



# Red Hat OpenShift Service on AWS 4

## Tutorials

Red Hat OpenShift Service on AWS tutorials



# Red Hat OpenShift Service on AWS 4 Tutorials

---

Red Hat OpenShift Service on AWS tutorials

## Legal Notice

Copyright © 2024 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<sup>®</sup> is the registered trademark of Linus Torvalds in the United States and other countries.

Java<sup>®</sup> is a registered trademark of Oracle and/or its affiliates.

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

MySQL<sup>®</sup> is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js<sup>®</sup> 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<sup>®</sup> 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.

## Abstract

Tutorials on creating your first Red Hat OpenShift Service on AWS (ROSA) cluster.

## Table of Contents

|  |           |
|--|-----------|
| <b>CHAPTER 1. TUTORIALS OVERVIEW</b> .....   | <b>8</b>  |
| <b>CHAPTER 2. TUTORIAL: ROSA WITH HCP ACTIVATION AND ACCOUNT LINKING</b> .....   | <b>9</b>  |
| 2.1. PREREQUISITES   | 9         |
| 2.2. SUBSCRIPTION ENABLEMENT AND AWS ACCOUNT SETUP   | 9         |
| 2.3. AWS AND RED HAT ACCOUNT AND SUBSCRIPTION LINKING  | 12        |
| 2.4. ROSA WITH HCP CLUSTER DEPLOYMENT USING THE CLI  | 17        |
| 2.5. ROSA WITH HCP CLUSTER DEPLOYMENT USING THE WEB CONSOLE  | 19        |
| <b>CHAPTER 3. TUTORIAL: VERIFYING PERMISSIONS FOR A ROSA STS DEPLOYMENT</b> .....  | <b>22</b> |
| 3.1. PREREQUISITES   | 22        |
| 3.2. VERIFYING ROSA PERMISSIONS  | 22        |
| 3.3. USAGE INSTRUCTIONS  | 22        |
| <b>CHAPTER 4. TUTORIAL: USING AWS WAF AND AMAZON CLOUDFRONT TO PROTECT ROSA WORKLOADS</b>                                  | <b>24</b> |
| 4.1. PREREQUISITES   | 24        |
| 4.1.1. Environment setup   | 24        |
| 4.2. SETTING UP THE SECONDARY INGRESS CONTROLLER   | 24        |
| 4.2.1. Configure the AWS WAF   | 26        |
| 4.3. CONFIGURE AMAZON CLOUDFRONT   | 27        |
| 4.4. DEPLOY A SAMPLE APPLICATION   | 29        |
| 4.5. TEST THE WAF  | 29        |
| 4.6. ADDITIONAL RESOURCES  | 30        |
| <b>CHAPTER 5. TUTORIAL: USING AWS WAF AND AWS ALBS TO PROTECT ROSA WORKLOADS</b> .....                                     | <b>31</b> |
| 5.1. PREREQUISITES   | 31        |
| 5.1.1. Environment setup   | 31        |
| 5.1.2. AWS VPC and subnets   | 31        |
| 5.2. DEPLOY THE AWS LOAD BALANCER OPERATOR   | 32        |
| 5.3. DEPLOY A SAMPLE APPLICATION   | 35        |
| 5.3.1. Configure the AWS WAF   | 36        |
| 5.4. ADDITIONAL RESOURCES  | 38        |
| <b>CHAPTER 6. TUTORIAL: DEPLOYING OPENSIFT API FOR DATA PROTECTION ON A ROSA CLUSTER</b> ..                                | <b>39</b> |
| 6.1. PREPARE AWS ACCOUNT   | 39        |
| 6.2. DEPLOY OADP ON THE CLUSTER  | 41        |
| 6.3. PERFORM A BACKUP  | 45        |
| 6.4. CLEANUP   | 47        |
| <b>CHAPTER 7. TUTORIAL: AWS LOAD BALANCER OPERATOR ON ROSA</b> .....   | <b>49</b> |
| 7.1. PREREQUISITES   | 49        |
| 7.1.1. Environment   | 49        |
| 7.1.2. AWS VPC and subnets   | 50        |
| 7.2. INSTALLATION  | 50        |
| 7.3. VALIDATING THE DEPLOYMENT   | 53        |
| 7.4. CLEANING UP   | 55        |
| <b>CHAPTER 8. TUTORIAL: CONFIGURING MICROSOFT ENTRA ID (FORMERLY AZURE ACTIVE DIRECTORY) AS AN IDENTITY PROVIDER</b> ..... | <b>56</b> |
| 8.1. PREREQUISITES   | 56        |
| 8.2. REGISTERING A NEW APPLICATION IN ENTRA ID FOR AUTHENTICATION  | 56        |
| 8.3. CONFIGURING THE APPLICATION REGISTRATION IN ENTRA ID TO INCLUDE OPTIONAL AND GROUP CLAIMS                             | 60        |

|   |           |
|---|-----------|
| Configuring optional claims   | 61        |
| Configuring group claims (optional)   | 63        |
| 8.4. CONFIGURING THE RED HAT OPENSIFT SERVICE ON AWS CLUSTER TO USE ENTRA ID AS THE IDENTITY PROVIDER | 64        |
| 8.5. GRANTING ADDITIONAL PERMISSIONS TO INDIVIDUAL USERS AND GROUPS                                   | 65        |
| Granting additional permissions to individual users   | 65        |
| Granting additional permissions to individual groups  | 66        |
| 8.6. ADDITIONAL RESOURCES   | 66        |
| <b>CHAPTER 9. TUTORIAL: USING AWS SECRETS MANAGER CSI ON ROSA WITH STS</b>                            | <b>67</b> |
| 9.1. PREREQUISITES  | 67        |
| Additional environment requirements   | 67        |
| 9.2. DEPLOYING THE AWS SECRETS AND CONFIGURATION PROVIDER   | 68        |
| 9.3. CREATING A SECRET AND IAM ACCESS POLICIES  | 68        |
| 9.4. CREATE AN APPLICATION TO USE THIS SECRET   | 70        |
| 9.5. CLEAN UP   | 71        |
| <b>CHAPTER 10. TUTORIAL: USING AWS CONTROLLERS FOR KUBERNETES ON ROSA</b>                             | <b>72</b> |
| 10.1. PREREQUISITES   | 72        |
| 10.2. SETTING UP YOUR ENVIRONMENT   | 72        |
| 10.3. PREPARING YOUR AWS ACCOUNT  | 72        |
| 10.4. INSTALLING THE ACK S3 CONTROLLER  | 73        |
| 10.5. VALIDATING THE DEPLOYMENT   | 75        |
| 10.6. CLEANING UP   | 75        |
| <b>CHAPTER 11. TUTORIAL: DEPLOYING THE EXTERNAL DNS OPERATOR ON ROSA</b>                              | <b>76</b> |
| 11.1. PREREQUISITES   | 76        |
| 11.2. SETTING UP YOUR ENVIRONMENT   | 76        |
| 11.3. SECONDARY INGRESS CONTROLLER SETUP  | 77        |
| 11.4. PREPARING YOUR AWS ACCOUNT  | 78        |
| 11.5. INSTALLING THE EXTERNAL DNS OPERATOR  | 80        |
| 11.6. DEPLOYING A SAMPLE APPLICATION  | 81        |
| <b>CHAPTER 12. TUTORIAL: DYNAMICALLY ISSUING CERTIFICATES USING THE CERT-MANAGER OPERATOR ON ROSA</b> | <b>83</b> |
| 12.1. PREREQUISITES   | 83        |
| 12.2. SETTING UP YOUR ENVIRONMENT   | 83        |
| 12.3. PREPARING YOUR AWS ACCOUNT  | 84        |
| 12.4. INSTALLING THE CERT-MANAGER OPERATOR  | 85        |
| 12.5. CREATING A CUSTOM DOMAIN INGRESS CONTROLLER   | 87        |
| 12.6. CONFIGURING DYNAMIC CERTIFICATES FOR CUSTOM DOMAIN ROUTES                                       | 90        |
| 12.7. DEPLOYING A SAMPLE APPLICATION  | 91        |
| 12.8. TROUBLESHOOTING DYNAMIC CERTIFICATE PROVISIONING  | 92        |
| <b>CHAPTER 13. TUTORIAL: ASSIGNING A CONSISTENT EGRESS IP FOR EXTERNAL TRAFFIC</b>                    | <b>93</b> |
| 13.1. SETTING YOUR ENVIRONMENT VARIABLES  | 93        |
| 13.2. ENSURING CAPACITY   | 93        |
| 13.3. CREATING THE EGRESS IP RULES  | 94        |
| 13.4. ASSIGNING AN EGRESS IP TO A NAMESPACE   | 94        |
| 13.5. ASSIGNING AN EGRESS IP TO A POD   | 95        |
| 13.5.1. Labeling the nodes  | 95        |
| 13.5.2. Reviewing the egress IPs  | 96        |
| 13.6. VERIFICATION  | 96        |
| 13.6.1. Deploying a sample application  | 96        |

---

|   |            |
|---|------------|
| 13.6.2. Testing the namespace egress  | 97         |
| 13.6.3. Testing the pod egress  | 98         |
| 13.6.4. Optional: Testing blocked egress  | 99         |
| 13.7. CLEANING UP YOUR CLUSTER  | 99         |
| <b>CHAPTER 14. TUTORIAL: UPDATING COMPONENT ROUTES WITH CUSTOM DOMAINS AND TLS CERTIFICATES</b> | <b>101</b> |
| 14.1. PREREQUISITES   | 101        |
| 14.2. SETTING UP YOUR ENVIRONMENT   | 101        |
| 14.3. FIND THE CURRENT ROUTES   | 101        |
| 14.4. CREATE A VALID TLS CERTIFICATE FOR EACH COMPONENT ROUTE                                   | 103        |
| 14.5. ADD THE CERTIFICATES TO THE CLUSTER AS SECRETS  | 103        |
| 14.6. FIND THE HOSTNAME OF THE LOAD BALANCER IN YOUR CLUSTER                                    | 104        |
| 14.7. ADD COMPONENT ROUTE DNS RECORDS TO YOUR HOSTING PROVIDER                                  | 104        |
| 14.8. UPDATE THE COMPONENT ROUTES AND TLS SECRET USING THE ROSA CLI                             | 104        |
| 14.9. RESET THE COMPONENT ROUTES TO THE DEFAULT USING THE ROSA CLI                              | 105        |
| <b>CHAPTER 15. GETTING STARTED WITH ROSA</b>  | <b>107</b> |
| 15.1. TUTORIAL: WHAT IS ROSA  | 107        |
| 15.1.1. Key features of ROSA  | 107        |
| 15.1.2. ROSA and Kubernetes   | 107        |
| 15.1.3. Basic responsibilities  | 108        |
| 15.1.4. Roadmap and feature requests  | 108        |
| 15.1.5. AWS region availability   | 108        |
| 15.1.6. Compliance certifications   | 108        |
| 15.1.7. Nodes   | 108        |
| 15.1.7.1. Worker nodes across multiple AWS regions  | 108        |
| 15.1.7.2. Minimum number of worker nodes  | 108        |
| 15.1.7.3. Underlying node operating system  | 108        |
| 15.1.7.4. Node hibernation or shut-down   | 108        |
| 15.1.7.5. Supported instances for worker nodes  | 108        |
| 15.1.7.6. Node autoscaling  | 109        |
| 15.1.7.7. Maximum number of worker nodes  | 109        |
| 15.1.8. Administrators  | 109        |
| 15.1.9. OpenShift versions and upgrades   | 109        |
| 15.1.10. Support  | 109        |
| 15.1.10.1. Limited support  | 109        |
| 15.1.11. Service-level agreement (SLA)  | 109        |
| 15.1.12. Notifications and communication  | 109        |
| 15.1.13. Open Service Broker for AWS (OBSA)   | 110        |
| 15.1.14. Offboarding  | 110        |
| 15.1.15. Authentication   | 110        |
| 15.1.16. SRE cluster access   | 110        |
| 15.1.17. Encryption   | 110        |
| 15.1.17.1. Encryption keys  | 110        |
| 15.1.17.2. KMS keys   | 110        |
| 15.1.17.3. Data encryption  | 110        |
| 15.1.17.4. etcd encryption  | 110        |
| 15.1.17.5. etcd encryption configuration  | 110        |
| 15.1.17.6. Multi-region KMS keys for EBS encryption   | 111        |
| 15.1.18. Infrastructure   | 111        |
| 15.1.19. Credential methods   | 111        |
| 15.1.20. Prerequisite permission or failure errors  | 111        |

---

|  |     |
|--|-----|
| 15.1.21. Storage   | 111 |
| 15.1.22. Using a VPC   | 111 |
| 15.1.23. Network plugin  | 111 |
| 15.1.24. Cross-namespace networking                                  | 111 |
| 15.1.25. Using Prometheus and Grafana                                | 112 |
| 15.1.26. Audit logs output from the cluster control-plane            | 112 |
| 15.1.27. AWS Permissions Boundary                                    | 112 |
| 15.1.28. AMI   | 112 |
| 15.1.29. Cluster backups   | 112 |
| 15.1.30. Custom domain   | 112 |
| 15.1.31. ROSA domain certificates                                    | 112 |
| 15.1.32. Disconnected environments                                   | 112 |
| 15.2. TUTORIAL: ROSA WITH AWS STS EXPLAINED                          | 113 |
| 15.2.1. Different credential methods to deploy ROSA                  | 113 |
| 15.2.1.1. Rosa with IAM Users  | 114 |
| 15.2.1.2. ROSA with STS  | 114 |
| 15.2.2. ROSA with STS security                                       | 114 |
| 15.2.3. AWS STS explained  | 114 |
| 15.2.4. Components specific to ROSA with STS                         | 114 |
| 15.2.5. Deploying a ROSA STS cluster                                 | 116 |
| 15.2.6. ROSA with STS workflow                                       | 116 |
| 15.2.7. ROSA with STS use cases                                      | 119 |
| 15.3. DEPLOYING A CLUSTER  | 119 |
| 15.3.1. Tutorial: Choosing a deployment method                       | 119 |
| 15.3.1.1. Deployment options   | 119 |
| 15.3.2. Tutorial: Simple CLI guide                                   | 120 |
| 15.3.2.1. Prerequisites  | 120 |
| 15.3.2.2. Creating account roles                                     | 120 |
| 15.3.2.3. Deploying the cluster                                      | 120 |
| 15.3.3. Tutorial: Detailed CLI guide                                 | 120 |
| 15.3.3.1. CLI deployment modes                                       | 120 |
| 15.3.3.2. Deployment workflow  | 121 |
| 15.3.3.3. Automatic mode   | 121 |
| 15.3.3.3.1. Creating account roles                                   | 121 |
| 15.3.3.3.2. Creating a cluster                                       | 122 |
| 15.3.3.3.2.1. Default configuration                                  | 123 |
| 15.3.3.3.3. Checking the installation status                         | 124 |
| 15.3.3.4. Manual Mode  | 124 |
| 15.3.3.4.1. Creating account roles                                   | 124 |
| 15.3.3.4.2. Creating a cluster                                       | 125 |
| 15.3.3.4.3. Creating Operator roles                                  | 127 |
| 15.3.3.4.4. Creating the OIDC provider                               | 128 |
| 15.3.3.4.5. Checking the installation status                         | 128 |
| 15.3.3.5. Obtaining the Red Hat Hybrid Cloud Console URL             | 128 |
| 15.3.4. Tutorial: Simple UI guide                                    | 129 |
| 15.3.4.1. Prerequisites  | 129 |
| 15.3.4.2. Creating account roles                                     | 129 |
| 15.3.4.3. Creating Red Hat OpenShift Cluster Manager roles           | 129 |
| 15.3.5. Tutorial: Detailed UI guide                                  | 130 |
| 15.3.5.1. Deployment workflow  | 130 |
| 15.3.5.2. Creating account wide roles                                | 130 |
| 15.3.5.3. Associating your AWS account with your Red Hat account     | 131 |
| 15.3.5.4. Creating and associating an OpenShift Cluster Manager role | 133 |



|   |            |
|---|------------|
| 15.3.5.4.1. Other OpenShift Cluster Manager role creation options       | 134        |
| 15.3.5.5. Creating an OpenShift Cluster Manager user role               | 134        |
| 15.3.5.6. Creating account roles  | 135        |
| 15.3.5.7. Confirming successful account association                     | 135        |
| 15.3.5.8. Creating the cluster  | 136        |
| 15.3.5.8.1. Networking  | 137        |
| 15.3.5.8.2. Cluster roles and policies                                  | 137        |
| 15.3.5.8.3. Cluster updates   | 137        |
| 15.3.5.8.4. Reviewing and creating your cluster                         | 137        |
| 15.3.5.8.5. Monitoring the installation progress                        | 137        |
| 15.3.5.9. Basic OpenShift Cluster Manager Role                          | 138        |
| 15.3.5.9.1. Creating Operator roles                                     | 138        |
| 15.3.5.9.2. Creating the OIDC provider                                  | 139        |
| 15.4. TUTORIAL: CREATING AN ADMIN USER                                  | 140        |
| 15.5. TUTORIAL: SETTING UP AN IDENTITY PROVIDER                         | 141        |
| 15.5.1. Setting up an IDP with GitHub                                   | 141        |
| 15.5.2. Granting other users access to the cluster                      | 145        |
| 15.6. TUTORIAL: GRANTING ADMIN PRIVILEGES                               | 146        |
| 15.6.1. Using the ROSA CLI  | 146        |
| 15.6.2. Using the Red Hat OpenShift Cluster Manager UI                  | 147        |
| 15.7. TUTORIAL: ACCESSING YOUR CLUSTER                                  | 148        |
| 15.7.1. Accessing your cluster using the CLI                            | 148        |
| 15.7.2. Accessing the cluster via the Hybrid Cloud Console              | 150        |
| 15.8. TUTORIAL: MANAGING WORKER NODES                                   | 150        |
| 15.8.1. Creating a machine pool   | 151        |
| 15.8.1.1. Creating a machine pool with the CLI                          | 151        |
| 15.8.1.2. Creating a machine pool with the UI                           | 151        |
| 15.8.2. Scaling worker nodes  | 154        |
| 15.8.2.1. Scaling worker nodes using the CLI                            | 154        |
| 15.8.2.2. Scaling worker nodes using the UI                             | 155        |
| 15.8.2.3. Adding node labels  | 155        |
| 15.8.3. Mixing node types   | 156        |
| 15.9. TUTORIAL: AUTOSCALING   | 157        |
| 15.9.1. Enabling autoscaling for an existing machine pool using the CLI | 157        |
| 15.9.2. Enabling autoscaling for an existing machine pool using the UI  | 158        |
| 15.10. TUTORIAL: UPGRADING YOUR CLUSTER                                 | 158        |
| 15.10.1. Manually upgrading your cluster using the CLI                  | 158        |
| 15.10.2. Manually upgrading your cluster using the UI                   | 159        |
| 15.10.3. Setting up automatic recurring upgrades                        | 159        |
| 15.11. TUTORIAL: DELETING YOUR CLUSTER                                  | 160        |
| 15.11.1. Deleting a ROSA cluster using the CLI                          | 160        |
| 15.11.2. Deleting a ROSA cluster using the UI                           | 161        |
| 15.12. TUTORIAL: OBTAINING SUPPORT                                      | 161        |
| 15.12.1. Adding support contacts  | 161        |
| 15.12.2. Contacting Red Hat for support using the UI                    | 162        |
| 15.12.3. Contacting Red Hat for support using the support page          | 162        |
| <b>CHAPTER 16. DEPLOYING AN APPLICATION</b> .....                       | <b>165</b> |
| 16.1. TUTORIAL: DEPLOYING AN APPLICATION                                | 165        |
| 16.1.1. Introduction  | 165        |
| 16.1.1.1. Lab overview  | 165        |
| 16.2. TUTORIAL: DEPLOYING AN APPLICATION                                | 165        |
| 16.2.1. Prerequisites   | 165        |

