



# APACHE CASSANDRA™ NODETOOL/MANAGEMENT REFERENCE

## ***What is Cassandra?***

Apache Cassandra is a high performance, extremely scalable, fault tolerant (i.e. no single point of failure), distributed post-relational database solution. Cassandra combines all the benefits of Google Bigtable and Amazon Dynamo to handle the types of database management needs that traditional RDBMS vendors cannot support. DataStax is the leading worldwide commercial provider of Cassandra products, services, support, and training.

## ***How do I install Cassandra?***

Downloading and installing Cassandra is very easy. Downloads of Cassandra are available via the DataStax web site at: [www.datastax.com/download](http://www.datastax.com/download).

For installation guidance, please see our online documentation at: [www.datastax.com/docs/1.0/getting\\_started/index](http://www.datastax.com/docs/1.0/getting_started/index) and [www.datastax.com/docs/1.0/install/index](http://www.datastax.com/docs/1.0/install/index). You can also view a guided video tutorial for installing a simple Cassandra and DataStax OpsCenter setup at: [www.datastax.com/resources/tutorials](http://www.datastax.com/resources/tutorials).

## ***What client libraries/drivers can I use with Cassandra?***

There are a number of CQL (Cassandra Query Language) drivers and native client libraries available for most all popular development languages (e.g. Java, Ruby, etc.) All drivers and client libraries can be downloaded from: [www.datastax.com/download/clientdrivers](http://www.datastax.com/download/clientdrivers).

# APACHE CASSANDRA™ NODETOOL/MANAGEMENT REFERENCE

```
bin/nodetool -h HOSTNAME [-p JMX_PORT ] COMMAND...
```

**CFSTATS** Prints statistics for every keyspace and column family

**CFHISTOGRAMS** Displays latency statistics for a column family  
cfhistograms [keyspace][column family]

**CLEANUP** Removes keys no longer belonging to node  
cleanup [keyspace][column family]

**CLEARSNAPSHOT** Deletes all backup snapshots

**COMPACT** Performs compaction  
compact [keyspace][column family]

**DECOMMISSION** Takes node down and moves data to next node

**DRAIN** Flushes all memtables and stops write operations

**FLUSH** Flushes all memtables to disk and clears commit log  
flush [keyspace][column family]

**GETCOMPACTIONTHRESHOLD** Gets compaction threshold for a column family  
getcompactionthreshold [keyspace][column family]

**INFO** Displays basic monitoring info on the node

**INVALIDATEKEYCACHE** Removes/cleans out the key cache  
invalidatekeycache [keyspace][column family]

**INVALIDATEROWCACHE** Removes/cleans out the row cache  
invalidaterowcache [keyspace][column family]

**JOIN** Includes the node in the ring

**LOADBALANCE** Moves the node to a new token so that it will split the range of whatever token currently has the highest load

**MOVE** Moves node to new token  
move <new token>

**NETSTATS** Displays network monitoring info for the node

**RING** Displays general information about the ring

**REMOVETOKEN** Removes a specified token  
removetoken status | force | token

**REPAIR** Begins an anti-entropy node repair  
repair keyspace [column family] [-pr]

**SCRUB** Rebuilds specified SSTables for column family  
scrub [keyspace][column family]

**SETCACHECAPACITY** Sets the size of the key/row cache  
setcachecapacity keyspace column family  
key\_cache\_capacity row\_cache\_capacity

**SETCOMPACTIONTHRESHOLD** Sets the compaction threshold  
setcompactionthreshold column family min\_  
threshold [max\_threshold]

**SNAPSHOT** Does an online backup of the node's data  
snapshot [snapshot name]

**TPSTATS** Prints the number of active, pending, and completed tasks for each of the thread pools that Cassandra uses for stages of operations

**VERSIONS** Displays the software version

## **ABOUT DATASTAX**

DataStax is the primary developer of and worldwide leader in Apache Cassandra. DataStax provides free and commercial software products including DataStax Enterprise, a distributed, scalable, and highly available database platform, based on Apache Cassandra, that is designed for the management of both real-time and analytic workloads. The company has over 100 customers, including leaders such as Netflix, Cisco, Rackspace and Constant Contact, and spanning verticals including web, financial services, telecommunications, logistics and government.

DataStax is backed by industry leading investors, including Lightspeed Venture Partners and Crosslink Capital and is based in Burlingame, CA with offices in Austin, TX and Stamford, CT.

For more information and software downloads, visit [www.datastax.com](http://www.datastax.com).