



## **MLflow**

### **NetApp Solutions**

NetApp  
August 14, 2024

# 目錄

MLflow .....	1
MLflow 部署 .....	1
資料集對模型的追蹤性、可透過 NetApp 和 MLflow 進行 .....	2

# MLflow

## MLflow 部署

本節說明在 Kubernetes 叢集中部署 MLflow 時必須完成的工作。



您可以在 Kubernetes 以外的平台上部署 MLflow。在 Kubernetes 以外的平台上部署 MLflow 不屬於本解決方案的範圍。

### 先決條件

在您執行本節所述的部署練習之前、我們假設您已經執行下列工作：

1. 您已經擁有有效的Kubernetes叢集。
2. 您已在 Kubernetes 叢集中安裝並設定 NetApp Astra Trident。如需 Astra Trident 的詳細資訊、請 "[Astra Trident文件](#)"參閱。

### 安裝Helm

MLflow 是使用知名的 Kubernetes 套件管理程式 Helm 來部署。在部署 MLflow 之前、您必須先在 Kubernetes 控制節點上安裝 Helm。若要安裝 Helm、請遵循 "[安裝說明](#)" 正式 Helm 文件中的。

### 設定預設Kubernetes StorageClass

在部署 MLflow 之前、您必須在 Kubernetes 叢集中指定預設 StorageClass。若要在叢集中指定預設 StorageClass、請遵循一節中所述的指示 "[Kubeflow部署](#)" 進行。如果您已在叢集內指定預設StorageClass、則可以跳過此步驟。

### 部署 MLflow

一旦符合先決條件、您就可以使用船舵圖表開始進行 MLflow 部署。

設定 **MLflow Helm** 圖表部署。

在使用 Helm 圖表部署 MLflow 之前、我們可以設定部署以使用 NetApp Trident 儲存類別、並使用 **config.yaml** 檔案變更其他參數以符合我們的需求。\* 組態 .yaml\* 檔案的範例可在以下網址找到：<https://github.com/bitnami/charts/blob/main/bitnami/mlflow/values.yaml>



您可以在 config.yaml 檔案中的 \* global.defaultStorageClass\* 參數下設定 Trident storageClass (例如 storageClass : 「 ontp-FlexVol 」)。

### 安裝 Helm 圖表

使用下列命令、可以使用 MLflow 的自訂 \* 組態 .yaml\* 檔案來安裝 Helm 圖表：

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



此命令會透過提供的 \* config.yaml\* 檔案、在自訂組態的 Kubernetes 叢集上部署 MLflow  
◦ MLflow 會部署在指定的命名空間中、並透過 Kubernetes 為發行提供隨機版本名稱。

## 檢查部署

在完成 Helm 圖表的部署之後、您可以使用下列項目來檢查服務是否可供存取：

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



將 **jupyterhub** 替換為部署期間使用的命名空間。

您應該會看到下列服務：

NAME	TYPE	CLUSTER-IP	EXTERNAL-IP
PORT(S)            AGE			
mlflow-1719843029-minio 80/TCP,9001/TCP    25d	ClusterIP	10.233.22.4	<none>
mlflow-1719843029-postgresql 5432/TCP            25d	ClusterIP	10.233.5.141	<none>
mlflow-1719843029-postgresql-hl 5432/TCP            25d	ClusterIP	None	<none>
mlflow-1719843029-tracking 30002:30002/TCP    25d	NodePort	10.233.2.158	<none>



我們編輯 config.yaml 檔案、使用 NodePort 服務存取連接埠 30002 上的 MLflow。

## 存取 MLflow

當所有與 MLflow 相關的服務都啟動並執行後、您就可以使用指定的 NodePort 或負載平衡器 IP 位址（例如）來存取它。<http://10.61.181.109:30002>

## 資料集對模型的追蹤性、可透過 NetApp 和 MLflow 進行

```
https://github.com/NetApp/netapp-dataops-
toolkit/tree/main/netapp_dataops_k8s["適用於Kubernetes的NetApp
DataOps工具套件"^]可搭配 MLflow
的實驗追蹤功能使用、以實作資料集對模型或工作區對模型的追蹤。
```

若要實作資料集對模型或工作區對模型的可追蹤性、只需在訓練過程中使用 DataOps Toolkit 建立資料集或工作區 Volume 的快照、如下所示的程式碼片段範例所示。此程式碼會將資料磁碟區名稱和快照名稱儲存為與您登入 MLflow 實驗追蹤伺服器的特定訓練路跑相關的標籤。

```
...
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)

    ...
```

## 版權資訊

Copyright © 2024 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變更本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

## 商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。