



Customer Data Platforms: BUILD, BUY... OR BOTH?

Customer Data Platforms: Build, Buy ...or Both?



INTRODUCTION

Build or buy?

01



There's an existential question in IT when it comes to implementing a data platform. How do we want data engineers to spend their time?

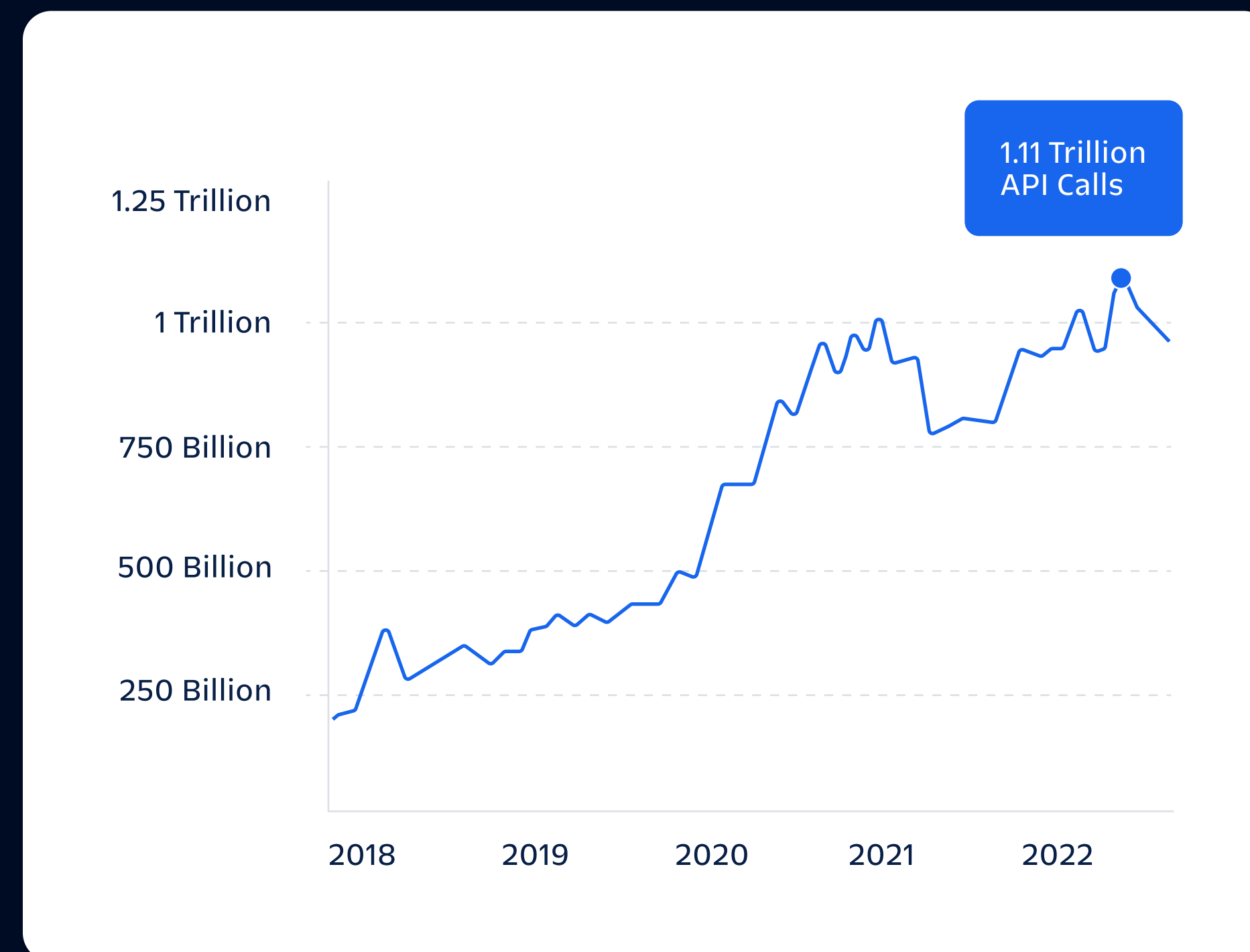
This isn't a cut-and-dried decision to make. Teams often run the gamut of evaluation criteria that includes cost, maintenance, control, and scalability. And while it's important to consider each of these points methodically, businesses don't have the luxury of indecision.

In recent years, there's been a cross-industry consensus that "data is the new capital." It's a prerequisite for personalization, product development, machine learning – the list goes on.

And as the world embraces our new digital-first era – at breakneck speed, in part, due to Covid-19 – the sheer amount of data being created has skyrocketed.

More data should be a boon for businesses, right? Well, sort of.

