

searce° |  hightouch

Composing a **BigQuery** **CDP** Tailored to Your Organization

Authored by: [Searce](#) and [Hightouch](#)



Introduction

Organizations that want to collect, store, and then distribute customer data to all of their business tools historically relied on all-in-one Customer Data Platforms ([CDPs](#)). Traditional CDPs democratized data to non-technical teams, making it far easier for them to use that data for use cases like personalized lifecycle marketing across channels like email and SMS, and targeted advertising on major ad platforms.

These pre-packaged platforms, such as Segment and MParticle, grew in popularity throughout the 2010s for CMOs that wanted a one-stop-shop for their business customer data. In that same time period, however, data teams increasingly adopted cloud data warehouses, such as Google BigQuery, to store and model all data from across their organizations. Too often, you'll find marketers relying on a CDPs narrow silo of data, separate from the majority of data that the rest of their peers leverage in the warehouse.

Luckily, there's a solution: rather than purchasing a separate CDP, you can create a [Composable CDP](#)—directly on your existing architecture, such as Google BigQuery. This allows you to invest in your data opportunistically, adding what you need for each incremental use case, rather than investing in a massive platform launch all at once. In this article, we'll explain a strategy you can follow to compose your own CDP on your data infrastructure.

The Foundation for Customer Data Platforms: Data Storage

A “data platform” of any sort must include data storage. Traditional CDPs collect data into their own storage from which they sync that data to downstream tools.

Unless you're a brand new startup, you inevitably already have customer data stored somewhere. A composable CDP can adapt to your existing data infrastructure: there are plenty of ways to model your existing customer data and activate it to your tools.

That said, some data storage options are better than others. The major cloud data warehouses offer best-in-class scalability, reliability, and more to support the full range of data you may ultimately need to collect and model in the future.

BigQuery is an extensively managed serverless data warehousing solution, offering users the ability to execute intricate queries and derive valuable data insights, all while eliminating the need for infrastructure management, server configuration, and scalability concerns.

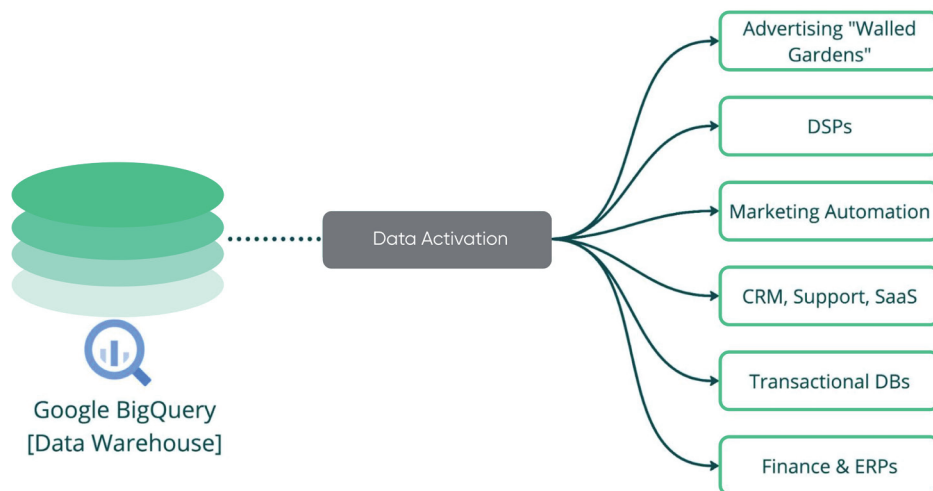
Thanks to BigQuery's powerful and distributed analysis engine, it can swiftly retrieve terabytes of data in mere seconds and process petabytes of data in just minutes. Additionally, BigQuery delivers an unmatched level of flexibility by separating the computing engine responsible for data analysis from your storage choices. With BigQuery, you have the option to either store and analyze your data directly within the platform or leverage BigQuery to assess and analyze data no matter where it's stored.

Put Existing Data to Work for Immediate Value

The whole point of any CDP is to use customer data. Traditional CDPs have prebuilt connections to hundreds of business tools to get data from their storage into the hands of team members throughout your company. To accomplish this using your own data storage, you need a Data Activation product, which allows you to sync data to your tools at scale.

