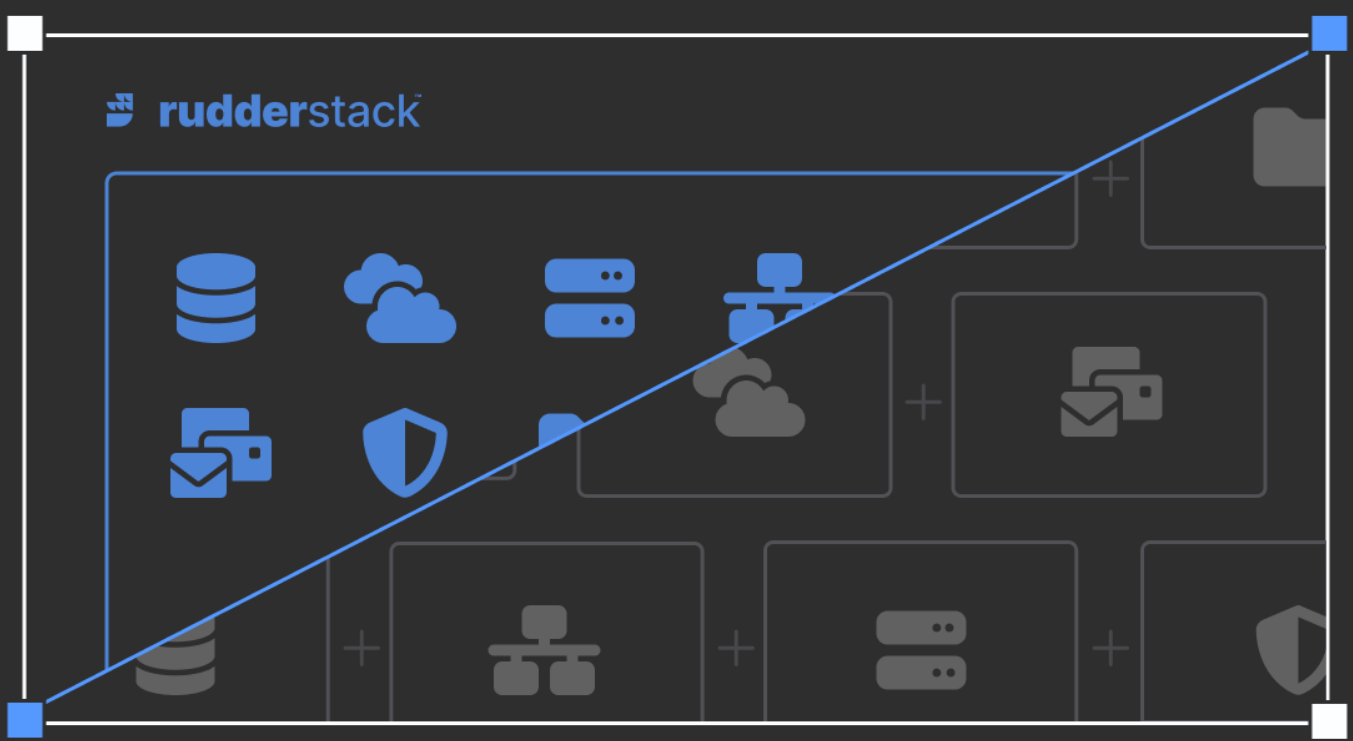
rudderstack™

The Definitive Guide

3 architectures for a cost-efficient data stack



3 architectures for a cost-efficient data stack

Table of contents

Three architectures to make your data stack more efficient in 2023

In 2023, data teams are working hard to make every project and every piece of data count in the face of uncertain economic times.

Companies are accelerating projects that leverage data to become more revenue efficient, like personalization and churn reduction, but they are also scrutinizing expenses and ensuring that every piece of technology is pulling its weight in driving value.

Over the past several months we've worked with many of our customers to evaluate their stacks and find opportunities for efficiency.

