



# **NVIDIA AI Enterprise と NetApp 、 VMware** NetApp artificial intelligence solutions

NetApp  
August 18, 2025

# 目次

NVIDIA AI Enterprise とNetApp、VMware	1
NVIDIA AI Enterprise とNetApp、VMware	1
テクノロジーの概要	1
NVIDIA AIエンタープライズ	2
NVIDIA GPU クラウド (NGC)	2
VMware vSphere	2
NetApp ONTAP	2
NetApp DataOps ツールキット	3
アーキテクチャ	4
初期セットアップ	5
前提条件	5
NVIDIA AI Enterprise Host ソフトウェアをインストールする	5
NVIDIA NGCソフトウェアを使用する	5
セットアップ	5
使用例 - TensorFlow トレーニングジョブ	8
詳細情報の入手方法	10
謝辞	11

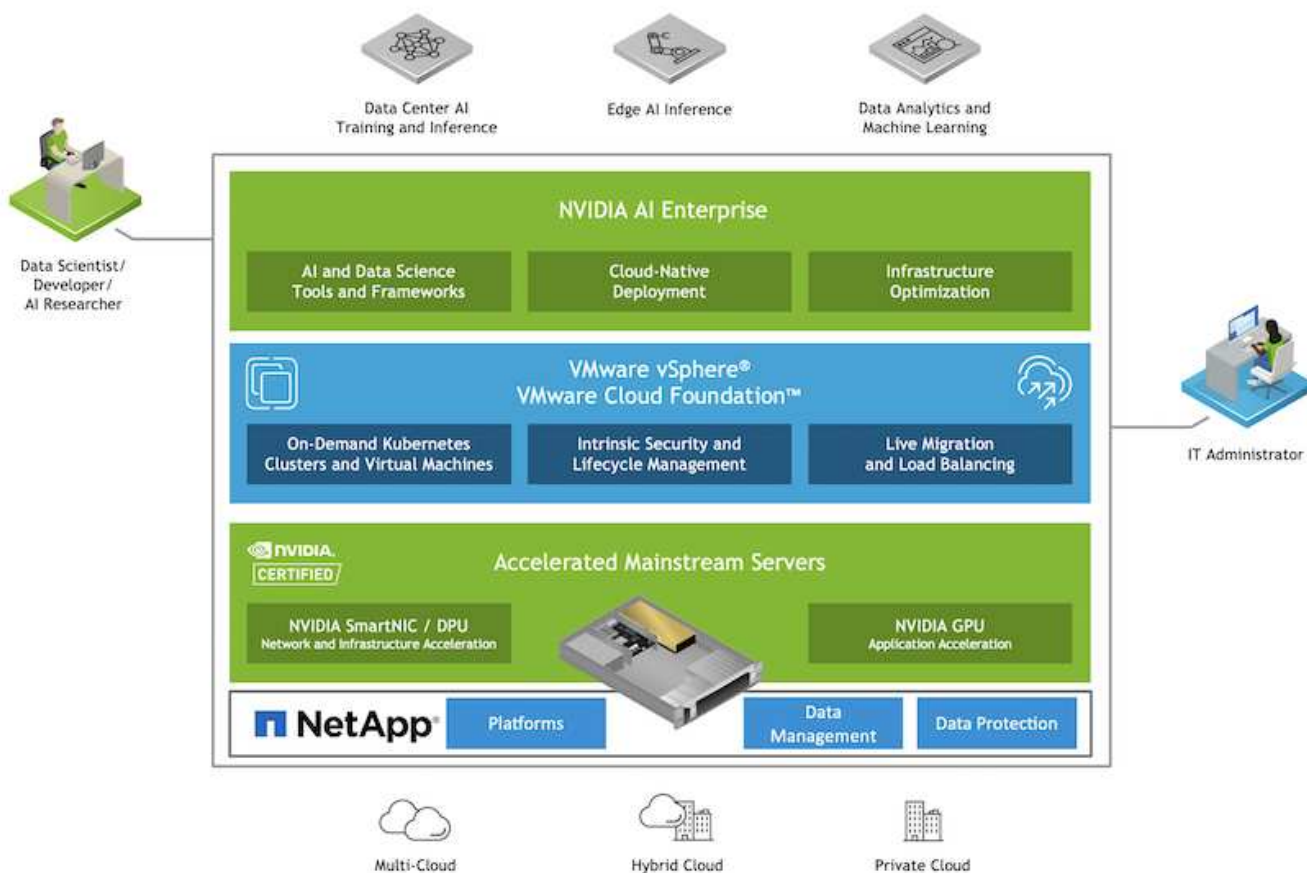
# NVIDIA AI Enterprise と NetApp、VMware

## NVIDIA AI Enterprise と NetApp、VMware

マイク・オグルスピー、NetApp

IT アーキテクトや管理者にとって、AI ツールは複雑で馴染みのない場合があります。さらに、多くの AI プラットフォームはエンタープライズ対応ではありません。NetAppとVMware の技術を活用したNVIDIA AI Enterprise は、合理化されたエンタープライズクラスの AI アーキテクチャを提供するために開発されました。

NVIDIA AI Enterprise は、NVIDIA認定システムを搭載した VMware vSphere で実行できるようにNVIDIAによって最適化、認定、サポートされている、エンドツーエンドのクラウド ネイティブな AI およびデータ分析ソフトウェアスイートです。このソフトウェアは、最新のハイブリッドクラウド環境での AI ワークロードのシンプルかつ迅速な導入、管理、スケーリングを可能にします。NetAppとVMware を搭載したNVIDIA AI Enterprise は、シンプルで使い慣れたパッケージでエンタープライズクラスの AI ワークロードとデータ管理を提供します。



## テクノロジーの概要

このセクションでは、NetAppおよびVMwareを使用したNVIDIA AI Enterprise のテクノロジーの概要を説明します。