|   |     |
|---|-----|
| 16.3. TUTORIAL: DEPLOYING AN APPLICATION                | 165 |
| 16.3.1. Lab overview                                    | 166 |
| 16.3.1.1. Lab resources                                 | 166 |
| 16.3.1.2. About the OSToy application                   | 170 |
| 16.3.1.3. OSToy Application Diagram                     | 170 |
| 16.3.1.4. Understanding the OSToy UI                    | 170 |
| 16.4. TUTORIAL: DEPLOYING AN APPLICATION                | 171 |
| 16.4.1. Deploying the OSToy application with Kubernetes | 171 |
| 16.4.1.1. Retrieving the login command                  | 171 |
| 16.4.1.2. Creating a new project                        | 172 |
| 16.4.1.2.1. Using the CLI                               | 172 |
| 16.4.1.2.2. Using the web console                       | 173 |
| 16.4.1.3. Deploying the back-end microservice           | 173 |
| 16.4.1.4. Deploying the front-end service               | 173 |
| 16.4.1.5. Getting the route                             | 174 |
| 16.4.1.6. Viewing the application                       | 174 |
| 16.5. TUTORIAL: NETWORKING                              | 175 |
| 16.5.1. Intra-cluster networking                        | 175 |
| 16.6. TUTORIAL: PERSISTENT VOLUMES FOR CLUSTER STORAGE  | 177 |
| 16.6.1. Using persistent volumes                        | 177 |
| 16.6.1.1. Viewing a persistent volume claim             | 177 |
| 16.6.1.2. Storing your file                             | 177 |
| 16.6.1.3. Crashing the pod                              | 178 |
| 16.6.1.4. Confirming persistent storage                 | 178 |
| 16.6.1.5. Ending the session                            | 180 |
| 16.6.2. Additional resources                            | 180 |



## CHAPTER 1. TUTORIALS OVERVIEW

Step-by-step tutorials from Red Hat experts to help you get the most out of your Managed OpenShift cluster.

In an effort to make this Cloud Expert tutorial content available quickly, it may not yet be tested on every supported configuration.

## CHAPTER 2. TUTORIAL: ROSA WITH HCP ACTIVATION AND ACCOUNT LINKING

This tutorial describes the process for activating Red Hat OpenShift Service on AWS (ROSA) with hosted control planes (HCP) and linking to an AWS account, before deploying the first cluster.



### IMPORTANT

If you have received a private offer for the product, make sure to proceed according to the instructions provided with the private offer before following this tutorial. The private offer is designed either for a case when the product is already activated, which replaces an active subscription, or for first time activations.

## 2.1. PREREQUISITES

- Make sure to log in to the Red Hat account that you plan to associate with the AWS account where you have activated ROSA with HCP in previous steps.
- Only a single AWS account that will be used for service billing can be associated with a Red Hat account. Typically an organizational AWS account that has other AWS accounts, such as developer accounts, linked would be the one that is to be billed, rather than individual AWS end user accounts.
- Red Hat accounts belonging to the same Red Hat organization will be linked with the same AWS account. Therefore, you can manage who has access to creating ROSA with HCP clusters on the Red Hat organization account level.

## 2.2. SUBSCRIPTION ENABLEMENT AND AWS ACCOUNT SETUP

1. Activate the ROSA with HCP product at the AWS console page by clicking the **Get started** button:

| ROSA service fee pricing (US) |  |
|-------------------------------|--|
| Control plane                 | \$0.25 per hour <sup>1</sup>           |
| Worker nodes (hourly)         | \$0.17/4 vCPU per hour <sup>2</sup>    |
| Worker nodes (annually)       | \$1.28 per node per month <sup>3</sup> |

If you have activated ROSA before but did not complete the process, you can click the button and complete the account linking as described in the following steps.

2. Confirm that you want your contact information to be shared with Red Hat and enable the service:

[ROSA](#) > [Get Started](#)

## Verify ROSA prerequisites Info

This page verifies if your account meets the prerequisites to create a Red Hat OpenShift Service on AWS (ROSA) cluster.

### ROSA enablement Info


ROSA is jointly managed by AWS and Red Hat. Enable ROSA to create a connection with Red Hat, which is required for metering and billing.

**After enabling ROSA, you can create two types of clusters :**

- ROSA classic: cluster control plane infrastructure hosted in your AWS account.
- ROSA with hosted control planes (HCP): cluster control plane infrastructure hosted in Red Hat-owned AWS account.

You choose which control plane model to use when you create your cluster. [Learn more](#)

I agree to share my AWS account number and email address with Red Hat. This information is used for service metering and technical support outreach. You do not incur any fee when you enable ROSA.

[Enable ROSA with HCP and ROSA classic](#) 

Last checked on: February 14, 2024 at 14:18 (UTC)

- You will not be charged by enabling the service in this step. The connection is made for billing and metering that will take place only after you deploy your first cluster. This could take a few minutes.

3. After the process is completed, you will see a confirmation:

[ROSA](#) > [Get Started](#)

## Verify ROSA prerequisites Info

This page verifies if your account meets the prerequisites to create a Red Hat OpenShift Service on AWS (ROSA) cluster.

### ROSA enablement Info

ROSA is jointly managed by AWS and Red Hat. Enable ROSA to create a connection with Red Hat, which is required for metering and billing.

**After enabling ROSA, you can create two types of clusters :**

- ROSA classic: cluster control plane infrastructure hosted in your AWS account.
- ROSA with hosted control planes (HCP): cluster control plane infrastructure hosted in Red Hat-owned AWS account.

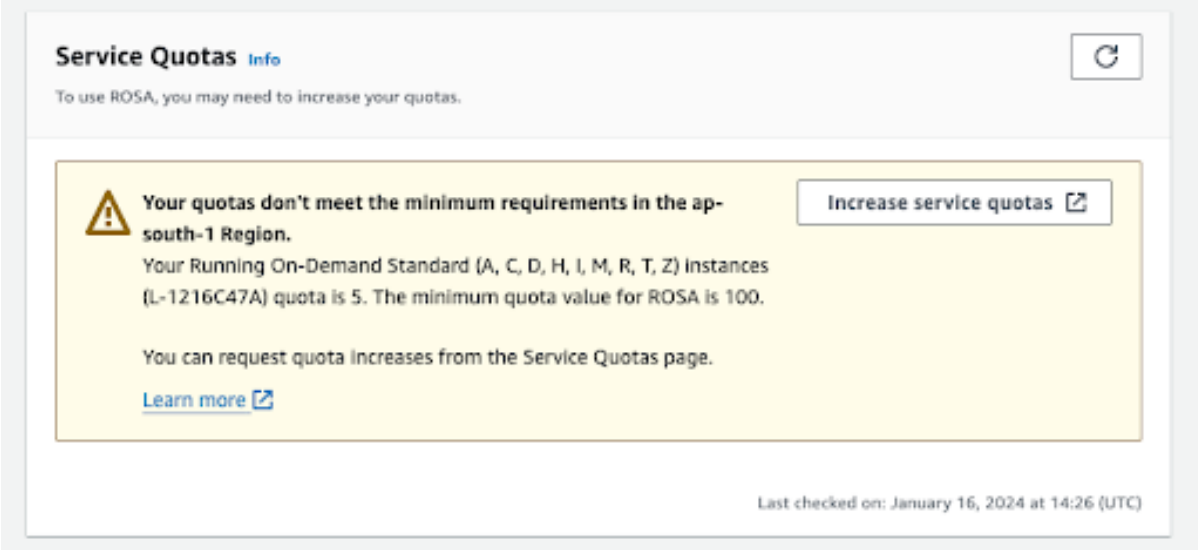
You choose which control plane model to use when you create your cluster. [Learn more](#)

You previously enabled ROSA HCP and ROSA classic.

Last checked on: February 14, 2024 at 14:06 (UTC)

**► AWS Organizations administrators: enable ROSA classic across your organization**

4. Other sections on this verification page show the status of additional prerequisites. In case any of these prerequisites are not met, a respective message is shown. Here is an example of insufficient quotas in the selected region:



The screenshot shows the AWS Service Quotas console. At the top, it says "Service Quotas" with an "Info" link and a refresh button. Below that, a message states: "To use ROSA, you may need to increase your quotas." A yellow warning box contains the following text: "Your quotas don't meet the minimum requirements in the ap-south-1 Region. Your Running On-Demand Standard (A, C, D, H, I, M, R, T, Z) instances (L-1216C47A) quota is 5. The minimum quota value for ROSA is 100. You can request quota increases from the Service Quotas page. Learn more" with an external link icon. To the right of the warning box is a button labeled "Increase service quotas" with an external link icon. At the bottom right of the console, it says "Last checked on: January 16, 2024 at 14:26 (UTC)".

- a. Click the **Increase service quotas** button or use the **Learn more** link to get more information about the about how to manage service quotas. In the case of insufficient quotas, note that quotas are region-specific. You can use the region switcher in the upper right corner of the web console to re-run the quota check for any region you are interested in and then submit service quota increase requests as needed.
5. If all the prerequisites are met, the page will look like this:

[ROSA](#) > [Get Started](#)

## Verify ROSA prerequisites [Info](#)

This page verifies if your account meets the prerequisites to create a Red Hat OpenShift Service on AWS (ROSA) cluster.

**ROSA enablement** [Info](#)  
 ROSA is jointly managed by AWS and Red Hat. Enable ROSA to create a connection with Red Hat, which is required for metering and billing.

After enabling ROSA, you can create two types of clusters:

- ROSA classic: cluster control plane infrastructure hosted in your AWS account.
- ROSA with hosted control planes (HCP): cluster control plane infrastructure hosted in Red Hat-owned AWS account.

You choose which control plane model to use when you create your cluster. [Learn more](#)

You previously enabled ROSA HCP and ROSA classic.

Last checked on: February 14, 2024 at 14:09 (UTC)

► **AWS Organizations administrators: enable ROSA classic across your organization**

**Service Quotas** [Info](#)  
 To use ROSA, you may need to increase your quotas.

Your quotas meet the requirements for ROSA.

Last checked on: February 14, 2024 at 14:09 (UTC)

**ELB service-linked role** [Info](#)

ROSA uses the Elastic Load Balancing (ELB) service-linked role to call AWS services on your behalf. If your account doesn't have this role, the role is created for you.

AWSServiceRoleForElasticLoadBalancing already exists. [View the role](#)

Last checked on: February 14, 2024 at 14:09 (UTC)

**Next steps**

Choose Continue to Red Hat to complete the steps for these prerequisites.

- [AWS and Red Hat account linking](#) [Info](#)
- [AWS account-wide role creation](#) [Info](#)

Once finished, you can then create clusters on the Red Hat console.

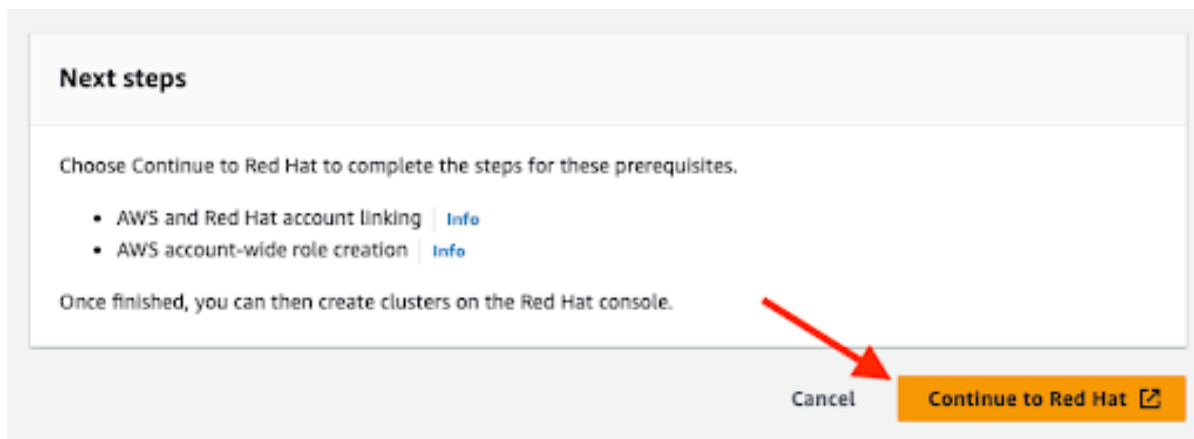
Cancel [Continue to Red Hat](#)

The ELB service-linked role is created for you automatically. You can click any of the small **Info** blue links to get contextual help and resources.

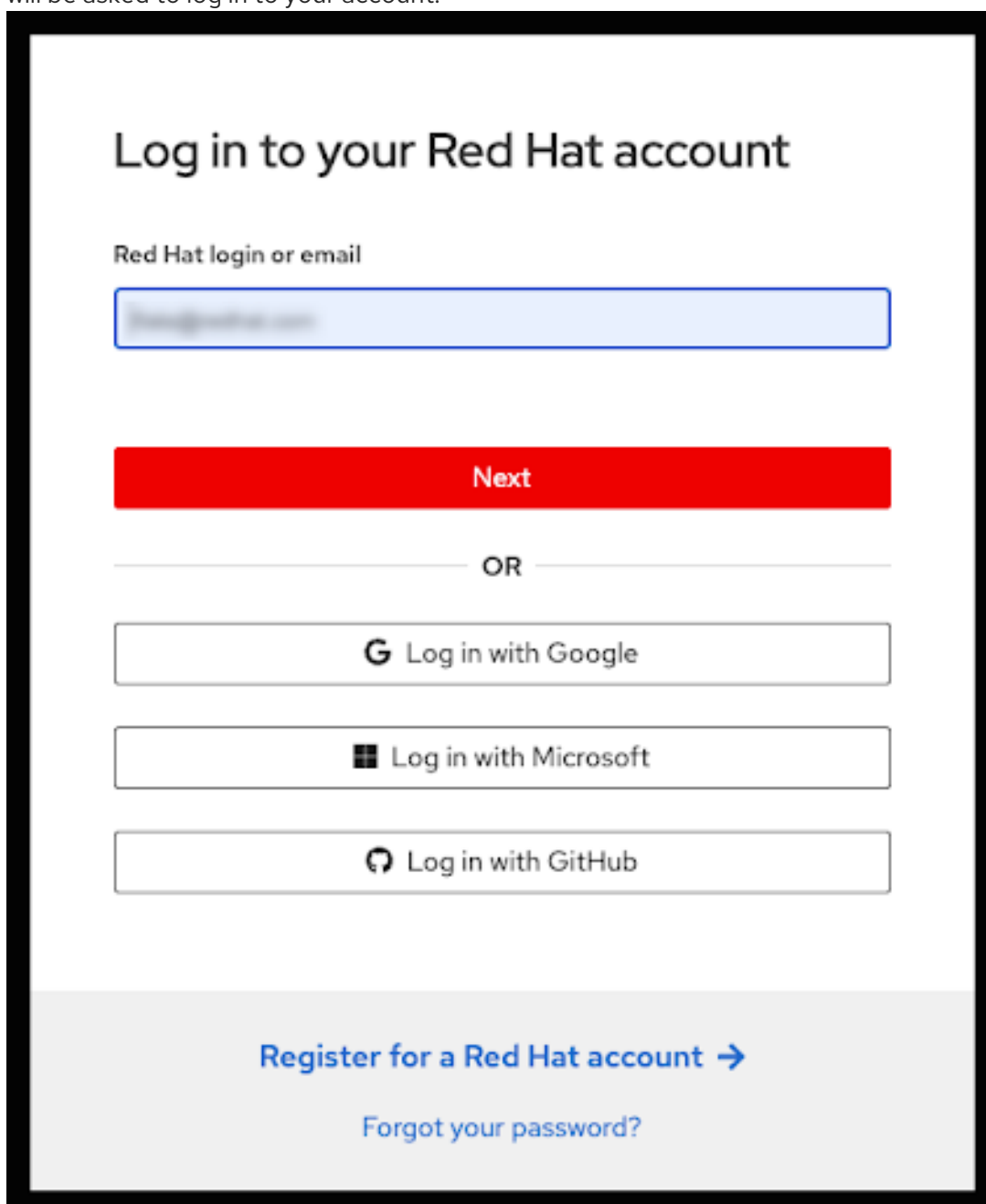
## 2.3. AWS AND RED HAT ACCOUNT AND SUBSCRIPTION LINKING

1. Click the orange **Continue to Red Hat** button to proceed with account linking:



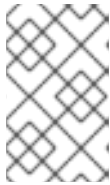


2. If you are not already logged in to your Red Hat account in your current browser's session, you will be asked to log in to your account:



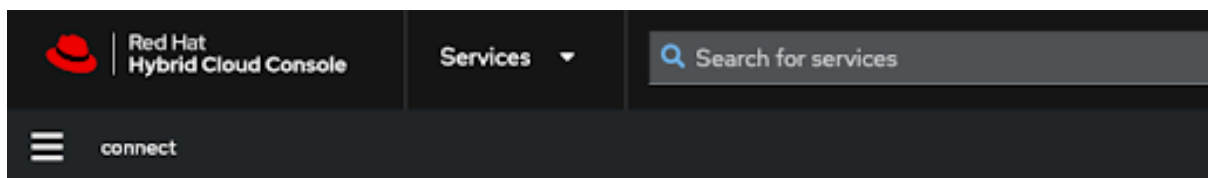
- You can also register for a new Red Hat account or reset your password on this page.

