CLICKHOUSE

OPEN-SOURCE COLUMN-ORIENTED DATABASE MANAGEMENT SYSTEM THAT ALLOWS GENERATING ANALYTICAL DATA REPORTS IN REAL TIME

Created by Yegor Andreenko / @f1yegor

WHOAMI

- Yegor Andreenko
- @f1yegor
- Software Engineer at Crobox
- Data Storages lover

CROBOX

combines consumer psychology with machine-learning to influence (purchase) decisions

In one line of code

HISTORY

Yandex development

- MyISAM 2008-2011
- Metrage 2010-now
- OLAPServer 2009-2013
- ClickHouse 2013-now

REQUIREMENTS

- Be able to work with Big Data - 11,4 trillion rows(2015)
- Linear scalability - from 60 to 394 in 2 years, 1 Tb/s of uncompressed data
- High performance (benchmarks)
- Enough functionality for tools of web analytics. SQL dialect

COMPETITORS

- Commercial OLAP systems on-premise

  HP Vertica, Actian Vector, Actian Matrix, EXASol, Sybase IQ, etc.

  Difference: open-source and free

- Cloud solutions

  Amazon Redshift and Google BigQuery.

  Difference: can run you on own hardware

COMPETITORS

- Hadoop solutions

  Cloudera Impala, Spark SQL, Facebook Presto, Apache Drill.

  Difference: web user-service, hadoop-less, geo-distribution

- Open-source OLAP DBMS

  InfiniDB, MonetDB, LucidDB.

  Difference: they are inmature or abandoned, non-distributed

COMPETITORS

- Open-source analytical systems that is not Relational OLAP DBMS

  Metamarkets Druid, Apache Kylin.

  Difference: ClickHouse doesn't require preaggregation, has SQL support

WHAT IS CLICKHOUSE?

ClickHouse is a columnar DBMS for OLAP.

In a "normal" row-oriented DBMS, data is stored in this order

```
5123456789123456789  1  Eurobasket - Greece - Bosnia and
Herzegovina  1  2011-09-01 01:03:02  6274717

5234985259563631958  0  Consulting, Tax assessment, Accounting,
Law  13  2011-09-02 01:03:20  6320881

5234985259563623423  2  Accounting, Tax assessment  3  2011-09-04
01:03:445  6320890
```

WHAT IS CLICKHOUSE?

In a column-oriented DBMS, data is stored like this:

```
WatchID:    5385521489354350662     5385521490329509958     538552
1489953706054     5385521490476781638     5385521490583269446
5385521490218868806     5385521491437850694   5385521491090174022
     5385521490792669254     5385521490420695110     5385521491532
181574     5385521491559694406     5385521491459625030     5385521
492275175494   5385521492781318214      5385521492710027334     53
855214929556153O2     5385521493708759110     5385521494506434630
    5385521493104611398
JavaEnable: 1       0       1       0       0       0       1
  0       1       1       1       1       1       1       0
1       0       0       1       1
Title:      Yandex  Announcements - Investor Relations - Yandex
  Yandex — Contact us — Moscow    Yandex — Mission         Ru
Yandex — History — History of Yandex    Yandex Financial Releases
- Investor Relations - Yandex Yandex — Locations      Yandex Board
 of Directors - Corporate Governance - Yandex       Yandex — Techn
ologies
GoodEvent:  1       1       1       1       1       1       1
  1       1       1       1       1       1       1       1
1       1       1       1       1
EventTime:  2016-05-18 05:19:20     2016-05-18 08:10:20     2016-0
5-18 07:38:00     2016-05-18 01:13:08     2016-05-18 00:04:06
2016-05-18 04:21:30     2016-05-18 00:34:16     2016-05-18 07:35:4
9     2016-05-18 11:41:59     2016-05-18 01:13:32
...
```

DISTINCTIVE FEATURES OF CLICKHOUSE

- True column-oriented DBMS.
- Data compression.
- Disk storage of data(not only RAM).
- Parallel processing on multiple cores.
- Distributed processing on multiple servers.
- Indexes.
- Vector engine.

DISTINCTIVE FEATURES OF CLICKHOUSE

- SQL support.

  NULLs are not supported. All the functions have different names.
  However, this is a declarative query language based on SQL that can't be differentiated from SQL in many instances.
  JOINs are supported. Subqueries are supported in FROM, IN, JOIN clauses.

DISTINCTIVE FEATURES OF CLICKHOUSE

- Real-time data updates.
- Suitable for online queries.
- Support for approximated calculations.
- Support for nested data structures. Support for arrays as data types.
- Support for restrictions on query complexity, along with quotas.
- Data replication and support for data integrity on replicas.
  Uses asynchronous multimaster replication.

INTERFACES

1. HTTP

   ```
   $ curl 'http://localhost:8123/'
   $ wget -O- -q 'http://localhost:8123/?query=SELECT 1'
   ```

2. JDBC driver
3. Third-party client libraries

   There exist third-party client libraries for Python, PHP, Go, Node.js, Perl.

4. Native interface (TCP)
5. Command-line client

TABLE ENGINES

- TinyLog

  for small intermediate batches

- Memory

  stores data in RAM, in uncompressed form.

- ....
- Replicated

TABLE ENGINES

- Distributed

```
<remote_servers>
 <logs>
  <shard>
    <weight>1</weight>
        <internal_replication>false</internal_replication>
        <replica>
         <host>example01-01-1</host>
         <port>9000</port>
        </replica>
        <replica>
         <host>example01-01-2</host>
         <port>9000</port>
        </replica>
   </shard>
   <shard>
```

TABLE ENGINES

- MergeTree

```
Example without sampling support:
MergeTree(EventDate, (CounterID, EventDate), 8192)

Example with sampling support:
MergeTree(EventDate, intHash32(UserID), (CounterID, EventDate, intHash32(UserID)
```

- CollapsingMergeTree

  The change log makes it possible to incrementally calculate almost any statistics. To do this, we need to consider "new" rows with a plus sign, and "old" rows with a minus sign. In other words, incremental calculation is possible for all statistics whose algebraic structure contains an operation for taking the inverse of an element.

TABLE ENGINES

- SummingMergeTree
- AggregatingMergeTree
- ReplacingMergeTree (not documented yet)
  00325_replacing_merge_tree.sql

EXPORT/IMPORT FORMATS

- Native
- CSV
- JSON
- JSONEachRow
- TSV

DEMO

DOCUMENTATION

- Tutorial
- Reference
- Forum
- Benchmarks

QA

Yegor Andreenko

@f1yegor