## NVIDIA AIエンタープライズ

NVIDIA AI Enterprise は、NVIDIA認定システムを搭載した VMware vSphere で実行できるようにNVIDIAによって最適化、認定、サポートされている、エンドツーエンドのクラウド ネイティブな AI およびデータ分析ソフトウェア スイートです。このソフトウェアは、最新のハイブリッド クラウド環境での AI ワークロードのシンプルかつ迅速な導入、管理、スケーリングを可能にします。

## NVIDIA GPU クラウド (NGC)

NVIDIA NGC は、AI 実践者が AI ソリューションを開発できるように、GPU に最適化されたソフトウェアのカタログをホストしています。また、モデルトレーニング用のNVIDIA Base Command、モデルの展開と監視用のNVIDIA Fleet Command、独自の AI ソフトウェアに安全にアクセスして管理するための NGC Private Registry など、さまざまな AI サービスへのアクセスも提供します。また、NVIDIA AI Enterprise のお客様は、NGC ポータルを通じてサポートをリクエストできます。

## VMware vSphere

VMware vSphere は、データセンターを CPU、ストレージ、ネットワークリソースを含む統合コンピューティング インフラストラクチャへと変革する VMware の仮想化プラットフォームです。vSphere はこれらのインフラストラクチャを統合されたオペレーティング環境として管理し、その環境に参加するデータセンターを管理するためのツールを管理者に提供します。

vSphere の 2 つのコア コンポーネントは、ESXi と vCenter Server です。ESXiは、管理者が仮想マシンと仮想アプライアンスを作成および実行するための仮想化プラットフォームです。vCenter Serverは、管理者がネットワークに接続された複数のホストを管理し、ホストリソースをプールするためのサービスです。

## NetApp ONTAP

NetAppの最新世代のストレージ管理ソフトウェアであるONTAP 9 により、企業はインフラストラクチャを最新化し、クラウド対応のデータセンターに移行できるようになります。ONTAP は業界をリードするデータ管理機能を活用し、データの保存場所に関係なく、単一のツール セットでデータの管理と保護を可能にします。また、エッジ、コア、クラウドなど、必要な場所にデータを自由に移動することもできます。ONTAP 9 には、データ管理を簡素化し、重要なデータを高速化および保護し、ハイブリッド クラウド アーキテクチャ全体で次世代のインフラストラクチャ機能を実現する多数の機能が含まれています。

