Instacart Taps Confluent to Put (and Keep) Its Data in Motion During Unprecedented Growth

Headquarters
San Francisco, CA

Industry
Grocery

Challenge
Design and maintain a resilient IT infrastructure that can ensure continued seamless grocery delivery during a period of unprecedented growth.

Solution
Confluent Cloud for a real-time data platform that unlocks the full value of streaming data and empowers data visibility, agility, and flexibility across a rapidly growing organization.

Results
- Better inventory management via real-time data
- Reduced TCO
- Improved fraud detection
- Faster execution

Instacart is the leading online grocery platform in North America, partnering with nearly 700 national, regional, and local retailers to offer delivery and pickup services from more than 65,000 stores.

At the onset of the COVID-19 pandemic in early 2020, Instacart experienced its highest customer demand in history. Its customer-facing applications, backend systems, and business processes were all put to the test as millions of people turned to Instacart as an essential service to safely get the groceries and goods they needed.

"We experienced about 10 years of growth in six weeks," said Dusty Pearce, Vice President of Engineering at Instacart. "How do we help these users who've maybe never used an app or done e-commerce before? How do we rapidly change our product to help them get what they need? How do we onboard 500,000 shoppers across the United States in six weeks? All of these challenges wrapped up into one was just a really wild ride."

Instacart required a true cloud-native service that could elastically scale and deliver the resilience needed for an always-on service to support its growing business.

They chose Confluent's platform for data in motion to connect the disparate systems and applications required to update inventory in real time and support Instacart's end-to-end order fulfillment process from the time an item is added to a cart to when it's delivered to a customer.

By leveraging Confluent's fully managed cloud service to support its underlying data infrastructure, Instacart is now able to better allocate valuable technical resources toward innovating on core products and offerings for customers, shoppers, and partners.

> "Ultimately it comes down to support. When I call Confluent, I know I'm already going to be talking to a Kafka expert. They don't have to go find someone. And that matters when minutes or even seconds are on the line for uptime."
>
> **Dusty Pearce**
> Vice President of Engineering, Instacart

Building the infrastructure to support millions of customers

Few companies grow as fast as Instacart did in 2020, going from being available in about 1,500 stores in 30 states to more than 28,000 stores nationwide and in Canada, and serving more than 85% of U.S. households.

With essentially no roadmap for managing such a swift and large transformation, the engineering team at Instacart had to think both fast and in real time to keep their data discoverable and accurate, and to uphold Instacart's mission of making it as easy as possible for American and Canadian families to put food on their table at a time when physically going to the grocery store had, for many, become a scary prospect.

"It was really a nightly exercise of, 'What can we do to create headroom for the wave, the bigger wave of people coming tomorrow?'" Pearce said. "And that was a big part of what my infrastructure team was working on, primarily with data systems. Data systems are the ones that are the most brittle, but at the same time, our product team and our executive team were looking at all of these new users."

Going beyond Kafka for massive scale

Already familiar with Apache Kafka® because he'd used it at previous jobs, Pearce knew Kafka was the technology Instacart needed to adopt to keep their massive inventory system running smoothly during this unprecedented time.

"I'm very familiar with (Kafka) and understand the platform, and I knew that it was a platform I could trust," Pearce said. "However, I didn't have the time that I had before to build a team and build this from the ground up. I had to jump in and just start leveraging streaming systems almost immediately."

Instacart had begun to use Kafka for a few initial use cases, including Customer 360, relying on their cloud provider for Kafka-related managed services. "It worked to some degree, but we had struggles with support," Pearce said. "We also lacked the internal expertise, so we really needed to find a partner who could help us push this platform forward at the speed that we needed to move without me waiting a year or two years to assemble all of the staff in a really competitive market. It just felt like it was the right time."

Pearce and his team chose Confluent Cloud and immediately began to weave Confluent into its IT infrastructure to support key inventory management and other use cases.

"The famous line I've always used is when you call your service provider, you don't want the support desk to go find their Kafka expert," Pearce said. "So Confluent was a natural choice for our Kafka infrastructure."

The results:
The right product at the right time for every customer

Better inventory management via real-time data. "Confluent augments the experience in a way that allows us to use data to more accurately track inventory," Pearce said. "By using data in motion and real-time data, the likelihood of you getting not only the apple that you asked for, but a really good apple, goes up tremendously."

Reduced TCO. "It's funny to say this because we're a big engineering team, but we're always under-resourced," said Nate Kupp, Director of Engineering, Data at Instacart. "For me to go hire a bunch of engineers to babysit Kafka, I don't have the ability to go do that. And being able to offload those concerns is just such a relief for us and lets us focus on delivering value to the organization and not worrying about ops and other overhead."

Improved fraud detection. "When I think of our first few wins that we've got with Confluent, the one that stands out to me is fraud," Kupp said. "We plugged in, spun up Kafka and Confluent, spun up the team around this, and in a couple of months we were able to put a system in production that's solving this for the team. That was awesome."

Faster execution. "Confluent has really been an accelerant to the work that my team does," Kupp said. "I talked earlier about taking the ops burden off my team, but I think actually, Confluent really unlocks a faster execution for my team on a number of dimensions. Just things like spinning up Kafka clusters and getting prototypes up and running very quickly, Confluent has been really helpful there. Those sort of exercises might take a long time for my team to do if we were doing this on vanilla open source Kafka. With Confluent, we can turn that around very quickly."

What's next for Instacart's data infrastructure?

For the future use of Confluent, Pearce and Kupp said they plan to focus on making Confluent the backbone of Instacart's data architecture to further enable the current use cases and also unlock the full value of data in motion across the organization.

"We've been thinking a lot about things like fraud, where we want to be able to have real-time fraud detection and be able to build that into the product," Kupp said. "Or our availability algorithms—how do we understand if a particular item is available at a particular store? Those sorts of things. There's so many opportunities like that across the company, and I think a lot of it comes with moving to streaming from these batch systems that we had before."

Ultimately, Kupp said, the vision for Confluent is to use it as a data mesh and central nervous system for all of Instacart. "As we look across our distributed organization, there are data ownership concerns associated with our data mesh strategy and the Confluent features that enable and empower all of our organizations to own their own data," Pearce said. "And then there are specific ways that we can incorporate real-time data into software development. And then there are other parts of our business, like our ads business, that want to move quickly, want to understand user behavior, and try to inspire users with content they value."

For Pearce, unlocking the full value of Confluent within Instacart really comes down to tapping into the full potential of Instacart's talented engineers to improve the customer experience. "You will start to see more and more software development teams thinking about reasoning and how we can leverage real-time data to build features and to return value back to the customer in ways that I think weren't possible in the older batch data model," he said. "I think Confluent will unlock the creativity of a lot of really smart people who really understand their customers, and that can't be anything but good."

Ultimately, Pearce said, Confluent will become the backbone for keeping data in motion through Instacart's continued growth. "It was really about how do we partner with someone who knows a lot about this space, has a lot of exposure to keeping data in motion, and understands the data mesh strategy we've been working on," he said. "Now we have to operate and run the software and company at this much bigger scale; Confluent has been and will be critical to us succeeding as this much larger enterprise."

Learn More About Instacart

https://instacart.com