

PigInteroperability

This page describes how Pig interoperates with other platforms, such as HBase and Hive.

Pig and Cassandra

Cassandra has Pig support for loading and storing starting with Cassandra 0.7. It requires Pig 0.7+. The CassandraStorage class (for loading and storing) is found in the Cassandra src and it's currently in a contrib module (<http://cassandra.apache.org/download/>). Details are found in the readme there. More hadoop support details are found here: <http://wiki.apache.org/cassandra/HadoopSupport>. A project that looks to ease integration between the two is Pygmalion: <https://github.com/jeromatron/pygmalion/>.

Pig and HBase

In Pig 0.6 and before, the built in HBaseStorage can be used to load data from Hbase. Work is ongoing to enhance this loader and make it a storage function also. See <http://issues.apache.org/jira/browse/PIG-1205>

Pig and Hive RCFiles

The HiveColumnarLoader, available as part of PiggyBank in Pig 0.7.0.

Pig and Voldemort

The Pig LoadFunc for Voldemort.

See <http://github.com/rsumbaly/voldemort/blob/hadoop/contrib/hadoop/src/java/voldemort/hadoop/pig/VoldemortStore.java>

Pig and Avro

[AvroStorage](#), an Avro LoadFunc/StoreFunc in piggybank. See <https://issues.apache.org/jira/browse/PIG-1748>