About Brisk

Brisk is an open-source Hadoop and Hive distribution developed by DataStax that utilizes Apache Cassandra for its core services and storage. Brisk provides Hadoop MapReduce capabilities using CassandraFS, an HDFS-compatible storage layer inside Cassandra. By replacing HDFS with CassandraFS, users are able to leverage their current MapReduce jobs on Cassandra's peer-to-peer, fault-tolerant, and scalable architecture. Brisk is also able to support dual workloads, allowing you to use the same cluster of machines for both real-time applications and data analytics without having to move the data around between systems.

Brisk is now available via Apache license v2.0. The DataStax team welcomes your valued feedback.

Release Contents

* *Brisk is comprised of the following components. For component-specific information, refer to their respective release notes and documentation.*

- Apache Hadoop 0.20.203.0 + (HADOOP-7172, HADOOP-5759, HADOOP-7255)
- Cassandra 0.8.1 + (CASSANDRA-2683)
- Apache Hive 0.7
- Apache Pig 0.8.3

New Features in Brisk 1.0 Beta 2

The following new features have been added in this release:

Feature	Description
BRISK-12	Apache Pig Integration. See the DataStax Documentation for more information about using Pig in Brisk.
BRISK-89	Job Tracker Failover. See the DataStax Documentation for more information about using the new `brisktool movejt` command.
BRISK-207	New Snappy Compression Codec built on Google Snappy is now used internally for automatic CassandraFS block compression.
BRISK-180	Automap Cassandra Column Families to Hive Tables in the Brisk Hive Metastore.
BRISK-152	Add a second HDFS layer in CassandraFS for long-term data storage. This is needed because the blocks column family in CFS requires frequent compactions - Hadoop uses it during MapReduce processing to store small files and temporary data. Compaction cleans this temporary data up after it is not needed anymore. Now there is the `cfs:///` and `cfs-archive:///` endpoints within CFS. The blocks column family in `cfs-`

Feature	Description
	`archive:///` has compaction disabled to improve performance for static data stored in CFS.

Major Fixes in Brisk 1.0 Beta 2

Brisk 1.0 Beta 2 also incudes the following major fixes. For details on all fixes in Beta 2, see the Brisk Jira Project Web site:

Issue	Description
BRISK-126	Remove multiple slf4j warnings
BRISK-203	Use `batchMutate` instead of `insert` in `HiveCassandraOutputFormat`
BRISK-219	Cassandra super columns not mapping in Hive
BRISK-220	Improve performance of `hadoop fs -ls`
CASSANDRA-2683	Compaction issue causing secondary index corruption.

Open Issues

For a description of the open issues in Brisk, see the Brisk Jira Project Web site.

Upgrading from Beta 1 to Beta 2

Perform a rolling upgrade by performing the following steps on each node in your Brisk cluster, one node at a time.

1. Flush the commit log: `nodetool drain`
2. Stop any client applications.
3. Stop the Brisk service: `service brisk stop`
4. Upgrade the Brisk packages to Beta 2.
   On RedHat Systems: `yum upgrade brisk-full brisk-demos`
   On Debian Systems: `apt-get upgrade brisk-full brisk-demos`
   For Binary Installs: Download and unpack the tar file and update $BRISK_HOME and $PATH to point to the new location.
5. Restart Brisk: `service brisk start`