Replicated Database Engine in Clickhouse

Val Baturin
Supervisor: Alexey Milovidov

June 5, 2020

Outline

Results

clickhouse-client

```
CREATE DATABASE database_name
ENGINE = Replicated('zookeeper_path', 'replica_name')
```

Results

- Replicated metadata
- DDL log written to Zookeeper
- Metadata snapshots[1]
- Integration tests

---

[1]Tests are not sufficient

Database Engine

- initialization of set of known tables
- checking existence of a table and getting a table object
- retrieving a list of all tables
- creating and dropping tables
- renaming tables and moving between databases.

Database Engine: Hierarchy

IDatabase

DatabaseOnDisk

DatabaseAtomic

DatabaseReplicated

Data in DBMS

Raw bytes / String

Row

Table

Database

Query interpretation: Context

clickhouse-client -n

```
DESC TABLE ONE_TABLE;
USE DIFFERENT_DATABASE;
DESC TABLE ANOTHER_TABLE;
```

clickhouse-client -n

```
DESC TABLE default.ONE_TABLE;
DESC TABLE DIFFERENT_DATABASE.ANOTHER_TABLE;
```

Query interpretation: Interpreter

- executeQuery
- 35 interpreters
- e.g. InterpreterCreateQuery

Query interpretation: Parser

- parseQuery
- AST
- TokenIterator
- Lexer

Replication Algorithm

SSM (Single Source Of Metadata)

Replicated Log

Hybrid

Implementation gist

- propose query
- ReplicatedBackgroundPool task
- setCurrentDatabase
- AST modification
- ReplicatedTables are simplified on creation

Questions

Ask Me Anyting