



Union.aiとNetApp FlexCacheを使用したハイブリッドAIトレーニングの導入

NetApp artificial intelligence solutions

NetApp
December 04, 2025

目次

Union.aiとNetApp FlexCacheを使用したハイブリッドAIトレーニングの導入.....	1
概要	1
顧客ユースケース: ハイブリッドクラウド AI トレーニング	1
お客様のメリット	1
プラグインの有効化 – 前提条件	1
リファレンスアーキテクチャ	2
ステップ1: FlexCacheボリュームを作成する	2
ステップ2: <i>Trident</i> を構成する	2
ステップ3: <i>Union.ai</i> ワークフローをデプロイする	3
PVCからのデータをロードしてトレーニングする	5
ステップ4: 統合の検証	5
セキュリティに関する考慮事項	5
監視と最適化	5
関連リンク	6
まとめ	6
Union.ai - コンパニオンガイド	6

Union.aiとNetApp FlexCacheを使用したハイブリッドAIトレーニングの導入

NetApp FlexCacheとTrident for Kubernetesストレージプロビジョニングを使用したUnion.aiオーケストレーションを使用して、ハイブリッドAIトレーニング環境を展開する方法を学びます。

David Espejo 氏、Union.ai Sathish Thyagarajan 氏、NetApp

概要

Union.aiのハイブリッドオーケストレーションプラットフォームは、NetApp ONTAPおよびFlexCacheとシームレスに統合され、AI/MLトレーニングワークフローを加速します。このソリューションにより、AIトレーニングワークロードにクラウドベースのGPUコンピューティングを活用しながら、データをオンプレミスで安全に維持できます。NetApp FlexCacheは、必要なデータのみがクラウドにキャッシュされるようにし、効率的で安全かつスケーラブルなハイブリッドAI/MLパイプラインを実現します。

顧客ユースケース: ハイブリッドクラウドAIトレーニング

- オンプレミスデータ: コンプライアンスとセキュリティのためにNetApp ONTAPに保存されます。
- クラウドコンピューティング: EKS/GKE/AKSでのスケーラブルなGPUトレーニング。
- AI/MLオーケストレーション: Union.aiは、環境全体のデータ処理とトレーニングを調整します。
- ストレージプロビジョニング: NetApp TridentはPVC/PVプロビジョニングを自動化します。

お客様のメリット

- NetApp ONTAPのスケールアウト機能を使用して、大規模なデータセットでAIワークロードを実行します。
- NetAppのハイブリッドクラウド機能を使用して、オンプレミスとクラウド間でデータを移動および同期します。
- FlexCacheを使用してオンプレミスのデータをクラウドに素早くキャッシュします。
- Union.aiは、バージョン管理、系統追跡、成果物管理により、環境間のオーケストレーションを簡素化します。
- 機密データをオンプレミスに保持しながら、クラウドでトレーニングを実行します。

プラグインの有効化 – 前提条件

要件	詳細
ONTAPバージョン	ONTAP 9.7+ (FlexCacheライセンスは不要)
FlexCacheライセンス	ONTAP 9.6以前で必須
Kubernetes	オンプレミスおよびクラウドクラスター(EKS/GKE/AKS)

Trident	オンプレミスとクラウド クラスタの両方にインストール可能
ユニオン.ai	コントロール プレーンがデプロイされている (Union Cloud またはセルフホスト)
ネットワーク	クラスタ間接続 (ONTAPクラスタが分離されている場合)
権限	ONTAPおよびKubernetes クラスターへの管理者アクセス。 <input type="checkbox"/> 正しいONTAP認証情報 (例: vsadmin) を使用する
Union.ai は初めてですか?	このドキュメントの最後にあるガイドを参照してください

リファレンスアーキテクチャ

次の図は、ハイブリッド AI トレーニング用にNetAppストレージと統合された Union.ai コントロール プレーンを示しています。

[Union.ai とNetAppによるハイブリッド AI トレーニング アーキテクチャ] | *ai-hybrid-union-001.png*

- Union.ai コントロール プレーン: ワークフローをオーケストレーションし、データの移動を管理し、NetApp API と統合します。
- NetApp ONTAP + FlexCache: オンプレミスからクラウドへの効率的なデータ キャッシュを提供します。
- ハイブリッド トレーニング クラスター: トレーニング ジョブは、オンプレミスからキャッシュされたデータを使用して、クラウド K8s クラスター (EKS など) で実行されます。

ステップ1: FlexCacheボリュームを作成する

ONTAPシステムマネージャの使用

- [ストレージ] > [ボリューム] に移動します。
- [Add]をクリックします。
- [その他のオプション]を選択します。
- リモート ボリュームのキャッシュとして追加を有効にします。
- ソース (オンプレミス) ボリュームと宛先 (クラウド) ボリュームを選択します。
- QoS またはパフォーマンス レベルを定義します (オプション)。
- [Create]をクリックします。

権限またはアグリゲートの問題によりNetApp DataOps Toolkit が機能しない場合は、ONTAP System Manager または CLI を使用してFlexCacheボリュームを直接作成します。

ステップ2: Trident を構成する

両方のクラスターにTridentをインストールします。

"Tridentインストール ガイド"

