DeltaStream

Stream Fast, Spend Less

How modern data teams are cutting cloud costs
and simplifying pipelines with DeltaStream

Faster Stream

Table of Contents

Introduction

Why Streaming Is No Longer Optional

Today's data pipelines weren't designed for the speed and scale modern businesses demand. Batch architectures, once the backbone of analytics workflows, are now sources of latency, operational overhead, and spiraling cloud bills. Teams spend more time fixing broken jobs than building new features; and more money running batch loads that often replicate stale logic downstream.

This isn't just a cost problem. It's an architectural problem. And it's pushing teams to reconsider how and where data gets processed.

A streaming-first, shift-left approach offers an alternative: one where data is filtered, validated, and transformed before it hits the warehouse. Instead of bloating compute layers like Snowflake or Databricks, you reserve them for what they do best. And with DeltaStream, governance isn't a casualty of speed—it's embedded at the edge, in real time.

In this guide, you'll learn how a streaming-first model reduces costs, improves agility, and simplifies pipelines without sacrificing governance.

The Real Cost of Batch

For many teams, batch processing is the default. Not because it's optimal, but because it's familiar. But that comfort hides real costs.

Batch jobs are typically scheduled on fixed intervals, not triggered by need. They process everything, even stale or duplicate data. When jobs fail (and they do), they often cascade, delaying critical analytics and triggering expensive re-runs.

The deeper issue is that batch pipelines tend to shift responsibility downstream. Data typically lands in its raw form, without immediate processing. Cleaning is postponed until later stages, and governance is often handed off to a separate team or tool altogether.

This delayed responsibility creates several downstream problems. Compute resources sit idle during off-peak cycles, unnecessarily burning through credits. Rigid job schedules slow time-to-insight, making it harder for teams to respond quickly to changing needs. Operational efficiency suffers as teams are forced to constantly monitor, rerun, and debug late-stage jobs. Meanwhile, platforms like Snowflake and Databricks are overburdened, used for basic filtering and transformation tasks that should have been handled earlier in the data pipeline.

The bottom line: teams pay cloud bills not just for meaningful work, but for inefficiencies baked into the pipeline.

Compute waste across batch pipeline stages

Ingestion → Processing → Cleaning → Training
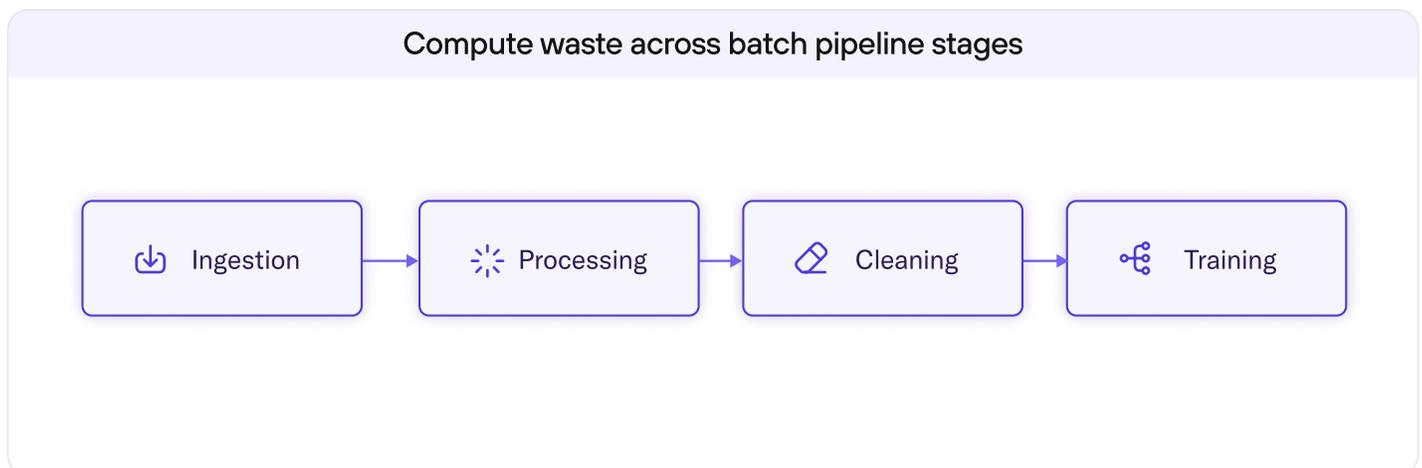

What "Streaming-First" Really Means

Streaming-first isn't just a speed play. It's a design decision that prioritizes simplicity, efficiency, and precision.

