



Kubeflow Deployment

NetApp Solutions

NetApp
October 20, 2023

This PDF was generated from https://docs.netapp.com/us-en/netapp-solutions/ai/aicp_kubeflow_deployment_overview.html on October 20, 2023. Always check docs.netapp.com for the latest.

Table of Contents

- Kubeflow Deployment 1
 - Prerequisites 1
 - Set Default Kubernetes StorageClass 1
 - Use NVIDIA DeepOps to Deploy Kubeflow 2

Kubeflow Deployment

This section describes the tasks that you must complete to deploy Kubeflow in your Kubernetes cluster.

Prerequisites

Before you perform the deployment exercise that is outlined in this section, we assume that you have already performed the following tasks:

1. You already have a working Kubernetes cluster, and you are running a version of Kubernetes that is supported by Kubeflow. For a list of supported versions, see the [official Kubeflow documentation](#).
2. You have already installed and configured NetApp Trident in your Kubernetes cluster as outlined in [Trident Deployment and Configuration](#).

Set Default Kubernetes StorageClass

Before you deploy Kubeflow, you must designate a default StorageClass within your Kubernetes cluster. The Kubeflow deployment process attempts to provision new persistent volumes using the default StorageClass. If no StorageClass is designated as the default StorageClass, then the deployment fails. To designate a default StorageClass within your cluster, perform the following task from the deployment jump host. If you have already designated a default StorageClass within your cluster, then you can skip this step.

1. Designate one of your existing StorageClasses as the default StorageClass. The example commands that follow show the designation of a StorageClass named `ontap-ai- flexvols-retain` as the default StorageClass.



The `ontap-nas-flexgroup` Trident Backend type has a minimum PVC size that is fairly large. By default, Kubeflow attempts to provision PVCs that are only a few GBs in size. Therefore, you should not designate a StorageClass that utilizes the `ontap-nas-flexgroup` Backend type as the default StorageClass for the purposes of Kubeflow deployment.

```
$ kubectl get sc
NAME                                PROVISIONER                        AGE
ontap-ai-flexgroups-retain         csi.trident.netapp.io            25h
ontap-ai-flexgroups-retain-iface1  csi.trident.netapp.io            25h
ontap-ai-flexgroups-retain-iface2  csi.trident.netapp.io            25h
ontap-ai-flexvols-retain           csi.trident.netapp.io            3s
$ kubectl patch storageclass ontap-ai-flexvols-retain -p '{"metadata":
{"annotations":{"storageclass.kubernetes.io/is-default-class":"true"}}}'
storageclass.storage.k8s.io/ontap-ai-flexvols-retain patched
$ kubectl get sc
NAME                                PROVISIONER                        AGE
ontap-ai-flexgroups-retain         csi.trident.netapp.io            25h
ontap-ai-flexgroups-retain-iface1  csi.trident.netapp.io            25h
ontap-ai-flexgroups-retain-iface2  csi.trident.netapp.io            25h
ontap-ai-flexvols-retain (default) csi.trident.netapp.io            54s
```

Use NVIDIA DeepOps to Deploy Kubeflow

NetApp recommends using the Kubeflow deployment tool that is provided by NVIDIA DeepOps. To deploy Kubeflow in your Kubernetes cluster using the DeepOps deployment tool, perform the following tasks from the deployment jump host.



Alternatively, you can deploy Kubeflow manually by following the [installation instructions](#) in the official Kubeflow documentation

1. Deploy Kubeflow in your cluster by following the [Kubeflow deployment instructions](#) on the NVIDIA DeepOps GitHub site.
2. Note down the Kubeflow Dashboard URL that the DeepOps Kubeflow deployment tool outputs.

```
$ ./scripts/k8s/deploy_kubeflow.sh -x
...
INFO[0007] Applied the configuration Successfully!
filename="cmd/apply.go:72"
Kubeflow app installed to: /home/ai/kubeflow
It may take several minutes for all services to start. Run 'kubectl get
pods -n kubeflow' to verify
To remove (excluding CRDs, istio, auth, and cert-manager), run:
./scripts/k8s_deploy_kubeflow.sh -d
To perform a full uninstall : ./scripts/k8s_deploy_kubeflow.sh -D
Kubeflow Dashboard (HTTP NodePort): http://10.61.188.111:31380
```

