Solution Brief

AI OS - The ML/AI Cockpit

Automate MLOps, DevOps, and DataOps for data scientists and ML teams

Key Features

Data, models, and infrastructure have fused together to create one automated pipeline for data scientists

- Provide onboarding and caching of frequently used datasets
- Attach models directly from the NVIDIA GPU Cloud (NGC) catalog with one click
- Leverage Red Hat OpenShift as a Kubernetes managed platform to simplify and automate machine learning (ML) infrastructure

Industry-first solution for ML dataset caching

- Create the optimal proximity of datasets to compute
- Boost productivity
- Evict tightly managed datasets
- Work in the cloud with NetApp® Cloud Volumes ONTAP® and on the premises with NetApp storage systems

Automate your ML Kubernetes infrastructure

- Attach resources to models with one click, with all other MLOps semantics hidden
- Partner with managed Kubernetes leaders, such as Red Hat OpenShift, to offer one single, unified control plane for data scientists

Challenge

ML development through deployment relies on ad-hoc tools, plug-ins, scripts, and a myriad of siloed tools that are impeding organizations, large and small, from streamlining ML development. Both the infrastructure and the application layers are going through fundamental changes. The complexities of MLOps also creates a new subcategory of DataOps, which encompasses the processes for ingesting, engineering, cleaning, and prepping the data for ML pipelines.

The Solution

NetApp and cnvrg.io have partnered to deliver an AI/ML data science pipeline solution that is streamlined and drives productivity and efficiency. The solution incorporates industry-leading Kubernetes managed clusters (for example, Red Hat OpenShift), cached datasets for extreme performance, and the one-click attachments of models to datasets with NVIDIA NGC integration.

It's not uncommon to have hundreds of datasets feeding models. However, those datasets may live far away from the compute that is training the models, such as in the public cloud or in a data lake. With NetApp and cnvrg.io, you can cache the needed datasets (and/or their versions) and make sure that they're located in the ONTAP® node attached to the GPU cluster or CPU cluster that is exercising the training. Once the needed datasets are cached, they can be used multiple times by different team members. Caching creates the following business advantages:

- **Increased productivity.** Datasets are ready to be used in seconds rather than hours.
- **Improved sharing and collaboration.** Cached datasets can be authorized and used by multiple teams in the same compute cluster connected to the NetApp cached data.
- **Reduced cost.** Models are pulling the datasets from the cache, not from remote storage, which may require payments per download.
- **Facilitation of hybrid cloud**. The cache can present the on-premises mirror storage for the data lake residing in the cloud.

NetApp®

cnvrg.io

Technology partner.

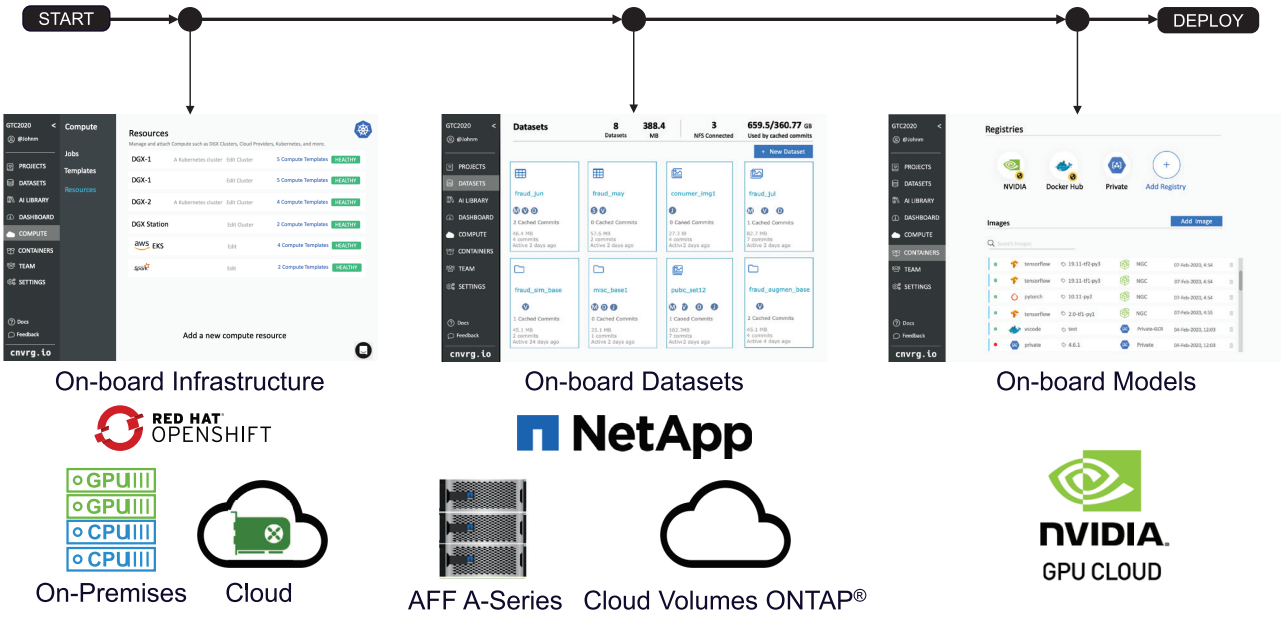

Figure 1) NetApp ONTAP and cnvrg.io AI/ML workflow.

About cnvrg.io

cnvrg.io is an AI OS, transforming the way enterprises manage, scale and accelerate AI and data science development from research to production. The code-first platform is built by data scientists, for data scientists and offers flexibility to run on-premise or cloud. With Model-management, MLOps and continual machine learning solutions, cnvrg.io brings top of the line technology to data science teams so they can spend less time on DevOps and focus on the real magic - algorithms. Since using cnvrg.io, teams across industries have gotten more models to production resulting in increased business value.

About NetApp

NetApp is the leader in cloud data services, empowering global organizations to change their world with data. Together with our partners, we are the only ones who can help you build your unique data fabric. Simplify hybrid multicloud and securely deliver the right data, services, and applications to the right people at the right time. Learn more at **www.netapp.com**.