aws marketplace

ClickHouse

Reviews, tips, and advice from real users






Powered by PeerSpot

Contents

Product Recap

ClickHouse

ClickHouse Recap

ClickHouse is renowned for its speed, scalability, and real-time query performance. Its compatibility with SQL standards enhances flexibility while enabling integration with popular tools.

ClickHouse leverages a column-based architecture for efficient data compression and real-time analytics. It seamlessly integrates with tools like Kafka and Tableau and is effective in handling large datasets due to its cost-efficient aggregation capabilities. With robust data deduplication and strong community backing, users can access comprehensive documentation and up-to-date functionality. However, improvements in third-party integration, cloud deployment, and handling of SQL syntax differences are noted, impacting ease-of-use and migration from other databases.

What features make ClickHouse outstanding?

- Speed and Scalability: Enables rapid deployment and scalability, particularly on Kubernetes.
- SQL Compatibility: Aligns with SQL standards for flexible database management.
- Column-based Structure: Offers efficient data compression and quick query performance.
- Integration Support: Works with Kafka, Tableau, and other popular platforms.
- Data Deduplication: Manages complex data duplication efficiently.

What benefits should users consider?

- Cost-Efficiency: Reduces expenses while managing large datasets effectively.
- Real-Time Analytics: Provides fast data processing for real-time insights.
- Flexibility: Offers multiple integration points and compatibility options.
- Robust Documentation: Backed by an active development community, ensuring comprehensive support.

ClickHouse is deployed in sectors like telecommunications for passive monitoring and is beneficial for data analytics, logging Clickstream data, and as an ETL engine. Organizations harness it for machine learning applications when combined with GPT. With the ability to be installed independently, it's an attractive option for avoiding cloud service costs.

Valuable Features

Excerpts from real customer reviews on PeerSpot:

- “ClickHouse is very easy to use; one of the good features is that it has joins, which were not present in Druid, and Druid was quite expensive, especially with our applications at Sam's Club utilizing ClickHouse very quickly.”

Aswini Atibudhi

Distinguished AI Leader at Walmart Global Tech at Walmart

- “ClickHouse is open source with no vendor lock-in, providing excellent freedom to choose any vendor without restrictions.”

ArpitShah

Software Analyst at CLSA

- “The best thing about the tool is that I can set it up on my computer and run queries without depending on the cloud. This is why I use it every day.”

Dmitriy Yugin

Senior Data/Web Analyst at Raiffeisen Bank

"It's easier to work with big data and calculations using the product."

Anton P

Senior Data Engineer at Brightika, Inc.

"We faced a challenge with deploying ClickHouse onto Kubernetes. Recently, we've been using ClickHouse Cloud, and the main issue is the high cost of the cloud service. The pricing isn't very competitive, especially for startups. I would instead buy a server and self-host if I have enough disk space. Besides that, ClickHouse has done very well, with clear goals and effective execution."

Eden Chen

Senior Software Engineer at a energy/utilities company with 1,001-5,000 employees

"ClickHouse is much faster than traditional databases like MySQL and MongoDB. Its column-row searching strategy makes it very efficient. With ClickHouse, we can manage multiple databases, automatically insert data from other databases and delete data as needed. It supports real-time query performance, allowing simultaneous data insertion and retrieval. ClickHouse has improved significantly over the past two years, adding more functions and queries, as well as top functionality."

Parveen Yadav

Software Engineer at Activant Solutions Pvt Ltd, Jaipur

"The main feature of ClickHouse is the optimizer because we had too many records to deduplicate, and the optimizer took this task by itself."

Ayham Al-Adm

Full-stack Web Developer at a tech services company with 51-200 employees

What users had to say about valuable features:

"It's easier to work with big data and calculations using the product. For example, we can easily calculate metrics around terabytes of data using ClickHouse. The dictionaries help us to make our analysts' jobs faster and easier and give value to the business faster.."

Anton P

Senior Data Engineer at Brightika, Inc.

Read full review

"ClickHouse is a user-friendly solution that tries to be compatible with SQL standards. It also tries to provide command-line tools, very nice formatting, libraries. And of course blazing speed it is main selling point of this technology."

Andrei Kochemirovskii

Lead Data Engineer at Omni Verse

Read full review

ClickHouse is very easy to use; one of the good features is that it has joins, which were not present in Druid, and Druid was quite expensive, especially with our applications at Sam's Club utilizing ClickHouse very quickly.ClickHouse deserves a rating of 9 when compared to competitors, particularly Druid, which is stable but comes with higher costs and subpar support. ClickHouse proves to be more lightweight, offering low latency and high throughput, along with joins, making it especially good for log and metrics handling.

