



Proof Trading Reduces TCO 70% and Creates High-Performance, High-Transparency Trading Platform with SingleStore on AWS

1B/sec

Reading 1 billion bits/second

35M in 80sec

Has executed 35 million trades in 80 seconds

-70% TCO

70% total cost of ownership savings

500K/sec

Saving \$1,000+/mo. DB cost + massive opportunity cost

719M

records ingested in 24 minutes

“SingleStore offers fast analytics while having rich transactional support, and it’s SQL-compliant to the extreme, which enables us to load our trades and quotes at more than 500,000 records/second.”



Marcio Moreno, CTO, Proof Trading

Proof Trading, founded in 2019, is an innovative institutional equities execution platform that boldly declares “Finance is a bloated industry” front-and-center on its website. The team behind this execution-only broker-dealer is working to close the gap between market principles and the actual trading experiences of long-term investors. Proof is building algorithms to navigate the market on behalf of its institutional investor clients and to provide unprecedented transparency with products that offer both high performance and high accountability.

Challenges/Goals

With decades of finance industry experience, Proof Trading’s team set its sights on addressing some of the biggest issues it encountered in the US equities market.

Finance is a bloated industry. Its signature piece, the US equities market, is full of contradictions. It is home to new microwave communications towers and low-latency software – implemented through old manual processes that have persisted despite decades of technological revolution,” explained Proof Trading. “It is a place that on the one hand is designed to facilitate public price discovery through fair competition, but on the other, the pricing of its own services is opaque and carries inherent conflicts,” such as a lack of transparency around processes that may be most convenient for the business as opposed to most beneficial for the client.

Proof Trading’s vision was to build from the ground up a flexible, state-of-the-art trading platform in the cloud. Proof Trading chose to build its own solution to avoid the constraints associated with licensing or partnering with an existing or off-the-shelf platform. The company wants to be a broker-dealer that is unconflicted and different from most of the others. Its company code is to be transparent with our customers and the market, and as part of that it publishes everything, from how its algorithms work to the tools it uses, eventually it plans to open-source everything.

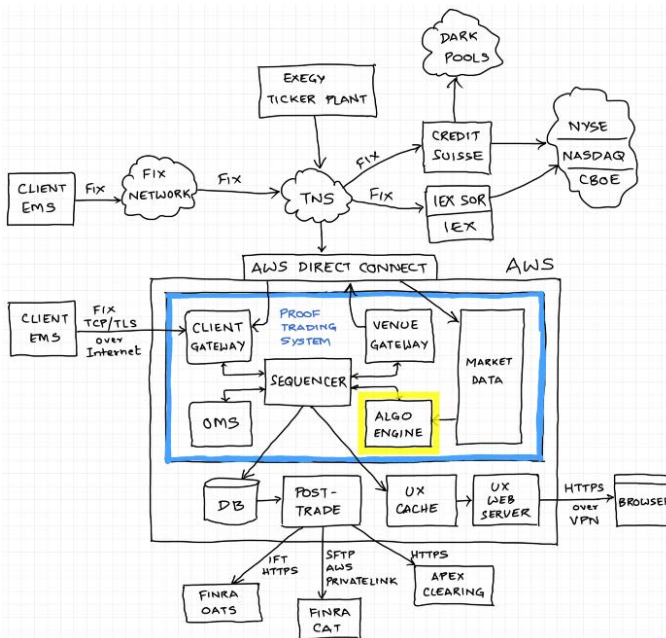


IMAGE / Proof Trading: How the algos work. [Why it publishes its algo design](#)



Technology Requirements

Proof stores every single trade it makes in the market, every field, for every market data product its clients are in. It then runs regressions on top of it to keep the open data model. Proof needed its data platform of choice to support this data-intensive approach, and with data growing by 50-100GB of uncompressed data and a billion records per day the selected database would have to deliver a robust set of features, including:

- Real-time data ingestion capable of handling 250,000 to one million events per second for multiple hours a day
- Fast in-memory aggregations with window functions, upserts, and other cross-sectional aggregations
- In-memory tables for fast analytics on working sets of data, and optimized on-disk tables to enhance performance
- A mature product that offers tools to set up, manage, and monitor deployments, including multi-node clusters
- Cost-effectiveness for a cost-conscious startup looking for something free or relatively inexpensive
- Good documentation and community support

Why SingleStore?

