



Beispiel Kubeflow Operations und Tasks

NetApp Solutions

NetApp
May 03, 2024

This PDF was generated from https://docs.netapp.com/de-de/netapp-solutions/ai/aicp_example_kubeflow_operations_and_tasks_overview.html on May 03, 2024. Always check docs.netapp.com for the latest.

Inhalt

- Beispiel Kubeflow Operations und Tasks 1
 - Beispiel Kubeflow Operations und Tasks 1
 - Bereitstellen eines Jupyter Notebook Workspace für Data Scientist oder Entwickler 1
 - Beispiel für Notebooks und Pipelines. 8

Beispiel Kubeflow Operations und Tasks

Dieser Abschnitt enthält Beispiele für verschiedene Vorgänge und Aufgaben, die Sie mit Kubeflow ausführen möchten.

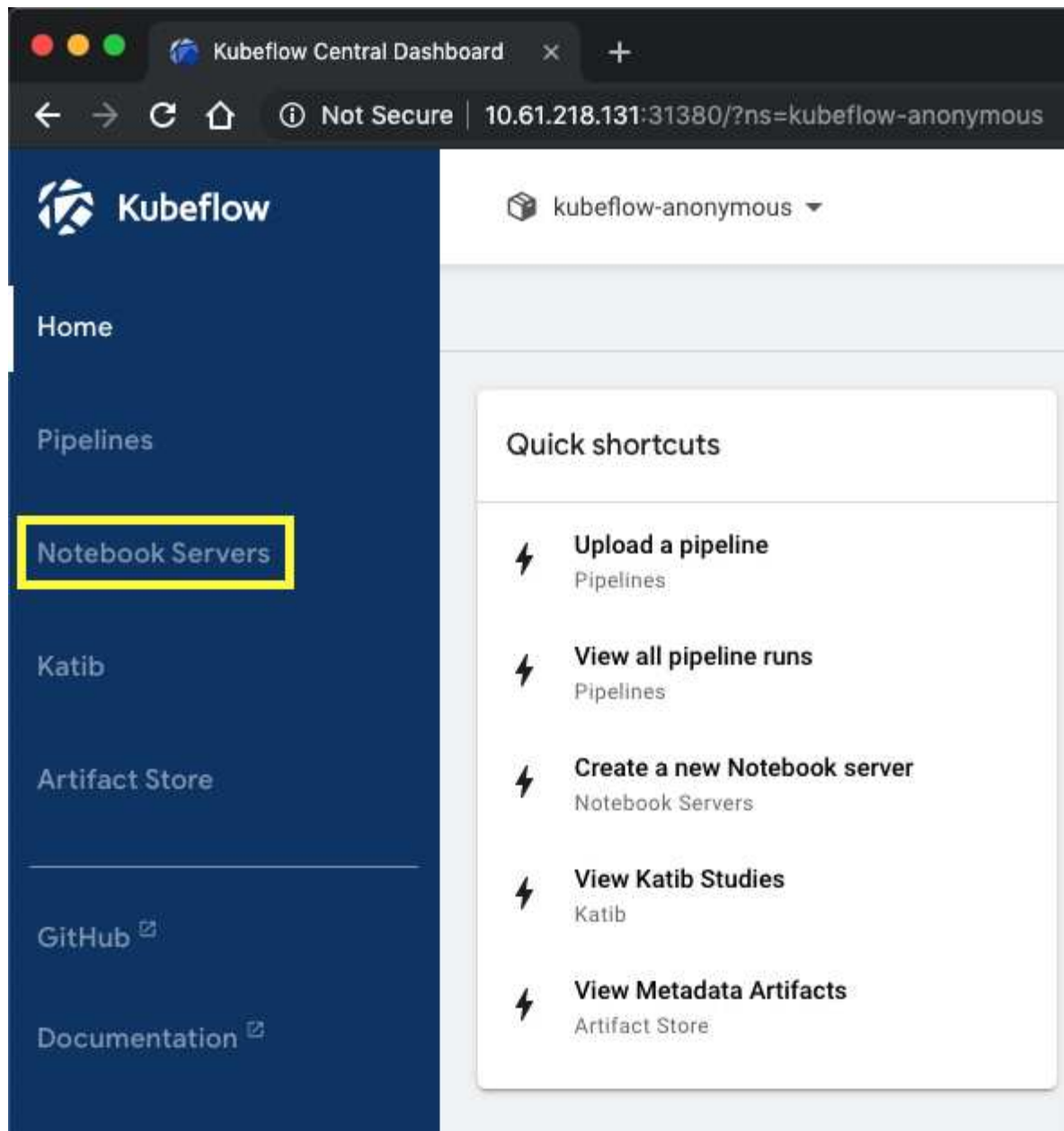
Beispiel Kubeflow Operations und Tasks

