Data & Drinks

Building Next-Gen Data Systems with Apache DataFusion

Andrew Lamb
InfluxData, Staff Engineer, Apache DataFusion PMC Chair

Mehmet Ozan Kabak
CEO & Co-founder of Synnada, Apache DataFusion PMC

Jan Kaul
Founder & CEO of Dashbook

Xomnia
Data & AI - Strategy to Execution

APACHE DATAFUSION™

Thursday, January 23rd 2025
17:30 at Xomnia

A leading **Data & AI Consultancy in the Netherlands**

From strategy…

Define data & AI strategies

Analytics Translators

…to execution

Build data platforms and data infrastructure

Data Engineers
Data Architects

Develop data & AI solutions

Analytics Engineers
Data Scientists
ML Engineers

vodafone Ziggo · ENEXIS · DELTA FIBER · HEMA · Ahold Delhaize · KLM · ARCADIS · allianders · Rabobank · X

Open positions

- Account managers
- Support Engineer
- Analytics Engineers
- Data Engineers
- ML Engineers
- Data Scientists
- Data Solutions Architect

Data & Drinks

Building Next-Gen Data Systems with Apache DataFusion

Andrew Lamb
InfluxData, Staff Engineer, Apache DataFusion PMC Chair

Mehmet Ozan Kabak
CEO & Co-founder of Synnada, Apache DataFusion PMC

Jan Kaul
Founder & CEO of Dashbook

Xomnia
Data & AI - Strategy to Execution

APACHE DATAFUSION™

Thursday, January 23rd 2025
17:30 at Xomnia

influxdata®

APACHE DATAFUSION™

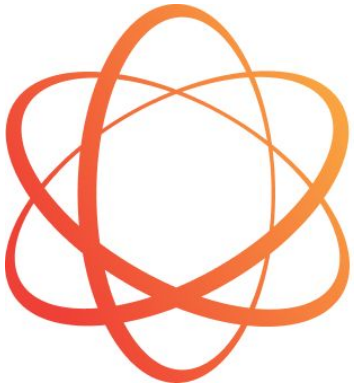
Intro to DataFusion: Technology, Community, and Not Quite Enough Time

Xomnia

Andrew Lamb
Staff Engineer, InfluxData

January 23, 2025, Data & Drinks: Building Next-Gen Data Systems with Apache DataFusion