3. Confirm that all pods deployed within the Kubeflow namespace show a STATUS of Running and confirm that no components deployed within the namespace are in an error state. It may take several minutes for all pods to start.

```
$ kubectl get all -n kubeflow
```

NAME			READY
STATUS	RESTARTS	AGE	
pod/admission-webhook-bootstrap-stateful-set-0			1/1
Running	0	95s	
pod/admission-webhook-deployment-6b89c84c98-vrtbh			1/1
Running	0	91s	
pod/application-controller-stateful-set-0			1/1
Running	0	98s	
pod/argo-ui-5dcf5d8b4f-m2wn4			1/1
Running	0	97s	
pod/centraldashboard-cf4874ddc-7hcr8			1/1
Running	0	97s	
pod/jupyter-web-app-deployment-685b455447-gjhh7			1/1
Running	0	96s	
pod/katib-controller-88c97d85c-kgq66			1/1
Running	1	95s	
pod/katib-db-8598468fd8-5jw2c			1/1
Running	0	95s	
pod/katib-manager-574c8c67f9-wtrf5			1/1
Running	1	95s	
pod/katib-manager-rest-778857c989-fjbzn			1/1
Running	0	95s	
pod/katib-suggestion-bayesianoptimization-65df4d7455-qthmw			1/1
Running	0	94s	
pod/katib-suggestion-grid-56bf69f597-98vwn			1/1
Running	0	94s	
pod/katib-suggestion-hyperband-7777b76cb9-9v6dq			1/1
Running	0	93s	
pod/katib-suggestion-nasrl-77f6f9458c-2qzxq			1/1
Running	0	93s	
pod/katib-suggestion-random-77b88b5c79-164j9			1/1
Running	0	93s	
pod/katib-ui-7587c5b967-nd629			1/1
Running	0	95s	
pod/metacontroller-0			1/1
Running	0	96s	
pod/metadata-db-5dd459cc-swzkm			1/1
Running	0	94s	
pod/metadata-deployment-6cf77db994-69fk7			1/1
Running	3	93s	
pod/metadata-deployment-6cf77db994-mpbjt			1/1
Running	3	93s	
pod/metadata-deployment-6cf77db994-xg7tz			1/1
Running	3	94s	
pod/metadata-ui-78f5b59b56-qb6kr			1/1

```

Running    0          94s
pod/minio-758b769d67-1lvdr                      1/1
Running    0          91s
pod/ml-pipeline-5875b9db95-g8t2k                 1/1
Running    0          91s
pod/ml-pipeline-persistenceagent-9b69ddd46-bt9r9  1/1
Running    0          90s
pod/ml-pipeline-scheduledworkflow-7b8d756c76-7x56s 1/1
Running    0          90s
pod/ml-pipeline-ui-79ffd9c76-fcwpd               1/1
Running    0          90s
pod/ml-pipeline-viewer-controller-deployment-5fdc87f58-b2t9r 1/1
Running    0          90s
pod/mysql-657f87857d-15k9z                       1/1
Running    0          91s
pod/notebook-controller-deployment-56b4f59bbf-8bvnr 1/1
Running    0          92s
pod/profiles-deployment-6bc745947-mrdkh           2/2
Running    0          90s
pod/pytorch-operator-77c97f4879-hmlrv            1/1
Running    0          92s
pod/seldon-operator-controller-manager-0          1/1
Running    1          91s
pod/spartakus-volunteer-5fdfddb779-17qkm         1/1
Running    0          92s
pod/tensorboard-6544748d94-nh8b2                 1/1
Running    0          92s
pod/tf-job-dashboard-56f79c59dd-6w59t            1/1
Running    0          92s
pod/tf-job-operator-79cbfd6dbc-rb58c             1/1
Running    0          91s
pod/workflow-controller-db644d554-cwrnb          1/1
Running    0          97s

```

NAME	TYPE
CLUSTER-IP	AGE
service/admission-webhook-service	ClusterIP
10.233.51.169 <none> 443/TCP	97s
service/application-controller-service	ClusterIP
10.233.4.54 <none> 443/TCP	98s
service/argo-ui	NodePort
10.233.47.191 <none> 80:31799/TCP	97s
service/centraldashboard	ClusterIP
10.233.8.36 <none> 80/TCP	97s
service/jupyter-web-app-service	ClusterIP
10.233.1.42 <none> 80/TCP	97s
service/katib-controller	ClusterIP

