OSA CON 24

Designing a Lakehouse *for product engineers*

Zhou Sun
co-founder & CEO, Mooncake Labs

November 19-21, 2024

“I don’t care.”

every product engineer ever

that’s it. thanks for attending.

The lakehouse today is polarising…

(instead of unitying :) hope you get the joke)

Data stored in Files in S3

↓

Full table semantics
(Iceberg/Delta)

↓

SQL + Python
single-node + distributed

"The Lake(house) will be the OLAP DBMS archetype for the next 10 years"

you, me, Andy Pavlo, Michael Stonebraker

I don't understand this

↓

I can't get an Iceberg table

↓

My app isn't any better

"This is just for big data people
I really don't get it."

our friends at Clay, Standard Metrics, Inkeep

so, what do product engineers really want?

Fast in-app analytics

Embeddings for vector search

Processed & Clean tables

one data model (ORM) + serverless execution + on their transactional data

a lakehouse might be what they need…

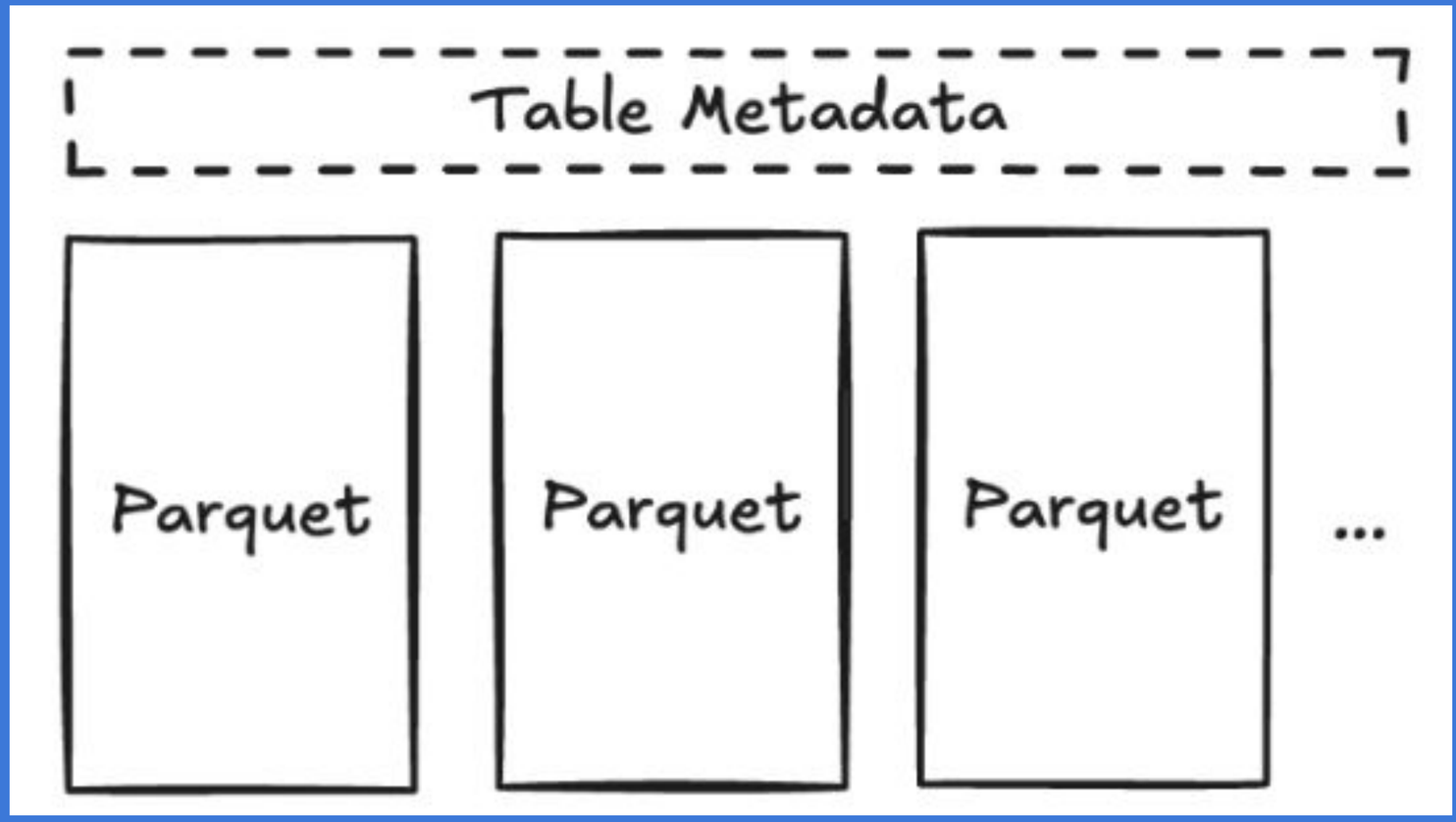
But, not today's lakehouse.

