



Red Hat OpenShift Data Science 1

管理用户和用户资源

学习如何在 Red Hat OpenShift Data Science 中管理用户权限和环境

Red Hat OpenShift Data Science 1 管理用户和用户资源

学习如何在 Red Hat OpenShift Data Science 中管理用户权限和环境

法律通告

Copyright © 2023 Red Hat, Inc.

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.

摘要

学习如何在 Red Hat OpenShift Data Science 中管理用户权限和环境

目录

第 1 章 使用数据收集	3
1.1. OPENSIFT DATA SCIENCE 的使用数据收集通告	3
1.2. 启用使用数据收集	3
1.3. 禁用使用数据收集	4
第 2 章 用户权限概述	5
第 3 章 用户类型	6
第 4 章 定义 OPENSIFT DATA SCIENCE ADMIN 和用户组	7
第 5 章 为 OPENSIFT DATA SCIENCE 添加用户	8
5.1. 将用户添加到专用的 OPENSIFT DATA SCIENCE 用户组	8
5.2. 其他资源	9
第 6 章 查看 OPENSIFT DATA SCIENCE 用户	10
第 7 章 删除用户和用户资源	11
7.1. 从 AMAZON EBS 备份存储数据	11
7.2. 从 GOOGLE PERSISTENT DISK 备份存储数据	12
7.3. 停止其他用户拥有笔记本服务器	13
7.4. 撤销对 JUPYTER 的用户访问权限	14
7.5. 删除用户后清除	15
第 8 章 为 OPENSIFT DATA SCIENCE 用户分配其他资源	17
第 9 章 在 OPENSIFT DATA SCIENCE 中启用 GPU 支持	18
第 10 章 为集群配置默认 PVC 大小	19
第 11 章 为集群恢复默认 PVC 大小	20
第 12 章 管理笔记本服务器	21
12.1. 访问 JUPYTER 管理界面	21
12.2. 启动由其他用户拥有的笔记本服务器	21
12.3. 访问其他用户拥有的笔记本服务器	22
12.4. 停止其他用户拥有笔记本服务器	23
12.5. 停止闲置的 NOTEBOOK	24
12.6. 配置自定义笔记本镜像	24
第 13 章 备份存储数据	27
13.1. 从 AMAZON EBS 备份存储数据	27
13.2. 从 GOOGLE PERSISTENT DISK 备份存储数据	28

第 1 章 使用数据收集

Red Hat OpenShift Data Science 管理员可以选择是否允许红帽收集集群中的 Red Hat OpenShift Data Science 使用数据。收集这些数据可让红帽监控并改进我们的软件和支持。有关红帽收集的数据的详细信息，请参阅 [OpenShift Data Science 数据收集通告](#)。

当在 OpenShift Dedicated 集群上安装 OpenShift Data Science 时，会默认启用数据收集。

有关在集群中禁用此数据收集的说明，请参阅[禁用使用数据收集](#)。如果您在集群中禁用了数据收集，并且希望再次启用它，请参阅[启用使用数据收集](#)。

1.1. OPENSIFT DATA SCIENCE 的使用数据收集通告

在您使用此红帽产品时，红帽可能会收集您所使用的软件的使用数据。通过这些数据，红帽可以监控软件并改进红帽产品和支持，包括识别、故障排除和响应影响用户的问题。

红帽收集哪些信息？

软件中的工具监控各种指标，这些信息将向红帽传输。指标包括如下信息：

- 有关产品仪表板中启用的应用程序的信息。
- 使用的部署大小（即分配的 CPU 和内存资源）。
- 有关从产品仪表板访问的文档资源的信息。
- 使用的笔记本镜像的名称（即 Minimal Python、Standard Data Science 和其他镜像）。
- 在初始用户登录过程中生成的随机识别符，用于将数据与特定用户名相关联。
- 有关组件、功能和扩展的使用情况信息。

第三方服务提供商

红帽使用某些第三方服务提供商收集遥测数据。

安全性

红帽采用技术和组织措施来保护使用数据。

个人数据

红帽不会收集个人信息。如果红帽发现个人信息被意外收集，红帽将根据红帽的隐私声明删除此类个人信息并处理此类个人信息。有关红帽隐私实践的更多信息，请参阅红帽[隐私声明](#)。

启用和禁用使用数据

您可以按照禁用使用数据收集或启用使用数据集中的说明 [禁用使用数据收集](#) 或 [启用使用数据收集](#)。

1.2. 启用使用数据收集

Red Hat OpenShift Data Science 管理员可以选择是否允许红帽收集集群中的 Red Hat OpenShift Data Science 使用数据。当在 OpenShift Dedicated 集群上安装 OpenShift Data Science 时，会默认启用数据收集。如果您之前禁用了数据收集，您可以按照下列步骤重新启用它。

先决条件

- 您已登陆到 Red Hat OpenShift Data Science。
- 您是 OpenShift Dedicated 中的 OpenShift Data Science 管理员组的一部分。

步骤

1. 从 OpenShift Data Science 仪表板中，点 **Settings** → **Cluster settings**。
2. 找到**使用数据收集**部分。
3. 选择 **Allow collection usage data** 复选框。
4. 点 **Save Changes**。

验证

- 更新设置时会显示通知：**设置更改已保存**。

其他资源

- [OpenShift Data Science 的使用数据收集通告](#)

1.3. 禁用使用数据收集

Red Hat OpenShift Data Science 管理员可以选择是否允许红帽收集集群中的 Red Hat OpenShift Data Science 使用数据。当在 OpenShift Dedicated 集群上安装 OpenShift Data Science 时，会默认启用数据收集。

您可以按照以下步骤禁用数据收集。

先决条件

- 您已登录到 Red Hat OpenShift Data Science。
- 您是 OpenShift Dedicated 中的 OpenShift Data Science 管理员组的一部分。

步骤

1. 从 OpenShift Data Science 仪表板中，点 **Settings** → **Cluster settings**。
2. 找到**使用数据收集**部分。
3. 取消选中 **Allow collection usage data** 复选框。
4. 点 **Save Changes**。

