



Red Hat OpenShift Data Science 1

OpenShift Data Science のインストール

Red Hat OpenShift Cluster Manager を使用した OpenShift Dedicated クラスターの
アドオンとしての Red Hat OpenShift Data Science のインストール

Red Hat OpenShift Data Science 1 OpenShift Data Science のインストール

Red Hat OpenShift Cluster Manager を使用した OpenShift Dedicated クラスターのアドオンとしての Red Hat OpenShift Data Science のインストール

Enter your first name here. Enter your surname here.

Enter your organisation's name here. Enter your organisational division here.

Enter your email address here.

法律上の通知

Copyright © 2022 | You need to change the HOLDER entity in the en-US/Installing_OpenShift_Data_Science.ent file |.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux[®] is the registered trademark of Linus Torvalds in the United States and other countries.

Java[®] is a registered trademark of Oracle and/or its affiliates.

XFS[®] is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL[®] is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js[®] is an official trademark of Joyent. Red Hat is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack[®] Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

概要

Red Hat Cluster Manager を使用して、Red Hat OpenShift Data Science を OpenShift Dedicated クラスターのアドオンとしてインストールします。

目次

はじめに	3
第1章 OPENSIFT DATA SCIENCE のアーキテクチャー	4
第2章 OPENSIFT DATA SCIENCE のデプロイの概要	6
第3章 OPENSIFT DATA SCIENCE の要件	7
第4章 OPENSIFT DEDICATED のアイデンティティプロバイダーの設定	8
4.1. OPENSIFT DATA SCIENCE のアイデンティティ管理オプション	9
第5章 OPENSIFT DEDICATED の管理者ユーザーの追加	10
第6章 OPENSIFT DEDICATED への OPENSIFT DATA SCIENCE のインストール	11
第7章 OPENSIFT DATA SCIENCE での GPU サポートの有効化	13
第8章 ユーザーによるインスタンスアドレスの共有	16
第9章 関連情報	17

はじめに

本リリースに関連するサービスおよびライフサイクルの情報については、以下のドキュメントを参照してください。

- [OpenShift Data Science Service Definition](#)
- [OpenShift Data Science Life Cycle](#)

第1章 OPENSIFT DATA SCIENCE のアーキテクチャー

Red Hat OpenShift Data Science は、完全に Red Hat が管理するクラウドサービスであり、Red Hat OpenShift Dedicated および Red Hat OpenShift Service on Amazon Web Services (ROSA) 上のアドオンとして利用できます。

OpenShift Data Science は、以下のコンポーネントとサービスを統合します。

- サービス層:

OpenShift Data Science ダッシュボード

OpenShift Data Science 環境で利用可能でインストールされているアプリケーションと、チュートリアル、クイックスタートの例、ドキュメントなどの学習リソースを表示する、お客様向けのダッシュボード。

JupyterHub (Red Hat 管理)

データサイエンティストが独自のノートブックサーバー環境を構成し、JupyterLab で機械学習モデルを開発できるようにする Red Hat 管理アプリケーション。

- 管理層:

Red Hat OpenShift Data Science Operator

OpenShift Data Science の一部であるすべてのコンポーネントとサブ Operator をデプロイおよび維持するメタ Operator。

サービスの監視

Alertmanager、Prometheus、Grafana は連携して、OpenShift Data Science からメトリクスを収集し、モニターリングと請求の目的で役立つ方法でそれらのメトリクスを整理および表示します。

- データ層:

JupyterHub データベース

JupyterHub のすべてのユーザーデータを保存するデータベース。

Cluster Manager に OpenShift Data Science アドオンをインストールすると、以下の新しいプロジェクトが作成されます。

- **redhat-ods-operator** プロジェクトには、OpenShift Data Science Operator が含まれていません。
- **redhat-ods-applications** プロジェクトは、OpenShift Data Science のダッシュボードおよびその他の必要なコンポーネントをインストールします。
- **redhat-ods-monitoring** プロジェクトには、監視と請求のためのサービスが含まれています。
- **rhods-notebooks** プロジェクトは、ノートブック環境がデフォルトでデプロイされる場所です。

データサイエンティストは、機械学習モデルを使用する OpenShift アプリケーション用の追加プロジェクトを作成する必要があります。

ダッシュボード上のアプリケーションのカードで特に指示されていない限り、OpenShift Data Science アドオンに関連付けられた namespace に独立系ソフトウェアベンダー (ISV) アプリケーションをインストールしないでください。

