



例： **Kubeflow** の操作とタスク NetApp Solutions

NetApp
April 10, 2024

目次

例：Kubeflow の操作とタスク	1
例：Kubeflow の操作とタスク	1
データサイエンティストまたは開発者向けに Jupyter Notebook Workspace をプロビジョニングします... 1 使用	
ノートブック PC とパイプラインの例	9

例： **Kubeflow** の操作とタスク

このセクションでは、Kubeflow を使用して実行したいさまざまな操作とタスクの例を示します。

例： **Kubeflow** の操作とタスク

このセクションでは、Kubeflow を使用して実行したいさまざまな操作とタスクの例を示します。

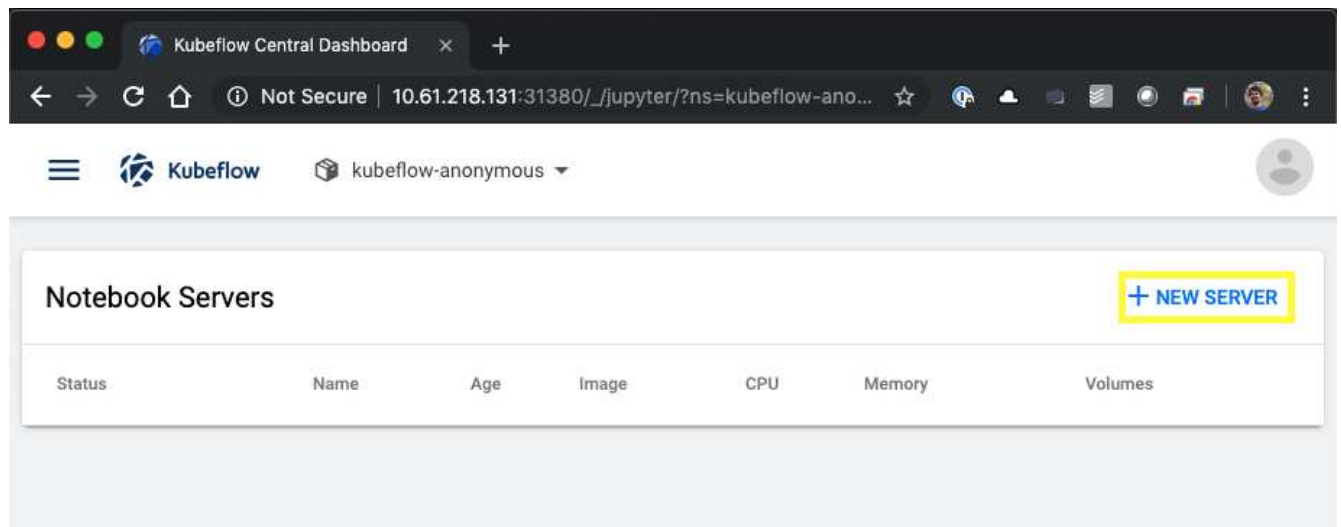
データサイエンティストまたは開発者向けに **Jupyter Notebook Workspace** をプロビジョニングします 使用

Kubeflow は、データサイエンティストのワークスペースとして機能する、新しい Jupyter Notebook サーバの迅速なプロビジョニングを可能にします。Kubeflow を使用して新しい Jupyter Notebook サーバをプロビジョニングするには、次のタスクを実行します。Kubeflow コンテキスト内の Jupyter Notebook の詳細については、[を参照してください](#) "[Kubeflow の公式ドキュメント](#)"。

1. Kubeflow 中央ダッシュボードのメインメニューで Notebook Servers をクリックして、Jupyter Notebook サーバ管理ページに移動します。



2. 新しい Jupyter Notebook サーバをプロビジョニングするには、New Server をクリックします。



3. 新しいサーバに名前を付け、サーバのベースにする Docker イメージを選択し、サーバで予約する CPU と RAM の容量を指定します。[名前空間] フィールドが空白の場合は、ページヘッダーの [名前空間の選択] メニューを使用して名前空間を選択します。選択したネームスペースがネームスペースフィールドに自動的に入力されます。

次の例では 'kubeflow-anonymous' ネームスペースが選択されていますまた、Docker イメージ、CPU、RAM のデフォルト値も使用できます。

Name

Specify the name of the Notebook Server and the Namespace it will belong to.

Name: Namespace:

Image

A starter Jupyter Docker Image with a baseline deployment and typical ML packages.

☐ Custom Image

Image:

CPU / RAM

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).