Dieser Abschnitt enthält Beispiele für verschiedene Vorgänge und Aufgaben, die Sie mit Kubeflow ausführen möchten.

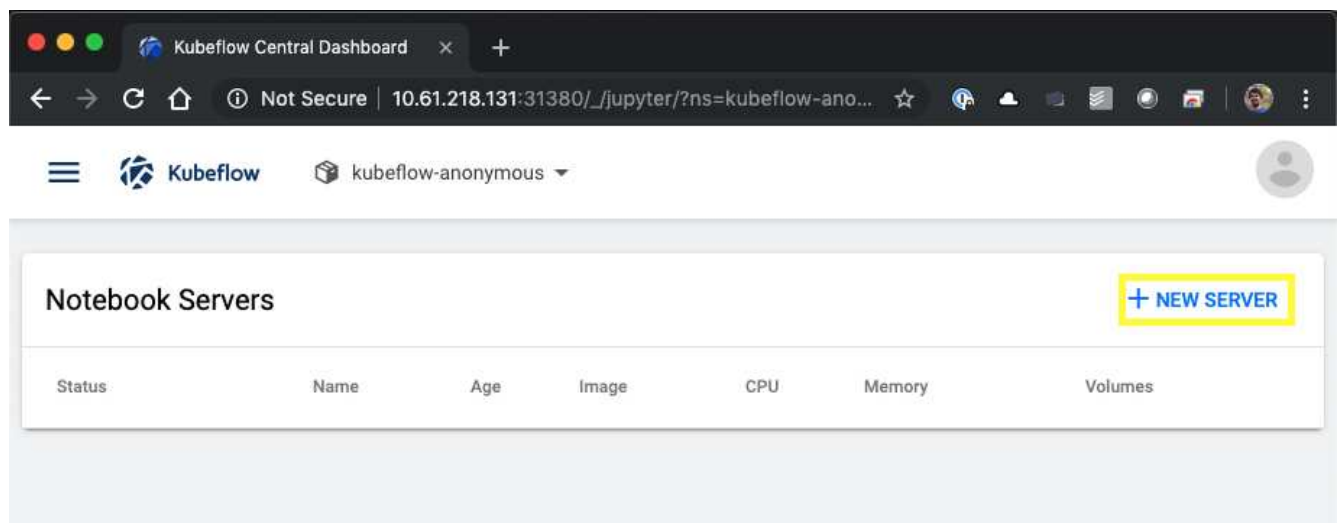
Bereitstellen eines Jupyter Notebook Workspace für Data Scientist oder Entwickler

Kubeflow ist in der Lage, neue Jupyter Notebook-Server schnell als Data Scientist-Workspaces bereitzustellen. Um einen neuen Jupyter Notebook-Server mit Kubeflow bereitzustellen, führen Sie die folgenden Aufgaben aus. Weitere Informationen zu Jupyter Notebooks im Kubeflow-Kontext finden Sie im ["Offizielle Dokumentation von Kubeflow"](#).

1. Klicken Sie im zentralen Kubeflow Dashboard im Hauptmenü auf Notebook-Server, um zur Jupyter Notebook-Serververwaltungsseite zu navigieren.

