Red Hat OpenShift Data Science self-managed 2.4

Installing and uninstalling OpenShift Data Science self-managed

Install and uninstall Red Hat OpenShift Data Science self-managed

Red Hat OpenShift Data Science self-managed 2.4 Installing and uninstalling OpenShift Data Science self-managed

Install and uninstall Red Hat OpenShift Data Science self-managed

Legal Notice

Abstract

Install and uninstall Red Hat OpenShift Data Science self-managed on your OpenShift Container Platform cluster.

Table of Contents

PREFACE

Learn how to use both the OpenShift command-line interface and web console to install Red Hat OpenShift Data Science self-managed on your OpenShift Container Platform cluster. To uninstall the product, learn how to use the recommended command-line interface (CLI) method.

> **NOTE**
>
> Red Hat recommends that you install only one instance of OpenShift Data Science on your cluster.
>
> Installing the Red Hat OpenShift Data Science Operator on the same cluster as the OpenShift Data Science add-on is not recommended or supported.

CHAPTER 1. ARCHITECTURE OF OPENSHIFT DATA SCIENCE SELF-MANAGED

Red Hat OpenShift Data Science self-managed is an Operator that is available on a self-managed environment, such as Red Hat OpenShift Container Platform.

OpenShift Data Science integrates the following components and services:

- At the service layer:

  OpenShift Data Science dashboard

  A customer-facing dashboard that shows available and installed applications for the OpenShift Data Science environment as well as learning resources such as tutorials, quick starts, and documentation. Administrative users can access functionality to manage users, clusters, notebook images, and model-serving runtimes. Data scientists can use the dashboard to create projects to organize their data science work.

  Model serving

  Data scientists can deploy trained machine-learning models to serve intelligent applications in production. After deployment, applications can send requests to the model using its deployed API endpoint.

  Data science pipelines

  Data scientists can build portable machine learning (ML) workflows with data science pipelines, using Docker containers. This enables your data scientists to automate workflows as they develop their data science models.

  Jupyter (self-managed)

  A self-managed application that allows data scientists to configure their own notebook server environment and develop machine learning models in JupyterLab.

  Distributed workloads

  Data scientists can use multiple nodes in parallel to train machine-learning models or process data more quickly. This approach significantly reduces the task completion time, and enables the use of larger datasets and more complex models.

  > **IMPORTANT**
  >
  > The distributed workloads feature is currently available in Red Hat OpenShift Data Science 2.4 as a Technology Preview feature only. Technology Preview features are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.
  >
  > For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope .

- At the management layer:

  The Red Hat OpenShift Data Science Operator

  A meta-operator that deploys and maintains all components and sub-operators that are part of OpenShift Data Science.

  Monitoring services

  Prometheus gathers metrics from OpenShift Data Science for monitoring purposes.

When you install the OpenShift Data Science Operator in the OpenShift Container Platform cluster, the following new projects are created:

- The **redhat-ods-operator** project contains the OpenShift Data Science operator.

- The **redhat-ods-applications** project installs the dashboard and other required components of OpenShift Data Science.

- The **redhat-ods-monitoring** project contains services for monitoring.

- The **rhods-notebooks** project is where notebook environments are deployed by default.

You or your data scientists must create additional projects for the applications that will use your machine learning models.

Do not install independent software vendor (ISV) applications in namespaces associated with OpenShift Data Science.

CHAPTER 2. OVERVIEW OF INSTALLING AND DEPLOYING OPENSHIFT DATA SCIENCE

Red Hat OpenShift Data Science is a platform for data scientists and developers of artificial intelligence (AI) applications. It provides a fully supported environment that lets you rapidly develop, train, test, and deploy machine learning models on-premises and/or in the public cloud.

OpenShift Data Science is provided as a managed cloud service add-on for Red Hat OpenShift or as self-managed software that you can install on-premise or in the public cloud on OpenShift.

For information on installing OpenShift Data Science as self-managed software, see Installing OpenShift Data Science self-managed.

There are two deployment options for Red Hat OpenShift Data Science as a managed cloud service add-on:

- **OpenShift Dedicated with a Customer Cloud Subscription on Amazon Web Services or Google Cloud Platform**
  OpenShift Dedicated is a complete OpenShift Container Platform cluster provided as a cloud service, configured for high availability, and dedicated to a single customer. OpenShift Dedicated is professionally managed by Red Hat and hosted on Amazon Web Services (AWS) or Google Cloud Platform (GCP). The Customer Cloud Subscription (CCS) model allows Red Hat to deploy and manage clusters into a customer's AWS or GCP account. Contact your Red Hat account manager to get OpenShift Dedicated through a CCS.

- **Red Hat OpenShift Service on AWS (ROSA)**
  ROSA is a fully-managed, turnkey application platform that allows you to focus on delivering value to your customers by building and deploying applications. You subscribe to the service directly from your AWS account.

Installing OpenShift Data Science as a managed cloud service involves the following high-level tasks:

1. Confirm that your OpenShift Dedicated cluster meets all requirements.

2. Configure an identity provider for your OpenShift cluster.

3. Add administrative users for your OpenShift cluster.

4. Subscribe to the Red Hat OpenShift Data Science Add-on.
   For OpenShift Dedicated with a CCS for AWS or GCP, get a subscription through Red Hat.

   For ROSA, get a subscription through the AWS Marketplace.

