



使用 Union.ai 和NetApp FlexCache部署混合 AI 训练

NetApp artificial intelligence solutions

NetApp
December 04, 2025

目录

使用 Union.ai 和 NetApp FlexCache 部署混合 AI 训练	1
概述	1
客户用例：混合云人工智能培训	1
客户价值	1
启用插件 - 前提条件	1
参考架构	2
步骤 1：创建 FlexCache 卷	2
步骤 2：配置 Trident	2
步骤 3：部署 Union.ai 工作流	3
加载并使用来自 PVC 的数据进行训练。	5
步骤 4：验证集成	5
安全注意事项	5
监测与优化	5
相关链接	6
结束语	6
Union.ai - 配套指南	6

使用 Union.ai 和NetApp FlexCache部署混合 AI 训练

了解如何使用 Union.ai 编排、 NetApp FlexCache和Trident进行 Kubernetes 存储配置，从而部署混合 AI 训练环境。

David Espejo, Union.ai Sathish Thyagarajan, NetApp

概述

Union.ai 的混合编排平台与NetApp ONTAP和FlexCache无缝集成，可加速 AI/ML 训练工作流程。该解决方案允许数据安全地保留在本地，同时利用基于云的 GPU 计算进行 AI 训练工作负载。 NetApp FlexCache确保仅将必要的数据缓存在云端，从而实现高效、安全且可扩展的混合 AI/ML 管道。

客户用例：混合云人工智能培训

- 本地数据：存储在NetApp ONTAP上，以确保合规性和安全性。
- 云计算：在 EKS/GKE/AKS 上进行可扩展的 GPU 训练。
- AI/ML 编排：Union.ai 协调跨环境的数据处理和训练。
- 存储配置： NetApp Trident可自动配置 PVC/PV。

客户价值

- 利用NetApp ONTAP 的横向扩展功能，在海量数据集上运行 AI 工作负载。
- 使用 NetApp 的混合云功能，在本地和云端之间移动和同步数据。
- 使用FlexCache快速将本地数据缓存到云端。
- Union.ai 通过版本控制、血缘跟踪和工件管理简化了跨环境的编排。
- 在云端执行培训，同时将敏感数据保留在本地。

启用插件 - 前提条件

要求	细节
ONTAP 版本	ONTAP 9.7+（无需FlexCache许可证）
FlexCache许可证	ONTAP 9.6 及更早版本需要此功能
Kubernetes	本地集群和云集群（EKS/GKE/AKS）
Trident	可安装在本地集群和云集群上
Union.ai	控制平面已部署（联合云或自托管）
网络连接	集群间连通性（如果ONTAP集群是独立的）

权限	ONTAP和 Kubernetes 集群的管理员权限。 □请使用正确的ONTAP凭据（例如，vsadmin）。
刚接触 Union.ai?	请参阅本文档末尾的配套指南。

参考架构

下图展示了 Union.ai 控制平面与NetApp存储的集成，用于混合 AI 训练。

[采用 Union.ai 和NetApp 的混合 AI 训练架构] | *ai-hybrid-union-001.png*

- Union.ai 控制平面：协调工作流、管理数据移动并与NetApp API 集成。
- NetApp ONTAP + FlexCache：提供从本地到云端的高效数据缓存。
- 混合训练集群：训练作业在云 K8s 集群（例如 EKS）中运行，数据从本地缓存。

步骤 1：创建FlexCache卷

使用ONTAP系统管理器

1. 导航至“存储”>“卷”。
2. 单击“添加”。
3. 选择更多选项。
4. 启用“添加为远程卷的缓存”功能。
5. 选择源卷（本地）和目标卷（云）。
6. 定义服务质量或性能级别（可选）。
7. 单击“创建”。

□如果由于权限或聚合问题导致NetApp DataOps Toolkit 无法正常工作，请直接使用ONTAP系统管理器或 CLI 创建FlexCache卷。

步骤 2：配置 Trident

在两个集群上都安装Trident：

□ "[Trident安装指南](#)"

创建Trident后端

```
apiVersion: trident.netapp.io/v1
kind: TridentBackendConfig
metadata:
  name: ontap-flexcache
spec:
  version: 1
  storageDriverName: ontap-nas
  managementLIF: <ONTAP-MGMT-IP>
  dataLIF: <ONTAP-DATA-IP>
  svm: <SVM-NAME>
  username: vsadmin
  password: <password>

Apply: kubectl apply -f backend-flexcache.yaml
```