DataFusion: Technology



APACHE DATAFUSION™

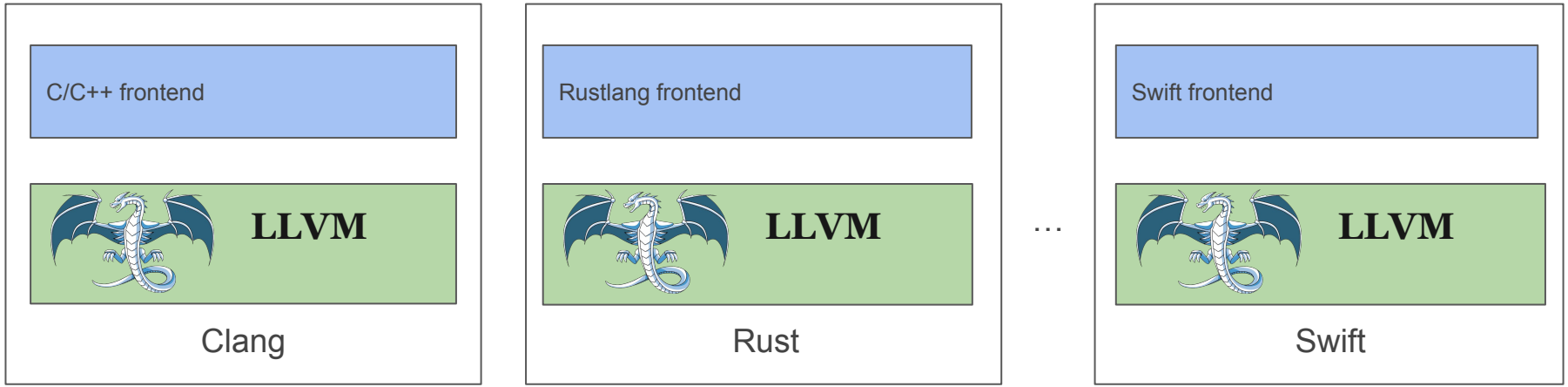
https://datafusion.apache.org/

Top Level Project, Apache Software Foundation

Apache 2.0 Licensed

 | © Copyright 2024, InfluxData

Analogy: DataFusion is LLVM for Databases

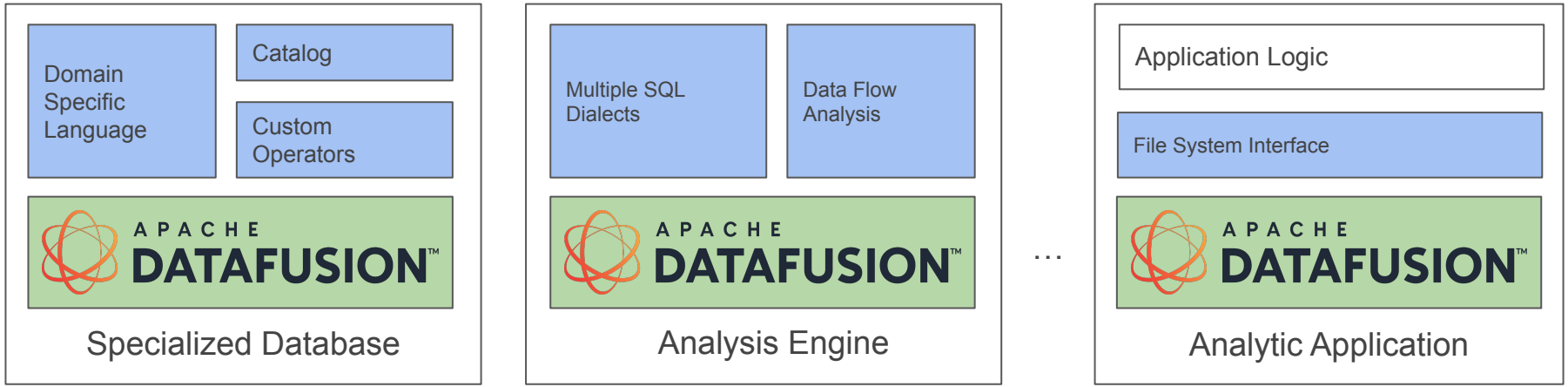


LLVM enabled innovation in programming languages:

- High quality reusable optimizer, code generator, debugger, lsp integration, etc.
- Focus on language design, ecosystem, libraries, etc

 | © Copyright 2024, InfluxData

Analogy: DataFusion is LLVM for Databases



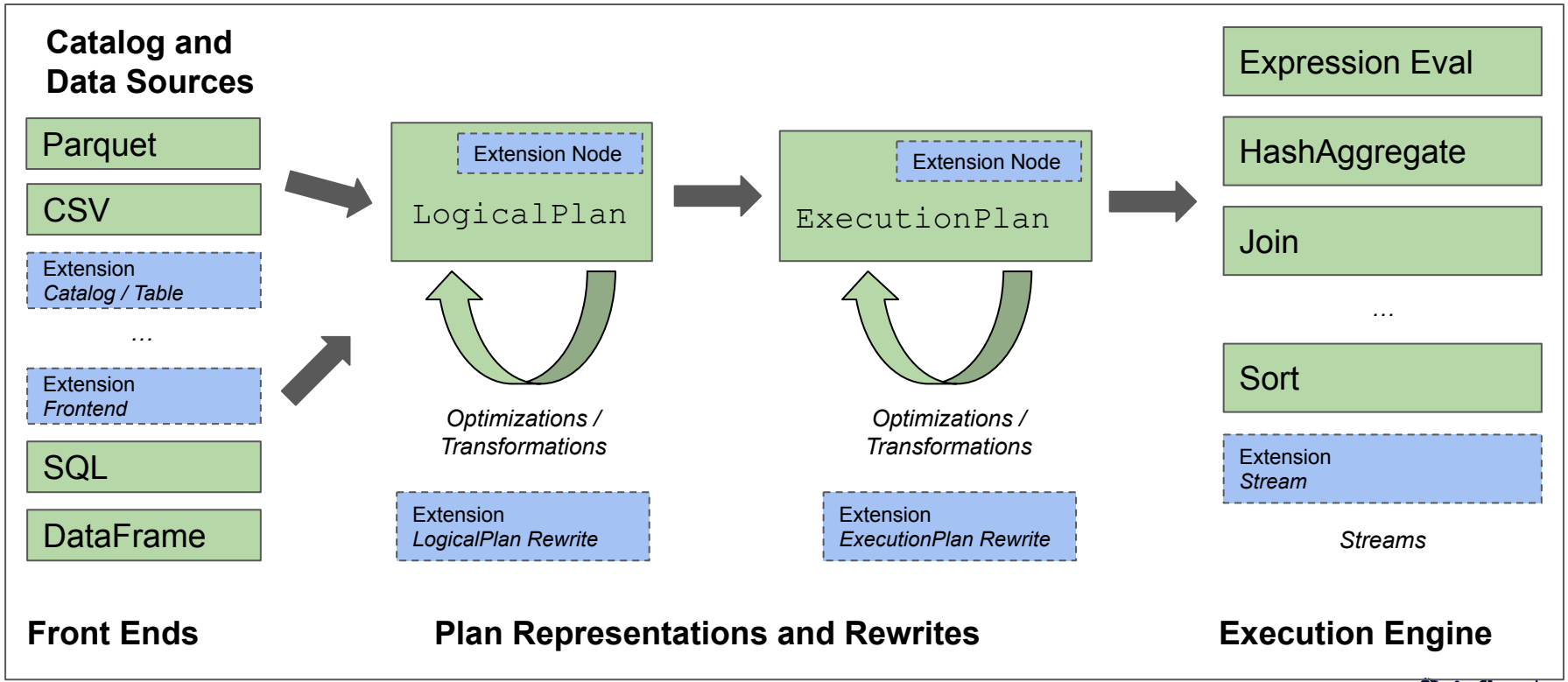
[DataFusion](#) enables innovation in data intensive systems