验证

- 更新设置时会显示通知：**设置更改已保存**。

其他资源

- [OpenShift Data Science 的使用数据收集通告](#)

第 2 章 用户权限概述

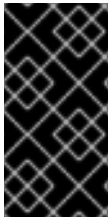
默认情况下，所有 OpenShift 用户都可以访问 Red Hat OpenShift 数据源。另外，OpenShift Dedicated 管理员组中的用户（默认为 **dedicated-admins**）自动具有 OpenShift Data Science 的管理员访问权限。

或者，您可以创建专用用户组来限制用户和管理员的 OpenShift Data Science 访问权限。因此，您必须决定是否要使用专用用户组限制对 OpenShift Data Science 部署的访问，而不是让所有 OpenShift 用户访问。

如果您决定限制访问，并且您已在配置的身份提供程序中定义用户组，您可以将这些用户组添加到您的 OpenShift Data Science 部署中。如果您决定在不从身份提供程序中添加这些组的情况下使用专用用户组，您必须在 OpenShift Data Science 中创建组，然后将适当的用户添加到其中。

在 OpenShift Dedicated、**cluster-admins** 和 **dedicated-admins** 中配置的用户组与专用的 OpenShift Data Science 用户组分开。有一些与 OpenShift Data Science 相关的操作，需要 **cluster-admins** 或 **dedicated-admins** 角色。这些操作包括：

- 如果您使用专用组，向 OpenShift Data Science 用户和组添加用户。
- 如果您使用专用组，从 OpenShift Data Science 用户和组中删除用户。
- 在 OpenShift Dedicated 中为用户管理自定义环境和存储配置，如 Jupyterbook 资源、ConfigMap 和持久性卷声明 (PVC)。



重要

虽然 OpenShift Data Science 用户及其组件通过 OpenShift 进行身份验证，但会话管理与身份验证分开。这意味着，注销 OpenShift Dedicated 或 OpenShift Data Science 不会影响在这些平台上运行的 Jupyter 会话。这意味着，当用户的权限改变时，此用户必须注销所有当前会话才能使更改生效。

第 3 章 用户类型

Red Hat OpenShift Data Science 有以下用户类型：

表 3.1. 用户类型

用户类型	权限
数据科学家	数据科学家可访问和使用 Red Hat OpenShift Data Science 的各个组件，如 Jupyter。
IT 操作管理员	除了数据科学家允许的操作外，IT 操作管理员还可以： <ul style="list-style-type: none">● 配置 Red Hat OpenShift Data Science 设置。● 访问和管理笔记本服务器。

其他资源

- [OpenShift Dedicated 集群管理](#)

第 4 章 定义 OPENSIFT DATA SCIENCE ADMIN 和用户组

默认情况下，具有 cluster admin 权限和 **dedicated-admins** 管理员组中的用户是 OpenShift Data Science 的管理员，但 OpenShift 中通过身份验证的所有用户均可访问 OpenShift Data Science。**集群管理员**是可以在 OpenShift 集群中任何项目中执行任何操作的超级用户。当使用本地绑定来绑定一个用户时，这些用户可以完全控制项目中每一资源的配额和所有操作。您还可以使用控制面板定义额外的 OpenShift Data Science admin 和用户组。

先决条件

- 您已登录到 Red Hat OpenShift Data Science，如 [Logging to OpenShift Data Science](#) 所述。
- 您是 OpenShift Dedicated 中的 OpenShift Data Science 管理员组的一部分。
- OpenShift Dedicated 中存在您要定义为 admin 和 user 组的组。

步骤

1. 在 OpenShift Data Science 仪表板中点 **Settings** → **User Management**。
2. 定义您的 OpenShift Data Science 管理组：在 **Data Science** 管理员组下，点文本框并选择 OpenShift 组。重复此过程来定义多个管理员组。
3. 定义 OpenShift Data Science 用户组：在 **Data Science** 用户组下，点文本框并选择 OpenShift 组。重复此过程以定义多个用户组。



重要

system:authenticated 设置允许 OpenShift 中通过身份验证的所有用户访问 OpenShift Data Science。

4. 点 **Save Changes**。

验证

- 管理员用户可以成功登录 OpenShift Data Science 并执行管理功能。
- 非管理员用户可以成功登录 OpenShift Data Science，并可访问和使用各个组件，如 Jupyter。

第 5 章 为 OPENSIFT DATA SCIENCE 添加用户

默认情况下，所有 OpenShift 用户都可以访问 Red Hat OpenShift 数据源。如果您使用这些默认权限设置，则不需要进一步操作。但是，如果您计划通过定义专用用户组来限制对 OpenShift 数据收集的访问，则必须通过向 Red Hat OpenShift Data Science 用户组、管理员组或两者添加用户帐户来授予用户访问 OpenShift Data Science 的权限。您可以使用默认组名称，或者指定身份供应商中已存在的组名称。

用户组提供对 Red Hat OpenShift Data Science 仪表板中的开发人员功能以及相关服务（如 Jupyter）的访问。

管理员组提供对 Red Hat OpenShift Data Science 仪表板和相关服务（如 Jupyter）中的开发人员和管理员功能的访问权限。

如果您限制对使用特殊用户组的访问，不在 OpenShift Data Science 用户或组的用户仍然可以查看仪表板，但无法使用关联的服务，如 Jupyter。它们还无法访问 **Cluster 设置** 页面。

要使用默认组名称，请参阅 [将用户添加到专用的 OpenShift Data Science 用户组](#)。这个方法易于设置，但您必须在 OpenShift Dedicated Web 控制台中手动配置用户列表。



重要

如果您使用 LDAP 作为身份提供程序，则需要配置到 OpenShift Dedicated 的 LDAP 同步。如需更多信息，请参阅[同步 LDAP 组](#)。

5.1. 将用户添加到专用的 OPENSIFT DATA SCIENCE 用户组

从 Red Hat OpenShift Data Science 1.8 开始，所有 OpenShift Dedicated 用户都默认能够访问 Red Hat OpenShift Data Science。另外，**dedicated-admins** 管理员组中的用户会自动具有 OpenShift Data Science 的管理员访问权限。1.8 之前的 OpenShift Data Science 版本包含专用的 OpenShift Data Science 管理员和用户组。为了进一步限制对版本 1.8 或更高版本中的 OpenShift Data Science 的访问，您可以继续创建专用的 OpenShift Data Science 管理员和用户组。

