

REDUCED DATA SCIENCE PLATFORM COSTS FOR ONRULE

CASE STUDY

COMPANY OVERVIEW

OnRule, a cloud-based Software-as-a-Service platform developed with compliance teams in mind, helps clients increase visibility and exercise better control over any product's regulatory compliance. OnRule simplifies the process of monitoring, communicating, and reporting product certification status across markets and products.

CHALLENGE

- Linearly increasing data center computing costs as business scaled.
- High GPU compute and data transfer costs for training datasets used in AI/ML solutions on GPU servers.
- On-demand AWS pricing that was suitable during the startup stage became increasingly expensive with growth.
- Always-on workloads further escalated costs, making operations cost-prohibitive.

SOLUTION

- Collaborated with the client to address network performance challenges for offshore development centers (ODCs) connecting to end customer networks.
- Designed a private cloud solution for enhanced security controls and visibility.
- Optimized GPU usage with an Acceleration on Demand solution, increasing server utilization and reducing overall GPU requirements.

SOLUTION (CONTD.)

- Provisioned required images with AI libraries and frameworks like TensorFlow.
- Implemented a memory caching solution for low-latency storage access.
- Built the solution using NVIDIA Tesla V100 series GPUs.
- Assisted with data migration from AWS to the private cloud.

BENEFITS

- 75% reduction in GPU server costs, saving over $1 million annually.
- Enabled future budget savings and reallocation to go-to-market initiatives.
- Achieved greater infrastructure control and improved client satisfaction with enhanced network connectivity.
- Completed implementation in days instead of months, accelerating savings.
- Utilized proven deployment models for rapid execution.