- High quality reusable SQL planner, optimizer, function library, vectorized operators, etc
- Focus on language design, data management, use case specific features

 | © Copyright 2024, InfluxData

influxdata®

Architecture



 | © Copyright 2024, InfluxData

influxdata®

Recognized Tech

Apache Arrow DataFusion: A Fast, Embeddable, Modular Analytic Query Engine

Andrew Lamb
InfluxData
Boston, MA, USA
alamb@apache.org

Yijie Shen
Space and Time
Irvine, CA, USA
yjshen@apache.org

Daniël Heres
Coralogix
Utrecht, The Netherlands
dheres@apache.org

Jayjeet Chakraborty
UC Santa Cruz
Santa Cruz, CA, USA
jayjeetc@ucsc.edu

Mehmet Ozan Kabak
Synnada
Austin, TX, USA
ozankabak@apache.org

Liang-Chi Hsieh
Apple
Seattle, WA, USA
viirya@apache.org

Chao Sun
Apple
Cupertino, CA, USA
sunchao@apache.org

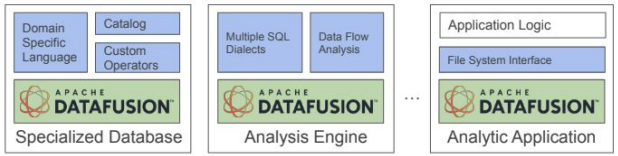


Figure 1: When building with DataFusion, system designers implement domain-specific features via extension APIs (blue), rather than re-implementing standard OLAP query engine technology (green).

ABSTRACT

Apache Arrow DataFusion[25] is a fast, embeddable, and extensible query engine written in Rust[76] that uses Apache Arrow[24] as its memory model. In this paper we describe the technologies on which it is built, and how it fits in long-term database implementation trends. We then enumerate its features, optimizations, architecture and extension APIs to illustrate the breadth of requirements of modern OLAP engines as well as the interfaces needed by systems built with them. Finally, we demonstrate open standards and extensible design do not preclude state-of-the-art performance using a series of experimental comparisons to DuckDB[66].

While the individual techniques used in DataFusion have been previously described many times, it differs from other industrial strength engines by providing competitive performance *and* an open architecture that can be customized using more than 10 major extension APIs. This flexibility has led to use in many commercial and open source databases, machine learning pipelines, and other data-intensive systems. We anticipate that the accessibility and versatility of DataFusion, along with its competitive performance, will further the proliferation of high-performance custom data infrastructures tailored to specific needs assembled from modular components[18, 61].

CCS CONCEPTS

• **Information systems** → **Database management system engines**; **Online analytical processing engines**; **DBMS engine architectures**; Relational database model; *Database query processing*; • **Software and its engineering** → *Abstraction, modeling and modularity*; *Software performance*; *Software usability*.

KEYWORDS

Database Systems, Modular Query Engines, Column Stores, OLAP, Vectorized Execution, Parallel Execution, API Design

ACM Reference Format:

Andrew Lamb, Yijie Shen, Daniël Heres, Jayjeet Chakraborty, Mehmet Ozan Kabak, Liang-Chi Hsieh, and Chao Sun. 2024. Apache Arrow DataFusion: A Fast, Embeddable, Modular Analytic Query Engine. In *Companion of the 2024 International Conference on Management of Data (SIGMOD-Companion '24), June 9–15, 2024, Santiago, AA, Chile.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3626246.3653368

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGMOD-Companion '24, June 9–15, 2024, Santiago, AA, Chile

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0422-2/24/06.
https://doi.org/10.1145/3626246.3653368

Apache Arrow DataFusion:

A Fast, Embeddable,

Modular Analytic Query

Engine

In [SIGMOD 2024]

Santiago
Chile

SIGMOD
PODS
2024

[We wrote this paper entirely in the open]

influxdata®

Top of the Line Performance

Speed (and underlying techniques) similar to other top engines such as ClickHouse + DuckDB

[Apache DataFusion is now the fastest single node engine for querying Apache Parquet files]

Apache DataFusion is now the fastest single node engine for querying Apache Parquet files

Posted on: Mon 18 November 2024 by Andrew Lamb, Staff Engineer at InfluxData

I am extremely excited to announce that Apache DataFusion is the fastest engine for querying Apache Parquet files in ClickBench. It is faster than DuckDB, chDB and Clickhouse using the same hardware. It also marks the first time a Rust-based engine holds the top spot, which has previously been held by traditional C/C++-based engines.