- Make sure to log in to the Red Hat account that you plan to associate with the AWS account where you have activated ROSA with HCP in previous steps.
  - Only a single AWS account that will be used for service billing can be associated with a Red Hat account. Typically an organizational AWS account that has other AWS accounts, such as developer accounts, linked would be the one that is to be billed, rather than individual AWS end user accounts.
  - Red Hat accounts belonging to the same Red Hat organization will be linked with the same AWS account. Therefore, you can manage who has access to creating ROSA with HCP clusters on the Red Hat organization account level.
3. Complete the Red Hat account linking after reviewing the terms and conditions:



#### NOTE

This step is available only if the logged-in Red Hat account, or the Red Hat organization managing the Red Hat account, was not linked to an AWS account before.



### Complete your account connection

Red Hat account number

AWS account ID

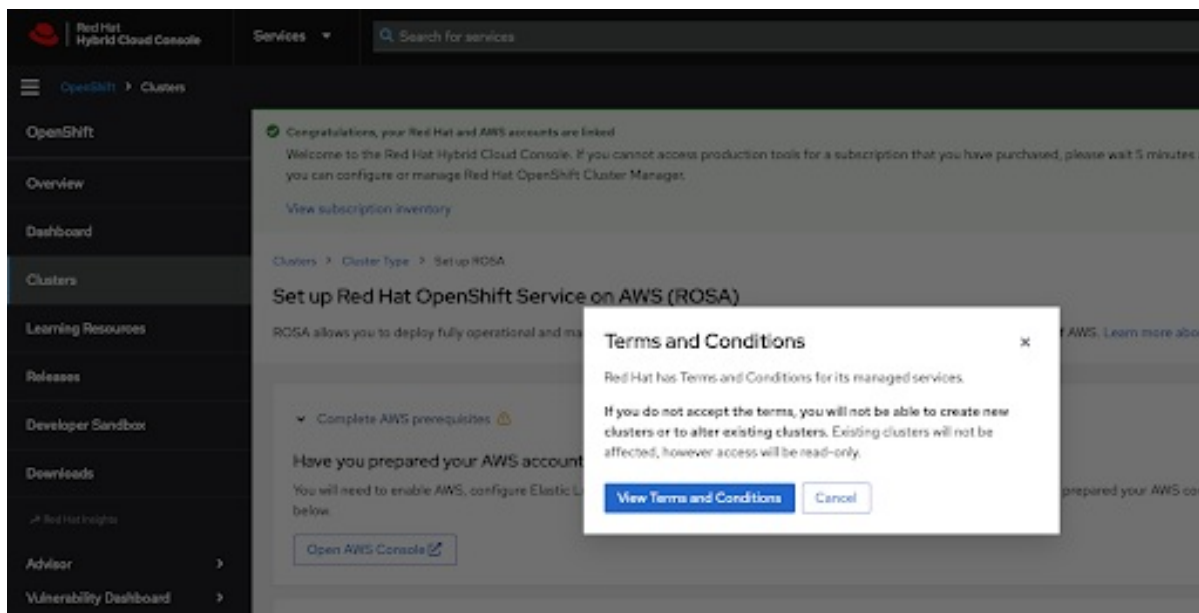
Subscription(s) Red Hat OpenShift Service on AWS with Hosted Control Plane

Terms and conditions \*

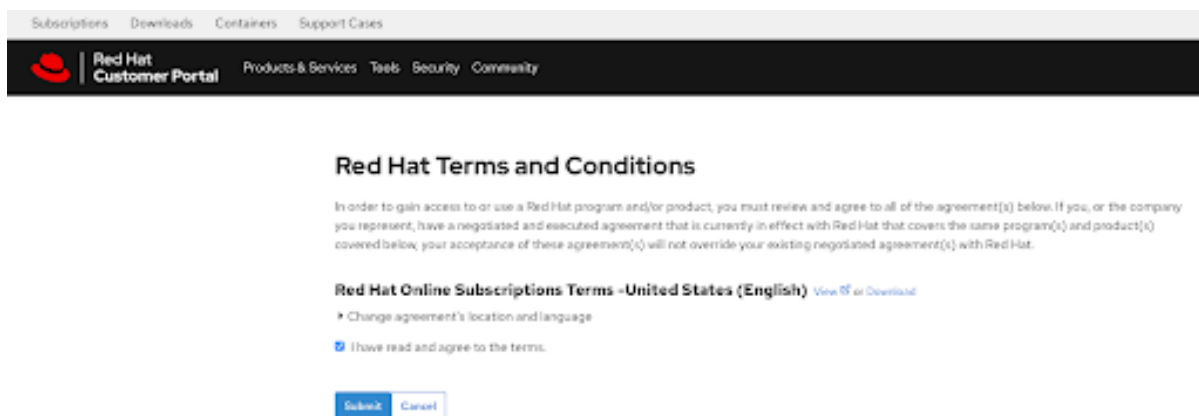
I have read and agreed to the [terms and conditions](#).

Both the Red Hat and AWS account numbers are shown on this screen.

4. Click the **Connect accounts** button if you agree with the service terms. If this is the first time you are using the Red Hat Hybrid Cloud Console, you will be asked to agree with the general managed services terms and conditions before being able to create the first ROSA cluster:



Additional terms that need to be reviewed and accepted will be shown after clicking the **View Terms and Conditions** button:



Submit your agreement once you have reviewed any additional terms when prompted at this time.

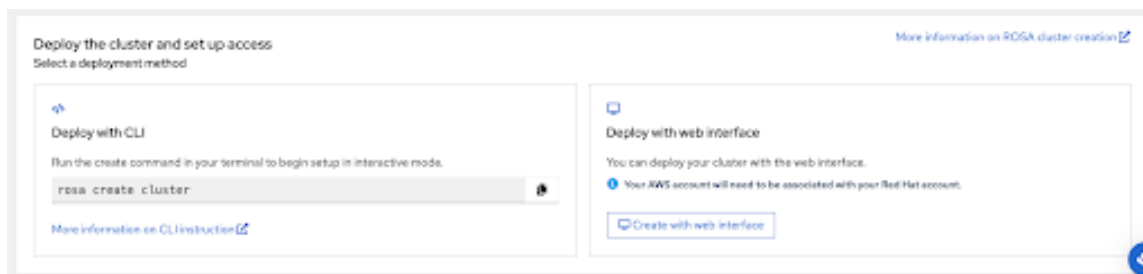
5. The Hybrid Cloud Console provides a confirmation that AWS prerequisites were completed and lists the first steps needed for cluster deployment:

The screenshot shows the Red Hat Hybrid Cloud Console interface. The left sidebar contains navigation links: OpenShift, Overview, Dashboard, Clusters (highlighted), Learning Resources, Releases, Developer Sandbox, Downloads, Red Hat Insights, Advisor, Vulnerability Dashboard, Subscriptions, and Cost Management. The main content area is titled 'Set up Red Hat OpenShift Service on AWS (ROSA)'. It features a progress indicator for 'Complete AWS prerequisites' and a section for 'Complete ROSA prerequisites'. Step 1 involves downloading and installing the ROSA and AWS command line tools (CLI) and adding them to the PATH. Step 2 involves logging in to the ROSA CLI with a Red Hat account token and creating AWS account roles and policies.

6. The following steps pertain to technical deployment of the cluster:

The screenshot details the 'Complete ROSA prerequisites' section. Step 1: Download and install the ROSA and AWS command line tools (CLI) and add it to your PATH. 1.1 Download the latest version of the ROSA CLI. The interface shows a dropdown menu for 'Linux' and 'x86\_64' with a 'Download the ROSA CLI' button. A link for 'Help with ROSA CLI setup' is provided. 1.2 Download the AWS CLI version 2. A link for 'Instructions to install the AWS CLI' is provided. Step 2: Log in to the ROSA CLI with your Red Hat account token and create AWS account roles and policies. 2.1 To authenticate, run this command: `rosa login --token="`. 2.2 To create the necessary account-wide roles and policies quickly, use the default auto method that's provided in the ROSA CLI: `rosa create account-roles --mode auto`. A note indicates that if manual creation is preferred, users should follow the provided instructions.

- It is possible that these steps will be performed on a different machine than where the service enablement and account linking were completed.
- As mentioned previously, any Red Hat account belonging to the Red Hat organization that was linked with the AWS account that activated the ROSA service will have access to creating a cluster and will be able to select the billing AWS account that was linked under this Red Hat organization previously.  
The last section of this page shows cluster deployment options, either using the **rosa** CLI or through the web console:



- The following steps describe cluster deployment using the **rosa** CLI.
- If you are interested in deployment using the web console only, you can skip to the *ROSA with HCP cluster deployment using the web console* section. However, note that the **rosa** CLI is required for certain tasks, such as creating the account roles. If you are deploying ROSA for the first time, follow these CLI steps until running the **rosa whoami** command, before skipping to the web console deployment steps.

## 2.4. ROSA WITH HCP CLUSTER DEPLOYMENT USING THE CLI

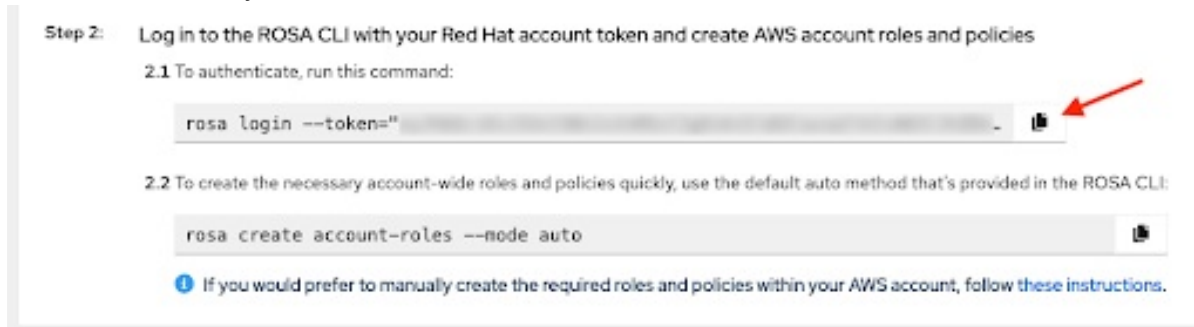
1. Click the **Download the ROSA CLI** button to download the ROSA command line interface (CLI) for your operating system and set it up as described in the [Help with ROSA CLI setup](#).



### IMPORTANT

Make sure that you have the most recent AWS CLI installed. See [Instructions to install the AWS CLI](#) for more information.

2. After the previous steps are completed, you can verify that both CLI are available by running the **rosa version**. This command shows an update notification if you are using an older version and **aws --version** commands in your terminal.
3. The prerequisite for creating a ROSA with HCP cluster is to log in using the **rosa** cli by the personalized command with your unique token shown under step 2.1. *To authenticate, run this command* on the web console. Use the **copy** button for easy copy and pasting of the command with full token into your terminal:



Do not share your unique token.

4. The final prerequisite before your first cluster deployment is making sure the necessary account-wide roles and policies are created. The **rosa** CLI can help with that by using the command shown under step 2.2. *To create the necessary account-wide roles and policies quickly...* on the web console. The alternative to that is manual creation of these roles and policies.
5. After logging in, creating the account roles, and verifying your identity using the **rosa whoami** command, your terminal will look similar to this:

```

➔ -- aws --version
aws-cli/2.15.11 Python/3.11.7 Darwin/23.2.0 source/arm64 prompt/off
➔ -- rosa version
1.2.33
I: Your ROSA CLI is up to date.
➔ -- rosa login --token="
[REDACTED]


I: Logged in as '[REDACTED]' on 'https://api.openshift.com'
➔ -- rosa create account-roles --mode auto
I: Logged in as '[REDACTED]' on 'https://api.openshift.com'
I: Validating AWS credentials...
I: AWS credentials are valid!
I: Validating AWS quota...
I: AWS quota ok. If cluster installation fails, validate actual AWS resource usage against https://docs.openshift.com/rosa/rosa_getting_start
ed/rosa-required-aws-service-quotas.html
I: Verifying whether OpenShift command-line tool is available...
I: Current OpenShift Client Version: 4.13.11
I: Creating account roles
I: By default, the create account-roles command creates two sets of account roles, one for classic ROSA clusters, and one for Hosted Control
Plane clusters.
In order to create a single set, please set one of the following flags: --classic or --hosted-cp
I: Creating classic account roles using 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-Installer-Role'
I: Created role 'ManagedOpenShift-Installer-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-Installer-Role'
I: Created role 'ManagedOpenShift-ControlPlane-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-ControlPlane-Role'
I: Created role 'ManagedOpenShift-Worker-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-Worker-Role'
I: Created role 'ManagedOpenShift-Support-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-Support-Role'
I: Creating hosted CP account roles using 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-HCP-RD-SA-Installer-Role'
I: Created role 'ManagedOpenShift-HCP-RD-SA-Installer-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-HCP-RD-SA-Installer-Role'
I: Created role 'ManagedOpenShift-HCP-RD-SA-Support-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-HCP-RD-SA-Support-Role'
I: Created role 'ManagedOpenShift-HCP-RD-SA-Worker-Role' with ARN 'arn:aws:iam::[REDACTED]:role/ManagedOpenShift-HCP-RD-SA-Worker-Role'
➔ -- rosa whoami
AWS ARN:          arn:aws:iam::[REDACTED]:user/[REDACTED]
AWS Account ID:  [REDACTED]
AWS Default Region: us-east-2
OCM API:         https://api.openshift.com
OCM Account Email: [REDACTED]
OCM Account ID:  [REDACTED]
OCM Account Name: [REDACTED]
OCM Account Username: [REDACTED]
OCM Organization External ID: [REDACTED]
OCM Organization ID: [REDACTED]
OCM Organization Name: [REDACTED]

```

6. Initiate the cluster deployment using the presented command. You can click the **copy** button again and paste the command in your terminal:

### Deploy the cluster and set up access


Select a deployment method

 **Deploy with CLI**

Run the create command in your terminal to begin setup in interactive mode.

```
rosa create cluster
```

[More information on CLI instruction](#)



7. To use a custom AWS profile, one of the non-default profiles specified in your `~/.aws/credentials`, you can add the `--profile <profile_name>` selector to the `rosa create cluster` command so that the command looks like `rosa create cluster --profile stage`. If no AWS CLI profile is specified using this option, the default AWS CLI profile will determine the AWS infrastructure profile into which the cluster is deployed. The billing AWS profile is selected in one of the following steps.
8. After entering a cluster name, you will be asked whether to use the hosted control plane. Select **yes**:

```

→ ~ rosa create cluster
I: Enabling interactive mode
? Cluster name: test-cluster-07
? Deploy cluster with Hosted Control Plane: [? for help] (y/N) y

```

9. When deploying a ROSA with HCP cluster, the billing AWS account needs to be specified:

```

→ ~ rosa create cluster
I: Enabling interactive mode
? Cluster name: test-cluster-07
? Deploy cluster with Hosted Control Plane: Yes
? Billing Account: [Use arrows to move, type to filter, ? for more help]
> [redacted] [Contract enabled]

```

- Only AWS accounts that were linked to the Red Hat organization that is currently used will be shown.
- The specified AWS account will be charged for using the ROSA service, regardless of whether the infrastructure AWS account is linked to it in the same AWS organization.
- You can see an indicator of whether the ROSA contract is enabled for a given AWS billing account or not.
- To select an AWS account that does not have the contract enabled, refer to the first few steps in this tutorial to enable the contract and allow the service charging, which is required for a successful cluster deployment.

10. In the following steps, you will specify technical details of the cluster that is to be deployed:

```

→ ~ rosa create cluster
I: Enabling interactive mode
? Cluster name: test-cluster-07
? Deploy cluster with Hosted Control Plane: Yes
? Billing Account: [redacted] [Contract enabled]
I: Using '[redacted]' as billing account.
I:
+-----+-----+
| Start Date      | Dec 08, 2023 |
| End Date        | Aug 17, 2024 |
| Number of vCPUs: | [redacted]    |
| Number of clusters: | [redacted]    |
+-----+-----+

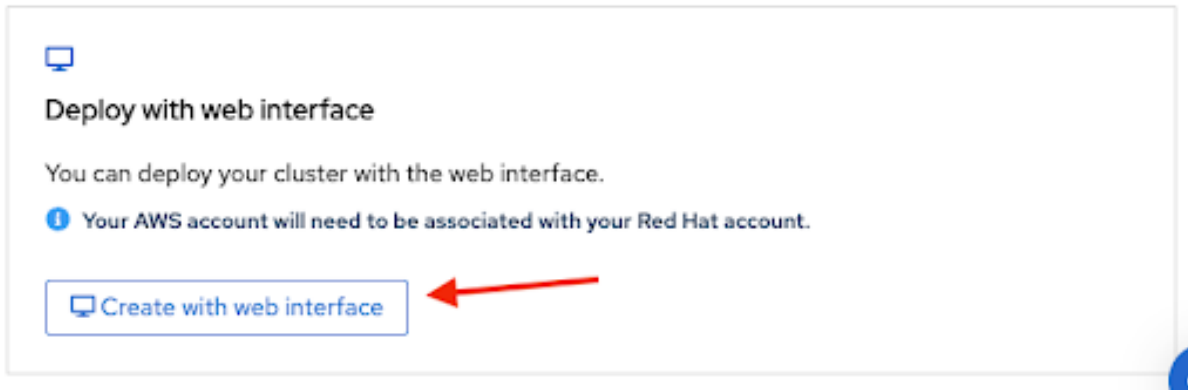
? OpenShift version (default = '4.14.8'): [Use arrows to move, type to filter, ? for more help]
> 4.14.8
4.14.7
4.14.6
4.14.5

```

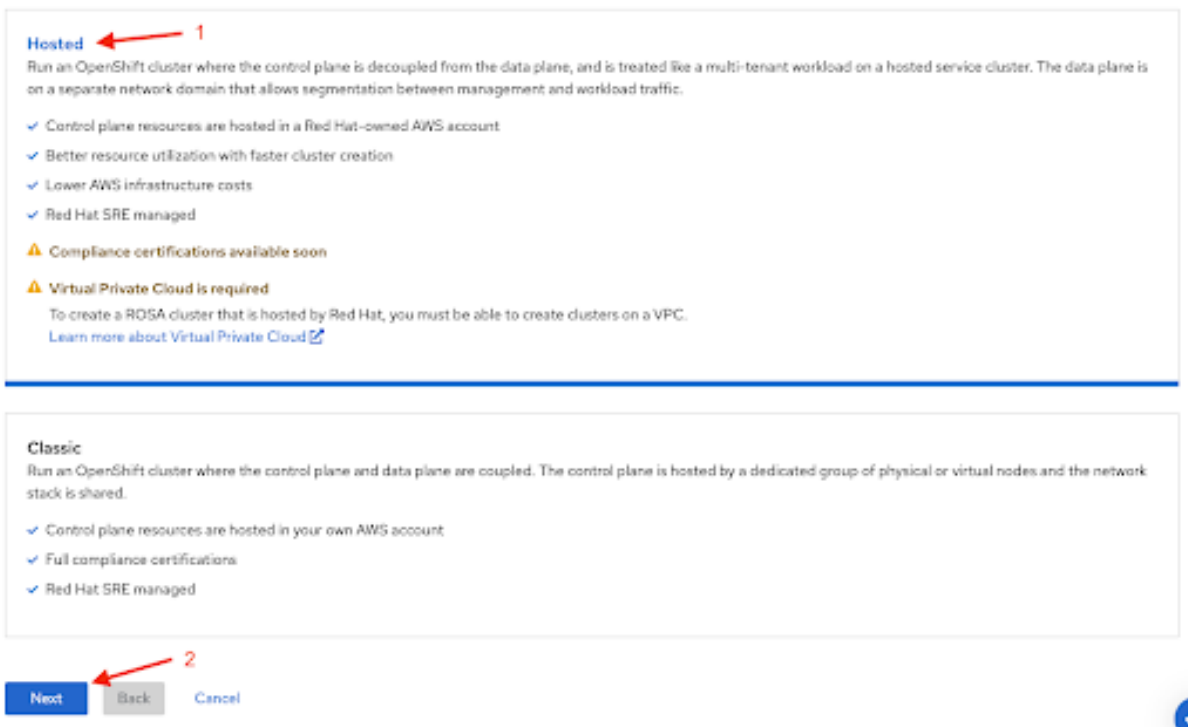
- These steps are beyond the scope of this tutorial. See [Creating ROSA with HCP clusters using the default options](#) for more details about how to complete the ROSA with HCP cluster deployment using the CLI.