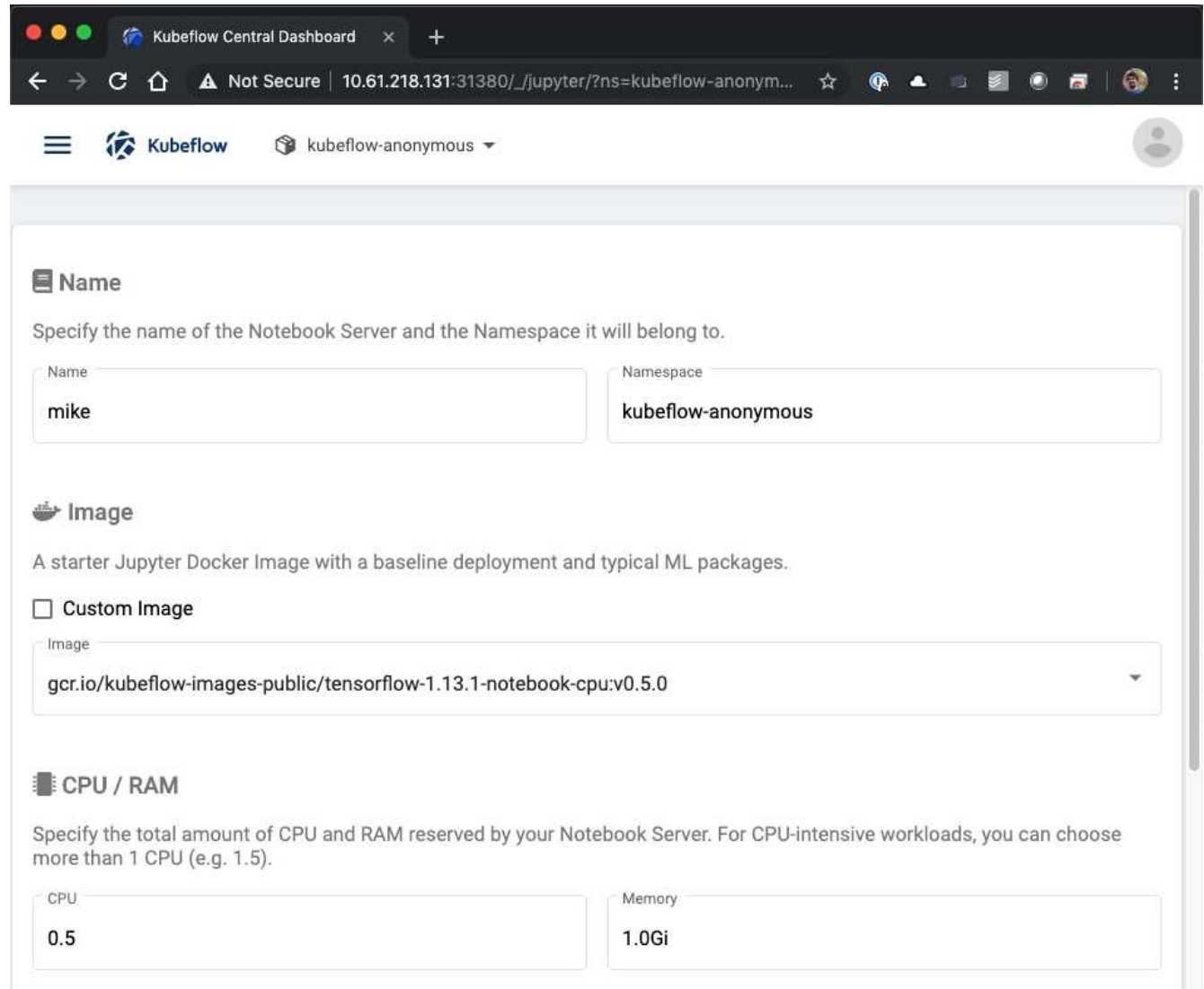


2. Klicken Sie auf Neuer Server, um einen neuen Jupyter Notebook-Server bereitzustellen.



3. Geben Sie Ihrem neuen Server einen Namen, wählen Sie das Docker-Image aus, auf dem Ihr Server basieren soll, und geben Sie die CPU- und RAM-Größe an, die von Ihrem Server reserviert werden soll. Wenn das Namespace-Feld leer ist, wählen Sie über das Menü Namespace auswählen in der Seitenüberschrift einen Namespace aus. Das Namespace-Feld wird dann automatisch mit dem ausgewählten Namespace gefüllt.

