



## **MLflow**

### **NetApp Solutions**

NetApp  
August 14, 2024

# 目录

MLflow .....	1
MLflow部署 .....	1
使用NetApp和MLflow实现数据集到模型的可跟踪性 .....	2

# MLflow

## MLflow部署

本节介绍在Kubernetes集群中部署MLflow所需完成的任务。



可以在Kubnetes以外的平台上部署MLflow。在Kubernetes以外的平台上部署MLflow不在此解决方案讨论范围之内。

### 前提条件

在执行本节所述的部署练习之前，我们假定您已执行以下任务：

1. 您已有一个工作正常的 Kubernetes 集群。
2. 您已在Kubernetes集群中安装并配置NetApp Asta三端存储。有关Astra Trident的详细信息，请参阅 "[Astra Trident 文档](#)"。

### 安装 Helm

MLflow使用Helm进行部署、Helm是Kubernetes常用的软件包管理器。在部署MLflow之前、您必须在Kubernetes控制节点上安装Helm。要安装Helm、请按照官方Helm文档中的进行操作 "[安装说明](#)"。

### 设置默认 Kubernetes StorageClass

在部署MLflow之前、您必须在Kubernetes集群中指定一个默认StorageClass。要在集群中指定默认StorageClass、请按照一节中所述的说明进行操作 "[Kubeflow 部署](#)"。如果已在集群中指定默认 StorageClass，则可以跳过此步骤。

### 部署MLflow

满足前提条件后、您可以使用Helm图表开始MLflow部署。

配置MLflow Helm Chart部署。

在使用Helm图表部署MLflow之前，我们可以使用\*config.yaml\*文件将部署配置为使用NetApp Trident存储类并根据需要更改其他参数。有关\*config.yaml\*文件的示例，请访问：<https://github.com/bitnami/charts/blob/main/bitnami/mlflow/values.yaml>



您可以在config.yaml文件中的\*. global.defaultStorageClass\*参数下设置Trident存储类(例如、存储类: "ONTAP FlexVol ")。

### 安装Helm图表

使用以下命令，可以将Helm图表与MLflow的自定义\*config.yaml\*文件一起安装：

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



命令通过提供的\*config.yaml\*文件在自定义配置中的Kubernetes集群上部署MLflow。MLflow部署在给定命名空间中、并通过Kubernetes为此版本随机指定版本名称。

## 检查部署

完成Helm图表部署后、您可以使用以下命令检查此服务是否可访问：

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



将\*jupyterhub\*替换为部署期间使用的命名空间。

您应看到以下服务：

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			



我们编辑了config.yaml文件以使用NodePort服务访问端口30002上的MLflow。

## 访问MLflow

当所有与MLflow相关的服务都启动并运行后，您可以使用给定的nodePort或负载均衡器IP地址(例如)访问它<http://10.61.181.109:30002>。

## 使用NetApp和MLflow实现数据集到模型的可跟踪性

```
https://github.com/NetApp/netapp-dataops-
toolkit/tree/main/netapp_dataops_k8s["适用于Kubernetes的NetApp DataOps
工具包"^]可与MLflow的实验跟踪功能结合使用、以实现数据集到模型或工作空间到模型的可追溯性
。
```

要实现数据集到模型或工作空间到模型的可追溯性、只需在训练过程中使用DataOps工具包创建数据集或工作空间卷的快照、如以下示例代码片段所示。此代码会将数据卷名称和快照名称保存为与您要记录到MLflow实验跟踪服务器的特定训练运行关联的标记。

```
...
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)

...
```

## 版权信息

版权所有 © 2024 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。