APACHE DATAFUSION™

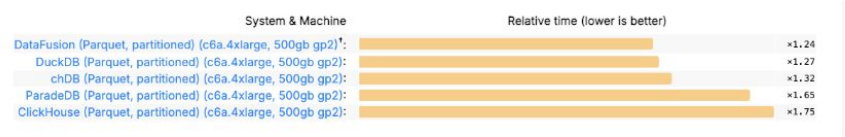


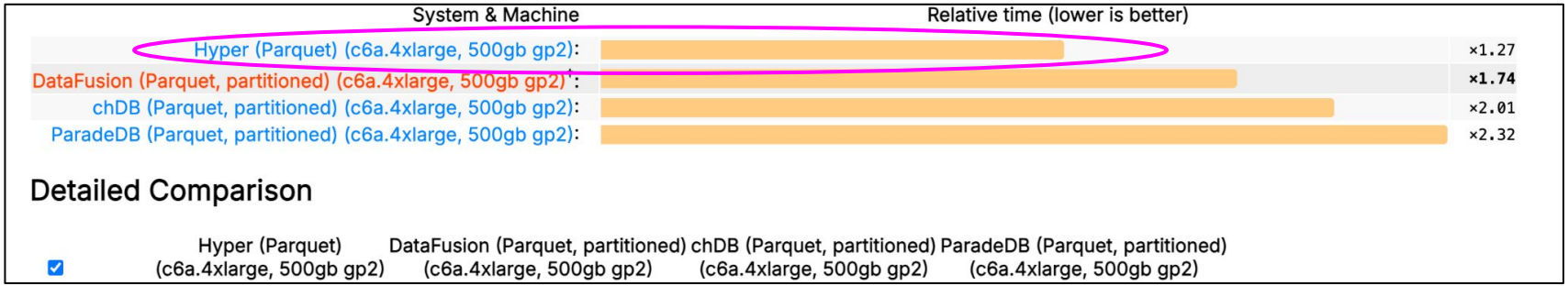
Figure 1: 2024-11-16 ClickBench Results for the 'hot'[^1] run against the partitioned 14 GB Parquet dataset (100 files, each ~140MB) on a `c6a.4xlarge` (16 CPU / 32 GB RAM) VM. Measurements are relative (`1.x`) to results using different hardware.

Best in class performance on Parquet is now available to anyone. DataFusion's open design lets you start quickly with a full featured Query Engine, including SQL, data formats, catalogs, and more, and then customize any behavior you need. I predict the continued emergence of new classes of data systems now that creators can focus the bulk of their innovation on areas such as query languages, system integrations, and data formats rather than trying to play catchup with core engine performance.

 | © Copyright 2024, InfluxData

Top of the Line Performance (is fleeting!)

ClickBench Results as of Jan 23, 2025

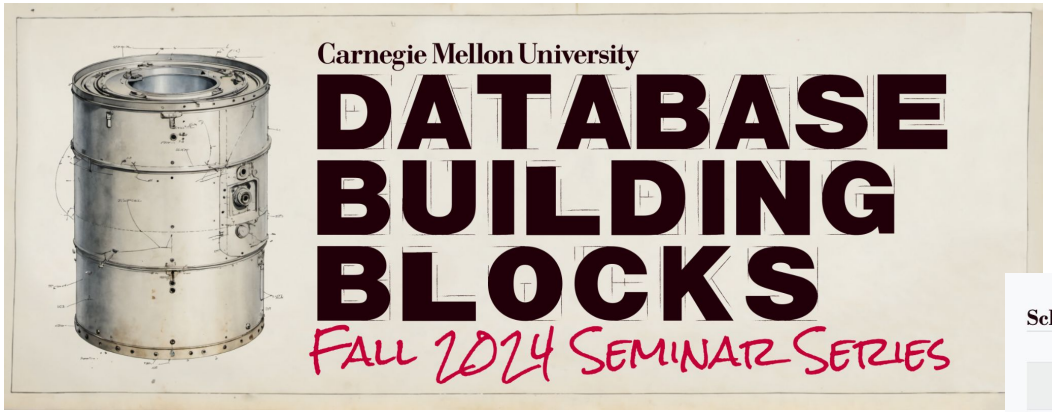


Looking for help rerunning ClickBench on 44 (and 45!)
[Update ClickBench benchmarks with DataFusion 44.0.0 #13983]