### データ管理を簡素化

データ管理は、AI アプリケーションと AI/ML データセットのトレーニングに適切なリソースが使用されるように、企業の IT 運用とデータサイエンティストにとって非常に重要です。NetAppテクノロジーに関する次の追加情報は、この検証の範囲外ですが、導入によっては関連する可能性があります。

ONTAPデータ管理ソフトウェアには、運用を合理化および簡素化し、総運用コストを削減するための次の機能が含まれています。

- インライン データ圧縮と拡張重複排除。データ圧縮によりストレージ ブロック内の無駄なスペースが削減され、重複排除により実効容量が大幅に増加します。これは、ローカルに保存されたデータとクラウドに階層化されたデータに適用されます。
- 最小、最大、および適応型サービス品質 (AQoS)。きめ細かなサービス品質 (QoS) 制御により、高度に共有された環境における重要なアプリケーションのパフォーマンス レベルを維持できます。
- NetAppFabricPool。Amazon Web Services (AWS)、Azure、NetApp StorageGRIDストレージ ソリューションなどのパブリックおよびプライベート クラウド ストレージ オプションへのコールド データの自動

階層化を提供します。FabricPoolの詳細については、以下を参照してください。"[TR-4598: FabricPoolのベストプラクティス](#)"。

## データの高速化と保護

ONTAP は優れたレベルのパフォーマンスとデータ保護を提供し、これらの機能を次のように拡張します。

- パフォーマンスと低レイテンシ。ONTAP は、可能な限り低いレイテンシで最高のスループットを提供します。
- データ保護：ONTAP は、すべてのプラットフォームにわたる共通管理を備えた組み込みのデータ保護機能を提供します。
- NetAppボリューム暗号化 (NVE)。ONTAP は、オンボードと外部キー管理の両方をサポートするネイティブのボリューム レベルの暗号化を提供します。
- マルチテナントと多要素認証。ONTAP は、最高レベルのセキュリティでインフラストラクチャ リソースを共有できるようにします。

## 将来を見据えたインフラ

ONTAP は、次の機能により、要求が厳しく常に変化するビジネス ニーズへの対応に役立ちます。

- シームレスなスケーリングと中断のない運用。ONTAP は、既存のコントローラおよびスケールアウト クラスタへの無停止の容量追加をサポートします。お客様は、コストのかかるデータ移行や停止なしで、NVMe や 32Gb FC などの最新テクノロジーにアップグレードできます。
- クラウド接続。ONTAP は、すべてのパブリック クラウドでソフトウェア定義ストレージ (ONTAP Select) とクラウド ネイティブ インスタンス (Google Cloud NetApp Volumes) のオプションを備えた、最もクラウドに接続されたストレージ管理ソフトウェアです。
- 新しいアプリケーションとの統合。ONTAP は、既存のエンタープライズ アプリケーションをサポートするのと同じインフラストラクチャを使用して、自律走行車、スマート シティ、インダストリー 4.0 などの次世代プラットフォームとアプリケーション向けにエンタープライズ グレードのデータ サービスを提供します。

