Design Considerations

Kevin Liew

May 15, 2016

1 RPC

How are ResultSets transferred? Batches contained in Thrift objects? Contiguous batches of TRowSet.

```
ie.
ResultSet{
    TRowSet batch1{
        TColumn1{row1, row2},
        TColumn2{row1, row2},
        columnCount = 2
    }
    TRowSet batch2{
        TColumn1{row3, row4},
        TColumn2{row3, row4},
        columnCount = 2
    }
}
```

except that the columns are in 'binaryColumns' rather than a list of TColumn after HIVE-12049

2 Interface

Should the preferred compressor be per-connection, or per-query?

Should the client send a preferred compressor plugin's jar file? Or should the client specify the compressor in the connection string and require that the server already have the plugin?

Location of exemplar code for sending jars from the client and receiving at the server?

How can we give the server more control over plugins? A list of disallowed compressors is not useful if we allow the client to send arbitrary jars with arbitrary entry-classes. If we instead have a list of allowed compressors, this list would require the server to have prior knowledge of all possible compressors. In this case, the server might as well be pre-loaded with those plugins and there is no need for the client to send the jars.

3 Compressor-Decompressor

How do we handle the case where the compressor throws an exception while handling an inner batch? Do we restart serialization and serialize the batch as an uncompressed column? What do we do with prior batches that were already serialized?)

Where does decompression occur in the client? Will we operate on (Encoded)ColumnBasedSet or directly on TRowSet?

We are no longer storing the compressor-name with each column. Will that introduce any issues?

hive.server2.thrift.resultset.max.fetch.size: "Max number of rows sent in one Fetch RPC call by the server to the client." If the client receives batch-by-batch, then we don't need to store the batch size in TEnColumn because the client will receive and decompress each batch separately instead of receiving one huge blob with all batches?