関連情報

目次

- [OpenShift Data Science のインストール](#)

第2章 OPENSIFT DATA SCIENCE のデプロイの概要

このセクションでは、データサイエンティスト向けに Red Hat OpenShift Data Science を開発環境およびテスト環境としてデプロイする方法を説明します。

OpenShift Data Science のインストールには主に、以下のタスクが必要です。

1. OpenShift Dedicated クラスターがすべての要件を満たしていることを確認します。
2. OpenShift Dedicated のアイデンティティプロバイダーを設定します。
3. OpenShift Dedicated の管理者ユーザーを追加します。詳細は、[OpenShift Data Science のユーザーの追加](#) を参照してください。
4. OpenShift Data Science アドオンをインストールします。詳細は、[OpenShift Dedicated への OpenShift Data Science のインストール](#) を参照してください。
5. OpenShift Data Science へのユーザーアクセスを提供するために、ユーザーおよび管理者グループを設定します。
6. OpenShift Data Science をデプロイした OpenShift Dedicated クラスターの URL をユーザーに提供します。

第3章 OPENSIFT DATA SCIENCE の要件

Red Hat OpenShift Data Science のサポートを受けるには、お使いの環境が特定の要件を満たす必要があります。

インストール要件

OpenShift Data Science を Red Hat OpenShift Dedicated クラスターまたは Red Hat OpenShift Service on Amazon Web Services (ROSA) クラスターにインストールする前に、以下の要件を満たす必要があります。

- **Red Hat カスタマーアカウント**
OpenShift Cluster Manager (<http://console.redhat.com/openshift>) に移動し、ログインするか、新しいアカウントに登録します。
- **製品サブスクリプション**
以下の製品およびアドオンのサブスクリプション:
 - Red Hat OpenShift Dedicated または ROSA
 - Red Hat OpenShift Data Science アドオン

新規サブスクリプションを購入する場合は、Red Hat アカウントマネージャーにお問い合わせください。アカウントマネージャーがない場合は <https://cloud.redhat.com/products/dedicated/contact/> でフォームに入力して要求してください。

- **OpenShift Dedicated クラスター**
OpenShift Dedicated クラスター [の作成についての OpenShift Dedicated ドキュメントに従って](#)、既存のクラスターを使用するか、または新規クラスターを作成します。

アドオンをインストールするときに OpenShift Data Science で使用できるように、クラスターには少なくとも2つのワーカーノードと少なくとも8つの CPU および 32 GiB の RAM (インスタンスタイプ **m5.2xlarge** 以上) が必要です。この要件を満たしていないとインストールプロセスが開始に失敗し、エラーが表示されます。

デフォルトで、クラスターは1つのマシンプールを使用して作成されます。最小要件に対応するために、追加のマシンプールを追加するか、またはデフォルトプールを変更することができます。ただし、最小リソース要件は、クラスター内の単一のマシンプールによって満たされる必要があります。複数のマシンプールのリソースを使用して要件を満たすことはできません。詳細は、[OpenShift Dedicated でのマシンプールの作成](#) について参照してください。

- **ROSA クラスターでは、AWS Identity および Access Management の認証情報**
AWS Security Token Service (STS) を使用する ROSA クラスターに OpenShift Data Science をインストールすることはできません。ROSA に OpenShift Data Science をインストールする場合は、AWS Identity and Access Management (IAM) 認証情報のみを使用する必要があります。STS なしでのデプロイに関するアドバイスについては、ROSA ドキュメントを参照してください：[Deploying ROSA without AWS STS](#) を参照してください。

OpenShift Dedicated の場合と同様に、デフォルトのマシンプールは ROSA クラスターのインストール時に作成されます。インストール後に、OpenShift Cluster Manager または ROSA CLI(rosa)を使用してクラスターの追加のマシンプールを作成するか、または追加のマシンプールを追加するか、または最小要件に対応するためにデフォルトプールを変更することができます。ただし、最小リソース要件は、クラスター内の単一のマシンプールによって満たされる必要があります。複数のマシンプールのリソースを使用して要件を満たすことはできません。詳細は、[ROSA でのマシンプールの作成](#) について参照してください。