## NetApp DataOps ツールキット

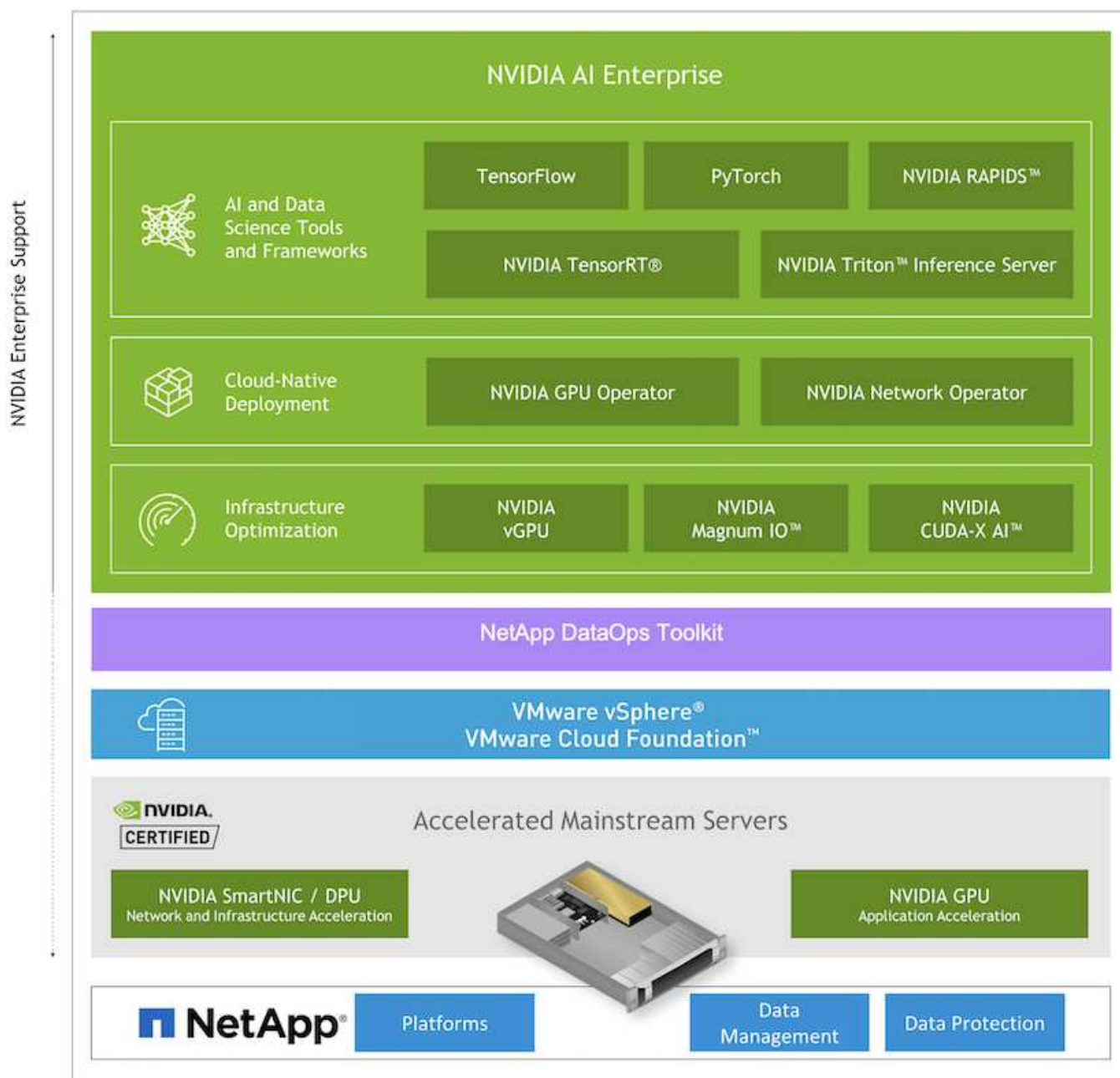
NetApp DataOps Toolkit は、高性能なスケールアウトNetAppストレージを基盤とする開発/トレーニング ワークスペースと推論サーバーの管理を簡素化する Python ベースのツールです。主な機能は次のとおりです。

- 高性能のスケールアウトNetAppストレージを活用した新しい大容量の JupyterLab ワークスペースを迅速にプロビジョニングします。
- エンタープライズ クラスのNetAppストレージを活用した新しいNVIDIA Triton Inference Server インスタンスを迅速にプロビジョニングします。
- 実験や迅速な反復を可能にするために、大容量の JupyterLab ワークスペースをほぼ瞬時に複製します。
- バックアップやトレーサビリティ/ベースライン作成のために、大容量の JupyterLab ワークスペースのスナップショットをほぼ瞬時に保存します。
- 大容量、高パフォーマンスのデータ ボリュームをほぼ瞬時にプロビジョニング、クローン作成、スナップショット作成します。

