Broadcast joins (formerly ‘Map join’) in Hive on Tez

Tez is providing a broadcast edge for map joins in hive. This edge requires that the execution phase of hive do the following:

1. Before the DAG is generated in hive, we need to modify the operator tree to create a map join operator. This includes setting up information such as the big table and join keys based on the type of join and statistics of the tables.
2. In the DAG generation phase we need to set up the broadcast edge based on the map join operator.
3. Build the hash tables from a stream of rows being sent by the broadcast edges.
4. Identify the edge corresponding to a particular table and associate the built hash tables from step 3. We need to use this information and set up to run according to the existing join operators.

The building of hash tables and the mapping of the table to the edge/hash table is done in the TezProcessor class. TezProcessor class is a new class that will replace the use of ExecMapper and ExecReducer in the Tez DAG.

With this change, broadcast join will support any subquery as the ‘small table’, and can run at any phase in the DAG (it can also run one of the ‘reduce’ steps).