Im folgenden Beispiel wird der verwendet `kubeflow-anonymous` Namespace ausgewählt. Darüber hinaus werden die Standardwerte für das Docker Image, die CPU und den RAM akzeptiert.



The screenshot shows the Kubeflow Central Dashboard in a web browser. The browser's address bar shows the URL `10.61.218.131:31380/_jupyter/?ns=kubeflow-anonym...`. The dashboard header includes the Kubeflow logo and a dropdown menu for the namespace, which is currently set to `kubeflow-anonymous`. The main content area is titled "Name" and contains a form for specifying the Notebook Server name and namespace. The "Name" field is filled with `mike` and the "Namespace" field is filled with `kubeflow-anonymous`. Below this, there is a section titled "Image" with a description: "A starter Jupyter Docker Image with a baseline deployment and typical ML packages." There is a checkbox for "Custom Image" which is currently unchecked. Below the checkbox, there is a dropdown menu for the image, which is currently set to `gcr.io/kubeflow-images-public/tensorflow-1.13.1-notebook-cpu:v0.5.0`. At the bottom, there is a section titled "CPU / RAM" with a description: "Specify the total amount of CPU and RAM reserved by your Notebook Server. For CPU-intensive workloads, you can choose more than 1 CPU (e.g. 1.5)." There are two input fields: "CPU" with the value `0.5` and "Memory" with the value `1.0Gi`.

4. Geben Sie die Einzelheiten zum Workspace-Volumen an. Wenn Sie ein neues Volume erstellen möchten, wird dieses Volume oder PVC mithilfe der Standard-StorageClass bereitgestellt. Da ein StorageClass mit Trident als Standard-StorageClass in diesem Abschnitt festgelegt wurde "[Kubeflow Deployment](#)", Das Volume oder die PVC wird mit Trident bereitgestellt. Dieses Volume wird automatisch als Standardarbeitsbereich im Jupyter Notebook Server-Container bereitgestellt. Alle Notebooks, die ein Benutzer auf dem Server erstellt, die nicht in einem separaten Datenvolumen gespeichert werden, werden automatisch in diesem Workspace-Volumen gespeichert. Daher bleiben die Notebooks auch bei Neustarts erhalten.

Workspace Volume

Configure the Volume to be mounted as your personal Workspace.

☐ Don't use Persistent Storage for User's home

Type	Name	Size	Mode	Mount Point
New	workspace-mike	10Gi	ReadWriteOnce	/home/jovyan

5. Hinzufügen von Daten-Volumes Im folgenden Beispiel wird ein vorhandenes PVC mit dem Namen „pb-fg-all“ angegeben und der Standard-Bereitstellungspunkt akzeptiert.

Data Volumes

Configure the Volumes to be mounted as your Datasets.

[+ ADD VOLUME](#)

Type	Name	Size	Mode	Mount Point
Existing	pb-fg-all	10Gi	ReadWriteOnce	/home/jovyan/data-vol-1

6. **Optional:** Anfrage, dass die gewünschte Anzahl von GPUs Ihrem Notebook-Server zugewiesen wird. Im folgenden Beispiel wird eine GPU angefordert.

Configurations

Extra layers of configurations that will be applied to the new Notebook. (e.g. Insert credentials as Secrets, set Environment Variables.)

Configurations

Extra Resources

Specify extra resources that might be needed in the Notebook Server.

☒ Enable Shared Memory

Extra Resources *

```
{"nvidia.com/gpu": 1}
```

Extra Resources available in the cluster (ex. NVIDIA GPUs)

LAUNCH **CANCEL**

7. Klicken Sie auf Starten, um Ihren neuen Notebook-Server bereitzustellen.
8. Warten Sie, bis Ihr Notebook-Server vollständig bereitgestellt ist. Dies kann einige Minuten dauern, wenn Sie noch keinen Server mit dem von Ihnen angegebenen Docker-Image bereitgestellt haben, da das Image heruntergeladen werden muss. Wenn Ihr Server vollständig bereitgestellt wurde, wird in der Spalte Status auf der Seite Jupyter Notebook Server Administration ein grünes Häkchen angezeigt.

