



Trino: The Data Synthesizer

Monica Miller

Developer Advocate @  Starburst

 @moni4489

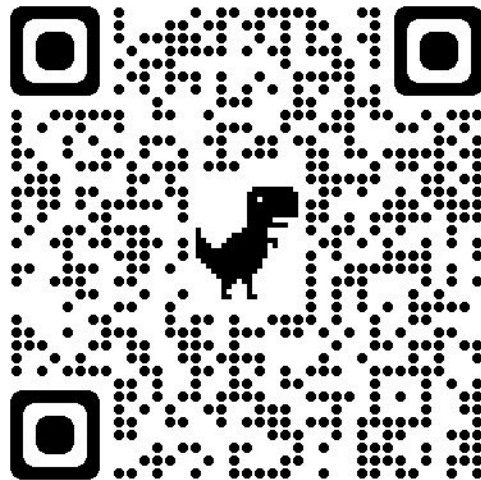
Monica Miller (she/her)

Data Engineer -> Developer Advocate 🥑

Trino Enthusiast

Passionate about Couch Driven Development

Connect with me @ link.monimiller.com



The Synthesizer

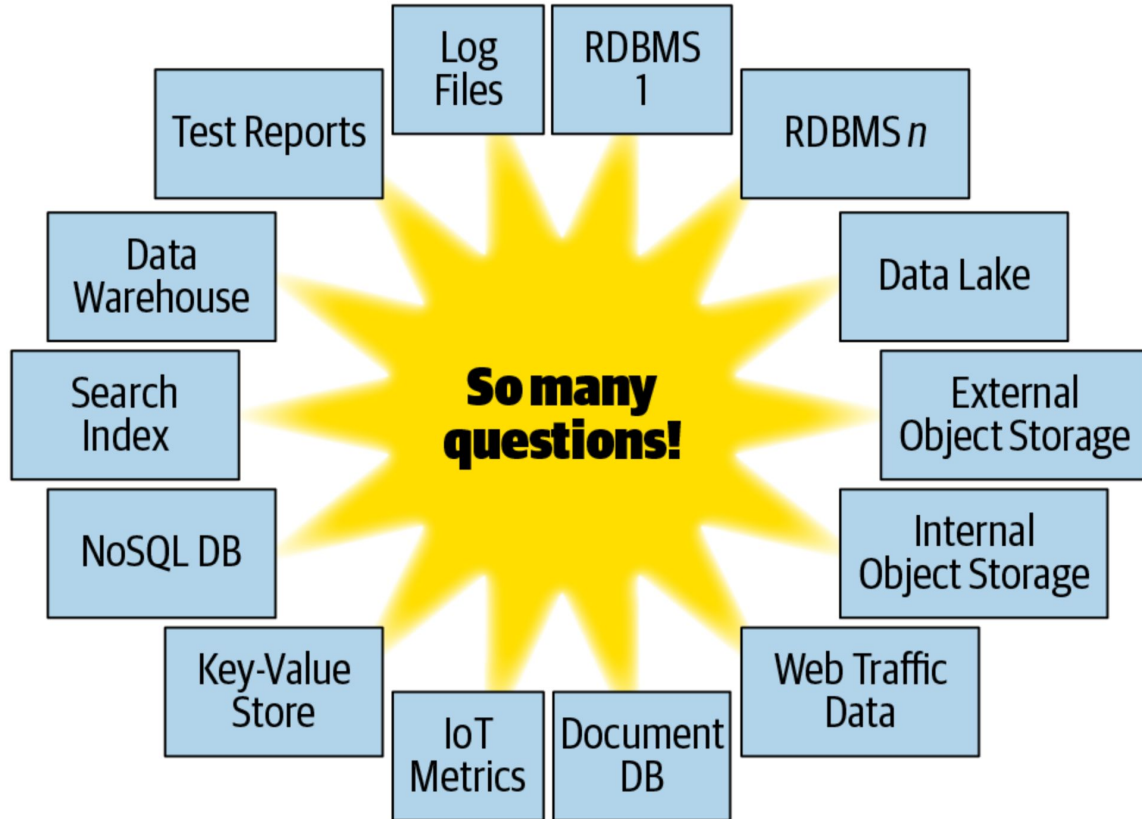
The Synthesizer is a musical instrument that creates, modifies, or combines different tones to produce a variety of sounds.