1. Write parquet files from your Postgres database.
2. Upload to S3.
3. Use pyiceberg/ delta-rs to build lake metadata.
4. Set-up an execution engine to query them. (Thanks DuckDB, at least you don't need to set-up Spark…)
5. Realize it's not very fast…
6. Just use a specialized database like ClickHouse
7. {Optional} unfriend the data person who suggested a Lakehouse...

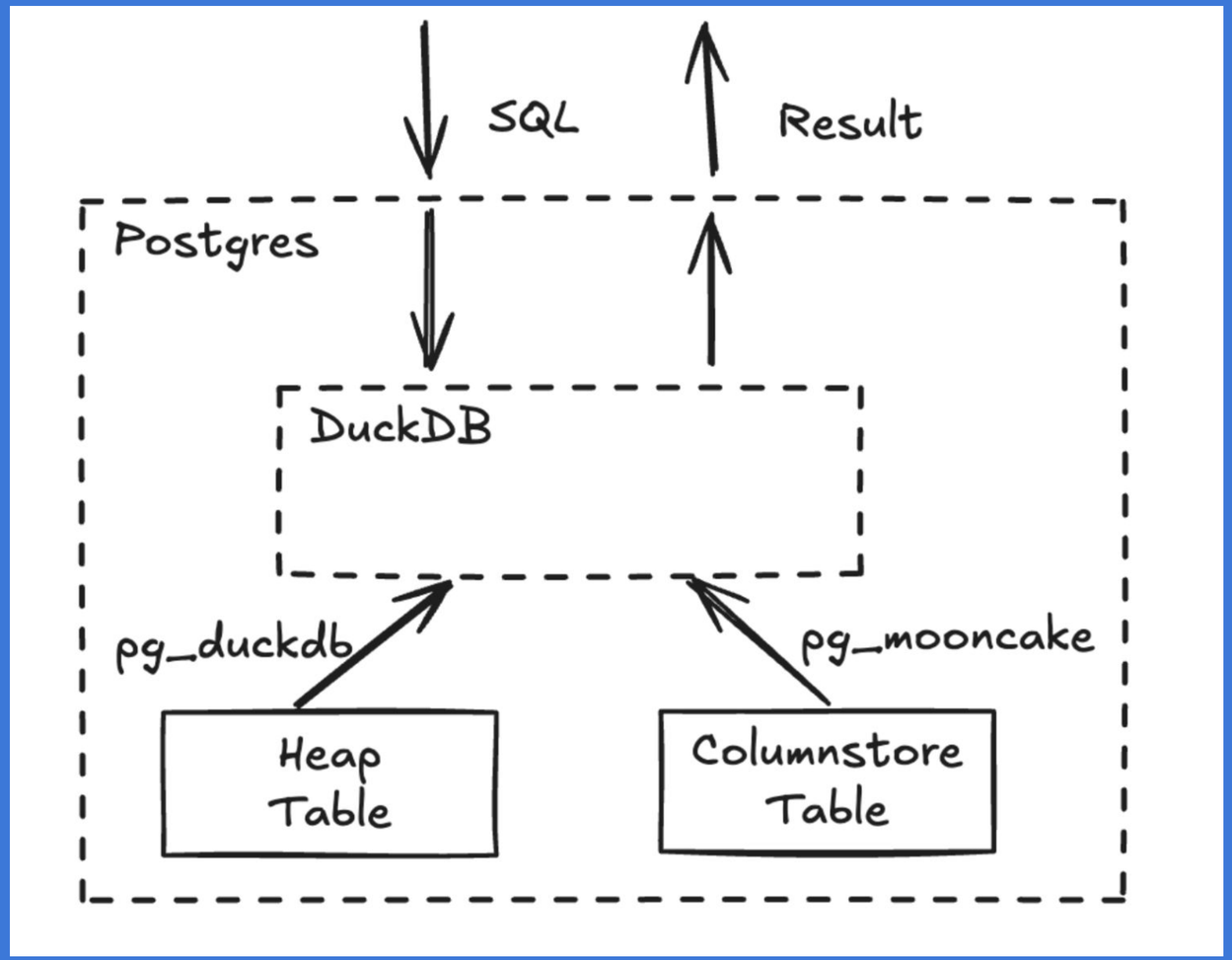
pg_mooncake

a columnstore in Postgres.
a delta/iceberg table outside.