 | © Copyright 2024, InfluxData

influxdata®

Trending



Who's who of DataFusion users

https://db.cs.cmu.edu/seminar2024/

Schedule

	Date	Speaker	Talk Title	Video
DATA FUSION	Sep 23	Andrew Lamb	Apache Arrow DataFusion: A Fast, Embeddable, Modular Analytic Query Engine	
DATAFUSION COMET	Sep 30	Andy Grove	Accelerating Apache Spark workloads with Apache DataFusion Comet	
ParadeDB	Oct 7	Philippe Noël	ParadeDB – Postgres for Search and Analytics	
Spice AI	Oct 21	Luke Kim	Accelerating Data and AI with Spice.ai Open-Source Software	
Exon	Oct 28	Trent Hauck	Exon: A Built for Purpose Bioinformatics Database	
Synnada	Nov 4	Mehmet Ozan Kabak	Towards "Unified" Compute Engines: Opportunities and Challenges	
influxdb	Nov 11	Paul Dix	Building InfluxDB 3.0 with the FDAP Stack: Apache Flight, DataFusion, Arrow and Parquet	
GlareDB	Nov 18	Sean Smith	Biting the Bullet: Rebuilding GlareDB from the Ground Up	
APACHE OpenDAL	Dec 2	Xuanwo	Apache OpenDAL: One Layer, All Storage	
GreptimeDB	Dec 9	Ruihang Xia	Implement, Integrate and Extend a Query Engine	

 | © Copyright 2024, InfluxData

Architecture

Design Goals:

- Work “out of the box” (fast time to awesome)
- Customize *everything* via APIs
- Architecturally Boring 🥱 (“Industrial best practice”)

Results for Users

- Quickly start with a basic, high-performance engine
- Specialize to suit their needs and available engineering capacity
- Easy to try out new ideas (operators, rewrites, etc)

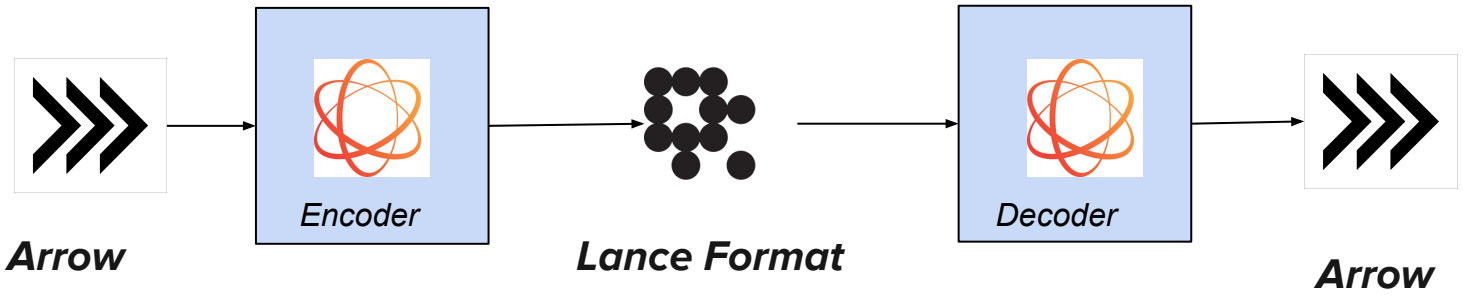
 | © Copyright 2024, InfluxData

Use Case: File Formats (Lance)

[Lance file format](#) uses DataFusion to implement pushdown filtering

Encoder uses DataFusion aggregators to calculate min/max (source)