## 2.5. ROSA WITH HCP CLUSTER DEPLOYMENT USING THE WEB CONSOLE

1. A cluster can be created using the web console by selecting the second option in the bottom section of the introductory **Set up ROSA** page:



- The first step when creating a ROSA cluster using the web console is the control plane selection. Make sure the **Hosted** option is selected before clicking the **Next** button:



- The next step **Accounts and roles** allows you specifying the infrastructure AWS account, into which the ROSA cluster will be deployed and where the resources will be consumed and managed:

### AWS infrastructure account

Select an AWS account that is associated with your Red Hat account or associate a new account. This account will contain the ROSA infrastructure.

Associated AWS infrastructure account \* ⓘ

XXXXXXXXXX

Refresh

How to associate a new AWS account

- Click the **How to associate a new AWS account** if you don't see the account into which you want to deploy the ROSA cluster for detailed information on how to create or link account roles for this association.
- The **rosa** CLI is used for this.



- If you are using multiple AWS accounts and have their profiles configured for the AWS CLI, you can use the **--profile** selector to specify the AWS profile when working with the **rosa** CLI commands.
4. The billing AWS account is selected in the immediately following section:

#### AWS billing account

This account will be charged for your subscription usage. You can select an already connected AWS account or sign in to a different AWS account that you want to connect to ROSA.

AWS billing account • ?

The screenshot displays the AWS billing account selection interface. At the top, there is a dropdown menu showing the current selected account and a 'Refresh' button. Below this is a search box labeled 'Filter by account ID'. A list of accounts is shown below the search box, with the first account highlighted and marked with a blue checkmark and the text 'Contract enabled'. At the bottom of the list, there is a button labeled 'Connect ROSA to a new AWS billing account' and a small icon with the text 'es'.

- Only AWS accounts that were linked to the Red Hat organization that is currently used will be shown.
- The specified AWS account will be charged for using the ROSA service, regardless of whether the infrastructure AWS account is linked to it in the same AWS organization.
- You can see an indicator whether the ROSA contract is enabled for a given AWS billing account or not.
- In case you would like to use an AWS account that does not have a contract enabled yet, you can either use the *Connect ROSA to a new AWS billing account* to reach the ROSA AWS console page, where you can activate it after logging in using the respective AWS account by following steps described earlier in this tutorial, or ask the administrator of the AWS account to do that for you.

The following steps past the billing AWS account selection are beyond the scope of this tutorial.

#### Additional resources

- For information on using the CLI to create a cluster, see [Creating a ROSA with HCP cluster using the CLI](#).
- See [this learning path](#) for more details on how to complete ROSA cluster deployment using the web console.

## CHAPTER 3. TUTORIAL: VERIFYING PERMISSIONS FOR A ROSA STS DEPLOYMENT

To proceed with the deployment of a ROSA cluster, an account must support the required roles and permissions. AWS Service Control Policies (SCPs) cannot block the API calls made by the installer or operator roles.

Details about the IAM resources required for an STS-enabled installation of ROSA can be found here: [About IAM resources for ROSA clusters that use STS](#)

This guide is validated for ROSA v4.11.X.

### 3.1. PREREQUISITES

- [AWS CLI](#)
- [ROSA CLI v1.2.6](#)
- [jq CLI](#)
- [AWS role with required permissions](#)

### 3.2. VERIFYING ROSA PERMISSIONS

To verify the permissions required for ROSA, we can run the script included in the following section without ever creating any AWS resources.

The script uses the **rosa**, **aws**, and **jq** CLI commands to create files in the working directory that will be used to verify permissions in the account connected to the current AWS configuration.

The AWS Policy Simulator is used to verify the permissions of each role policy against the API calls extracted by **jq**; results are then stored in a text file appended with **.results**.

This script is designed to verify the permissions for the current account and region.

### 3.3. USAGE INSTRUCTIONS

1. To use the script, run the following commands in a **bash** terminal (the **-p** option defines a prefix for the roles):

```
$ mkdir scratch
$ cd scratch
$ cat << 'EOF' > verify-permissions.sh
#!/bin/bash
while getopts 'p:' OPTION; do
  case "$OPTION" in
    p)
      PREFIX="$OPTARG"
      ;;
    ?)
      echo "script usage: $(basename \"$0\") [-p PREFIX]" >&2
      exit 1
      ;;
  esac
done
```

```

done
shift "$(($OPTIND -1))"
rosa create account-roles --mode manual --prefix $PREFIX
INSTALLER_POLICY=$(cat sts_installer_permission_policy.json | jq )
CONTROL_PLANE_POLICY=$(cat sts_instance_controlplane_permission_policy.json | jq)
WORKER_POLICY=$(cat sts_instance_worker_permission_policy.json | jq)
SUPPORT_POLICY=$(cat sts_support_permission_policy.json | jq)
simulatePolicy () {
  outputFile="${2}.results"
  echo $2
  aws iam simulate-custom-policy --policy-input-list "$1" --action-names $(jq '.Statement |
map(select(.Effect == "Allow"))|.Action | if type == "string" then . else .[] end' "$2" -r) --output
text > $outputFile
}
simulatePolicy "$INSTALLER_POLICY" "sts_installer_permission_policy.json"
simulatePolicy "$CONTROL_PLANE_POLICY"
"sts_instance_controlplane_permission_policy.json"
simulatePolicy "$WORKER_POLICY" "sts_instance_worker_permission_policy.json"
simulatePolicy "$SUPPORT_POLICY" "sts_support_permission_policy.json"
EOF
$ chmod +x verify-permissions.sh
$ ./verify-permissions.sh -p SimPolTest

```

2. After the script completes, review each results file to ensure that none of the required API calls are blocked:

```
$ for file in $(ls *.results); do echo $file; cat $file; done
```

The output will look similar to the following:

```

sts_installer_permission_policy.json.results
EVALUATIONRESULTS  autoscaling:DescribeAutoScalingGroups  allowed *
MATCHEDSTATEMENTS  PolicyInputList.1  IAM Policy
ENDPOSITION 6 195
STARTPOSITION 17 3
EVALUATIONRESULTS  ec2:AllocateAddress  allowed *
MATCHEDSTATEMENTS  PolicyInputList.1  IAM Policy
ENDPOSITION 6 195
STARTPOSITION 17 3
EVALUATIONRESULTS  ec2:AssociateAddress  allowed *
MATCHEDSTATEMENTS  PolicyInputList.1  IAM Policy
...

```



#### NOTE

If any actions are blocked, review the error provided by AWS and consult with your Administrator to determine if SCPs are blocking the required API calls.

## CHAPTER 4. TUTORIAL: USING AWS WAF AND AMAZON CLOUDFRONT TO PROTECT ROSA WORKLOADS

AWS WAF is a web application firewall that lets you monitor the HTTP and HTTPS requests that are forwarded to your protected web application resources.

You can use an Amazon CloudFront to add a Web Application Firewall (WAF) to your Red Hat OpenShift Service on AWS (ROSA) workloads. Using an external solution protects ROSA resources from experiencing denial of service due to handling the WAF.

### 4.1. PREREQUISITES

- A ROSA (HCP or Classic) cluster.
- You have access to the OpenShift CLI (**oc**).
- You have access to the AWS CLI (**aws**).

#### 4.1.1. Environment setup

- Prepare the environment variables:

```
$ export DOMAIN=apps.example.com 1
$ export AWS_PAGER=""
$ export CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}" | sed 's/-[a-z0-9]{5}$//')
$ export REGION=$(oc get infrastructure cluster -o=jsonpath="{.status.platformStatus.aws.region}")
$ export AWS_ACCOUNT_ID=$(aws sts get-caller-identity --query Account --output text)
$ export SCRATCH="/tmp/${CLUSTER}/cloudfront-waf"
$ mkdir -p ${SCRATCH}
$ echo "Cluster: ${CLUSTER}, Region: ${REGION}, AWS Account ID:
${AWS_ACCOUNT_ID}"
```

- 1 Replace with the custom domain you want to use for the **IngressController**.



#### NOTE

The "Cluster" output from the previous command might be the name of your cluster, the internal ID of your cluster, or the cluster's domain prefix. If you prefer to use another identifier, you can manually set this value by running the following command:

```
$ export CLUSTER=my-custom-value
```

### 4.2. SETTING UP THE SECONDARY INGRESS CONTROLLER

It is necessary to configure a secondary ingress controller to segment your external WAF-protected traffic from your standard (and default) cluster ingress controller.

#### Prerequisites

- A publicly trusted SAN or wildcard certificate for your custom domain, such as **CN=\*.apps.example.com**



## IMPORTANT

Amazon CloudFront uses HTTPS to communicate with your cluster's secondary ingress controller. As explained in the [Amazon CloudFront documentation](#), you cannot use a self-signed certificate for HTTPS communication between CloudFront and your cluster. Amazon CloudFront verifies that the certificate was issued by a trusted certificate authority.

## Procedure

1. Create a new TLS secret from a private key and a public certificate, where **fullchain.pem** is your full wildcard certificate chain (including any intermediaries) and **privkey.pem** is your wildcard certificate's private key.

### Example

```
$ oc -n openshift-ingress create secret tls waf-tls --cert=fullchain.pem --key=privkey.pem
```

2. Create a new **IngressController** resource:

### Example `waf-ingress-controller.yaml`

```
apiVersion: operator.openshift.io/v1
kind: IngressController
metadata:
  name: cloudfront-waf
  namespace: openshift-ingress-operator
spec:
  domain: apps.example.com 1
  defaultCertificate:
    name: waf-tls
  endpointPublishingStrategy:
    loadBalancer:
      dnsManagementPolicy: Unmanaged
      providerParameters:
        aws:
          type: NLB
          type: AWS
          scope: External
          type: LoadBalancerService
  routeSelector: 2
  matchLabels:
    route: waf
```

- 1** Replace with the custom domain you want to use for the **IngressController**.
- 2** Filters the set of routes serviced by the Ingress Controller. In this tutorial, we will use the **waf** route selector, but if no value was to be provided, no filtering would occur.

3. Apply the **IngressController**:

## Example

```
$ oc apply -f waf-ingress-controller.yaml
```

4. Verify that your IngressController has successfully created an external load balancer:

```
$ oc -n openshift-ingress get service/router-cloudfront-waf
```

## Example output

```
NAME                                TYPE           CLUSTER-IP      EXTERNAL-IP
PORT(S)                            AGE
router-cloudfront-waf               LoadBalancer   172.30.16.141
a68a838a7f26440bf8647809b61c4bc8-4225395f488830bd.elb.us-east-1.amazonaws.com
80:30606/TCP,443:31065/TCP          2m19s
```

### 4.2.1. Configure the AWS WAF

The [AWS WAF](#) service is a web application firewall that lets you monitor, protect, and control the HTTP and HTTPS requests that are forwarded to your protected web application resources, like ROSA.

1. Create a AWS WAF rules file to apply to our web ACL:

```
$ cat << EOF > ${SCRATCH}/waf-rules.json
[
  {
    "Name": "AWS-AWSManagedRulesCommonRuleSet",
    "Priority": 0,
    "Statement": {
      "ManagedRuleGroupStatement": {
        "VendorName": "AWS",
        "Name": "AWSManagedRulesCommonRuleSet"
      }
    },
    "OverrideAction": {
      "None": {}
    },
    "VisibilityConfig": {
      "SampledRequestsEnabled": true,
      "CloudWatchMetricsEnabled": true,
      "MetricName": "AWS-AWSManagedRulesCommonRuleSet"
    }
  },
  {
    "Name": "AWS-AWSManagedRulesSQLiRuleSet",
    "Priority": 1,
    "Statement": {
      "ManagedRuleGroupStatement": {
        "VendorName": "AWS",
        "Name": "AWSManagedRulesSQLiRuleSet"
      }
    },
    "OverrideAction": {
      "None": {}
    }
  }
]
```

```

    },
    "VisibilityConfig": {
      "SampledRequestsEnabled": true,
      "CloudWatchMetricsEnabled": true,
      "MetricName": "AWS-AWSManagedRulesSQLiRuleSet"
    }
  }
]
EOF

```

This will enable the Core (Common) and SQL AWS Managed Rule Sets.

2. Create an AWS WAF Web ACL using the rules we specified above:

```

$ WAF_WACL=$(aws wafv2 create-web-acl \
  --name cloudfront-waf \
  --region ${REGION} \
  --default-action Allow={} \
  --scope CLOUDFRONT \
  --visibility-config
SampledRequestsEnabled=true,CloudWatchMetricsEnabled=true,MetricName=${CLUSTER}-
waf-metrics \
  --rules file://${SCRATCH}/waf-rules.json \
  --query 'Summary.Name' \
  --output text)

```

### 4.3. CONFIGURE AMAZON CLOUDFRONT

1. Retrieve the newly created custom ingress controller's NLB hostname:

```

$ NLB=$(oc -n openshift-ingress get service router-cloudfront-waf \
  -o jsonpath='{.status.loadBalancer.ingress[0].hostname}')

```

2. Import your certificate into Amazon Certificate Manager, where **cert.pem** is your wildcard certificate, **fullchain.pem** is your wildcard certificate's chain and **privkey.pem** is your wildcard certificate's private key.



#### NOTE

Regardless of what region your cluster is deployed, you must import this certificate to **us-east-1** as Amazon CloudFront is a global AWS service.

#### Example

```

$ aws acm import-certificate --certificate file://cert.pem \
  --certificate-chain file://fullchain.pem \
  --private-key file://privkey.pem \
  --region us-east-1

```

3. Log into the [AWS console](#) to create a CloudFront distribution.
4. Configure the CloudFront distribution by using the following information:

**NOTE**

If an option is not specified in the table below, leave them the default (which may be blank).

| Option                         | Value  |
|--------------------------------|--|
| Origin domain                  | Output from the previous command <sup>[1]</sup>                        |
| Name                           | rosa-waf-ingress <sup>[2]</sup>  |
| Viewer protocol policy         | Redirect HTTP to HTTPS   |
| Allowed HTTP methods           | GET, HEAD, OPTIONS, PUT, POST, PATCH, DELETE                           |
| Cache policy                   | CachingDisabled  |
| Origin request policy          | AllViewer  |
| Web Application Firewall (WAF) | Enable security protections  |
| Use existing WAF configuration | true   |
| Choose a web ACL               | <b>cloudfront-waf</b>  |
| Alternate domain name (CNAME)  | *.apps.example.com <sup>[3]</sup>                                      |
| Custom SSL certificate         | Select the certificate you imported from the step above <sup>[4]</sup> |

1. Run **echo \${NLB}** to get the origin domain.
2. If you have multiple clusters, ensure the origin name is unique.
3. This should match the wildcard domain you used to create the custom ingress controller.
4. This should match the alternate domain name entered above.
5. Retrieve the Amazon CloudFront Distribution endpoint:

```
$ aws cloudfront list-distributions --query "DistributionList.Items[?Origins.Items[?DomainName=='${NLB}']].DomainName" --output text
```

6. Update the DNS of your custom wildcard domain with a CNAME to the Amazon CloudFront Distribution endpoint from the step above.