CPU: Memory:

- ワークスペースボリュームの詳細を指定します。新しいボリュームを作成するように選択した場合は、そのボリュームまたは PVC がデフォルトの StorageClass を使用してプロビジョニングされます。Trident を利用するストレージクラスがデフォルトとして指定されているため StorageClass "[Kubeflow の導入](#)" を指定した場合、ボリュームまたは PVC は Trident でプロビジョニングされます。このボリュームは、Jupyter Notebook Server コンテナ内のデフォルトワークスペースとして自動的にマウントされます。別のデータボリュームに保存されていないユーザーがサーバー上に作成したノートブックは、自動的にこのワークスペースボリュームに保存されます。そのため、ノートブックは再起動後も維持されます。

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:

- データボリュームを追加次に、「pb-fg-all」という名前の既存の PVC を指定し、デフォルトのマウントポイントを受け入れる例を示します。

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. * オプション：* 希望する数の GPU をノートブックサーバーに割り当てるよう要求します。次の例では、GPU が 1 つ要求されています。

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**

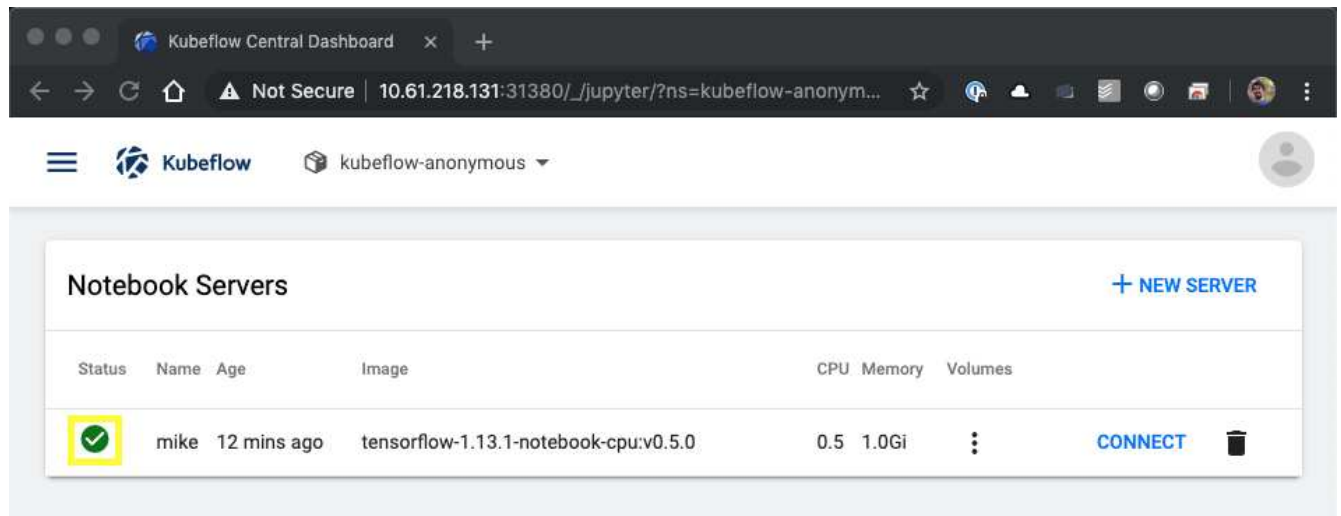
Extra Resources *

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

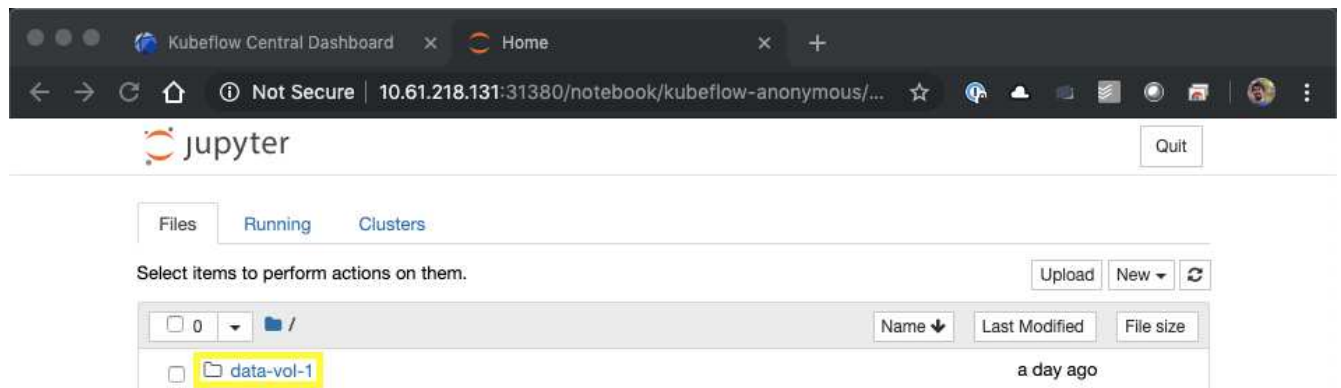
Extra Resources available in the cluster (ex. NVIDIA GPUs)