The Proof Trading team had extensive experience with kdb+, a column-based, in-memory, relational time series database commonly used in high-frequency trading. However, the high cost, steep learning curve, and lack of operational tools led Proof team to evaluate a number of other databases before ultimately choosing SingleStoreDB:

MySQL and PostgreSQL, the Traditional RDBMS Starting Points

Proof said it included MySQL only as a starting point, mostly to confirm that a traditional relational database management system (RDBMS) was not the right solution. MySQL did not support the type of advanced analytics or performance Proof Trading was looking for. When the team tested MySQL (the InnoDB engine), it could not keep up with batched inserts of 250,000 per second, and the insert speed slowed noticeably after a few million records were added to the table. PostgreSQL performed better than MySQL on the load tests, but performance was not good enough compared to the other contenders.

TimescaleDB: A Promising PostgreSQL Extension but Poor Performance

TimescaleDB, a PostgreSQL extension, was designed to have better insert rates and predictable query times with large data sizes compared to PostgreSQL, and it included a number of time series features. While it seemed promising at first, TimescaleDB kept running out of memory when Proof Trading tried to ingest 35 million trade records.

ClickHouse: Impressive Performance but Limited Query Language

ClickHouse checked most of the boxes Proof Trading was looking for and delivered impressive numbers in the team's performance testing, but a few dealbreakers that took it out of the running:

- The team would have to issue queries using a SQL-like query language with significant restrictions
- User-defined functions and Stored Procedures were not supported
- Only one SELECT statement could be issued per request.

Why SingleStore? (con't)

DolphinDB: Right Features, Wrong Cost

DolphinDB caught Proof Trading's attention since it was positioned as an alternative to kdb+. The team's testing experience was positive, but it ran into these drawbacks:

- ▽ Cost: DolphinDB was cheaper than kdb+ but still too expensive for the startup's budget.
- ▽ Non-standard language: the learning curve was less steep than kdb+ but still significant.
- ▽ Performance: the team encountered several crashes and out-of-memory errors.

EMR Cluster with Apache Spark: Experimenting with Flat Files Over Databases

At that point, Proof Trading considered working directly with flat files, testing an Elastic Map Reduce (EMR) cluster with Apache Spark on Amazon Web Services (AWS), however, it was too fragile and unstructured to use as a primary database.

SingleStoreDB: Database Found

Proof Trading had first discovered SingleStoreDB in 2019, when it was known as MemSQL, and found that it came the closest to meeting all of the team's requirements. SingleStoreDB is the distributed SQL database built to power data-intensive applications. With a unified data engine for transactional and analytical workloads, SingleStoreDB powers fast, real-time analytics and applications.

Proof started with a free commercial license for a starter cluster, feeling that would be enough to get Proof Trading to the point where it could consider paid products, and it soon did just that. "SingleStoreDB was the most balanced solution we found in our investigations. It is mature, easy to use, free (for now), fast, efficient, and interoperable with all the standard tools that we can hope for," said the Proof Trading team.

Other SingleStoreDB features that appealed to Proof Trading included:

- Unified data store offering fast analytics with rich transactional support.
- SQL compliance that comes with MySQL wire-compatibility which opens up access to integration with MySQL's vast ecosystem of tools.
- Extensive Procedural Language support with both user-defined functions (UDFs) and Stored Procedures.
- In-memory rowstore tables and on-disk columnstore tables with sharding, sorting, and compression.
- Parallelized, ultra-fast ingestion using multiple cores for load processes.
- Ease of use.

"SingleStoreDB was the most balanced solution we found in our investigations. It is mature, easy to use, fast, efficient, and interoperable with all the standard tools we could hope for."

— Proof Trading

Solution

SingleStore was among the easiest tools to work with in terms of deployment, administration, monitoring, cluster setup, and even loading and querying of data, Proof Trading said. The company started with [SingleStoreDB Self-Managed](#) and later moved to [SingleStoreDB Cloud](#), a fully managed solution, in 2023, running on AWS. Although the team can certainly manage a database, it wants to focus its efforts not on maintaining data infrastructure but on delivering strategic business value to the company, and SingleStoreDB Cloud frees it to do that.

The team uses SingleStoreDB for the trading system itself. It ingests data directly into SingleStore, running analytics using Python and queries via APIs. It uses RedisDB for user experience (UX) and OneTick for historical research.



