

# ClickHouse Update

Tanya Bragin, Product @ ClickHouse



# ClickHouse Update

## Current recent features & roadmap

1

ClickHouse Core

2

ClickHouse Ecosystem Integrations

3

ClickHouse Cloud



# ClickHouse Core

# Core recent features and roadmap

Github - Roadmap Discussion <https://github.com/ClickHouse/ClickHouse/issues/32513>

YouTube - Monthly Release Webinars <https://www.youtube.com/c/ClickHouseDB>

Sign up for next webinar on Oct 20!

## Completed

- ✓ Make clickhouse-keeper production ready
- ✓ Support for backup and restore
- ✓ AArch64 hardware support
- ✓ Replicated database engine
- ✓ Type inference for data import
- ✓ Lightweight DELETE

## In progress

- IP Semistructured Data
- IP SQL Compatibility Improvements
- IP Support for Transactions
- IP Resource Management
- IP JOIN Improvements



# Fast data onboarding and analytics

## Core recent features and roadmap

### Schema inference (22.1)

**Add data to ClickHouse without having to define the schema upfront**

- ✓ Schema inference
- ✓ Table structure auto detection
- ✓ Format autodetection

### Semi-structured data (22.3)

**Add support for semi-structured data sources like JSON**

- IP SQL/JSON support
- IP JSON Object Type (experimental)

### Faster queries

**Faster results for more use cases; never stop improving performance**

- ✓ Projections (production ready)
- ✓ Faster JOINS ("direct", "parallel hash", "full\_sorting\_merge")
- IP Faster JOINS ("grace hash")

# More standard SQL support

## Core recent features and roadmap

### Standard SQL commands

**More functionality and seamless migrations / ecosystem integrations**

- ✓ Support for GROUPING SETS
- ✓ Window functions inside expressions

### Lightweight deletes (22.8)

**Update & delete data frequently without impacting performance**

**IP** Lightweight DELETE statement, without need to specify ALTER, implemented via delete-masks (experimental)

### Transactions (22.4)

**Bring ACID properties to use cases important to ClickHouse users**

**IP** BEGIN TRANSACTION, COMMIT and ROLLBACK statements

Ex. use case: Atomic INSERTs to multiple tables, materialized views, partitions

*In Preview (works with MergeTree only)*



# Operations and administration

## Core recent features and roadmap

### Operations

**Operate ClickHouse more easily, reliably, and at lower cost**

- ✓ clickhouse-keeper (production ready)
- ✓ Flexible memory limits
- ✓ Backup & restore

### Resource efficiency

**Make your hardware go even further with ClickHouse**

- ✓ Sparse encoding for columns
- ✓ Dozens of performance optimizations in every release

### New platform support

**Operate ClickHouse on more platforms, e.g. ARM (AArch64)**

- ✓ ARM (AArch64) hardware support is now production ready. 100% functional tests are run in CI for AArch64.



# **ClickHouse Ecosystem Integrations**



**Guess how many projects on Github  
integrate with ClickHouse?**

10

100

1000

APACHE  
Spark



- **2324** repositories with ClickHouse in their name on Github (excluding forks)

- From which **154** repositories with >50 stars

\* As of Sept 4th, 2022

The screenshot shows a GitHub search interface with the query 'clickhouse' in the search bar. The left sidebar displays various filters: Repositories (2K), Code (280K), Commits (185K), Issues (52K), Discussions (410), Packages (85), Marketplace (3), Topics (48), Wikis (415), and Users (96). Below these, a 'Languages' section lists: Go (303), Python (287), Java (234), Shell (158), PHP (99), JavaScript (90), Dockerfile (79), C++ (65), TypeScript (48), and C# (43). The main content area shows '2,324 repository results' (highlighted with a red box) and a list of repositories. The first three results are: 1. 'artpaul/clickhouse-cpp' (ClickHouse C++ client library, 94 stars, C++, updated 17 Mar). 2. 'FriendsOfDoctrine/dbal-clickhouse' (Doctrine DBAL driver for ClickHouse database, 77 stars, PHP, GPL-3.0 license, updated 18 days ago). 3. 'enqueue/metabase-clickhouse-driver' (ClickHouse database driver for the Metabase business intelligence front-end, 333 stars, Clojure, AGPL-3.0 license, updated 26 Apr). Below these are 'go-graphite/carbon-clickhouse' (Graphite metrics receiver with ClickHouse as storage, 163 stars, Go, MIT license, updated 29 Mar) and 'go-graphite/graphite-clickhouse' (Graphite cluster backend with ClickHouse support, 188 stars, Go, MIT license, updated 2 days ago, 2 issues need help).

