NetApp AI Control Plane

Full-stack AI data and experiment management across a hybrid cloud

■ NetApp

Challenge

Artificial intelligence (AI) and machine learning (ML) are changing the game by helping organizations around the world to accelerate innovation and gain a competitive advantage. As organizations increase their use of AI and ML, they face three major challenges: workload scalability, deployment complexity, and data availability. There are many frameworks in the marketplace that attempt to improve scalability and simplify deployment, but most fail to address the critical challenge of data availability. Many also feature proprietary data platforms that lack proven, enterprise-class reliability and that don't scale across different sites and regions.

Solution

The NetApp® AI Control Plane enables you to unleash your AI and ML with a solution that offers extreme scalability, streamlined deployment, and nonstop data availability—when and where you need it. The AI Control Plane solution integrates Kubernetes and Kubeflow with a data fabric enabled by NetApp. Kubernetes, the industry-standard container orchestration platform for cloud-native deployments, enables workload scalability and portability. Kubeflow is an open-source machine learning platform that simplifies management and deployment, enabling your developers to do more data science in less time. A data fabric enabled by NetApp offers uncompromising data availability and portability to ensure that your data is accessible across the pipeline, from edge to core to cloud.

Key benefits

Scale seamlessly

- Leverage Kubernetes to scale AI workloads across regions and sites
- Enable workload portability

Streamline data science

- Simplify AI workload and data science workspace provisioning
- Quickly and easily define end-to-end AI workflows

Get the most out of your data

- Trust that your data is always available whenever and wherever it's needed, no matter where it resides
- Build your MLOps pipeline

Scale Seamlessly

With AI Control Plane, you can take advantage of Kubeflow, an innovative AI/ML Kubernetes-native framework that provides a standard and open platform for simplifying deployment of AI/ML workloads. Kubeflow abstracts the intricacies of Kubernetes, allowing data scientists and developers to focus on their work, not on data administration.

Kubeflow also includes Jupyter Notebooks, enabling team leads and administrators to provision and destroy Jupyter Notebook servers on demand. When Kubeflow is deployed as part of the NetApp AI Control Plane solution, data volumes, potentially containing petabytes of data, can be presented to data scientists as simple folders within a Jupyter workspace. Data scientists have instant access to all of their data from in a familiar interface.
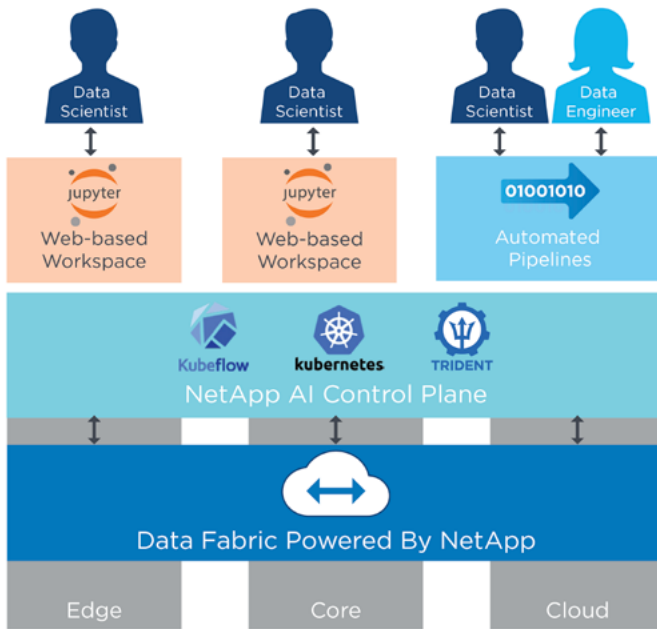
Figure 1) NetApp AI Control Plane architecture.

Streamline Data Science
Using NetApp Trident, NetApp's persistent storage provisioner for Kubernetes, data volumes stored in your data fabric can be presented to Kubernetes workloads in an enterprise-class, Kubernetes- and cloud-native format.

With Trident, developers can perform data management functions, such as triggering the creation of snapshot copies and creating volume clones, using standard Kubernetes API calls and commands directly from within a Kubeflow pipeline. This ability provides storage-efficient traceability, because a snapshot copy does not consume additional storage space until it starts to deviate from its source volume. The same developer can clone a volume that contains a trained AI/ML model. This clone can subsequently be used as a dev/test workspace or for A/B testing.

Kubeflow also allows data scientists to define end-to-end AI/ML workflows using a simple Python SDK. Your data scientists don't need to know how to define Kubernetes deployments or execute commands—they can continue to use familiar Python commands and accelerate time to production.

Get the Most Out of Your Data
NetApp enables you to build a data fabric that facilitates seamless data movement and access across edge, private cloud, and public cloud sites— all while providing enterprise-class data management and data protection capabilities. With a data fabric enabled by NetApp, you get enhanced data portability and scalability beyond what Kubernetes alone can provide. Your data scientists no longer have to wait for days while the datasets they need are being copied across sites, servers, or even workstations. You can trust that your data will be there whenever and wher-ever it's needed.

When your data fabric is paired with Kubernetes, AI/ML work-loads and petabytes of training data can be seamlessly scaled together across sites and regions. You no longer have to worry about tracking changes across multiple versions of the same datasets or the risk of losing track of a specific copy.

About NetApp
In a world full of generalists, NetApp is a specialist. We're focused on one thing, helping your business get the most out of your data. NetApp brings the enterprise-grade data services you rely on into the cloud, and the simple flexibility of cloud into the data center. Our industry-leading solutions work across diverse customer environments and the world's biggest public clouds.

As a cloud-led, data-centric software company, only NetApp can help build your unique data fabric, simplify and connect your cloud, and securely deliver the right data, services and applications to the right people—anytime, anywhere. www.netapp.com

■ NetApp