按照本节中的步骤将用户添加到专用的 OpenShift Data Science 管理员和用户组中。这个方法易于设置，但您必须在 OpenShift Dedicated web 控制台中手动管理用户列表。

先决条件

- 您已为 OpenShift Dedicated 配置了一个受支持的身份提供程序。
- 您是 OpenShift Dedicated 中的 **dedicated-admins** 用户组的一部分。
- 您已定义了 OpenShift Data Science 管理员组和用户组。

流程

1. 在 OpenShift Dedicated Web 控制台中，点 **User Management** → **Groups**。
2. 点您要添加用户到的组名称。
 - 对于管理用户，点 administrator 组，如 **rhods-admins**。
 - 对于普通用户，点用户组，例如：**rhods-users**。此时会出现该组的组详细信息页面。
3. 点 **Actions** → **Add Users**。

此时会出现 **Add Users** 对话框。

4. 在 **Users** 字段中，输入要添加到组中的相关用户名。
5. 点 **Save**。

验证

- 点每个组的 **Details** 选项卡，并确认 **Users** 部分包含您添加的用户名。

5.2. 其他资源

- [将用户添加到专用的 OpenShift Data Science 用户组](#)
- [访问 OpenShift Data Science 仪表盘](#)

第 6 章 查看 OPENSIFT DATA SCIENCE 用户

默认情况下，所有 OpenShift 用户都可以访问 Red Hat OpenShift 数据源。具有 **cluster-admin** 角色的用户在 OpenShift Data Science 中自动具有管理员访问权限。另外，OpenShift Dedicated 管理员组中的用户（默认为 **dedicated-admins**）自动具有 OpenShift Data Science 的管理员访问权限。但是，更改 **dedicated-admins** 组的名称会导致属于此组的用户不会保留其 OpenShift Data Science 管理员访问权限。但是，您可以创建专用用户组来限制用户和管理员的 OpenShift Data Science 访问权限。如果您定义了专用的 OpenShift Data Science 用户组，参照以下步骤，以便可以查看属于这些组的用户。

先决条件

- Red Hat OpenShift Data Science 用户组、管理员组或两组均存在。
- 您有 **cluster-admin** 角色，或者您是 **dedicated-admins** 管理员组的一部分。
- 您已为 OpenShift Dedicated 配置了一个受支持的身份提供程序。

步骤

1. 在 OpenShift Dedicated Web 控制台中，点 **User Management** → **Groups**。
2. 点包含您要查看的用户的组名称。
 - 对于管理用户，点您的管理员组的名称，如 **rhods-admins**。
 - 对于普通用户，点您的用户组的名称，如 **rhods-users**。

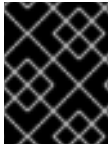
显示该组的组详细信息页面。

验证

- 在相关组的 **Users** 部分中，您可以查看有权访问 Red Hat OpenShift Data Science 的用户。

第 7 章 删除用户和用户资源

有权访问 OpenShift Dedicated 的用户，可以撤销用户对 Jupyter 的访问，并从 Red Hat OpenShift Data Science 中删除用户资源。



重要

要完全从 OpenShift Data Science 中删除用户，您必须将它们从 OpenShift 身份提供商的允许组中删除。

7.1. 从 AMAZON EBS 备份存储数据

红帽建议定期备份持久性卷声明 (PVC) 中的数据。在删除用户前备份您的数据，以及在卸载 OpenShift Data Science 之前备份您的数据非常重要，因为在卸载 OpenShift Data Science 时，所有 PVC 都会被删除。

先决条件

- 您有 OpenShift Cluster Manager 的凭证 (<https://console.redhat.com/openshift/>)。
- 具有 OpenShift Dedicated 集群的管理员访问权限。
- 您有部署 OpenShift Dedicated 集群的 Amazon Web Services (AWS) 帐户的凭证。

步骤

1. 确定您要备份的持久性卷 (PV) 的 ID。
 - a. 在 OpenShift Dedicated Web 控制台中，更改到 **Administrator** 视角。
 - b. 点 **Home** → **Projects**。
 - c. 点 **rhods-notebooks** 项目。
将打开项目的 **Details** 页面。
 - d. 点 **Inventory** 部分中的 **PersistentVolumeClaims**。
PersistentVolumeClaims 页面将打开。
 - e. 记录您要备份的持久性卷 (PV) 的 ID。



注意

您需要注意的持久性卷 (PV)，才能识别正确的 EBS 卷以便在 AWS 实例中备份。

2. 找到包含您要备份的 PV 的 EBS 卷。
请参阅 [Amazon Web Services 文档：创建 Amazon EBS 快照](#) 以了解更多信息。
 - a. 登录到 AWS (<https://aws.amazon.com>)，并确保您查看部署了 OpenShift Dedicated 集群的区域。
 - b. 点 **Services**。
 - c. 点 **Compute** → **EC2**。

- d. 在侧边导航中点 **Elastic Block Storage → Volumes**。
此时会打开 **Volumes** 页面。
 - e. 在搜索栏中，输入您之前记下的持久性卷 (PV) 的 ID。
卷页面会重新加载以显示搜索结果。
 - f. 点显示的卷，并验证任何 **kubernetes.io/create-for/pvc/namespace** 标签都包含值 **rhods-notebooks**，以及任何 **kubernetes.io/created-for/pvc/name** 标签是否与 EC2 卷所用于的持久性卷的名称匹配，例如，**jupyter-nb-user1-pvc**。
3. 备份包含持久性卷 (PV) 的 EBS 卷。
 - a. 右键点您要备份的卷，然后从列表中选择 **Create Snapshot**。
Create Snapshot 页面将打开。
 - b. 为卷输入 **Description**。
 - c. 点 **Create Snapshot**。
创建卷的快照。
 - d. 点 **Close**。