第4章 OPENSIFT DEDICATED のアイデンティティプロバイダーの設定

OpenShift Dedicated クラスターのアイデンティティプロバイダーを、ユーザーおよびグループを管理するように設定します。

重要

同じユーザー名が複数のプロバイダーに存在する場合に、複数の OpenShift アイデンティティプロバイダーを追加すると問題が生じる可能性があります。

MappingMethod が **claim** (アイデンティティプロバイダーのデフォルトマッピングメソッド) に設定され、複数のプロバイダーに同じユーザー名に関連付けられた認証情報がある場合は、OpenShift にログインするのに使用される最初のプロバイダーは、アイデンティティプロバイダーが設定される順序に関係なく、そのユーザーで機能する認証情報になります。

マッピング方法の詳細については、OpenShift Dedicated ドキュメントの [アイデンティティプロバイダーパラメーター](#) を参照してください。

前提条件

- OpenShift Cluster Manager の認証情報 (<https://console.redhat.com/openshift/>)。
- 既存の OpenShift Dedicated クラスター。

手順

1. OpenShift Cluster Manager にログインします (<https://console.redhat.com/openshift/>)。
2. **Clusters** をクリックします。 **Clusters** ページが開きます。
3. 設定するクラスターの名前をクリックします。
4. **Access control** タブをクリックします。
5. **Identity providers** をクリックします。
6. **Add identity provider** をクリックします。
 - a. **Identity Provider** 一覧からプロバイダーを選択します。
 - b. 選択した ID プロバイダーに関連する残りのフィールドに入力します。詳細は、[アイデンティティプロバイダーの設定](#) を参照してください。
7. **Confirm** をクリックします。

検証

- 設定されたアイデンティティプロバイダーが **Cluster details** ページの **Access control** タブに表示されるようになりました。

関連情報

- [アイデンティティプロバイダーの設定](#)

- [OpenShift Data Science のアイデンティティ管理オプション](#)
- [LDAP グループの同期](#)

4.1. OPENSIFT DATA SCIENCE のアイデンティティ管理オプション

Red Hat OpenShift Data Science は、Red Hat OpenShift Dedicated および Red Hat OpenShift Service on Amazon Web Services (ROSA) と同じ認証システムをサポートします。

詳細は、クラスターの適切なドキュメントを確認してください。

関連情報

- [OpenShift Dedicated でサポートされるアイデンティティプロバイダー](#)
- [ROSA でサポートされているアイデンティティプロバイダー](#)

第5章 OPENSIFT DEDICATED の管理者ユーザーの追加

データサイエンティストユーザー向けに OpenShift Data Science をインストールして設定する前に、管理ユーザーを定義する必要があります。管理ユーザーのみが OpenShift Data Science をインストールして設定できます。

前提条件

- OpenShift Cluster Manager の認証情報 (<https://console.redhat.com/openshift/>)。
- アイデンティティプロバイダーが設定された既存の OpenShift Dedicated クラスタ。

手順

1. OpenShift Cluster Manager にログインします (<https://console.redhat.com/openshift/>)。
2. **Clusters** をクリックします。 **Clusters** ページが開きます。
3. 設定するクラスタの名前をクリックします。
4. **Access control** タブをクリックします。
5. **Cluster Roles and Access** をクリックします。
6. **Cluster administrative users** で **Add user** ボタンをクリックします。
Add cluster user ポップオーバーが表示されます。
7. **User ID** フィールドにユーザー名を入力します。
8. ユーザーに適した **Group** を選択します。



重要

このユーザーがアイデンティティプロバイダーの既存のグループを使用して OpenShift Data Science アクセスを制御する必要がある場合は、**cluster-admin** を選択します。

これらのユーザータイプの詳細は、OpenShift Dedicated ドキュメントの [クラスタの管理](#) を確認してください。

9. **Add user** をクリックします。

検証

- ユーザー名と選択されたグループは、**クラスタ管理ユーザー** のリストに表示されます。

関連情報

- [OpenShift Dedicated クラスタの管理](#)

第6章 OPENSIFT DEDICATED への OPENSIFT DATA SCIENCE のインストール

Red Hat OpenShift Data Science は、Red Hat OpenShift Cluster Manager を使用して Red Hat OpenShift Dedicated クラスタにアドオンとしてインストールできます。