5. Install the OpenShift Data Science Add-on.

6. Access the OpenShift Data Science dashboard.

7. Optionally, enable graphics processing units (GPUs) in OpenShift Data Science to ensure that your data scientists can use compute-heavy workloads in their models.

Installing OpenShift Data Science involves the following high-level tasks:

1. Confirm that your OpenShift Container Platform cluster meets all requirements. See Requirements for OpenShift Data Science self-managed .

2. Add administrative users for OpenShift Container Platform. See Adding administrative users for OpenShift Container Platform.

3. Install the OpenShift Data Science Operator. See Installing the Red Hat OpenShift Data Science Operator.

4. Install OpenShift Data Science components. See Installing and managing Red Hat OpenShift Data Science components.

5. Configure user and administrator groups to provide user access to OpenShift Data Science. See Adding users.

6. Access the OpenShift Data Science dashboard. See Accessing the OpenShift Data Science dashboard.

7. Optionally, enable graphics processing units (GPUs) in OpenShift Data Science to ensure that your data scientists can use compute-heavy workloads in their models. See Enabling GPU support in OpenShift Data Science.

CHAPTER 3. REQUIREMENTS FOR OPENSHIFT DATA SCIENCE SELF-MANAGED

Your environment must meet certain requirements to receive support for Red Hat OpenShift Data Science.

Installation requirements

You must meet the following requirements before you are able to install OpenShift Data Science on your Red Hat OpenShift Container Platform cluster.

- **Product subscriptions**

  - A subscription for Red Hat OpenShift Data Science self-managed
    Contact your Red Hat account manager to purchase new subscriptions. If you do not yet have an account manager, complete the form at https://www.redhat.com/en/contact to request one.

- **An OpenShift Container Platform cluster 4.11 or greater**

  - Use an existing cluster or create a new cluster by following the OpenShift Container Platform documentation: OpenShift Container Platform installation overview .
    Your cluster must have at least 2 worker nodes with at least 8 CPUs and 32 GiB RAM available for OpenShift Data Science to use when you install the Operator. To ensure that OpenShift Data Science is usable, additional cluster resources are required beyond the minimum requirements.

  - A default storage class that can be dynamically provisioned must be configured.
    Confirm that a default storage class is configured by running the **oc get storageclass** command. If no storage classes are noted with **(default)** beside the name, follow the OpenShift Container Platform documentation to configure a default storage class: Changing the default storage class . For more information about dynamic provisioning, see Dynamic provisioning.

  - Open Data Hub must not be installed on the cluster.
    For more information about managing the machines that make up an OpenShift cluster, see Overview of machine management.

- **An identity provider configured for OpenShift Container Platform**
  Access to the cluster as a user with the **cluster-admin** role; the **kubeadmin** user is not allowed.

  Red Hat OpenShift Data Science supports the same authentication systems as Red Hat OpenShift Container Platform. See Understanding identity provider configuration for more information on configuring identity providers.

- **Internet access**
  Along with Internet access, the following domains must be accessible during the installation of OpenShift Data Science self-managed:

  - cdn.redhat.com

  - subscription.rhn.redhat.com

  - registry.access.redhat.com

  - registry.redhat.io

- quay.io

For CUDA-based images, the following domains must be accessible:

- ngc.download.nvidia.cn

- developer.download.nvidia.com

- **OpenShift Pipelines operator installation**

  - The Red Hat OpenShift Pipelines operator enables support for installation of pipelines in a self-managed environment.
    Before you use data science pipelines in OpenShift Data Science, you must install the Red Hat OpenShift Pipelines Operator. For more information, see Installing OpenShift Pipelines. If your deployment is in a disconnected self-managed environment, see Red Hat OpenShift Pipelines Operator in a restricted environment.

  - Before you can execute a pipeline in a disconnected environment, you must mirror any images used by your pipelines to a private registry.

  - You can store your pipeline artifacts in an Amazon Web Services (AWS) Simple Storage Service (S3) bucket to ensure that you do not consume local storage. To do this, you must first configure write access to your S3 bucket on your AWS account.
    If you do not have access to Amazon S3 storage, you must configure your own storage solution for use with pipelines.

CHAPTER 4. ADDING ADMINISTRATIVE USERS FOR OPENSHIFT CONTAINER PLATFORM

Before you can install and configure OpenShift Data Science for your data scientist users, you must define administrative users. Only users with the **cluster-admin** role can install and configure OpenShift Data Science.

For more information about creating a cluster admin user, see Creating a cluster admin .

CHAPTER 5. INSTALLING THE RED HAT OPENSHIFT DATA SCIENCE OPERATOR

This section shows how to install the Red Hat OpenShift Data Science Operator on your OpenShift Container Platform cluster using the command-line interface (CLI) and the OpenShift web console.

> **NOTE**
>
> If you want to upgrade from version 1 of OpenShift Data Science rather than performing a new installation, see Upgrading from version 1 of OpenShift Data Science .

> **NOTE**
>
> If your OpenShift cluster uses a proxy to access the Internet, you can configure the proxy settings for the Red Hat OpenShift Data Science Operator. See Overriding proxy settings of an Operator for more information.

5.1. INSTALLING THE RED HAT OPENSHIFT DATA SCIENCE OPERATOR BY USING THE CLI

The following procedure shows how to use the OpenShift command-line interface (CLI) to install the Red Hat OpenShift Data Science Operator on your OpenShift Container Platform cluster. You must install the Operator before you can install OpenShift Data Science components on the cluster.