验证

- 您创建的快照在 AWS 中的 **Snapshots** 页面中可见。

其他资源

- [Amazon Web Services 文档：创建 Amazon EBS 快照](#)

7.2. 从 GOOGLE PERSISTENT DISK 备份存储数据

红帽建议定期备份持久性卷声明 (PVC) 中的数据。在删除用户前备份您的数据，以及在卸载 OpenShift Data Science 之前备份您的数据非常重要，因为在卸载 OpenShift Data Science 时，所有 PVC 都会被删除。

先决条件

- 您有 OpenShift Cluster Manager 的凭证 (<https://console.redhat.com/openshift/>)。
- 具有 OpenShift Dedicated 集群的管理员访问权限。
- 您有部署 OpenShift Dedicated 集群的 Google Cloud Platform (GCP) 帐户的凭证。

步骤

1. 确定您要备份的持久性卷 (PV) 的 ID。
 - a. 在 OpenShift Dedicated Web 控制台中，更改到 **Administrator** 视角。
 - b. 点 **Home → Projects**。
 - c. 点 **rhods-notebooks** 项目。
将打开项目的 **Details** 页面。
 - d. 点 **Inventory** 部分中的 **PersistentVolumeClaims**。

PersistentVolumeClaims 页面将打开。

- e. 记录您要备份的持久性卷 (PV) 的 ID。
需要持久性卷(PV) ID 来识别正确的持久性磁盘以便在 GCP 实例中备份。
2. 找到包含您要备份的 PV 的持久磁盘。
 - a. 登录到 Google Cloud 控制台<https://console.cloud.google.com>，并确保您查看 OpenShift Dedicated 集群部署到的区域。
 - b. 单击导航菜单(ALLOWED)，然后单击 **Compute Engine**。
 - c. 在侧边导航中，在 **Storage** 下点 **Disks**。
Disks 页面将打开。
 - d. 在 **Filter query** 框中，输入您之前记下的持久性卷(PV)的 ID。
Disks 页面会重新加载以显示搜索结果。
 - e. 点显示的磁盘，并验证任何 **kubernetes.io/create-for/pvc/namespace** 标签都包含值 **rhods-notebooks**，以及任何 **kubernetes.io/created-for/pvc/name** 标签是否与用于持久磁盘的持久性卷的名称匹配，例如 **jupyterhub-nb-user1-pvc**。
 3. 备份包含持久性卷(PV)的持久磁盘。
 - a. 从顶部导航中选择 **CREATE SNAPSHOT**。
Create a snapshot 页面将打开。
 - b. 输入快照的唯一 **Name**。
 - c. 在 **Source disk** 下，验证您要备份的持久磁盘是否显示。
 - d. 根据需要更改任何可选设置。
 - e. 单击 **CREATE**。
创建持久磁盘的快照。

验证

- 您创建的快照在 GCP 中的 **Snapshots** 页面中可见。

其他资源

- [Google Cloud 文档：创建和管理磁盘快照](#)

7.3. 停止其他用户拥有笔记本服务器

管理员可以停止由其他用户拥有的笔记本服务器，以降低集群上的资源消耗，或作为从集群删除用户及其资源的一部分。

先决条件

- 如果您使用专用的 OpenShift Data Science 组，则作为管理员组的一部分（如 **rhods-admins**）。如果您不使用专用组，则作为 OpenShift Dedicated 管理员组的一部分。如需更多信息，请参阅[为 OpenShift Dedicated 添加管理用户](#)。
- 您已启动 Jupyter 应用程序。请参阅[启动 Jupyter 并启动笔记本服务器](#)。

- 要停止的笔记本服务器正在运行。

流程

1. 在启动 Jupyter 时打开的页面中，点 **Administration** 选项卡。
2. 停止一个或多个服务器。
 - 如果要停止一个或多个特定服务器，请执行以下操作：
 - i. 在 **Users** 部分中，找到笔记本服务器所属的用户。
 - ii. 要停止笔记本服务器，请执行以下操作之一：
 - 单击相关用户旁边的操作菜单(IANA)，然后选择 **Stop server**。
 - 点相关用户旁的 **View server**，然后点 **Stop notebook server**。此时会出现 **Stop server** 对话框。
 - iii. 点 **Stop server**。
 - 如果要停止所有服务器，请执行以下操作：
 - i. 点 **Stop all servers** 按钮。
 - ii. 点 **OK** 以确认停止所有服务器。

验证

- 当笔记本服务器停止后，每个服务器都会将 **Stop server** 链接更改为 **Start server** 链接。

7.4. 撤销对 JUPYTER 的用户访问权限

您可以通过 Jupyter 撤销用户对 Jupyter 的访问权限，以防止他们在集群中运行笔记本服务器和消耗资源，同时仍允许他们访问使用 OpenShift 身份提供商进行身份验证的 OpenShift Data Science 和其他服务。



重要

只有在限制使用特殊用户组对 OpenShift Data Science 的访问时才应执行这些步骤。要完全从 OpenShift Data Science 中删除用户，您必须将它们从 OpenShift 身份提供商的允许组中删除。

先决条件

- 您已停止了您要删除的用户拥有的任何笔记本服务器。
- 您是 OpenShift Dedicated 中的 **dedicated-admins** 用户组的一部分。
- 如果您使用专用的 OpenShift Data Science 用户组，该用户是 OpenShift Data Science 用户组、管理员组或两者的一部分。

步骤

1. 在 OpenShift Dedicated Web 控制台中，点 **User Management** → **Groups**。

2. 点您要从中删除该用户的组名称。
 - 对于管理用户，点您的管理员组的名称，如 **rhods-admins**。
 - 对于普通用户，点您的用户组的名称，如 **rhods-users**。

显示该组的组详细信息页面。

3. 在 **Details** 选项卡中的 **Users** 部分，找到您要删除的用户。
4. 点您要删除的用户旁的操作菜单 (⋮) 并点 **Remove user**。

