What's new in Presto

Budapest Data Forum 2020

Wojciech Biela, Sr Director, Engineering, Starburst
Piotr Findeisen, Presto Committer, Co-founder, Starburst

Wojciech Biela

Sr Director, Engineering @ Starburst

/in/biela

@ilfrin

wojciech.biela@starburstdata.com

Piotr Findeisen

Presto Committer, Co-founder @ Starburst

in/piotrfindeisen/

@findepi

piotr@starburstdata.com

Agenda

- Introduction
  - What is Presto
  - What is Starburst
- Presto in more detail
- What’s new in Presto land
- Q&A

What is Presto?

High performance MPP SQL engine

- Interactive ANSI SQL queries
- Proven scalability
- High concurrency

Community-driven open source project

Separation of compute & storage

- Scale storage & compute independently
- SQL-on-anything
- Federated queries

Deploy Anywhere

- Kubernetes
- Cloud
- On premises

Presto Users

presto

Facebook: 10,000+ of nodes, 1000s of users
Uber 2,000+ nodes, 160K+ queries daily

LinkedIn: 500+ nodes, 200K+ queries daily
Lyft: 400+ nodes, 100K+ queries daily

What is Starburst?

- Starburst Enterprise Presto - distribution
- Open core model
- Azure, AWS, GCP, On Premises, & Kubernetes
- Biggest open source Presto contributor

- Based in Boston, MA
- EMEA offices in Warsaw and London

Starburst

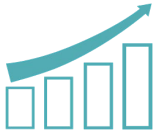
What is Starburst?

- Starburst Enterprise Presto - distribution
- Open core model
- Azure, AWS, GCP, On Premises, & Kubernetes
- Biggest open source Presto contributor