# アーキテクチャ

このソリューションは、NetApp、VMware、NVIDIA認定システムを採用した、実績のある使い慣れたアーキテクチャに基づいて構築されています。詳細については、次の表を参照してください。

コンポーネント	詳細
AIおよびデータ分析ソフトウェア	"VMware 向けNVIDIA AI エンタープライズ"
仮想化プラットフォーム	"VMware vSphere"
コンピューティングプラットフォーム	"NVIDIA認定システム"
データ管理プラットフォーム	"NetApp ONTAP"



# 初期セットアップ

このセクションでは、NetAppおよびVMwareでNVIDIA AI Enterpriseを利用するために実行する必要がある初期セットアップタスクについて説明します。

## 前提条件

このセクションで説明する手順を実行する前に、VMware vSphereとNetApp ONTAPがすでに導入されていることを前提としています。参照["NVIDIA AI エンタープライズ製品サポートマトリックス"](#)サポートされているvSphereバージョンの詳細については、こちらをご覧ください。参照["NetAppおよびVMware ソリューションのドキュメント"](#)VMware vSphereをNetApp ONTAPとともに導入する方法の詳細については、を参照してください。

## NVIDIA AI Enterprise Host ソフトウェアをインストールする

NVIDIA AI Enterprise Hostソフトウェアをインストールするには、["NVIDIA AI エンタープライズ クイック スタート ガイド"](#)。

# NVIDIA NGCソフトウェアを使用する

## セットアップ

このセクションでは、NVIDIA AI Enterprise 環境内でNVIDIA NGC エンタープライズソフトウェアを利用するために実行する必要がある初期セットアップタスクについて説明します。

## 前提条件

このセクションで説明する手順を実行する前に、NVIDIA AI Enterprise Hostソフトウェアが、["初期セットアップ"](#)ページ。

## vGPU を搭載した Ubuntu ゲスト VM を作成する

まず、vGPU を搭載した Ubuntu 20.04 ゲスト VM を作成する必要があります。vGPU を搭載した Ubuntu 20.04 ゲスト VM を作成するには、["NVIDIA AI エンタープライズ導入ガイド"](#)。

## NVIDIA ゲスト ソフトウェアのダウンロードとインストール

次に、前の手順で作成したゲスト VM 内に必要な NVIDIA ゲスト ソフトウェアをインストールする必要があります。ゲスト VM 内で必要な NVIDIA ゲスト ソフトウェアをダウンロードしてインストールするには、["NVIDIA AI エンタープライズ クイック スタート ガイド"](#)。



セクション 5.4 で概説されている検証タスクを実行する場合、ガイドの作成以降に CUDA コンテナ イメージが更新されているため、異なる CUDA コンテナ イメージ バージョン タグを使用する必要がある場合があります。検証では、「nvidia/cuda:11.0.3-base-ubuntu20.04」を使用しました。

## AI/分析フレームワークコンテナをダウンロード

次に、必要な AI または分析フレームワーク コンテナ イメージをNVIDIA NGC からダウンロードして、ゲスト VM 内で使用できるようにする必要があります。ゲストVM内でフレームワークコンテナをダウンロードするには、"[NVIDIA AI エンタープライズ導入ガイド](#)"。

## NetApp DataOps ツールキットのインストールと設定

次に、ゲスト VM 内に従来環境向けのNetApp DataOps Toolkit をインストールする必要があります。NetApp DataOps Toolkit を使用すると、ゲスト VM 内のターミナルから直接ONTAPシステム上のスケールアウト データ ボリュームを管理できます。ゲスト VM 内にNetApp DataOps Toolkit をインストールするには、次のタスクを実行します。

1. pip をインストールします。

```
$ sudo apt update
$ sudo apt install python3-pip
$ python3 -m pip install netapp-dataops-traditional
```

