



주피터허브 NetApp Solutions

NetApp
August 14, 2024

목차

| | |
|--|---|
| 주피터허브 | 1 |
| JupyterHub 배포 | 1 |
| JupyterHub와 함께 NetApp DataOps 툴킷을 사용할 수 있습니다 | 3 |

주피터허브

JupyterHub 배포

이 섹션에서는 Kubernetes 클러스터에 JupyterHub를 구축하기 위해 완료해야 하는 작업에 대해 설명합니다.



Kubernetes 이외의 플랫폼에 JupyterHub를 배포할 수 있습니다. Kubernetes 이외의 플랫폼에 JupyterHub를 구축하는 것은 이 솔루션의 범위를 벗어납니다.

필수 구성 요소

이 섹션에 요약된 배포 연습을 수행하기 전에 이미 다음 작업을 수행했다고 가정합니다.

1. Kubernetes 클러스터 작업이 이미 진행 중입니다.
2. Kubernetes 클러스터에 NetApp Astra Trident를 이미 설치하고 구성했습니다. Astra Trident에 대한 자세한 내용은 ["Astra Trident 문서"](#) 참조하십시오.

제어 장치를 설치합니다

JupyterHub는 Kubernetes의 인기 패키지 매니저인 Helm을 사용하여 배포됩니다. JupyterHub를 배포하기 전에 Kubernetes 제어 노드에 Helm을 설치해야 합니다. Helm을 설치하려면 공식 Helm 설명서의 ["설치 지침"](#) 참조하십시오.

기본 Kubernetes StorageClass를 설정합니다

JupyterHub를 배포하기 전에 Kubernetes 클러스터 내에 기본 StorageClass를 지정해야 합니다. 클러스터 내에서 기본 StorageClass를 지정하려면 섹션에 설명된 지침을 ["Kubeflow 구축"](#) 따릅니다. 클러스터 내에서 기본 StorageClass를 이미 지정한 경우에는 이 단계를 건너뛸 수 있습니다.

JupyterHub를 배포합니다

위의 단계를 완료한 후 JupyterHub를 배포할 준비가 되었습니다. JupyterHub 배포에는 다음 단계가 필요합니다.

JupyterHub 배포 구성

배포하기 전에 각 환경에 맞게 JupyterHub 배포를 최적화하는 것이 좋습니다. config.yaml* 파일을 만들고 Helm 차트를 사용하여 배포 중에 활용할 수 있습니다.

예제 * config.yaml * 파일은 [여기](https://github.com/jupyterhub/zero-to-jupyterhub-k8s/blob/HEAD/jupyterhub/values.yaml)에서 찾을 수 있습니다 <https://github.com/jupyterhub/zero-to-jupyterhub-k8s/blob/HEAD/jupyterhub/values.yaml>



이 config.yaml 파일에서 NetApp Trident StorageClass에 대한 * (singleuser.storage.dynamic.storageClass) * 매개 변수를 설정할 수 있습니다. 개별 사용자 작업 공간에 대한 볼륨을 프로비저닝하는 데 사용되는 스토리지 클래스입니다.

공유 볼륨을 추가하는 중입니다

모든 JupyterHub 사용자에게 대해 공유 볼륨을 사용하려면 * config.yaml * 을 적절하게 조정할 수 있습니다. 예를 들어 jupyterhub-shared-volume이라는 공유 PersistentVolumeClaim이 있는 경우 다음과 같이 모든 사용자 포드에서 /home/shared로 마운트할 수 있습니다.

```
singleuser:
  storage:
    extraVolumes:
      - name: jupyterhub-shared
        persistentVolumeClaim:
          claimName: jupyterhub-shared-volume
    extraVolumeMounts:
      - name: jupyterhub-shared
        mountPath: /home/shared
```



이 단계는 선택 사항이므로 필요에 따라 이러한 매개변수를 조정할 수 있습니다.

