

Solution Brief

Democratize AI with a NetApp Data Fabric and H2O Driverless AI

Machine learning models at scale on site or in the cloud

Key Benefits

Deliver AI at Scale

Scale from zero to 100TB deployments in a matter of seconds with a simple, easy-to-use cloud interface.

Reduce Risk

Instantaneous NetApp® Snapshot™ copies and cloning enable data scientists to experiment with datasets without risking data loss.

Dynamically Change Service Levels

Enhance performance or reduce operational expenditures (opex) with the ability to dynamically change storage service levels.

Build More Models in Less Time

Reduce the time that it takes to develop accurate, production-ready models by automating time-consuming data science tasks.

Train Across the Hybrid Cloud

Seamlessly move H2O Driverless AI workloads between on-premises NetApp ONTAP® AI infrastructure and NetApp Cloud Volumes Service.

The Challenge

Visionaries in virtually every industry are now looking for ways to apply artificial intelligence (AI). AI applications can include driving business value, creating new customer experiences, acting upon insights from real-time and analytic data, obtaining a 360-degree view of customer challenges, or anything in between. Unfortunately, many organizations still underestimate how much AI depends on an ability to manage, secure, and share vast quantities of data. A bottleneck at any point idles expensive infrastructure and increases costs. Data scientists waste valuable time on troubleshooting infrastructure and waiting for results.

The Solution

NetApp has partnered with H2O.ai to integrate NetApp Cloud Volumes Service, our cloud-native file storage service, and NetApp ONTAP AI with H2O Driverless AI. H2O Driverless AI is a machine learning (ML) platform that empowers data scientists to scale and to deliver trusted, production-ready models. It automates time-consuming data science tasks, including advanced feature engineering, model selection, and model deployment. Model deployment is streamlined with automatic scoring pipelines that include everything that you need to run the model in production.

Cloud Volumes Service is a fully managed, cloud-native file storage service that is based on proven NetApp ONTAP data management software. Cloud Volumes Service combines NetApp's vast file services expertise with the simplicity and flexibility of the biggest clouds, including Amazon Web Services (AWS), Azure, and Google Cloud Platform. ONTAP AI is a proven architecture that is powered by NVIDIA DGX supercomputers and NetApp cloud-connected all-flash storage. With ONTAP AI, you can streamline data access and accelerate AI with a NetApp Data Fabric that spans from the edge to the core to the cloud.

Driverless AI seamlessly integrates with Cloud Volumes Service, giving you an easy and convenient way to build trusted ML models at scale on any of the top-three public clouds. After a cloud volume has been configured and mounted by Cloud Volumes Service, it is seen as one of the data sources in the Driverless AI user interface. So, you get convenient access to data with a shared data management model. Cloud Volumes Service facilitates collaboration across multiple instances of Driverless AI that use the same cloud volume, so data scientists can share experiments, artifacts, and model updates. Integration of Driverless AI with ONTAP AI enables you to train workloads in the cloud and on the premises in a hybrid cloud model.

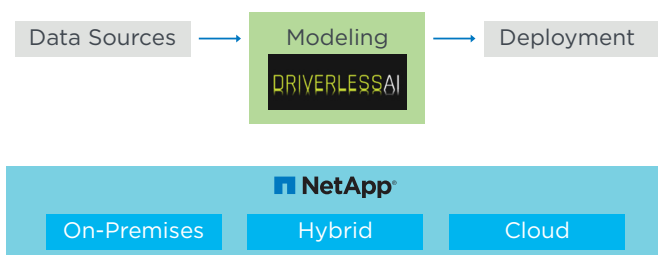


Figure 1) Build AI models at scale on site or in the cloud.

Deliver AI at Scale

H2O Driverless AI on NetApp Cloud Volumes Service and ONTAP AI empowers your data engineering team to develop business-changing applications with the industry's leading automatic ML platform. With H2O Driverless AI, you can scale from zero to 100TB deployments in a matter of seconds, with support for enterprise data access and connectors to the world's largest public clouds. Driverless AI's simple cloud-native service, with an easy-to-use interface and single payment model, makes it simple to deploy resources without worrying about spending more than you need to. And with no fee per read or write, you can save on API charges.

Reduce Risk

Migrate, replicate, and synchronize data across on-premises infrastructure and clouds, using NetApp Snapshot copies for data protection and restores and using NetApp FlexClone® technology for data cloning. Instantaneous Snapshot copies and cloning features enable you to experiment with datasets without risking data loss. If Driverless AI goes down, all your data and experiments are saved and are replicated with NetApp Snapshot technology.

Dynamically Change Service Levels

To enhance performance or to reduce opex, you can easily and conveniently provision and mount cloud volumes on Driverless AI with support for changing storage service levels. Leverage high storage performance with three service levels that you can change dynamically.

With the Driverless AI checkpointing feature, your data scientists can even restart jobs on alternate compute instances without impeding their progress. If a job is taking too long to complete, a data scientist can mount the cloud volume on a new, powerful compute node with GPUs that are provisioned and restarted

from a checkpoint. Alternatively, they can also downgrade to a lower compute instance to save on opex if necessary. With Cloud Volumes Service, data scientists can deploy and collaborate on multiple instances of Driverless AI.

Build More Models in Less Time

Reducing the time that it takes to develop accurate, production-ready models is critical to solving a large number of business challenges with AI. Driverless AI automates time-consuming data science tasks, including advanced feature engineering, model selection, hyperparameter tuning, model stacking, and model deployment.

The speed and scalability of Cloud Volumes Service enable you to test thousands of combinations and iterations to find the best model in minutes. NetApp Cloud Volumes Service gives you public cloud scalability. You can be confident that your Driverless AI workloads have the resources that they need, when they need it, so you can expand your modeling pipeline as business demands dictate.

Train Across the Hybrid Cloud

You can seamlessly move H2O Driverless AI workloads between on-premises ONTAP AI infrastructure and NetApp Cloud Volumes Service. Start in the cloud and move to an on-premises environment if you need more processing power and storage, or start on the premises and move data to the cloud for archiving.

About H2O.ai

H2O.ai is the open-source leader in AI, with a mission to democratize AI for everyone. H2O.ai is transforming the use of AI with software, with its category-creating visionary open-source machine learning platform, H2O. More than 18,000 companies use open-source H2O in mission-critical use cases for Finance, Insurance, Healthcare, Retail, Telco, Sales and Marketing. H2O Driverless AI uses AI to do AI in order to provide an easier, faster, and cost-effective means of implementing data science.

[Learn more.](#)

About NetApp

NetApp is the data authority for hybrid cloud. We provide a full range of hybrid cloud data services that simplify management of applications and data across cloud and on-premises environments to accelerate digital transformation. Together with our partners, we empower global organizations to unleash the full potential of their data to expand customer touchpoints, foster greater innovation and optimize their operations. For more information, visit www.netapp.com. #DataDriven