2. ゲスト VM ターミナルからログアウトし、再度ログインします。
3. NetApp DataOps ツールキットを構成します。この手順を完了するには、ONTAPシステムのAPIアクセスの詳細が必要になります。ストレージ管理者からこれらを取得する必要がある場合があります。



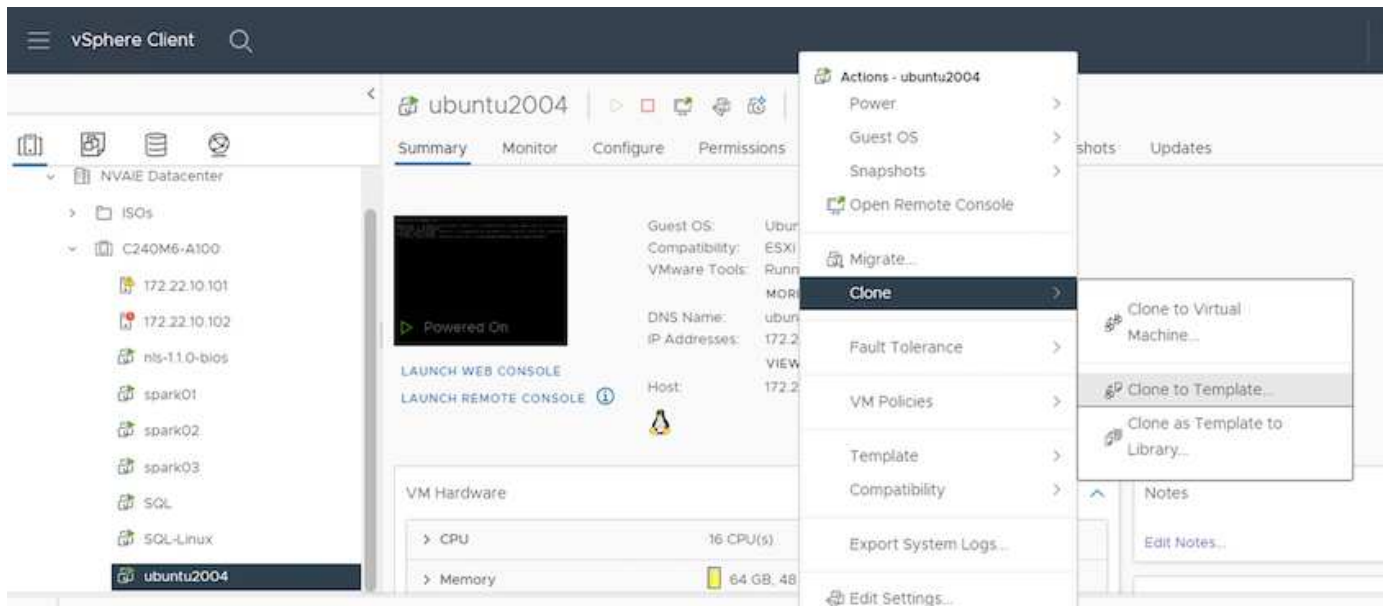
```
$ netapp_dataops_cli.py config

Enter ONTAP management LIF hostname or IP address (Recommendation: Use
SVM management interface): 172.22.10.10
Enter SVM (Storage VM) name: NVAIE-client
Enter SVM NFS data LIF hostname or IP address: 172.22.13.151
Enter default volume type to use when creating new volumes
(flexgroup/flexvol) [flexgroup]:
Enter export policy to use by default when creating new volumes
[default]:
Enter snapshot policy to use by default when creating new volumes
[none]:
Enter unix filesystem user id (uid) to apply by default when creating
new volumes (ex. '0' for root user) [0]:
Enter unix filesystem group id (gid) to apply by default when creating
new volumes (ex. '0' for root group) [0]:
Enter unix filesystem permissions to apply by default when creating new
volumes (ex. '0777' for full read/write permissions for all users and
groups) [0777]:
Enter aggregate to use by default when creating new FlexVol volumes:
aff_a400_01_NVME_SSD_1
Enter ONTAP API username (Recommendation: Use SVM account): admin
Enter ONTAP API password (Recommendation: Use SVM account):
Verify SSL certificate when calling ONTAP API (true/false): false
Do you intend to use this toolkit to trigger BlueXP Copy and Sync
operations? (yes/no): no
Do you intend to use this toolkit to push/pull from S3? (yes/no): no
Created config file: '/home/user/.netapp_dataops/config.json'.
```