Notebook Servers [+ NEW SERVER](#)

Status	Name	Age	Image	CPU	Memory	Volumes	
	mike	12 mins ago	tensorflow-1.13.1-notebook-cpu:v0.5.0	0.5	1.0Gi		CONNECT

9. Klicken Sie auf Verbinden, um eine Verbindung mit der neuen Server-Webschnittstelle herzustellen.
10. Vergewissern Sie sich, dass das in Schritt 6 angegebene Datensatz-Volumen auf dem Server installiert ist. Beachten Sie, dass dieses Volume standardmäßig im Standardarbeitsbereich bereitgestellt wird. Aus der Perspektive des Benutzers ist dies nur ein weiterer Ordner im Arbeitsbereich. Der Anwender, der wahrscheinlich ein Data Scientist ist und nicht ein Infrastruktur-Experte ist, muss nicht über Storage-Fachkenntnisse verfügen, um dieses Volumen zu nutzen.

Jupyter [Quit](#)

Files Running Clusters

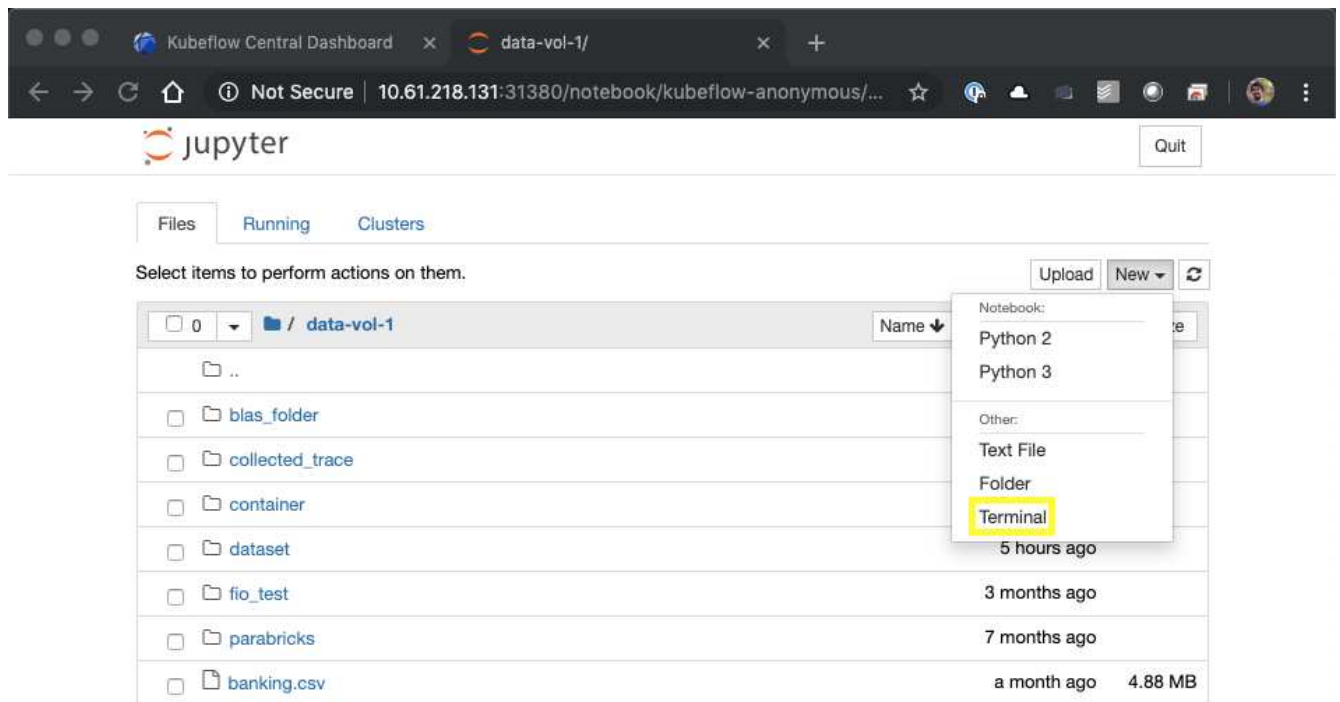
Select items to perform actions on them. [Upload](#) [New](#) [Refresh](#)