验证

- 检查 **Details** 选项卡中的 **Users** 部分，并确认您删除的用户是否不可见。
- 在 **rhods-notebooks** 项目中，检查 **Workload** → **Pods** 下，并确保没有针对此用户的笔记本服务器 pod。如果您可以看到名为 **jupyter-nb-<username>*** 的 pod，则代表您删除了这个 pod，请删除该 pod，以确保已删除用户不会消耗集群中的资源。
- 在数据科学项目中，检查数据科学项目列表。删除属于用户的任何项目。

7.5. 删除用户后清除

在删除了用户对 Red Hat OpenShift Data Science 或 Jupyter 的访问权限后，您还必须从 OpenShift Dedicated 中删除其相关的配置文件。建议您在删除其配置文件前备份用户数据。

先决条件

- (可选) 如果要完全删除用户对 OpenShift Data Science 的访问权限，已从身份提供程序中删除其凭证。
- 您已撤销了用户对 Jupyter 的访问权限。
- 您已从 Amazon EBS 或 Google Persistent Disk 备份用户的存储数据。
- 如果您使用专用的 OpenShift Data Science 组，则作为管理员组的一部分（如 **rhods-admins**）。如果您不使用专用组，则作为 OpenShift Dedicated 管理员组的一部分。如需更多信息，请参阅 [为 OpenShift Dedicated 添加管理用户](#)。
- 已登陆到 OpenShift Dedicated Web 控制台。
- 您已登陆到 OpenShift Data Science。

步骤

1. 删除用户的持久性卷声明 (PVC)。
 - a. 点 **Storage** → **PersistentVolumeClaims**。
 - b. 如果还没有选择，请从项目列表中选择 **rhods-notebooks** 项目。
 - c. 找到 **jupyter-nb-<username> PVC**。
将 **<username>** 替换为相关用户名。
 - d. 点操作菜单 (⋮)，从列表中选择 **Delete PersistentVolumeClaim**。

此时会出现 **Delete PersistentVolumeClaim** 对话框。

- e. 检查对话框，并确认您删除正确的 PVC。
 - f. 点 **Delete**。
2. 删除用户的 ConfigMap。
- a. 点 **Workloads** → **ConfigMaps**。
 - b. 如果还没有选择，请从项目列表中选择 **rhods-notebooks** 项目。
 - c. 找到 **jupyterhub-singleuser-profile-<username>** ConfigMap。
将 **<username>** 替换为相关用户名。
 - d. 点操作菜单 (⋮)，并从列表中选择 **Delete ConfigMap**。
此时会出现 **Delete ConfigMap** 对话框。
 - e. 检查对话框，并确认您删除正确的 ConfigMap。
 - f. 点 **Delete**。

验证

- 用户无法访问 Jupyter，并在尝试时看到 "Access permission needed" 信息。
- 用户的单用户配置集、持久性卷声明 (PVC) 和 ConfigMap 无法在 OpenShift Dedicated 中可见。

第 8 章 为 OPENSIFT DATA SCIENCE 用户分配其他资源

作为集群管理员，您可以将其他资源分配给集群以支持计算密集型数据科学工作。这包括增加集群中的节点数量并更改集群分配的机器池。

先决条件

- 您有用于在 OpenShift Cluster Manager 中管理集群的凭证 (<https://console.redhat.com/openshift/>)。有关在 OpenShift Cluster Manager 中配置管理访问权限的更多信息，请参阅在 [OpenShift Cluster Manager 中配置对集群的访问](#)。
- 如果您使用 NVIDIA GPU 增加机器池的大小，您的 OpenShift Dedicated 集群支持它们。
- 您有一个具有容量的 AWS 或 GCP 实例，以创建更大的容器大小。
- 对于计算密集型操作，您有足够容量的 AWS 或 GCP 实例，以适应最大容器大小 **XL**。

步骤

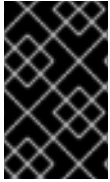
1. 登录到 OpenShift Cluster Manager (<https://console.redhat.com/openshift/>)。
2. 点 **Clusters**。
Clusters 页面将打开。
3. 点您要分配额外资源的集群名称。
4. 点 **Actions** → **Edit node count**。
5. 从列表中选择 **Machine 池**。
6. 从 **Node count** 列表中选择分配给机器池的节点数量。
7. 点 **应用**。

验证

- 您分配给集群的其他资源会显示在 **Machine Pools** 标签页中。

第 9 章 在 OPENSIFT DATA SCIENCE 中启用 GPU 支持

另外，为了确保数据科学家可以在其模型中使用计算密集型工作负载，您可以在 OpenShift Data Science 中启用图形处理单元(GPU)。要在 OpenShift 中启用 GPU，您必须安装 NVIDIA GPU Operator。作为安装 NVIDIA GPU Operator 的先决条件，您必须安装 Node Feature Discovery (NFD) Operator。有关如何安装这些 Operator 的详情，请参考 [Red Hat OpenShift Container Platform 上的 NVIDIA GPU Operator](#)。



重要

NVIDIA GPU 附加组件不再被支持。按照本章中的说明安装 NVIDIA GPU Operator。在安装 NVIDIA GPU Operator 之前，如果您的部署安装了 NVIDIA GPU 附加组件，请使用 OpenShift Cluster Manager 从集群中卸载 NVIDIA GPU 附加组件。

其他资源

- [NVIDIA 文档中的 Red Hat OpenShift Container Platform 上的 NVIDIA GPU Operator](#)。
- [Node Feature Discovery Operator](#)

第 10 章 为集群配置默认 PVC 大小

要配置如何在 OpenShift Data Science 集群中声明资源，您可以更改集群持久性卷声明 (PVC) 的默认大小，确保请求的存储与常见存储工作流匹配。PVC 是对集群中的资源请求，还可作为对资源的声明检查。

先决条件

- 您已登陆到 Red Hat OpenShift Data Science。
- 您是 OpenShift Dedicated 中的 OpenShift Data Science 管理员组的一部分。

更改此设置会重启 Jupyter pod，Jupyter pod 可能最多有 30 秒不可用。作为临时解决方案，建议您在组织的典型工作日外执行该操作。

步骤

1. 从 OpenShift Data Science 仪表板中，点 **Settings** → **Cluster settings**。
2. 在 **PVC 大小** 下，以 KB 为单位输入新大小。最小值为 1 GiB，最大大小为 16384 GiB。
3. 点 **Save Changes**。

