

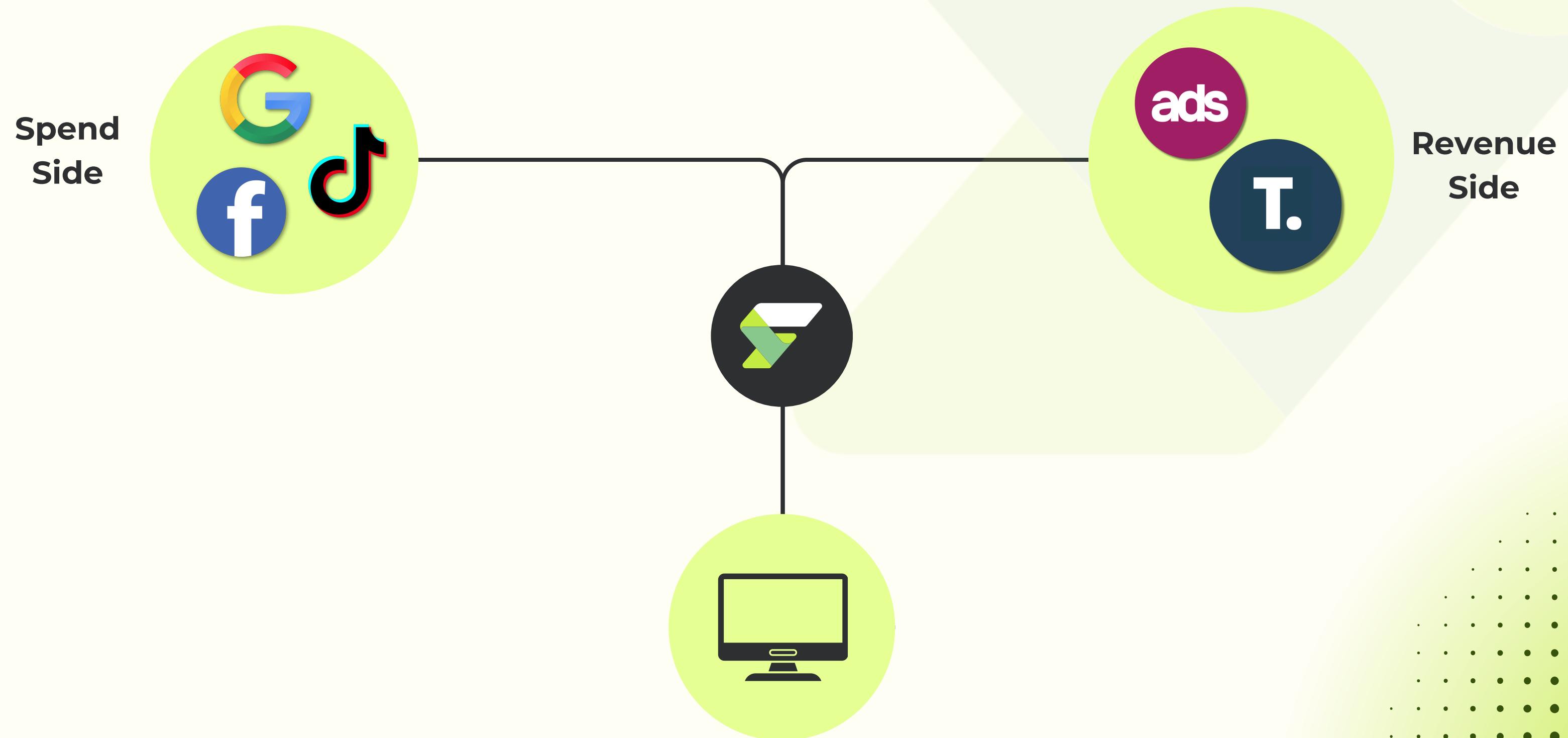
FabFunnel & ClickHouse

Delivering Real-Time Marketing Analytics

Overview

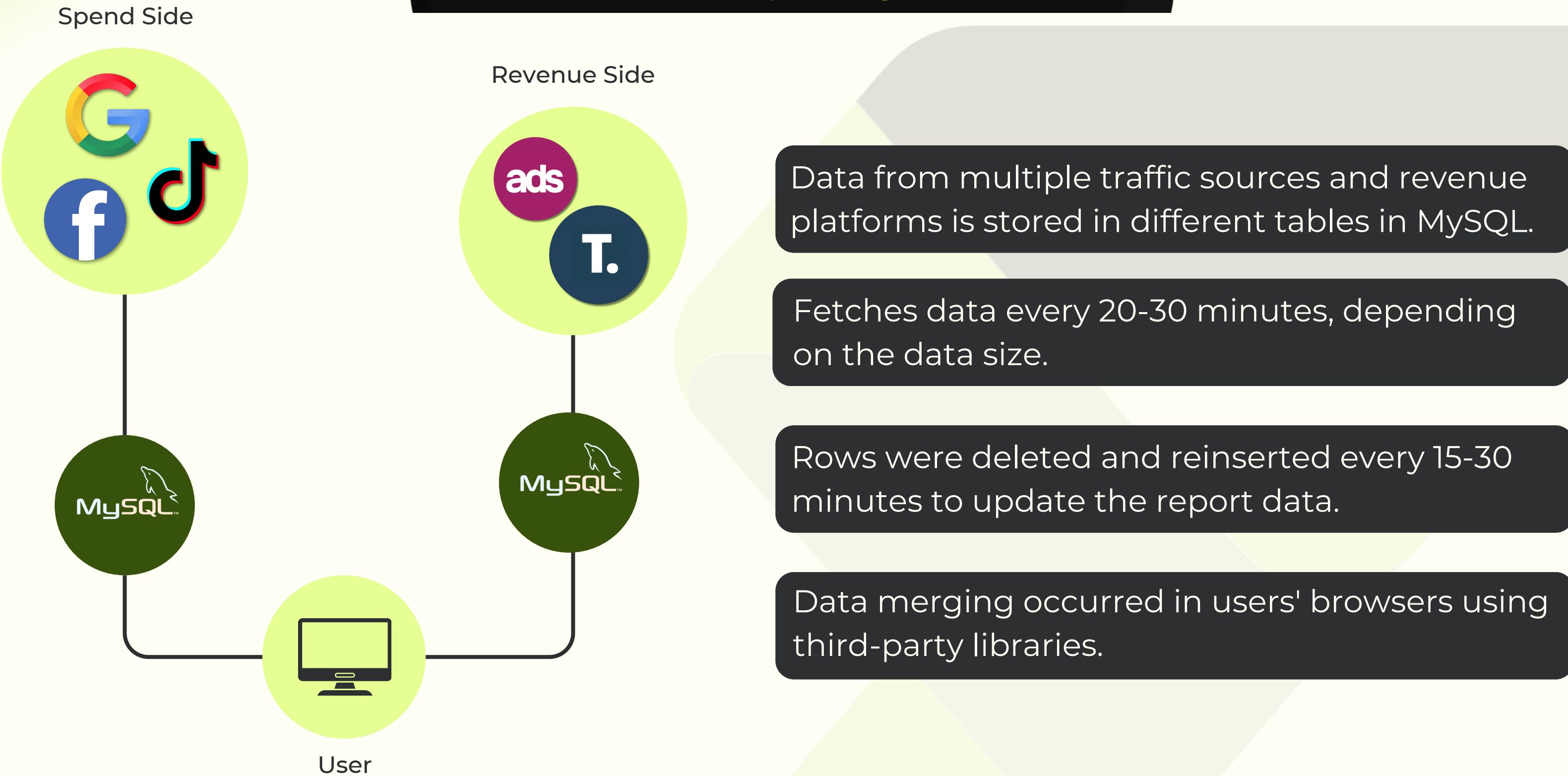
- **What is FabFunnel**
- **What we use to do**
- **Key Challenges**
- **What we wanted?**
- **What are we doing now**
- **Impact of ClickHouse on our Product**

what is FabFunnel



Previous Architecture

Our Previous Reporting Architecture



Key Challenges

Problem Statements

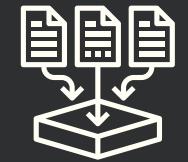
Key Issues



Heavy dependency on joins, causing high **CPU usage** and slow data retrieval.



Frequent updates from multiple sources caused **15-30 min** refresh delays.



Difficulty handling queries for **large date ranges**.



Merging data in browsers caused lag and crashes on **low-end devices**.

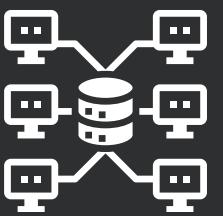
What we wanted?



Required a database supporting fast inserts and real-time analytical queries.