前提条件

- OpenShift Data Science のエンタイトルメントの購入。
- OpenShift Cluster Manager の認証情報 (<https://console.redhat.com/openshift/>)。
- OpenShift Dedicated クラスタへの管理者アクセス。

手順

1. OpenShift Cluster Manager にログインします (<https://console.redhat.com/openshift/>)。
2. **Clusters** をクリックします。
Clusters ページが開きます。
3. OpenShift Data Science をインストールするクラスタの名前をクリックします。
クラスタの **Details** ページが開きます。
4. **Add-ons** タブをクリックし、**Red Hat OpenShift Data Science** カードを見つけます。
5. **Install** をクリックします。**Configure Red Hat OpenShift Data Science** ペインが表示されま
す。
6. **Notification email** フィールドに、サービス停止などの Red Hat OpenShift Data Science の状
態についての重要なアラートを受け取るメールアドレスを入力します。
7. **Install** をクリックします。

検証

- OpenShift Cluster Manager では、クラスタの **Add-ons** タブで、OpenShift Data Science
カードに以下のいずれかの状態が表示されていることを確認します。
 - **Installing** - インストールが進行中です。これが **Installed** になるのを待ちます。これには
約 30 分かかります。
 - **Installed** - インストールが完了し、**View in console** ボタンが表示されていることを確認し
ます。
- OpenShift Dedicated で **Home** → **Projects** をクリックし、以下のプロジェクトの namespace
が表示され、**Active** として一覧表示されていることを確認します。
 - **redhat-ods-applications**
 - **redhat-ods-monitoring**
 - **redhat-ods-operator**
 - **rhods-notebooks**

関連情報

- [OpenShift Data Science のユーザーの追加](#)

第7章 OPENSIFT DATA SCIENCE での GPU サポートの有効化

データサイエンティストがモデルで計算量の多いワークロードを利用できるようにするために、OpenShift Data Science でグラフィックスプロセッシングユニット (GPU) を有効にすることができます。OpenShift Data Science で GPU を使用できるようにするには、OpenShift Data Science をインストールした後、NVIDIA GPU アドオンをインストールする必要があります。このアドオンは、クラスター内の GPU 対応ワーカーノードを見つけて有効にし、GPU インスタンスタイプを選択できるようにします。NVIDIA GPU アドオンをインストールし、クラスターに GPU 対応のワーカーノードがあることを確認したら、データサイエンティストは、JupyterHub で GPU 対応のノートブックの1つを、データサイエンスの仕事に必要な GPU の数とともに選択できます。

Red Hat は、[nvidia.com/gpu NoSchedule](https://nvidia.com/gpu/NoSchedule) ティントのある GPU ノード用に別のマシンプールを使用することをお勧めします。既存のマシンプールを編集してこのティントを追加する場合は、最初にマシンプールをゼロノードにスケールダウンしてから、マシンプールを必要なノード数まで増やす必要があります。これにより、新しいティントがマシンプール内のすべてのノードに確実に適用されます。マシンプール内のすべてのノードで一貫した動作を確保するために、Red Hat ではマシンプールのスケールを速やかに増加させることを推奨しています。ノードをゼロにスケールするとデプロイメントに悪影響を及ぼすため、Red Hat では、適切な時間を選択するときにサービスの使用パターンを考慮しながら、このアクションをできるだけ早く実行することを推奨しています。

前提条件

- OpenShift Cluster Manager の認証情報 (<https://console.redhat.com/openshift/>) がある。
- OpenShift Dedicated の **cluster-admins** ユーザーグループに属している。
- OpenShift Data Science および NVIDIA GPU アドオンの要件を満たすのに十分なリソースを含むクラスターをプロビジョニングしている。
- Red Hat OpenShift Data Science をインストールしてログインしている。
- OpenShift CLI (oc) をインストールしてログインしている必要がある。

手順

1. OpenShift Cluster Manager でクラスターにナビゲートします。
 - a. OpenShift Cluster Manager にログインします (<https://console.redhat.com/openshift/>)。
 - b. **Clusters** をクリックします。
Clusters ページが開きます。
 - c. OpenShift Data Science をインストールしたクラスターの名前をクリックします。
クラスターの **Details** ページが開きます。
2. GPU を備えたノード用のマシンプールを追加します。
 - a. **Machine pools** タブをクリックします。
 - b. **Add machine pool** ボタンをクリックします。
Add machine pool ウィンドウが開きます。
 - c. **Machine pool name** を指定します。
 - d. **ワーカーノードインスタンスタイプ** を設定します。インスタンスタイプが1つ以上の GPU を提供していることを確認してください。

