

Containerization for AI/ML Workloads

Industry: EdTech
Workloads: AI/ML
Solution: Containers



ABOUT THE CLIENT: PEACHJAR

Peachjar is a medium-size EdTech that produces eLearning and communication software solutions for K-12 schools and families. Its platform streamlines communication by distributing resources directly to parents.

CHALLENGE

Peachjar wanted a solution to run specialized machine learning and data science workloads on an elastically scalable, cost-efficient platform. Their existing system of VMs required excessive maintenance and was both difficult to scale and replicate. The orchestration of their TensorFlow and Spark workloads was handled through a custom-built solution that failed to meet Peachjar's needs and took capacity from their development teams.

Their manually deployed VM solution was detrimental to their testing efforts and increased the risk of delivering changes and new workloads to their production environment. Peachjar's existing solution was also cost inefficient. They required large and specialized instance sizes, and these resources remained idle for long periods of time, accumulating large cloud bills for resources that were severely underutilized.

WHY AWS

Peachjar's engineering and technology teams are comfortable with containerized environments and wanted to move their data science and machine learning workloads to a containerized solution. They did not, however, have the staffing in place to maintain a highly specialized Kubernetes cluster, and did not see value in maintaining the underlying infrastructure of the container orchestration platform.

AWS' Elastic Kubernetes Service was a perfect fit for Peachjar's needs. Other services such as CloudFormation, CodeCommit, CodeBuild, and CodePipeline also provided a highly duplicatable solution that allows the customer to reliably deliver the solution across environments easily and quickly.

WHY STRATUS10

Stratus10's expertise to deliver the containerized solution ensured that Peachjar's received the best combination of tools available. Their experience in delivering complex containerized solutions across the AWS ecosystem resulted in a smooth delivery of the EKS solution that met all scaling, cost, and performance targets.

SOLUTION

Stratus10 assisted Peachjar in migrating their ML and data science workloads from a static, VM-based solution to an elastically scalable containerized solution orchestrated by EKS. Argo was used to replace a custom task orchestration solution that was difficult to deploy and maintain, and multiple node groups were used to optimize performance for specialized data science jobs.



Stratus10 delivered a solution consisting of EKS, EFS, and Argo solutions deployed using CloudFormation. EKS was used as the container orchestration solution both because the customer was familiar with Kubernetes and because the managed control plane requires very little administration and maintenance. Stratus10 used multiple managed node groups to allow for the use of different instance sizes for the specialized jobs. A robust tagging and labeling solution was also created to ensure the Kubernetes NodeSelector solution could be used to target pods to specialized nodes. The Horizontal Pod Autoscaler and the Cluster Autoscaler were used to elastically scale both pods and nodes within the cluster.

Amazon Elastic File System (EFS) was mounted to all nodes, providing access to the data for all workloads and eliminating the need for tasks that were transferring the data between systems. The EFS mount was delivered as a shared secret making it very easy to mount into pods. Argo, Spark, and TensorFlow made up the primary components of the customer's machine learning solutions. Stratus10 assisted in the deployment and configuration of containerized versions of each of these components. Through CloudFormation, Peachjar could reliably deliver the ML and data science cluster to multiple environments. They can also easily tear down the cluster during idle times and quickly bring the entire solution back up when it's needed.

RESULTS AND BENEFITS

The machine learning and data science EKS solution delivered by Stratus10 met all performance and cost optimization goals set by Peachjar. Workloads are now delivered through containerized solutions the customer is familiar and comfortable with, and the time and resources required to maintain their ML systems has dropped significantly. By targeting autoscaling node groups for specialized workloads, Peachjar can now run their jobs on larger, faster infrastructure for less than the previous static solution. Finally, the solution is highly repeatable across their environments and Peachjar can more confidently release changes to their ML solution.

ABOUT STRATUS10

We're a team of AWS experts with Fortune 500 pedigree and an entrepreneurial mindset. We have refined the process of cloud adoption and optimization by including the human element in every step of the process, assuring the ideal solution for every client and workload.