- Based in Boston, MA
- EMEA offices in Warsaw and London

Starburst

About Starburst



600% Growth YoY

BUSINESS INSIDER

Named Open Source Startup to Watch 2020

100+ Enterprise Customers

NPS Score

Our Platform

ANSI SQL MPP Query Engine

On-Prem, or Cloud

High Concurrency

Massive Scale

Rapid Time to Insights

Low Cost of Ownership

Enterprise Grade Security

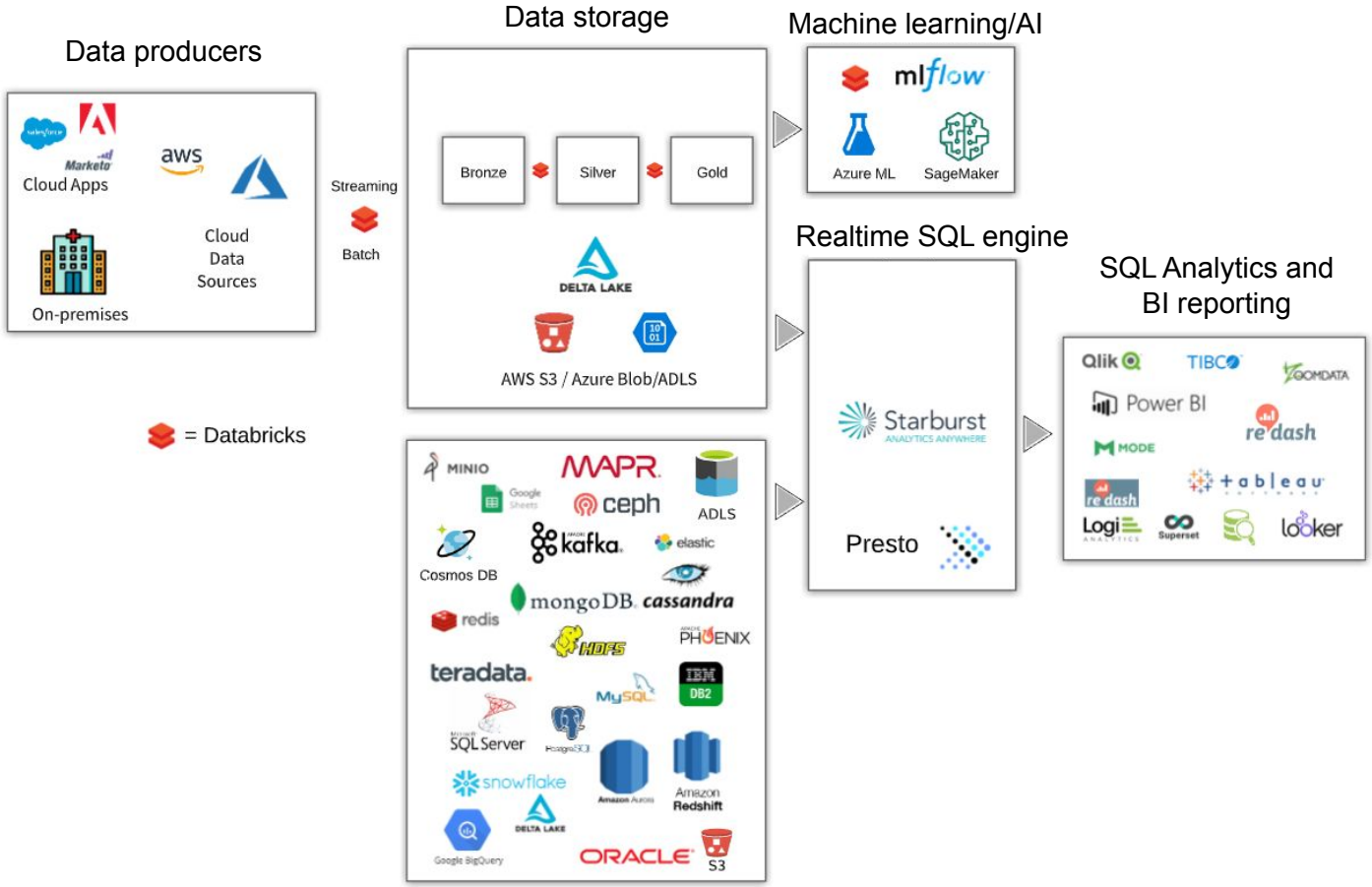
24x7 Expert Support

Starburst

Data sources

presto

Data Ingestion and Analytics Ecosystem

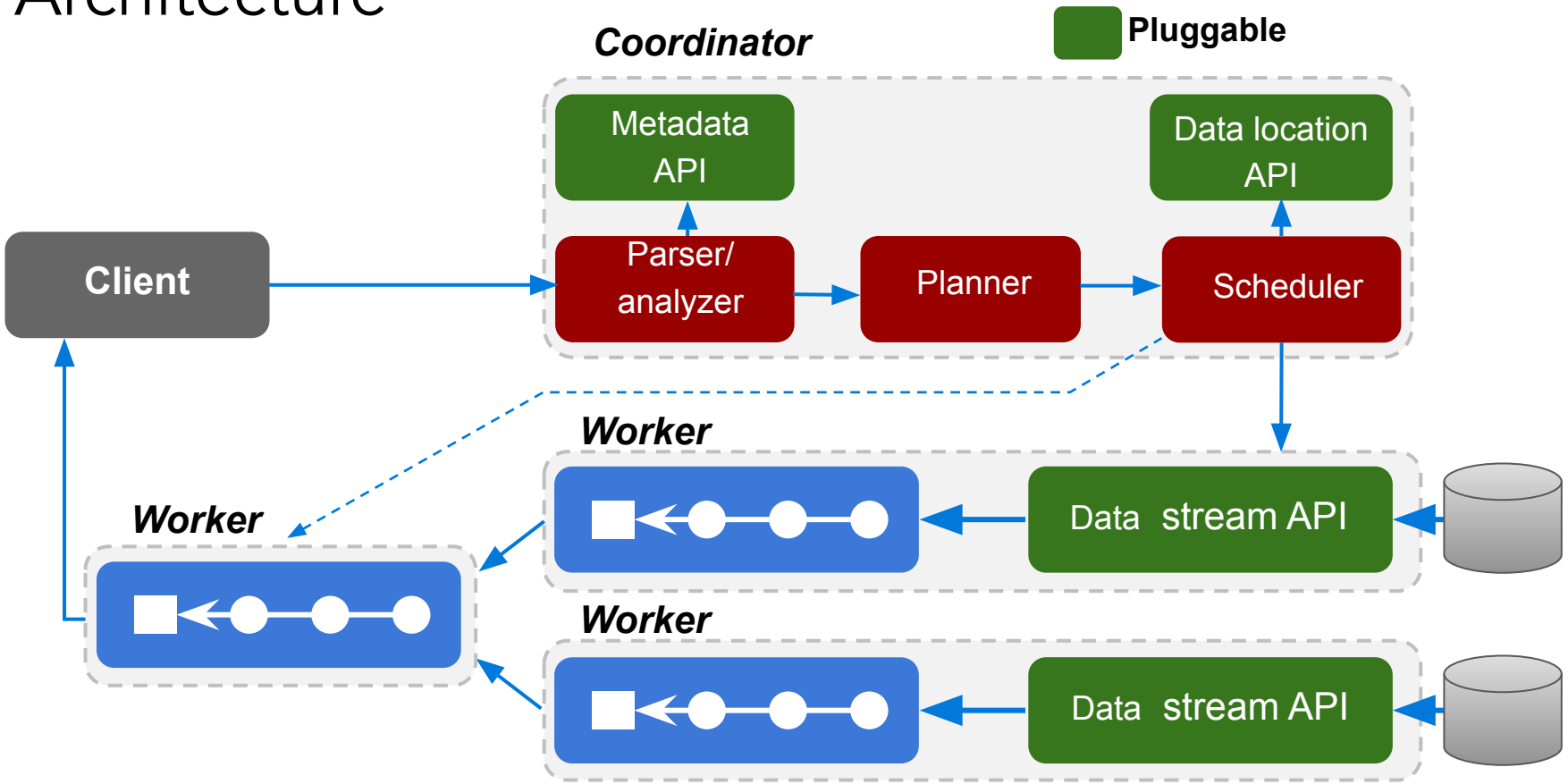


Presto in Production

- **Facebook**: 10,000+ of nodes, HDFS (ORC, RCFile), sharded MySQL, 1000s of users
- **Uber**: 2,000+ nodes (clusters on prem.) with 160K+ queries daily over HDFS (Parquet/ORC)
- **Twitter**: 2,000+ nodes (several clusters on premises and GCP), 20K+ queries daily (Parquet)
- **LinkedIn**: 500+ nodes, 200K+ queries daily over HDFS (ORC), and ~1000 users
- **Lyft**: 400+ nodes in AWS, 100K+ queries daily, 20+ PBs in S3 (Parquet)
- **Netflix**: 300+ nodes in AWS, 100+ PB in S3 (Parquet)
- **Yahoo! Japan**: 200+ nodes for HDFS (ORC), and ObjectStore
- **FINRA**: 120+ nodes in AWS, 4PB in S3 (ORC), 200+ users

Presto in detail

Architecture



Built for Performance

Query Execution Engine:

- MPP-style **pipelined** in-memory execution
- **Vectorized** data processing
- Runtime query **bytecode generation**
- Memory efficient **data structures**
- Multi-threaded **multi-core execution**
- Optimized readers for **columnar formats** (ORC and Parquet)
- Predicate and column projection **pushdown**