如果收到 401 未授权错误，请验证ONTAP用户是否具有足够的 API 权限，以及是否使用了正确的用户名 (vsadmin) 和密码。

定义存储类

```
apiVersion: storage.k8s.io/v1
kind: StorageClass
metadata:
  name: flexcache-sc
provisioner: csi.trident.netapp.io
parameters:
  backendType: "ontap-nas"
Apply:
  kubectl apply -f storageclass-flexcache.yaml
```

步骤 3：部署 Union.ai 工作流

Union 使用 PVC 将FlexCache卷挂载到训练作业中。

示例 PodTemplate

```
apiVersion: v1
kind: PodTemplate
metadata:
  name: netapp-podtemplate
  namespace: flytesnacks-development
  template:
    metadata:
      labels:
        default-storage: netapp
    spec:
      containers:
        - name: primary
      volumeMounts:
        - name: flexcache-storage
          mountPath: /data/flexcache
      volumes:
        - name: flexcache-storage
      persistentVolumeClaim:
        claimName: flexcache-pvc
```

示例工作流程

从联合导入任务、工作流

```
@task(pod_template="netapp-podtemplate")

def train_model(pvc_path: str):
```

加载并使用来自 PVC 的数据进行训练。

```
@workflow

def training_pipeline():
    train_model(pvc_path="/data/flexcache")
```

工会运营商将：

- 制作PVC管
- 挂载FlexCache卷
- 在云端 Kubernetes 集群中调度作业

步骤 4：验证集成

任务	验证
PVC支架	训练 pod 应该能够成功挂载 /data/flexcache。
数据访问	训练作业可以对FlexCache进行读写操作。
缓存行为	监控ONTAP中的缓存命中/未命中情况。确保聚合支持FlexCache
性能	验证训练工作负载的延迟和吞吐量

使用NetApp BlueXP或ONTAP CLI 监控性能。

安全注意事项

- 使用 VPC 端点为NetApp ONTAP 的FSx 启用
- 启用传输中和静态加密
- 申请基于角色的访问控制/身份与授权以访问ONTAP
- Union.ai 不会访问或存储客户数据

监测与优化

工具	目的
NetApp BlueXP	监控FlexCache 的使用情况和性能
Union.ai 用户界面	跟踪管道状态和指标
Trident日志	调试 PVC 或后端问题

可选增强功能

- 使用BlueXP API 自动创建FlexCache
- 使用 Union SDK 在训练前预热缓存
- 训练后添加批量推理或模型服务管道
- 如果 DataOps Toolkit 失败，则回退到通过系统管理器手动创建FlexCache。

故障排除

问题	解决
PVC卡在待定状态	检查Trident日志和后端配置
来自ONTAP API 的 401 未授权	使用 vsadmin 并验证权限
作业失败：没有合适的存储空间	确保ONTAP聚合支持FlexCache/ FabricPool
训练表现缓慢	检查缓存命中率和网络延迟
数据未同步	在ONTAP中验证FlexCache关系健康状况

后续步骤

1. 使用测试数据验证FlexCache
2. 部署 Union.ai 训练管道
3. 监控和优化性能
4. 记录客户特定设置

相关链接

- "[Union.ai 文档](#)"
- "[NetApp FlexCache概述](#)"
- "[Trident CSI 驱动程序](#)"
- "[适用于NetApp ONTAP 的 FSx](#)"

结束语

您现在拥有一个经过验证的混合 AI 训练环境，该环境使用 Union.ai 和NetApp FlexCache。训练作业可以在云端运行，同时安全高效地访问本地数据，而无需复制整个数据集或损害治理。

Union.ai - 配套指南

第一步：选择部署模式

方案A：联合云

- 访问： "[console.union.ai](#)"
- 创建组织 → 创建项目

选项B：自托管

- 跟随：<https://docs.union.ai/platform/latest/deployment/self-hosted/>["自托管指南"]
- 通过 Helm 部署：

helm repo add unionai <https://unionai.github.io/helm-charts/>

helm install union unionai/union -n union-system -f values.yaml

步骤 2：安装联合运算符

kubectl apply -f <https://raw.githubusercontent.com/unionai/operator/main/deploy/operator.yaml>

kubectl get pods -n union-system

步骤 3：安装 Union CLI

pip install unionai

工会登录

步骤 4：注册工作流程

联合项目创建混合人工智能

union register training_pipeline.py --project hybrid-ai

步骤 5：运行和监控

union run training_pipeline --project hybrid-ai

联合观察培训管道

查看日志 "[联合用户界面](#)"

步骤 6：注册计算集群（可选）

union cluster register --name cloud-k8s --kubeconfig ~/.kube/config

步骤7：追踪文物及传承

Union 会自动跟踪：

- 输入/输出参数
- 数据版本
- 日志和指标

- 执行谱系

版权信息

版权所有 © 2025 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。