In a streaming-first architecture:

- Data is processed as it arrives, not hours or days later
- Events are filtered, validated, and shaped on ingest
- The warehouse becomes an endpoint, not a processing hub

This approach minimizes redundant work and reduces the surface area for errors. But most importantly, it gives teams the ability to act early on data quality, structure, policy, and access rather than react after the fact.

DeltaStream enables this shift by acting as a live governance and transformation layer that sits directly on top of your event streams. It enforces schemas, applies policies, and lets teams shape data midstream using SQL before it ever hits the lakehouse.
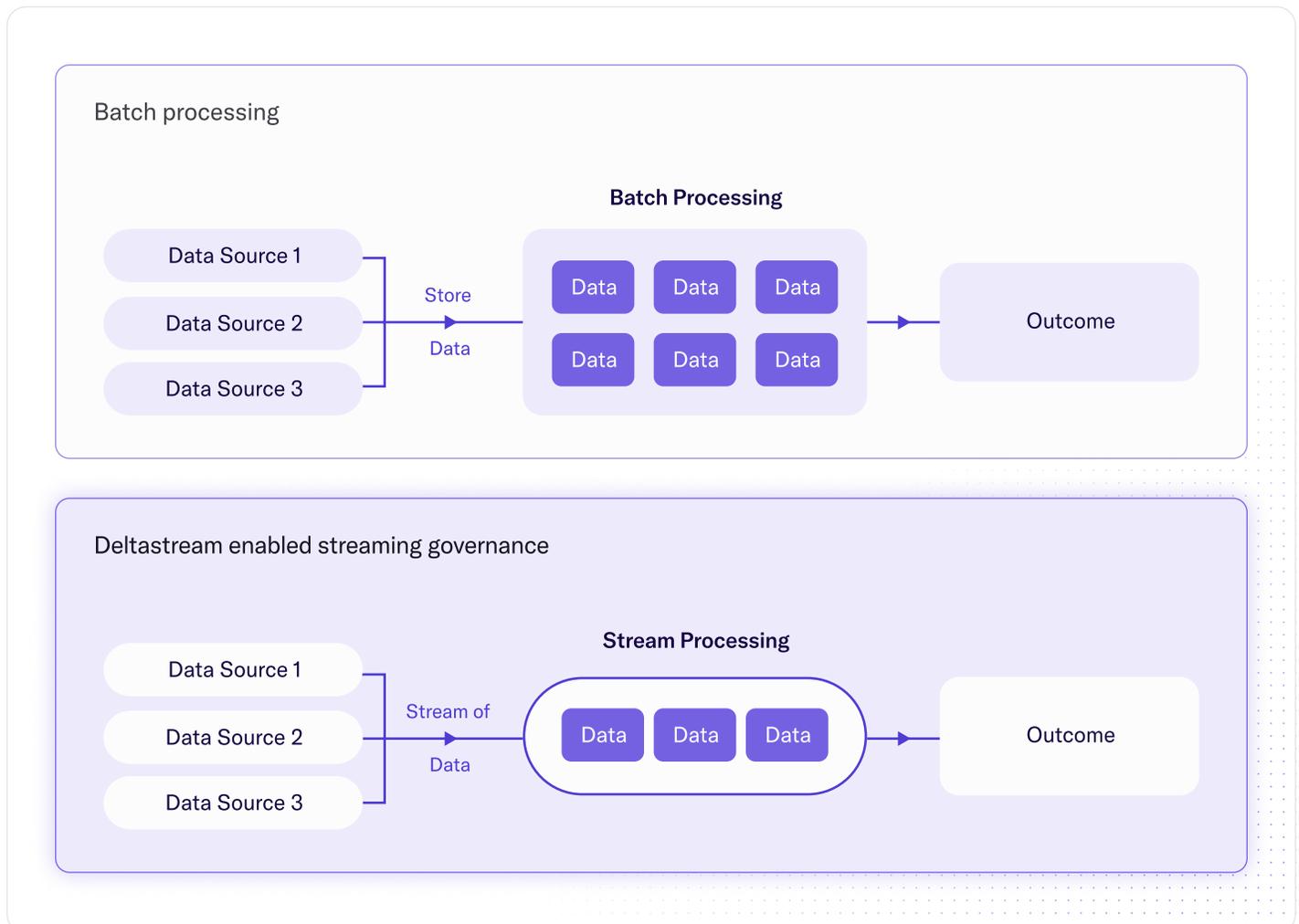
The Shift-Left Advantage