Data activation relies on a technology called Reverse ETL. ETL (extract, transform, load) is a popular process to load data into your data storage; reverse ETL is the opposite, the way to get data out of your data storage. Data activation tools allow you to declare where and how you want your data to be distributed, and they take care of all the underlying intricacies that come with mapping your data to various operational tools. These platforms can be very fast to set up to get ROI (Hightouch boasts that it takes just 23 minutes on average for customers to set up their first data sync) and are far more affordable to implement than a full traditional CDP.

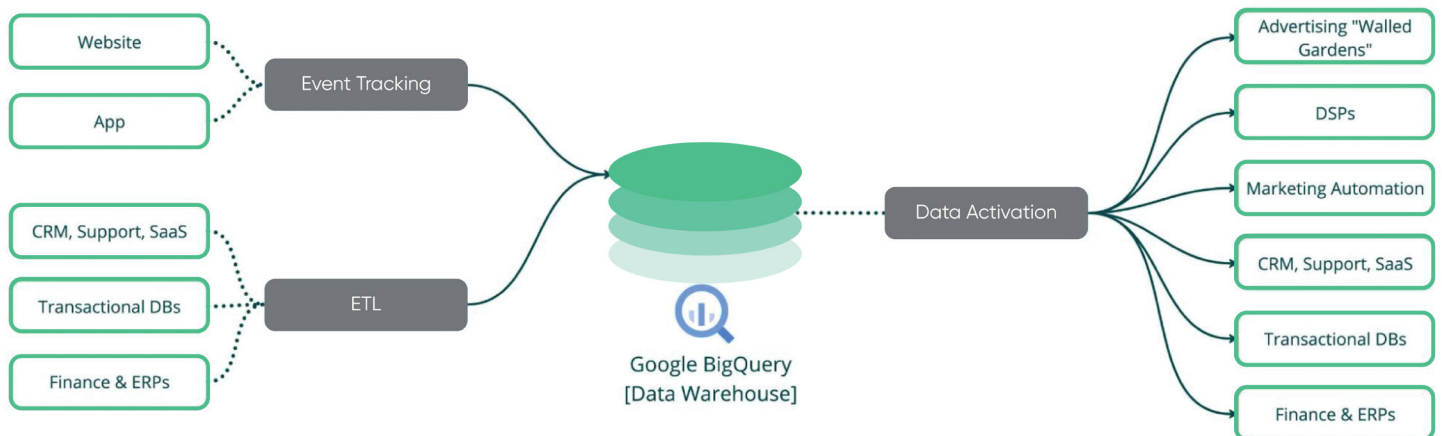
With the combination of your data storage and a Data Activation platform, you already have a lot of power at your disposal to accomplish your CDP use cases. You can sync customer lists, attributes, and actions to your email platform to start to personalize your marketing. You can sync suppression lists to ad platforms and send conversions directly to the Conversion APIs of major walled gardens like Meta and Google Ad Manager to optimize your campaigns. You can trigger automated alerts to company tools like Slack, update CRMs like Salesforce with new lead data, and so much more—you just need to start defining customer data use cases and prioritizing them.