验证

- 使用您配置的默认存储大小创建新的 PVC。

其他资源

- [了解持久性存储](#)

第 11 章 为集群恢复默认 PVC 大小

要更改 OpenShift Data Science 集群中使用的资源大小，您可以恢复集群持久性卷声明(PVC)的默认大小。

先决条件

- 您已登陆到 Red Hat OpenShift Data Science。
- 您是 OpenShift Dedicated 中的 OpenShift Data Science 管理员组的一部分。

步骤

1. 从 OpenShift Data Science 仪表板中，点 **Settings** → **Cluster settings**。
2. 点 **恢复默认** 恢复默认 PVC 大小为 20GiB。
3. 点 **Save Changes**。

验证

- 创建新的 PVC，其默认存储大小为 20 GiB。

其他资源

- [了解持久性存储](#)

第 12 章 管理笔记本服务器

12.1. 访问 JUPYTER 管理界面

您可以使用 Jupyter 管理界面控制 Red Hat OpenShift Data Science 环境中的笔记本服务器。

先决条件

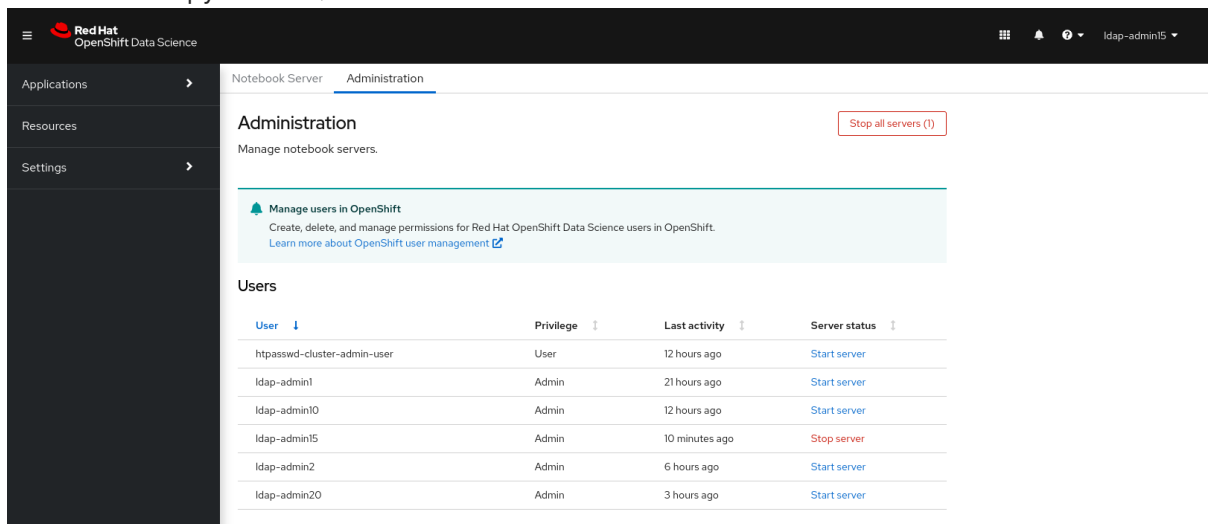
- 您是 OpenShift Dedicated 管理员组的一部分。如需更多信息，请参阅 [OpenShift Dedicated 添加管理用户](#)。

步骤

- 要从 OpenShift Data Science 访问 Jupyter 管理界面，请执行以下操作：
 - i. 在 OpenShift Data Science 中，在左侧菜单的 **Applications** 部分中，单击 **Enabled**。
 - ii. 找到 Jupyter 卡，点 **Launch application**。
 - iii. 在启动 Jupyter 时打开的页面中，点 **Administration** 选项卡。此时会打开 **Administration** 页面。
- 要从 JupyterLab 访问 Jupyter 管理界面，请执行以下操作：
 - i. 点 **File → Hub Control Panel**。
 - ii. 在 OpenShift Data Science 中打开的页面中，单击 **Administration** 选项卡。此时会打开 **Administration** 页面。

验证

- 您可以看到 Jupyter 管理界面。



12.2. 启动由其他用户拥有的笔记本服务器

管理员可以从 Jupyter 管理界面为另一现有用户启动笔记本服务器。

先决条件

- 您是 OpenShift Dedicated 管理员组的一部分。如需更多信息，请参阅 [OpenShift Dedicated 添加管理用户](#)。
- 您已启动 Jupyter 应用程序。请参阅 [启动 Jupyter 并启动笔记本服务器](#)。

步骤

1. 在启动 Jupyter 时打开的页面中，点 **Administration** 选项卡。
2. 在 **Administration** 选项卡中，执行以下操作：
 - a. 在 **Users** 部分中，找到您要启动的笔记本服务器的用户。
 - b. 点相关用户旁的 **Start server**。
 - c. 完成 **启动笔记本服务器** 页面。
 - d. 可选：如果需要，选择 **Start server in current tab**
 - e. 单击 **Start server**。
服务器启动后，您会看到以下行为之一：
 - 如果您之前选择了 **Start server in current tab** 则 JupyterLab 界面会在 Web 浏览器的当前标签页中打开。
 - 如果您之前没有选中 **Start server in current tab** 则启动服务器对话框会提示您在新浏览器标签页或当前标签页中打开服务器。
JupyterLab 接口根据您的选择打开。

验证

- 此时会打开 JupyterLab 接口。

其他资源

- [启动 Jupyter 并启动笔记本服务器](#)
- [笔记本服务器环境的选项](#)

12.3. 访问其他用户拥有的笔记本服务器

管理员可以访问由其他用户拥有的笔记本服务器，以更正配置错误或帮助数据科学家对其环境进行故障排除。

先决条件

- 您是 OpenShift Dedicated 管理员组的一部分。如需更多信息，请参阅 [OpenShift Dedicated 添加管理用户](#)。
- 您已启动 Jupyter 应用程序。请参阅 [启动 Jupyter 并启动笔记本服务器](#)。
- 您要访问的笔记本服务器正在运行。