**Example**



```
*.apps.example.com CNAME d1b2c3d4e5f6g7.cloudfront.net
```

## 4.4. DEPLOY A SAMPLE APPLICATION

1. Create a new project for your sample application by running the following command:

```
$ oc new-project hello-world
```

2. Deploy a hello world application:

```
$ oc -n hello-world new-app --image=docker.io/openshift/hello-openshift
```

3. Create a route for the application specifying your custom domain name:

### Example

```
$ oc -n hello-world create route edge --service=hello-openshift hello-openshift-tls \
--hostname hello-openshift.${DOMAIN}
```

4. Label the route to admit it to your custom ingress controller:

```
$ oc -n hello-world label route.route.openshift.io/hello-openshift-tls route=waf
```

## 4.5. TEST THE WAF

1. Test that the app is accessible behind Amazon CloudFront:

### Example

```
$ curl "https://hello-openshift.${DOMAIN}"
```

### Example output

```
Hello OpenShift!
```

2. Test that the WAF denies a bad request:

### Example

```
$ curl -X POST "https://hello-openshift.${DOMAIN}" \
-F "user='<script><alert>Hello</alert></script>'"
```

### Example output

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<HTML><HEAD><META HTTP-EQUIV="Content-Type" CONTENT="text/html; charset=iso-
8859-1">
<TITLE>ERROR: The request could not be satisfied</TITLE>
</HEAD><BODY>
```

```
<H1>403 ERROR</H1>
<H2>The request could not be satisfied.</H2>
<HR noshade size="1px">
Request blocked.
We can't connect to the server for this app or website at this time. There might be too much
traffic or a configuration error. Try again later, or contact the app or website owner.
<BR clear="all">
If you provide content to customers through CloudFront, you can find steps to troubleshoot
and help prevent this error by reviewing the CloudFront documentation.
<BR clear="all">
<HR noshade size="1px">
<PRE>
Generated by cloudfront (CloudFront)
Request ID: nFk9q2yB8jddI6FZOTjdliezx-FwZtr8xUQUNT75HThPlrALDxbag==
</PRE>
<ADDRESS>
</ADDRESS>
</BODY></HTML>
```

The expected result is a **403 ERROR**, which means the AWS WAF is protecting your application.

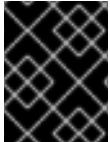
## 4.6. ADDITIONAL RESOURCES

- [Adding Extra Security with AWS WAF, CloudFront and ROSA | Amazon Web Services](#) on YouTube

## CHAPTER 5. TUTORIAL: USING AWS WAF AND AWS ALBS TO PROTECT ROSA WORKLOADS

AWS WAF is a web application firewall that lets you monitor the HTTP and HTTPS requests that are forwarded to your protected web application resources.

You can use an AWS Application Load Balancer (ALB) to add a Web Application Firewall (WAF) to your Red Hat OpenShift Service on AWS (ROSA) workloads. Using an external solution protects ROSA resources from experiencing denial of service due to handling the WAF.



### IMPORTANT

It is recommended that you use the more flexible [CloudFront method](#) unless you absolutely must use an ALB based solution.

## 5.1. PREREQUISITES

- Multiple availability zone (AZ) ROSA (HCP or Classic) cluster.



### NOTE

AWS ALBs require at least two *public* subnets across AZs, [per the AWS documentation](#). For this reason, only multiple AZ ROSA clusters can be used with ALBs.

- You have access to the OpenShift CLI (**oc**).
- You have access to the AWS CLI (**aws**).

### 5.1.1. Environment setup

- Prepare the environment variables:

```
$ export AWS_PAGER=""
$ export CLUSTER=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}")
$ export REGION=$(oc get infrastructure cluster -o=jsonpath="{.status.platformStatus.aws.region}")
$ export OIDC_ENDPOINT=$(oc get authentication.config.openshift.io cluster -o jsonpath='{.spec.serviceAccountIssuer}' | sed 's|^https://|')
$ export AWS_ACCOUNT_ID=$(aws sts get-caller-identity --query Account --output text)
$ export SCRATCH="/tmp/${CLUSTER}/alb-waf"
$ mkdir -p ${SCRATCH}
$ echo "Cluster: $(echo ${CLUSTER} | sed 's/-[a-z0-9]{5}$//'), Region: ${REGION}, OIDC Endpoint: ${OIDC_ENDPOINT}, AWS Account ID: ${AWS_ACCOUNT_ID}"
```

### 5.1.2. AWS VPC and subnets



## NOTE

This section only applies to clusters that were deployed into existing VPCs. If you did not deploy your cluster into an existing VPC, skip this section and proceed to the installation section below.

1. Set the below variables to the proper values for your ROSA deployment:

```
$ export VPC_ID=<vpc-id> 1
$ export PUBLIC_SUBNET_IDS=(<space-separated-list-of-ids>) 2
$ export PRIVATE_SUBNET_IDS=(<space-separated-list-of-ids>) 3
```

- 1** Replace with the VPC ID of the cluster, for example: **export VPC\_ID=vpc-04c429b7dbc4680ba**.
- 2** Replace with a space-separated list of the private subnet IDs of the cluster, making sure to preserve the (). For example: **export PUBLIC\_SUBNET\_IDS=(subnet-056fd6861ad332ba2 subnet-08ce3b4ec753fe74c subnet-071aa28228664972f)**.
- 3** Replace with a space-separated list of the private subnet IDs of the cluster, making sure to preserve the (). For example: **export PRIVATE\_SUBNET\_IDS=(subnet-0b933d72a8d72c36a subnet-0817eb72070f1d3c2 subnet-0806e64159b66665a)**.

2. Add a tag to your cluster's VPC with the cluster identifier:

```
$ aws ec2 create-tags --resources ${VPC_ID} \
--tags Key=kubernetes.io/cluster/${CLUSTER},Value=shared --region ${REGION}
```

3. Add a tag to your public subnets:

```
$ aws ec2 create-tags \
--resources ${PUBLIC_SUBNET_IDS} \
--tags Key=kubernetes.io/role/elb,Value='1' \
Key=kubernetes.io/cluster/${CLUSTER},Value=shared \
--region ${REGION}
```

4. Add a tag to your private subnets:

```
$ aws ec2 create-tags \
--resources ${PRIVATE_SUBNET_IDS} \
--tags Key=kubernetes.io/role/internal-elb,Value='1' \
Key=kubernetes.io/cluster/${CLUSTER},Value=shared \
--region ${REGION}
```

## 5.2. DEPLOY THE AWS LOAD BALANCER OPERATOR

The [AWS Load Balancer Operator](#) is used to install, manage and configure an instance of **aws-load-balancer-controller** in a ROSA cluster. To deploy ALBs in ROSA, we need to first deploy the AWS Load Balancer Operator.

1. Create a new project to deploy the AWS Load Balancer Operator into by running the following command:

■

```
$ oc new-project aws-load-balancer-operator
```

2. Create an AWS IAM policy for the AWS Load Balancer Controller if one does not already exist by running the following command:



#### NOTE

The policy is sourced from [the upstream AWS Load Balancer Controller policy](#). This is required by the operator to function.

```
$ POLICY_ARN=$(aws iam list-policies --query \
  "Policies[?PolicyName=='aws-load-balancer-operator-policy'].{ARN:Arn}" \
  --output text)
```

```
$ if [[ -z "${POLICY_ARN}" ]]; then
  wget -O "${SCRATCH}/load-balancer-operator-policy.json" \
    https://raw.githubusercontent.com/kubernetes-sigs/aws-load-balancer-
    controller/main/docs/install/iam_policy.json
  POLICY_ARN=$(aws --region "$REGION" --query Policy.Arn \
    --output text iam create-policy \
    --policy-name aws-load-balancer-operator-policy \
    --policy-document "file://${SCRATCH}/load-balancer-operator-policy.json")
fi
```

3. Create an AWS IAM trust policy for AWS Load Balancer Operator:

```
$ cat <<EOF > "${SCRATCH}/trust-policy.json"
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "${OIDC_ENDPOINT}:sub": ["system:serviceaccount:aws-load-balancer-operator:aws-
          load-balancer-operator-controller-manager", "system:serviceaccount:aws-load-balancer-
          operator:aws-load-balancer-controller-cluster"]
        }
      },
      "Principal": {
        "Federated": "arn:aws:iam::${AWS_ACCOUNT_ID}:oidc-provider/${OIDC_ENDPOINT}"
      },
      "Action": "sts:AssumeRoleWithWebIdentity"
    }
  ]
}
EOF
```

4. Create an AWS IAM role for the AWS Load Balancer Operator:

```
$ ROLE_ARN=$(aws iam create-role --role-name "${CLUSTER}-alb-operator" \
  --assume-role-policy-document "file://${SCRATCH}/trust-policy.json" \
  --query Role.Arn --output text)
```

5. Attach the AWS Load Balancer Operator policy to the IAM role we created previously by running the following command:

```
$ aws iam attach-role-policy --role-name "${CLUSTER}-alb-operator" \  
  --policy-arn ${POLICY_ARN}
```

6. Create a secret for the AWS Load Balancer Operator to assume our newly created AWS IAM role:

```
$ cat << EOF | oc apply -f -  
apiVersion: v1  
kind: Secret  
metadata:  
  name: aws-load-balancer-operator  
  namespace: aws-load-balancer-operator  
stringData:  
  credentials: |  
    [default]  
    role_arn = ${ROLE_ARN}  
    web_identity_token_file = /var/run/secrets/openshift/serviceaccount/token  
EOF
```

7. Install the AWS Load Balancer Operator:

```
$ cat << EOF | oc apply -f -  
apiVersion: operators.coreos.com/v1  
kind: OperatorGroup  
metadata:  
  name: aws-load-balancer-operator  
  namespace: aws-load-balancer-operator  
spec:  
  upgradeStrategy: Default  
  ---  
apiVersion: operators.coreos.com/v1alpha1  
kind: Subscription  
metadata:  
  name: aws-load-balancer-operator  
  namespace: aws-load-balancer-operator  
spec:  
  channel: stable-v1.0  
  installPlanApproval: Automatic  
  name: aws-load-balancer-operator  
  source: redhat-operators  
  sourceNamespace: openshift-marketplace  
  startingCSV: aws-load-balancer-operator.v1.0.0  
EOF
```

8. Deploy an instance of the AWS Load Balancer Controller using the operator:



#### NOTE

If you get an error here wait a minute and try again, it means the Operator has not completed installing yet.

```
$ cat << EOF | oc apply -f -
apiVersion: networking.olm.openshift.io/v1
kind: AWSLoadBalancerController
metadata:
  name: cluster
spec:
  credentials:
    name: aws-load-balancer-operator
  enabledAddons:
    - AWSWAFv2
EOF
```

9. Check the that the operator and controller pods are both running:

```
$ oc -n aws-load-balancer-operator get pods
```

You should see the following, if not wait a moment and retry:

```
NAME                                                    READY STATUS  RESTARTS  AGE
aws-load-balancer-controller-cluster-6ddf658785-pdp5d  1/1  Running  0         99s
aws-load-balancer-operator-controller-manager-577d9ffc9-w6zqn  2/2  Running  0         2m4s
```

### 5.3. DEPLOY A SAMPLE APPLICATION

1. Create a new project for our sample application:

```
$ oc new-project hello-world
```

2. Deploy a hello world application:

```
$ oc new-app -n hello-world --image=docker.io/openshift/hello-openshift
```

3. Convert the pre-created service resource to a NodePort service type:

```
$ oc -n hello-world patch service hello-openshift -p '{"spec":{"type":"NodePort"}}'
```

4. Deploy an AWS ALB using the AWS Load Balancer Operator:

```
$ cat << EOF | oc apply -f -
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: hello-openshift-alb
  namespace: hello-world
  annotations:
    alb.ingress.kubernetes.io/scheme: internet-facing
spec:
  ingressClassName: alb
  rules:
    - http:
        paths:
          - path: /
```

```

pathType: Exact
backend:
  service:
    name: hello-openshift
    port:
      number: 8080
EOF

```

5. Curl the AWS ALB Ingress endpoint to verify the hello world application is accessible:



#### NOTE

AWS ALB provisioning takes a few minutes. If you receive an error that says **curl: (6) Could not resolve host**, please wait and try again.

```

$ INGRESS=$(oc -n hello-world get ingress hello-openshift-alb -o
jsonpath='{.status.loadBalancer.ingress[0].hostname}')
$ curl "http://${INGRESS}"

```

#### Example output

```
Hello OpenShift!
```

### 5.3.1. Configure the AWS WAF

The [AWS WAF](#) service is a web application firewall that lets you monitor, protect, and control the HTTP and HTTPS requests that are forwarded to your protected web application resources, like ROSA.

1. Create a AWS WAF rules file to apply to our web ACL:

```

$ cat << EOF > ${SCRATCH}/waf-rules.json
[
  {
    "Name": "AWS-AWSManagedRulesCommonRuleSet",
    "Priority": 0,
    "Statement": {
      "ManagedRuleGroupStatement": {
        "VendorName": "AWS",
        "Name": "AWSManagedRulesCommonRuleSet"
      }
    },
    "OverrideAction": {
      "None": {}
    },
    "VisibilityConfig": {
      "SampledRequestsEnabled": true,
      "CloudWatchMetricsEnabled": true,
      "MetricName": "AWS-AWSManagedRulesCommonRuleSet"
    }
  },
  {
    "Name": "AWS-AWSManagedRulesSQLiRuleSet",
    "Priority": 1,

```



```

    "Statement": {
      "ManagedRuleGroupStatement": {
        "VendorName": "AWS",
        "Name": "AWSManagedRulesSQLiRuleSet"
      }
    },
    "OverrideAction": {
      "None": {}
    },
    "VisibilityConfig": {
      "SampledRequestsEnabled": true,
      "CloudWatchMetricsEnabled": true,
      "MetricName": "AWS-AWSManagedRulesSQLiRuleSet"
    }
  }
]
EOF

```

This will enable the Core (Common) and SQL AWS Managed Rule Sets.

2. Create an AWS WAF Web ACL using the rules we specified above:

```

$ WAF_ARN=$(aws wafv2 create-web-acl \
  --name ${CLUSTER}-waf \
  --region ${REGION} \
  --default-action Allow={} \
  --scope REGIONAL \
  --visibility-config
SampledRequestsEnabled=true,CloudWatchMetricsEnabled=true,MetricName=${CLUSTER}-
waf-metrics \
  --rules file://${SCRATCH}/waf-rules.json \
  --query 'Summary.ARN' \
  --output text)

```

3. Annotate the Ingress resource with the AWS WAF Web ACL ARN:

```

$ oc annotate -n hello-world ingress.networking.k8s.io/hello-openshift-alb \
  alb.ingress.kubernetes.io/wafv2-acl-arn=${WAF_ARN}

```

4. Wait for 10 seconds for the rules to propagate and test that the app still works:

```

$ curl "http://${INGRESS}"

```

### Example output

```

Hello OpenShift!

```

5. Test that the WAF denies a bad request:

```

$ curl -X POST "http://${INGRESS}" \
  -F "user='<script><alert>Hello</alert></script>'"

```

### Example output

```

■

```

```
<html>
<head><title>403 Forbidden</title></head>
<body>
<center><h1>403 Forbidden</h1></center>
</body>
</html>
```

**NOTE**

Activation of the AWS WAF integration can sometimes take several minutes. If you do not receive a **403 Forbidden** error, please wait a few seconds and try again.

The expected result is a **403 Forbidden** error, which means the AWS WAF is protecting your application.

## 5.4. ADDITIONAL RESOURCES

- [Adding Extra Security with AWS WAF, CloudFront and ROSA | Amazon Web Services](#) on YouTube

## CHAPTER 6. TUTORIAL: DEPLOYING OPENSIFT API FOR DATA PROTECTION ON A ROSA CLUSTER



### IMPORTANT

This content is authored by Red Hat experts, but has not yet been tested on every supported configuration.

### Prerequisites

- A [ROSA classic cluster](#)

### Environment

- Prepare the environment variables:



### NOTE

Change the cluster name to match your ROSA cluster and ensure you are logged into the cluster as an Administrator. Ensure all fields are outputted correctly before moving on.

```
$ export CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}" | sed 's/-[a-z0-9]{5}$//')
$ export ROSA_CLUSTER_ID=$(rosa describe cluster -c ${CLUSTER_NAME} --output json | jq -r .id)
$ export REGION=$(rosa describe cluster -c ${CLUSTER_NAME} --output json | jq -r .region.id)
$ export OIDC_ENDPOINT=$(oc get authentication.config.openshift.io cluster -o jsonpath='{.spec.serviceAccountIssuer}' | sed 's|^https://||')
$ export AWS_ACCOUNT_ID=`aws sts get-caller-identity --query Account --output text`
$ export CLUSTER_VERSION=`rosa describe cluster -c ${CLUSTER_NAME} -o json | jq -r .version.raw_id | cut -f -2 -d '.'`
$ export ROLE_NAME="${CLUSTER_NAME}-openshift-oadp-aws-cloud-credentials"
$ export AWS_PAGER=""
$ export SCRATCH="/tmp/${CLUSTER_NAME}/oadp"
$ mkdir -p ${SCRATCH}
$ echo "Cluster ID: ${ROSA_CLUSTER_ID}, Region: ${REGION}, OIDC Endpoint: ${OIDC_ENDPOINT}, AWS Account ID: ${AWS_ACCOUNT_ID}"
```

## 6.1. PREPARE AWS ACCOUNT

1. Create an IAM Policy to allow for S3 Access:

```
$ POLICY_ARN=$(aws iam list-policies --query "Policies[?PolicyName=='RosaOadpVer1'].{ARN:Arn}" --output text)
if [[ -z "${POLICY_ARN}" ]]; then
$ cat << EOF > ${SCRATCH}/policy.json
{
"Version": "2012-10-17",
"Statement": [
{
```

```

"Effect": "Allow",
"Action": [
  "s3:CreateBucket",
  "s3>DeleteBucket",
  "s3:PutBucketTagging",
  "s3:GetBucketTagging",
  "s3:PutEncryptionConfiguration",
  "s3:GetEncryptionConfiguration",
  "s3:PutLifecycleConfiguration",
  "s3:GetLifecycleConfiguration",
  "s3:GetBucketLocation",
  "s3:ListBucket",
  "s3:GetObject",
  "s3:PutObject",
  "s3>DeleteObject",
  "s3:ListBucketMultipartUploads",
  "s3:AbortMultipartUpload",
  "s3:ListMultipartUploadParts",
  "ec2:DescribeSnapshots",
  "ec2:DescribeVolumes",
  "ec2:DescribeVolumeAttribute",
  "ec2:DescribeVolumesModifications",
  "ec2:DescribeVolumeStatus",
  "ec2:CreateTags",
  "ec2:CreateVolume",
  "ec2:CreateSnapshot",
  "ec2>DeleteSnapshot"
],
"Resource": "*"
}
]]
EOF
$ POLICY_ARN=$(aws iam create-policy --policy-name "RosaOadpVer1" \
--policy-document file://${SCRATCH}/policy.json --query Policy.Arn \
--tags Key=rosa_openshift_version,Value=${CLUSTER_VERSION}
Key=rosa_role_prefix,Value=ManagedOpenShift
Key=operator_namespace,Value=openshift-oadp Key=operator_name,Value=openshift-oadp
\
--output text)
fi
$ echo ${POLICY_ARN}

```

2. Create an IAM Role trust policy for the cluster:

```

$ cat <<EOF > ${SCRATCH}/trust-policy.json
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "Federated": "arn:aws:iam::${AWS_ACCOUNT_ID}:oidc-provider/${OIDC_ENDPOINT}"
    },
    "Action": "sts:AssumeRoleWithWebIdentity",
    "Condition": {
      "StringEquals": {
        "${OIDC_ENDPOINT}:sub": [

```

```

        "system:serviceaccount:openshift-adp:openshift-adp-controller-manager",
        "system:serviceaccount:openshift-adp:velero"]
    }
  }
}}
}
EOF
$ ROLE_ARN=$(aws iam create-role --role-name \
"${ROLE_NAME}" \
--assume-role-policy-document file://${SCRATCH}/trust-policy.json \
--tags Key=rosa_cluster_id,Value=${ROSA_CLUSTER_ID}
Key=rosa_openshift_version,Value=${CLUSTER_VERSION}
Key=rosa_role_prefix,Value=ManagedOpenShift
Key=operator_namespace,Value=openshift-adp Key=operator_name,Value=openshift-oadp \
--query Role.Arn --output text)

$ echo ${ROLE_ARN}

```

3. Attach the IAM Policy to the IAM Role:

```

$ aws iam attach-role-policy --role-name "${ROLE_NAME}" \
--policy-arn ${POLICY_ARN}

```

## 6.2. DEPLOY OADP ON THE CLUSTER

1. Create a namespace for OADP:

```

$ oc create namespace openshift-adp

```

2. Create a credentials secret:

```

$ cat <<EOF > ${SCRATCH}/credentials
[default]
role_arn = ${ROLE_ARN}
web_identity_token_file = /var/run/secrets/openshift/serviceaccount/token
EOF
$ oc -n openshift-adp create secret generic cloud-credentials \
--from-file=${SCRATCH}/credentials

```

3. Deploy the OADP Operator:



### NOTE

There is currently an issue with version 1.1 of the Operator with backups that have a **PartiallyFailed** status. This does not seem to affect the backup and restore process, but it should be noted as there are issues with it.

