

|||| ClickHouse

Meetup

WeWork Enam Sambhav (BKC)

Mumbai, India

February 1 at 10:30 AM IST



Meet the ClickHouse crew



Paul Davis

AVP Sales APJ



Rakesh Puttaswamy

Solution Architect



Anshul Mathur

Account Executive



Lalit Mishra

Country Manager, India



Piyush Vasudev

Support Engineering Manager



Kunal Gupta

Senior Software Engineer



Amogh Bharadwaj

Software Engineer



Kevin Biju

Software Engineer

Tech Talks



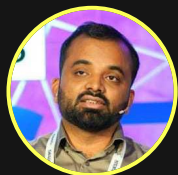
Introduction to ClickHouse

Rakesh Puttaswamy, Solution Architect
@ ClickHouse



The telemetry data platform: Breaking Down Operational Silos

Aditya Godbole, CTO @ Last9



Adopting ClickHouse @ Your Work

Sachin Tripathi, Senior Data Engineer @ Earnin

Last9

earn in



ClickHouse

A Quick Overview

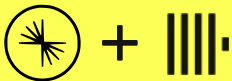
2025

 ClickHouse

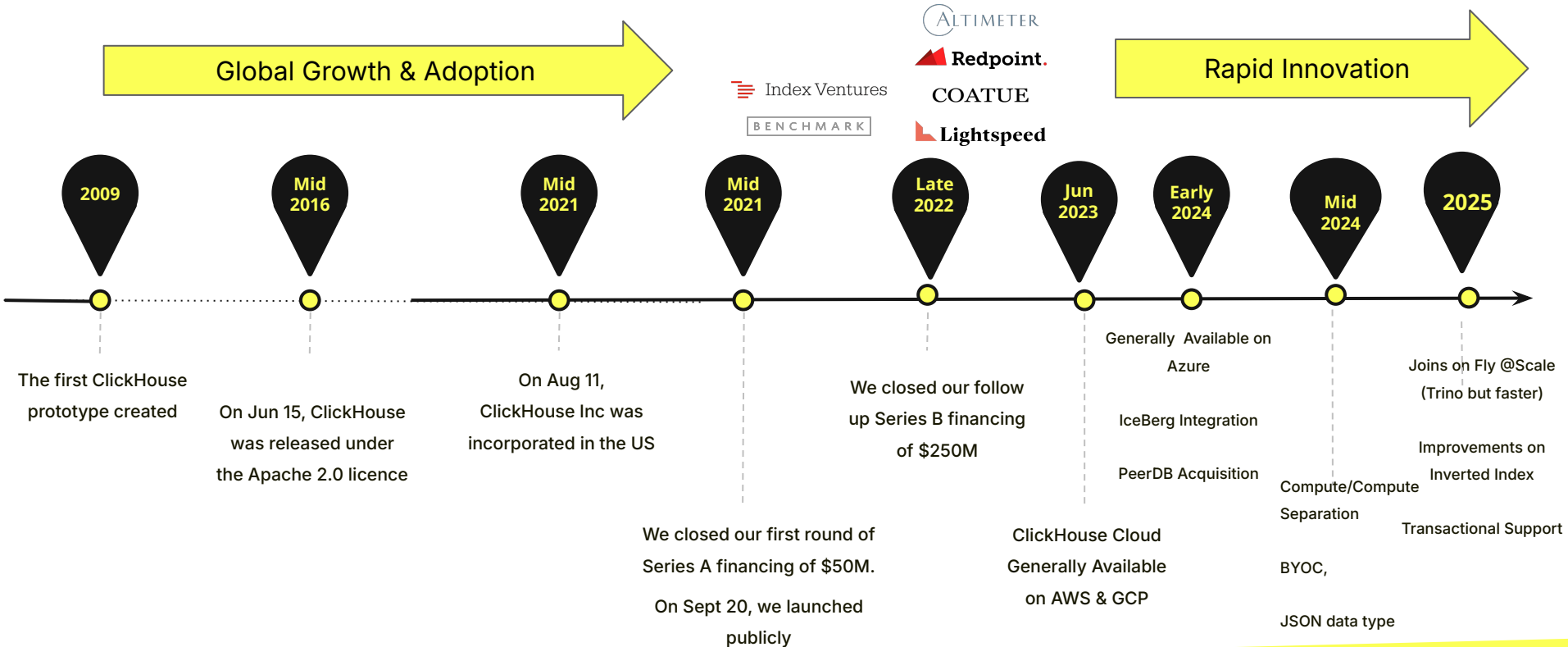
Rakesh Puttaswamy

Solution Architect @ ClickHouse

 [rakesh-p-gowda](https://www.linkedin.com/in/rakesh-p-gowda)