Here are three effective approaches we've seen data teams use to leverage RudderStack for making their function more cost-efficient in terms of both time and technology.

Time is money: eliminate low-value integrations engineering

The number of tools used by every team in modern companies is exploding (the average company has 75 tools!).

This often translates into custom integrations work for data teams across the stack, from ingesting data to creating custom pipelines that send that data to whatever downstream tools need it.

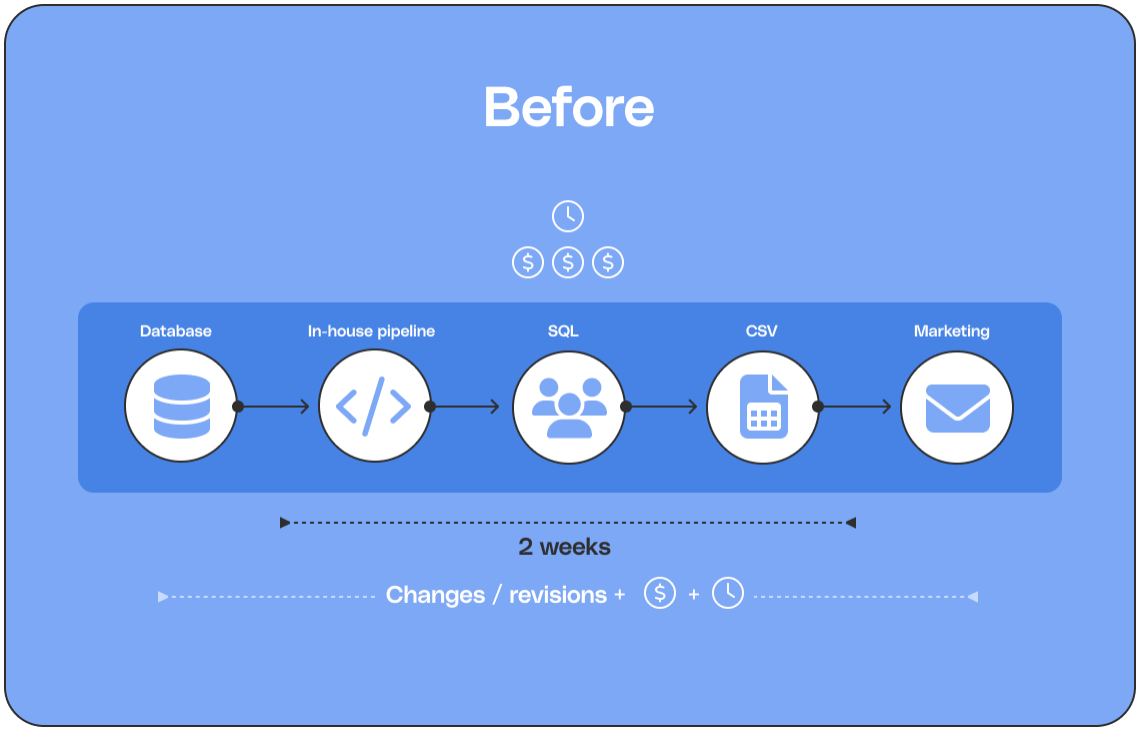
Here's a specific example from one of our customers (an international furniture retailer):

The problem

The data team spent a significant amount of time responding to marketing requests for adding new dimensions to user profiles for email campaigns, which required:

- Collecting data that was fragmented across the stack
- Working with analysts to join data and produce the desired data points
- Managing and running custom-built pipelines to get that data into the marketing team's email automation tool

Because of time delays, the marketing team would often rely on CSV exports, which they cleaned and loaded directly into their tool, which introduced constant challenges related to data discrepancies and file size limitations.