```

$ cat << EOF | oc create -f -
apiVersion: operators.coreos.com/v1
kind: OperatorGroup
metadata:
  generateName: openshift-adp-
  namespace: openshift-adp

```

```

name: oadp
spec:
  targetNamespaces:
  - openshift-adp
---
apiVersion: operators.coreos.com/v1alpha1
kind: Subscription
metadata:
  name: redhat-oadp-operator
  namespace: openshift-adp
spec:
  channel: stable-1.2
  installPlanApproval: Automatic
  name: redhat-oadp-operator
  source: redhat-operators
  sourceNamespace: openshift-marketplace
EOF

```

- Wait for the Operator to be ready:

```
$ watch oc -n openshift-adp get pods
```

### Example output

```

NAME                                READY STATUS RESTARTS AGE
openshift-adp-controller-manager-546684844f-qqjhn 1/1   Running 0      22s

```

- Create Cloud Storage:

```

$ cat << EOF | oc create -f -
apiVersion: oadp.openshift.io/v1alpha1
kind: CloudStorage
metadata:
  name: ${CLUSTER_NAME}-oadp
  namespace: openshift-adp
spec:
  creationSecret:
    key: credentials
    name: cloud-credentials
  enableSharedConfig: true
  name: ${CLUSTER_NAME}-oadp
  provider: aws
  region: $REGION
EOF

```

- Check your application's storage default storage class:

```
$ oc get pvc -n <namespace> ❶
```

- Enter your application's namespace.

### Example output

```

NAME      STATUS VOLUME                                     CAPACITY ACCESS MODES
STORAGECLASS AGE
applog Bound pvc-351791ae-b6ab-4e8b-88a4-30f73caf5ef8 1Gi      RWO          gp3-
csi       4d19h
mysql Bound pvc-16b8e009-a20a-4379-accb-bc81fedd0621 1Gi      RWO          gp3-
csi       4d19h

```

```
$ oc get storageclass
```

### Example output

```

NAME          PROVISIONER          RECLAIMPOLICY VOLUMEBINDINGMODE
ALLOWVOLUMEEXPANSION AGE
gp2           kubernetes.io/aws-efs Delete           WaitForFirstConsumer true
4d21h
gp2-csi       ebs.csi.aws.com      Delete          WaitForFirstConsumer true
4d21h
gp3           ebs.csi.aws.com      Delete          WaitForFirstConsumer true
4d21h
gp3-csi (default) ebs.csi.aws.com      Delete          WaitForFirstConsumer true
4d21h

```

Using either gp3-csi, gp2-csi, gp3 or gp2 will work. If the application(s) that are being backed up are all using PV's with CSI, include the CSI plugin in the OADP DPA configuration.

### 7. CSI only: Deploy a Data Protection Application:

```

$ cat << EOF | oc create -f -
apiVersion: oadp.openshift.io/v1alpha1
kind: DataProtectionApplication
metadata:
  name: ${CLUSTER_NAME}-dpa
  namespace: openshift-adp
spec:
  backupImages: true
  features:
    dataMover:
      enable: false
  backupLocations:
  - bucket:
    cloudStorageRef:
      name: ${CLUSTER_NAME}-oadp
    credential:
      key: credentials
      name: cloud-credentials
    prefix: velero
    default: true
    config:
      region: ${REGION}
  configuration:
    velero:
      defaultPlugins:
      - openshift
      - aws

```

```
- csi
restic:
  enable: false
EOF
```

**NOTE**

If you run this command for CSI volumes, you can skip the next step.

## 8. Non-CSI volumes: Deploy a Data Protection Application:

```
$ cat << EOF | oc create -f -
apiVersion: oadp.openshift.io/v1alpha1
kind: DataProtectionApplication
metadata:
  name: ${CLUSTER_NAME}-dpa
  namespace: openshift-adp
spec:
  backupImages: true
  features:
    dataMover:
      enable: false
  backupLocations:
  - bucket:
      cloudStorageRef:
        name: ${CLUSTER_NAME}-oadp
      credential:
        key: credentials
        name: cloud-credentials
      prefix: velero
      default: true
      config:
        region: ${REGION}
  configuration:
    velero:
      defaultPlugins:
      - openshift
      - aws
    restic:
      enable: false
  snapshotLocations:
  - velero:
      config:
        credentialsFile: /tmp/credentials/openshift-adp/cloud-credentials-credentials
        enableSharedConfig: 'true'
        profile: default
        region: ${REGION}
      provider: aws
EOF
```





## NOTE

- In OADP 1.1.x ROSA STS environments, the container image backup and restore (**spec.backupImages**) value must be set to **false** as it is not supported.
- The Restic feature (**restic.enable=false**) is disabled and not supported in ROSA STS environments.
- The DataMover feature (**dataMover.enable=false**) is disabled and not supported in ROSA STS environments.

## 6.3. PERFORM A BACKUP



## NOTE

The following sample hello-world application has no attached persistent volumes. Either DPA configuration will work.

1. Create a workload to back up:

```
$ oc create namespace hello-world
$ oc new-app -n hello-world --image=docker.io/openshift/hello-openshift
```

2. Expose the route:

```
$ oc expose service/hello-openshift -n hello-world
```

3. Check that the application is working:

```
$ curl `oc get route/hello-openshift -n hello-world -o jsonpath='{.spec.host}'`
```

### Example output

```
Hello OpenShift!
```

4. Back up the workload:

```
$ cat << EOF | oc create -f -
apiVersion: velero.io/v1
kind: Backup
metadata:
  name: hello-world
  namespace: openshift-adp
spec:
  includedNamespaces:
  - hello-world
  storageLocation: ${CLUSTER_NAME}-dpa-1
  ttl: 720h0m0s
EOF
```

5. Wait until the backup is done:

```
$ watch "oc -n openshift-adp get backup hello-world -o json | jq .status"
```

### Example output

```
{
  "completionTimestamp": "2022-09-07T22:20:44Z",
  "expiration": "2022-10-07T22:20:22Z",
  "formatVersion": "1.1.0",
  "phase": "Completed",
  "progress": {
    "itemsBackedUp": 58,
    "totalItems": 58
  },
  "startTimestamp": "2022-09-07T22:20:22Z",
  "version": 1
}
```

6. Delete the demo workload:

```
$ oc delete ns hello-world
```

7. Restore from the backup:

```
$ cat << EOF | oc create -f -
apiVersion: velero.io/v1
kind: Restore
metadata:
  name: hello-world
  namespace: openshift-adp
spec:
  backupName: hello-world
EOF
```

8. Wait for the Restore to finish:

```
$ watch "oc -n openshift-adp get restore hello-world -o json | jq .status"
```

### Example output

```
{
  "completionTimestamp": "2022-09-07T22:25:47Z",
  "phase": "Completed",
  "progress": {
    "itemsRestored": 38,
    "totalItems": 38
  },
  "startTimestamp": "2022-09-07T22:25:28Z",
  "warnings": 9
}
```

9. Check that the workload is restored:

```
$ oc -n hello-world get pods
```

### Example output

```
NAME                READY STATUS RESTARTS AGE
hello-openshift-9f885f7c6-kdjbj 1/1   Running 0      90s
```

```
$ curl -k -o /dev/null -s -H 'Host: hello-openshift-9f885f7c6-kdjbj' http://localhost:8080
```

### Example output

```
Hello OpenShift!
```

- For troubleshooting tips please refer to the OADP team's [troubleshooting documentation](#)
- Additional sample applications can be found in the OADP team's [sample applications directory](#)

## 6.4. CLEANUP

- Delete the workload:

```
$ oc delete ns hello-world
```

- Remove the backup and restore resources from the cluster if they are no longer required:

```
$ oc delete backup hello-world
$ oc delete restore hello-world
```

- To delete the backup/restore and remote objects in s3:

```
$ velero backup delete hello-world
$ velero restore delete hello-world
```

- Delete the Data Protection Application:

```
$ oc -n openshift-adp delete dpa ${CLUSTER_NAME}-dpa
```

- Delete the Cloud Storage:

```
$ oc -n openshift-adp delete cloudstorage ${CLUSTER_NAME}-oadp
```



### WARNING

If this command hangs, you might need to delete the finalizer:

```
$ oc -n openshift-adp patch cloudstorage ${CLUSTER_NAME}-oadp -p
'{"metadata":{"finalizers":null}}' --type=merge
```

6. Remove the Operator if it is no longer required:

```
$ oc -n openshift-adp delete subscription oadp-operator
```

7. Remove the namespace for the Operator:

```
$ oc delete ns redhat-openshift-adp
```

8. Remove the Custom Resource Definitions from the cluster if you no longer wish to have them:

```
$ for CRD in `oc get crds | grep velero | awk '{print $1}'`; do oc delete crd $CRD; done  
$ for CRD in `oc get crds | grep -i oadp | awk '{print $1}'`; do oc delete crd $CRD; done
```

9. Delete the AWS S3 Bucket:

```
$ aws s3 rm s3://${CLUSTER_NAME}-oadp --recursive  
$ aws s3api delete-bucket --bucket ${CLUSTER_NAME}-oadp
```

10. Detach the Policy from the role:

```
$ aws iam detach-role-policy --role-name "${ROLE_NAME}" \  
--policy-arn "${POLICY_ARN}"
```

11. Delete the role:

```
$ aws iam delete-role --role-name "${ROLE_NAME}"
```

## CHAPTER 7. TUTORIAL: AWS LOAD BALANCER OPERATOR ON ROSA



### IMPORTANT

This content is authored by Red Hat experts, but has not yet been tested on every supported configuration.

### TIP

Load Balancers created by the AWS Load Balancer Operator cannot be used for [OpenShift Routes](#), and should only be used for individual services or ingress resources that do not need the full layer 7 capabilities of an OpenShift Route.

The [AWS Load Balancer Controller](#) manages AWS Elastic Load Balancers for a Red Hat OpenShift Service on AWS (ROSA) cluster. The controller provisions [AWS Application Load Balancers \(ALB\)](#) when you create Kubernetes Ingress resources and [AWS Network Load Balancers \(NLB\)](#) when implementing Kubernetes Service resources with a type of LoadBalancer.

Compared with the default AWS in-tree load balancer provider, this controller is developed with advanced annotations for both ALBs and NLBs. Some advanced use cases are:

- Using native Kubernetes Ingress objects with ALBs
- Integrate ALBs with the AWS Web Application Firewall (WAF) service
- Specify custom NLB source IP ranges
- Specify custom NLB internal IP addresses

The [AWS Load Balancer Operator](#) is used to install, manage and configure an instance of **aws-load-balancer-controller** in a ROSA cluster.

## 7.1. PREREQUISITES



### NOTE

AWS ALBs require a multi-AZ cluster, as well as three public subnets split across three AZs in the same VPC as the cluster. This makes ALBs unsuitable for many PrivateLink clusters. AWS NLBs do not have this restriction.

- [A multi-AZ ROSA classic cluster](#)
- BYO VPC cluster
- AWS CLI
- OC CLI

### 7.1.1. Environment

- Prepare the environment variables:
  -

```

$ export AWS_PAGER=""
$ export ROSA_CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}" | sed 's/-[a-z0-9]{5}$//')
$ export REGION=$(oc get infrastructure cluster -o=jsonpath="{.status.platformStatus.aws.region}")
$ export OIDC_ENDPOINT=$(oc get authentication.config.openshift.io cluster -o jsonpath='{.spec.serviceAccountIssuer}' | sed 's|^https://|')
$ export AWS_ACCOUNT_ID=$(aws sts get-caller-identity --query Account --output text)
$ export SCRATCH="/tmp/${ROSA_CLUSTER_NAME}/alb-operator"
$ mkdir -p ${SCRATCH}
$ echo "Cluster: ${ROSA_CLUSTER_NAME}, Region: ${REGION}, OIDC Endpoint: ${OIDC_ENDPOINT}, AWS Account ID: ${AWS_ACCOUNT_ID}"

```

## 7.1.2. AWS VPC and subnets



### NOTE

This section only applies to clusters that were deployed into existing VPCs. If you did not deploy your cluster into an existing VPC, skip this section and proceed to the installation section below.

1. Set the below variables to the proper values for your ROSA deployment:

```

$ export VPC_ID=<vpc-id>
$ export PUBLIC_SUBNET_IDS=<public-subnets>
$ export PRIVATE_SUBNET_IDS=<private-subnets>
$ export CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}")

```

2. Add a tag to your cluster's VPC with the cluster name:

```

$ aws ec2 create-tags --resources ${VPC_ID} --tags
Key=kubernetes.io/cluster/${CLUSTER_NAME},Value=owned --region ${REGION}

```

3. Add a tag to your public subnets:

```

$ aws ec2 create-tags \
--resources ${PUBLIC_SUBNET_IDS} \
--tags Key=kubernetes.io/role/elb,Value=" \
--region ${REGION}

```

4. Add a tag to your private subnets:

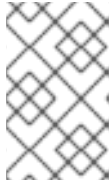
```

$ aws ec2 create-tags \
--resources "${PRIVATE_SUBNET_IDS}" \
--tags Key=kubernetes.io/role/internal-elb,Value=" \
--region ${REGION}

```

## 7.2. INSTALLATION

1. Create an AWS IAM policy for the AWS Load Balancer Controller:

**NOTE**

The policy is sourced from [the upstream AWS Load Balancer Controller policy](#) plus permission to create tags on subnets. This is required by the operator to function.

```
$ oc new-project aws-load-balancer-operator
$ POLICY_ARN=$(aws iam list-policies --query \
  "Policies[?PolicyName=='aws-load-balancer-operator-policy'].{ARN:Arn}" \
  --output text)
$ if [[ -z "${POLICY_ARN}" ]]; then
  wget -O "${SCRATCH}/load-balancer-operator-policy.json" \
    https://raw.githubusercontent.com/rh-mobb/documentation/main/content/rosa/aws-load-
balancer-operator/load-balancer-operator-policy.json
  POLICY_ARN=$(aws --region "$REGION" --query Policy.Arn \
    --output text iam create-policy \
    --policy-name aws-load-balancer-operator-policy \
    --policy-document "file://${SCRATCH}/load-balancer-operator-policy.json")
fi
$ echo $POLICY_ARN
```

2. Create an AWS IAM trust policy for AWS Load Balancer Operator:

```
$ cat <<EOF > "${SCRATCH}/trust-policy.json"
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Condition": {
        "StringEquals" : {
          "${OIDC_ENDPOINT}:sub": ["system:serviceaccount:aws-load-balancer-operator:aws-
load-balancer-operator-controller-manager", "system:serviceaccount:aws-load-balancer-
operator:aws-load-balancer-controller-cluster"]
        }
      },
      "Principal": {
        "Federated": "arn:aws:iam::$AWS_ACCOUNT_ID:oidc-provider/${OIDC_ENDPOINT}"
      },
      "Action": "sts:AssumeRoleWithWebIdentity"
    }
  ]
}
EOF
```

3. Create an AWS IAM role for the AWS Load Balancer Operator:

```
$ ROLE_ARN=$(aws iam create-role --role-name "${ROSA_CLUSTER_NAME}-alb-operator" \
  --assume-role-policy-document "file://${SCRATCH}/trust-policy.json" \
  --query Role.Arn --output text)
$ echo $ROLE_ARN

$ aws iam attach-role-policy --role-name "${ROSA_CLUSTER_NAME}-alb-operator" \
  --policy-arn $POLICY_ARN
```

- 4. Create a secret for the AWS Load Balancer Operator to assume our newly created AWS IAM role:

```
$ cat << EOF | oc apply -f -
apiVersion: v1
kind: Secret
metadata:
  name: aws-load-balancer-operator
  namespace: aws-load-balancer-operator
stringData:
  credentials: |
    [default]
    role_arn = $ROLE_ARN
    web_identity_token_file = /var/run/secrets/openshift/serviceaccount/token
EOF
```

- 5. Install the AWS Load Balancer Operator:

```
$ cat << EOF | oc apply -f -
apiVersion: operators.coreos.com/v1
kind: OperatorGroup
metadata:
  name: aws-load-balancer-operator
  namespace: aws-load-balancer-operator
spec:
  upgradeStrategy: Default
---
apiVersion: operators.coreos.com/v1alpha1
kind: Subscription
metadata:
  name: aws-load-balancer-operator
  namespace: aws-load-balancer-operator
spec:
  channel: stable-v1.0
  installPlanApproval: Automatic
  name: aws-load-balancer-operator
  source: redhat-operators
  sourceNamespace: openshift-marketplace
  startingCSV: aws-load-balancer-operator.v1.0.0
EOF
```

- 6. Deploy an instance of the AWS Load Balancer Controller using the operator:



#### NOTE

If you get an error here wait a minute and try again, it means the Operator has not completed installing yet.