Today, the average business uses well over 100 software apps, which means data is often trapped in team-specific silos. In fact, a Forrester survey found that fragmentation was the top challenge companies faced when trying to leverage their customer data for sales and marketing efforts.



This August alone Twilio Segment processed over 1.1 trillion API calls, the highest ever-recorded in a single calendar month on our platform.

This has made customer data platforms the new necessity in your tech stack. As businesses make a concerted effort to become more customer-centric, there needs to be a scalable data infrastructure in place that can not only collect data reliably, but stitch together data points into a complete customer profile, and activate this downstream in marketing and sales tools.

This makes the question of whether to build or buy a CDP even more high stakes. But...why is it always framed as an either/or?

At Segment, we've supported businesses that have opted to purchase a CDP and those that were looking for more of a hybrid model (a data platform partially built in-house, and partially outsourced).

The reality is that businesses today have three viable options when putting together a plan of action for their CDP:

- Start from scratch and build a CDP yourself.
- Buy a CDP off the shelf.
- Opt for a hybrid built-and-bought data platform.

In this guide, we consider the nuances of each option and provide a framework for making the right CDP decision for your business.



In a survey of over 4,000 decision-makers, 73% said a CDP will be critical to their customer experience efforts.

[Source](#)

TABLE OF CONTENTS

Introduction: Build or buy?	2
Building a CDP in-house	5
Why did ClearScore replace their homegrown platform with Segment?	9
Deciding what to build, and what to buy	11
Wrapping up	16
Recommended reading	17

Building a CDP in-house

The first option for businesses when it comes to implementing a CDP is to start from scratch and do it themselves.

Perhaps the most notable benefit to this approach is complete control. Data engineers have the flexibility to customize how data is stored, collected, and consolidated into identity graphs. Engineers can prioritize building features that are unique to the business, rather than be at the whims of a vendor's product roadmap.

Because, if data is an asset, then how it's used is a competitive differentiator – which makes a bespoke customer data platform all the more compelling.

But, there's a catch: time-to-value, total cost of ownership (TCO), and capability trade offs. Building and maintaining an in-house data platform is an ongoing project that requires a large investment of time and resources.

Starting out, many companies believe they have the capability inhouse to build an ID graph for identity resolution, or a system that scales and enables real-time data delivery (as just two examples). But this type of work is both complex and costly. While the first iteration of a data platform may seem easy, it's just the beginning of getting to a viable solution for the business – and many soon realize they're out of their depth.

Cost of Delay

Definition

Cost of Delay is the economic value lost by delaying the completion of a project by a set period of time.

Calculation

$(\text{fraction of year delayed}) * (\text{expected annual benefit})$

For example, if the project is expected to provide \$1M dollars in annual benefit, a one quarter delay would be:

$.25 * \$1,000,000 = \$250,000$ per quarter delayed

Calculating the Cost of Delayed IT Projects.
Source

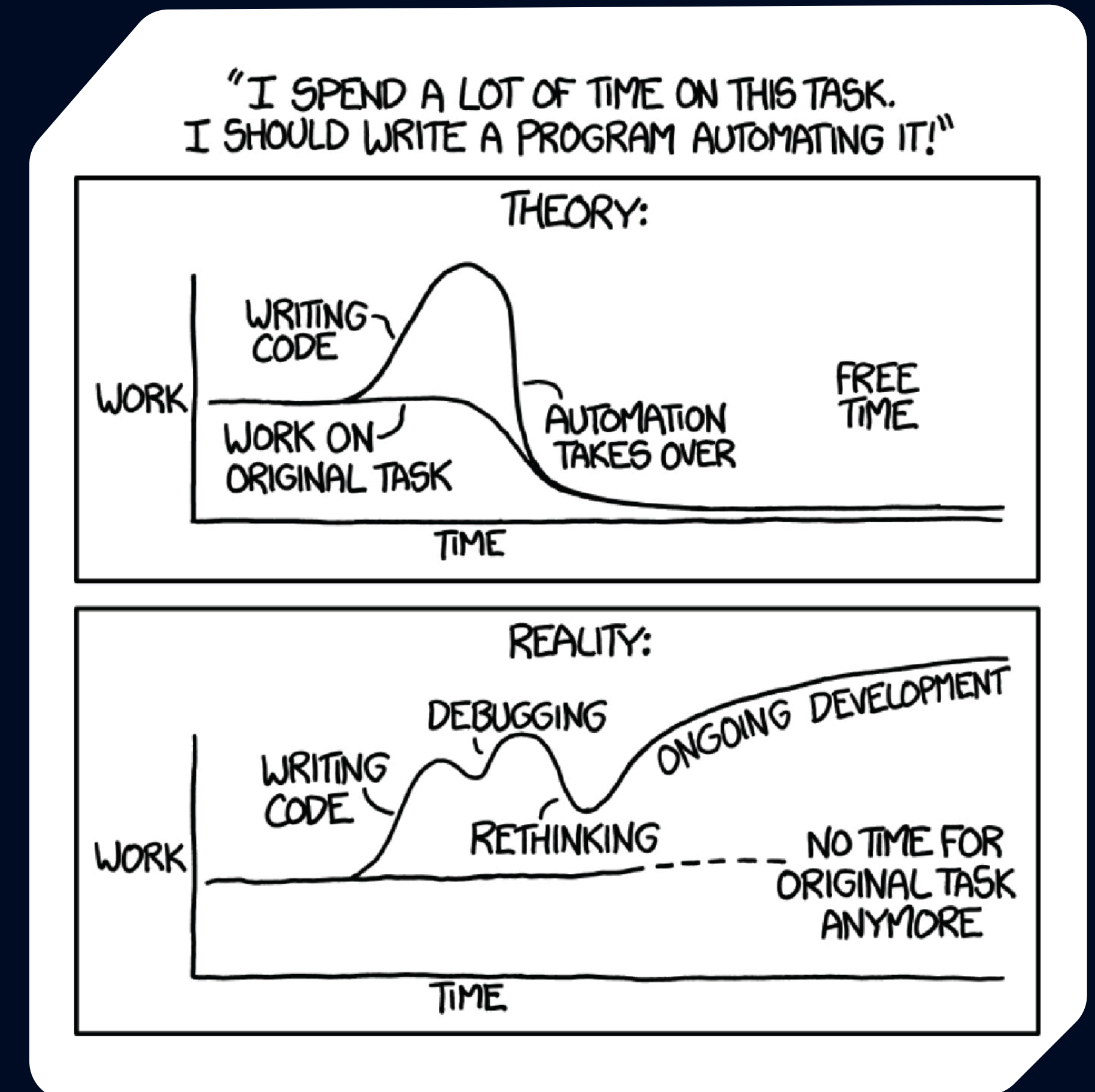
Businesses need to ask themselves:

- Can we afford to wait while engineers ramp up an in-house platform?
- Will ongoing maintenance distract engineers from working on our core differentiators (i.e. what's the opportunity cost? What happens when something breaks?)
- Will building in-house slow down decision-making? (e.g. as systems get larger, and more complex, so do the teams working on them, which can be difficult to coordinate)

Remember that with complete control comes complete responsibility. Maintaining ETL pipelines, setting up new integrations, complying with evolving privacy regulations, operating and hosting costs – this will all fall squarely on the shoulders of your engineering team when building a CDP in-house.

Pat Ottum, an Enterprise Account Executive at Segment notes, “It’s amazing how many companies think they’ve built a data platform,” but then realize key elements are lacking (like the ability to comply with GDPR or handle data deletion requests from users).

But, if your business has the resources in terms of time, budget, and engineering talent, this ongoing maintenance may not be an issue. In that case, the question of, “Can our business wait?” isn’t as applicable as, “Can a vendor match the pace of our innovation?” (To paraphrase David Marble, Product Manager at Facebook and former Group Product Manager at Intuit.)



Calculating the Cost of Delayed IT Projects.
[Source](#)

Buying a CDP

Rather than build in-house, businesses have the option to buy a cloud-based customer data platform from a vendor.

This sidesteps the months-long process of building a platform internally, and offloads the manual work of maintaining ETL pipelines and setting up integrations. This results in more bandwidth for engineers, and faster data democratization (giving non-technical teams easier access to customer data). With low-to-no-code integrations and features, non-technical teams are able to leverage customer insights (often in real-time), without distracting engineers from the product roadmap.

Fast implementation, strategically reallocating engineering resources, a more agile tech stack – these are all huge selling points for buying a CDP.



But there are potential challenges to consider. If you do purchase a CDP from a vendor, be sure to choose an open, flexible platform that can easily coexist with in-house build. There are vendors that will specify the way data is held and maintained in their platform, which makes it difficult to switch to a new vendor should your business ever decide to do so.

Evaluating different CDP vendors?
Check out our [CDP Buyer's Guide](#).

A study published in the Harvard Business Review, which analyzed 1,471 IT projects, found one in six projects had a cost overrun of 200% on average and a schedule overrun of almost 70%.

Source

ClearScore is a leading financial technology business in the UK focused on helping people achieve greater financial well being, by offering free access to credit reports and scores.

A large part of their success is rooted in their data strategy. Their home-grown system extracted event data from a number of different sources and pushed it into a centralized database – allowing their team to analyze and track user behavior and drive rich customer analytics.