	Name	Last Modified	File size
<input type="checkbox"/>	/		
<input type="checkbox"/>	data-vol-1	a day ago	



11. Öffnen Sie ein Terminal und, vorausgesetzt, dass ein neues Volume in Schritt 5 angefordert wurde, führen Sie aus `df -h`. So bestätigen Sie, dass ein neues persistentes Volume mit Trident-Bereitstellung als Standardarbeitsbereich gemountet wird.

Das Standard-Workspace-Verzeichnis ist das Basisverzeichnis, mit dem Sie angezeigt werden, wenn Sie zum ersten Mal auf die Web-Schnittstelle des Servers zugreifen. Auf diesem persistenten Volume durch Trident werden daher alle mit der Web-Schnittstelle erstellten Artefakte gespeichert.




```

$ df -h
Filesystem                                Size  Used Avail
Use% Mounted on
overlay                                  439G   34G  382G
9% /
tmpfs                                     64M    0   64M
0% /dev
tmpfs                                    252G    0  252G
0% /sys/fs/cgroup
/dev/sda2                                439G   34G  382G
9% /etc/hosts
192.168.11.11:/trident_pvc_3dcfe7e5_d5a9_11e9_9b9d_00505681a82d 10G  320K   10G
1% /home/jovyan
tmpfs                                    252G    0  252G
0% /dev/shm
192.168.11.11:/pb_fg_all                  10T   10T   47G
100% /home/jovyan/data-vol-1
tmpfs                                    252G   12K  252G
1% /run/secrets/kubernetes.io/serviceaccount
tmpfs                                    252G   12K  252G
1% /proc/driver/nvidia
tmpfs                                     51G   4.9M   51G
1% /run/nvidia-persistenced/socket
udev                                     252G    0  252G
0% /dev/nvidia5
tmpfs                                    252G    0  252G
0% /proc/acpi
tmpfs                                    252G    0  252G
0% /proc/scsi
tmpfs                                    252G    0  252G
0% /sys/firmware
$

```

12. Führen Sie mit dem Terminal aus `nvidia-smi` Um zu bestätigen, dass dem Notebook-Server die korrekte Anzahl an GPUs zugewiesen wurde. Im folgenden Beispiel wurde dem Notebook-Server eine GPU zugewiesen, wie in Schritt 7 angefordert.

```

$ nvidia-smi
Fri Sep 13 13:52:15 2019
+-----+
| NVIDIA-SMI 410.104                Driver Version: 410.104                CUDA Version: N/A                |
+-----+
| GPU   Name                               Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
+-----+-----+
|  0  Tesla V100-SXM2...    On           | 00000000:86:00:0 Off |                    0 |
|N/A   38C    P0     46W / 300W |  0MiB / 32480MiB |      0%      Default |
+-----+-----+

+-----+
| Processes:                               GPU Memory |
|  GPU       PID    Type    Process name                        Usage      |
+-----+-----+
| No running processes found               |
+-----+
$

```

Beispiel für Notebooks und Pipelines

Der "[Das NetApp Data Science Toolkit für Kubernetes](#)" Kann in Verbindung mit Kubeflow verwendet werden. Die Verwendung des NetApp Data Science Toolkit und Kubeflow bietet folgende Vorteile:

- Data Scientists können erweiterte NetApp Datenmanagement-Vorgänge direkt über ein Jupyter Notebook ausführen.
- Erweiterte Datenmanagement-Vorgänge von NetApp können mithilfe des Kubeflow Pipelines in automatisierte Workflows integriert werden.

Siehe "[Kubeflow Beispiele](#)" Im Abschnitt zum NetApp Data Science Toolkit GitHub Repository finden Sie weitere Informationen zur Verwendung des Toolkit mit Kubeflow.

Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtsinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFTE SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGEND EINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter <http://www.netapp.com/TM> aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.