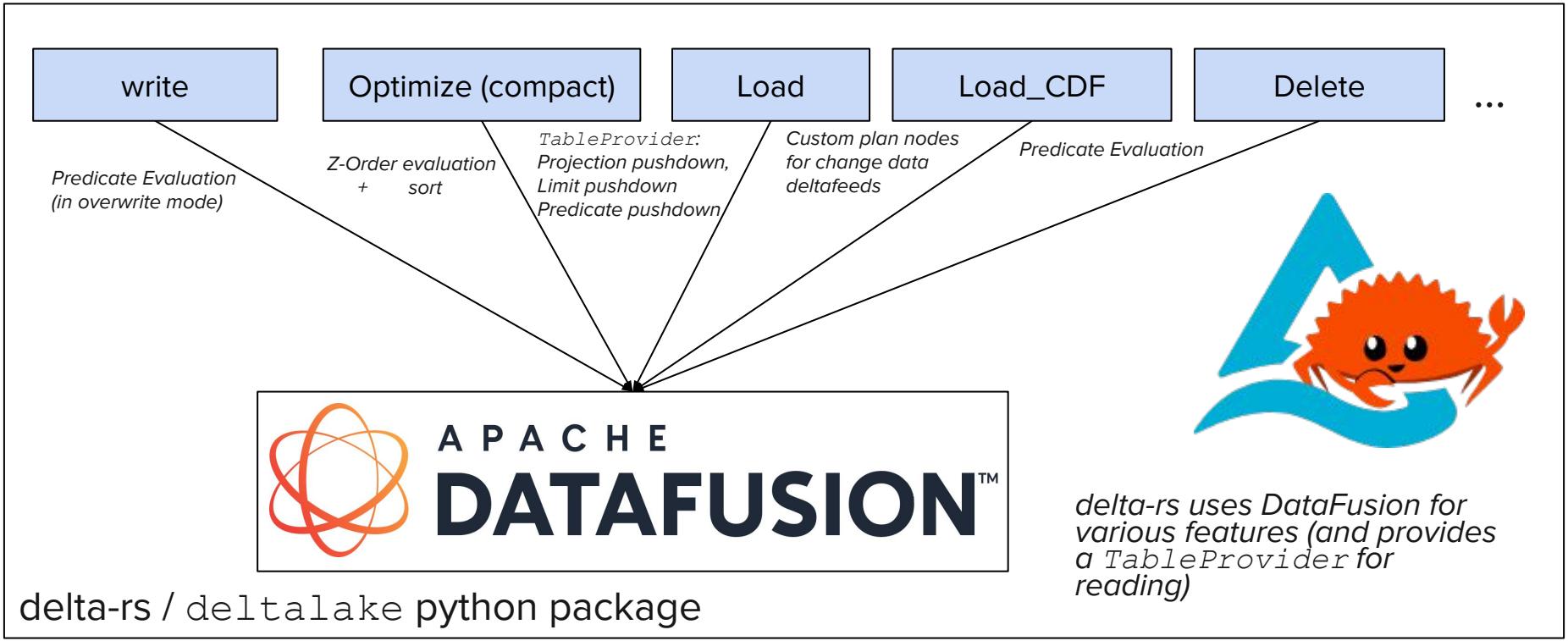
Decoder uses DataFusion `Expr` simplification to calculate zone pruning (source)



Courtesy of Weston Pace

 | © Copyright 2024, InfluxData

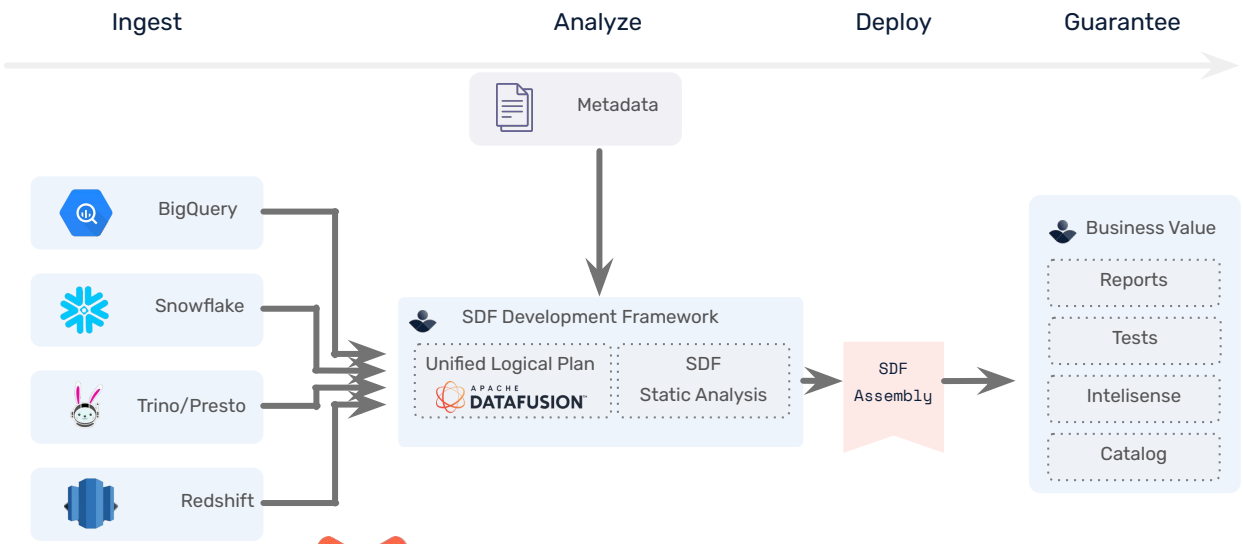
Use Case: Table Formats



Source (link)

© Copyright 2024, InfluxData

Use Case: SQL Analysis (frontend)



Courtesy of ~~SDF~~ / Lukas Shute

SDF uses complete ANTLR Grammars to define many SQL dialects - notably, proprietary ones like Snowflake.

All SQLs compile to a unified Intermediate Representation: the Datafusion Logical Plan.

This gives SDF Executable Semantics.