A columnstore table implementation



in Postgres: a columnstore table



Metadata in PG catalog table

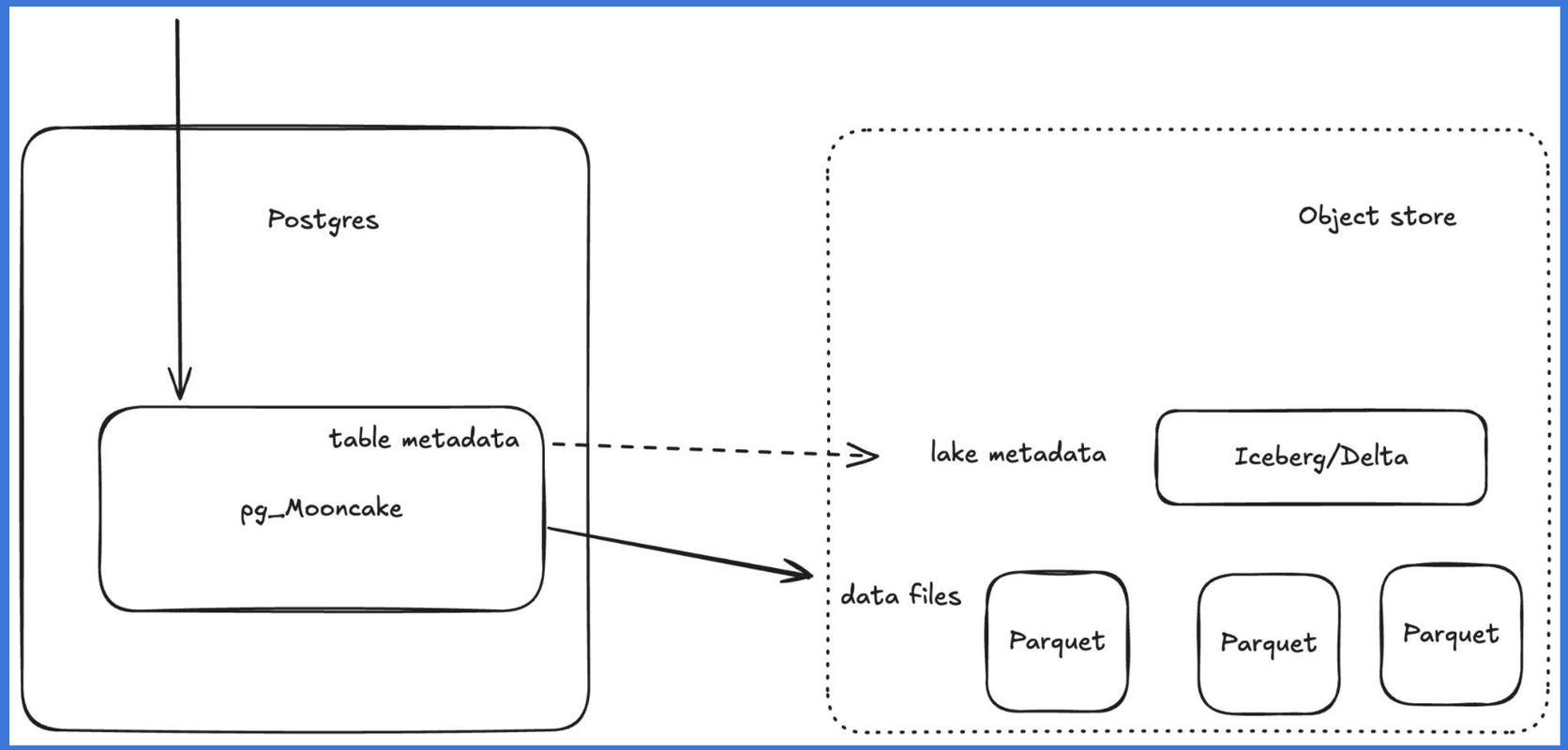
+

DuckDB execution

+

full table semantics