Prerequisites

- You have a running OpenShift Container Platform cluster, version 4.11 or greater, configured with a default storage class that can be dynamically provisioned.

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

- You have downloaded and installed the OpenShift command-line interface (CLI). See Installing the OpenShift CLI.

Procedure

1. Open a new terminal window.

2. In the OpenShift command-line interface (CLI), log in to your OpenShift Container Platform cluster as a cluster administrator, as shown in the following example:

   ```
   $ oc login <openshift_cluster_url> -u <admin_username> -p <password>
   ```

3. Create a namespace for installation of the Operator by performing the following actions:

   a. Create a namespace YAML file, for example, **rhods-operator-namespace.yaml**.

      ```
      apiVersion: v1
      kind: Namespace
      metadata:
        name: redhat-ods-operator ❶
      ```

1   **redhat-ods-operator** is the recommended namespace for the Operator.

b. Create the namespace in your OpenShift Container Platform cluster.

```
$ oc create -f rhods-operator-namespace.yaml
```

You see output similar to the following:

```
namespace/redhat-ods-operator created
```

4. Create an operator group for installation of the Operator by performing the following actions:

a. Create an **OperatorGroup** object custom resource (CR) file, for example, **rhods-operator-group.yaml**.

```
apiVersion: operators.coreos.com/v1
kind: OperatorGroup
metadata:
  name: rhods-operator
  namespace: redhat-ods-operator 1
```

1   You must specify the same namespace that you created earlier in this procedure.

b. Create the **OperatorGroup** object in your OpenShift Container Platform cluster.

```
$ oc create -f rhods-operator-group.yaml
```

You see output similar to the following:

```
operatorgroup.operators.coreos.com/rhods-operator created
```

5. Create a subscription for installation of the Operator by performing the following actions:

a. Create a **Subscription** object CR file, for example, **rhods-operator-subscription.yaml**.

```
apiVersion: operators.coreos.com/v1alpha1
kind: Subscription
metadata:
  name: rhods-operator
  namespace: redhat-ods-operator 1
spec:
  name: rhods-operator
  channel: stable 2
  source: redhat-operators
  sourceNamespace: openshift-marketplace
```

1   You must specify the same namespace that you created earlier in this procedure.

2   For **channel**, select a value of **stable**, **embedded**, or **alpha**. These subscription channels are described as follows:

stable

> In the **stable** channel, Red Hat provides updates for the Operator approximately every three weeks. The **stable** channel is intended for production use and provides functionally complete, generally available features (in addition to early-access features where noted in the documentation) that are supported with Red Hat production service level agreements (SLAs).

embedded

> The **embedded** channel provides updates for products that integrate Red Hat OpenShift Data Science. This includes IBM watsonx.ai. If this specific use case does not apply to your organization, choose **stable**. The **embdedded** channel is intended for production use and provides functionally complete, generally available features (in addition to early-access features where noted in the documentation) that are supported with Red Hat production service level agreements (SLAs).

alpha

> The **alpha** channel is intended for development use only. The channel provides development builds and early-access features.

> NOTE
>
> The development builds and early-access features that the **alpha** channel provides are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. Early-access features enable customers to test functionality and provide feedback during the development process. For production environments, choose **embedded** or **stable**, based on the preceding descriptions.
>
> For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope . For more information about the support scope of Red Hat Developer Preview features, see Developer Preview Features Support Scope .

> NOTE
>
> The **beta** channel is a legacy channel that will be removed in a future release. Do not select the **beta** channel for a new installation of the Operator.
>
> For more information about the lifecycle associated with each of the available subscription channels, see Red Hat OpenShift Data Science self-managed Life Cycle.

b. As described in the preceding step, ensure that the subscription channel you specify is appropriate for your organization's requirements.

c. Create the **Subscription** object in your OpenShift Container Platform cluster to install the Operator.

```
$ oc create -f rhods-operator-subscription.yaml
```

You see output similar to the following:

```
subscription.operators.coreos.com/rhods-operator created
```

Verification

- In the OpenShift Container Platform web console, click **Operators → Installed Operators** and confirm that the Red Hat OpenShift Data Science Operator shows one of the following statuses:

  - **Installing** - installation is in progress; wait for this to change to **Succeeded**. This might take several minutes.

  - **Succeeded** - installation is successful.

- In the web console, click **Home → Projects** and confirm that the following project namespaces are visible and listed as **Active**:

  - **redhat-ods-applications**

  - **redhat-ods-monitoring**

  - **redhat-ods-operator**

Additional resources

- [Installing and managing Red Hat OpenShift Data Science components](#)

- [Adding users](#)

- [Adding Operators to a cluster](#)

5.2. INSTALLING THE RED HAT OPENSHIFT DATA SCIENCE OPERATOR BY USING THE WEB CONSOLE

The following procedure shows how to use the OpenShift Container Platform web console to install the Red Hat OpenShift Data Science Operator on your cluster. You must install the Operator before you can install OpenShift Data Science components on the cluster.

Prerequisites

- You have a running OpenShift Container Platform cluster, version 4.11 or greater, configured with a default storage class that can be dynamically provisioned.