Marcio Moreno

Chief Technology Officer, Proof Trading

Among the SingleStoreDB Cloud features Proof Trading leverages are:

- **Workspaces:** a Workspace is an independent deployment of compute resources that can be used to run a workload, and each Workspace can attach and detach databases as needed for flexibility and scalability. When the team creates its ingest pipeline, it creates a number of Workspaces with whatever queries it wants. If some queries are getting slow, the team can simply upsize the Workspace, attach the database, and query from there.
- **Unlimited (Bottomless) Storage:** Unlimited Storage is the separation of storage and compute for SingleStore. The system persists data files such as snapshots, logs, and blobs to S3 (or comparable blob storage, or even NFS) asynchronously. Blobs not being queried can be deleted from SingleStore node's local disk, allowing the cluster to hold more data than available disk, thus making the cluster's storage bottomless.

The rest of the platform's technology stack includes:

- Cloud Platform: AWS
- Operating System: Amazon Linux 2
- Programming languages: Java SE, Python, Typescript
- DevOps: Jira, Bitbucket, Confluence/Notion
- Clustering: None for the trading system, AWS Elastic Container Service (ECS) for UX
- Operational Tools: AWS CloudFormation, Ansible, Jenkins, DataDog
- UX Technologies: Node.js, React/Redux, AG Grid

While the Proof Trading System runs inside a private network in the AWS Cloud, much of the equities ecosystem it needs to interface with is still deployed in proprietary data centers in New Jersey. The team selected AWS Region US-East-1 in Northern Virginia and 3.5 milliseconds of network latency away from its major trading centers.

Outcomes

Reducing TCO by 70% with Suspend and Resume Functionality

Proof Trading also uses SingleStoreDB Cloud's suspend and resume functionality to suspend clusters while not in use, which reduces the total cost of ownership (TCO) significantly.

Proof's servers are only needed during business hours – they only run for 10 hours a day and only on weekdays, so it only pays for 50 hours in a week, or less than 30% of what the cost would be if it ran the servers 24x7. The ability to stop and start its databases with SingleStore saves it a lot of money.

The entire trading system, which includes the databases and EC2 servers, costs Proof Trading less than \$10,000 per month.

Ingesting 500,000 Records and Reading 1B Bits Per Second

"SingleStore offers fast analytics while having rich transactional support and is SQL-compliant to the extreme, which enables us to load our trades and quotes at more than 500,000 records per second," said the team. Proof Trading has loaded 35 million trades in 80 seconds and 719 million quotes in 24 minutes at this ingestion rate with SingleStore. The team also found it can read data from SingleStoreDB at over one billion bits per second (Gbps) with an external tool and the Java Database Connectivity (JDBC) interface.

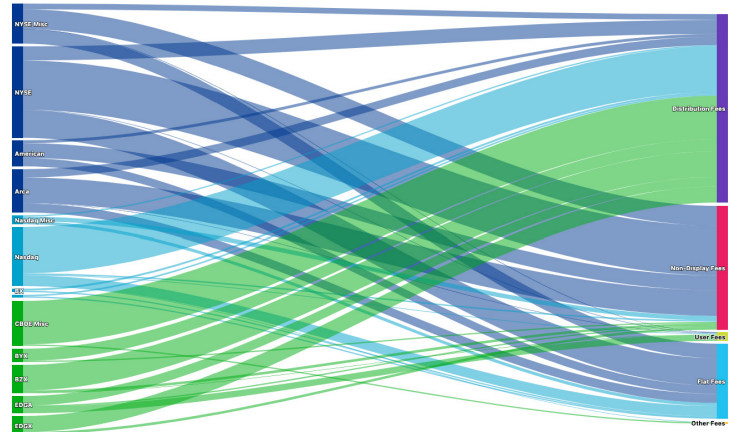


IMAGE / Proof Trading: a small slice of what is available on its [Market Data Fee Calculator](#)

Managing Massive Data Volumes and Reducing Cost with Compression

SingleStoreDB's compression capabilities mean that the loaded data only takes up as much space as Proof Trading's g-zipped source files. With daily data growth rates averaging 50-100GB, having efficient compression lowers the storage burden and cost.



SingleStore is helping companies compete and win across every vertical

[Learn More >](#)

We hope you find this SingleStore impact story a worthwhile read, and before you leave this story we want to highly recommend a collection of others: the [Proof Trading home page](#) with links to highly informative documents titled rejecting the black box, how to build a broker, and many more.