HBaseMetastoreDevelopmentGuide

Guide for contributors to the metastore on hbase development work. Umbrella JIRA - HIVE-9452

This work is discontinued and the code is removed in release 3.0.0 (HIVE-17234).

Building

You will need to download the source for Tephra and build it from the develop branch. You need Tephra 0.5.1-SNAPSHOT. You can get Tephra from Cask's github. Switch to the branch develop and doing 'mvn install' will build the version you need.

Setup for running hive against hbase metastore -

Once you've built the code from the HBase metastore branch (hbase-metastore), here's how to make it run against HBase:

1. Install HBase, preferably HBase 1.1.1 as that's what is being used for testing.
2. Copy following jars into $HBASE_HOME/lib
   a. hive-common-.*.jar
   b. hive-metastore-.*.jar
   c. hive-serde-.*.jar
3. Setup HBase, http://hbase.apache.org/book.html#quickstart I run it in stand alone mode, so you have to set a couple of values in hbase-site.xml for this to work.
4. Set HADOOP_HOME if you're not in a cluster where hadoop is already on your path.
5. Start HBase: $HBASE_HOME/bin/start-hbase.sh
6. Set it up so that HBase jars and conf file are picked up by Hive
   a. export HIVE_AUX_JARS_PATH=$HBASE_HOME/lib/
   b. export AUX_CLASSPATH=$HBASE_HOME/conf
7. Create the metastore tables in HBase: `hive --service hbaseschematool --install`
8. Configure Hive to use HBase as its metastore, in hive-site.xml:

```
<property>
    <name>hive.metastore.rawstore.impl</name>
    <value>org.apache.hadoop.hive.metastore.hbase.HBaseStore</value>
</property>
<property>
    <name>hive.metastore.fastpath</name>
    <value>true</value>
</property>
```

Now start Hive as normal, all should just work.

Notes

- Currently (as of Apr 8 2015) we have not tested the HBase metastore with the metastore service. We don't know if it works or not. We have tested it with the command line client and HiveServer2.
- Not all Hive operations have been tested. Insert, select, create table, drop table, create database, add partition, drop partition have been tested. Other features may not work.
- Once we switched to HBase 1.1.1, Tephra no longer works. You'll need to use the VanillaHBaseConnection (which is the default) until we get Tephra working again.

Importing metadata from rdbms to hbase

Set hive-site.xml so that it has necessary jdo properties for rdbms, and setup Hive to use hbase for metadata storage as documented above

The following command will import the metadata from the rdbms to hbase:

hive --service hbaseimport

Design Docs

Overall Approach

Hbase execution plans for RawStore partition filter condition