```
$ cat << EOF | oc apply -f -
apiVersion: networking.olm.openshift.io/v1
kind: AWSLoadBalancerController
metadata:
  name: cluster
```



```
spec:
  credentials:
    name: aws-load-balancer-operator
EOF
```

7. Check the that the operator and controller pods are both running:

```
$ oc -n aws-load-balancer-operator get pods
```

You should see the following, if not wait a moment and retry:

```
NAME                                     READY STATUS  RESTARTS  AGE
aws-load-balancer-controller-cluster-6ddf658785-pdp5d      1/1   Running  0         99s
aws-load-balancer-operator-controller-manager-577d9ffcb9-w6zqn 2/2   Running  0         2m4s
```

## 7.3. VALIDATING THE DEPLOYMENT

1. Create a new project:

```
$ oc new-project hello-world
```

2. Deploy a hello world application:

```
$ oc new-app -n hello-world --image=docker.io/openshift/hello-openshift
```

3. Configure a NodePort service for the AWS ALB to connect to:

```
$ cat << EOF | oc apply -f -
apiVersion: v1
kind: Service
metadata:
  name: hello-openshift-nodeport
  namespace: hello-world
spec:
  ports:
    - port: 80
      targetPort: 8080
      protocol: TCP
  type: NodePort
  selector:
    deployment: hello-openshift
EOF
```

4. Deploy an AWS ALB using the AWS Load Balancer Operator:

```
$ cat << EOF | oc apply -f -
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: hello-openshift-alb
  namespace: hello-world
annotations:
```

```

    alb.ingress.kubernetes.io/scheme: internet-facing
spec:
  ingressClassName: alb
  rules:
  - http:
    paths:
    - path: /
      pathType: Exact
      backend:
        service:
          name: hello-openshift-nodeport
          port:
            number: 80
EOF

```

5. Curl the AWS ALB Ingress endpoint to verify the hello world application is accessible:



#### NOTE

AWS ALB provisioning takes a few minutes. If you receive an error that says **curl: (6) Could not resolve host**, please wait and try again.

```

$ INGRESS=$(oc -n hello-world get ingress hello-openshift-alb \
  -o jsonpath='{.status.loadBalancer.ingress[0].hostname}')
$ curl "http://${INGRESS}"

```

#### Example output

```
Hello OpenShift!
```

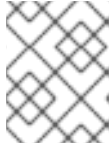
6. Deploy an AWS NLB for your hello world application:

```

$ cat << EOF | oc apply -f -
apiVersion: v1
kind: Service
metadata:
  name: hello-openshift-nlb
  namespace: hello-world
  annotations:
    service.beta.kubernetes.io/aws-load-balancer-type: external
    service.beta.kubernetes.io/aws-load-balancer-nlb-target-type: instance
    service.beta.kubernetes.io/aws-load-balancer-scheme: internet-facing
spec:
  ports:
  - port: 80
    targetPort: 8080
    protocol: TCP
  type: LoadBalancer
  selector:
    deployment: hello-openshift
EOF

```

7. Test the AWS NLB endpoint:

**NOTE**

NLB provisioning takes a few minutes. If you receive an error that says **curl: (6) Could not resolve host**, please wait and try again.

```
$ NLB=$(oc -n hello-world get service hello-openshift-nlb \
-o jsonpath='{.status.loadBalancer.ingress[0].hostname}')
$ curl "http://${NLB}"
```

**Example output**

```
Hello OpenShift!
```

**7.4. CLEANING UP**

1. Delete the hello world application namespace (and all the resources in the namespace):

```
$ oc delete project hello-world
```

2. Delete the AWS Load Balancer Operator and the AWS IAM roles:

```
$ oc delete subscription aws-load-balancer-operator -n aws-load-balancer-operator
$ aws iam detach-role-policy \
  --role-name "${ROSA_CLUSTER_NAME}-alb-operator" \
  --policy-arn $POLICY_ARN
$ aws iam delete-role \
  --role-name "${ROSA_CLUSTER_NAME}-alb-operator"
```

3. Delete the AWS IAM policy:

```
$ aws iam delete-policy --policy-arn $POLICY_ARN
```

# CHAPTER 8. TUTORIAL: CONFIGURING MICROSOFT ENTRA ID (FORMERLY AZURE ACTIVE DIRECTORY) AS AN IDENTITY PROVIDER

You can configure Microsoft Entra ID (formerly Azure Active Directory) as the cluster identity provider in Red Hat OpenShift Service on AWS (ROSA).

This tutorial guides you to complete the following tasks:

1. Register a new application in Entra ID for authentication.
2. Configure the application registration in Entra ID to include optional and group claims in tokens.
3. Configure the Red Hat OpenShift Service on AWS cluster to use Entra ID as the identity provider.
4. Grant additional permissions to individual groups.

## 8.1. PREREQUISITES

- You created a set of security groups and assigned users by following [the Microsoft documentation](#).

## 8.2. REGISTERING A NEW APPLICATION IN ENTRA ID FOR AUTHENTICATION

To register your application in Entra ID, first create the OAuth callback URL, then register your application.

### Procedure

1. Create the cluster's OAuth callback URL by changing the specified variables and running the following command:



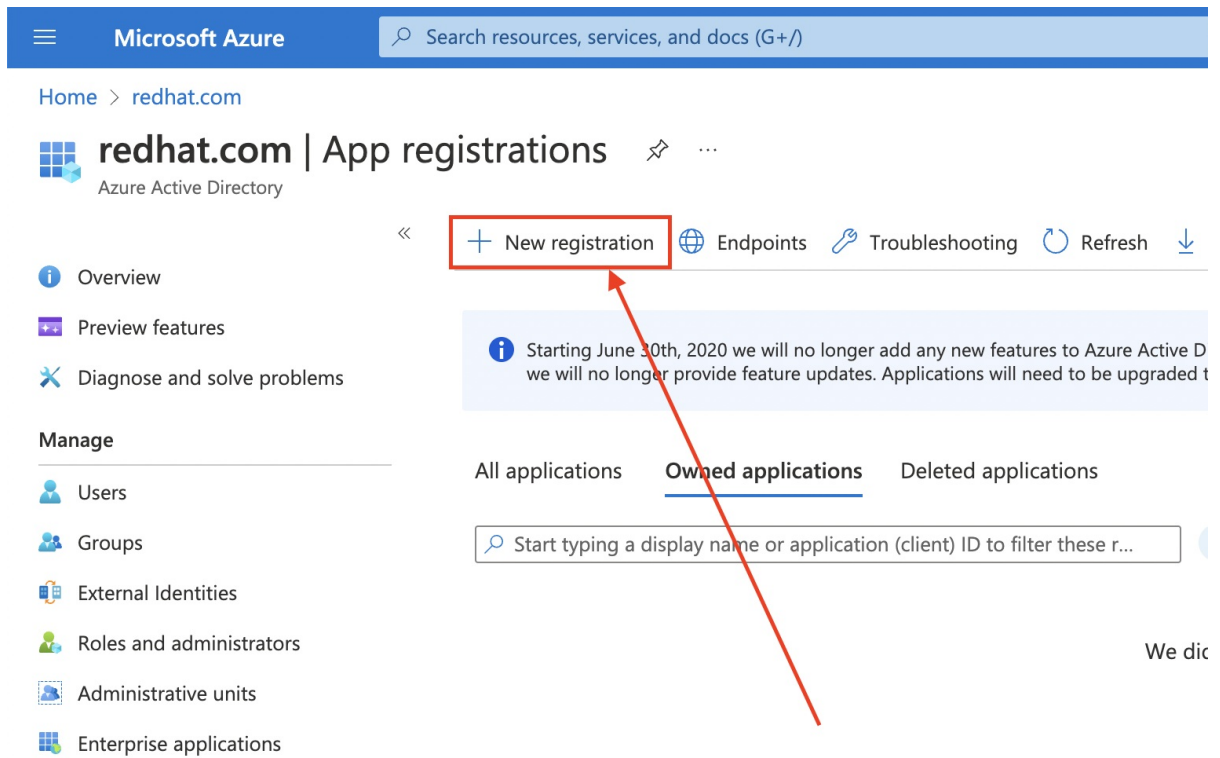
#### NOTE

Remember to save this callback URL; it will be required later in the process.

