



JupyterHub

NetApp Solutions

NetApp
August 14, 2024

This PDF was generated from https://docs.netapp.com/zh-cn/netapp-solutions/ai/aicp_jupyterhub_deployment.html on August 14, 2024. Always check docs.netapp.com for the latest.

目录

JupyterHub	1
JupyterHub部署	1
将NetApp数据操作工具包与JupyterHub结合使用	3

JupyterHub

JupyterHub部署

本节介绍在Kubernetes集群中部署JupyterHub必须完成的任务。



可以在Kubernetes以外的平台上部署JupyterHub。在Kubernetes以外的平台上部署JupyterHub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 Astra三端存储。有关Astra Trident的详细信息，请参阅 "[Astra Trident 文档](#)"。

安装 Helm

JupyterHub是使用Helm部署的、Helm是Kubernetes常用的软件包管理器。在部署JupyterHub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JupyterHub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JupyterHub

完成上述步骤后、即可开始部署JupyterHub。JupyterHub部署需要执行以下步骤：

配置JupyterHub部署

在部署之前、最好针对您的环境优化JupyterHub部署。您可以创建一个*config.yaml文件，并在部署期间使用Helm图表来使用它。

可以在中找到*config.yaml文件的示例 <https://github.com/jupyterhub/zero-to-jupyterhub-k8s/blob/HEAD/jupyterhub/values.yaml>



在此config.yaml文件中，您可以为NetApp Trident存储类设置*(singleuser.storage.dynamic . storageClass)*参数。此存储类将用于为各个用户工作空间配置卷。

正在添加共享卷

如果要对所有JupyterHub用户使用共享卷，可以相应地调整*config.yaml。例如、如果您有一个名为jupyterhub-shared-volume的共享持久卷、则可以在所有用户Pod中将其挂载为/home / shared、如下所示：

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



这是一个可选步骤、您可以根据需要调整这些参数。

使用Helm Chart部署JupyterHub

使Helm了解JupyterHub Helm图表存储库。

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

此命令显示的输出应如下所示：

```
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!☐
```

现在、通过从包含config.yaml的目录运行以下命令来安装由config.yaml配置的图表：

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



在此示例中：

<helm-release-name>已设置为my-jupyterhub、此名称将是JupyterHub版本的名称。<k8s-namespace>设置为my-命名空间、即要安装JupyterHub的命名空间。如果命名空间尚不存在、则使用—cree-命名空间标志创建命名空间。--values"标志指定包含所需配置选项的config.yaml文件。

检查部署

在运行第2步时、您可以使用以下命令查看正在创建的Pod：

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

等待集线器和代理Pod进入运行状态。

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

访问JupyterHub

查找可用于访问JupyterHub的IP。运行以下命令、直到proxy-public服务的外部IP可用、如示例输出中所示。



我们在config.yaml文件中使用了节点端口服务、您可以根据您的设置(例如负载均衡器)调整您的环境。

```
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
AGE				
1m				

要使用JupyterHub、请在浏览器中输入代理公共服务的外部IP。

将NetApp数据操作工具包与JupyterHub结合使用

```
https://github.com/NetApp/netapp-dataops-toolkit/tree/main/netapp_dataops_k8s["适用于Kubernetes的NetApp DataOps工具包"]可以与JupyterHub结合使用。通过将NetApp数据操作工具包与JupyterHub结合使用、最终用户可以直接从Jupyter笔记本电脑中创建卷快照、以实现工作空间备份和/或数据集到模型的可追溯性。
```

初始设置

在将DataOps工具包与JupyterHub结合使用之前、您必须为JupyterHub分配给单个用户Jupyter笔记本服务器Pod的Kubernetes服务帐户授予适当的权限。JupyterHub将使用JupyterHub Helm图表配置文件中的变量指定的服务帐户 `singleuser.serviceAccountName`。

为DataOps工具包创建集群角色

首先、创建一个名为"netapp-dataops"的集群角色、该角色具有创建卷快照所需的Kubernetes API权限。

```
$ 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
```

将集群角色分配给笔记本服务器服务帐户

创建一个角色绑定、将"NetApp-dataops-Snaps"集群角色分配给相应命名空间中的相应服务帐户。例如、如果您将JupyterHub安装在"jupyterhub"命名空间中、并通过变量指定了"default"服务帐户
singleuser.serviceAccountName、则应将"NetApp-dataops-Snapshots"集群角色分配给"jupyterhub"命名空间中的"default"服务帐户、如以下示例所示。

```

$ 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

```

在Jupyter笔记本中创建卷快照

现在、JupyterHub用户可以使用NetApp数据操作工具包直接从Jupyter笔记本中创建卷快照、如以下示例所示。

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.
```

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。