Helm Chart로 JupyterHub를 배포합니다

Helm이 JupyterHub Helm 차트 리포지토리를 인식하도록 합니다.

```
helm repo add jupyterhub https://hub.jupyter.org/helm-chart/
helm repo update
```

다음과 같이 출력이 표시됩니다.

```
Hang tight while we grab the latest from your chart repositories...
...Skip local chart repository
...Successfully got an update from the "stable" chart repository
...Successfully got an update from the "jupyterhub" chart repository
Update Complete. ☐ Happy Helming!☐
```

이제 config.yaml이 포함된 디렉터리에서 이 명령을 실행하여 config.yaml로 구성된 차트를 설치합니다.

```
helm upgrade --cleanup-on-fail \
  --install my-jupyterhub jupyterhub/jupyterhub \
  --namespace my-namespace \
  --create-namespace \
  --values config.yaml
```



이 예에서

<helm-release-name>는 내-jupyterhub로 설정되어 있으며, 이것은 당신의 JupyterHub 릴리스의 이름입니다. <k8s-namespace> 는 JupyterHub 를 설치할 네임스페이스인 my-namespace 로 설정됩니다. --create-namespace 플래그는 네임스페이스가 없는 경우 이 네임스페이스를 만드는 데 사용됩니다. 값 플래그는 원하는 구성 옵션이 포함된 config.yaml 파일을 지정합니다.

배포를 확인합니다

2단계가 실행되는 동안 다음 명령을 통해 Pod가 생성되는 것을 볼 수 있습니다.

```
kubectl get pod --namespace <k8s-namespace>
```

허브 및 프록시 포드가 실행 상태로 전환될 때까지 기다립니다.

| NAME | READY | STATUS | RESTARTS | AGE |
|-----------------------|-------|---------|----------|-----|
| hub-5d4ffd57cf-k68z8 | 1/1 | Running | 0 | 37s |
| proxy-7cb9bc4cc-9bdlp | 1/1 | Running | 0 | 37s |

JupyterHub에 액세스합니다

JupyterHub에 액세스하는 데 사용할 수 있는 IP를 찾습니다. 예제 출력과 같이 프록시 공용 서비스의 external-IP를 사용할 수 있을 때까지 다음 명령을 실행합니다.



우리는 우리의 config.yaml 파일에서 NodePort 서비스를 사용, 당신은 당신의 설정에 따라 환경에 맞게 조정 할 수 있습니다 (예: 로드 밸런서).

```
kubectl --namespace <k8s-namespace> get service proxy-public
```

| NAME | TYPE | CLUSTER-IP | EXTERNAL-IP | PORT(S) |
|--------------|----------|---------------|---------------|--------------|
| proxy-public | NodePort | 10.51.248.230 | 104.196.41.97 | 80:30000/TCP |

JupyterHub를 사용하려면 브라우저에 프록시 공용 서비스의 외부 IP를 입력합니다.

JupyterHub와 함께 NetApp DataOps 툴킷을 사용할 수 있습니다

는 "Kubernetes용 NetApp DataOps 툴킷" JupyterHub 와 함께 사용할 수 있습니다. 최종 사용자는 NetApp DataOps Toolkit을 JupyterHub와 함께 사용하여 Jupyter Notebook 내에서 직접 작업 공간 백업 및/또는 데이터 세트-모델 추적 기능을 위한 볼륨 스냅샷을 생성할 수 있습니다.

초기 설정

DataOps Toolkit을 JupyterHub와 함께 사용하려면 먼저 JupyterHub가 개별 사용자 Jupyter Notebook Server Pod에 할당하는 Kubernetes 서비스 계정에 적절한 권한을 부여해야 합니다. JupyterHub는 JupyterHub Helm 차트 구성 파일의 변수에 의해 지정된 서비스 계정을 `singleuser.serviceAccountName` 사용합니다.

