

nClouds | AWS Case Studies

An AI Innovator

nClouds helps an AI innovator achieve **72% AWS cost savings** by replacing Cloudera with Amazon S3, AWS Glue, and Amazon Kinesis.

Industry

AI, Computer Software, Media

Location

United States

Challenge

Optimize data analytics costs without sacrificing performance.

Featured Services

Amazon Athena, Amazon Simple Storage Service (Amazon S3), AWS Glue, Amazon Kinesis, nOps

About AI Innovator

The company is a fast-growth provider of an AI-powered software platform in the media space. It was founded in 2015 and has offices across the US.

Benefits Summary



72% monthly cost savings.



Fast data ETL.



Better consumability and customization of AI-powered data.

CHALLENGE

Challenge: Optimize data analytics costs without sacrificing performance.

An AWS Well-Architected Review revealed to the company that data analytics running on the Cloudera platform was driving high costs. While they were satisfied with the performance of the Cloudera platform, they wanted to reduce their data analytics costs without sacrificing performance.

Why AWS and nClouds

The company turned to nClouds, a Premier Consulting Partner and Well-Architected Partner in the AWS Partner Network, to collaborate on an [AWS Well-Architected Review](#) based on a set of architectural best practices for reviewing critical workloads running on Amazon Web Services (AWS). The Review revealed opportunities for infrastructure improvement, including cost optimization, security, operational excellence, and reliability.

After the Review, based on nClouds' AWS technical expertise in data and analytics, the company asked nClouds to help them optimize costs by migrating from the Cloudera platform to an AWS architecture featuring AWS managed services for data analytics with pay-as-you-go pricing.

 <p>Premier Consulting Partner</p> <ul style="list-style-type: none"> AWS CloudFormation Data & Analytics Competency DevOps Competency Immersion Day Partner Marketplace Seller 	 <p>Premier Consulting Partner</p> <ul style="list-style-type: none"> Migration Competency MSP Partner Public Sector Partner SaaS Competency Well-Architected Partner Program
--	--

And, the company asked nClouds to partner with them to enhance security. They implemented security best practices such as enabling MFA on the root; implementing life cycle, MFA, and password policies for all users; disabling or removing all inactive keys; restricting public access to virtual private cloud (VPC) resources; checking for Identity and Access Management (IAM) policies attached to users; and monitoring resources for encryption at rest.

AI Innovator leverages several Amazon services:

- **Amazon API Gateway** - Makes it easy for the company's developers to create, publish, maintain, monitor, and secure APIs at any scale.
- **Amazon Athena** - An interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to manage, and the company pays only for the queries that are run.
- **Amazon Comprehend** - A machine learning-powered service that makes it easy to find insights and relationships in text.
- **Amazon DynamoDB** - A key-value and document database that delivers single-digit millisecond performance at any scale.
- **Amazon ElastiCache** - Makes it easy for the company to scale out, scale in, and scale up to meet fluctuating application demands.
- **Amazon Elasticsearch Service** - A fully managed service that makes it easy for the company to deploy, secure, and operate Elasticsearch at scale with zero downtime.
- **Amazon Kinesis Data Firehose** - Part of the Kinesis streaming data platform, it's a fully managed service (so that the company doesn't need to write applications or manage resources) that delivers real-time streaming data.
- **Amazon S3** - A flexible way to store and retrieve data, providing the company with cost optimization, access control, and compliance.
- **Amazon Fargate** - Enables the company to run containers without having to manage servers or clusters.
- **Amazon Glue** - A fully managed extract, transform, and load (ETL) service that makes it easy for the company to prepare and load their data for analytics.
- **Amazon Lambda** - Enables the company to run code without provisioning or managing servers. They only pay for the compute time consumed – there is no charge when code is not running.

The company's solution stack also includes:

- **nOps** - A SaaS cloud management platform for AWS that facilitated the company's AWS Well-Architected Review with automated discovery and enables them to optimize cost and manage security and compliance for their AWS resources.

nClouds' Solution Architecture for AI Innovator

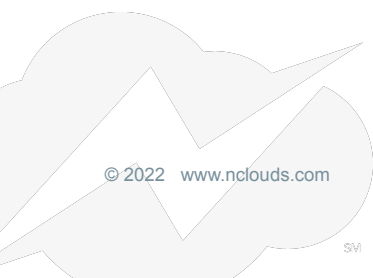
The company engaged with nClouds to help them optimize their data analytics costs by migrating from the Cloudera platform to an AWS infrastructure featuring AWS Glue, Amazon Athena, and Amazon S3.