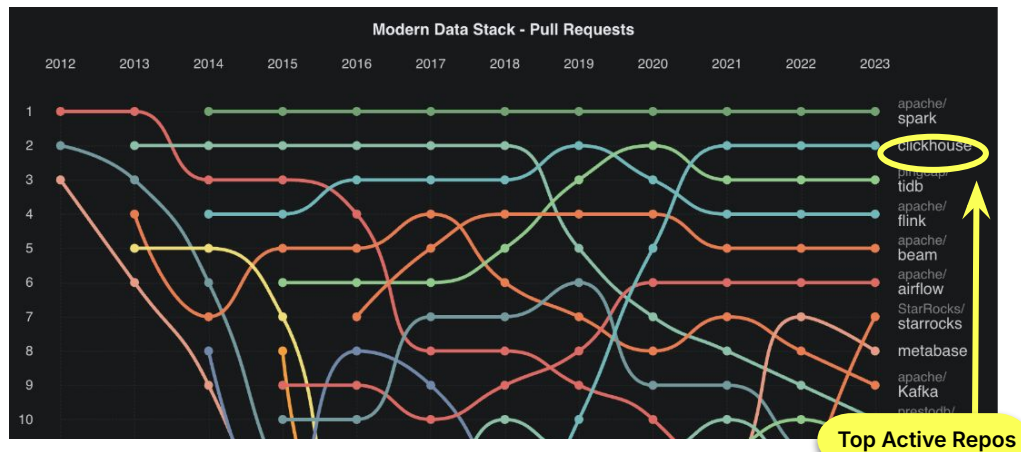


The ClickHouse Journey so far.....



ClickHouse Open Source

- ✓ 38.7k Stars
- ✓ 6.1k Forks
- ✓ 1.3k Contributors
- ✓ 100k+ Commits
- ✓ 114k active community members



ClickHouse Cloud



Bloomberg

- ✓ 300+ customers
- ✓ 150 active trials
- ✓ Processed 64+ billion queries, 28 quadrillion records, and 580 PiB of data since launch

NETFLIX

Uber



Deutsche Bank



NGINX



ROKT



What is ClickHouse?

Your (soon-to-be) favorite database!

Open source **column-oriented** **distributed** **OLAP** database

Since 2009
31,000+ GitHub stars
1300+ contributors
500+ releases

Best for aggregations
Files per column
Sorting and indexing
Background merges

Replication
Sharding
Multi-master
Cross-region

Analytics use cases
Aggregations
Visualization
Mostly immutable data

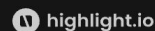


Use cases



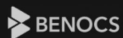
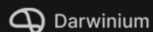
Real-time Analytics

Power interactive applications and dashboards that analyze and aggregate large amounts of data on the fly. Run complex internal analytics in ms, not mins or hrs.



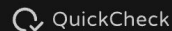
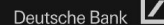
Logs, events, traces

Monitor with confidence your logs, events, and traces. Detect anomalies, fraud, network or infrastructure issues, and more.



Business intelligence

Interactively slice and dice your data for analysis, reporting, and building internal applications. Evaluate user behaviors, ad and media perf, market dynamics, and more.



ML and Gen AI

Execute fast and efficient vector search. Plug-and-play Generative AI models from any provider. Use lightning-fast aggregations to power model training at petabyte scale.



Key Features

Some of the cool things ClickHouse can do

1 Speaks SQL

Most SQL-compatible UIs, editors, applications, frameworks will just work!

2 Lots of writes

Up to several million writes per second - per server. "1 billion writes per second"

3 Distributed

Replicated and sharded, largest known cluster is 4000 servers.

4 Highly efficient storage

Lots of encoding and compression options - e.g. 20x from uncompressed CSV.

5 Very fast queries

Scan and process even billions of rows per second and use vectorized query execution.

6 Joins and lookups

Allows separating fact and dimension tables in a star schema.



Creating tables

ClickHouse SQL Basics

```
CREATE TABLE customers
(
    name      String,
    age       UInt8,
    address    Array(String),
    city       LowCardinality(String),
    created    DateTime,
    type       Enum8(...),
    attr       JSON
)
ENGINE = MergeTree
ORDER BY (city, name, type)
```

■ Other engines:

- ◇ ReplacingMergeTree
- ◇ CollapsingMergeTree
- ◇ AggregatingMergeTree

■ Integration engines:

- ◇ Kafka, RabbitMQ
- ◇ MySQL, PostgreSQL, MongoDB
- ◇ JDBC, ODBC
- ◇ S3, HDFS
- ◇ EmbeddedRocksDB

- Adding Replicating- in front makes an engine replicate data (e.g. ReplicatingMergeTree)



Inserting directly

ClickHouse SQL Basics

```
INSERT INTO people VALUES ('Obi-Wan Kenobi', 57, ...) ('Yoda', 900, ...)  
(...)
```

- Batching is important! Either batch yourself or turn on asynchronous inserts:

```
SET async_insert = true  
INSERT INTO people VALUES ('Obi-Wan Kenobi', 57, ...)  
INSERT INTO people VALUES ('Yoda', 900, ...)  
INSERT INTO people VALUES (...)
```