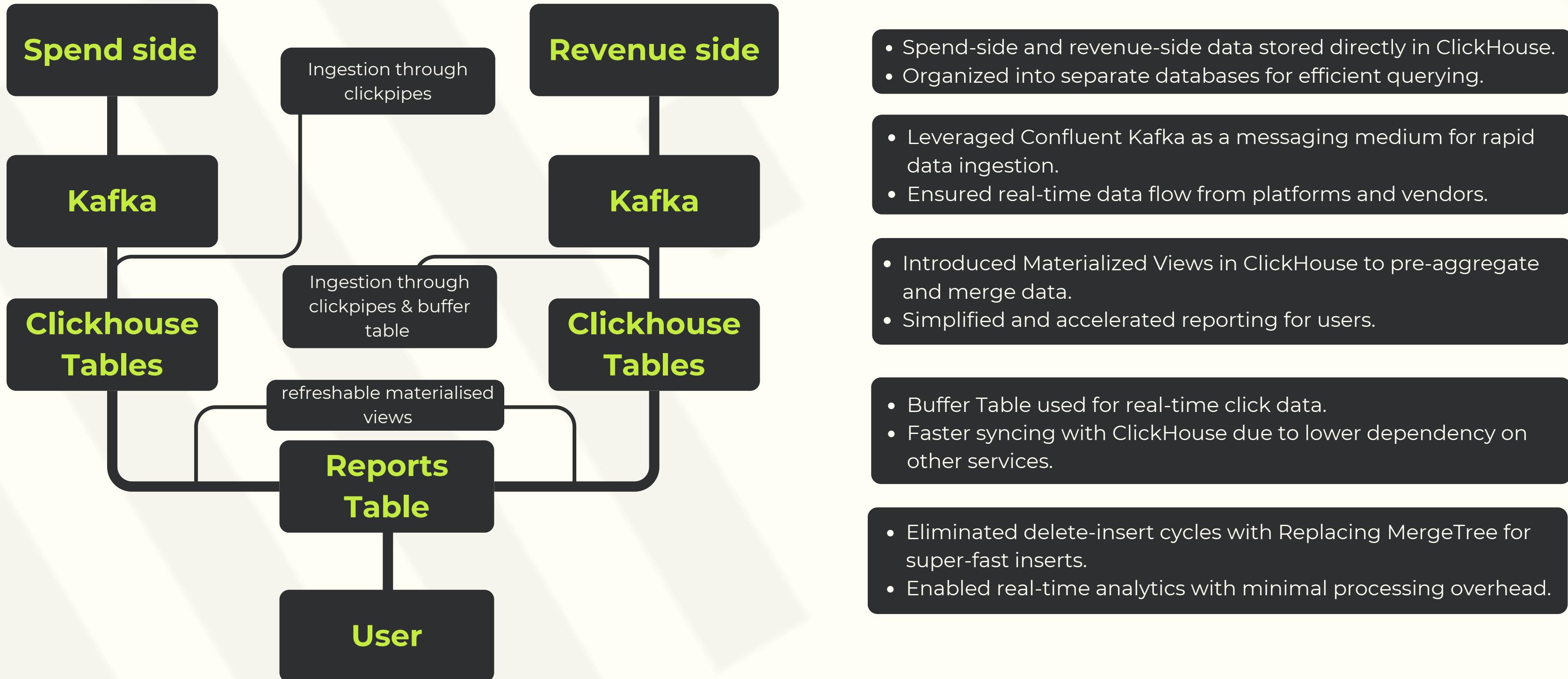
Easy migration due to SQL-like syntax; quickly adopted by the team.



Columnar database structure efficiently handled large volumes of denormalized data.

Transforming Reporting for **Speed, Scalability, and** **Real-Time Insights**

New Architecture with ClickHouse



Impact of ClickHouse on Our Reporting

Speed and Efficiency

Achieved **super-fast inserts** with Replacing MergeTree and eliminated delete-insert cycles, ensuring seamless data updates.

Real-Time Analytics

Enabled **up-to-the-minute insights** with significantly reduced data latency, empowering users to make timely decisions.

Concurrent Queries

Handled high volumes of **concurrent queries efficiently**, enhancing user experience with low-latency responses.

Impact of ClickHouse on Our Reporting

Advanced Aggregation and Reporting

Simplified reporting using **Materialized Views** for **pre-aggregated data**, reducing query complexity and improving response times.

Scalability and Cost Optimization

Ensured scalability to manage **growing data volumes** with **fixed costs** and **minimal infrastructure overhead**.

Performance Metrics (Feb)

- 19.2 million queries processed.
- 11.67 trillion rows read, totaling 605.99 TB of data.
- 5.15 billion rows written, consuming 1.32 TB of storage.
-

Query Execution Efficiency

- Queries processed ~111 million rows in under 1.1 seconds.
- Queries handled between 1.68 GB - 2.12 GB of data per execution in under a second.



Thank You!

TAKEAWAYS

- No **UPDATE/DELETE**: ClickHouse is append-only (mostly). Avoid update-heavy workflows.
- Integrated **ClickPipes** for seamless data flow, reducing delays and improving system efficiency.
- Consider pre-aggregating and storing results if needed.

Q&A

Open the floor for questions.



Anmol Jain
SDE-2 (Full Stack
Developer)



Siddhant Gaba
SDE-2 (Python
Developer)