Aswini Atibudhi
Distinguished AI Leader at Walmart Global Tech at Walmart

Read full review

One of the most valuable features of ClickHouse is that it is open source without vendor lock-in, allowing me the freedom to choose any vendor for the database. It offers numerous out-of-the-box analytical functions, eliminating the need for complex coding. The performance of ClickHouse aligns with its claims, being highly efficient and used by large organizations like Uber and Zomato. The deployment process is straightforward, and it is scalable both vertically and in distributed systems via the cloud.

ArpitShah
Software Analyst at CLSA

Read full review

“The main feature of ClickHouse is the optimizer, if we had too many records to deduplicate, the optimizer took this task by itself. The second valuable feature of the solution is its performance. It's not easy when we talk about five or six gigabytes of one table of data.

Then, if you have to generate too many KPIs, charts, lines, and reports, it's not easy to deal with all of these with just one engine and tool. ClickHouse was really nice in this respect, and we had no problem with its performance..”

Ayham Al-Adm

Full-stack Web Developer at a tech services company with 51-200 employees

Read full review

“The aggregation capability is a valuable feature. It's highly efficient, allowing us to review entire transaction histories and user activities in the market. We've tried MongoDB, Postgres, MariaDB, and BigQuery, but ClickHouse is the most cost-efficient solution for collecting data at high speeds with minimal cost. We even used ClickHouse Cloud for a month, and it proved to be a great setup, especially for startups looking to handle big data. For example, if there is a need for 2-4 terabytes of data and around 40 billion rows with reasonable computing speed and latency, ClickHouse is ideal.

Regarding the real-time query performance of ClickHouse, when using an API server to query it, I achieved query results in less than twenty milliseconds in some of my experiments with one billion rows. However, it depends on the scenario since ClickHouse has limitations in handling mutations.

Additionally, one of ClickHouse's strengths is its compression capability. Our experimental server has only four terabytes, and ClickHouse effectively compresses data, allowing us to store large amounts of data at high speed. This compression efficiency is a significant advantage of using ClickHouse..”

Eden Chen

Senior Software Engineer at a energy/utilities company with 1,001-5,000 employees

Read full review

Other Solutions Considered

I previously used Postgres, which started slowing down with massive amounts of data. I evaluated over twelve databases, starting with TiDB, but found ClickHouse to be the best fit after considering options like DuckDB. I initially preferred Postgres for its comprehensive features, but it couldn't handle the data scale.

ArpitShah
Software Analyst at CLSA

Read full review

“We chose ClickHouse because we needed to move away from cloud-based solutions due to risks in Russia. We considered other options, such as Postgres, NoSQL databases, Hadoop, and Hive, before deciding on ClickHouse..”

Dmitriy Yugin
Senior Data/Web Analyst at Raiffeisen Bank

Read full review

“The company decided to use ClickHouse because mobile networks produce enormous amounts of data—millions of timestamped vectors, each representing a measurement, which total billions of rows per month. Initially, they used MySQL, but as the data volume grew, MySQL couldn't handle the load. Therefore, they switched to ClickHouse..”

Spyros Almpanis
Backend Software Engineer at a tech vendor with 51-200 employees

Read full review

Use Case

“We use the solution not just as an analytics database but also as a data warehouse. A lot of our internal services communicate with the tool. Our analytics and ML teams also use the solution. It's the hub of our company..”

Anton P
Senior Data Engineer at Brightika, Inc.

Read full review

The main use case for ClickHouse is as a data warehouse for handling large volumes of data that exceed the capabilities of traditional databases like Postgres. I use it for creating dashboards and performing analytical tasks such as determining the total number of orders, average order value, and evaluations and ratios for various stores. I deploy ClickHouse both on the cloud provided by ClickHouse itself and on-premises for IoT and similar data tasks.

ArpitShah
Software Analyst at CLSA

Read full review

"I used ClickHouse to collect data, put it in the database, and then analyze it to find insights. The main advantage is that I can install it on my computer instead of using cloud-based solutions, so I don't have to pay for every query like with Google or Amazon cloud services.."

Dmitriy Yugin
Senior Data/Web Analyst at Raiffeisen Bank

Read full review

"I use ClickHouse to collect and analyze data from Ethereum. We primarily use it for data classification and occasionally for machine learning with GPT, but that's minimal. The primary use case is classification; sometimes, we use it for applications similar to OLTP scenarios. All of our data is stored in ClickHouse. We are customers of ClickHouse, not partners. It's an easy tool to use if you know SQL databases.."

