DesignDocs

Hive Design Documents

⚠ Proposals that appear in the "Completed" and "In Progress" sections should include a link to a JIRA ticket

Completed

- Views (HIVE-1143)
- Partitioned Views (HIVE-1941)
- Storage Handlers (HIVE-705)
- HBase Integration
- HBase Bulk Load
- Locking (HIVE-1293)
- Indexes (HIVE-417)
- Bitmap Indexes (HIVE-1803)
- Filter Pushdown (HIVE-279)
- Table-level Statistics (HIVE-1361)
- Dynamic Partitions
- Binary Data Type (HIVE-2380)
- Decimal Precision and Scale Support
- HCatalog (formerly Howl)
- HiveServer2 (HIVE-2935)
- Column Statistics in Hive (HIVE-1362)
- List Bucketing (HIVE-3026)
- Group By With Rollup (HIVE-2397)
- Enhanced Aggregation, Cube, Grouping and Rollup (HIVE-3433)
- Optimizing Skewed Joins (HIVE-3086)
- Correlation Optimizer (HIVE-2206)
- Hive on Tez (HIVE-4660)
    - Hive-Tez Compatibility
- Vectorized Query Execution (HIVE-4160)
- Cost Based Optimizer in Hive (HIVE-5775)
- Atomic Insert/Update/Delete (HIVE-5317)
- Transaction Manager (HIVE-5843)
- SQL Standard based secure authorization (HIVE-5837)
- Hybrid Hybrid Grace Hash Join (HIVE-9277)
- LLAP Daemons (HIVE-7926)
- Support for Hive Replication (HIVE-7973)

In Progress

- Column Level Top K Statistics (HIVE-3421)
- Hive on Spark (HIVE-7292)
- Hive on Spark: Join Design (HIVE-7613)
- Improve ACID Performance – download docx file (HIVE-14035, HIVE-14199, HIVE-14233)
- Query Results Caching (HIVE-18513)
- Default Constraint (HIVE-18726)
- Different TIMESTAMP types (HIVE-21348)
- Support SAML 2.0 authentication (HIVE-24543)

Proposed

- Spatial Queries
- Theta Join (HIVE-556)
- theta join and cross product proposal - rough draft
- JDBC Storage Handler
- MapJoin Optimization
- Proposal to standardize and expand Authorization in Hive
- Dependent Tables (HIVE-3466)
- AccessServer
- Type Qualifiers in Hive
- MapJoin & Partition Pruning (HIVE-5119)
- Updatable Views (HIVE-1143)
- Phase 2 of Replication Development (HIVE-14841)
- Subqueries in SELECT (HIVE-16091)
- DEFAULT keyword (HIVE-19059)
- Hive remote databases/tables

Incomplete

- Authorization (Committed but not secure/deployable – see Disclaimer)

Abandoned

- Hive across Multiple Data Centers (Physical Clusters)
- Metastore on HBase (HIVE-9452)

Other

- Security Notes
- Hive Outer Join Behavior
- Metastore ER Diagram