

Complex legacy infrastructure and multicloud data warehouse unwieldy to manage

InMobi is all about targeted advertising and helping brands reach and engage consumers in a meaningful and cost-effective way. But to accurately deliver relevant ads, you need data — a lot of it. Over time, they extended their on-premises Hadoop system by adding multiple cloud data warehouses to address various problems. However, as the amount of data the company needed to process increased exponentially (20TB of data per hour), InMobi continued to build on top of their legacy system, resulting in a multicloud data warehouse that had a host of challenges: It was overly complex, had outages, was extremely costly as they scaled, and led to data silos that limited data sharing and collaboration. The team at InMobi realized that if they continued with this system in place, it would slow down their ability to innovate and keep precious engineering resources locked into maintenance mode.

“The data infrastructure that we built worked, but it created significant complexity and overhead that pulled our focus away from our own core products,” said Madhan Sundaram, Senior Director of Platform Engineering at InMobi. “We needed our talented engineers to create more value for our customers, and to do that, we needed a simpler and more unified system.”

What InMobi wanted was a single system that could solve multiple problems. To accomplish this goal, the company needed to consolidate their disjointed systems into a single platform to free up engineers to focus on higher-value tasks such as developing ML and large language models. That’s why the team looked to the Databricks Data Intelligence Platform to unify their data warehousing and AI workloads on a single platform.

Migration to lakehouse results in unified data, analytics and AI

Although InMobi has a team of fully capable engineers, they learned a valuable lesson in productivity and operational agility. “We found that we were spending more time maintaining our environment, which was hurting our ability to support and collaborate with the business,” explained Sundaram. “We wanted to be more strategic and identify ways to operationalize data more efficiently.” After careful evaluation of their current multicloud data warehouse and the prospects of building in-house, they determined that the Databricks Data Intelligence Platform clearly aligned best with their goals — to improve developer productivity through reduced infrastructure complexity while achieving the best possible price/performance.

Once they made their selection, the team partnered with Databricks to start the migration planning process. With over a decade of customizations built on top of existing systems and over 1 petabyte of data, InMobi knew the migration was going to be complex. On the ETL side alone, there were 150 pipelines and eight teams involved with migrating from open source Apache Spark™. They also needed to migrate approximately 300 reporting dashboards to ensure there was no disruption of information delivery to suppliers and customers. To help navigate this process, Databricks worked closely with InMobi — which had allocated two in-house engineers per team to provide support — and Databricks’ implementation partner Celebal Technologies on the optimizations needed for a seamless migration.

As a result, Databricks was able to help InMobi migrate from their multicloud data warehouse to the lakehouse with ease. “Databricks gives us the edge in terms of being able to utilize industry-first features that help us establish technical differentiation,” said Sundaram. “Their expertise came through during the migration as they helped us optimize our compute resources for cost and performance. The ownership that Databricks took to drive the optimizations and how transparent and educational they were was commendable.”