So, why did they switch to **Twilio Segment**?

In 2016, ClearScore was ready to start expanding globally, prompting the question: what would it take to maintain and scale our homegrown solution to support new markets?

ClearScore knew that successful global expansion depended on:

- 1. Seamless data segregation.**
- 2. The ability for each country manager to adopt the tools necessary to analyze local data.**
- 3. Company-wide insights delivered in real time.**

ClearScore realized that working with Segment could lessen the burden on engineers during this high-growth stage.

USE CASE

Customer data infrastructure
Data stack modernization
Privacy and compliance

INDUSTRY

Financial Services and Insurance

HEADQUARTERS

London, England

REGION

EMEA

PRODUCTS

Connections

Segment gave ClearScore the ability to quickly set up separate workspaces and flexibly integrate with marketing and analytics tools – which helped fast-track their expansion.

By offering a consistent method of collecting data across back-end systems, websites, and mobile apps, getting data into ClearScore's analytics platform became simpler and more cost-effective.

“Rather than extracting data from multiple platforms, we now take an ‘integrate once, extract to anywhere’ approach. This reduces the maintenance overhead of gathering data and ensures that quality exists further up the data stream.”

Klaus Thorup,
Chief Technology Officer, Clearscore

ClearScore tested their first international market with Twilio Segment in June 2017 when they launched in South Africa. Following the success of this pilot, ClearScore continued on with this global expansion.

- **3x cost savings** by leveraging Segment's platform instead of building in-house.
- **25% of engineering resources** can now be preserved for building and innovating the

ClearScore product instead of the ongoing upkeep required of a home-grown data system.

“We didn't really have real-time data available until Segment. We used to have to wait 24 hours to be able to get our hands on the data and analyze it. Segment surfaces any issues in the customer experience that might hurt monetization and has allowed us to make better business decisions faster.”

Bruce Wood,
Data Principal at ClearScore



Deciding what to build, and what to buy

The pros and cons of building vs. buying a data platform are enough to send any business into analysis paralysis. But, what if this whole conversation was just perpetuating a false dichotomy?

There's been a paradigm shift in how businesses are starting to think about the customer data infrastructure. Why can't you have the benefits of building a data platform and working with a vendor?

Across industries, differentiation is necessary to survival – businesses need to keep building to stay competitive. But that doesn't mean that it's necessary to build everything.

A few years ago, Jeff Bezos offered up a new catchphrase for Amazon's web-scale services: "We build muck so you don't have to."

The "muck" that Bezos described has also been referred to as "undifferentiated heavy lifting" by Amazon's CTO. This is the necessary but arduous IT work that needs to be done, but doesn't actually contribute to the business's competitive differentiators.

Crucially, it takes up a lot of time: Bezos noted that their developers were spending roughly ~70% of their time on this type of work (i.e., server hosting, bandwidth management, scaling and managing physical growth. etc.)

So, Amazon started packaging and selling this undifferentiated heavy lifting (i.e. Amazon S3, EC2, etc.) to offer businesses the, "ability to compete based on the quality of their ideas rather than on their ability to create their own muck."



“Companies like Netflix and the Amazon’s of the world have the privilege of earning and making billions of dollars a year and are able to hire data engineers to focus on core data problems. What we’ve found is that building and maintaining data pipelines isn’t core to 99% of businesses today.”

Artem Ovsyannikov,
Senior Account Executive at Segment

Investing in API-first platforms (like Twilio Segment) allows businesses to build on top of an out-of-the-box infrastructure. It becomes more of a partnership rather than a full on replacement, and allows the vendor to tackle some of the harder infrastructure problems that can become a million-dollar problem (or a ten-million dollar problem) if left unsolved.

This includes handling rising cloud costs as you scale, and guaranteeing at-least-once delivery to ensure there’s never any lost data.

Instead of giving developers a complete solution – which, as Twilio CEO Jeff Lawson notes, “makes a lot of assumptions about what problem your customer needs solved” – they’re given building blocks to more quickly create their own, customized product.

“There’s a Darwinian evolution going on right now, which is if one company in a competitive sector hires software developers, listens to their customers, and starts building a better customer experience, they’re going to win. Once one starts doing it, they all need to start doing it. So it’s not really build versus buy anymore, it’s build versus die.”

Jeff Lawson,
CEO of Twilio,
Source



Intuit's framework for a hybrid data platform

Intuit specializes in financial software, offering platforms and products like TurboTax, QuickBooks, and Mint to approximately 50 million customers worldwide.

A few years back, Intuit was evaluating how to construct their ideal data platform. They had a list of the problems they wanted to solve, both for internal teams and their customers. This ranged from establishing consistent data tracking, to ultimately giving marketing and sales teams access to self-serve analytics, and of course, personalizing the end-to-end customer journey.