## ゲストVMテンプレートを作成する

最後に、ゲスト VM に基づいて VM テンプレートを作成する必要があります。このテンプレートを使用すると、NVIDIA NGC ソフトウェアを利用するためのゲスト VM をすばやく作成できます。

ゲスト VM に基づいて VM テンプレートを作成するには、VMware vSphere にログインし、ゲスト VM 名を右クリックして、「クローン」を選択し、「テンプレートにクローン...」を選択して、ウィザードに従います。



## 使用例 - TensorFlow トレーニングジョブ

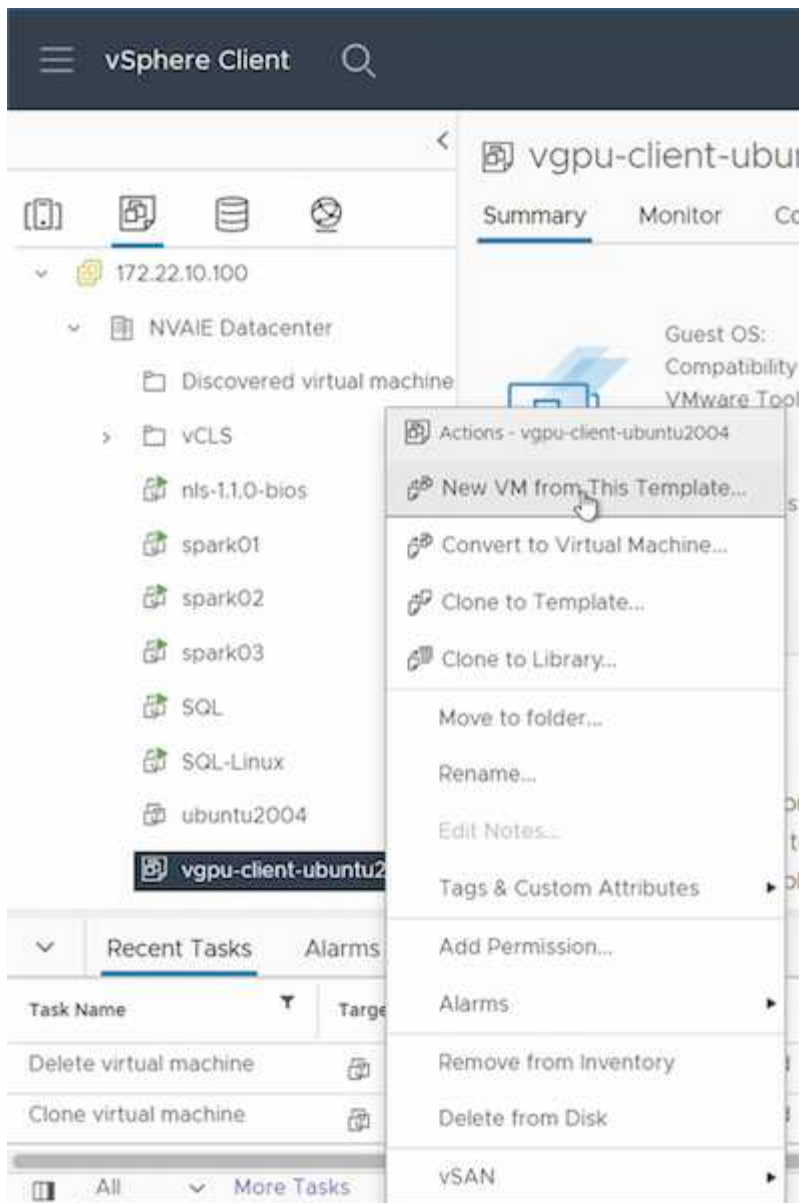
このセクションでは、NVIDIA AI Enterprise 環境内で TensorFlow トレーニング ジョブを実行するために実行する必要があるタスクについて説明します。