10.233.25.226	<none>	443/TCP	96s
service/katib-db			ClusterIP
10.233.33.151	<none>	3306/TCP	97s
service/katib-manager			ClusterIP
10.233.46.239	<none>	6789/TCP	96s
service/katib-manager-rest			ClusterIP
10.233.55.32	<none>	80/TCP	96s
service/katib-suggestion-bayesianoptimization			ClusterIP
10.233.49.191	<none>	6789/TCP	95s
service/katib-suggestion-grid			ClusterIP
10.233.9.105	<none>	6789/TCP	95s
service/katib-suggestion-hyperband			ClusterIP
10.233.22.2	<none>	6789/TCP	95s
service/katib-suggestion-nasrl			ClusterIP
10.233.63.73	<none>	6789/TCP	95s
service/katib-suggestion-random			ClusterIP
10.233.57.210	<none>	6789/TCP	95s
service/katib-ui			ClusterIP
10.233.6.116	<none>	80/TCP	96s
service/metadata-db			ClusterIP
10.233.31.2	<none>	3306/TCP	96s
service/metadata-service			ClusterIP
10.233.27.104	<none>	8080/TCP	96s
service/metadata-ui			ClusterIP
10.233.57.177	<none>	80/TCP	96s
service/minio-service			ClusterIP
10.233.44.90	<none>	9000/TCP	94s
service/ml-pipeline			ClusterIP
10.233.41.201	<none>	8888/TCP, 8887/TCP	94s
service/ml-pipeline-tensorboard-ui			ClusterIP
10.233.36.207	<none>	80/TCP	93s
service/ml-pipeline-ui			ClusterIP
10.233.61.150	<none>	80/TCP	93s
service/mysql			ClusterIP
10.233.55.117	<none>	3306/TCP	94s
service/notebook-controller-service			ClusterIP
10.233.10.166	<none>	443/TCP	95s
service/profiles-kfam			ClusterIP
10.233.33.79	<none>	8081/TCP	92s
service/pytorch-operator			ClusterIP
10.233.37.112	<none>	8443/TCP	95s
service/seldon-operator-controller-manager-service			ClusterIP
10.233.30.178	<none>	443/TCP	92s
service/tensorboard			ClusterIP
10.233.58.151	<none>	9000/TCP	94s
service/tf-job-dashboard			ClusterIP

```

10.233.4.17      <none>          80/TCP          94s
service/tf-job-operator          ClusterIP
10.233.60.32    <none>          8443/TCP        94s
service/webhook-server-service   ClusterIP
10.233.32.167   <none>          443/TCP         87s
NAME                                                    READY    UP-
TO-DATE    AVAILABLE    AGE
deployment.apps/admission-webhook-deployment          1/1      1
1           97s
deployment.apps/argo-ui                               1/1      1
1           97s
deployment.apps/centraldashboard                     1/1      1
1           97s
deployment.apps/jupyter-web-app-deployment            1/1      1
1           97s
deployment.apps/katib-controller                     1/1      1
1           96s
deployment.apps/katib-db                             1/1      1
1           97s
deployment.apps/katib-manager                        1/1      1
1           96s
deployment.apps/katib-manager-rest                   1/1      1
1           96s
deployment.apps/katib-suggestion-bayesianoptimization 1/1      1
1           95s
deployment.apps/katib-suggestion-grid                1/1      1
1           95s
deployment.apps/katib-suggestion-hyperband           1/1      1
1           95s
deployment.apps/katib-suggestion-nasrl               1/1      1
1           95s
deployment.apps/katib-suggestion-random              1/1      1
1           95s
deployment.apps/katib-ui                             1/1      1
1           96s
deployment.apps/metadata-db                          1/1      1
1           96s
deployment.apps/metadata-deployment                  3/3      3
3           96s
deployment.apps/metadata-ui                          1/1      1
1           96s
deployment.apps/minio                                1/1      1
1           94s
deployment.apps/ml-pipeline                          1/1      1
1           94s
deployment.apps/ml-pipeline-persistenceagent         1/1      1