SDF's transformation layer statically analyzes many logical plans at once for correctness and generates rich metadata.

| © Copyright 2024, InfluxData

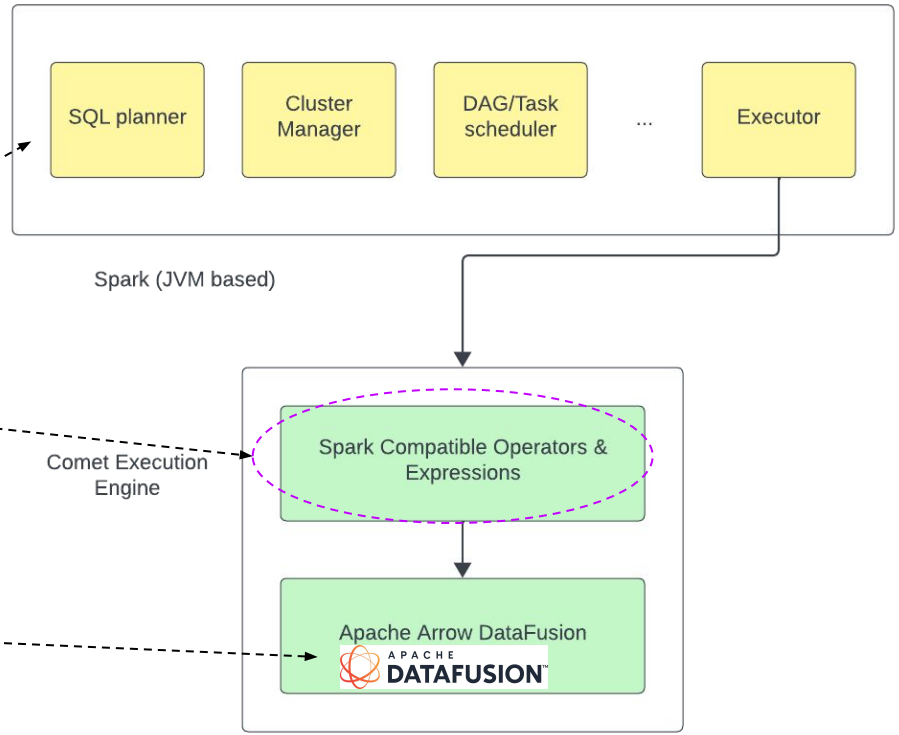
Use Case: Execution Engine

APACHE
DATAFUSION COMET™

Use Spark Planner / Executor machinery

Integration Layer with Spark

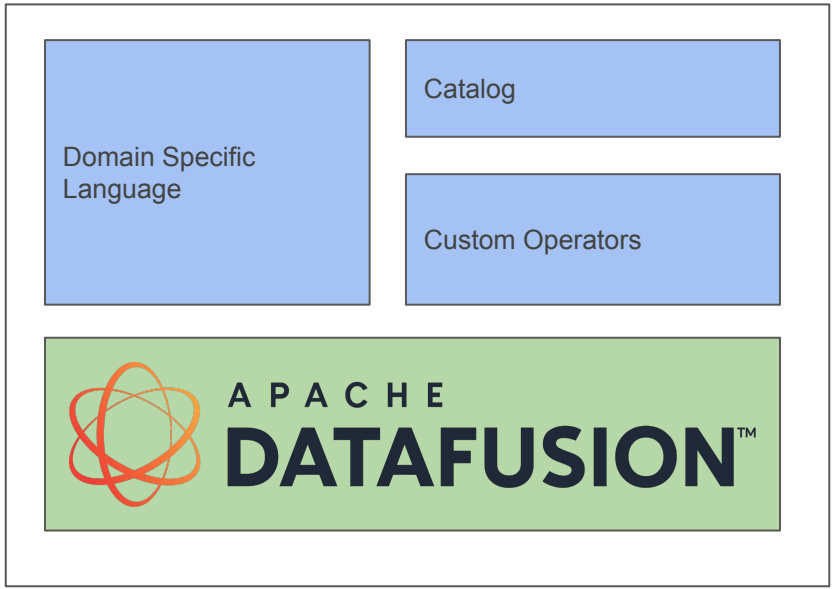
DataFusion's ExecutionPlan Streams



 | © Copyright 2024, InfluxData

influxdata®

Use Case: Specialized Database Systems



Examples

- GreptimeDB (observability)
- InfluxDB 3.0 (timeseries)
- LanceDB (AI)
- HoraeDB (observability)
- Coralogix (observability)
- Exon (bioinformatics)
- Arroyo (streaming)
- Synnada (streaming/AI)
- ...

See more: Apache DataFusion documentation

 | © Copyright 2024, InfluxData

DataFusion: Community + Not Quite Enough Time

“To achieve great things, two things are needed; **a plan**, and **not quite enough time**.”