Reading data

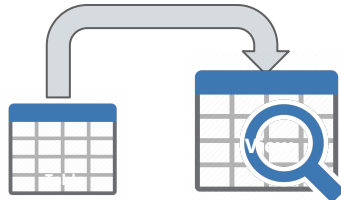
ClickHouse SQL Basics

- Saving a query as a view (no data movement):



```
CREATE VIEW view AS SELECT ... FROM table ...
```

- Continuously processing new data from a table into another table:



```
CREATE MATERIALIZED VIEW view AS SELECT avg(...) FROM table GROUP BY ...
```

```
CREATE MATERIALIZED VIEW view REFRESH EVERY 1 MINUTE TO table_name AS ..
```

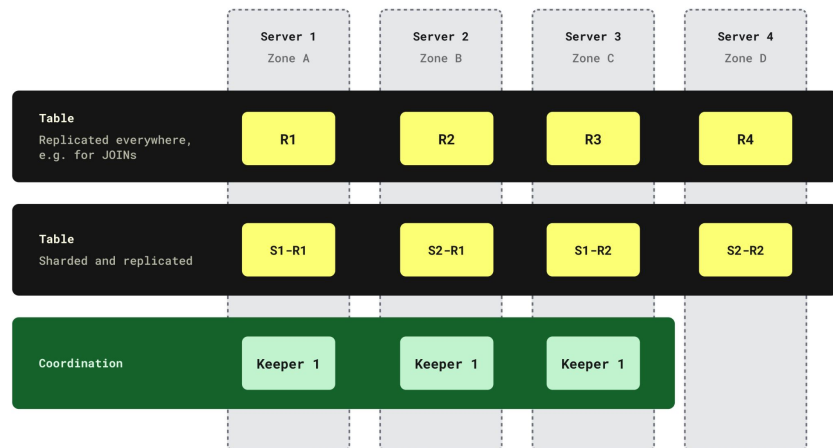


ClickHouse

Self-managed

- ✓ Open-source
- ✓ Flexible architecture
- ✓ Efficient and robust
- ✓ Support contracts available

Sample self-managed architecture

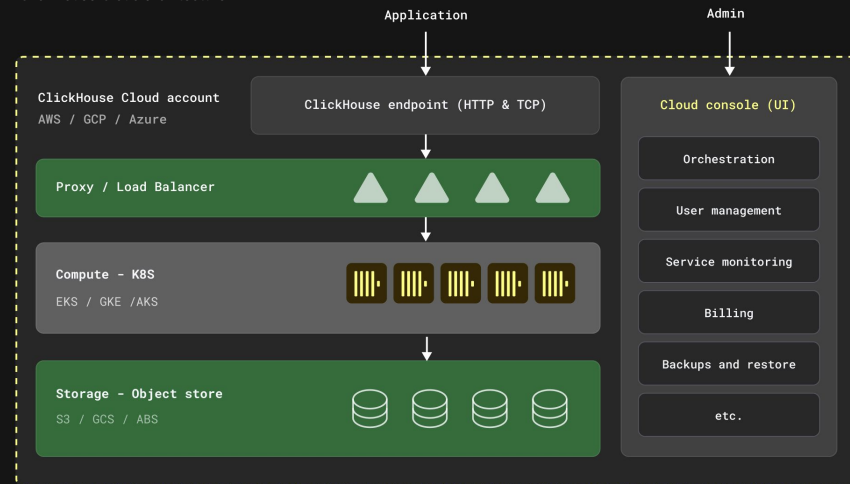


ClickHouse

Cloud

- ✓ Easy to use
 - ✓ Feature-rich
 - ✓ Fast
 - ✓ Scalable
 - ✓ Reliable
 - ✓ PAYG
- Managed for you
Cloud-first features & tooling
Automatically maximizes performance/efficiency
Scale seamlessly
Ensure reliability
SaaS usage and capacity based pricing

ClickHouse Cloud architecture



Let's do a quiz!

ClickHouse Quiz

What is ClickHouse?

A OLTP

B Key-Value

C Document

D OLAP

Pub Quiz

Is ClickHouse...

A

**Fully
consistent**

B

**Eventually
consistent**

Pub Quiz

Is ClickHouse...

A

Schema

B

Schemaless

Pub Quiz

Is ClickHouse...

A

Immutable

B

Mutable

Pub Quiz

Why is ClickHouse fast?

A

C++

B

Sorting

C

SIMD

D

Lots of love

E

**All of the
Above**

Pub Quiz

How much data can you store in ClickHouse?

A

Gigabytes

B

Terabytes

C

Petabytes

D

Exabytes

KEEP IN TOUCH!



clickhouse.com/slack



[@clickhousedb](https://twitter.com/clickhousedb)



[@clickhouseinc](https://www.linkedin.com/company/clickhouseinc) #clickhouseDB



[clickhouse](https://github.com/clickhouse)

 ClickHouse

QUERY OPTIMIZATION WORKSHOP

February 12, 2025 - 9:00 AM IST



clickhou.se/clickhouse-queryoptimise-workshop

Thank you!