Headquartered in Toronto, Canada, with more than 80,000 employees in offices around the world, the Toronto-Dominion Bank and its subsidiaries are collectively known as TD Bank Group (TD). TD offers a full range of financial products and services to more than 25 million customers worldwide through three key business lines: Canadian Retail, including TD Canada Trust; U.S. Retail, including TD Bank, America’s Most Convenient Bank; and Wholesale Banking, including TD Securities. TD Bank Group is one of the 10 largest banks in the U.S., with deep roots in the community dating back more than 150 years.

Creating a cloud-based infrastructure to understand customer behavior

TD has historically been a highly federated company, making the collection of data for meaningful 360-degree customer analytics and regulatory compliance difficult and costly.

Although TD Bank is 60 years old, it’s determined to remain innovative and be a leader in understanding its customers, anticipating their needs, and guiding them through their financial journey. To do that, the bank has developed the capability to collect vast amounts of data about customer transactions and the customer experience across channels.

“We seek feedback through multiple channels to understand customer needs and then use these insights to drive performance improvements,” says Joseph DosSantos VP, Enterprise Information Management, TD Bank Group.

The infrastructure the bank has built to deliver exceptional experiences includes a cloud data lake with data on customer behavior and interests, internal and external data, and structured and unstructured data.

DosSantos explains: “We went across the organization to gather operational data about customer interactions, transactions and feedback. We’re also mapping non-banking data, such as data that shows how consumers are interacting and how they’re using their phones or how they are communicating on social media. Finally, we’re also starting to incorporate vendor technologies that can capture audio and video recordings of customer interactions and complete experience. All this is then pieced together to produce a journey map.”

By applying complex analytics to that data, TD Bank is able to create unique, tailor-made experiences for each customer.

DosSantos says a cloud infrastructure enables the bank to give a wide range of people on-demand access to the data and tools they need—350 registered users access the platform for predictive modeling and analytics, and 1,000 more get reports and data directly from the bank’s Hadoop distribution. “That wouldn’t be possible with conventional on-prem platforms,” he says.

Why Talend?

Prior to the beginning of its big data journey, TD Bank had about 16 petabytes of data; now, their big data footprint alone is 3 petabytes. According to DosSantos, Talend has greatly reduced the cost of analyzing and delivering actionable insights from that data. “If you think about the need for data transformations conventionally delivered in a lake, they require massive amounts of coding, done by expensive MapReduce coders,” he says. “The code they develop is difficult to change.

INDUSTRY
- Financial Services

INFORMATION
- HQ: Canada
- 10,001+ employees

USE CASE
- Customer experience

CHALLENGE
- Creating a cloud-based infrastructure to understand customer behavior

TALEND PRODUCT USED
- Talend Big Data

PARTNER
- Cloudera

RESULTS
- Petabytes of data
- Cost of operation 50x cheaper per gigabyte
- $13 million cost avoidance

Whether our customers are visiting a branch or using a tablet, phone, computer or wearable device, our big data cloud-based infrastructure enables us to predict their needs and guide them through their financial journey.

Joseph DosSantos, VP, Enterprise Information Management, TD Bank Group.
By using Talend, we can have that coding done much more simply by staff members with conventional ETL skills, and we fundamentally abstract our logic of transformation. That delivers $13 million in cost avoidance and the code is portable, so we can deploy it across different platforms and different clouds.

DosSantos adds TD Bank has a multi-cloud strategy. “We're predominantly a private-cloud enterprise,” he says, “and banks in North America tend to be more risk-averse to the cloud journey. So we're taking a step-by-step approach, starting with the private cloud, then progressing to working with different public cloud providers.”

For Hadoop distribution, TD Bank uses Cloudera, and DosSantos says that “the partnership between Talend and Cloudera has been excellent. The IT benefits we receive from Talend and Hadoop are significant: The cost of operation for our data is on the order of 50 times cheaper per gigabyte, so we're seeing massive savings from the Hadoop architecture. For reporting, we have a BI in Hadoop strategy. We think that's revolutionary, and really important for cost-effectiveness. Talend is the software that enables the transformations in that system.”

Building customer-centricity across the organization

Banks need to work with the customer for building a sustained relationship based on providing education, personal guidance, and the right products and services through the right channels.

“We want to use our data and analytics to understand what our customers are trying to do and help them get there,” says DosSantos. “Our analytics have helped us focus on what savings or credit products our customers need to finance a child’s education, buy a first home, or reach another financial goal.”

The bank has also used the data lake to create innovative digital products like the TD MySpend app that helps customers track their monthly spending and improve their spending habits. “Connecting the data lake with the app enables us to answer questions ‘right now’ for customers, and that requires on-demand data, elasticity on the part of the platform, and flexibility from IT,” says DosSantos. “There’s no other way to do it than via the cloud.”

DosSantos says that deploying a cloud-based, big data infrastructure has given TD Bank a competitive edge. “Whereas it used to take months for us to develop a conformed data model in a warehouse, we’re now able to do lots of innovative things in the lake,” he says. “Our speed of delivery has increased dramatically. We think of it as data agility. The data access we have now enables us to develop and deliver innovative services for our customers quickly and cost-effectively.”