- Leonard Bernstein (according to the internet)

Who Controls Project / Roadmap

- DataFusion development is NOT directly funded
  - Users contribute together (including engineers paid by companies who build using DataFusion)
  - No one paid full time to work on DataFusion
- ⇒ Always a bit short on time, especially:
  - Reviews, Documentation, Ticket Triage
- Pro: forced prioritization

Plug: Join us to have fun and have a say 

 | © Copyright 2024, InfluxData

Community

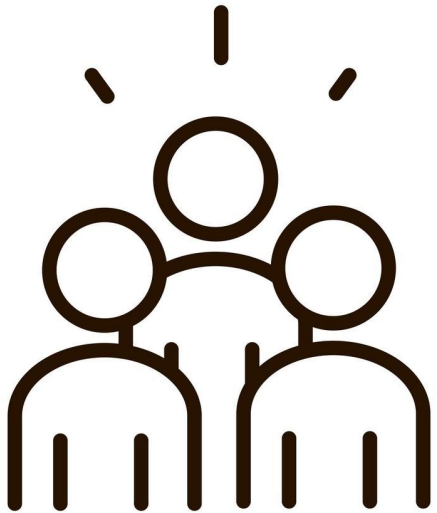
Not started / donated by a company: founded by Andy Grove

Community:

- 694 distinct contributors*
- 14 PMC members + 42 Committers
- Distinct contributors 96 in 43.0.0, 94 in 44.0.0

Velocity:

- Monthly releases for the last 3 years
- Multiple commits a day (😅 still!)



* Caveat: some distortion due to tortured git history

 | © Copyright 2024, InfluxData

Non profit governance of open source communities

“Community over Code” - The Apache Way

 | © Copyright 2024, InfluxData

Apache: Benefits for DataFusion

- ⇒ Predictable Foundation
- **Stable License**: (ASL 20 years old) low risk of changes, (ahem OpenTofu)
- **Communication**: Predictable and open (if slow)
- **Multi-Vendor Participation**: Shared investment reduces individual risk
- **Long Term Maintenance**: Hedged against life changes, corporate strategy shifts, VC funding cycles
- ⭐⭐⭐⭐⭐: Works far better than could be reasonably expected

 | © Copyright 2024, InfluxData

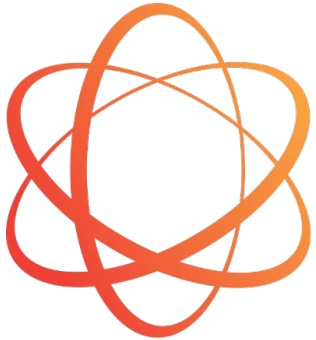
My Personal / Professional Goal

1,000+ projects!

(Used to be a crazy number I just made up. Not so crazy anymore...)

 | © Copyright 2024, InfluxData

Thank You: On with the Talks



APACHE DATAFUSION™

https://datafusion.apache.org/
https://github.com/apache/datafusion

Communication (Slack/Discord/etc)

 | © Copyright 2024, InfluxData

influxdata®

THANK YOU

Backup

 | © Copyright 2024, InfluxData

DataFusion / Query Engine: Input / Output

```
SELECT status, COUNT(1)
FROM http_api_requests_total
WHERE path = '/api/v2/write'
GROUP BY status;
```

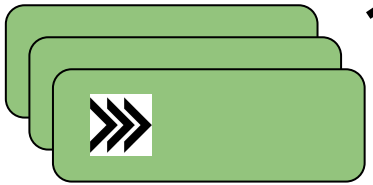
SQL Query

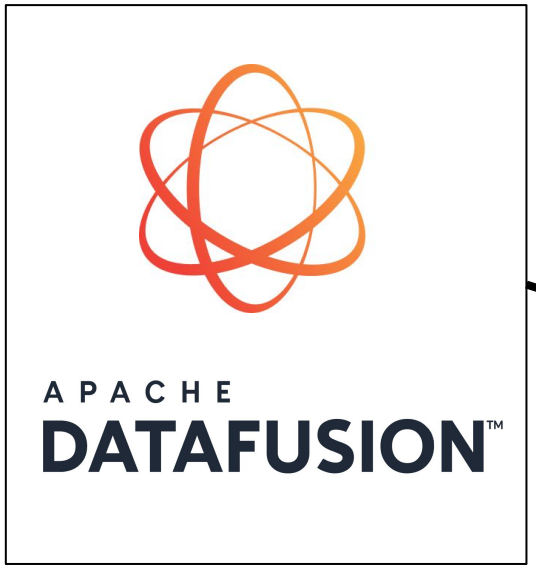
```
ctx.read_table("http")?
  .filter(...)?
  .aggregate(..)?;
```

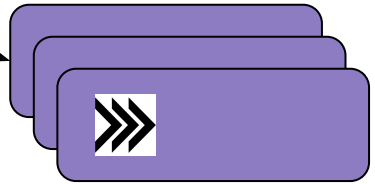
DataFrame

Catalog information:
tables, schemas, etc

Data Batches







Data Batches

 | © Copyright 2024, InfluxData

influxdata®