DataOps 툴킷에 대한 클러스터 역할 생성

먼저, 볼륨 스냅샷을 생성하는 데 필요한 Kubernetes API 권한이 있는 'netapp-dataops'라는 클러스터 역할을 생성합니다.

```
$ vi clusterrole-netapp-dataops-snapshots.yaml
---
kind: ClusterRole
apiVersion: rbac.authorization.k8s.io/v1
metadata:
  name: netapp-dataops-snapshots
rules:
- apiGroups: [""]
  resources: ["persistentvolumeclaims", "persistentvolumeclaims/status",
"services"]
  verbs: ["get", "list"]
- apiGroups: ["snapshot.storage.k8s.io"]
  resources: ["volumesnapshots", "volumesnapshots/status",
"volumesnapshotcontents", "volumesnapshotcontents/status"]
  verbs: ["get", "list", "create"]

$ kubectl create -f clusterrole-netapp-dataops-snapshots.yaml
clusterrole.rbac.authorization.k8s.io/netapp-dataops-snapshots created
```

노트북 서버 서비스 계정에 클러스터 역할을 할당합니다

'netapp-dataops-snapshots' 클러스터 역할을 해당 네임스페이스의 적절한 서비스 계정에 할당하는 역할 바인딩을 생성합니다. 예를 들어 'jupyterhub' 네임스페이스에 JupyterHub를 설치하고 변수를 통해 'default' 서비스 계정을 지정한 경우 `singleuser.serviceAccountName` 다음 예와 같이 'jupyterhub' 네임스페이스의 'default' 서비스 계정에 'netapp-dataops-snapshots' 클러스터 역할을 할당합니다.

```
$ vi rolebinding-jupyterhub-netapp-dataops-snapshots.yaml
---
kind: RoleBinding
apiVersion: rbac.authorization.k8s.io/v1
metadata:
  name: jupyterhub-netapp-dataops-snapshots
  namespace: jupyterhub # Replace with you JupyterHub namespace
subjects:
- kind: ServiceAccount
  name: default # Replace with your JupyterHub
singleuser.serviceAccountName
  namespace: jupyterhub # Replace with you JupyterHub namespace
roleRef:
  kind: ClusterRole
  name: netapp-dataops-snapshots
  apiGroup: rbac.authorization.k8s.io

$ kubectl create -f ./rolebinding-jupyterhub-netapp-dataops-snapshots.yaml
rolebinding.rbac.authorization.k8s.io/jupyterhub-netapp-dataops-snapshots
created
```

Jupyter Notebook 내에서 볼륨 스냅샷을 생성합니다

이제 JupyterHub 사용자는 NetApp DataOps Toolkit을 사용하여 다음 예에 표시된 대로 Jupyter 노트북 내에서 직접 볼륨 스냅샷을 생성할 수 있습니다.

Execute NetApp DataOps Toolkit operations within JupyterHub

This notebook demonstrates the execution of NetApp DataOps Toolkit operations from within a Jupyter Notebook running on JupyterHub

Install NetApp DataOps Toolkit for Kubernetes (only run once)

Note: This cell only needs to be run once. This is a one-time task

```
[ ]: %pip install --user netapp-dataops-k8s
```

Import NetApp DataOps Toolkit for Kubernetes functions

```
[1]: from netapp_dataops.k8s import list_volumes, list_volume_snapshots, create_volume_snapshot
```

Create Volume Snapshot for User Workspace Volume

The following example shows the execution of a "create volume snapshot" operation for my user workspace volume.

```
[2]: jupyterhub_namespace = "jupyterhub"
my_user_workspace_vol = "claim-moglesby"

create_volume_snapshot(namespace=jupyterhub_namespace, pvc_name=my_user_workspace_vol, print_output=True)

Creating VolumeSnapshot 'ntap-dsutil.20240726002955' for PersistentVolumeClaim (PVC) 'claim-moglesby' in namespace 'jupyterhub'.
VolumeSnapshot 'ntap-dsutil.20240726002955' created. Waiting for Trident to create snapshot on backing storage.
Snapshot successfully created.
```

저작권 정보

Copyright © 2024 NetApp, Inc. All Rights Reserved. 미국에서 인쇄된 본 문서의 어떠한 부분도 저작권 소유자의 사전 서면 승인 없이는 어떠한 형식이나 수단(복사, 녹음, 녹화 또는 전자 검색 시스템에 저장하는 것을 비롯한 그래픽, 전자적 또는 기계적 방법)으로도 복제될 수 없습니다.