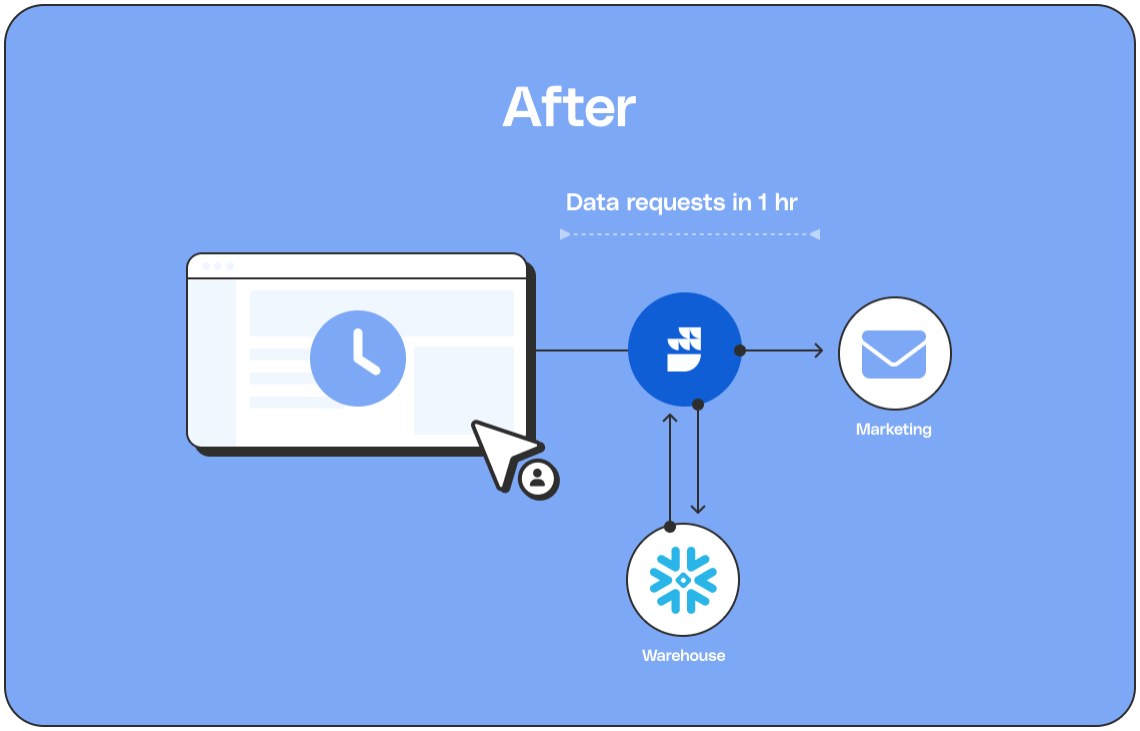
Old architecture:

- Brittle in-house pipelines
- Complex queries to join basic data
- 1-2 weeks to get new data points into marketing's platform

- Manual, error-prone data cleaning in spreadsheets

The solution:

This particular data team implemented a data integration layer using RudderStack's out-of-the-box integrations, eliminating all of their custom in-house solutions and manual data wrangling. The data team is now focused on helping marketing implement use cases key to thriving in 2023.



New architecture:

- Data is sent directly to the email automation tool
- Data is sent directly to the data warehouse
- Reverse ETL enables fulfillment of new data requests in 1 hour

Money is money: replace legacy CDPs with data infrastructure

10 years ago, legacy CDPs that stored your data made sense because data warehouses were still relatively new and cost-prohibitive for many companies.

