

INCREASING BUSINESS AGILITY TENFOLD WITH CLOUD-BASED DATA ARCHITECTURES

20X

Reduced
infrastructure cost

Impact

- Consistent overview of data across regions, systems and infrastructures
- Reduced infrastructure cost by 20x
- 10x faster time to market for new or updated applications
- Ephemeral capacities offered by Cloudbreak
- Ability to create on-demand big data clusters, respond faster to new business requirements

Santander Group is a leading retail and commercial bank founded and based in Spain. Ranked as the 1st bank in the euro zone by market capitalization, the organization runs a variety of diversified businesses around the world, with a clear focus on digital transformation. With an array of new digital solutions, Santander Group is setting the benchmark for technology innovation for the best customer experience.

The banking industry has recently undergone a wave of disruption, with the emergence of new technologies, banks and players creating an environment for innovation and change. As services are gradually transitioning to digital platforms, for the fastest and most accurate decision-making, the industry has realized the potential of analytics to remain competitive and generate revenue.

With data seen as the main currency of this digital transformation, SCIB focused on big data. The goal was to be able to extract value from a variety of datasets to create new products, services or even business models that bring value to the end-user and address compliancy requirements.

Gearing up with big data analytics

Working with 142 million clients globally and processing a large number of transactions for big clients, Santander Group is responsible for managing a wide variety of datasets across multiple locations, systems and infrastructures. However, with a focus on digital innovation that has paved the way for blockchain, digital payments and artificial intelligence applications, the bank was experiencing a sharp increase in the amount of data being generated by those services all subject to local data and compliance legislations.

Although traditional data-marts and data warehouses have been playing a strategic role in the analysis of data, the sheer amount of information generated by the Group could no longer be processed reliably and within reasonable timeframes. With data growth showing no signs of abating, the bank was on the lookout for a scalable big data infrastructure able to fuel an increasing number of business applications across on premise and cloud-based environments.

“Aside from our legacy architectures, our organization was facing a number of challenges as a result of extensive financial services operations” said Javier Nieto Centeno, Architecture & Innovation IT Expert, Santander Group. “With many applications running across our Corporate and Investment Banking divisions all subject to stringent financial regulation and auditors, we needed to rely on a system that would allow us to consolidate our datasets and interrogate them on demand.”

“Cloudera professional services has been instrumental as we began our cloud journey. By setting up brand new interfaces and extending already existing ones, the team has been able to address all the various requirements needed for the data lake to run on our IaaS.”

Javier Nieto Centeno, Architecture & Innovation IT Expert, Santander Group

Matching the pace of disruption with data analysis at scale

In order to obtain visibility and provide clear consistency across the Group's data landscapes, the institution partnered with Cloudera to lay the foundation of a big data architecture. With the initial use case geared towards uncovering new business opportunities, the bank introduced a variety of open source tools – such as Flink, Hazelcast, Kafka, Drools, Spark and NiFi – for daily and intra-daily data processing.

“When working with large enterprises on valuable contracts, often the key differentiator is how smoothly you are able to integrate your systems with the client's systems that are already in place” continued Nieto Centeno. “Various kinds of third-party systems are installed in many countries, coexisting with our centralized system, transmitting information both ways. With increasing pressure from a regulatory and compliance standpoint, standardizing our information in a data lake was vital to satisfy auditors and streamline data access for many of our divisions.”

Its first approach, on SCIB, into big data was operating with a number of on-premise clusters, where data storage and processing were interdependent. However, with the adoption of Object Storage Systems, the Group realized that ephemeral clusters could provide the flexibility necessary to better handle certain workloads and address a variety of business cases, which very often only required the analysis of vast amounts of data across a relatively short period of time.

As a result, the Group deployed Cloudbreak, with a view of simplifying cluster provisioning whilst optimizing cloud resource usage by seamlessly adjusting the cluster as workload and activity changes.

Harnessing global big data with the creation of an on-demand Data Lake

Santander is now able to create on-demand big data clusters to respond faster to new business requirements, cheaply and efficiently. Specifically, moving away from File Based towards Object Storage Systems has allowed the Group to lower the cost of storage. This new cloud-based cluster architecture pattern has also enabled Santander to reduce time-to-market of new applications as much as ten times, meaning that the bank is able to deploy new business applications – or newer versions of existing ones – faster.

Nieto Centeno stated that “breaking away from HDFS has enabled us not only to speed up application delivery, but also provide the ideal compliance framework for regulators to access historical data that would have otherwise been harder to access. This is a fundamental change within the banking industry and, as datasets keep on growing, we are confident that our newly established cloud infrastructure will enable us to fend against the strictest European compliance challenges.”

Since the deployment of Cloudbreak, the Group has been able to lower the costs of its hybrid big data infrastructure by using cloud resources as and when needed. This meant that once the project was over, the instance could be closed down to lower operational expenditure and avoid peaks. From the selection of a specific blueprint – whether custom-drawn or pre-selected within Cloudbreak – clusters can now be deployed in a few easy steps to run on-demand analytics, reporting or other mission critical functionalities that the bank requires at any given time.

Banking on cloud technologies for the years to come

Building a common architecture for all these systems was a large task for the bank, but now Santander has a common repository where all data can be stored and consolidated. The Group intends to build on this success by continuing to deploy cloud-native business applications to run on top of its Big Data infrastructure, as well as move its legacy applications to the cloud.

“Cloudera professional services has been instrumental as we began our cloud journey. By setting up brand new interfaces and extending already existing ones, the team has been able to address all the various requirements needed for the data lake to run on our IaaS” commented Nieto. “While there is still some work to be done, our organization has laid out the foundation of a long-term data strategy and is determined to make significant progress towards meeting our data-driven goals.”