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

Procedure

1. Log in to the OpenShift Container Platform web console as a cluster administrator.

2. In the web console, click **Operators → OperatorHub**.

3. On the **OperatorHub** page, locate the Red Hat OpenShift Data Science Operator.

   a. Scroll through available Operators or type **Red Hat OpenShift Data Science** into the **Filter by keyword** box to find the Red Hat OpenShift Data Science Operator.

4. Select the Operator to display additional information.

5. Read the information about the Operator and click **Install**.

6. For **Update channel**, select **stable**, **embedded**, or **alpha**. These subscription channels are described as follows:

   stable

   > In the **stable** channel, Red Hat provides updates for the Operator approximately every three weeks. The **stable** channel is intended for production use and provides functionally complete, generally available features (in addition to early-access features where noted in the documentation) that are supported with Red Hat production service level agreements (SLAs).

   embedded

   > The **embedded** channel provides updates for products that integrate Red Hat OpenShift Data Science. This includes IBM watsonx.ai. If this specific use case does not apply to your organization, choose **stable**. The **embdedded** channel is intended for production use and provides functionally complete, generally available features (in addition to early-access features where noted in the documentation) that are supported with Red Hat production service level agreements (SLAs).

   alpha

   > The **alpha** channel is intended for development use only. The channel provides development builds and early-access features.

   > **NOTE**
   >
   > The development builds and early-access features that the **alpha** channel provides are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. Early-access features enable customers to test functionality and provide feedback during the development process. For production environments, choose **embedded** or **stable**, based on the preceding descriptions.
   >
   > For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope . For more information about the support scope of Red Hat Developer Preview features, see Developer Preview Features Support Scope .

   > **NOTE**
   >
   > The **beta** channel is a legacy channel that will be removed in a future release. Do not select the **beta** channel for a new installation of the Operator.
   >
   > For more information about the lifecycle associated with each of the available subscription channels, see Red Hat OpenShift Data Science self-managed Life Cycle.

7. For **Installation mode**, observe that the only available value is **All namespaces on the cluster (default)**. This installation mode makes the Operator available to all namespaces in the cluster.

8. For **Installed Namespace**, select **redhat-ods-operator (Operator recommended)**.

9. Under **Update approval**, select either **Automatic** or **Manual**.

10. Click **Install**.
    An installation pane opens. When the installation finishes, a check mark appears beside the Operator name in the installation pane.

Verification

- In the OpenShift Container Platform web console, click **Operators → Installed Operators** and confirm that the Red Hat OpenShift Data Science Operator shows one of the following statuses:

  - **Installing** - installation is in progress; wait for this to change to **Succeeded**. This might take several minutes.

  - **Succeeded** - installation is successful.

- In the web console, click **Home → Projects** and confirm that the following project namespaces are visible and listed as **Active**:

  - **redhat-ods-applications**

  - **redhat-ods-monitoring**

  - **redhat-ods-operator**

Additional resources

- [Installing and managing Red Hat OpenShift Data Science components](#)

- [Adding users](#)

- [Adding Operators to a cluster](#)

CHAPTER 6. UPGRADING FROM VERSION 1 OF OPENSHIFT DATA SCIENCE

As a cluster administrator, you can configure either automatic or manual upgrade from version 1 to version 2 of the Red Hat OpenShift Data Science Operator.

6.1. OVERVIEW OF UPGRADING OPENSHIFT DATA SCIENCE SELF-MANAGED

As a cluster administrator, you can configure either automatic or manual upgrades for the Red Hat OpenShift Data Science Operator.

- If you configure automatic upgrades, when a new version of the Red Hat OpenShift Data Science Operator is available, Operator Lifecycle Manager (OLM) automatically upgrades the running instance of your Operator without human intervention.

- If you configure manual upgrades, when a new version of the Red Hat OpenShift Data Science Operator is available, OLM creates an update request.
  A cluster administrator must manually approve the update request to update the Operator to the new version. See Manually approving a pending Operator upgrade for more information about approving a pending Operator upgrade.

- By default, the Red Hat OpenShift Data Science Operator follows a sequential update process. This means that if there are several versions between the current version and the version that you plan to upgrade to, Operator Lifecycle Manager (OLM) upgrades the Operator to each of the intermediate versions before it upgrades it to the final, target version. If you configure automatic upgrades, OLM automatically upgrades the Operator to the latest available version, without human intervention. If you configure manual upgrades, a cluster administrator must manually approve each sequential update between the current version and the final, target version.

- When you upgrade OpenShift Data Science, the upgrade process reloads any models that you previously deployed using ModelMesh, which is the default component for model-serving. The models are not available until this reload is complete.

- Red Hat supports the current release version and three previous release versions of OpenShift Data Science self-managed. For more information, see the Red Hat OpenShift Data Science self-managed Life Cycle knowledgebase article.

Additional resources

- Operator Lifecycle Manager workflow

6.2. CONFIGURING THE UPGRADE STRATEGY FOR OPENSHIFT DATA SCIENCE

As a cluster administrator, you can configure either an automatic or manual upgrade strategy for the Red Hat OpenShift Data Science Operator.

IMPORTANT

- By default, the Red Hat OpenShift Data Science Operator follows a sequential update process. This means that if there are several versions between the current version and the version that you intend to upgrade to, Operator Lifecycle Manager (OLM) upgrades the Operator to each of the intermediate versions before it upgrades it to the final, target version. If you configure automatic upgrades, OLM automatically upgrades the Operator to the *latest* available version, without human intervention. If you configure manual upgrades, a cluster administrator must manually approve each sequential update between the current version and the final, target version.

