

AWS ProServe Hadoop Cloud Migration for Property and Casualty Insurance Leader

soft**serve**

Client Background

ur client is a leader in property and casualty insurance, group benefits and mutual funds. With more than 200 years of expertise, the company is widely recognized for its service excellence, sustainability practices, trust and integrity.

Business Challenge

ur client's existing analytics processes were based on an on-prem Teradata Hadoop Appliance that had an expiring support contract. The struggle with free space at the data storage infrastructure was putting business processes associated with Hadoop at risk.

The environment then was at >80% capacity with no room to scale, and our client needed to quickly move its multi-user analytics infrastructure out from on-prem Hadoop and build an AWS cloud-native analytics data platform. The business goal was to decommission the platform and migrate 400 TB of data along with analytical processes to the cloud, while minimizing costs associated with data storage and third party support, and enabling unlimited infrastructure scalability.

Project Description

ur client began its cloud migration to AWS with another vendor. However, due to this vendor's lack of AWS experience, our client struggled with infrastructure setup issues, crossed planned deadlines, and thus had fallen far behind schedule.

AWS reached out SoftServe to step in to the project as an AWS ProServe to get the migration project back on track, validate the target AWS architecture provided by the previous vendor, and help with issues resolution. SoftServe determined that one of the major technical requirements for the AWS-based solution was mimicking the existing Hortonworks Data Platform based Teradata Hadoop Appliance setup in terms of security, data analysis, and data exploration tools.

The cluster needed to be secured and auditable, support Kerberos/ActiveDirectory integration with SSO for end user interfaces, and support Hive/HDFS/HBase/Yarn authorization via Apache Ranger, as well as fine-grained authorization for EMRFS. End users such as data scientists, data analyst, developers, and power users had to be able to use toolsets like Spark, Hive on LLAP, HBase/Phoenix, Sqoop, Autosys, Atlas, and Ranger on Amazon EMR.

Customer Challenge

he data platform architecture, designed by the client, was based on a combination of customized Amazon EMR persistent and transient clusters with shared data storage on EMRFS, as shown on Figure 1.

Our client faced deployment and configuration challenges while integrating Ranger, Hive on LLAP, Atlas, and edge nodes with Amazon EMR and on-prem security systems, including EMRFS authorization configuration. This kind of integration is not complicated, however it requires deep domain expertise in big data, cloud, and security. These integration challenges, in conjunction with strict timelines, would jeopardize project success, and thus our client made the decision to involve the AWS ProServ team.



Partner Solution

oftServe experts were involved to accelerate the implementation of a new Hadoop/ EMR cluster. Because the target architecture had not yet been proved and implementation was in an early stage, SoftServe experts recommended two scenarios:

Primary Scenario: Re-platforming - Migration from Hortonworks/HDP to AWS EMR, (see Figure 1) with further optimization of analytical workloads

Alternative Scenario: Two-phased approach: Rehosting (Phase 1)—migration from HDP onprem to HDP on AWS (Figure 2) in short terms and re-architecture data platform (Phase 2), leveraging AWS cloud-native services. The alternative scenario implied immediate TCO reduction in Phase 1 by migrating the Hortonworks Data Platform from on-prem to Amazon EC2 instances using the CloudBreak solution and leveraging existing in-house HDP skills and a sped-up configuration of Ambarimanaged Hadoop services, while migrating data from HDFS to guarded S3 buckets. To achieve maximum TCO optimization (33-52%) with cloud-native services after Phase 2 it was recommended to migrate from Hadoop of EC2 instances to Amazon EMR and leverage Presto on Amazon EMR for interactive ad-hoc analytics instead of Hive on LLAP.



n order to meet deadlines associated with expiration of a Hortonworks support contract, the decision was made to go with Primary Scenario.

SoftServe customized a default Amazon EMR setup, integrated native and third party components with security and audit, designed the data transfer process from the on-prem data platform to the AWS-based platform. As an AWS ProServe, SoftServe provided consulting services and guided the clients team and performed the following major technical tasks:

- Apache Ranger set up and fine-grain authorization for Hive/HBase/Yarn/Kafka configured
- EMRFS authorization based on IAM roles to ActiveDirectory groups mapping configured
- Hive on LLAP for ad-hoc gueries performance set up
- Apache Atlas for data governance set up and integrated with data processing components
- All default and custom Amazon EMR components integrated with our client's security systems

Value Delivered

oftServe brought the client's migration project back on-track providing a bespoke Hadoop EMR cluster customized and secured according to the client's requirements in a timely manner. Customerspecific performance requirements were met after fine-tuning of the Hadoop cluster.



case study | AWS ProServe Hadoop Cloud Migration for Property and Casualty Insurance Leader

ABOUT US

SoftServe is a digital authority that advises and provides at the cutting-edge of technology. We reveal, transform, accelerate, and optimize the way enterprises and software companies do business. With expertise across healthcare, retail, energy, financial services, and more, we implement end-to-end solutions to deliver the innovation, quality, and speed that our clients' users expect.

SoftServe delivers open innovation, from generating compelling new ideas, to developing and implementing transformational products and services.

Our work and client experience is built on a foundation of empathetic, human-focused experience design that ensures continuity from concept to release.

We empower enterprises and software companies to (re)identify differentiation, accelerate solution development, and vigorously compete in today's digital economy. No matter where you are in your journey.

Visit our **website**, **blog**, **Facebook**, **Twitter**, and **LinkedIn** pages.

NORTH AMERICAN HQ

201 W 5th Street, Suite 1550 Austin, TX 75703 +1 866 687 358

1 University Avenue Suite 11-112 Toronto, ON M5J 2P1 +1 647 948 7638

EUROPEAN HQ

14 New Street London EC2M 4HE +44 (0) 800 302 9436

info@softserveinc.com www.softserveinc.com

soft**serve**