Tridentバックエンドを作成する

```
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
```

401 Unauthorized エラーが発生した場合は、ONTAPユーザーに十分な API 権限があり、正しいユーザー名 (vsadmin) とパスワードが使用されていることを確認してください。

ストレージクラスの定義

```
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
```

ステップ3：Union.aiワークフローをデプロイする

Union は PVC を使用してFlexCacheボリュームをトレーニング ジョブにマウントします。

PodTemplateの例

```
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
```

ワークフローの例

ユニオンインポートタスク、ワークフローから

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

def train_model(pvc_path: str):
```

PVCからのデータをロードしてトレーニングする

```
@workflow

def training_pipeline():

    train_model(pvc_path="/data/flexcache")
```

ユニオンオペレーターは次のことを行います。

- PVCを作成する
- FlexCacheボリュームをマウントする
- クラウドK8sクラスターでジョブをスケジュールする

ステップ4: 統合の検証

タスク	検証
PVCマウント	トレーニングポッドは /data/flexcache を正常にマウントする必要があります
データ アクセス	トレーニングジョブはFlexCacheから読み書きできる
キャッシュの動作	ONTAPのキャッシュ ヒット/ミスを監視します。アグリゲートがFlexCacheをサポートしていることを確認する
パフォーマンス	トレーニングワークロードのレイテンシとスループットを検証する

パフォーマンスを監視するには、 NetApp BlueXPまたはONTAP CLI を使用します。

セキュリティに関する考慮事項

- FSx for NetApp ONTAPに VPC エンドポイントを使用する
- 転送中および保存中の暗号化を有効にする
- ONTAPアクセスに RBAC/IAM を適用する
- Union.aiは顧客データにアクセスしたり保存したりしません

監視と最適化

道具	目的
NetAppBlueXP	FlexCacheの使用状況とパフォーマンスを監視する
Union.ai UI	パイプラインのステータスとメトリクスを追跡する
Tridentログ	PVC またはバックエンドの問題をデバッgingする

オプションの機能強化

- BlueXP APIを使用してFlexCacheの作成を自動化する
- トレーニング前にUnion SDKを使用してキャッシュをウォームアップする
- トレーニング後にバッチ推論またはモデルサービングパイプラインを追加する
- DataOps Toolkit が失敗した場合は、System Manager 経由で手動でFlexCacheを作成します。

トラブルシューティング

問題	解決
PVCが保留中	Tridentのログとバックエンドの設定を確認する
401 ONTAP APIからの不正アクセス	vsadminを使用して権限を確認する
ジョブ失敗: 適切なストレージがありません	ONTAPアグリゲートがFlexCache/ FabricPoolをサポートしていることを確認する
トレーニングパフォーマンスの低下	キャッシュヒット率とネットワーク遅延を確認する
データが同期されない	ONTAPでFlexCache関係の健全性を検証する

次のステップ

- テストデータでFlexCacheを検証する
- Union.ai トレーニング パイプラインをデプロイする
- パフォーマンスを監視および最適化する
- 顧客固有の設定を文書化する

関連リンク

- "Union.ai ドキュメント"
- "NetApp FlexCache の概要"
- "TridentCSI ドライバー"
- "NetApp ONTAP 用 FSx"

まとめ

これで、Union.ai とNetApp FlexCacheを使用した検証済みのハイブリッド AI トレーニング環境が完成しました。トレーニング ジョブは、データセット全体を複製したりガバナンスを損なうことなく、オンプレミス データに安全かつ効率的にアクセスしながらクラウドで実行できます。

Union.ai - コンパニオンガイド

ステップ1: 導入モデルの選択

オプションA：ユニオンクラウド

- 訪問："コンソール.ユニオン.ai"

- ・組織を作成 → プロジェクトを作成

オプションB: セルフホスト

- ・フォローする：<https://docs.union.ai/platform/latest/deployment/self-hosted/>["セルフホストガイド"]
- ・Helm 経由でデプロイ：

helmリポジトリにunionaiを追加 <https://unionai.github.io/helm-charts/>

```
helmインストール union unionai/union -n union-system -f values.yaml
```

ステップ2: Union Operatorをインストールする

□kubectl を適用 -f<https://raw.githubusercontent.com/unionai/operator/main/deploy/operator.yaml>[]

kubectl get pods -n ユニオンシステム

□

ステップ3: Union CLIをインストールする

□pip インストール unionai

組合ログイン

□

ステップ4: ワークフローを登録する

□ユニオンプロジェクトがハイブリッドAIを作成

ユニオンレジスタ training_pipeline.py --project hybrid-ai

□

ステップ5: 実行と監視

□union run training_pipeline --project hybrid-ai

ユニオンウォッチトレーニングパイプライン

□ログを表示する "ユニオンUI"

ステップ6: コンピューティングクラスターの登録(オプション)

□ ユニオンクラスターレジスタ --name cloud-k8s --kubeconfig ~/.kube/config

ステップ7: アーティファクトと系統を追跡する

Unionは自動的に次のものを追跡します:

- ・入力/出力パラメータ

- ・データバージョン
- ・ログとメトリクス
- ・実行系統

著作権に関する情報

Copyright © 2025 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を隨時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5225.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。