Data as a Resource

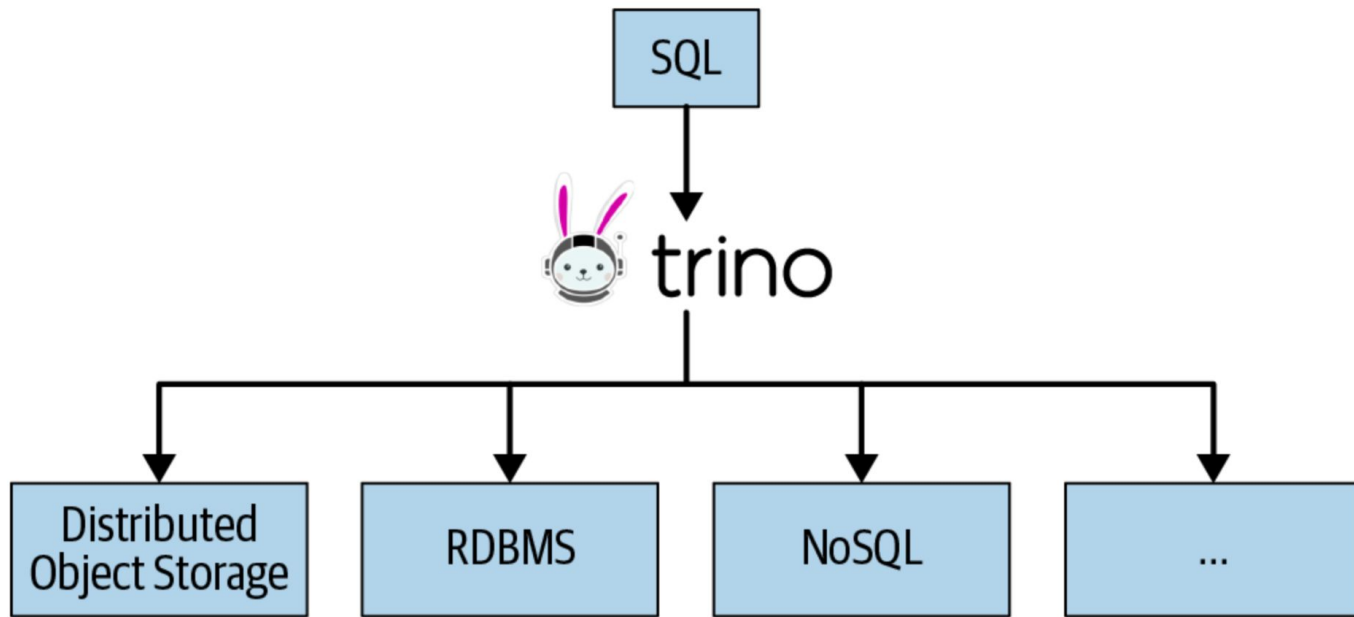


Silos Are Among Us



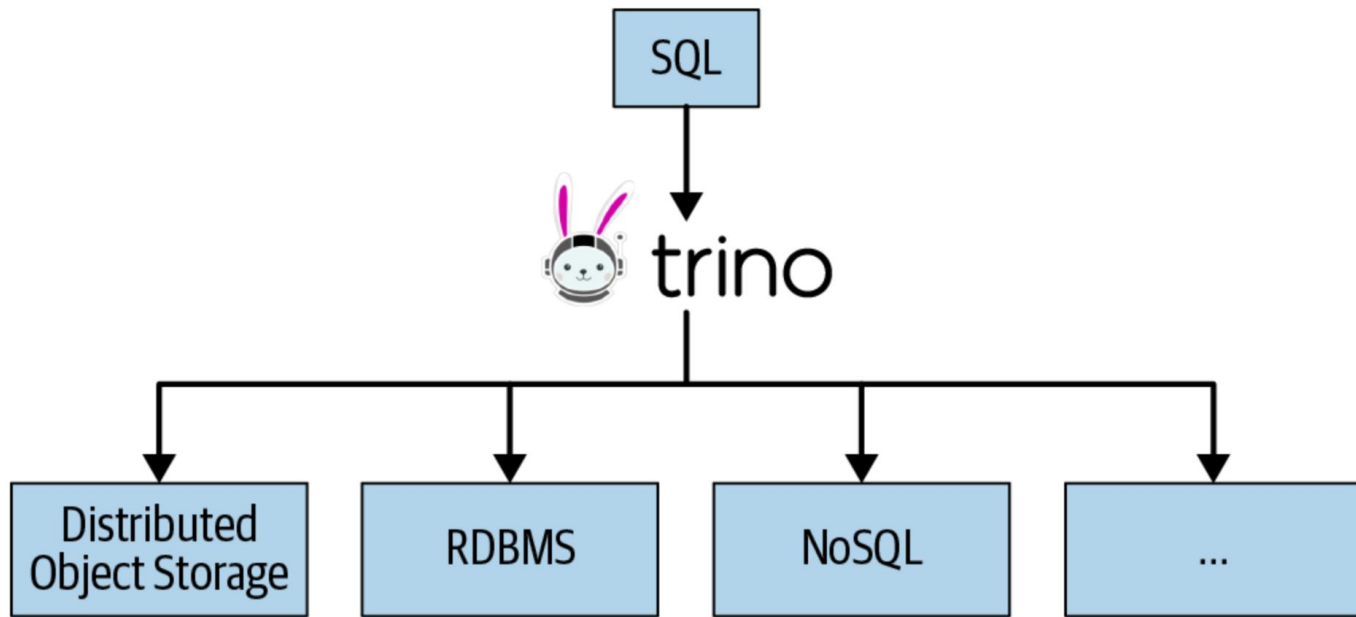
Trino: The Data Synthesizer

Trino is a fast distributed SQL query engine designed to query large data sets distributed over one or more heterogeneous data sources.



Trino: The Data Synthesizer

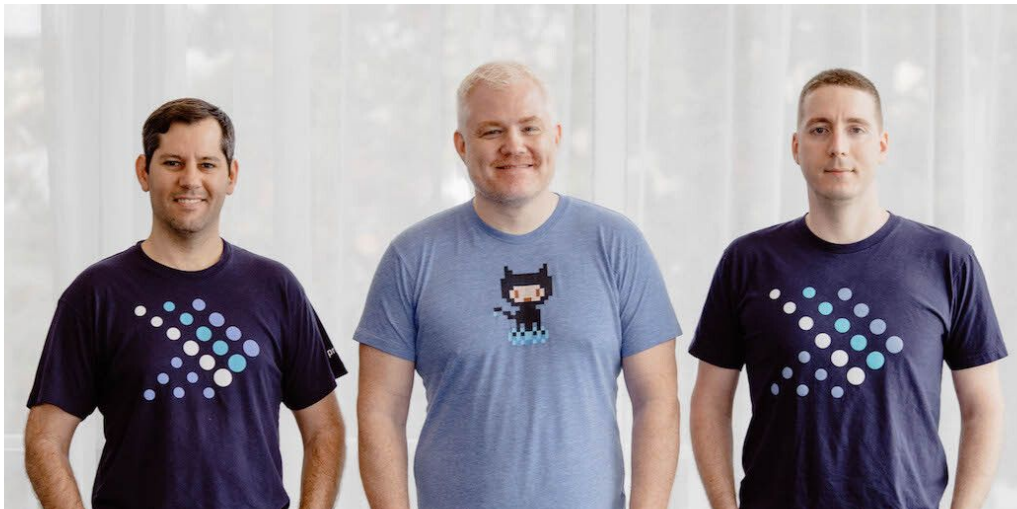
Trino is a fast distributed SQL query engine designed to query large data sets distributed over one or more heterogeneous data sources.



Presto! Your queries are fast!



Martin Traverso, Dain Sundstrom, and David Phillips created Presto in 2012. It aimed to solve for the slow queries of Hive at Facebook and eventually many more.



