to Cassandra. Do not use the Thrift port 9160.

Now lets install [Chiton](http://github.com/driftx/chiton), a GTK GUI written in Python that will allow us to view the data stored in Cassandra.

mkdir chiton-temp

cd chiton-temp

git clone git://github.com/driftx/chiton.git

export VERSIONER\_PYTHON\_PREFER\_32\_BIT=yes

In order to fulfill a Chiton dependecies we also need

* Twisted 8.1.0 or later
* Thrift (latest svn)
* PyGTK 2.14 or later
* simplejson
* Telephus

**Definations/Introduction to CQL ( Cassandra Query Language)**

**Column**

Column data structure which consists of column name and column value.

**Column Family**

Column family is set of columns. It is similar to row in RDBMS table. I will explain more detail about difference between Column Family and row in RDBMS later. Column Family has a key which identify each row in data set. Each row has a number of Columns.

**KeySpace**

Keyspace is logical set of column family for management perspective. It doesn’t impact data structure.

**Super Column & Super Column Family**

As I mentioned earlier, column value can have a column itself. (Similar to Java Hashtable can have ValueObject class as a ‘Object’ type)