

MyDuckSQL

Long-term goal

Learning how to embed duckdb into MySQL, as a MySQL storage engine

Short-term goal

Faking it: duckdb but connectable via MySQL wire protocol

Can be accessed via:

- BI tools, such as: metabase, redash, etc
- SQL client: dbeaver, mysql workbench
- Currently only can be accessed without auth via: `mysql -h 127.0.0.1`

Context / who is this for

Imagine you are: a single person or a very small team (1-3 person) of data scientists that own and maintain a self-hosted small data warehouse for your own team. Currently, you use MySQL 5, but there are many inconveniences, mainly you cannot do window functions and your query is slow for analytics purposes.

If you use:

- MySQL 5: no window functions, slow query, table lock
- MySQL 8 / PostgreSQL: have window functions. Query is still slow because of manual index. table lock

If you think a **single-node analytics database is enough for most cases**, you might be interested in MyDuckSQL.

Closest alternative

- Singlestore (can single-node, uses MySQL wire protocol, can delete rows unlike some analytics DB that cannot, self-hostable, not open source)
- Aliyun AnalyticDB (multi-node, uses MySQL wire protocol, can delete rows unlike some analytics DB that cannot, cloud only, not open source)

Who is this not for

If you are a big team of data analysts, you better go with a proper analytics database, such as BigQuery, Athena, Clickhouse, etc

How to run

`python server.py`

Disclaimer

- **Definitely causes data loss. Experimental use only. Do not use in production**
- MyDuckSQL is only a fun learning project

Credits

- Bastian Blank (waldiTM) <https://github.com/waldiTM/python-mysqlproto>