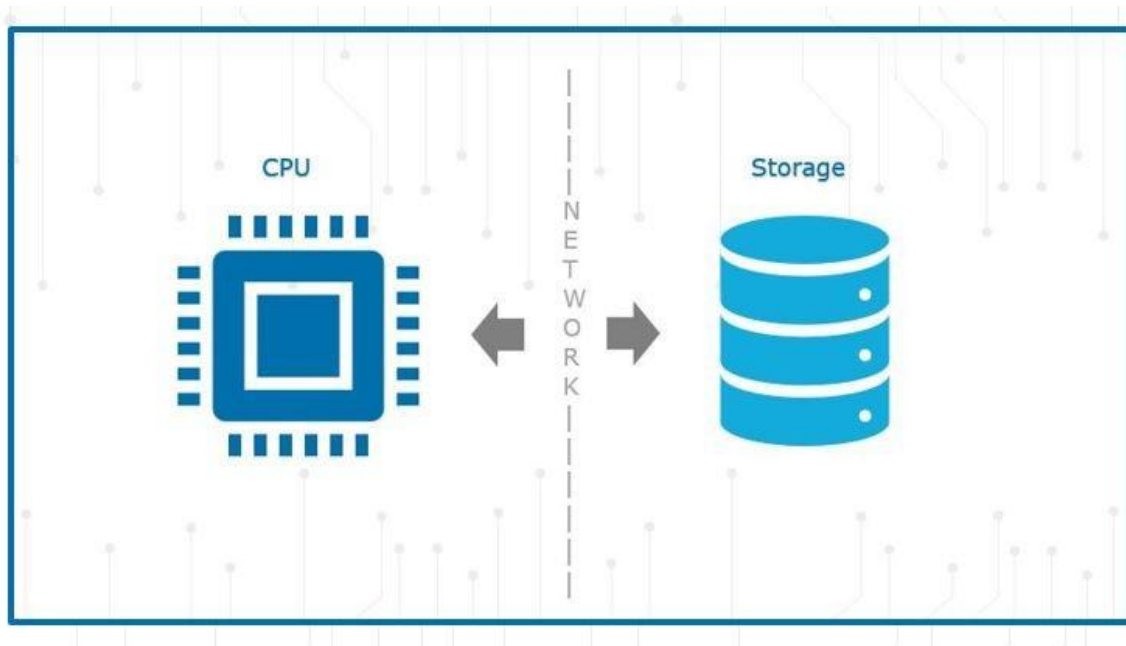
Presto >>> Trino

Development Philosophies both then and now:

- Open source with neutral governance model
- Fast, interactive analytics
- Correct results
- Standards based (ANSI SQL, JDBC, etc..)