```


1	93s			
deployment.apps/ml-pipeline-scheduledworkflow		1/1		1
1	93s			
deployment.apps/ml-pipeline-ui		1/1		1
1	93s			
deployment.apps/ml-pipeline-viewer-controller-deployment		1/1		1
1	93s			
deployment.apps/mysql		1/1		1
1	94s			
deployment.apps/notebook-controller-deployment		1/1		1
1	95s			
deployment.apps/profiles-deployment		1/1		1
1	92s			
deployment.apps/pytorch-operator		1/1		1
1	95s			
deployment.apps/spartakus-volunteer		1/1		1
1	94s			
deployment.apps/tensorboard		1/1		1
1	94s			
deployment.apps/tf-job-dashboard		1/1		1
1	94s			
deployment.apps/tf-job-operator		1/1		1
1	94s			
deployment.apps/workflow-controller		1/1		1
1	97s			
NAME				
DESIRED	CURRENT	READY	AGE	
replicaset.apps/admission-webhook-deployment-6b89c84c98				1
1	1	97s		
replicaset.apps/argo-ui-5dcf5d8b4f				1
1	1	97s		
replicaset.apps/centraldashboard-cf4874ddc				1
1	1	97s		
replicaset.apps/jupyter-web-app-deployment-685b455447				1
1	1	97s		
replicaset.apps/katib-controller-88c97d85c				1
1	1	96s		
replicaset.apps/katib-db-8598468fd8				1
1	1	97s		
replicaset.apps/katib-manager-574c8c67f9				1
1	1	96s		
replicaset.apps/katib-manager-rest-778857c989				1
1	1	96s		
replicaset.apps/katib-suggestion-bayesianoptimization-65df4d7455				1
1	1	95s		
replicaset.apps/katib-suggestion-grid-56bf69f597				1

1	1	95s		
replicaset.apps/katib-suggestion-hyperband-7777b76cb9			1	
1	1	95s		
replicaset.apps/katib-suggestion-nasrl-77f6f9458c			1	
1	1	95s		
replicaset.apps/katib-suggestion-random-77b88b5c79			1	
1	1	95s		
replicaset.apps/katib-ui-7587c5b967			1	
1	1	96s		
replicaset.apps/metadata-db-5dd459cc			1	
1	1	96s		
replicaset.apps/metadata-deployment-6cf77db994			3	
3	3	96s		
replicaset.apps/metadata-ui-78f5b59b56			1	
1	1	96s		
replicaset.apps/minio-758b769d67			1	
1	1	93s		
replicaset.apps/ml-pipeline-5875b9db95			1	
1	1	93s		
replicaset.apps/ml-pipeline-persistenceagent-9b69ddd46			1	
1	1	92s		
replicaset.apps/ml-pipeline-scheduledworkflow-7b8d756c76			1	
1	1	91s		
replicaset.apps/ml-pipeline-ui-79ffd9c76			1	
1	1	91s		
replicaset.apps/ml-pipeline-viewer-controller-deployment-5fdc87f58			1	
1	1	91s		
replicaset.apps/mysql-657f87857d			1	
1	1	92s		
replicaset.apps/notebook-controller-deployment-56b4f59bbf			1	
1	1	94s		
replicaset.apps/profiles-deployment-6bc745947			1	
1	1	91s		
replicaset.apps/pytorch-operator-77c97f4879			1	
1	1	94s		
replicaset.apps/spartakus-volunteer-5fdfddb779			1	
1	1	94s		
replicaset.apps/tensorboard-6544748d94			1	
1	1	93s		
replicaset.apps/tf-job-dashboard-56f79c59dd			1	
1	1	93s		
replicaset.apps/tf-job-operator-79cbfd6dbc			1	
1	1	93s		
replicaset.apps/workflow-controller-db644d554			1	
1	1	97s		
NAME			READY	AGE

