Updated 05/17/2024

Sign up

ClickHouse training

3 days (21 hours)

Presentation

ClickHouse is a column-oriented open-source software package for online analytical processing (OLAP). ClickHouse was developed in C++ by the Russian company Yandex. The system is optimized for high-performance generation of analytical reports from real-time data. It can process over a billion rows and tens of gigabytes of data per second on a single server.

ClickHouse executes more queries at the same time, tests more hypotheses, combines your data in lots of new ways, analyzes your data from new angles, discovers new dimensions. It uses all available hardware to its full potential, with peak query processing performance of over 2 TB per second after decompression.

For customers wishing to build and operate ClickHouse technology, we offer training courses specifically designed for database developers, database administrators (DBAs) and all data specialists.

Your training can be tailored specifically to your needs, covering the most important and relevant topics for your use case.

As with all our training courses, this one will introduce you to the latest stable LTS version and its new features (ClickHouse 23.8 LTS at the time of writing).

Objectives

- Manage large volumes of data
- Generate real-time data analysis reports
- Managing your data warehouse (EDD) / Data-Warehouse (DWH)

Target audience

Developers, DBAs, Architects, Systems administrators, DevOps

Prerequisites

Basic knowledge of a Unix system

ClickHouse training program

Introducing ClickHouse

- OLAP
- Column-oriented database (DBMS)
- Competitors
- Compromise
- History
- Documentation & Ecosystem

Installation and configuration

- Single-server installation on Ubuntu
- Launch / Configuration / Debug
- Client and driver interfaces (Python, JavaScript, PHP, Go, Rust, C#)

Applications with ClickHouse

- Tables and Engines
- Types and schematics
- ClickHouse SQL
- Dictionaries
- View and Materialized Views
- Distribution and Replication

Operating ClickHouse in cluster mode

- Application diagram (Zookeeper / CHKeeper)
- ClickHouse cluster on bare-metal
- A ClickHouse cluster with Kubernetes
- Monitoring
- Common errors and remedies

- Update

Further information

- Using arrays
- Approximate calculation and sampling
- SQL query optimization
- MergeTree engine extensions (Collapse / Replace/ Aggregate)
- Using Materialized views

Companies concerned

This course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.