



**Mortgage  
Processing  
Software  
Company -  
Data Analytics  
Platform utilizing  
DevOps**



## Contents

---

<b>About the Client .....</b>	<b>3</b>
<b>Problem Statement .....</b>	<b>3</b>
<b>What We Implemented .....</b>	<b>3</b>
<b>Outcomes and Benefits .....</b>	<b>5</b>

## About the Client

---

The customer is a software company that processes 35% of U.S. mortgage applications. The services are based on a software as a service model (SaaS) that streamlines and automates the process of originating and funding new mortgage loans and facilitating regulatory compliance

## Problem Statement

---

The customer wanted to build a multi-tenant analytics platform on AWS that will enable their customers get real time mortgage data utilizing DevOps tools and services. The platform needed to enable their customers with self-service search, BI and real time Analytics. Apart from this the data platform, they needed to open cross organization opportunities for data monetization exploration.

## What We Implemented

---

The Persistent team utilized AWS services and DevOps tools such as EC2, Batch, Lambda, CloudFormation, RDS, S3, SNS, SQS, ECS, Autoscaling, Load Balancing, ElasticCache, Route53, and ACM. Persistent also implemented security and compliance on the AWS infrastructure using VPC, IAM, GuardDuty, Inspector, and KMS and implemented monitoring solutions using CloudWatch, Elasticsearch and Kibana.

### DevOps

#### Infrastructure as Code (IaC) and Automation:

The infrastructure necessary for the platform is provisioned through infrastructure-as-code using CloudFormation to automate the creation and deployment. The automated process builds the infrastructure and applications in a repeatable and consistent manner and saves a significant amount of manual provisioning time.

Shell, Node.js and Python scripting is used extensively in the infrastructure code for leveraging the AWS CLI commands and passing the parameters to the CloudFormation templates on runtime. CloudWatch events and rules are used along with the lambda to automate scheduled start and stop of multiple services and instances. CloudFormation has helped to achieve the repeatability and setup an identical infrastructure on AWS by avoiding the manual errors.

**CI/CD and Source control:**

GITHUB is used to store all the source code and binaries and Jenkins is used for the Continuous Integration and Continuous Delivery/Deployment. Continuous integration is setup with GIT webhooks and Jenkins to automatically build the source code and deploy as soon as a push request is made on GIT branch. Infrastructure deployment is also setup with Jenkins and CloudFormation on AWS with a parameterized Jenkins pipeline.

**Management:**

The Persistent DevOps team manages all different environments on different AWS account for the customer and data platform deployment. Different AWS accounts and VPCs are used for keeping the lower and production environments isolated at every stage of the software development lifecycle. Persistent also uses AWS CloudWatch, CloudTrail, config and autoscaling for managing different AWS accounts, applications and environments.

**Storage:**

Appropriate storage options are used for different use cases. S3 is used for multiple use cases like storing JSON files from DataConnect platform, storing the PostgreSQL aggregation dumps and storing the configurations files for different environments. AWS RDS with PostgreSQL engine is used where the data is stored in structured/relational format. DynamoDB is also used to maintain the user login and session information.

**Security:**

Persistent Systems is responsible for the security and compliance on the AWS infrastructure using VPC, NACL, IAM, GuardDuty, Inspector, and KMS. Aggregated data is stored securely in the RDS database in private subnets and also implemented the KMS encrypted username and passwords for connecting the databases from application. All the AWS users have MFA enabled and integrated with the Active Directory. IAM policies are implemented based on the least privileges principal. To ensure the data security in transit, Amazon certificate manager is used for leveraging SSL certificates with the load balancer to enable SSL communication for all the applications. KMS keys are also used for secrets and config file encryption.

**Continuous Logging and monitoring:**

VPC flow logs and CloudTrail events are monitored by an in-house InfoSec team regularly for activities which need to be reviewed. All the suspicious events are tracked and monitored continuously, and suitable actions are taken for the reported incidents.

Persistent implemented a continuous monitoring system where all the running services will send their logs to the CloudWatch and the events will be continuously streamed to the Elasticsearch for better monitoring and query. We have also used CloudWatch event Metrics and Alarms to monitor the activities and to alert when the thresholds are breached.

**Alerts:**

CloudWatch alarms and event rules are used for sending email/slack alerts when the thresholds are breached for specific services. We have also setup the CloudWatch custom metrics for application errors and used SNS with Lambda to send alerts to the appropriate channels.

**Disaster Recovery and Failover:**

Persistent has implemented the push button disaster recovery setup for all applications. We have also implemented an on-demand DR setup which will quickly deploy the resources in a recovery region in case of a region failure. AWS RDS has multi availability zone setup and Read replica in a different region for handling availability zone failures and region failures. Read replicas are promoted as the master DB instance in case of Primary region failure where multi availability zone setup ensures the self-healing approach for a single availability zone failure.

DynamoDB global tables are used for having DynamoDB data available across multiple region where we can failover from a region failover. RTO of 45-60 minutes have been considered for the downtime and restore. No data loss is expected since data replication is enabled in another region.

**Cost optimization:**

Persistent has implemented a cost-effective solution on AWS by following AWS best practices such as AWS cost allocation tagging practices for all the services and their underneath resources such as EC2, lambda, and RDS. The team leveraged spot instances and on demand instances for smaller and temporary operations. Autoscaling scheduled actions feature was used to terminate the instances when they are not in use on daily basis. Some of the cost control measures are made available to the teams so that they can bring up environments when they need, and they will be brought down by a schedule. CloudWatch log retention feature and S3 lifecycle features are used for log archival and data archival respectively. In some cases, there are multiplexed multiple environments on common resources to eliminate the need for having multiple RDS instances.

## Outcomes and Benefits

---

Persistent built and implemented a multi-tenant data lake platform using a modern cloud data stack for streaming and batch mode data analytics. The DataConnect product has been released and adopted by 75+ customers. Persistent utilized their expertise in implementing AWS and DevOps solutions leveraging AWS services like EC2, Batch, Lambda, CloudFormation, RDS, S3, SNS, SQS, ECS, Autoscaling, Load Balancing, ElasticCache, Route53, and ACM and helped the customer to build the cloud-native data application leveraging AWS data and analytics services. Utilizing CloudFormation has helped to achieve the repeatability and setup an identical infrastructure on AWS by avoiding the manual errors.



## About Persistent Systems

Persistent Systems is a global company specializing in solutions, software products, and technology development services. Persistent Systems builds innovation that drives the business of our customer's enterprises with software at the core of their digital transformation. To learn more, visit <http://www.persistent.com>

### India

#### Persistent Systems Limited

Bhageerath, 402,  
Senapati Bapat Road, Pune 411 016  
Tel: +91 (20) 6703 0000  
Fax: +91 (20) 6703 009  
Email: [info@persistent.com](mailto:info@persistent.com)

### USA

#### Persistent Systems Inc

2055 Laurelwood Road, Suite 210,  
Santa Clara, CA 95054  
Tel: +1 (408) 216 7010  
Fax: +1 (408) 451 9177  
Email: [info@persistent.com](mailto:info@persistent.com)