Eden Chen
Senior Software Engineer at a energy/utilities company with 1,001-5,000 employees

Read full review

"Our company had about nine platforms, each with its own database and data. We had to gather all these data in one database and just one table. We used Apache Superset to integrate this database with the business intelligence tool. We had too many choices or options initially for the database engine.

We initially tested a database, and its performance was good. When we tried ClickHouse, we switched to it immediately because the performance was really amazing. When we had a huge amount of data, about five or six gigabytes in just one table, and we used ClickHouse to deduplicate some duplicated entries or records.."

Ayham Al-Adm

Full-stack Web Developer at a tech services company with 51-200 employees

Read full review

I have experience in ClickHouse, and we also use Apache Druid, which has corporate support from Druid, along with data products in Hadoop. We are currently exploring many platforms such as GMI, TKI, and Vertex.I use ClickHouse as a merchant side portal, especially when we started exploring how to use the data, which was coming from multiple sources such as logs, mainframe, Teradata, and many file systems that come to the data lake. The real-time challenge was joining the data and providing more analytical queries for our merchants, who work throughout the year to improve GMB, sales, and ensure the right quantity of items is ordered at the right time. That's the challenge for the merchants, and we aim for fast analytical queries on larger databases, which is why we selected ClickHouse as our columnar OLAP database supporting real-time analytics with its own SQL interface.We have installed both local Docker versions, which are quite scalable, and usually connect with BI tools such as Grafana, Superset, and Tableau while utilizing materialized views, DDLs partitions, and many other connectors with Python, such as ClickHouse connectors and drivers. It's exciting to see how ClickHouse has evolved, and we are evaluating ClickHouse Cloud while also having the on-premises version.We are already a customer of ClickHouse, with Sam's Club utilizing it on the merchant side while also exploring ClickHouse for consumers, primarily for user analytics, metrics, and streaming data analysis in ad tech. Additionally, we use custom analysis and metrics for fraud detection in payments and ad campaign metrics, with various teams utilizing it for ad campaign management and user behavior analytics, particularly on e-commerce sites focusing on customer behavior. It's extensively used due to its low latency, fast aggregations, and excellent OLAP columnar storage, featuring quick joins and real-time data visibility, making ClickHouse very appealing to us.

Aswini Atibudhi
Distinguished AI Leader at Walmart Global Tech at Walmart

Read full review

Setup

The setup process involves configuring and preparing the product or service for use, which may include tasks such as installation, account creation, initial configuration, and troubleshooting any issues that may arise. Below you can find real user quotes about the setup process.

"Installing the tool was easy. I used a Windows laptop with WSL and followed the documentation instructions. I didn't have any issues with the installation.."

Dmitriy Yugin

Senior Data/Web Analyst at Raiffeisen Bank

Read full review

"Regarding the initial installation, setup, and deployment, I can say it's easy for someone with my engineering skills. I prefer managing the installation myself rather than relying on out-of-the-box solutions. ."

Timur D.

ClickHouse DBA at a computer software company with 51-200 employees

Read full review

The initial setup for ClickHouse is relatively easier compared to Flink; however, for newcomers, it is quite challenging. I find it easier in terms of API with single-node setups through Yum or apt taking only a couple of minutes to install. Planning cluster setups is a bit complex, primarily an admin task, and while a single-node setup is easy, managing ClickHouse Cloud is extremely easy. Creating clusters can vary from moderate to difficult based on the scale, typically from 5 to 10 nodes, depending on the use case.

Aswini Atibudhi
Distinguished AI Leader at Walmart Global Tech at Walmart

Read full review

"The initial setup is easy for me because I have a lot of experience. Generally, when I mentor someone to deploy the whole cluster, it is difficult for them. It will not be hard to deploy one server or node because ClickHouse's team has shown a great demo of deploying it as a service on Linux. It gets harder if we want to tune some small sequences to get more performance. Real-time models require experience. We can open some community chats and find help. The deployment can be done in a couple of minutes. It's very fast.."

Anton P
Senior Data Engineer at Brightika, Inc.

Read full review

“The initial setup is pretty difficult since we deployed it in-house. We didn't use the cloud subscription, so we have to handle the deployment very carefully.

The challenge was deploying it and having the replication concept working. Another challenging feature is persistent volumes. You have to make sure the data is available on all clusters; otherwise, if one cluster goes down, you'll lose all your data. It's better to have it replicated.

We first used the cloud subscription, but we saw a possibility to reduce costs, so we tried deploying the open-source ClickHouse on-premises. That saved us money, but we didn't get all the features that come with the subscription..”

