

# Michael Thoun

## Lead Database Administrator/Data Engineer

Senior-level data professional with a broad range of expertise in database administration, data engineering, and data architecture. Creative problem solver that specializes in architecting and building scalable core data infrastructure such as data lakes and pipelines. Over a decade of experience in financial markets. Very strong background in 24/7 support, DataOps, DevOps/SRE, and systems engineering.

**Specialization:** Database Administration, Data Engineering, Data Architecture, Data Strategy and Governance

**Databases:** SQL Server, RDS, Aurora, PostgreSQL, MySQL, Oracle, Redshift, Vertica, MongoDB, Cassandra

**Distributed Computing:** EMR, Athena, Hadoop, Hive, Presto, Spark, Treasure Data

**Real-time/Streaming Data:** Kinesis, Kafka, Storm

**Data Pipelines:** Airflow, Luigi

**ETL:** Glue, SSIS, Talend, Pentaho

**Analysis:** Jupyter Notebook, Sagemaker, Pandas, NumPy, scikit-learn, matplotlib, R Studio

**Visualization:** Tableau, QuickSight, Redash, Superset, Kibana, Grafana

**Infrastructure-As-Code:** Terraform, SaltStack

**CI/CD:** Jenkins, CircleCI, Bamboo

**Other:** Git, Docker, S3, Glacier, EC2, ECS, Lambda, Cloudwatch, Salesforce

**Languages:** SQL, Python, Java, C#, C++, OCaml

**Financial Markets:** High Frequency Trading, Market Microstructure, Quantitative Strategies, Volatility, Crypto

## Experience

### Feb 2017 - present **Lead Database Administrator/Data Engineer/Architect**

#### **Outcome Health**

- ▶ Owner of Outcome Health's database systems and data infrastructure.
- ▶ Architected Outcome Health's data lake, a centralized source-of-truth for all of the company's data, storing both structured and unstructured data at scale, agnostic to source and format.
- ▶ Designed and built Outcome Health's data pipeline, leveraging technologies such as Airflow, Kinesis, and Hadoop for batch and streaming ingestion/processing of data in the data lake from upstream sources such as internal services, traditional database systems (PostgreSQL, MySQL, Redshift), APIs (Salesforce, MixPanel, Mojo, Broadsign), flat files, and more.
- ▶ Designed a holistic enterprise data model, joining data elements from all aspects of Outcome Health's business into a single cohesive view.
- ▶ Worked with developers to design application data models and troubleshoot inefficient database interactions.
- ▶ Devised and executed a database consolidation plan that reduced monthly AWS RDS cost by approximately 33%.
- ▶ Implemented standardized naming and security schemes for all database systems.
- ▶ Incorporated database query performance monitoring via Solarwinds Database Performance Analyzer and RDS Performance Insights.
- ▶ Brought Redash to Outcome Health, democratizing data and allowing users to easily create self-service BI and dashboards.
- ▶ Leveraged Jupyter notebook for data analysis to discover and correct data quality and synchronization issues.
- ▶ Created a process for generating Salesforce cases for the Network Health team to address network and deployment issues.
- ▶ Responsible for 24/7 support of databases, data pipelines/ETLs, and other core data infrastructure.

## Profile

### Address

437 W Division St #601  
Chicago, IL 60610

### Phone

630-299-9481

### E-mail

mikethoun@gmail.com

### Website

mikethoun.com

## Social

[linkedin.com/in/mikethoun](https://www.linkedin.com/in/mikethoun)

[twitter.com/mikethoun](https://twitter.com/mikethoun)

[github.com/mikethoun](https://github.com/mikethoun)

## Education

2005

### DePaul University

Information Technology

## Certifications

### Microsoft MCSE Data Platform

License E506-9378

### Microsoft MCSA SQL Server

License E416-7597

### Red Hat Certified Engineer

Certification ID 1111-095-220

### Red Hat Certified Systems Admin

Certification ID 1111-095-220

### Apache Cassandra Administrator

Datastax License V2.1-0128

## Courses

### Machine Learning

Stanford University on Coursera

License 3A8KUNVBVB3S

### R Programming

Johns Hopkins University on Coursera

License ZGGYTSMZAXT2

Jan 2009 -  
Feb 2017

## **Lead Database Administrator**

### ***Spot Trading LLC***

- ▶ Principal database administrator responsible for architecting, deploying, and managing multiple SQL and NoSQL database systems including Microsoft SQL Server, MongoDB, Apache Cassandra, HP Vertica, Amazon Redshift, and Amazon RDS in an environment totaling over 75 terabytes in size.
- ▶ Planned and executed numerous environment-wide database upgrades and migrations including the upgrade of the SQL Server environment from SQL 2005/2008 to SQL 2012, the migration of SQL Server instances from Failover Clustering to Always On Availability Groups, the migration of on-premises SQL Server and MongoDB infrastructure into AWS, the routine application of patches and service packs, and more.
- ▶ Primary point person at the firm for troubleshooting and resolving critical data quality issues, possessing both technical and domain knowledge.
- ▶ Implemented database backup, maintenance, and disaster recovery solutions.
- ▶ Reviewed and deployed thousands of database changes into the production environment with no occurrences of major outages or system downtime that would have put the firm's trading activities and positions at risk.
- ▶ Worked one-on-one with developers on a daily basis to improve the performance and efficiency of their applications' database interactions.
- ▶ Monitored database systems for slow and inefficient queries, identifying bottlenecks and resolving issues via performance tuning and query optimization by leveraging tools such as SQL Profiler, DPA, and Database Designer.
- ▶ Used Python, Powershell, and SSIS expertise to automate routine database management tasks and ETLs.
- ▶ Devised, implemented, and enforced a set of database standards across the environment, including a process for the submission, review, approval, and deployment of all production database changes.
- ▶ Gained in-depth knowledge of various database engine internals.
- ▶ Supported all databases systems on a 24/7 basis, including on-call and working overnight to resolve major issues to ensure that all systems were up by market open so the firm was not at risk.
- ▶ Initially hired as a systems engineer, responsible for building and managing CentOS servers in high performance, low latency environment.

Jun 2005 -  
Dec 2008

## **Infrastructure Engineer II**

### ***CME Group***

- ▶ Supported physical layer changes within CME Group's state-of-the-art high density data centers as an engineer on the Data Center and Computing Infrastructure team.
- ▶ Developed and managed DCCI's internal ASP.NET web application used for planning changes and tracking physical assets and network cable paths within CME data centers.
- ▶ Assisted with the design and execution of data center build outs in compliance with TIA-942 standards.
- ▶ Worked comprehensively with structured cabling systems, specifically, the troubleshooting of cable plant issues with category 5, 6, and single-mode/multi-mode fiber optic cable runs.
- ▶ Worked closely with manufacturers such as Panduit, Belden, and Systimax to research new physical infrastructure technologies and reported findings to managers on the feasibility of their use within CME data centers.
- ▶ Gained end-to-end project management experience including vendor and contractor management.
- ▶ Represented the DCCI team at weekly capacity planning meetings, reporting data center capacity and growth projections to upper management.
- ▶ Acquired extensive knowledge of different enterprise hardware platforms from Cisco, HP, Sun, Dell, IBM, EMC, Copan, Netscout, and others.
- ▶ Experience supporting mission-critical, time-sensitive changes and emergency fixes 24/7 as a member of the data center on-call rotation.