INDUSTRY
Software engineering

TECHNOLOGIES
Amazon Web Services (AWS), Amazon Elastic Container Services (ECS), Docker software container platform, Google Cloud, Google Cloud Machine Learning Engine, Google OAuth, Rancher container management platform

BUSINESS NEED
A U.S.-based software engineering company needed a "pipeline" to seamlessly connect its machine learning algorithms in Google Cloud to its client's reference database in the AWS cloud.

SOLUTION
Pythian custom-built a solution that allowed the company to span a cluster of Docker containers across the two clouds, complete with time-limited authentication to give its data scientists access to AWS services from within Google Cloud.

RESULTS
The company now has a reusable platform with DevOps support for multi-cloud infrastructures to assist its data scientist in building and deploying machine learning solutions at scale.

BRIDGING TWO CLOUDS TO BRING MACHINE LEARNING TO CROP DISEASE MANAGEMENT

Crop disease is difficult to diagnose, traditionally requiring plant pathologists to visit farmers' fields to physically examine plant specimens. Machine learning can accelerate this manual process by quickly analyzing and deriving insights from massive volumes of data—for example, by allowing photos from a farmer's phone to be checked against thousands of comparative crop images, serving up diagnostics and treatment options for keeping crops healthy in near real-time.

A U.S-based software engineering company was tasked with developing exactly such a solution on behalf of a global agricultural biotechnology firm. To make it work, the company called on the multi-cloud experts at Pythian to connect the cloud hosting its machine learning model to the cloud hosting its client's reference data.

BUSINESS NEED

As it was developing its mobile app for farmers, the software engineering company leveraged the advanced image analytics of Google Cloud Vision API, powered by the pattern-recognition capabilities of the Google Cloud Machine Learning Engine.

There was just one catch: while its machine learning model was based in Google, the data it was analyzing was not. For the app to work, the model had to cross-reference the many different datapoints in a farmer's photo—crop color, texture, and decay patterns—against the agricultural biotechnology firm's reference library of some 50,000 photos, which was stored on an entirely different platform: Amazon Web Services (AWS).

To complete the app, the company's data scientists needed to seamlessly connect the Google and Amazon clouds in a way that ensured fast and secure movement of data between the two platforms.

Pythian
love your data®

> " *Cloud vendors make it easy to manage security within their own cloud. But when it comes to dealing with security to another cloud, there's no ready-to-go solution there. We had to build our own custom solution to deal with that."*
>
> *Pierig LeSaux, Pythian*

ABOUT PYTHIAN

Pythian is a global IT company that helps businesses leverage disruptive data technologies to better compete. Our services and software solutions unleash the power of cloud, data and analytics to drive better business outcomes. Our 20 years in data, commitment to hiring the best talent, and our deep technical and business expertise allow us to meet our client promise of using technology to deliver the best outcomes faster.

WORLDWIDE OFFICES

Ottawa, Canada
New York City, USA
London, England
Sydney, Australia
San Francisco, USA
Hyderabad, India

SOLUTION

Pythian drew on its in-depth knowledge of Google Cloud to create a "pipeline" between Google and AWS, making it possible for photos received by the agricultural biotechnology firm's AWS platform to be seamlessly pushed over to Google Cloud for analysis by the software engineering company's machine learning model. Through this pipeline, the Google-based model could seamlessly access the reference library housed on AWS—and then instantly send its findings back to AWS and onward to the farmer.

Pythian custom-built this solution using an open-source Docker cluster orchestration tool called Rancher, which allowed the company's data scientists to span a cluster of Docker containers across both the AWS and Google clouds.

To ensure the pipeline is secure as data travels between the two clouds, Pythian also built a custom multi-cloud authentication solution through Google OAuth, which provides services running in AWS with time-limited access to the Google Cloud.

RESULT

With the flexible and easy-to-use deployment pipeline built by Pythian to connect the Google and AWS clouds, the software engineering company was able to ensure its data scientists had secure access to the agricultural biotechnology company's dataset during app development and testing. Pythian's multi-cloud expertise helped the company establish a reusable platform that enables efficient data and model management, along with DevOps support for multi-cloud infrastructure to assist its data scientists in performing data cleansing, learning, and service deployment.

The final version of the mobile app—powered by the Pythian pipeline—will make it possible for farmers to get instant diagnostics and recommendations on the health of their crops, ultimately helping to improve crop quality and yields.

Pythian
love your data®