Outside Postgres - Lakehouse Table



Parquet in S3

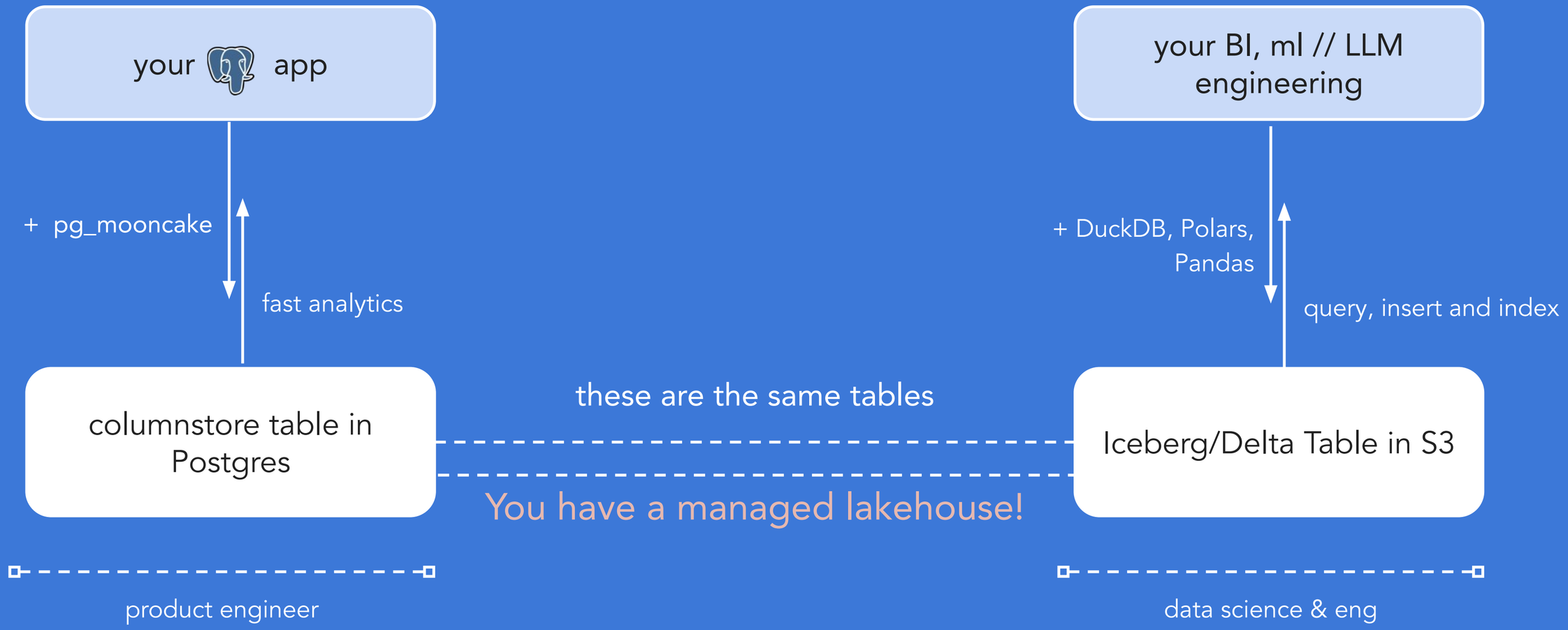
+

Delta/Iceberg metadata

+

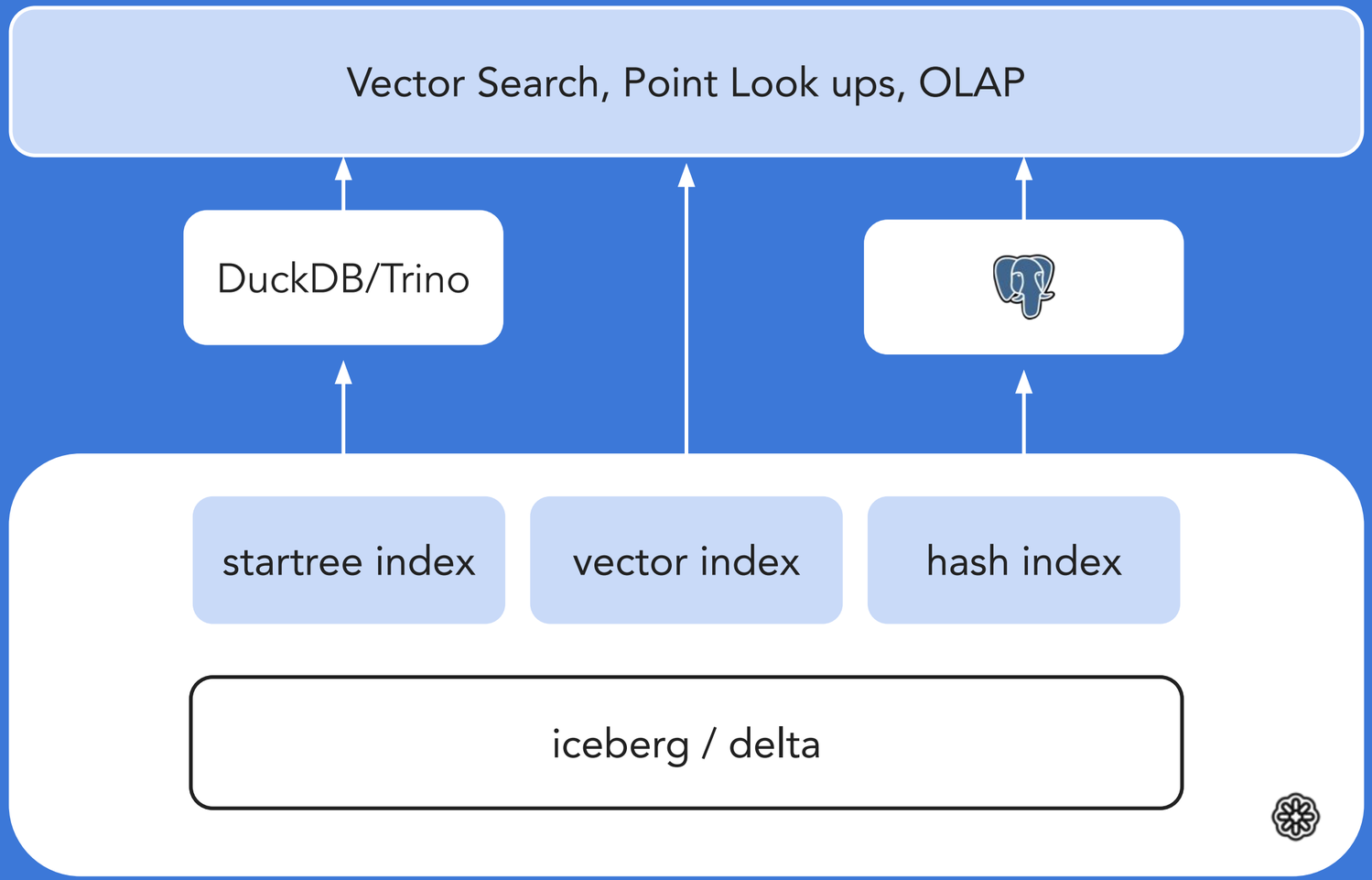
Queryable by any engine