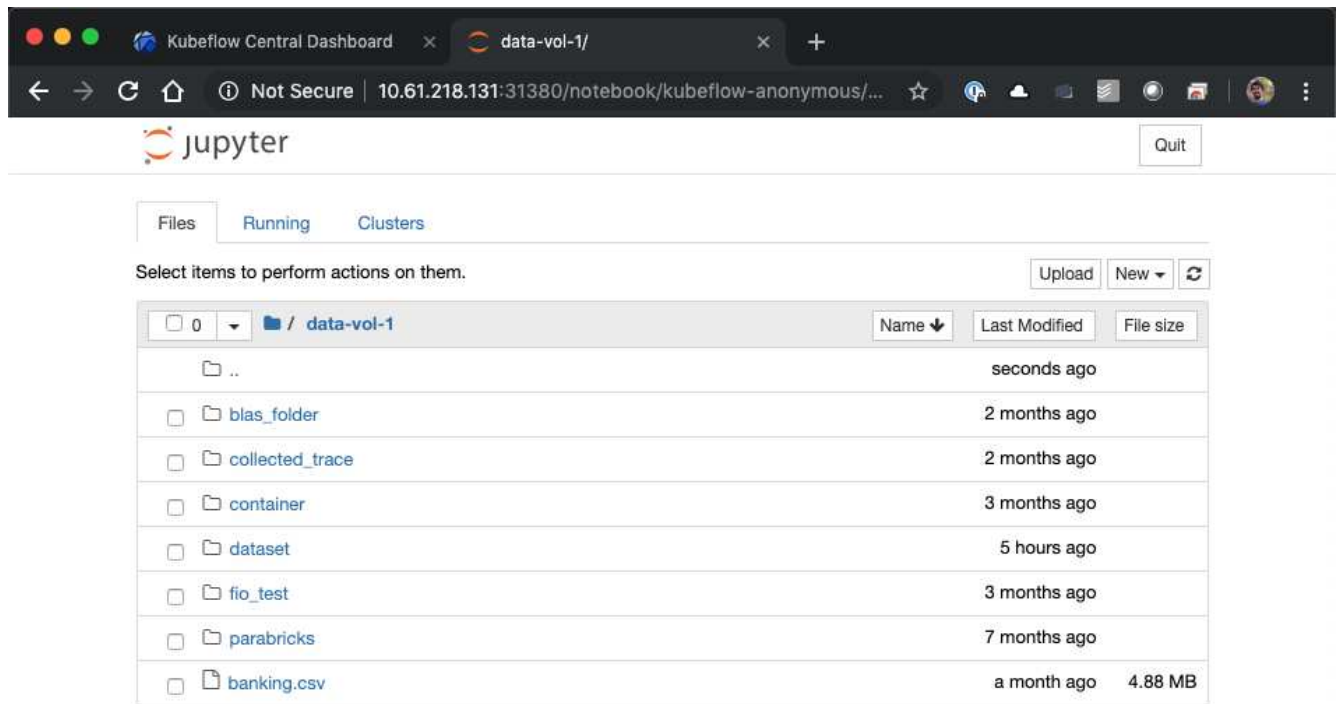
[LAUNCH](#) [CANCEL](#)

7. [起動] をクリックして、新しいノートブックサーバーをプロビジョニングします。
8. ノートブックサーバーが完全にプロビジョニングされるまで待ちます。指定した Docker イメージを使用してサーバをプロビジョニングしたことがない場合は、イメージのダウンロードが必要になるため、これには数分かかることがあります。サーバーが完全にプロビジョニングされると、Jupyter Notebook サーバー管理ページの [ステータス] 列に緑色のチェックマークが表示されます。