Verified user

Read full review

Software Development Engineer II at a financial services firm with 10,001+ employees

Customer Service and Support

“The cloud services support is excellent. Their support team is very timely and helpful, and even if you encounter any bugs, they assist you quickly. Compared to other services I've used, ClickHouse's support is very helpful. Even if you don't know much about databases or ClickHouse, their support will help resolve any issues..”

Eden Chen
Senior Software Engineer at a energy/utilities company with 1,001-5,000 employees

Read full review

“The product provides a lot of community support. It is useful. We can also contact a private company that provides support for ClickHouse. The solution has a community chat in Telegram that works well. We find solutions easily when the problem is already mentioned by someone. In rare cases, the issues stay unresolved because of NDA..”

Anton P
Senior Data Engineer at Brightika, Inc.

Read full review

"I have never interacted with the solution's technical support because I usually use the open-source version of ClickHouse. You can post your issue on GitHub at any time, and you will usually get a response.."

Andrei Kochemirovskii
Lead Data Engineer at Omni Verse

Read full review

"I have some experience talking with the tech support team. It was an open-source project at one point, so I used community resources for help. The best way to communicate with them was through their program channel, which had support available in both English and Russian.."

Timur D.
ClickHouse DBA at a computer software company with 51-200 employees

Read full review

The support team has its own community support on platforms such as Slack Overflow and ClickHouse Slack. Commercially, the company provides enterprise support, especially for Sam's Club through ClickHouse Cloud. We utilize AVN ClickHouse, which is effectively managed by AVN, providing bug fixes and developing new functionalities along with architecture reviews. I appreciate their 24/7 support which is beneficial, although those using open source might face some challenges. Overall, the enterprise support is quite good.

Aswini Atibudhi
Distinguished AI Leader at Walmart Global Tech at Walmart

Read full review

“I have talked to the ClickHouse support team before. They have a support group on Telegram messenger where you can ask technical questions. I often asked about working with tables and views and making sophisticated calculations. But now, I don't have any issues, so I don't need to ask for support.

I was satisfied with the support. Many people in the support group try to really understand your problem and help, not just dismiss it. If something isn't possible due to database limitations, they try to help you look at the situation differently..”

Dmitriy Yugin
Senior Data/Web Analyst at Raiffeisen Bank

Read full review

Other Advice

For the right use cases, I would rate ClickHouse eight to eight point five out of ten. However, it is not suitable as a primary database for startups due to the lack of transactional support. For companies with massive data struggling with query speed and facing high costs from vendor lock-ins, ClickHouse is an excellent choice.

ArpitShah
Software Analyst at CLSA

Read full review

"We do not use the real-time features much. Usually, we work with big data. We do not need to work with big data in real-time. We use CatBoost with ClickHouse. I always recommend the tool to others based on their requirements. If you have trouble with your queries and think that ClickHouse is slow, please review your queries. Overall, I rate the solution a ten out of ten.."

Anton P
Senior Data Engineer at Brightika, Inc.

Read full review

" I would recommend ClickHouse to others. If they have large datasets, ClickHouse is much more cost-effective and efficient than BigQuery. For example, running a query on one billion rows in BigQuery took a few minutes and was very expensive, whereas ClickHouse could do it in less than five seconds at a much lower cost.

I don't use AI in ClickHouse, but I use full-text search, and it's mighty. There's no

significant gap when migrating from other SQL databases to ClickHouse, though you must learn some specific syntax. If you are familiar with databases and know how to code and design systems, using ClickHouse should be straightforward.

Overall, I would rate ClickHouse an eight out of ten.."

Eden Chen
Senior Software Engineer at a energy/utilities company with 1,001-5,000 employees

Read full review

"The tool is open-source, so you don't need to pay for the software itself. However, you need to consider hardware costs and maintenance. A small company can install it on a company computer. For larger companies, you might need to hire a team for maintenance and consider data safety and privacy issues.

Integrating ClickHouse with other tools in our data stack was easy. It has native connections to many tools, such as Google and Amazon cloud solutions, and can easily connect with other databases.

For beginners, the ease of use depends on your background. If you're familiar with relational databases, it's easy. If not, you might need to read the documentation or ask for support.."

Dmitriy Yugin
Senior Data/Web Analyst at Raiffeisen Bank

Read full review

"If you're considering using ClickHouse for the first time, my advice would depend on how much data you need to handle. For most scenarios where big data isn't involved, I don't think it's a good idea to use ClickHouse. SQL Server, MySQL, or

PostgreSQL are well-documented and supported. The software you need to access these databases will be readily available. So, I don't see any reason to use ClickHouse for small to medium-scale scenarios.