步骤

1. 在启动 Jupyter 时打开的页面中，点 Administration 选项卡。
2. 在 Administration 页面中，执行以下操作：
 - a. 在 Users 部分中，找到笔记本服务器所属的用户。
 - b. 点相关用户旁的 View server。
 - c. 在笔记本服务器控制面板页面中，单击 Access notebook server。

验证

- 用户的笔记本服务器在 JupyterLab 中打开。

12.4. 停止其他用户拥有笔记本服务器

管理员可以停止由其他用户拥有的笔记本服务器，以降低集群上的资源消耗，或作为从集群删除用户及其资源的一部分。

先决条件

- 如果您使用专用的 OpenShift Data Science 组，则作为管理员组的一部分（如 rhods-admins）。如果您不使用专用组，则作为 OpenShift Dedicated 管理员组的一部分。如需更多信息，请参阅 [OpenShift Dedicated 添加管理用户](#)。
- 您已启动 Jupyter 应用程序。请参阅 [启动 Jupyter 并启动笔记本服务器](#)。
- 要停止的笔记本服务器正在运行。

步骤

1. 在启动 Jupyter 时打开的页面中，点 Administration 选项卡。
2. 停止一个或多个服务器。
 - 如果要停止一个或多个特定服务器，请执行以下操作：
 - i. 在 Users 部分中，找到笔记本服务器所属的用户。
 - ii. 要停止笔记本服务器，请执行以下操作之一：
 - 单击相关用户旁边的操作菜单(IANA)，然后选择 Stop server。
 - 点相关用户旁的 View server，然后点 Stop notebook server。此时会出现 Stop server 对话框。
 - iii. 点 Stop server。
 - 如果要停止所有服务器，请执行以下操作：
 - i. 点 Stop all servers 按钮。
 - ii. 点 OK 以确认停止所有服务器。

验证

- 当笔记本服务器停止后，每个服务器都会将 Stop server 链接更改为 Start server 链接。

12.5. 停止闲置的 NOTEBOOK

您可以通过停止闲置（无需登录用户）的 Notebook 服务器来减少 OpenShift Data Science 部署中的资源使用量。当集群中资源需求很高时，这非常有用。默认情况下，闲置 Notebook 不会在特定时间限制后停止。



注意

如果您将集群设置配置为在指定时间限制后从集群断开所有用户，那么这个设置优先于闲置 Notebook 时间限制。当用户的会话持续时间达到集群范围时间限制时，会在集群注销。

先决条件

- 您已登录到 Red Hat OpenShift Data Science。
- 您是 OpenShift Dedicated 中的 OpenShift Data Science 管理员组的一部分。

步骤

1. 从 OpenShift Data Science 仪表板中，点 Settings → Cluster settings。
2. 在 Stop idle notebooks 下，选择 Stop idle notebooks after。
3. 以小时和分钟为单位输入时间限制，用于指定空闲 notebooks 被停止的时间。
4. 点 Save Changes。

验证

- notebook-controller-culler-config ConfigMap 位于 Workloads → ConfigMaps 页面中的 redhat-ods-applications 项目中，包含以下 culling 配置设置：
 - ENABLE_CULLING: 指定是否启用或禁用 culling 功能（默认为 false）。
 - IDLENESS_CHECK_PERIOD : 轮询频率以检查手册的最后已知活动（以分钟为单位）。
 - CULL_IDLE_TIME: 将不活跃笔记本扩展为零（以分钟为单位）的最大分配时间。
- 闲置笔记本会在您设置的时间限制停止。

12.6. 配置自定义笔记本镜像





除了由红帽和独立软件供应商 (ISV) 提供和支持的笔记本镜像外，您还可以配置针对项目特定要求的自定义笔记本镜像。

红帽支持您在部署 OpenShift Data Science 中添加自定义笔记本镜像，并确保它们可用于在创建笔记本服务器时进行选择。但是，红帽不支持您的自定义笔记本镜像的内容。也就是说，如果您的自定义笔记本镜像可在笔记本服务器创建过程中选择，但没有创建可用的笔记本服务器，红帽不提供支持修复您的自定义笔记本镜像。

先决条件

- 您已登陆到 Red Hat OpenShift Data Science。
- 您是 OpenShift Dedicated 中的 `dedicated-admins` 用户组的一部分。
- 您的自定义笔记本镜像存在于镜像 registry 中，可访问。

步骤

1. 从 OpenShift Data Science 仪表板中，点 `Settings` → `Notebook 镜像`。
此时会出现 `Notebook image settings` 页面。以前导入的 `notebook` 镜像会被显示。要在包含相关 `notebook` 镜像的行中启用或禁用之前导入的 `notebook` 镜像，请点 `Enabled` 列中的切换。
2. 点 `Import new image`。或者，如果未找到之前导入的镜像，请点 `Import image`。
此时会出现 `Import Notebook 镜像` 对话框。
3. 在 `Repository` 字段中，输入包含 `notebook` 镜像的存储库的 URL。
4. 在 `Name` 字段中，为 `notebook` 输入相应的名称。
5. 在 `Description` 字段中，为 `notebook` 电脑镜像输入适当的描述。
6. 可选：将软件添加到 `notebook` 镜像。导入完成后，软件将添加到 `notebook` 镜像的 `meta-data` 中，并显示在 `Jupyter` 服务器创建页面中。
 - a. 点 `Software` 选项卡。
 - b. 点 `添加软件按钮`。
 - c. 点 `Edit` ()。
 - d. 输入软件名称。
 - e. 输入软件版本。
 - f. 点 `Confirm` () 确认您的条目。
 - g. 要添加其他软件，请点 `Add software`，填写相关字段并确认您的条目。
7. 可选：将软件包添加到 `notebook` 镜像。导入完成后，软件包将添加到 `notebook` 镜像的 `meta-data` 中，并显示在 `Jupyter` 服务器创建页面中。
 - a. 点 `Packages` 选项卡。
 - b. 点 `Add package 按钮`。
 - c. 点 `Edit` ()。
 - d. 输入软件包名称。
 - e. 输入软件包版本。
 - f. 点 `Confirm` () 确认您的条目。
 - g. 要添加附加软件包，请点 `Add package`，完成相关字段并确认您的条目。