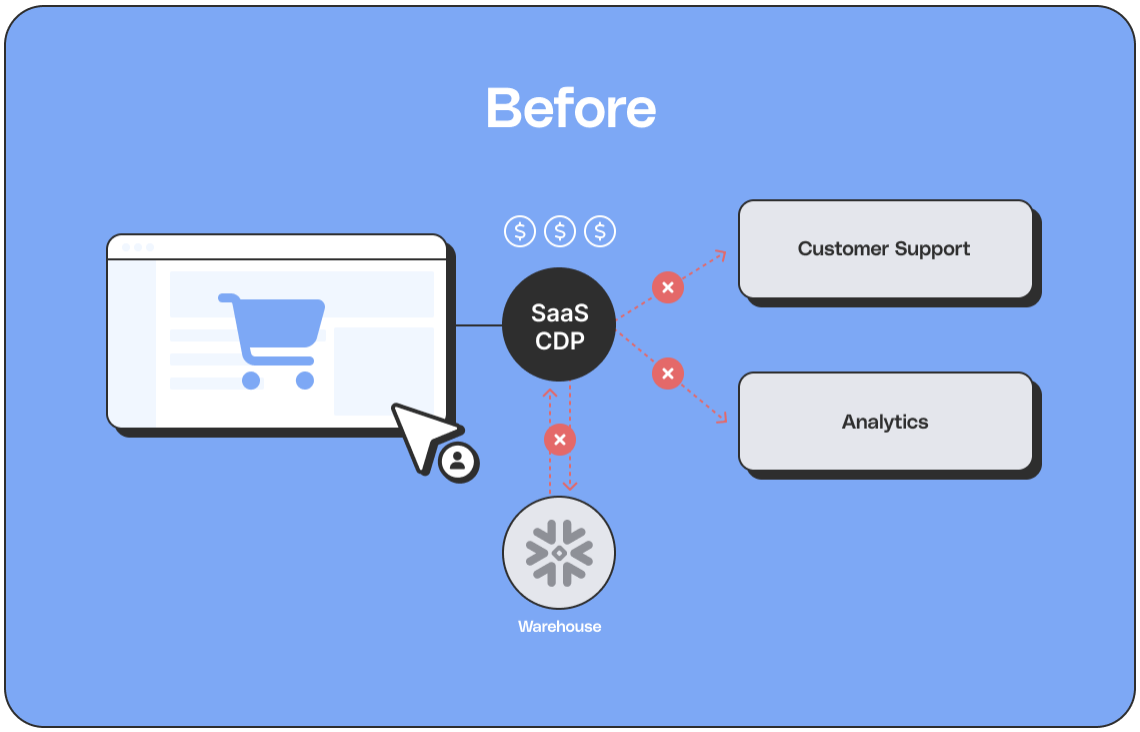
Today, the cloud data warehouse has become the most scalable, efficient, and secure place to centralize your data and ship complex projects like ID resolution and multi-channel attribution—all without creating another silo with an expensive CDP.

Let's take a look at an example:

A national shoe brand wanted to grow their eCommerce business with improved customer experiences, and knew that ID resolution was a crucial step to personalization and other use cases.

The problem:

Their legacy CDP vendor was costly, inflexible, and created a separate data set that kept teams siloed.



Old architecture:

