Safeguarding data investments during a seamless transition from Hadoop to Google Cloud

Company overview

Scotiabank, headquartered in Toronto, Ontario, is one of North America's largest financial institutions, with approximately $1.4 trillion in assets, while serving over 10 million customers in Canada, and employing over 90,000 colleagues globally. Scotiabank continues to grow and includes a significant presence in the U.S.

Scotiabank is focused on

Modernizing the Hadoop estate

Adopting a Hybrid Data Infrastructure during cloud transition

Long term cloud shift

Objectives

By migrating to Google Cloud, Scotiabank aims to unlock next–generation analytics capabilities while maintaining the integrity and performance of their existing Hadoop infrastructure during the transition.

Challenge(s):

The bank couldn't take a big bang approach to their cloud aspirations, as analysis of every application, Spark workload and map reduce job was required to determine priorities and to evaluate whether any remediation was required when running on a cloud infrastructure. A phased approach was required to systematically move the artifacts into Google. The bank had initially looked at the DistCp utility to assist them in moving data sets into the cloud, however as DistCp runs as a standard MapReduce job competing for resources

with other important workloads, it was felt that this additional overhead would not be acceptable in an infrastructure already struggling with compute capacity. The DistCp solution also required multiple executions to capture ongoing data changes which again added to the resourcing overload.

Solution(s):

Cirata Data migrator was chosen as the primary transfer tool for the following reasons:

- Petabyte scale data movement — The bank has 15PB of data to contend with and Data migrator was chosen as the only tool that can manage this volume of data.
- Data Migrator's ability to automate data migration to Google with no business disruption.
- No manual or additional coding requirements for the data migration.
- Complete flexibility in the direction of data movement and the data source and targets.

Results

Data Migrator enabled Scotiabank to:

- Reduce the overall time needed to perform Data migration (estimated anywhere from 12 to 16 months time savings).
- Reduce the number of resources that will be needed to perform the current and future migrations.
- Reduce costs and risks associated with the migration project.
- Faster time to value by providing Google analytics with rich data sources.
- Mitigate compliance risk within a regulatory banking environment.