The concept of "shift left" means finding and fixing issues earlier in the lifecycle when they're cheaper and easier to address. From a data engineering standpoint, it refers to the application of logic upstream, at the edge, instead of relying on downstream clean-up.

This shift has practical benefits:

- Lower compute costs by reducing transformation workloads in Snowflake
- Faster analytics by delivering structured data earlier
- Simplified pipelines by eliminating redundant stages and tools
- Stronger governance by enforcing policies where the data originates

DeltaStream supports this "shift left" approach by making it easy to:

- Write reusable SQL-based transformations on the stream
- Apply schema validation before bad data lands
- Enforce access controls at the topic or column level

The further upstream you catch problems, the less they cost to fix.

Shift Left

📄 **Source** → ((•)) **Streaming Validation + Transformation** → 🏬 **Data Warehouse**
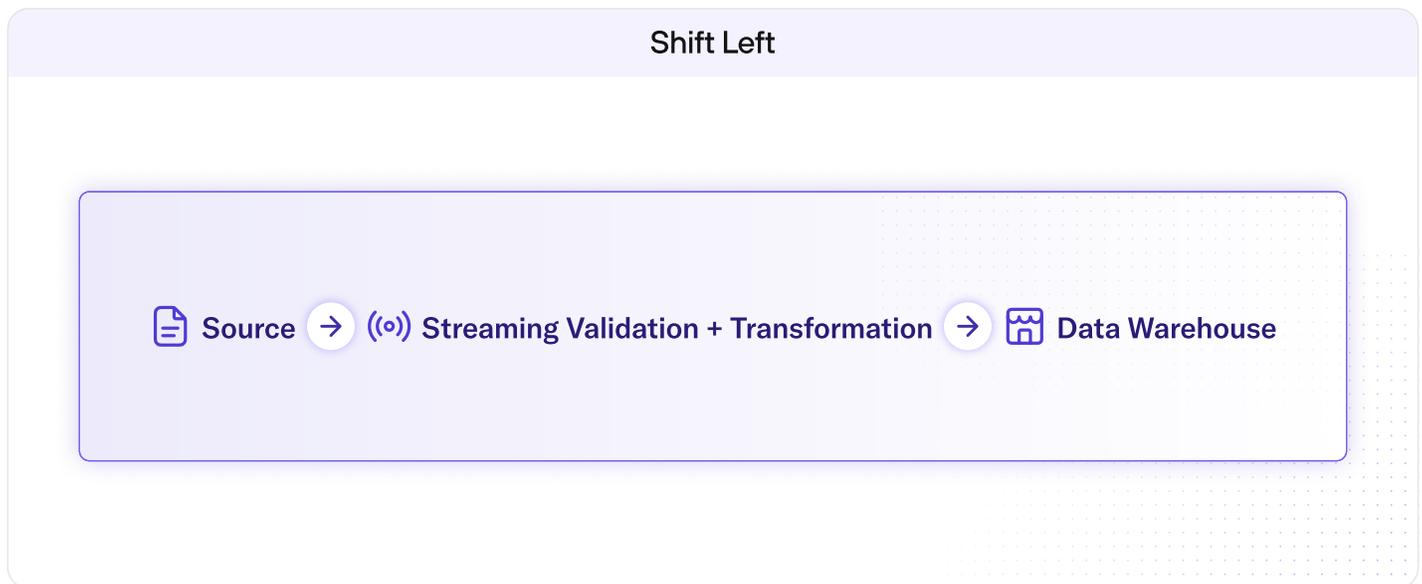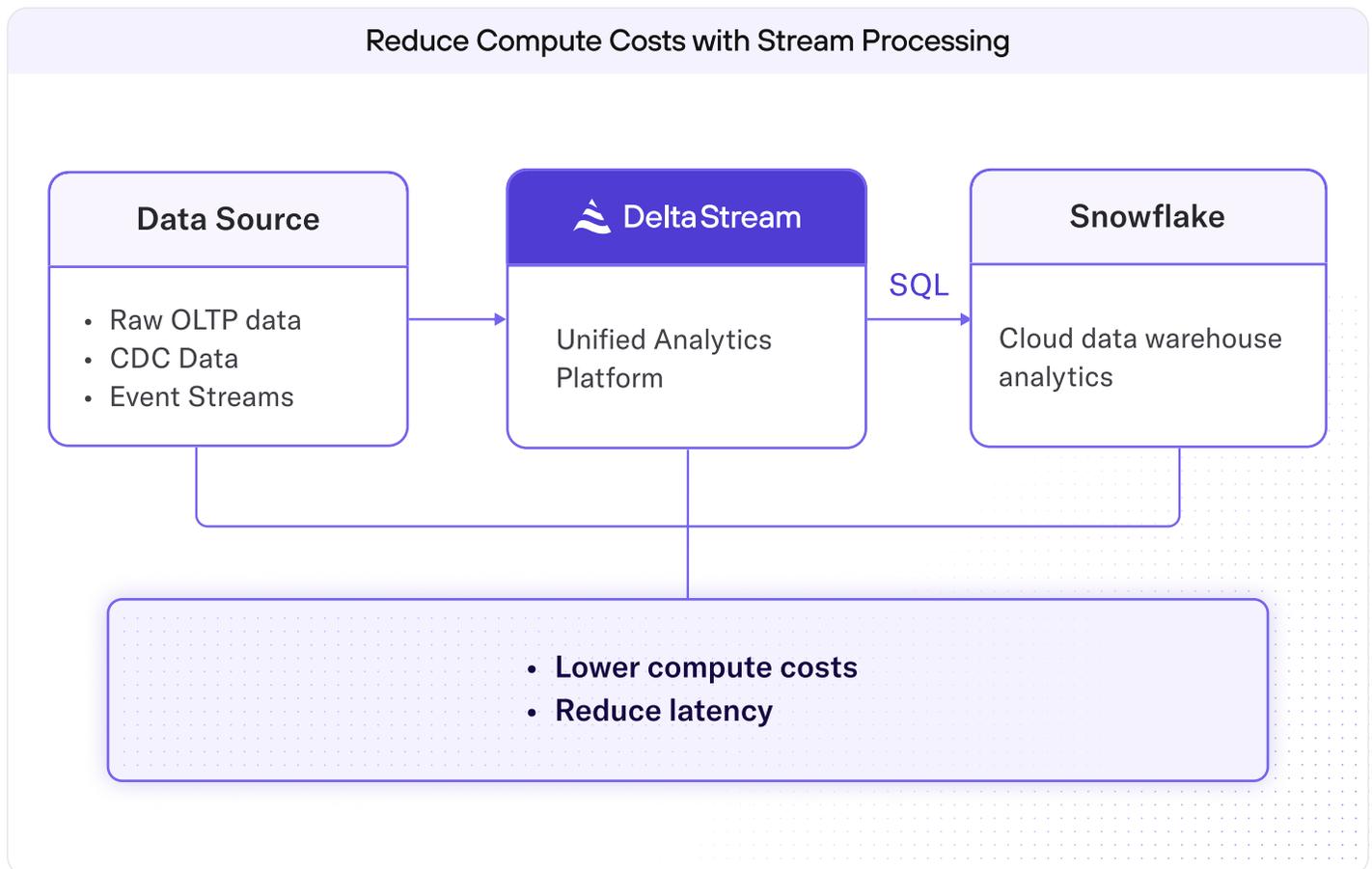
Introducing DeltaStream Fusion

DeltaStream Fusion brings structure to the chaos of streaming data.

Instead of requiring multiple tools for stream ingestion, processing, and governance, Fusion offers a unified control plane for everything in motion. It connects producers, applies policies, performs transformations, and delivers data to your lakehouse or other downstream systems.