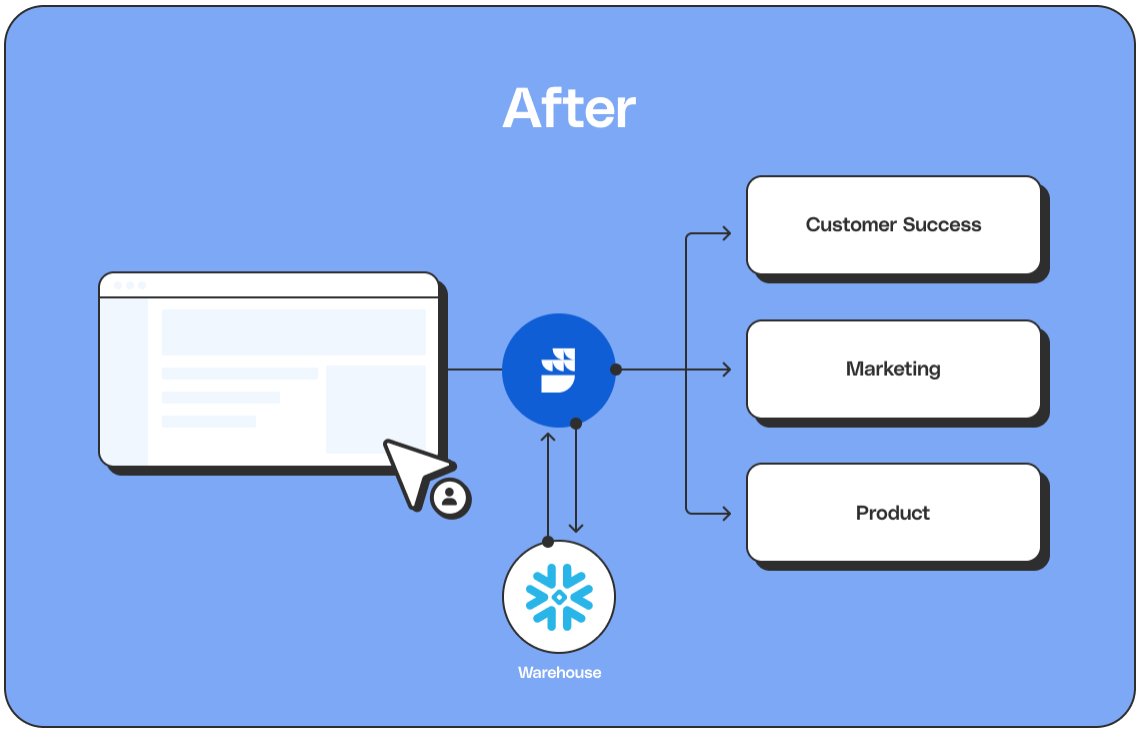
- No ID resolution
- Limited use cases and data integrations
- Limitations in syncing internal data into vendor platform

- no Reverse ETL
- Trigger mechanisms mostly based on order data

The solution:

They replaced their legacy CDP with their own data warehouse (Snowflake), and RudderStack's Event Stream and Reverse ETL pipelines, giving them far more control, scalability, and flexibility at a fraction of the price.

Plus, they improved security and compliance because the legacy CDP no longer kept a copy of their data.



New architecture:

- Warehouse is the single source of truth
- ID resolution with all warehouse data and compute power
- Pricing based on usage of owned infrastructure, low to no maintenance cost
- Full ownership of data - secure and compliant
- Advanced features for data sharing, APIs and code-based transformations

By replacing their expensive legacy CDP vendor, the data team was able to extend their Snowflake investment and help the business deepen customer relationships with consistent data across their entire stack, and personalized experiences across the customer journey.

Flexibility is money: move faster with adaptable infrastructure

In 2023, taking advantage of opportunities quickly will be key to staying ahead of the competition. For data teams, that means fueling ideas and initiatives with data and helping teams get them live as quickly as possible.

New projects often create new requirements, which often means either some sort of custom engineering work or buying a new tool to do the job, both of which will be extremely hard to justify in the current environment.

We've seen our customers use RudderStack's flexible infrastructure to build creative solutions that help other teams move fast while avoiding unnecessary costs.

Here are two examples.

Using Transformations to enrich data payloads with 3rd-party geography and account information

There are entire companies focused on data enrichment and their services can be pricey, especially as add-ons for CRMs and marketing tools.

There are plenty of powerful, cost-effective APIs out there, though, and our customers use RudderStack Transformations to hit those APIs and enrich event payloads in the pipeline before they are delivered to downstream tools.

