What's new in Trino this summer

Martin Traverso, Starburst

Hosted by Starburst

By the numbers

- 14 releases

- 2822 commits
  - On pace for the most active year in Trino development ever

- 138 unique individual contributors
  - Including 75 first-time contributors in 2024

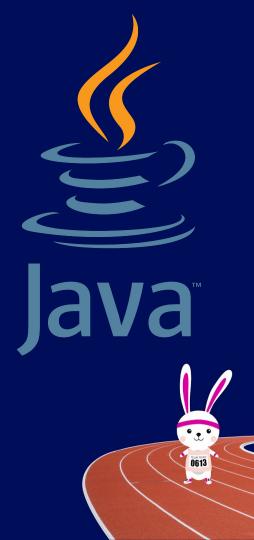**Trino is everywhere, and it's growing faster than ever...**

Trino Fest 2024

Hosted by ✳ Starburst

Some more community numbers...

- **Developers from 50+ companies/organizations**
  - **Including Google, Apple, Amazon, Facebook, and Microsoft**
- **~50 monthly active contributors**
- **12,500 Slack members**
- **5,000 followers on LinkedIn**
- **Nearly 10,000 stars on GitHub**
  - **If you haven't given us a star yet, go do that!**

Trino Fest 2024

Hosted by Starburst

Using latest Java

- Moved from 17, 21, and now to 22 as requirement

- Enables use of new language features

- Cleaner code for maintenance and improvements

- Ongoing GC and performance updates

- FFI for external library use

- Vector (SIMD) APIs

News from [Project Hummingbird](#)

More performance improvement shipped:

- Parquet reader with vectorized decoding

- Columnar filter evaluation

- Flat hash tables for GROUP BY

- Aircompressor with FFI for Zstd, LZ4, and Snappy (upcoming)

Dynamic catalogs progress

- **Lifecycle and shutdown improvements**

- **Revamped file system support**

  - **Directly handles file systems in S3, GCS, and Azure**

  - **Removes dependencies on Hadoop and related memory issues**

- **Catalog storage SPI**

- **Support for ALTER CATALOG and SHOW CATALOG (in progress)**

- **Secrets management (upcoming)**

Connector improvements

- **New OpenSearch and Snowflake connectors**

- **Iceberg: support for Snowflake catalog, views with REST catalog**

- **Hive: Glue V2**

- **OpenLineage support**

Dedicated talks about Delta Lake and Hudi later today…

Trino Gateway updates

- Access control, authentication, and other security improvements

- New web UI

- New router modules

- New Docker container and Helm chart

- And a new logo ...

And as always, tons of small improvements...

- Improve performance by not generating redundant predicates.
- Improve performance of queries with selective joins on partition columns.
- Improve query performance for queries that don't use table statistics.
- Improve performance of queries with filters on `ROW` columns stored in Parquet files.
- Improve performance of the `arrays_overlap` function.
- Improve latency for queries with file system caching enabled.
- Improve latency for queries on tables with checkpoints.
- Improve performance when reading JSON files.
- Improve storage table cleanup when creating a materialized view fails.
- Improve planning time for queries with a large number of joins.
- Improve performance of reading table comments.
- Improve performance of queries with `ORDER BY ... LIMIT` clause, or `avg`, `count(distinct)`, `stddev`, or `stddev_pop` aggregation functions when the computation can be pushed down to the underlying database.
- Improve performance of writing data files.

- Improve performance of queries selecting only catalog, schema, and name from the `system.metadata.materialized_views` table.
- Improve performance of queries with `timestamp(n)` values.
- Improve performance of `INSERT` statements into partitioned tables when the `retry_policy` configuration property is set to `TASK`.
- Improve performance of `ORDER BY` queries with `LIMIT` on large data sets.
- Improve performance of queries containing the `rank` or `row_number` window functions.
- Improve performance and memory usage when equality delete files are used.
- Improve performance of correlated queries with `EXISTS`.
- Improve performance when listing schemas.
- Improve performance when pushing down equality predicates on textual types.
- Improve performance of reading from Parquet files.
- Improve performance of writing to Parquet files.

Trino Fest 2024

Hosted by Starburst

Roadmap

- HTTP/2

- Client protocol

- Plan fragment caching

- Incremental materialized views

- Branching and tagging

- Exasol, Vertica, and LanceDB connectors

- Workload management

Wishlist (need contributors)

- **Temporary tables**

- **PIVOT**

- **SQL:2023 JSON type**

- **Memory connector**

- **Server discovery**

Community events

- Trino Contributor Congregation tomorrow!

- [Trino Community Broadcast](#) up to 60 episodes, with four more already planned

- Monthly Trino Contributor Call

- Tons of small, vendor-sponsored Trino events and meetups

Trino Fest 2024

Hosted by ⚹ Starburst

Questions?

Thank you to our sponsors!

Starburst

ALLUXIO

star⚡tree

upsolver

ONEHOUSE

aws

Trino Fest 2024

Hosted by Starburst