With Fusion, teams get:

- A unified platform for batch, stream, and real-time
- Native lakehouse support (Snowflake, Databricks, etc.)
- Built-in governance without extra layers or tools required
- Operational simplicity with fewer systems to maintain

It's streaming-native, but designed to meet the realities of hybrid stacks.

Reduce Compute Costs with Stream Processing

Data Source	DeltaStream	Snowflake
• Raw OLTP data • CDC Data • Event Streams	Unified Analytics Platform	Cloud data warehouse analytics

SQL

- **Lower compute costs**
- **Reduce latency**

From Benchmark to Business Case

DeltaStream's architectural bet has been validated in real-world benchmarks.
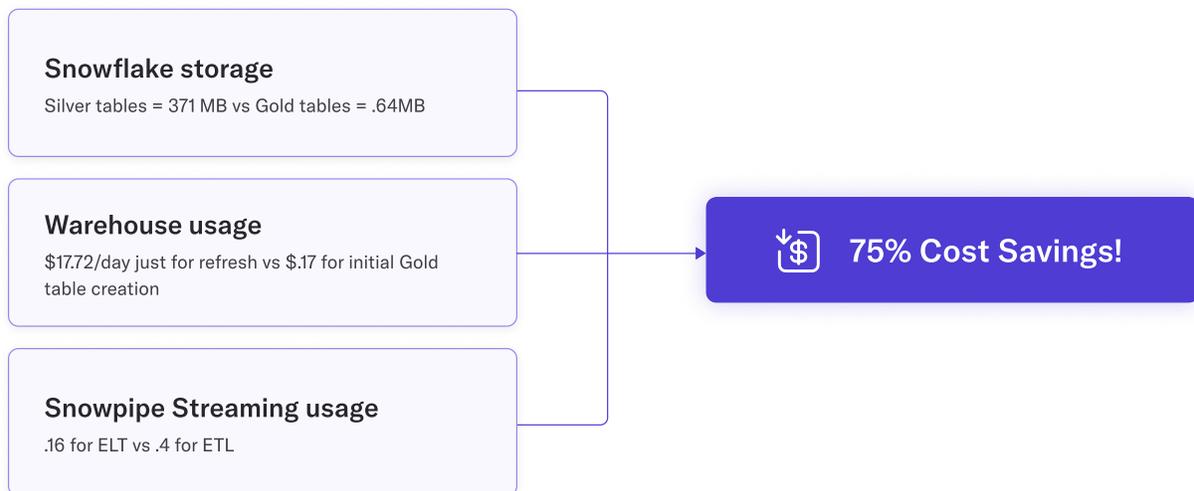In one test using NYC Taxi data, DeltaStream reduced Snowflake compute costs by 75% and improved time-to-insight by 4x.

These gains weren't from tuning queries or resizing clusters. They came from doing less downstream work because data was filtered, validated, and transformed at the edge.

Translating that into business impact:

- Less time spent managing jobs
- Fewer resources burned on bad data
- More confidence in what lands in your lakehouse
- Faster insights with less infrastructure

If your warehouse is overloaded, your pipeline is sprawling or your costs keep climbing, streaming-first isn't a future goal. It's your next step.

Snowflake storage
Silver tables = 371 MB vs Gold tables = .64MB

Warehouse usage
$17.72/day just for refresh vs $.17 for initial Gold table creation

Snowpipe Streaming usage
.16 for ELT vs .4 for ETL

75% Cost Savings!

Learn more about how we ran our benchmark

Read about how shifting left with DeltaStream can cut Snowflake costs by 75%

Read the blog →

Conclusion

Stream Smarter, Spend Less

The batch-first model worked when speed wasn't critical and complexity could be tolerated. But today, waiting to govern data after it lands leads to delays, bloated pipelines, and rising costs.

A streaming-first, shift-left approach changes that. It moves governance to the edge, reduces downstream processing, and gives teams access to clean, trusted data earlier in the workflow. Instead of scaling infrastructure to handle complexity, you reduce the complexity altogether.

DeltaStream helps teams make this transition. It applies policies as data flows, transforms records before they land, and unifies governance across real-time and batch without disrupting existing systems.

The result is a leaner architecture, lower cloud spend, and faster insights. It's a practical way to simplify your stack and future-proof your data operations.

Let's talk about what streaming-first could look like for your team.

DeltaStream