Geolocation based on IP

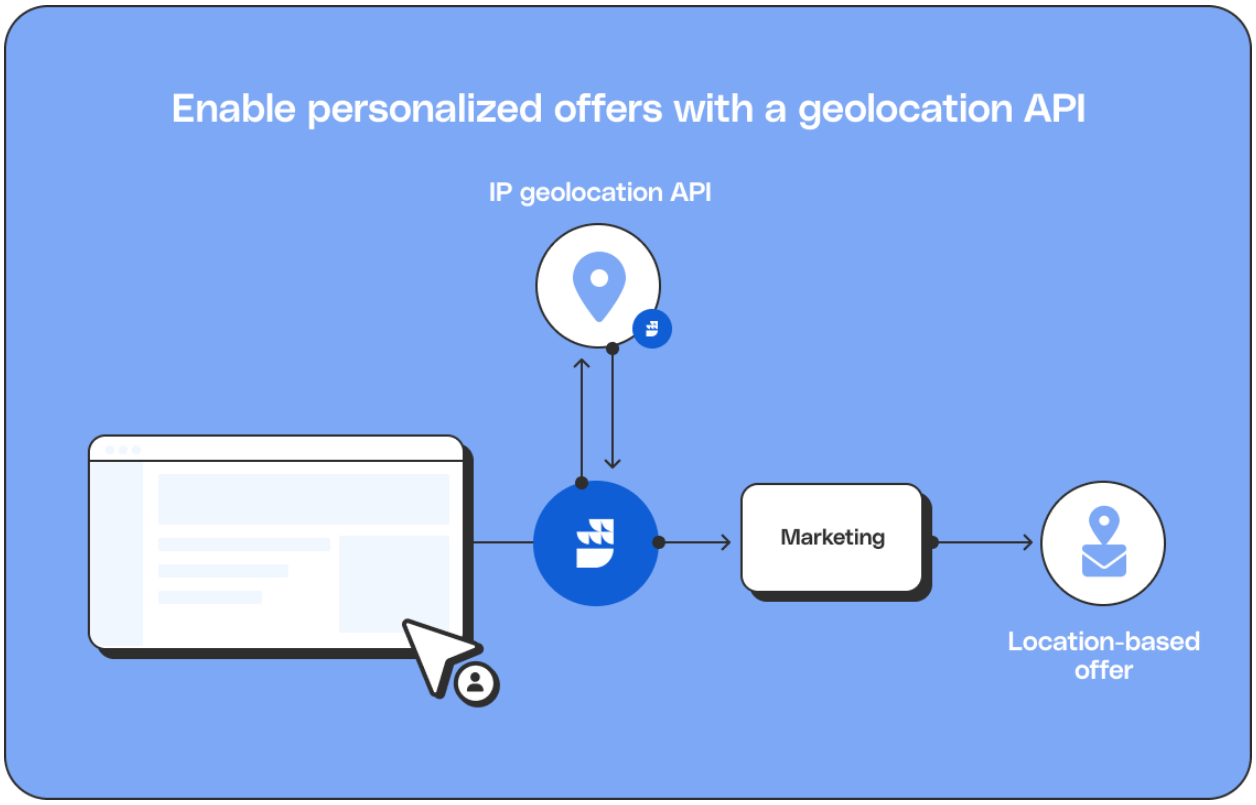
Problem

A customer marketing team figured out that delivering localized offers by region significantly increased conversion rate, but they didn't have addresses for many of their users.

Solution

Their data team leveraged a cost-effective geolocation API to solve the problem. In a RudderStack Transformation, they passed the user's IP address to the service and appended the returned region to the payload, which was passed into a custom field in the marketing platform.

The marketing team was then able to automatically segment users into regional lists and trigger location-based offers in real-time.



Bonus: the data team was also able to scrub the actual IP address from the payload before delivery to the marketing tool to maintain PII compliance.