- **Cost-Based Optimizer**

What’s new in Presto land

Baseline - Starburst Presto 312 (Jun'19)

- Open source Presto features:
  - Hive parallel connector (HDFS, S3, ADLS, GCS)
    - ORC, Parquet, RCFile, AVRO, SeqFile, JSON, CSV, Text
  - SQL Server connector
  - PostgreSQL connector
  - MySQL connector
  - Cassandra connector (parallel)
  - MongoDB connector
  - ...

- Starburst:
  - BigQuery connector
  - Oracle connector
  - Generic JDBC
  - Ranger & Sentry integration

Starburst Presto 323 (Nov'19)

- Open source Presto features:
  - Add support for `OFFSET, FETCH FIRST` syntax (incl. `WITH TIES`)
  - Allow connectors to provide view definitions.
  - Improve performance of queries using UNNEST
  - Support reading from and writing to Hadoop encryption zones (Hadoop KMS)
  - Collect column statistics on write by default
  - Export JMX statistics for various JDBC and connector operations
  - Support for simple column and row field dereference pushdown into connectors
  - Improve performance of information schema tables
  - Elasticsearch connector (parallel), add support for nested fields, dynamic schema discovery
  - Hive connector
    - Improve performance of S3 object listing & GCS reading
    - Allow inserting into bucketed, unpartitioned and partitioned tables
    - Add support for instance credentials and custom credentials providers for the Glue metastore
    - Add support for Azure Data Lake (adl) file system
    - Add caching support for Glue metastore
  - Kinesis connector
  - MemSQL connector