- When you upgrade OpenShift Data Science, the upgrade process reloads any models that you previously deployed using ModelMesh, which is the default component for model-serving. The models are not available until this reload is complete.

Prerequisites

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

- The Red Hat OpenShift Data Science Operator is installed.

Procedure

1. Log in to the OpenShift Container Platform cluster web console as a cluster administrator.

2. In the **Administrator** perspective, in the left menu, select **Operators → Installed Operators**.

3. Click the **Red Hat OpenShift Data Science** Operator.

4. Click the **Subscription** tab.

5. Under **Update approval**, click the pencil icon and select one of the following update strategies:

   - **Automatic**: New updates are installed as soon as they become available.

   - **Manual**: A cluster administrator must approve any new update before installation begins.

6. Click **Save**.

Additional resources

- For more information about the subscription channels that are available in version 2 of the Red Hat OpenShift Data Science Operator, see Installing the Red Hat OpenShift Data Science Operator.

- For more information about upgrading Operators that have been installed using OLM, see Updating installed Operators in the OpenShift Container Platform documentation.

6.3. CLEANING UP UNUSED RESOURCES FROM VERSION 1 OF RED HAT OPENSHIFT DATA SCIENCE

Version 1 of OpenShift Data Science created a set of Kubeflow Deployment Definition (that is, **KfDef**) custom resource instances on your OpenShift Container Platform cluster for various components of OpenShift Data Science. When you upgrade to version 2, these resources are no longer used and

require manual removal from your cluster. The following procedures shows how to remove unused **KfDef** instances from your cluster by using both the OpenShift command-line interface (CLI) and the web console.

6.3.1. Removing unused resources by using the CLI

The following procedure shows how to remove unused **KfDef** instances from the **redhat-ods-applications**, **redhat-ods-monitoring**, and **rhods-notebooks** projects in your OpenShift Container Platform cluster by using the OpenShift command-line interface (CLI). These resources become unused after you upgrade from version 1 to version 2 of OpenShift Data Science.

Prerequisites

- You upgraded from version 1 to version 2 of OpenShift Data Science.

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

Procedure

1. Open a new terminal window.

2. In the OpenShift command-line interface (CLI), log in to your on your OpenShift Container Platform cluster as a cluster administrator, as shown in the following example:

   ```
   $ oc login <openshift_cluster_url> -u system:admin
   ```

3. Delete any **KfDef** instances that exist in the **redhat-ods-applications** project.

   ```
   $ oc delete kfdef --all -n redhat-ods-applications --ignore-not-found || true
   ```

   For any **KfDef** instance that is deleted, the output is similar to the following example:

   ```
   kfdef.kfdef.apps.kubeflow.org "rhods-dashboard" deleted
   ```

   TIP

   If deletion of a **KfDef** instance fails to finish, you can force deletion of the object using the information in the "Force individual object removal when it has finalizers" section of the following Red Hat solution article: https://access.redhat.com/solutions/4165791.

4. Delete any **KfDef** instances in the **redhat-ods-monitoring** and **rhods-notebooks** projects by entering the following commands:

   ```
   $ oc delete kfdef --all -n redhat-ods-monitoring --ignore-not-found || true
   $ oc delete kfdef --all -n rhods-notebooks --ignore-not-found || true
   ```

Verification

- Check whether all **KfDef** instances have been removed from the **redhat-ods-applications**, **redhat-ods-monitoring**, and **rhods-notebooks** projects.

  ```
  $ oc get kfdef --all-namespaces
  ```

Verify that you see no **KfDef** instances listed in the **redhat-ods-applications**, **redhat-ods-monitoring**, or **rhods-notebooks** projects.

6.3.2. Removing unused resources by using the web console

The following procedure shows how to remove unused **KfDef** instances from the **redhat-ods-applications**, **redhat-ods-monitoring**, and **rhods-notebooks** projects in your OpenShift Container Platform cluster by using the OpenShift web console. These resources become unused after you upgrade from version 1 to version 2 of OpenShift Data Science.

Prerequisites

- You upgraded from version 1 to version 2 of OpenShift Data Science.

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

Procedure

1. Log in to the OpenShift Container Platform web console as a cluster administrator.

2. In the web console, click **Administration → CustomResourceDefinitions**.

3. On the **CustomResourceDefinitions** page, click the **KfDef** custom resource definition (CRD).

4. Click the **Instances** tab.
   The page shows all **KfDef** instances on the cluster.

5. Take note of any **KfDef** instances that exist in the **redhat-ods-applications**, **redhat-ods-monitoring**, and **rhods-notebooks** projects. These are the projects that you will clean up in the remainder of this procedure.

6. To delete a **KfDef** instance from the **redhat-ods-applications**, **redhat-ods-monitoring**, or **rhods-notebooks** project, click the action menu (⋮) beside the instance and select **Delete KfDef** from the list.

7. To confirm deletion of the instance, click **Delete**.

   TIP

   If deletion of a **KfDef** instance fails to finish, you can force deletion of the object using the information in the "Force individual object removal when it has finalizers" section of the following Red Hat solution article: https://access.redhat.com/solutions/4165791.
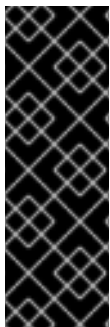
8. Repeat the preceding steps to delete all remaining **KfDef** instances that you see in the **redhat-ods-applications**, **redhat-ods-monitoring**, and **rhods-notebooks** projects.