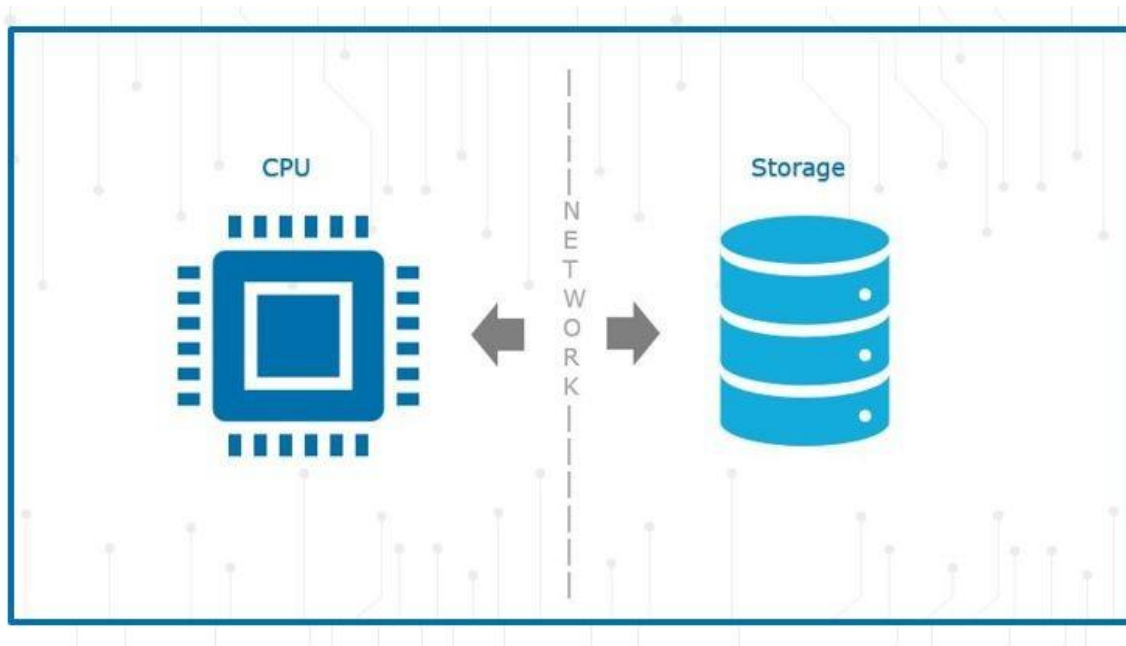
Trino Differentiators

Separates Storage and Compute



Trino Differentiators

Separates Storage and Compute



Trino Differentiators

Based on ANSI SQL

Data Analysts, Data Engineers, and Data Scientists hold the power



Faster time to insight

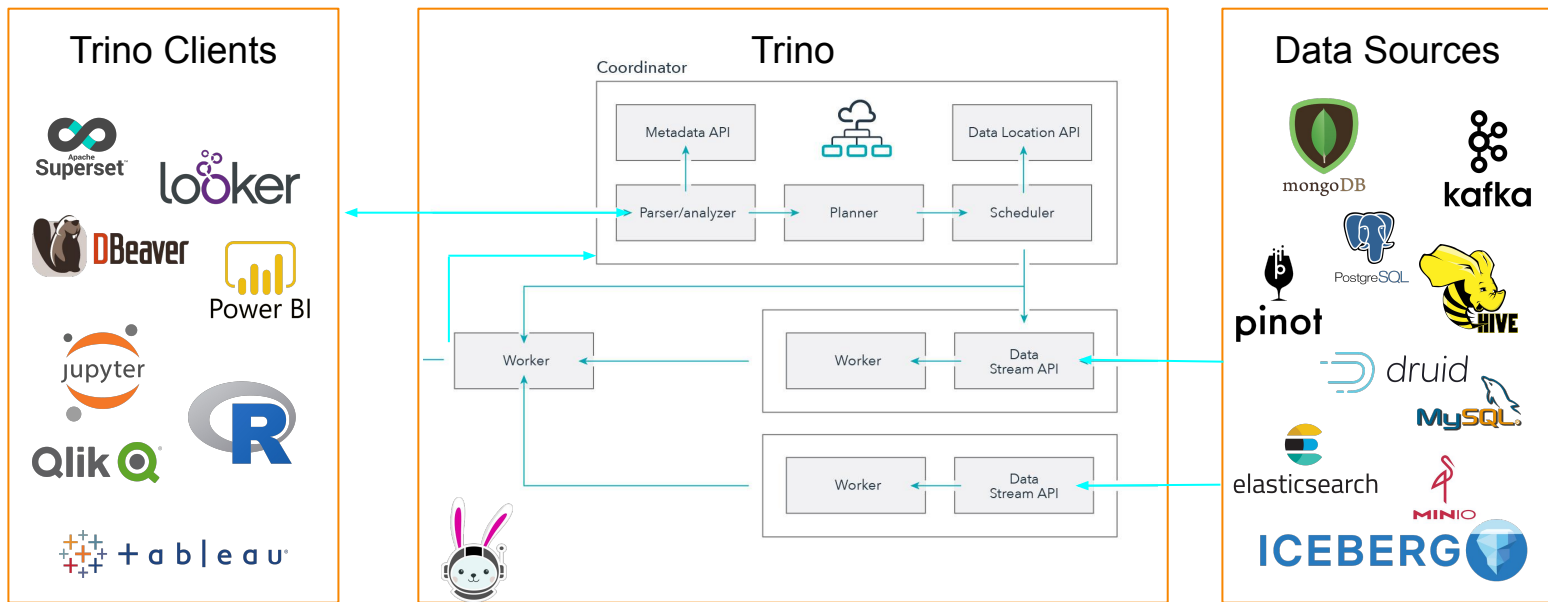
Trino Differentiators

Cost Based Optimizations

- Pushdown Optimization
- Distributed Query Planning
- Join Ordering

Common Use Cases

Interactive Analytics



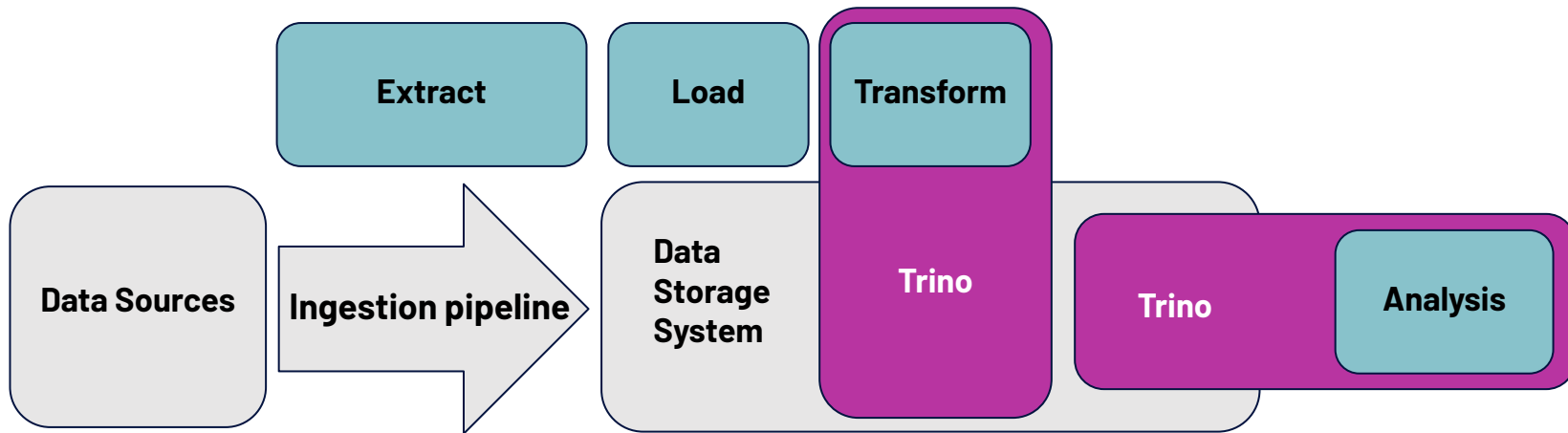
Common Use Cases

Data Lake Analytics

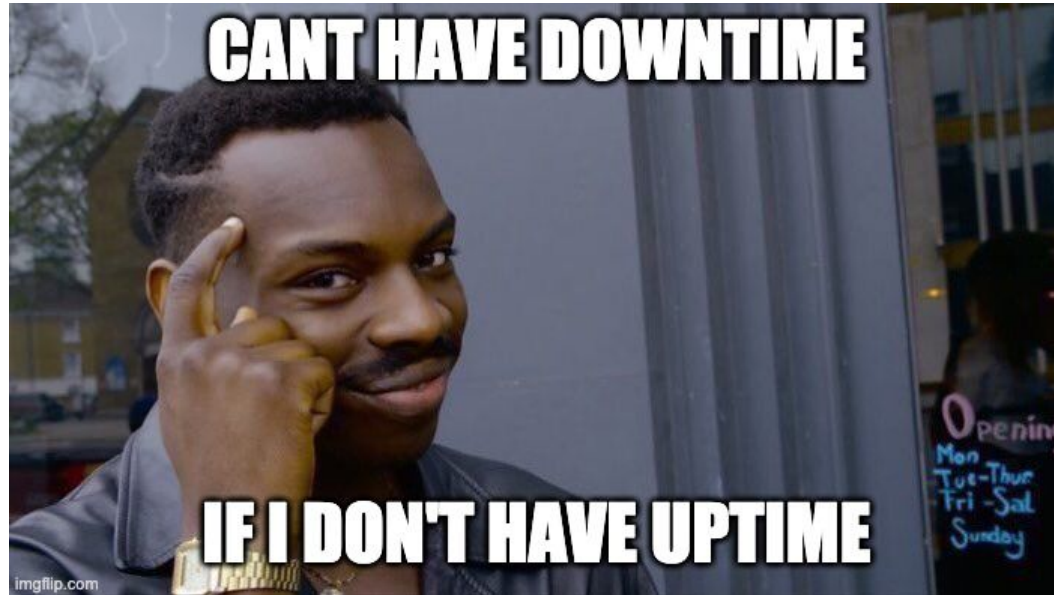


Batch Analytics

Batch ETL/ELT processing across disparate systems



Great! Why do I care?



Trino can:

- Automate processes that historically have been manual and slow
- Operate and evolve applications quickly and reliably
- Help engineers independently accomplish tasks that normally require help from other teams