### 前提条件

このセクションで説明する手順を実行する前に、ゲストVMテンプレートが、["セットアップ"](#)ページ。

### テンプレートからゲストVMを作成する

まず、前のセクションで作成したテンプレートから新しいゲスト VM を作成する必要があります。テンプレートから新しいゲスト VM を作成するには、VMware vSphere にログインし、テンプレート名を右クリックして、「このテンプレートから新しい VM を作成...」を選択し、ウィザードに従います。



## データボリュームの作成とマウント

次に、トレーニング データセットを保存する新しいデータ ボリュームを作成する必要があります。NetApp DataOps Toolkit を使用すると、新しいデータ ボリュームをすばやく作成できます。次の例のコマンドは、容量が 2 TB の「imagenet」という名前のボリュームの作成を示しています。

```
$ netapp_dataops_cli.py create vol -n imagenet -s 2TB
```

データ ボリュームにデータを入力する前に、それをゲスト VM 内にマウントする必要があります。NetApp DataOps Toolkit を使用して、データ ボリュームをすばやくマウントできます。次のサンプルコマンドは、前の手順で作成されたボリュームのマウントを示しています。

```
$ sudo -E netapp_dataops_cli.py mount vol -n imagenet -m ~/imagenet
```

データボリュームを入力する

新しいボリュームがプロビジョニングされマウントされたら、トレーニング データセットをソースの場所から取得し、新しいボリュームに配置できます。これには通常、S3 または Hadoop データ レイクからデータを取得することが含まれ、場合によってはデータ エンジニアの支援が必要になることもあります。

## TensorFlow トレーニングジョブを実行する

これで、TensorFlow トレーニング ジョブを実行する準備が整いました。TensorFlow トレーニング ジョブを実行するには、次のタスクを実行します。

1. NVIDIA NGC エンタープライズ TensorFlow コンテナ イメージをプルします。

```
$ sudo docker pull nvcr.io/nvaie/tensorflow-2-1:22.05-tf1-nvaie-2.1-py3
```

2. NVIDIA NGC エンタープライズ TensorFlow コンテナのインスタンスを起動します。'-v' オプションを使用して、データ ボリュームをコンテナに接続します。

```
$ sudo docker run --gpus all -v ~/imagenet:/imagenet -it --rm  
nvcr.io/nvaie/tensorflow-2-1:22.05-tf1-nvaie-2.1-py3
```

3. コンテナ内で TensorFlow トレーニング プログラムを実行します。次のサンプル コマンドは、コンテナ イメージに含まれているサンプル ResNet-50 トレーニング プログラムの実行を示しています。

```
$ python ./nvidia-examples/cnn/resnet.py --layers 50 -b 64 -i 200 -u  
batch --precision fp16 --data_dir /imagenet/data
```

## 詳細情報の入手方法

このドキュメントに記載されている情報の詳細については、次のドキュメントや Web サイトを参照してください。

- NetApp ONTAP データ管理ソフトウェア — ONTAP 情報ライブラリ

<http://mysupport.netapp.com/documentation/productlibrary/index.html?productID=62286>

- NetApp DataOps ツールキット

["https://github.com/NetApp/netapp-dataops-toolkit"](https://github.com/NetApp/netapp-dataops-toolkit)

- NVIDIA AI Enterprise と VMware

["https://www.nvidia.com/en-us/data-center/products/ai-enterprise/"](https://www.nvidia.com/en-us/data-center/products/ai-enterprise/)

## 謝辞

- Bobby Oommen 氏、シニア マネージャー、 NetApp
- Ramesh Isaac、 NetAppシステム管理者
- ロニー・ダニエル、 NetAppテクニカル マーケティング エンジニア

## 著作権に関する情報

Copyright © 2025 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。