



MLflow

NetApp Solutions

NetApp
August 14, 2024

目次

MLflow	1
MLflowの展開	1
NetAppおよびMLflowによるデータセットからモデルへのトレーサビリティ	2

MLflow

MLflowの展開

このセクションでは、KubernetesクラスタにMLflowを導入するために完了する必要があるタスクについて説明します。



MLflowはKubernetes以外のプラットフォームに導入することも可能です。Kubernetes以外のプラットフォームへのMLflowの導入は、このソリューションの範囲外です。

前提条件

このセクションで説明する導入の演習を行う前に、次の作業をすでに実行していることを前提としています。

1. Kubernetes クラスタをすでに使用している。
2. KubernetesクラスタにNetApp Astra Tridentをインストールして設定しておきます。Astra Tridentの詳細については、を参照してください "[Astra Trident のドキュメント](#)"。

Helm をインストールします

MLflowは、Kubernetesで人気のあるパッケージマネージャであるHelmを使用して導入されます。MLflowを導入する前に、Kubernetes制御ノードにHelmをインストールする必要があります。Helmをインストールするには、公式のHelmドキュメントのを実行し "[インストール手順](#)" ます。

デフォルトの Kubernetes StorageClass を設定します

MLflowを導入する前に、Kubernetesクラスタ内でデフォルトのStorageClassを指定する必要があります。クラスタ内でデフォルトのストレージクラスを指定するには、の手順に従ってください "[Kubeflow の導入](#)"。クラスタ内ですでにデフォルトの StorageClass を指定している場合は、この手順を省略できます。

MLflowの展開

前提条件が満たされたら、Helmチャートを使用してMLflowの導入を開始できます。

MLflow Helmチャート展開を設定します。

Helmチャートを使用してMLflowを導入する前に、NetApp Tridentストレージクラスを使用するように導入を構成し、* config.yaml *ファイルを使用して他のパラメータをニーズに合わせて変更できます。config.yaml*ファイルの例は、次の場所にあります。 <https://github.com/bitnami/charts/blob/main/bitnami/mlflow/values.yaml>



Trident storageClassは、config.yamlファイルの* global.defaultStorageClass*パラメータで設定できます（例：storageClass："ONTAP-FlexVol"）。

Helmチャートのインストール

次のコマンドを使用して、HelmチャートをMLflow用のカスタム* config.yaml *ファイルとともにインストールできます。

```
helm install oci://registry-1.docker.io/bitnamicharts/mlflow -f
config.yaml --generate-name --namespace jupyterhub
```



このコマンドは、提供されている* config.yaml *ファイルを使用して、カスタム構成内のKubernetesクラスタにMLflowをデプロイします。MLflowは指定された名前空間にデプロイされ、リリースのためにKubernetes経由でランダムなリリース名が与えられます。

展開を確認

Helmチャートの導入が完了したら、次のコマンドを使用してサービスにアクセスできるかどうかを確認できます。

```
kubectl get service -n jupyterhub
```



*jupyterhub*を、導入時に使用した名前空間に置き換えます。

次のサービスが表示されます。

NAME	TYPE	CLUSTER-IP	EXTERNAL-IP
mlflow-1719843029-minio	ClusterIP	10.233.22.4	<none>
80/TCP, 9001/TCP			
mlflow-1719843029-postgresql	ClusterIP	10.233.5.141	<none>
5432/TCP			
mlflow-1719843029-postgresql-hl	ClusterIP	None	<none>
5432/TCP			
mlflow-1719843029-tracking	NodePort	10.233.2.158	<none>
30002:30002/TCP			



config.yamlファイルを編集し、NodePortサービスを使用してポート30002のMLflowにアクセスしました。

アクセスMLflow

MLflowに関連するすべてのサービスが起動して実行されたら、指定したNodePortまたはLoadBalancer IPアドレス（例：）を使用してMLflowにアクセスできます <http://10.61.181.109:30002>。

NetAppおよびMLflowによるデータセットからモデルへのトレーサビリティ

は、"Kubernetes向けNetApp DataOpsツールキット" データセットからモデル、またはワークスペースからモデルへのトレーサビリティを実装するために、MLflowの実験追跡

機能と組み合わせて使用できます。

データセット間、またはワークスペース間のトレーサビリティを実装するには、次のコード例に示すように、DataOps Toolkitをトレーニングの一部として使用して、データセットまたはワークスペースボリュームのスナップショットを作成します。このコードは、データボリューム名とスナップショット名を、MLflow実験追跡サーバーにロギングする特定のトレーニング実行に関連付けられたタグとして保存します。

```
...
from netapp_dataops.k8s import create_volume_snapshot

with mlflow.start_run() :
    ...

    namespace = "my_namespace" # Kubernetes namespace in which dataset
    volume PVC resides
    dataset_volume_name = "project1" # Name of PVC corresponding to
    dataset volume
    snapshot_name = "run1" # Name to assign to your new snapshot

    # Create snapshot
    create_volume_snapshot(
        namespace=namespace,
        pvc_name=dataset_volume_name,
        snapshot_name=snapshot_name,
        printOutput=True
    )

    # Log data volume name and snapshot name as "tags"
    # associated with this training run in mlflow.
    mlflow.set_tag("data_volume_name", dataset_volume_name)
    mlflow.set_tag("snapshot_name", snapshot_name)

    ...
```

著作権に関する情報

Copyright © 2024 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随時、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 5252.227-19（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权を有し、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权については、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。