CHAPTER 7. INSTALLING AND MANAGING RED HAT OPENSHIFT DATA SCIENCE COMPONENTS

The following procedures show how to use the command-line interface (CLI) and OpenShift Container Platform web console to install and manage components of Red Hat OpenShift Data Science on your OpenShift Container Platform cluster.

7.1. INSTALLING RED HAT OPENSHIFT DATA SCIENCE COMPONENTS BY USING THE CLI

The following procedure shows how to use the OpenShift command-line interface (CLI) to install specific components of Red Hat OpenShift Data Science on your OpenShift Container Platform cluster.

> **IMPORTANT**
>
> The following procedure describes how to create and configure a **DataScienceCluster** object to install Red Hat OpenShift Data Science components as part of a *new* installation. However, if you upgraded from version 1 of OpenShift Data Science, the upgrade process automatically created a default **DataScienceCluster** object. To inspect the default **DataScienceCluster** object and change the installation status of Red Hat OpenShift Data Science components, see Updating the installation status of Red Hat OpenShift Data Science components by using the web console.

Prerequisites

- The Red Hat OpenShift Data Science Operator is installed on your OpenShift Container Platform cluster. See Installing the Red Hat OpenShift Data Science Operator .

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

- You have downloaded and installed the OpenShift command-line interface (CLI). See Installing the OpenShift CLI.

Procedure

1. Open a new terminal window.

2. In the OpenShift command-line interface (CLI), log in to your on your OpenShift Container Platform cluster as a cluster administrator, as shown in the following example:

   ```
   $ oc login <openshift_cluster_url> -u <admin_username> -p <password>
   ```

3. Create a **DataScienceCluster** object custom resource (CR) file, for example, **rhods-operator-dsc.yaml**.

   ```
   apiVersion: datasciencecluster.opendatahub.io/v1
   kind: DataScienceCluster
   metadata:
     name: default-dsc
   spec:
     components:
       codeflare:
         managementState: "Removed"
   ```

```
      dashboard:
        managementState: "Removed"
      datasciencepipelines:
        managementState: "Removed"
      kserve:
        managementState: "Removed"
      modelmeshserving:
        managementState: "Removed"
      ray:
        managementState: "Removed"
      workbenches:
        managementState: "Removed"
```

4. In the **spec.components** section of the CR, for each OpenShift Data Science component shown, set the value of the **managementState** field to either **Managed** or **Removed**. These values are defined as follows:

   Managed

   The Operator actively manages the component, installs it, and tries to keep it active. The Operator will upgrade the component only if it is safe to do so.

   Removed

   The Operator actively manages the component but does not install it. If the component is already installed, the Operator will try to remove it.

   > **IMPORTANT**
   >
   > - The KServe component is part of a composite model-serving stack that is a Limited Availability feature. Limited Availability means that you can install and receive support for the feature only with specific approval from Red Hat. Without such approval, the feature is unsupported. Before enabling the KServe component, you must install some other, dependent Operators, including the Red Hat OpenShift Serverless and Red Hat OpenShift Service Mesh Operators. For more information, see Serving large language models .
   >
   > - You cannot install both the KServe and Model Mesh Serving components.
   >
   > - The CodeFlare and KubeRay components are Technology Preview features only. Technology Preview features are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process. For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope.
   >
   > - To learn how to configure the distributed workloads feature that uses the CodeFlare and KubeRay components, see Configuring distributed workloads.

5. Create the **DataScienceCluster** object in your OpenShift Container Platform cluster to install the specified OpenShift Data Science components.

```
$ oc create -f rhods-operator-dsc.yaml
```
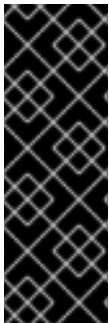
You see output similar to the following:

> datasciencecluster.datasciencecluster.opendatahub.io/default created

Verification

- In the OpenShift Container Platform web console, click **Workloads → Pods**. In the **Project** list at the top of the page, select **redhat-ods-applications**. In the applications namespace, confirm that there are running pods for each of the OpenShift Data Science components that you installed.

- In the web console, click **Operators → Installed Operators** and then perform the following actions:

    - Click the Red Hat OpenShift Data Science Operator.

    - Click the **Data Science Cluster** tab and select the **DataScienceCluster** object called **default-dsc**.

    - Select the **YAML** tab.

    - In the **installedComponents** section, confirm that the components you installed have a status value of **true**.

7.2. INSTALLING RED HAT OPENSHIFT DATA SCIENCE COMPONENTS BY USING THE WEB CONSOLE

The following procedure shows how to use the OpenShift Container Platform web console to install specific components of Red Hat OpenShift Data Science on your cluster.

IMPORTANT

The following procedure describes how to create and configure a **DataScienceCluster** object to install Red Hat OpenShift Data Science components as part of a *new* installation. However, if you upgraded from version 1 of OpenShift Data Science, the upgrade process automatically created a default **DataScienceCluster** object. To inspect the default **DataScienceCluster** object and change the installation status of Red Hat OpenShift Data Science components, see Updating the installation status of Red Hat OpenShift Data Science components by using the web console.