Filter	Count
Repositories	2K
Code	280K
Commits	185K
Issues	52K
Discussions	410
Packages	85
Marketplace	3
Topics	48
Wikis	415
Users	96

Language	Count
Go	303
Python	287
Java	234
Shell	158
PHP	99
JavaScript	90
Dockerfile	79
C++	65
TypeScript	48
C#	43











Repository	Description	Stars	Language	License	Updated
artpaul/clickhouse-cpp	ClickHouse C++ client library	94	C++		17 Mar
FriendsOfDoctrine/dbal-clickhouse	Doctrine DBAL driver for ClickHouse database	77	PHP	GPL-3.0 license	18 days ago
enqueue/metabase-clickhouse-driver	ClickHouse database driver for the Metabase business intelligence front-end	333	Clojure	AGPL-3.0 license	26 Apr
go-graphite/carbon-clickhouse	Graphite metrics receiver with ClickHouse as storage	163	Go	MIT license	29 Mar
go-graphite/graphite-clickhouse	Graphite cluster backend with ClickHouse support	188	Go	MIT license	2 days ago

# Ecosystem Integrations Team

## Support core, partner, and community integrations





### Core integrations

Core integrations are built or maintained by ClickHouse. These integrations are supported by ClickHouse and live in the ClickHouse GitHub organization.

Name	Logo	Category	Description	Resources
Amazon S3		Data ingestion	Import from, export to, and transform S3 data in flight with ClickHouse built in S3 functions and clickhouse-local.	<a href="#">Documentation</a>
ClickHouse Client		SQL client	ClickHouse Client is the native command-line client for ClickHouse.	<a href="#">Documentation</a>
DBever		SQL client	Free multi-platform database administration tool. Connects to Clickhouse through JDBC driver.	<a href="#">Documentation</a>
dbt		Data ingestion	Use dbt (data build tool) to transform data in ClickHouse by simply writing select statements. dbt puts the T in ELT.	<a href="#">Documentation</a>
Go		Language client	The Go client uses the native interface for a performant, low-overhead means of connecting to ClickHouse.	<a href="#">Documentation</a>
Java		Language client	The Java client is an async, lightweight, and low-overhead library for ClickHouse.	<a href="#">Documentation</a>
Kafka		Data ingestion	Bidirectional integration with Apache Kafka, the open-source distributed event streaming platform.	<a href="#">Documentation</a>
Node.JS		Language client	The official Node.js client for connecting to ClickHouse.	<a href="#">Documentation</a>
Python		Language client	A suite of Python packages for connecting Python to ClickHouse.	<a href="#">Documentation</a>
Superset		Visualization client	Explore and visualize your ClickHouse data with Preset or Apache Superset.	<a href="#">Documentation</a>



### Partner integrations

Partner integrations are built or maintained, and supported by, third-party software vendors.

Name	Logo	Category	Description
Arctype		SQL client	Arctype is a fast and easy-to-use database client for writing queries, building dashboards, and sharing data.
DataGrip		SQL client	DataGrip is a powerful database IDE with dedicated support for ClickHouse.
Deepnote		Data visualization	Deepnote is a collaborative Jupyter-compatible data notebook built for teams to discover and share insights.
Grafana		Data visualization	With Grafana you can create, explore and share all of your data through dashboards.

### Community integrations

Community integrations are built or maintained and supported by community members. No direct support is available besides the public github repositories and community Slack channels.

Name	Logo	Category	Description
Airbyte		Data ingestion	Use Airbyte, to create ELT data pipelines with more than 140 connectors to load and sync your data into ClickHouse.
Apache Spark		Data ingestion	Spark ClickHouse Connector is a high performance connector built on top of Spark DataSource V2.

# Ecosystem Integrations status and roadmap

	Available	Next
<b>Data Ingestion / Orchestration</b>	<ul style="list-style-type: none"> <li>✓ S3</li> <li>✓ Kafka</li> <li>✓ DBT (<b>recent</b>)</li> <li>✓ Airbyte (<b>recent</b>)</li> </ul>	<ul style="list-style-type: none"> <li>● Kafka Connect (<b>early access</b>)</li> <li>● Vector</li> <li>● Decodable (partner)</li> <li>● Metaplane (partner)</li> </ul>
<b>Data visualization</b>	<ul style="list-style-type: none"> <li>✓ Grafana (partner) (<b>v2.0</b>)</li> <li>✓ Superset / Preset</li> <li>✓ Deepnote (partner) (<b>recent</b>)</li> <li>✓ HEX (partner) (<b>recent</b>)</li> <li>✓ Metabase (community)</li> </ul>	<ul style="list-style-type: none"> <li>● Tableau</li> <li>● AWS Quicksight (partner)</li> <li>GCP Looker</li> <li>Azure Power BI</li> </ul>
<b>Language clients</b>	<ul style="list-style-type: none"> <li>✓ Go</li> <li>✓ Java</li> <li>✓ Python</li> <li>✓ Node.js (<b>new</b>)</li> </ul>	<ul style="list-style-type: none"> <li>.NET</li> <li>Ruby</li> <li>...</li> </ul>
<b>SQL clients</b>	<ul style="list-style-type: none"> <li>✓ DBeaver</li> <li>✓ Datagrip</li> <li>✓ Arctype (partner)</li> </ul>	<ul style="list-style-type: none"> <li>SQL Workbench</li> <li>...</li> </ul>