David Marble, who was a Group Product Manager at Intuit during this CDP evaluation process, offered insight into their decision-making process. The team knew that they wanted their data platform to be a hybrid of build-and-buy – but separating what should be done internally vs. what should be outsourced was more nuanced.

Intuit started with the five overarching requirements they needed from a CDP.

Collection

“Data in”, APIs/SDKs, batch/streaming support, schema validation, filters, transformation

Profile Enrichment

Options to add dynamic attributes (aggregations, count, “last of”, etc.), add 3rd party data

Distribution / Orchestration

“Data out”, batch/streaming, APIs/SDKs, action orchestration

Identity Graph & Profiles

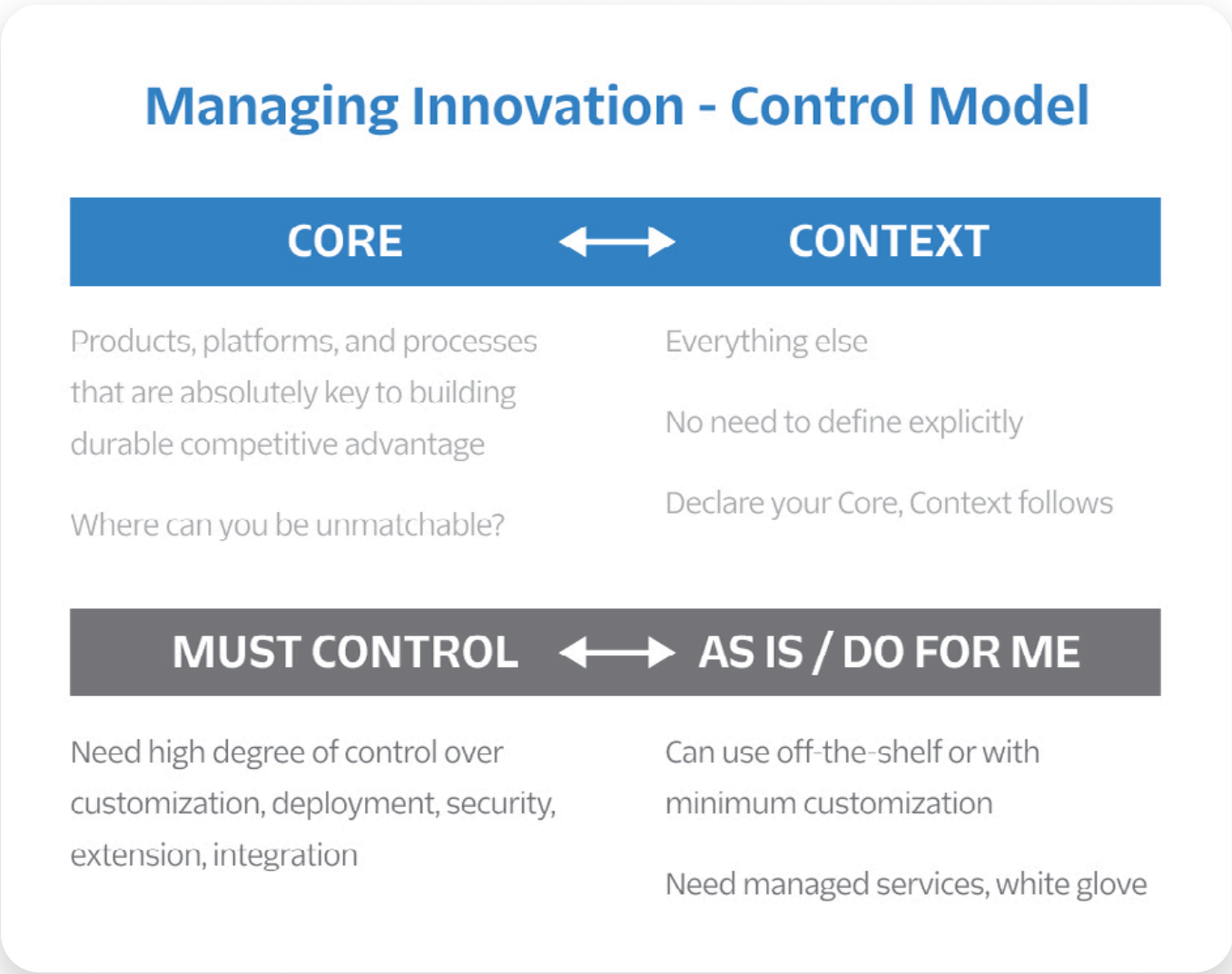
Graph of identities, identifier stitching approach, profile creation with basic attributes