I don't think you'll find it any more difficult than other databases, apart from the SQL syntax, which is a bit different. The most challenging part with ClickHouse is dealing with the large amounts of data it handles without overloading your server. I don't think the database itself is difficult to use. However, I was primarily accessing data from it and don't have much experience with setting it up or feeding it data.

I rate the overall solution a nine out of ten. .”

Spyros Almpanis
Backend Software Engineer at a tech vendor with 51-200 employees

Read full review

“For about six gigabytes, we took about two seconds to fetch all data at the maximum performance. Otherwise, it was really nice to have a medium CPU or database engine and resources. We don't have a really huge server; it's just traditional servers and traditional resources.

ClickHouse is not a straightforward tool for anyone to use. Users need some time to switch from traditional things to study new concepts.

We had just one client, Apache Superset. Apache Superset connects with just one connection but with too many requests. We had about 20 to 30 reports on the same page, and they work concurrently.

The solution’s documentation is amazing.

I would recommend the solution to other users. ClickHouse is the first step to the next generation of databases. When we deal with this amount of data and this performance, I think it's a respected technology.

Overall, I rate the solution a nine out of ten..”

Ayham Al-Adm
Full-stack Web Developer at a tech services company with 51-200 employees

Read full review

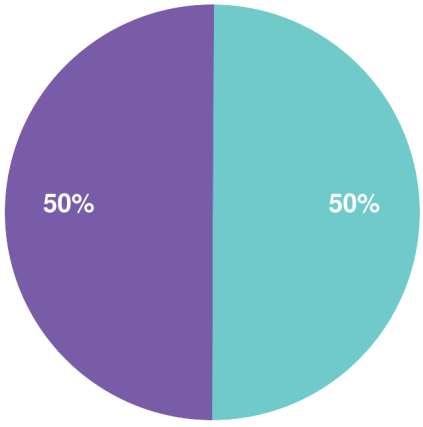
Top Industries

by visitors reading reviews

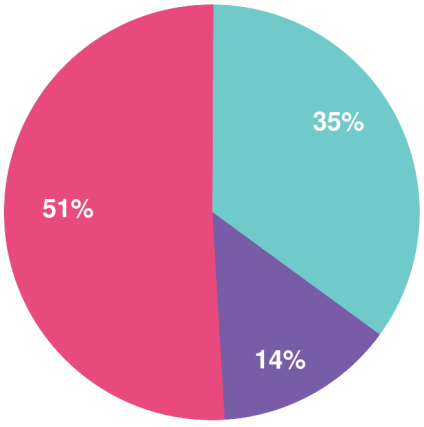
Computer Software Company	22%
Financial Services Firm	17%
Educational Organization	10%
Manufacturing Company	9%

Company Size

by reviewers



by visitors reading reviews



Large Enterprise Midsize Enterprise Small Business

About this buyer's guide

Thanks for downloading this PeerSpot report.

The summaries, overviews and recaps in this report are all based on real user feedback and reviews collected by PeerSpot's team. Every reviewer on PeerSpot has been authenticated with our triple authentication process. This is done to ensure that every review provided is an unbiased review from a real user.

Get a custom version of this report... Personalized for you!

Please note that this is a generic report based on reviews and opinions from the collective PeerSpot community. We offer a customized report of solutions recommended for you based on:

- Your industry
- Company size
- Which solutions you're already considering

The customized report will include recommendations for you based on what other people like you are using and researching.

Answer a few questions in our short wizard to get your customized report.
Get your personalized report here

About PeerSpot

PeerSpot is the leading review site for software running on AWS and other platforms. We created PeerSpot to provide a trusted platform to share information about software, applications, and services. Since 2012, over 22 million people have used PeerSpot to choose the right software for their business.

PeerSpot helps tech professionals by providing:

- A list of products recommended by real users
- In-depth reviews, including pros and cons
- Specific information to help you choose the best vendor for your needs

Use PeerSpot to:

- Read and post reviews of products
- Access over 30,000 buyer's guides and comparison reports
- Request or share information about functionality, quality, and pricing

Join PeerSpot to connect with peers to help you:

- Get immediate answers to questions
- Validate vendor claims
- Exchange tips for getting the best deals with vendor

Visit PeerSpot: www.peerspot.com

PeerSpot

244 5th Avenue, Suite R-230 • New York, NY 10001

reports@peerspot.com

+1 646.328.1944