## Legend

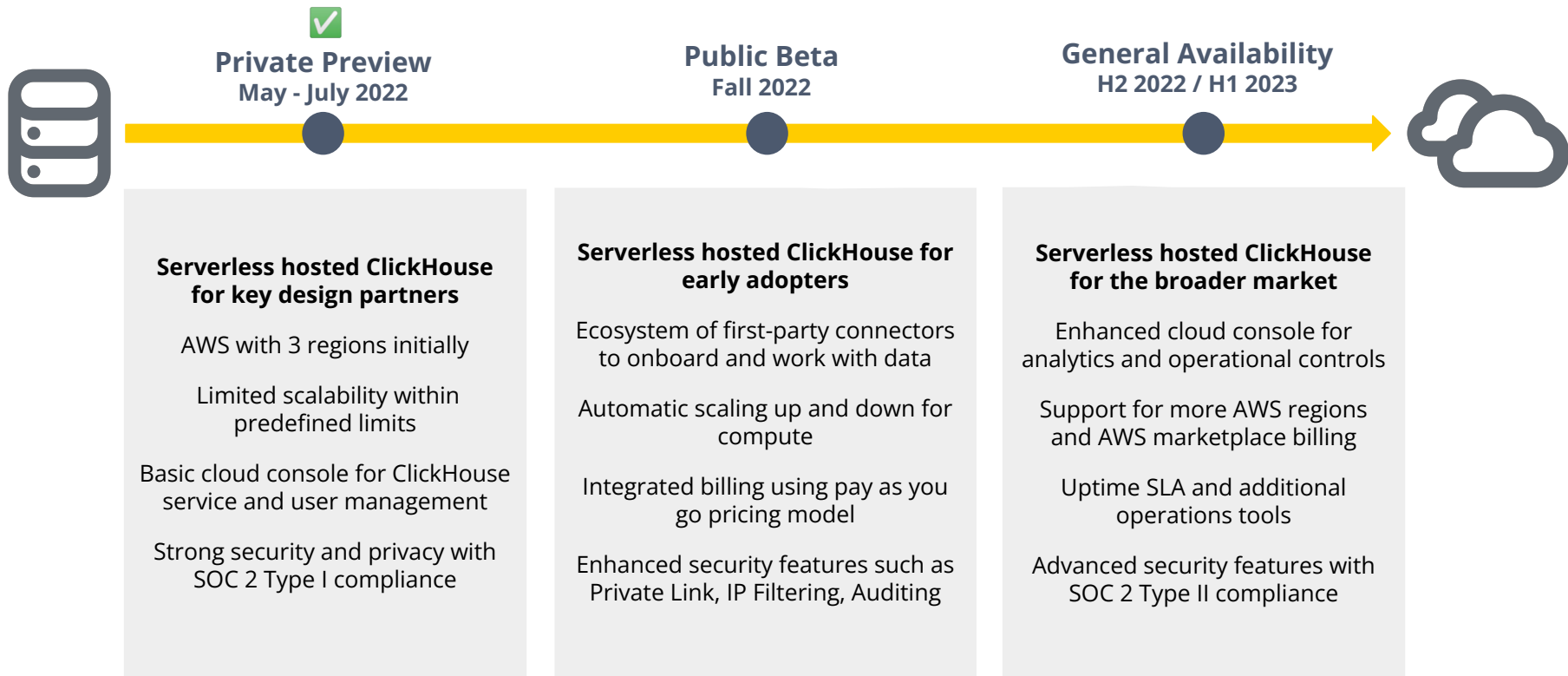
- ✓ Completed
- In development



The background of the image features a dark blue-grey gradient. On the left side, there are two overlapping circles of a slightly lighter shade of blue-grey. The text "ClickHouse Cloud" is positioned on the left, overlapping the circles.

**ClickHouse Cloud**

# ClickHouse Cloud Milestones



# DEMO

ClickHouse Cloud



# ClickHouse Cloud

## Expansion to more cloud providers

### GCP - Private preview for select customers in late H2 2022

- Iowa (us-central1)
- Amsterdam/Netherlands (europe-west4)
- Singapore (asia-southeast1)
- Oregon (us-west1)
- Frankfurt (europe-west3)
- S. Carolina (us-east1) or Virginia (us-east4)



### Azure - Targeted to H1 2023

- Please register interest in private preview





<https://clickhouse.com/company/events/aws-reinvent>

## Event

# AWS re:Invent

Core members of the ClickHouse team -- including 2 of our founders -- will be at re:Invent from November 29 to December 3. We are available on the show floor, but are also determining interest in holding an event during the time there.

Interested in meeting up with us and talking about ClickHouse?  
Please fill out the form and let us know!

First name

Last name

Email

Register to event

## Hosted by



**Aaron Katz**  
CEO



**Yury Izrailevsky**  
President, Product & Technology



**Tanya Bragin**  
Vice President, Product



**Mike Hayes**  
Vice President, Sales





**Thank you!**