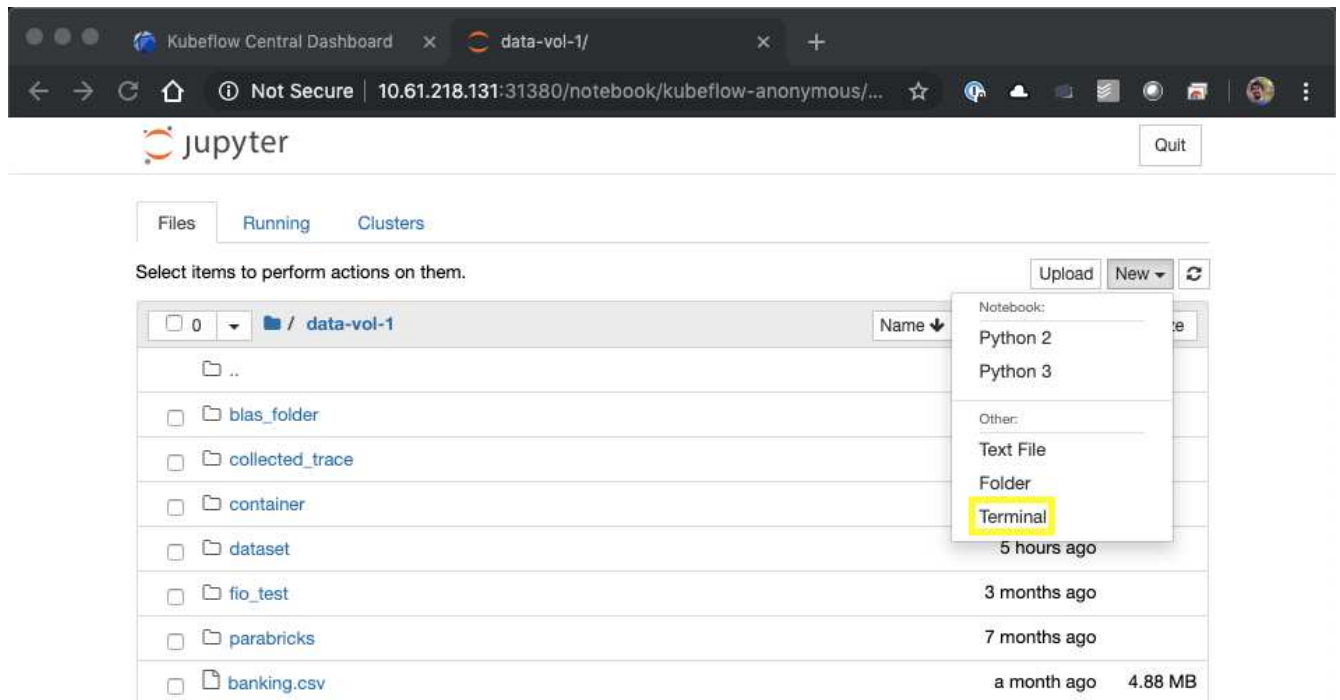
9. [接続] をクリックして、新しいサーバー Web インターフェイスに接続します。
10. 手順 6 で指定したデータセットボリュームがサーバにマウントされていることを確認します。デフォルトでは、このボリュームはデフォルトのワークスペース内にマウントされます。ユーザーの観点から見ると、これはワークスペース内の別のフォルダーにすぎません。データサイエンティストで、インフラのエキスパートではないユーザは、このボリュームを使用するためにストレージの専門知識を持っている必要はありません。





11. ターミナルを開き、手順 5 で新しいボリュームが要求された場合は、「df-h」を実行して、Trident でプロビジョニングされた新しい永続ボリュームがデフォルトのワークスペースとしてマウントされていることを確認します。

デフォルトのワークスペースディレクトリは、サーバーの Web インターフェイスに最初にアクセスしたときに表示されるベースディレクトリです。そのため、Web インターフェイスを使用して作成したアーティファクトは、Trident でプロビジョニングされた永続ボリュームに保存されます。



```

$ 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
$

```

- ターミナルを使用して「nvidia-smi」を実行し、ノートブックサーバーに正しい数の GPU が割り当てられていることを確認します。次の例では、手順 7 で要求されたように、1 つの GPU がノートブックサーバーに割り当てられています。

```

$ 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%         0 |
|N/A   38C    P0     46W / 300W | 0MiB / 32480MiB |           | Default |
+-----+-----+

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

```

ノートブック PC とパイプラインの例

。 ["NetApp Data Science Toolkit for Kubernetes"](#) Kubeflow と組み合わせて使用できます。NetApp Data Science Toolkit と Kubeflow を使用すると、次のようなメリットがあります。

- データサイエンティストは、Jupyter Notebook 内から、ネットアップの高度なデータ管理操作を直接実行できます。
- Kubeflow Pipelines フレームワークを使用すると、ネットアップの高度なデータ管理処理を自動化されたワークフローに組み込むことができます。

を参照してください ["Kubeflow の例"](#) ツールキットと Kubeflow を使用する場合は詳細については、NetApp Data Science Toolkit GitHub リポジトリのセクションを参照してください。

著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータ ソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用权を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用权については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/TM>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。