Prerequisites

- The Red Hat OpenShift Data Science Operator is installed on your OpenShift Container Platform cluster. See Installing the Red Hat OpenShift Data Science Operator .

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

Procedure

1. Log in to the OpenShift Container Platform web console as a cluster administrator.

2. In the web console, click **Operators → Installed Operators** and then click the Red Hat OpenShift Data Science Operator.

3. Create a **DataScienceCluster** object to install OpenShift Data Science components by performing the following actions:

   a. Click the **Data Science Cluster** tab.

   b. Click **Create DataScienceCluster**.

   c. For **Configure via**, select **YAML view**.
      An embedded YAML editor opens showing a default custom resource (CR) for the **DataScienceCluster** object.

   d. In the **spec.components** section of the CR, for each OpenShift Data Science component shown, set the value of the **managementState** field to either **Managed** or **Removed**. These values are defined as follows:

      Managed

         The Operator actively manages the component, installs it, and tries to keep it active. The Operator will upgrade the component only if it is safe to do so.

      Removed

         The Operator actively manages the component but does not install it. If the component is already installed, the Operator will try to remove it.

      IMPORTANT

      - The KServe component is part of a composite model-serving stack that is a Limited Availability feature. Limited Availability means that you can install and receive support for the feature only with specific approval from Red Hat. Without such approval, the feature is unsupported. Before enabling this component, you must install some other, dependent Operators, including the Red Hat OpenShift Serverless and Red Hat OpenShift Service Mesh Operators. For more information, see Serving large language models.

      - You cannot install both the KServe and Model Mesh Serving components.

      - The CodeFlare and KubeRay components are Technology Preview features only. Technology Preview features are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process. For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope.

      - To learn how to configure the distributed workloads feature that uses the CodeFlare and KubeRay components, see Configuring distributed workloads.

4. Click **Create**.

Verification

- On the **DataScienceClusters** page, click the **default-dsc** object and then perform the following actions:

  - Select the **YAML** tab.

  - In the **installedComponents** section, confirm that the components you installed have a status value of **true**.

- In the OpenShift Container Platform web console, click **Workloads → Pods** and then perform the following actions:

  - In the **Project** list at the top of the page, select the **redhat-ods-applications** project.

  - In the project, confirm that there are running pods for each of the OpenShift Data Science components that you installed.

7.3. UPDATING THE INSTALLATION STATUS OF RED HAT OPENSHIFT DATA SCIENCE COMPONENTS BY USING THE WEB CONSOLE

The following procedure shows how to use the OpenShift Container Platform web console to update the installation status of components of Red Hat OpenShift Data Science on your OpenShift Container Platform cluster.

IMPORTANT

If you upgraded from version 1 to version 2 of OpenShift Data Science, the upgrade process automatically created a default **DataScienceCluster** object and enabled several components of OpenShift Data Science. In this case, the following procedure shows how to do the following:

- Change the installation status of the existing Red Hat OpenShift Data Science components

- Add additional components to the **DataScienceCluster** object that were not available in version 1 of OpenShift Data Science

Prerequisites

- The Red Hat OpenShift Data Science Operator is installed on your OpenShift Container Platform cluster.

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

Procedure

1. Log in to the OpenShift Container Platform web console as a cluster administrator.

2. In the web console, click **Operators → Installed Operators** and then click the Red Hat OpenShift Data Science Operator.

3. Click the **Data Science Cluster** tab.

4. On the **DataScienceClusters** page, click the **default** object.

5. Click the **YAML** tab.

An embedded YAML editor opens showing the custom resource (CR) file for the **DataScienceCluster** object.

6. In the **spec.components** section of the CR, for each OpenShift Data Science component shown, set the value of the **managementState** field to either **Managed** or **Removed**. These values are defined as follows:

Managed

The Operator actively manages the component, installs it, and tries to keep it active. The Operator will upgrade the component only if it is safe to do so.

Removed

The Operator actively manages the component but does not install it. If the component is already installed, the Operator will try to remove it.
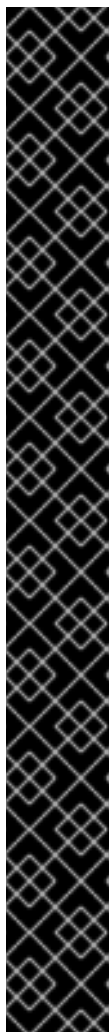
IMPORTANT

- If they are not already present in the CR file, you can install the CodeFlare, KServe, and KubeRay features by adding components called **codeflare**, **kserve**, and **ray** to the **spec.components** section of the CR and setting the **managementState** field for the components to **Managed**.

- The KServe component is part of a composite model-serving stack that is a Limited Availability feature. Limited Availability means that you can install and receive support for the feature only with specific approval from Red Hat. Without such approval, the feature is unsupported. Before enabling this component, you must install some other, dependent Operators, including the Red Hat OpenShift Serverless and Red Hat OpenShift Service Mesh Operators. For more information, see Serving large language models .

- You cannot install both the KServe and Model Mesh Serving components.

- The CodeFlare and KubeRay components are Technology Preview features only. Technology Preview features are not supported with Red Hat production service level agreements (SLAs) and might not be functionally complete. Red Hat does not recommend using them in production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process. For more information about the support scope of Red Hat Technology Preview features, see Technology Preview Features Support Scope.