Collecting and Improving Data for New Use Cases

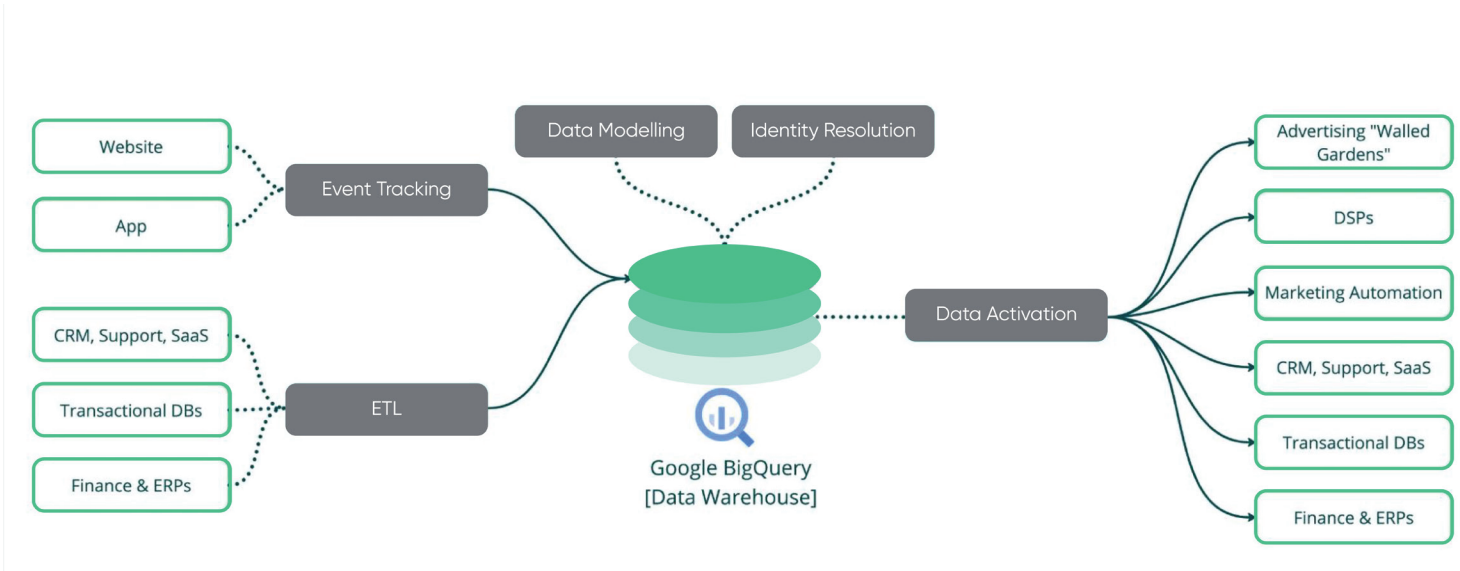
Once you're getting value from your data and closing the loop between business users and the customer information you already have, you'll likely start to think of new use cases you don't yet have the right data for.

One part of this gap may be collecting the right data in the first place. As with anything else in the Composable CDP, you can build a roadmap of your use cases and add new data ingestion opportunistically for whatever will get you the quickest value. Maybe you need to collect new digital events. If so, event tracking products offer SDKs that you can implement on your website or app to gather user interaction data. Pro Tip: You'll want to pick a solution that loads these events directly into your data warehouse rather than just into a traditional CDP's separate data silo, of course. If you need to ingest data from other sources such as Salesforce or your other business tools, you can use popular ETL solutions like Fivetran.



A trickier problem that can block your customer data use cases is data quality. Notably, there's the challenge of identity resolution, which is the process of deduplicating users and unifying all of their activities across devices, browsers, and offline interactions into a 360-degree view of the customer. Traditional CDPs offer identity resolution as part of their bundled solution within their separate data silos, and historically, this was one of their major differentiators and reasons that marketers relied on them. Luckily, now identity resolution can be solved directly in the cloud data warehouse with purpose-built tools or with SQL. This allows you to continue to use your own data infrastructure for your Composable CDP, unifying all of the customer data you've already collected.

If you have a dedicated data team, they'll have other steps to take in a warehouse using tools like dbt to cleanse data, build more robust calculations for analytics, and build predictive models to support your business optimization like expected customer lifetime value (LTV). If you're already activating data from your warehouse, you'll know the ROI that data improvements there can drive, and can prioritize these deeper data investments accordingly.



Getting Started

By prioritizing the use cases a customer data platform can unlock, you can build incrementally on your data architecture to accomplish them. The key to success for any data strategy is to start with the end in mind: what do you want to accomplish with your customer data? If you do this, you can work through your use cases one by one with your existing data infrastructure, rather than starting from scratch with a traditional CDP.

To learn more about tools that support the Composable CDP on Google BigQuery in particular, check out [Hightouch's BigQuery CDP offering](#).

Searce's data and analytics solutions are designed to help you make data-driven decisions, improve operational efficiency, and drive business growth. We offer a suite of services, from data modernization to advanced analytics and predictive modeling, so that you can unlock the hidden potential within your data and turn it into a competitive advantage.

What sets us apart is our commitment to understanding your unique needs and tailoring our solutions to fit your organization's specific goals. As a partner of choice for companies across the globe, we combine the latest technologies, machine learning, and artificial intelligence with expert human insights to provide a holistic approach to data and analytics.