Segmentation / Querying

Options for defining and discovering audiences helpful to engagement, personalization, and analysis

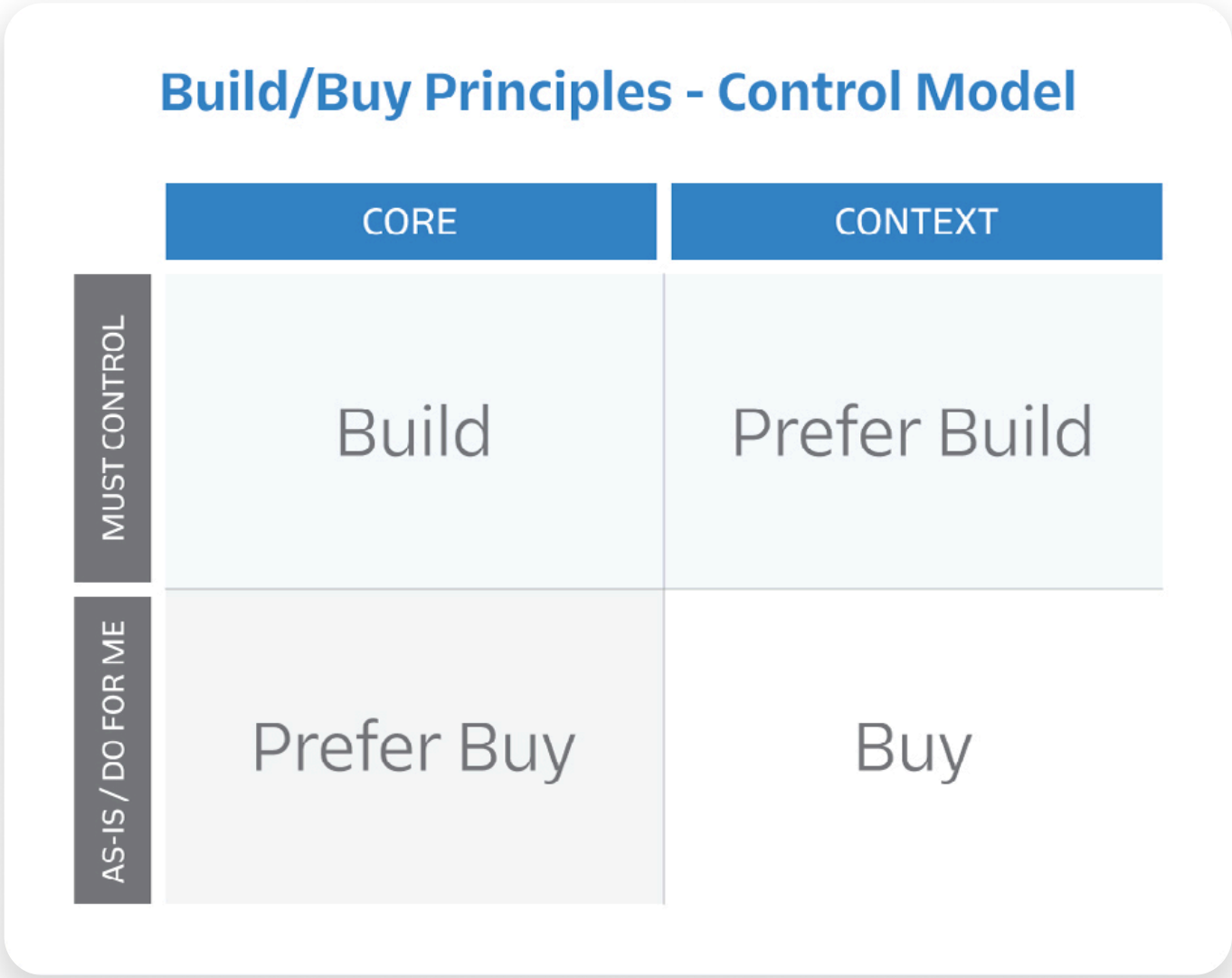
Then, their team borrowed Geoffrey Moore’s model for managing innovation (tweaking the framework slightly). Moore’s model splits criteria between two axes, the first being “Core” vs. “Context.” Whereas “core” refers to the products and platforms that are essential to the business (and serve as a competitive advantage), “context” refers to everything outside that description.

As for the second axis, Intuit drew a line between what they absolutely needed to control internally versus what could be done by a vendor.