The Mooncake Experience

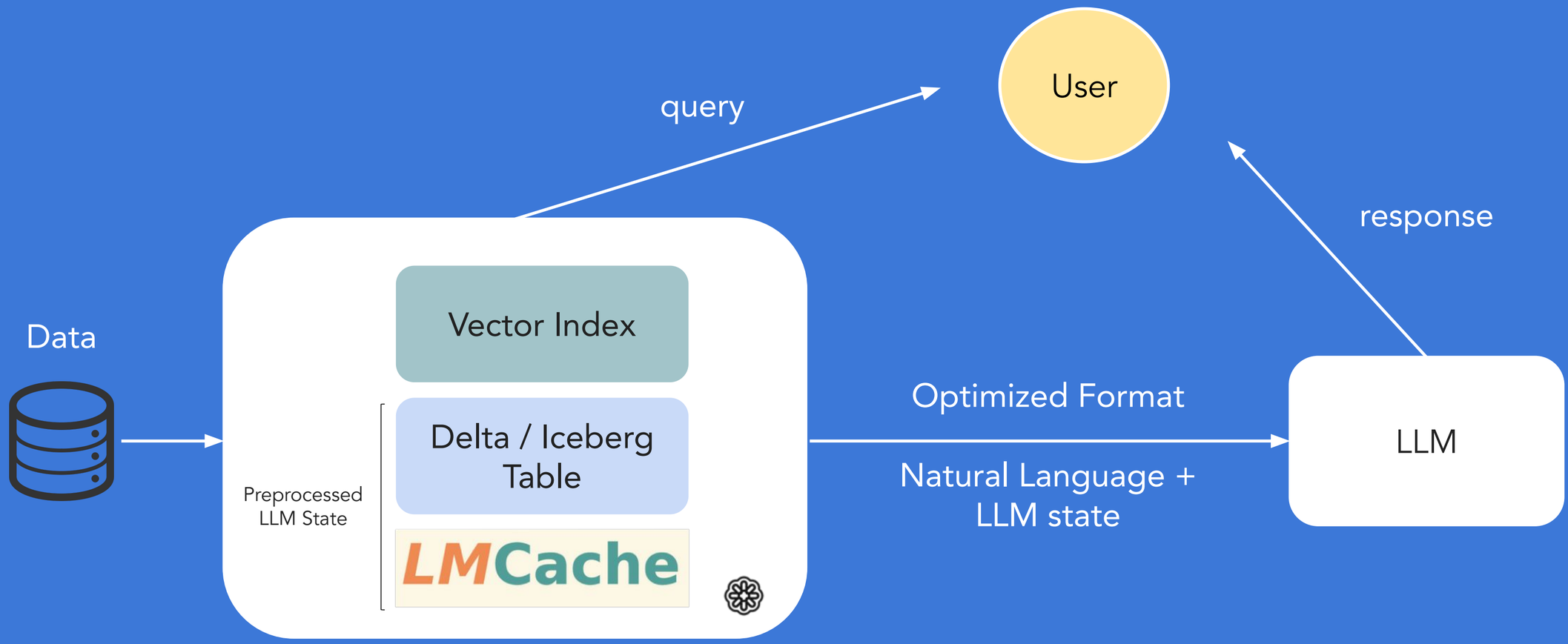


We believe in the Lakehouse…
and more workloads will come to the Lakehouse….

Mooncake Indexes → serve apps directly from the Lake



LMCache + Mooncake → Fast inference on large context



Mooncake is a research lab on the modern Lakehouse.

Questions / Join us: founders@mooncakelabs.com