Amazon Glue:

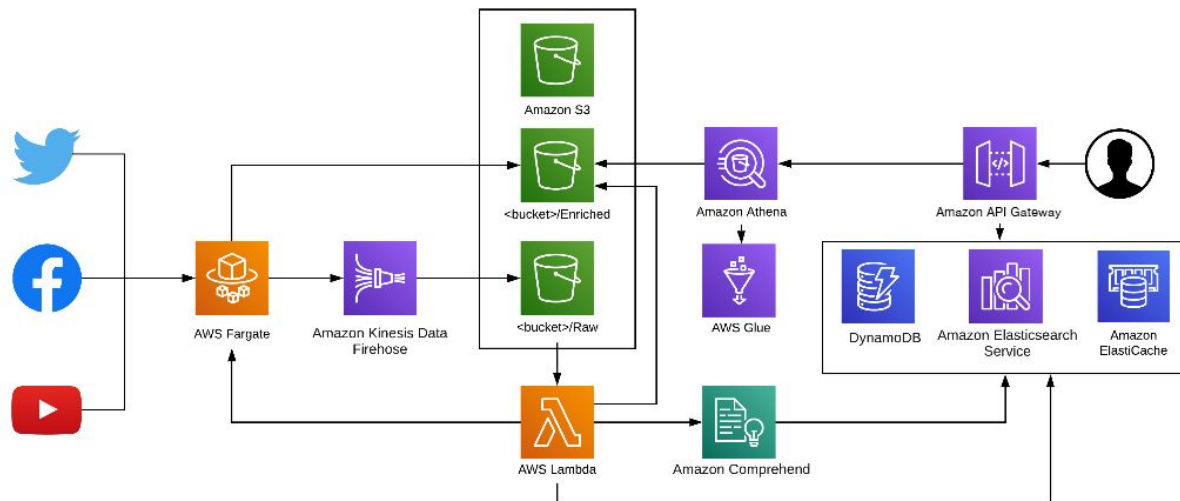
AWS Glue ETL service prepares and loads the company's data for analytics. It structures different formats of datasets in a standard way and puts a metadata layer on top of it (the AWS Glue metadata catalog) so that all data can be accessed in the same way. As a managed service, there are no resources to manage, no upfront costs, and there's no charge for startup or shutdown time.

Amazon Athena:

Once the data is in the AWS Glue metadata catalog, it is sent to Amazon Athena, which provides significant cost savings and performance gains by compressing, partitioning, or converting data to reduce the amount of data that Athena must scan to execute a query.



High-level architecture diagram:



The Benefits

Teaming with AI Innovator, nClouds migrated data analytics from Cloudera to AWS. The project has yielded numerous benefits:



72% monthly cost savings.

With AWS Glue, the company pays only for the time their ETL job takes to run. Amazon Athena charges the company only for the queries that they run based on the amount of data scanned by each query. Amazon S3 saves storage costs via S3 Intelligent-Tiering, and data access costs are reduced (by as much as 80%) by using Amazon S3 Select.



Fast extract, transform, load (ETL) of data.

Using AWS Glue, the company's data is immediately searchable, queryable, and available for ETL. It automates much of the effort in building, maintaining, and running ETL jobs. AWS Glue is serverless, so there is no infrastructure to provision or manage.



Better data consumability and customization.

AWS Glue takes datasets in different formats and converts them into a single, query-optimized format that can be easily consumed by various analytical tools.

Endnote

Amazon Web Services, Inc. (2018). Amazon S3 Select Is Now Generally Available.
<https://aws.amazon.com/about-aws/whats-new/2018/04/amazon-s3-select-is-now-generally-available/>

About nClouds

nClouds is a certified, award-winning provider of AWS and DevOps consulting and implementation services. We partner with our customers, as extensions of their teams, to build and manage modern infrastructure solutions that deliver innovation faster. We leap beyond the status quo.

Copyright © 2022 nClouds, Inc. All rights reserved