Using Webhooks to hydrate internal APIs and tooling

Problem

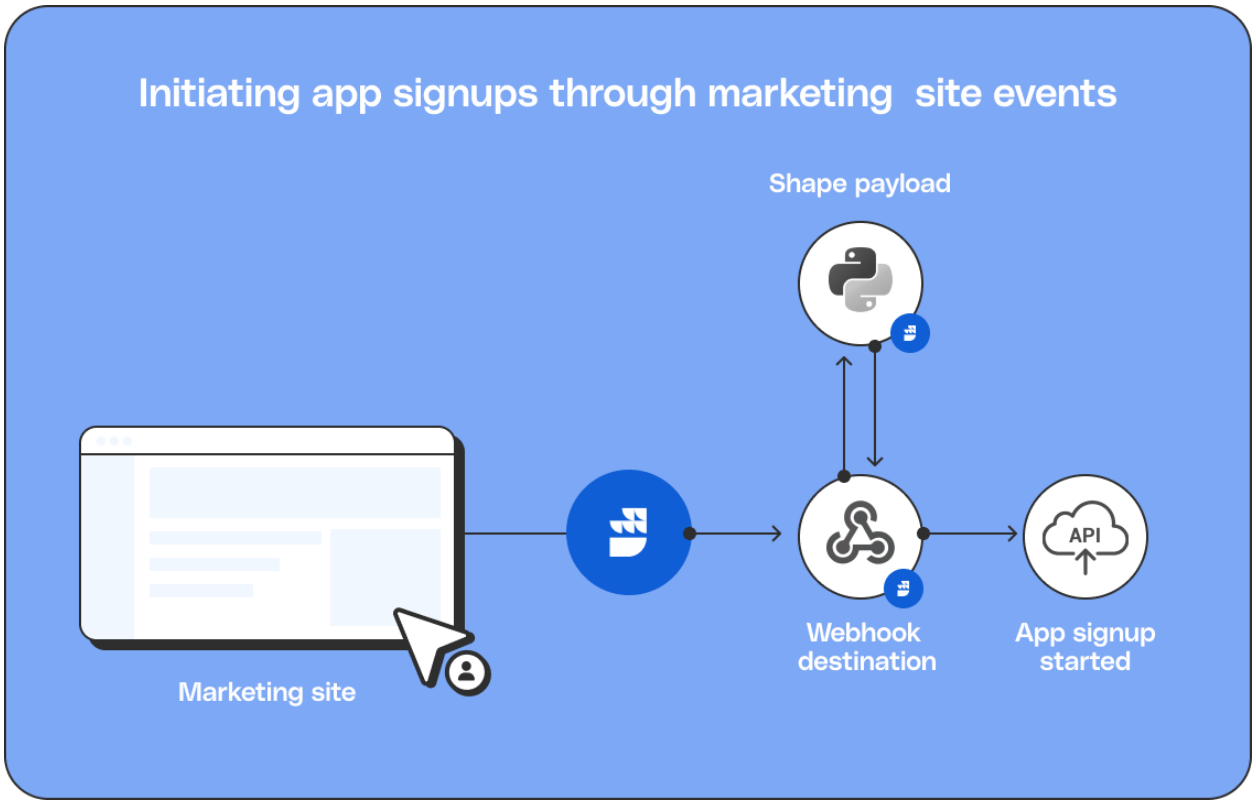
A customer product team was experimenting with various sign-up experiences and learned through A/B testing that there were certain pages on the marketing site and in their documentation that were likely to convert very high.

The problem for the product team was engineering resources: there were higher priorities than building a custom signup flow into the marketing site itself.

Solution

The product team brought this problem to the data engineering team, who had recently implemented RudderStack. The data team also knew that the app exposed an API for account creation.

Their solution was to use RudderStack's Webhook destination combined with a Transformation to initiate the signup process without requiring any work from the engineering team, which was a huge win for everyone.



Here's how it works:

- Payload comes into RudderStack via the marketing site source
- The payload hits a Transformation to shape it correctly for the internal signup API (this team uses Python, but you can also write Transformations in JavaScript)
- The Webhook destination forwards the payload to the signup endpoint
- The signup process is initiated and the user gets an email from the app

---

Ready to make your data stack more efficient and ship more powerful use cases? Reach out to our team to get a demo of RudderStack.

RudderStack is the warehouse-first, customer data platform built for developers. We take a new approach to building and operating your customer data infrastructure, making it easy to collect, unify, transform, and store customer data as well as securely route it to a wide range of marketing, analytics, sales, and product tools. Over 18,000 sites and apps run RudderStack including Crate & Barrel, Acorns, Hinge, Stripe, Allbirds, and more.

www.rudderstack.com

@rudderstack