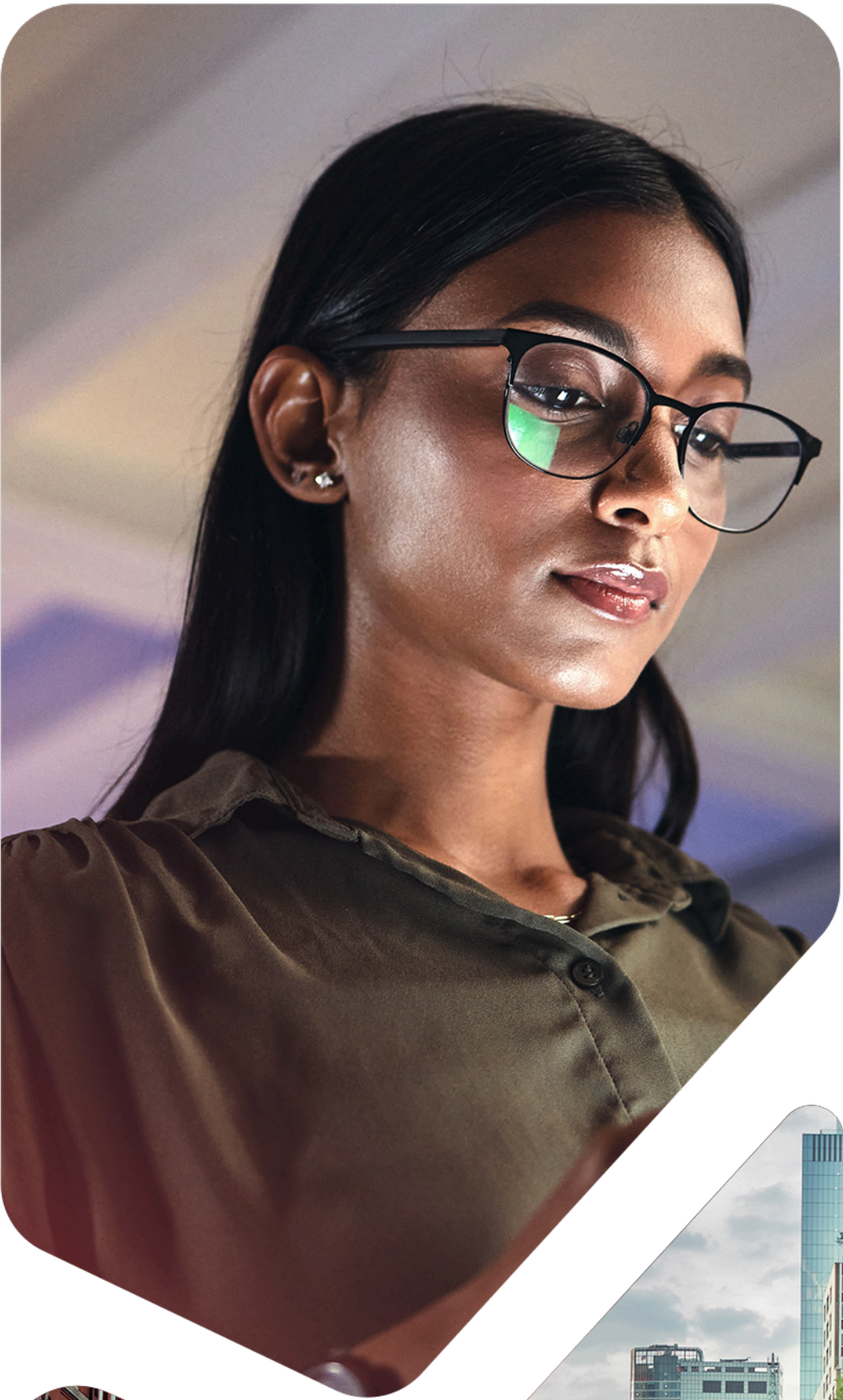


Source

What followed was conversations across the company to determine what Intuit wanted to build or prefer to build, versus buy or prefer to buy.



Source



Ultimately, Intuit settled on buying a CDP to help with data collection, distribution, and orchestration. But when it came to their identity graphs, adding dynamic attributes (e.g. average invoice, “last of,” etc.), and segmentation, that internal financial data was too sensitive – and core to the business – to put into an external platform.