- Starburst:
  - IBM DB2 connector
  - Snowflake parallel connector
  - Teradata parallel connector
  - Starburst Secrets
  - Azure Data Lake Storage (ADLS) Gen 2
  - Power BI DirectQuery connector

Starburst Presto 332 (Apr'20)

- Open source Presto features:
  - Support for late materialization to join operations
  - Allow analyzing a subset of table columns (rather than all columns)
  - Add Password File Authentication
  - Allow using multiple system access controls
  - Handle common disk failures during spill
  - Enable cost based join reordering and join type selection optimizations by default
  - Add support for large prepared statements
  - Add support for row filtering and column masking (SPI)
  - Hive connector
    - Support reading from insert-only transactional tables
    - Add support for Azure WASB, ADLS Gen1 (ADL) and ADLS Gen2 (ABFS) file systems
    - Add support for Hive full ACID tables
    - Add flexible S3 Security Mapping, allowing for separate credentials or IAM roles for specific users or buckets/paths
  - BigQuery parallel connector
  - Google Sheets connector

- Starburst:
  - S3 Caching (preview)
  - IBM Cloud Object Storage (COS) support (parallel)
  - Databricks Delta Lake native connector (parallel)
  - Oracle/Exadata Parallel Connector
  - Global Security
  - Okta integration
  - Event Logger (audit)
  - Amazon IAM role passthrough

Starburst Presto 338 (Jul'20)

- Open source Presto features:
  - Add pluggable Certificate Authenticator
  - Add framework to support aggregation pushdown into connectors
  - Optimized Parquet writer
  - Druid connector
  - Pinot connector
  - Prometheus Connector
  - Dynamic filtering framework

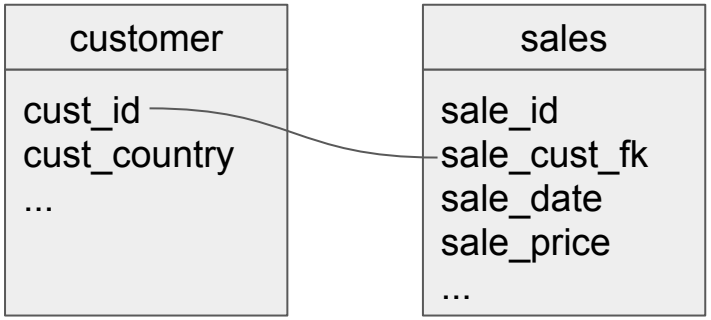
- Starburst:
  - Cloudera Data Platform (CDP) 7.1 connector
  - MapR (GA) connector
  - Caching for HDFS, ADLS, GCS, S3

Starburst Presto 34x (Oct'20)

- Open source Presto features:
  - Add support for writing Bloom filters in ORC files
  - MongoDB
    - Allow querying Azure Cosmos DB
  - Experimental support for recursive queries (WITH RECURSIVE)
  - Add support for variable-precision TIME/TIMESTAMP [WITH TIME ZONE] types
  - Iceberg Connector
  - Add Salesforce password authentication
  - TBD

- Starburst:
  - Greenplum parallel connector
  - SAP Hana connector
  - Dynamic filtering for RDBMS connectors
  - Aggregation pushdown for RDBMS connectors
  - TBD

Static partition pruning



sales table partitioned by sale_date

Static (planning time) partition pruning saves the day!

Only 31 partitions of sales table will be scanned

```
SELECT cust_country, sum(sale_price)
FROM customer JOIN sales ON cust_id = sale_cust_fk
WHERE sale_date >= date '2012-08-01' and sale_date <= date '2012-08-31'
GROUP BY cust_country
```

Dynamic partition pruning

sales table partitioned by sale_date_fk

Dynamic (runtime) partition pruning saves the day!

Only ~31 partitions of sales table will be scanned

```
SELECT cust_country, sum(sale_price)
FROM customer
    JOIN sales ON cust_id = sale_cust_fk
    JOIN date_dim ON sale_date_fk = date_id
WHERE date_date >= date '2012-08-01' and date_date <= date '2012-08-31'
GROUP BY cust_country
```

We’re hiring!

Q&A