- To learn how to configure the distributed workloads feature that uses the CodeFlare and KubeRay components, see Configuring distributed workloads.

7. Click **Save**.

Verification

- On the **DataScienceClusters** page, click the **default-dsc** object and then perform the following actions:

  - Select the **YAML** tab.

  - In the **installedComponents** section, confirm that the components you installed have a status value of **true**.

- In the OpenShift Container Platform web console, click **Workloads → Pods**. In the **Project** list at the top of the page, select **redhat-ods-applications**. In the applications namespace, confirm that there are running pods for each of the OpenShift Data Science components that you have installed.

CHAPTER 8. UNINSTALLING RED HAT OPENSHIFT DATA SCIENCE SELF-MANAGED

This section shows how to use the OpenShift command-line interface (CLI) to uninstall the Red Hat OpenShift Data Science Operator and any OpenShift Data Science components installed and managed by the Operator.

> **NOTE**
>
> Using the CLI is the recommended way to uninstall the Operator. Depending on your version of OpenShift Container Platform, using the web console to perform the uninstallation might not prompt you to uninstall all associated components. This could leave your cluster in an unclear state.

8.1. UNINSTALLING RED HAT OPENSHIFT DATA SCIENCE SELF-MANAGED BY USING THE CLI

The following procedure shows how to use the OpenShift command-line interface (CLI) to uninstall the Red Hat OpenShift Data Science Operator and any OpenShift Data Science components installed and managed by the Operator.

Prerequisites

- You have cluster administrator privileges for your OpenShift Container Platform cluster.

- You have downloaded and installed the OpenShift command-line interface (CLI). See Installing the OpenShift CLI.

- You have backed up the persistent disks or volumes used by your persistent volume claims (PVCs).

Procedure

1. Open a new terminal window.

2. In the OpenShift command-line interface (CLI), log in to your OpenShift Container Platform cluster as a cluster administrator, as shown in the following example:

   $ oc login *<openshift_cluster_url>* -u system:admin

3. Create a **ConfigMap** object for deletion of the Red Hat OpenShift Data Science Operator.

   $ oc create configmap delete-self-managed-odh -n redhat-ods-operator

4. To delete the **rhods-operator**, set the **addon-managed-odh-delete** label to **true**.

   $ oc label configmap/delete-self-managed-odh api.openshift.com/addon-managed-odh-delete=true -n redhat-ods-operator

5. When all objects associated with the Operator are removed, delete the **redhat-ods-operator** project.

   a. Set an environment variable for the **redhat-ods-applications** project.

```
$ PROJECT_NAME=redhat-ods-applications
```

b. Wait until the **redhat-ods-applications** project has been deleted.

```
$ while oc get project $PROJECT_NAME &> /dev/null; do
echo "The $PROJECT_NAME project still exists"
sleep 1
done
echo "The $PROJECT_NAME project no longer exists"
```

When the **redhat-ods-applications** project has been deleted, you see the following output.

```
The redhat-ods-applications project no longer exists
```

c. When the **redhat-ods-applications** project has been deleted, delete the **redhat-ods-operator** project.

```
$ oc delete namespace redhat-ods-operator
```

Verification

- Confirm that the **rhods-operator** subscription no longer exists.

```
$ oc get subscriptions --all-namespaces | grep rhods-operator
```

- Confirm that the following projects no longer exist.

    - **redhat-ods-applications**

    - **redhat-ods-monitoring**

    - **redhat-ods-operator**

    - **rhods-notebooks**

```
$ oc get namespaces | grep -e redhat-ods* -e rhods*
```

NOTE

The **rhods-notebooks** project was created only if you installed the workbenches component of OpenShift Data Science. See Installing and managing Red Hat OpenShift Data Science components.

CHAPTER 9. ACCESSING THE OPENSHIFT DATA SCIENCE DASHBOARD

After you have installed OpenShift Data Science and added users, you can access the URL for your OpenShift Data Science console and share the URL with the users to let them log in and work on their data models.

Prerequisites

- You have installed OpenShift Data Science on your OpenShift Container Platform cluster.

- You have added at least one user to the user group for OpenShift Data Science.

Procedure

1. Log in to OpenShift Container Platform web console.

2. Click the application launcher ( ).

3. Right-click on **Red Hat OpenShift Data Science** and copy the URL for your OpenShift Data Science instance.

4. Provide this instance URL to your data scientists to let them log in to OpenShift Data Science.

Verification

- Confirm that you and your users can log in to OpenShift Data Science by using the instance URL.

Additional resources

- [Logging in to OpenShift Data Science](#)

- [Adding users](#)

CHAPTER 10. ENABLING GPU SUPPORT IN OPENSHIFT DATA SCIENCE

Optionally, to ensure that your data scientists can use compute-heavy workloads in their models, you can enable graphics processing units (GPUs) in OpenShift Data Science. To enable GPUs on OpenShift, you must install the NVIDIA GPU Operator. As a prerequisite to installing the NVIDIA GPU Operator, you must install the Node Feature Discovery Operator. For information about how to install these operators, see NVIDIA GPU Operator on Red Hat OpenShift Container Platform in the NVIDIA documentation.

IMPORTANT

Follow the instructions in this chapter only if you want to enable GPU support in an unrestricted self-managed environment. To enable GPU support in a disconnected self-managed environment, see Enabling GPU support in OpenShift Data Science instead.