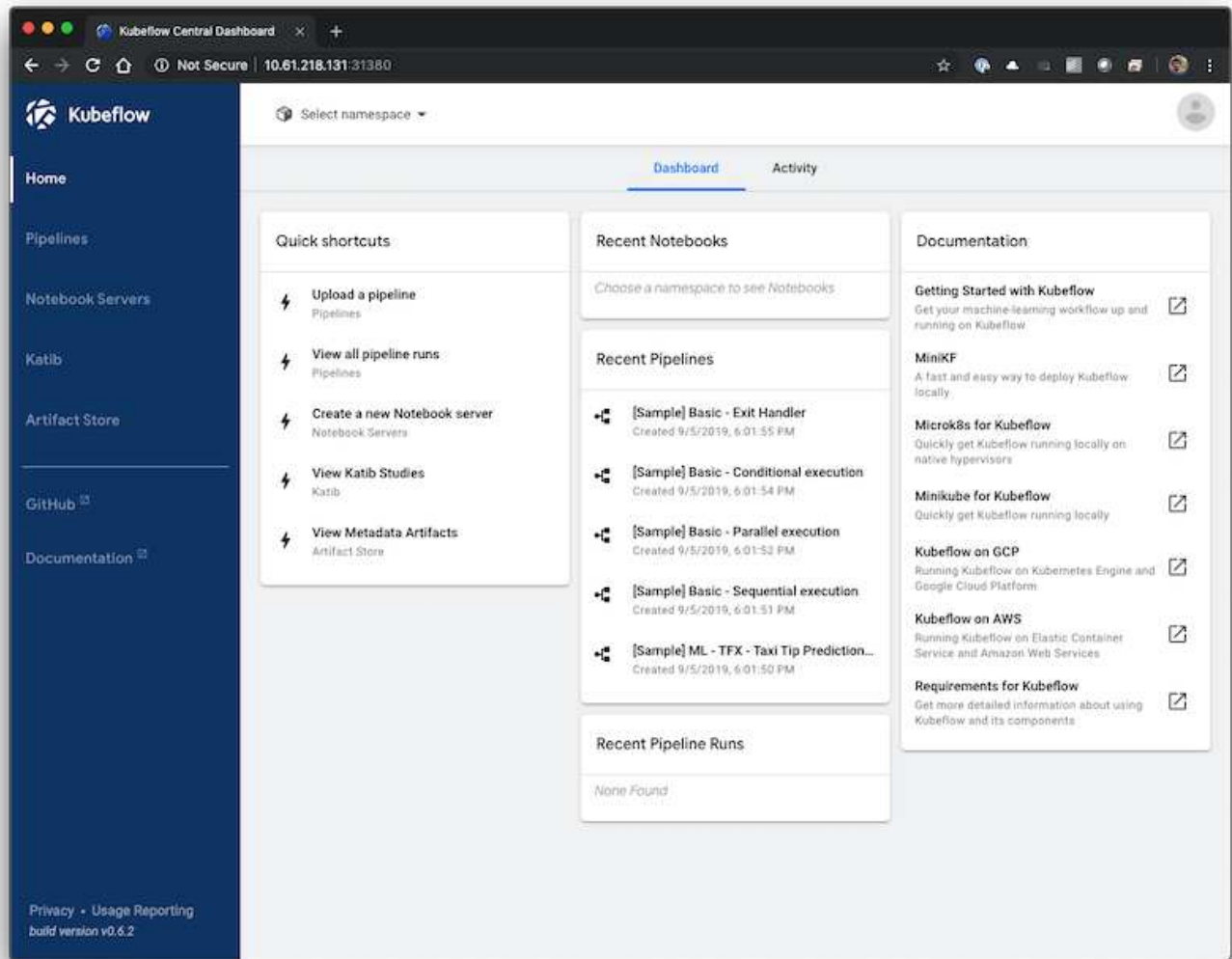
```

statefulset.apps/admission-webhook-bootstrap-stateful-set 1/1 97s
statefulset.apps/application-controller-stateful-set 1/1 98s
statefulset.apps/metacontroller 1/1 98s
statefulset.apps/seldon-operator-controller-manager 1/1 92s
$ kubectl get pvc -n kubeflow
NAME                                STATUS    VOLUME                                     CAPACITY   ACCESS MODES   STORAGECLASS          AGE
katib-mysql                         Bound    pvc-b07f293e-d028-11e9-9b9d-00505681a82d  10Gi       RWO             ontap-ai-flexvols-retain 27m
metadata-mysql                     Bound    pvc-b0f3f032-d028-11e9-9b9d-00505681a82d  10Gi       RWO             ontap-ai-flexvols-retain 27m
minio-pv-claim                     Bound    pvc-b22727ee-d028-11e9-9b9d-00505681a82d  20Gi       RWO             ontap-ai-flexvols-retain 27m
mysql-pv-claim                     Bound    pvc-b2429afd-d028-11e9-9b9d-00505681a82d  20Gi       RWO             ontap-ai-flexvols-retain 27m

```

4. In your web browser, access the Kubeflow central dashboard by navigating to the URL that you noted down in step 2.

The default username is `admin@kubeflow.org`, and the default password is `12341234`. To create additional users, follow the instructions in the [official Kubeflow documentation](#).



Next: Example Kubeflow Operations and Tasks.

Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.