8. 点 Import。

验证

- 您导入的 notebook 镜像会在 Notebook 镜像设置页面的表中显示。
- 在 Jupyter 的 启动一个笔记本服务器 页面中可选择您的自定义笔记本镜像。

其他资源

- [管理镜像流](#)
- [了解构建配置](#)

第 13 章 备份存储数据

13.1. 从 AMAZON EBS 备份存储数据

红帽建议定期备份持久性卷声明 (PVC) 中的数据。在删除用户前备份您的数据，以及在卸载 OpenShift Data Science 之前备份您的数据非常重要，因为在卸载 OpenShift Data Science 时，所有 PVC 都会被删除。

先决条件

- 您有 OpenShift Cluster Manager 的凭证 (<https://console.redhat.com/openshift/>)。
- 具有 OpenShift Dedicated 集群的管理员访问权限。
- 您有部署 OpenShift Dedicated 集群的 Amazon Web Services (AWS) 帐户的凭证。

步骤

1. 确定您要备份的持久性卷 (PV) 的 ID。
 - a. 在 OpenShift Dedicated Web 控制台中，更改到 Administrator 视角。
 - b. 点 Home → Projects。
 - c. 点 rhods-notebooks 项目。
将打开项目的 Details 页面。
 - d. 点 Inventory 部分中的 PersistentVolumeClaims。
PersistentVolumeClaims 页面将打开。
 - e. 记录您要备份的持久性卷 (PV) 的 ID。



注意

您需要注意的持久性卷 (PV)，才能识别正确的 EBS 卷以便在 AWS 实例中备份。

2. 找到包含您要备份的 PV 的 EBS 卷。
请参阅 [Amazon Web Services 文档：创建 Amazon EBS 快照](#) 以了解更多信息。
 - a. 登录到 AWS (<https://aws.amazon.com>)，并确保您查看部署了 OpenShift Dedicated 集群的区域。
 - b. 点 Services。
 - c. 点 Compute → EC2。
 - d. 在侧边导航中点 Elastic Block Storage → Volumes。
此时会打开 Volumes 页面。
 - e. 在搜索栏中，输入您之前记下的持久性卷 (PV) 的 ID。
卷页面会重新加载以显示搜索结果。

- f. 点显示的卷，并验证任何 `kubernetes.io/create-for/pvc/namespace` 标签都包含值 `rhods-notebooks`，以及任何 `kubernetes.io/created-for/pvc/name` 标签是否与 EC2 卷所用于的持久性卷的名称匹配，例如，`jupyter-nb-user1-pvc`。
3. 备份包含持久性卷 (PV) 的 EBS 卷。
 - a. 右键点您要备份的卷，然后从列表中选择 `Create Snapshot`。`Create Snapshot` 页面将打开。
 - b. 为卷输入 `Description`。
 - c. 点 `Create Snapshot`，创建卷的快照。
 - d. 点 `Close`。

验证

- 您创建的快照在 AWS 中的 `Snapshots` 页面中可见。

其他资源

- [Amazon Web Services 文档：创建 Amazon EBS 快照](#)

13.2. 从 GOOGLE PERSISTENT DISK 备份存储数据

红帽建议定期备份持久性卷声明 (PVC) 中的数据。在删除用户前备份您的数据，以及在卸载 OpenShift Data Science 之前备份您的数据非常重要，因为在卸载 OpenShift Data Science 时，所有 PVC 都会被删除。

先决条件

- 您有 OpenShift Cluster Manager 的凭证 (<https://console.redhat.com/openshift/>)。
- 具有 OpenShift Dedicated 集群的管理员访问权限。
- 您有部署 OpenShift Dedicated 集群的 Google Cloud Platform (GCP) 帐户的凭证。

步骤

1. 确定您要备份的持久性卷 (PV) 的 ID。
 - a. 在 OpenShift Dedicated Web 控制台中，更改到 `Administrator` 视角。
 - b. 点 `Home` → `Projects`。
 - c. 点 `rhods-notebooks` 项目。
将打开项目的 `Details` 页面。
 - d. 点 `Inventory` 部分中的 `PersistentVolumeClaims`。
`PersistentVolumeClaims` 页面将打开。
 - e. 记录您要备份的持久性卷 (PV) 的 ID。
需要持久性卷 (PV) ID 来识别正确的持久性磁盘以便在 GCP 实例中备份。
2. 找到包含您要备份的 PV 的持久磁盘。

- a. 登录到 Google Cloud 控制台<https://console.cloud.google.com>，并确保您查看 OpenShift Dedicated 集群部署到的区域。
 - b. 单击导航菜单(ALLOWED)，然后单击 Compute Engine。
 - c. 在侧边导航中，在 Storage 下点 Disks。
Disks 页面将打开。
 - d. 在 Filter query 框中，输入您之前记下的持久性卷(PV)的 ID。
Disks 页面会重新加载以显示搜索结果。
 - e. 点显示的磁盘，并验证任何 `kubernetes.io/create-for/pvc/namespace` 标签都包含值 `rhods-notebooks`，以及任何 `kubernetes.io/created-for/pvc/name` 标签是否与用于持久磁盘的持久性卷的名称匹配，例如 `jupyterhub-nb-user1-pvc`。
3. 备份包含持久性卷(PV)的持久磁盘。
 - a. 从顶部导航中选择 CREATE SNAPSHOT。
Create a snapshot 页面将打开。
 - b. 输入快照的唯一 Name。
 - c. 在 Source disk 下，验证您要备份的持久磁盘是否显示。
 - d. 根据需要更改任何可选设置。
 - e. 单击 CREATE。
创建持久磁盘的快照。

验证

- 您创建的快照在 GCP 中的 Snapshots 页面中可见。

其他资源

- [Google Cloud 文档：创建和管理磁盘快照](#)