Which results in:

- Less strain on the infrastructure
- Proper utilization of resources
- Data Democratization

Great! Why do I care? Part - 2

We are looking for cool developers who want to join the community and learn more!



Trino on Kubernetes

Kubernetes helps Trino

- Scale up/scale down
- Support worker failover

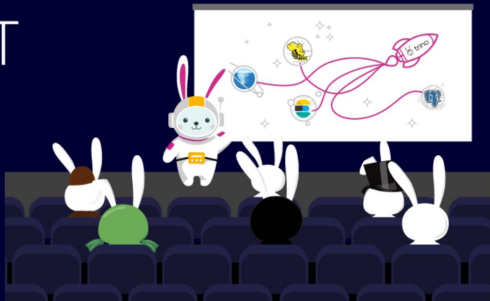
Our helm chart is available here: <https://github.com/trinodb/charts>

Community

Hosted by  Starburst

 trino SUMMIT

NOVEMBER 10TH, 2022



- Trino Slack
 - <https://trino.io/slack.html>
 - Trino Community Broadcast
 - <https://trino.io/broadcast/>
 - Trino Summit (Virtual & Free) on November 10th
 - <https://www.starburst.io/info/trinosummit>
 - Contribute to the project
 - <https://trino.io/development/>
-

Connect with me
@
link.monimiller.com



Thank you

Please give us a ★ on
github.com/trinodb/trino