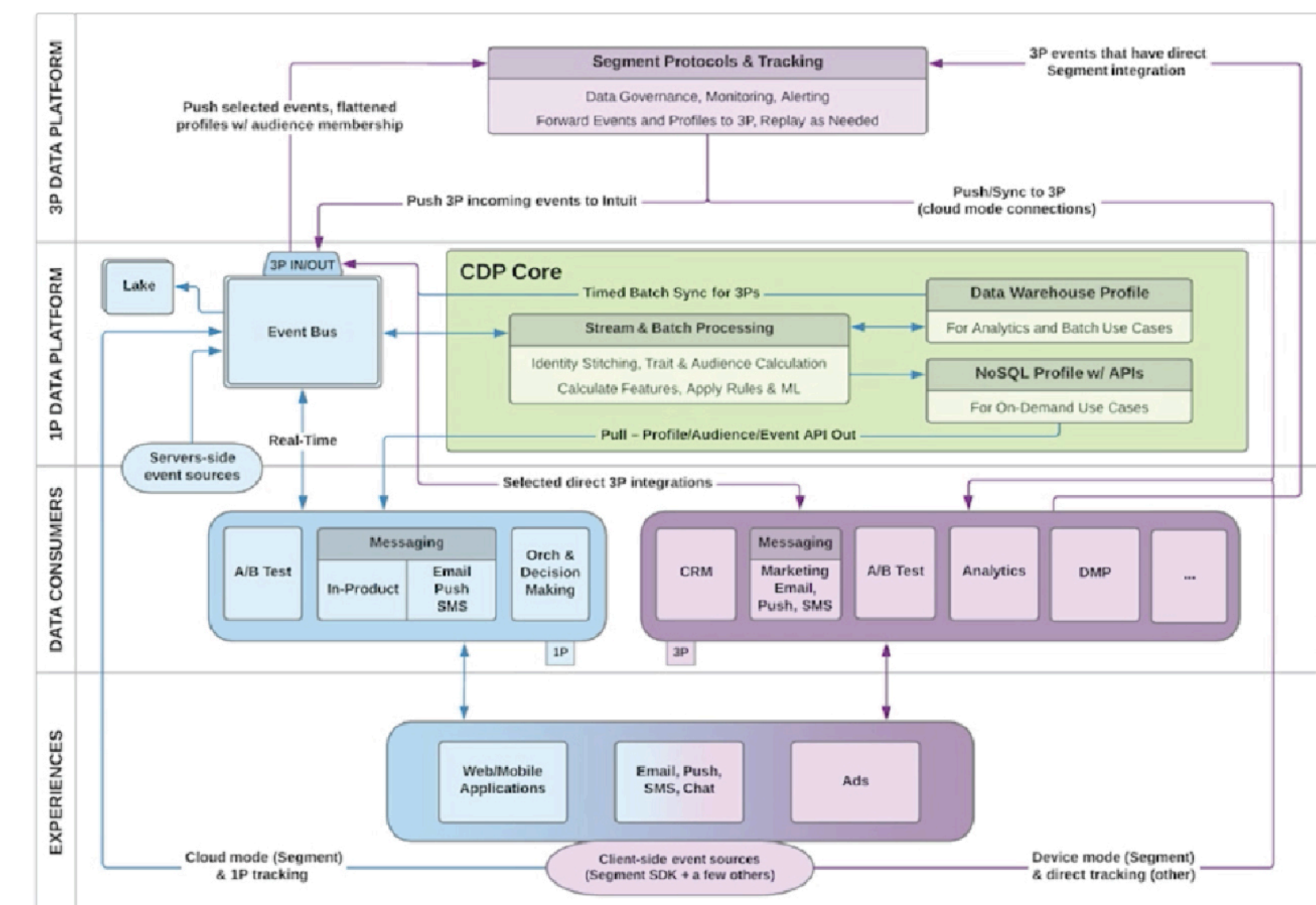
You can watch David Marble’s full Synapse talk [here](#).

Intuit CDP

Build / Built

Build (CDP Core)

Buy / Bought



Here’s the infrastructure they ultimately settled on:
[Source](#)

Wrapping up

To understand what would be most beneficial to build, and what makes sense to buy, consider your starting point, available resources, and both short- and long-term goals. We've compiled a list of the most salient criteria to consider:

- What is “core” to your business? What features, services, assets, etc. are your competitive differentiators that you want complete control over?
- What are you comfortable outsourcing? What work is fundamental to your business, but doesn't require complete oversight?
- Do you have the engineering resources to handle ongoing maintenance and operational costs of building in-house? Do you have the skills internally to reach CDP maturity (machine learning algorithms, ID graphs, etc.)?
- Do you have the budget to handle operational costs?
- Are you in a position to scale successfully?
- Are you ready to handle evolving privacy regulations (especially in different regions, for global companies)?

These types of questions will help you filter out the components of your CDP that absolutely need to be built in-house vs. those that actually make more sense to outsource. Because, there isn't a line in the sand between the two options. To stay competitive, it makes sense to harness the benefits of both.

**Want to learn more
about the right blend of
build-and-buy
for your business?**

[Schedule a demo today.](#)



THANKS FOR READING

If you would like to learn more about what Twilio can do for your business,
please [contact the Segment sales team](#).

