HRDBMS: A NewSQL Database for Analytics
Jason Arnold, Boris Glavic, Ioan Raicu

Motivation

**Challenges**
- MPP relational databases don’t scale well past medium-sized clusters
- Big Data platform databases (Hive, Spark SQL, Azure SQL, Dremel) have poor per node performance

**Contributions**
- Combine the best of MPP relational databases with the best of Big Data platform databases
- Take advantage of decades of query optimization research
- Take advantage of scalability of Big Data platforms
- Create a custom Big Data platform for HRDBMS based on lessons learned from other such databases
- Fully support DML and ACID

HRDBMS Cost-Based Optimizer

- Phase 1: Traditional relational optimizations
- Phase 2: Conversion to workflow
- Phase 3: Basic workflow optimizations
- Phase 4: Advanced workflow optimization
- Phase 5: Indexes and miscellaneous optimizations

HRDBMS Architecture

**HRDBMS clusters contain coordinator nodes and worker nodes**

**Coordinator Nodes**
- Accept client connections
- Maintain system metadata (what data is where, schema info, indexes, etc)
- Query planning and workflow generation
- Manage query execution

**Worker Nodes**
- Contain user table/index data
- Workflow execution

Experimental Results

- All testing was done with TPC-H scale factor 100 (100GB)
- All nodes are Amazon EC2 m3.2xlarge

**Distributed Execution Engine**

Custom execution engine modeled after Map/Reduce

- Pipelined execution
- Materialize only when necessary
- Non-blocking shuffle
- Intra-task parallelism
- Can chain Map/Combiner phases together without a Reduce
- More flexible network communication structure (hierarchical shuffle)
- Data stored in local filesystem on worker nodes / No HDFS
- Worker nodes contain large number of disks
- All disks read in parallel

*4 nodes*