NetApp이 저작권을 가진 자료에 있는 소프트웨어에는 아래의 라이선스와 고지사항이 적용됩니다.

본 소프트웨어는 NetApp에 의해 '있는 그대로' 제공되며 상품성 및 특정 목적에의 적합성에 대한 명시적 또는 묵시적 보증을 포함하여(이에 제한되지 않음) 어떠한 보증도 하지 않습니다. NetApp은 대체품 또는 대체 서비스의 조달, 사용 불능, 데이터 손실, 이익 손실, 영업 중단을 포함하여(이에 국한되지 않음), 이 소프트웨어의 사용으로 인해 발생하는 모든 직접 및 간접 손해, 우발적 손해, 특별 손해, 징벌적 손해, 결과적 손해의 발생에 대하여 그 발생 이유, 책임론, 계약 여부, 엄격한 책임, 불법 행위(과실 또는 그렇지 않은 경우)와 관계없이 어떠한 책임도 지지 않으며, 이와 같은 손실의 발생 가능성이 통지되었다 하더라도 마찬가지입니다.

NetApp은 본 문서에 설명된 제품을 언제든지 예고 없이 변경할 권리를 보유합니다. NetApp은 NetApp의 명시적인 서면 동의를 받은 경우를 제외하고 본 문서에 설명된 제품을 사용하여 발생하는 어떠한 문제에도 책임을 지지 않습니다. 본 제품의 사용 또는 구매의 경우 NetApp에서는 어떠한 특허권, 상표권 또는 기타 지적 재산권이 적용되는 라이선스도 제공하지 않습니다.

본 설명서에 설명된 제품은 하나 이상의 미국 특허, 해외 특허 또는 출원 중인 특허로 보호됩니다.

제한적 권리 표시: 정부에 의한 사용, 복제 또는 공개에는 DFARS 252.227-7013(2014년 2월) 및 FAR 52.227-19(2007년 12월)의 기술 데이터-비상업적 품목에 대한 권리(Rights in Technical Data -Noncommercial Items) 조항의 하위 조항 (b)(3)에 설명된 제한사항이 적용됩니다.

여기에 포함된 데이터는 상업용 제품 및/또는 상업용 서비스(FAR 2.101에 정의)에 해당하며 NetApp, Inc.의 독점 자산입니다. 본 계약에 따라 제공되는 모든 NetApp 기술 데이터 및 컴퓨터 소프트웨어는 본질적으로 상업용이며 개인 비용만으로 개발되었습니다. 미국 정부는 데이터가 제공된 미국 계약과 관련하여 해당 계약을 지원하는 데에만 데이터에 대한 전 세계적으로 비독점적이고 양도할 수 없으며 재사용이 불가능하며 취소 불가능한 라이선스를 제한적으로 가집니다. 여기에 제공된 경우를 제외하고 NetApp, Inc.의 사전 서면 승인 없이는 이 데이터를 사용, 공개, 재생산, 수정, 수행 또는 표시할 수 없습니다. 미국 국방부에 대한 정부 라이선스는 DFARS 조항 252.227-7015(b)(2014년 2월)에 명시된 권한으로 제한됩니다.

상표 정보

NETAPP, NETAPP 로고 및 <http://www.netapp.com/TM>에 나열된 마크는 NetApp, Inc.의 상표입니다. 기타 회사 및 제품 이름은 해당 소유자의 상표일 수 있습니다.