- e. 少なくとも1つの **ワーカーノード数** を設定します。
- f. **Edit node labels and taints** をクリックして、**Node labels** セクションを展開します。
- g. **Taints** の下に、**nvidia.com/gpu** の **Key** と、**NoSchedule** の **Effect** を使用してテイントを追加します。**Value** は、**true**などの任意の文字列に設定できます。



注記

テイントを設定する際は、テイントが誤字脱字なく正しく宣言されていることを確認する必要があります。

- h. **Add machine pool** をクリックします。
マシンプールが作成されます。
- i. 指定した **Taint** がマシンプールの **Details** ページに表示されていることを確認します (例:**nvidia.com/gpu=true:NoSchedule**)。
 - a. **Add-ons** タブをクリックします。
 - b. **NVIDIA GPU Operator** カードをクリックします。
 - c. **Install** をクリックします。

検証

- OpenShift Cluster Manager で、クラスターの **Add-ons** タブで、NVIDIA GPU Operator がインストールされていることを確認します。
- OpenShift Dedicated Web コンソールの **Compute** → **Nodes** で、新しいマシンプールの各ノードに **nvidia.com/gpu** テイントが設定されていることを確認します (例:**nvidia.com/gpu=true:NoSchedule**)。
 - **Workloads** → **ConfigMaps** ページの **redhat-ods-applications** プロジェクトにある **jupyterhub-singleuser-profiles** ConfigMap には、次の **NoSchedule** 許容値が含まれています。

```
gpuTypes:
- type: gpu_one
  node_tolerations:
  - key: provider
    operator: Equal
    value: gpu-node
    effect: NoSchedule
  # This is the default NoSchedule toleration that is supported by the NVIDIA GPU operator
- type: nvidia_gpu
  node_tolerations:
  - key: "nvidia.com/gpu"
    operator: Exists
    effect: NoSchedule
```

- GPU 対応機能が Red Hat OpenShift Data Science で利用可能であることを確認します。
 - nvidia-device-plugin-validator ログを確認して検証します。OpenShift CLI で、次のコマン

ドを入力します。

```
oc logs nvidia-device-plugin-validator-<alpha-numeric-string> -n redhat-gpu-operator
```

ここで、**<alpha-numeric-string>** は、ランダムに生成された英数字の文字列になります。

検証が正常に行われると、以下の応答が返されます。

```
device-plugin validation is successful
```

- Red Hat では、GPU 対応モデルが Red Hat OpenShift Data Science で正常に実行できることを確認するために、サンプル GPU アプリケーションを実行することを推奨しています。詳細は [Running a sample GPU application](#) を参照してください。
- 関連する Pod 内で **nvidia-smi** コマンドを実行して、サンプルプロジェクトの GPU 使用率をテストします。詳細は、[Getting information about the GPU](#) を参照してください。

関連情報

- [OpenShift Dedicated で利用可能なアドオンサービス](#)
- [NVIDIA GPU Operator](#)
- [テイントおよび容認を使用したインフラストラクチャーノードのワークロードのバインディング](#)


第8章 ユーザーによるインスタンスアドレスの共有

ユーザーを Red Hat OpenShift Data Science に追加した後に、インスタンスアドレスをそれらのユーザーと共有し、それらをデータモデルでログインして作業できるようにします。

前提条件

- OpenShift Data Science が OpenShift Dedicated クラスターにインストールされている。
- OpenShift Data Science のユーザーグループに 1 人以上のユーザーを追加している。

手順

1. OpenShift Dedicated Web コンソールにログインします。
2. アプリケーションランチャー () をクリックします。
3. **Red Hat OpenShift Data Science** を右クリックし、OpenShift Data Science インスタンスの URL をコピーします。
4. OpenShift Data Science にログインできるようにするため、このインスタンス URL をデータサイエンティストに提供します。

検証

- 自身およびユーザーがインスタンス URL を使用して OpenShift Data Science にログインできることを確認します。

関連情報

- [OpenShift Data Science へのログイン](#)

第9章 関連情報

- [OpenShift Data Science のアイデンティティ管理オプション](#)
- [OpenShift Data Science のユーザーの追加](#)