```
$ domain=$(rosa describe cluster -c <cluster_name> | grep "DNS" | grep -oE  
  \S+.openshiftapps.com')  
$ echo "OAuth callback URL: https://oauth-openshift.apps.$domain/oauth2callback/AAD"
```

The "AAD" directory at the end of the OAuth callback URL must match the OAuth identity provider name that you will set up later in this process.

2. Create the Entra ID application by logging in to the Azure portal, and select the [App registrations blade](#). Then, select **New registration** to create a new application.



3. Name the application, for example **openshift-auth**.
4. Select **Web** from the *Redirect URI* dropdown and enter the value of the OAuth callback URL you retrieved in the previous step.
5. After providing the required information, click **Register** to create the application.

Microsoft Azure

Home > redhat.com | App registrations >

## Register an application

**\* Name**  
The user-facing display name for this application (this can be changed later).

**Supported account types**  
Who can use this application or access this API?

Accounts in this organizational directory only (redhat.com only - Single tenant)

Accounts in any organizational directory (Any Azure AD directory - Multitenant)

Accounts in any organizational directory (Any Azure AD directory - Multitenant) and personal Microsoft accounts (e.g. Skype, Xbox)

Personal Microsoft accounts only

[Help me choose...](#)

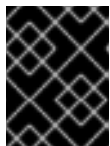
**Redirect URI (optional)**  
We'll return the authentication response to this URI after successfully authenticating the user. Providing this now is optional and it can be changed later, but a value is required for most authentication scenarios.

Register an app you're working on here. Integrate gallery apps and other apps from outside your organization by adding from [Enterprise applications](#)

By proceeding, you agree to the [Microsoft Platform Policies](#)

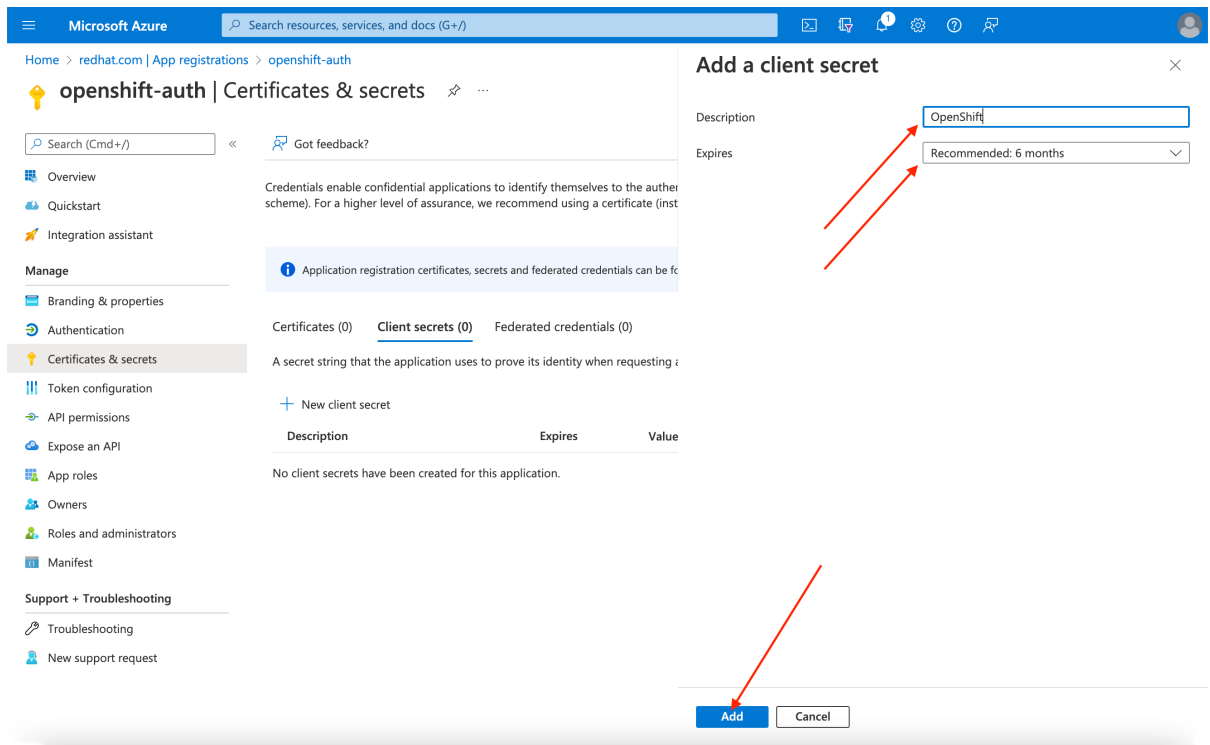
6. Select the **Certificates & secrets** sub-blade and select **New client secret**.

- Complete the requested details and store the generated client secret value. This secret is required later in this process.

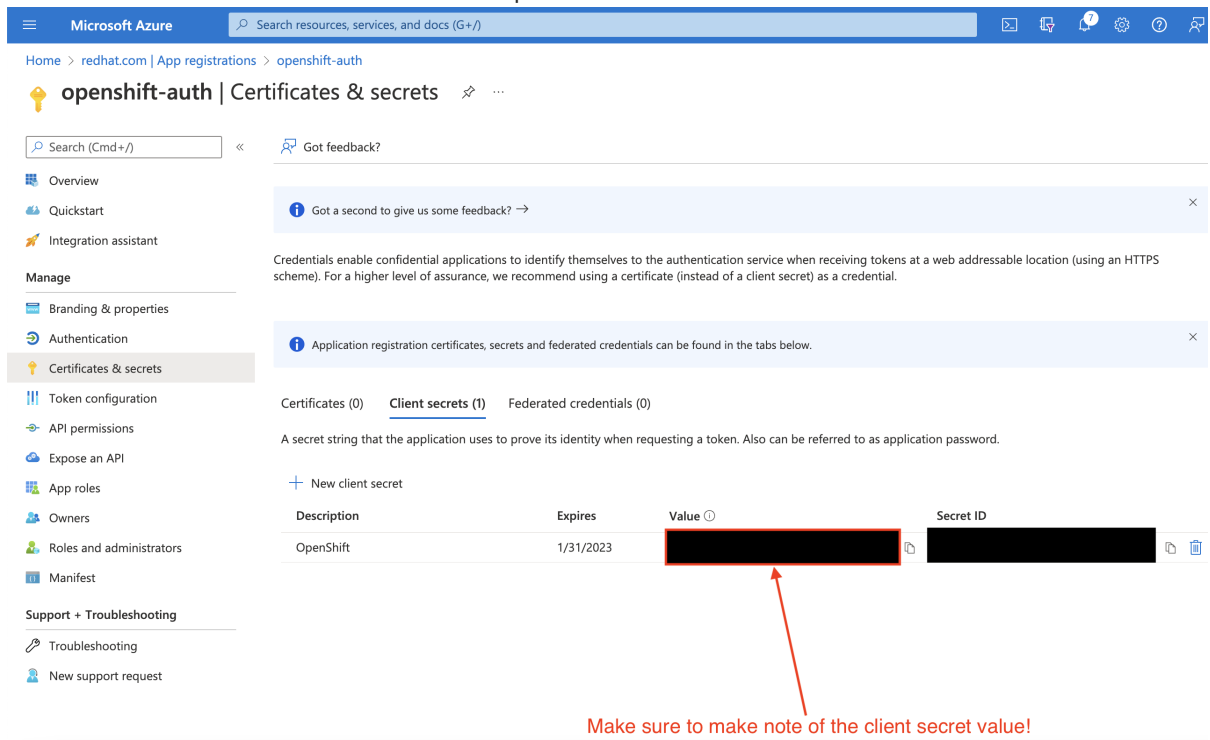


**IMPORTANT**

After initial setup, you cannot see the client secret. If you did not record the client secret, you must generate a new one.



8. Select the **Overview** sub-blade and note the **Application (client) ID** and **Directory (tenant) ID**. You will need these values in a future step.



### 8.3. CONFIGURING THE APPLICATION REGISTRATION IN ENTRA ID TO INCLUDE OPTIONAL AND GROUP CLAIMS

So that Red Hat OpenShift Service on AWS has enough information to create the user’s account, you must configure Entra ID to give two optional claims: **email** and **preferred\_username**. For more information about optional claims in Entra ID, see [the Microsoft documentation](#).

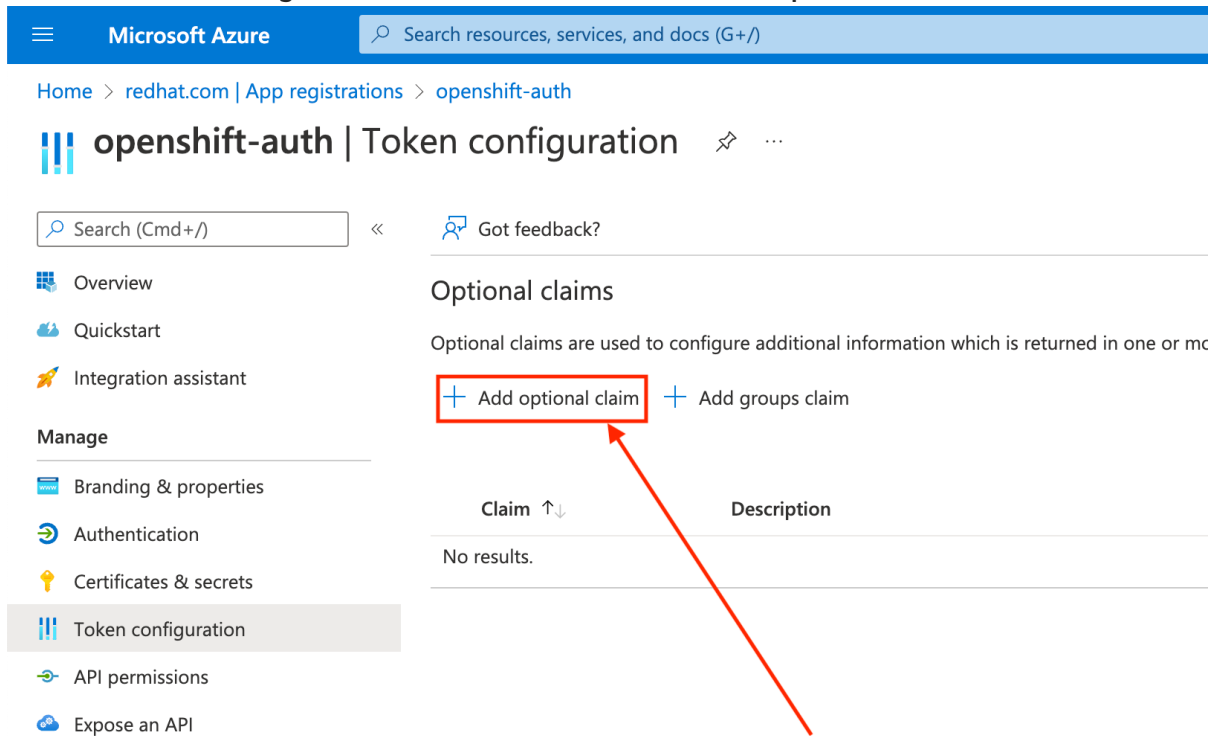
In addition to individual user authentication, Red Hat OpenShift Service on AWS provides group claim functionality. This functionality allows an OpenID Connect (OIDC) identity provider, such as Entra ID, to offer a user’s group membership for use within Red Hat OpenShift Service on AWS.



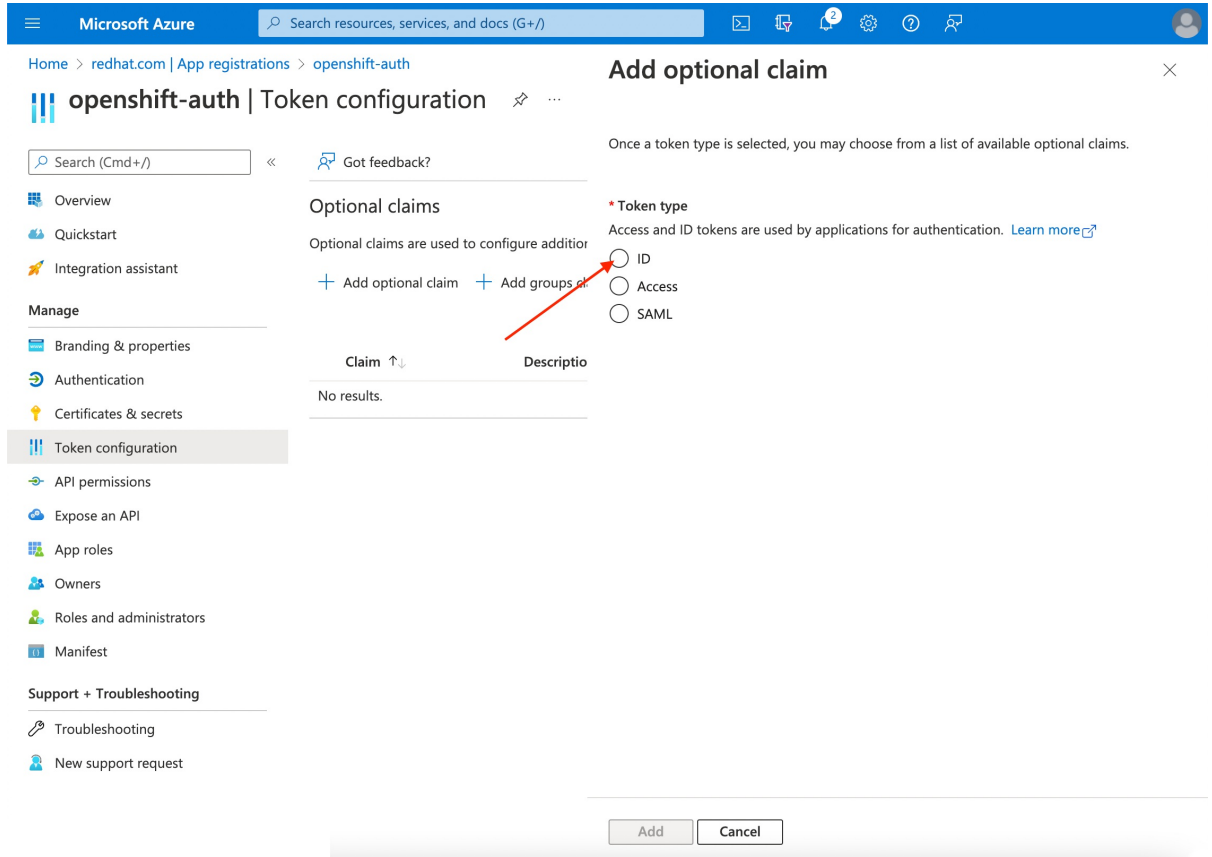
## Configuring optional claims

You can configure the optional claims in Entra ID.

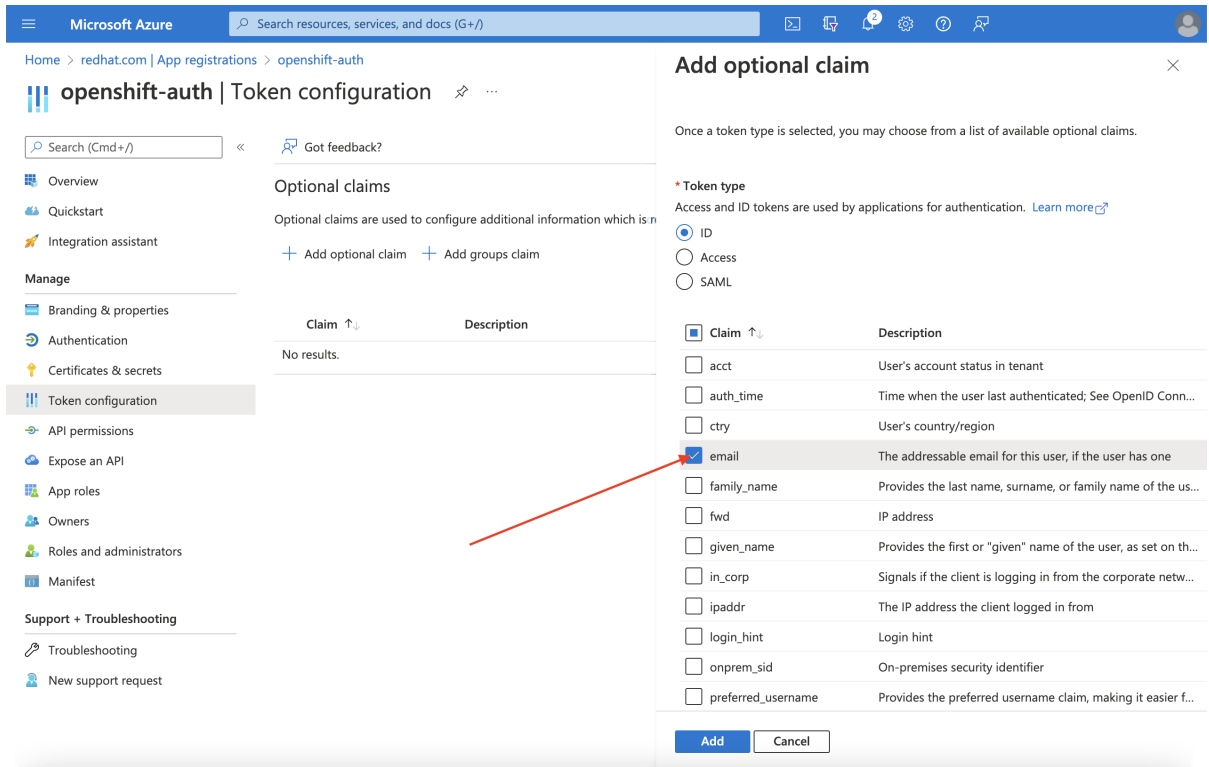
1. Click the **Token configuration** sub-blade and select the **Add optional claim** button.



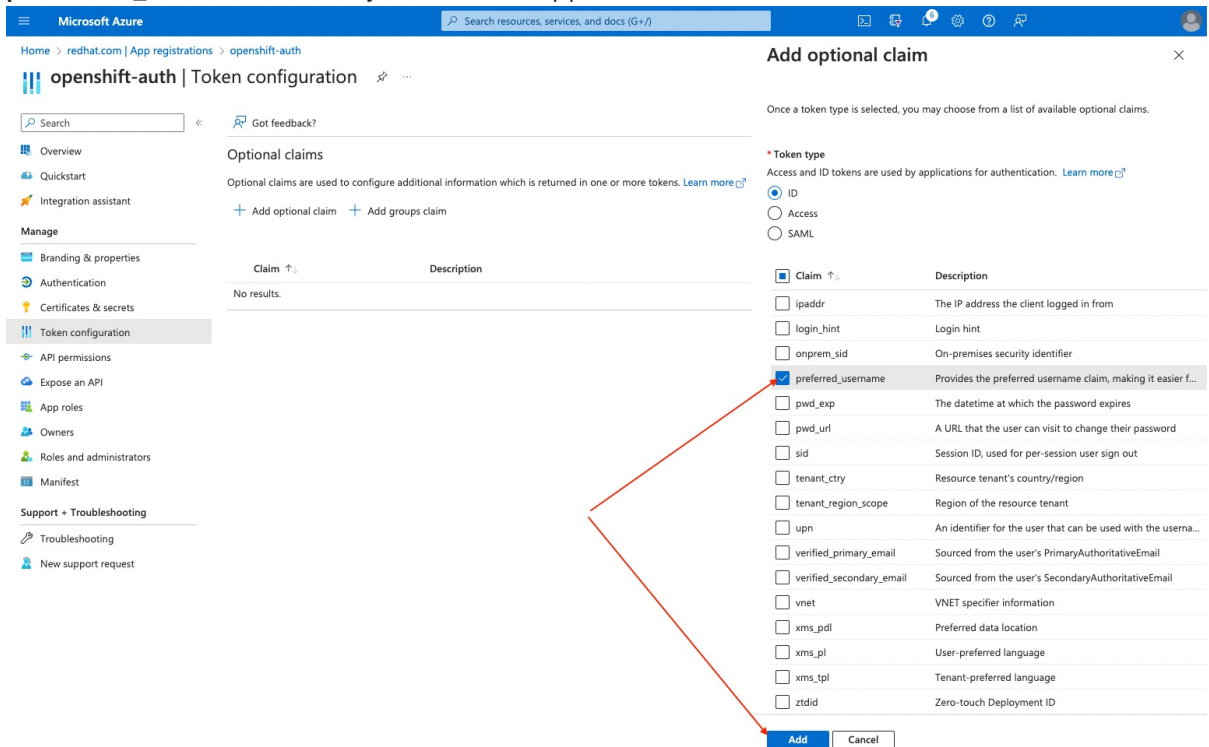
2. Select the **ID** radio button.



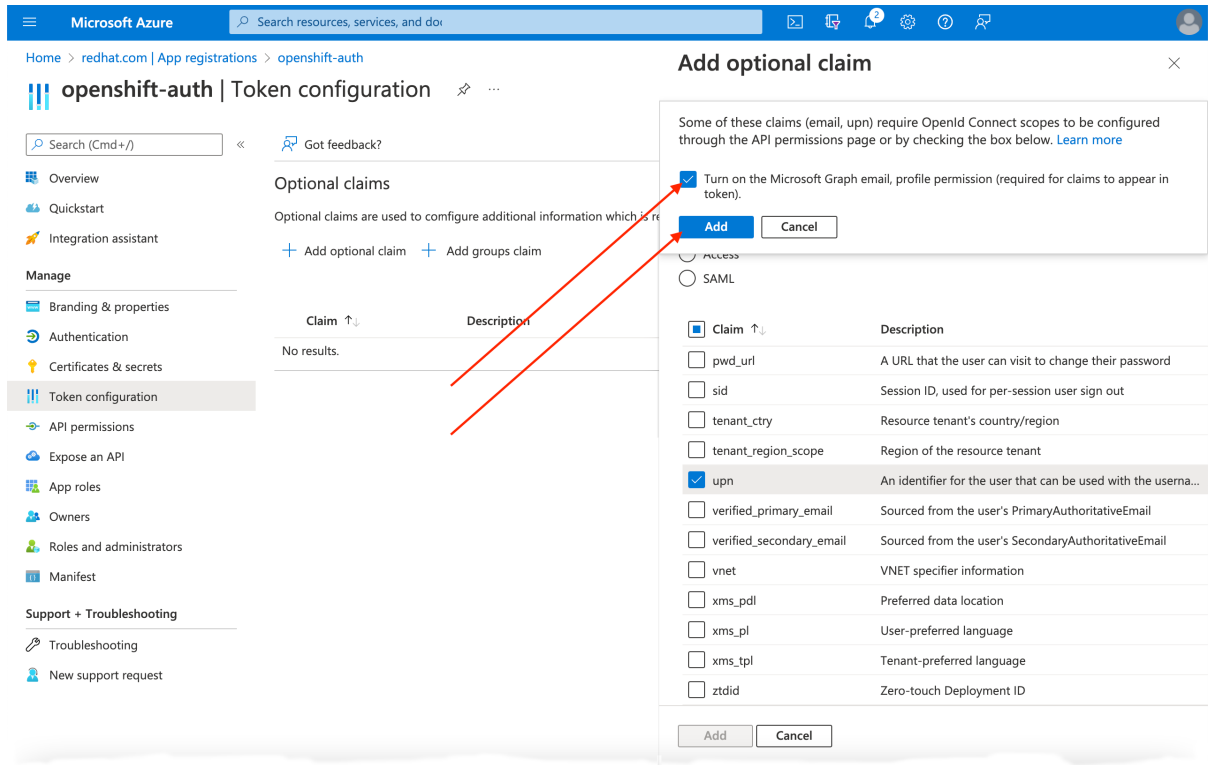
3. Select the **email** claim checkbox.



4. Select the **preferred\_username** claim checkbox. Then, click **Add** to configure the **email** and **preferred\_username** claims your Entra ID application.



5. A dialog box appears at the top of the page. Follow the prompt to enable the necessary Microsoft Graph permissions.

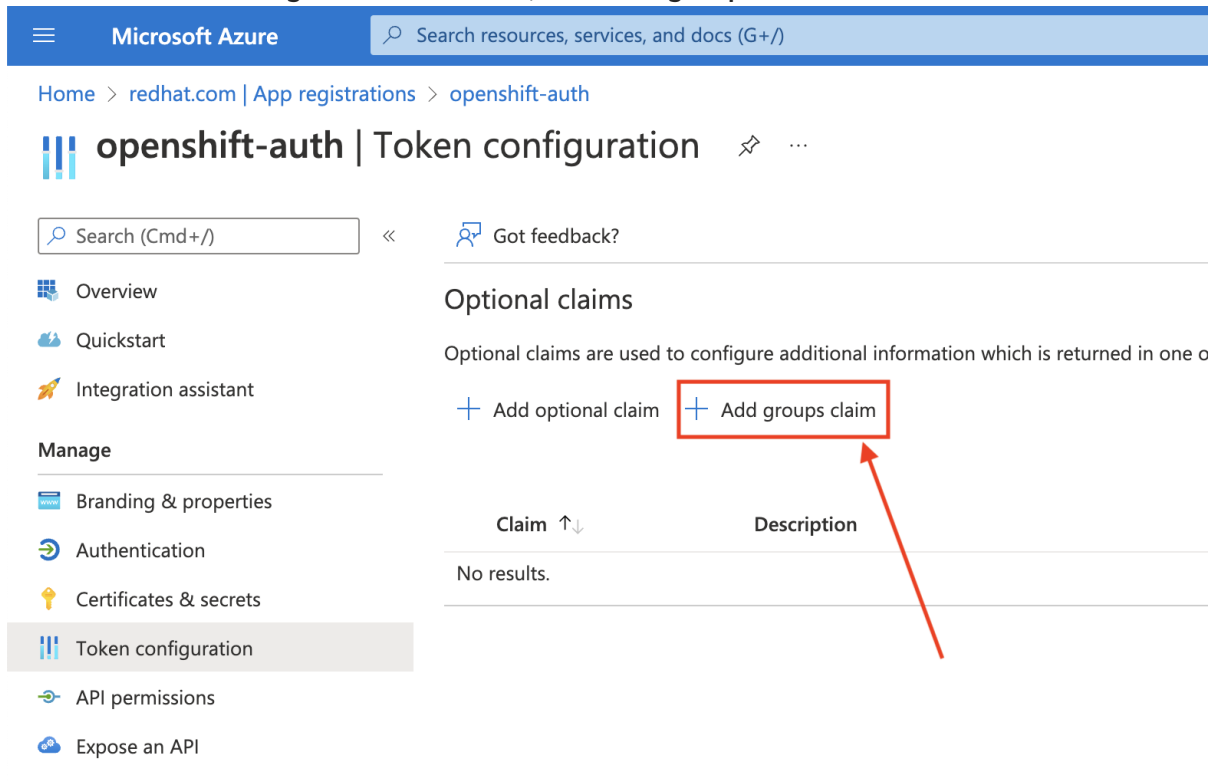


## Configuring group claims (optional)

Configure Entra ID to offer a groups claim.

### Procedure

1. From the **Token configuration** sub-blade, click **Add groups claim**.



2. To configure group claims for your Entra ID application, select **Security groups** and then click the **Add**.



## NOTE

In this example, the group claim includes all of the security groups that a user is a member of. In a real production environment, ensure that the groups that the group claim only includes groups that apply to Red Hat OpenShift Service on AWS.

The screenshot shows the Azure portal interface for configuring an application registration. The main pane is titled 'Edit groups claim'. Under the heading 'Select group types to include in Access, ID, and SAML tokens.', the 'Security groups' checkbox is checked, while 'Directory roles', 'All groups (includes distribution lists but not groups assigned to the application)', and 'Groups assigned to the application' are unchecked. Below this, there are sections for 'Customize token properties by type' with expandable options for ID, Access, and SAML. At the bottom of the configuration pane, there are 'Add' and 'Cancel' buttons. A red arrow points from the 'Add' button back to the 'Optional claims' table in the left pane, which lists 'email' and 'upn' claims.

## 8.4. CONFIGURING THE RED HAT OPENSIFT SERVICE ON AWS CLUSTER TO USE ENTRA ID AS THE IDENTITY PROVIDER

You must configure Red Hat OpenShift Service on AWS to use Entra ID as its identity provider.

Although ROSA offers the ability to configure identity providers by using OpenShift Cluster Manager, use the ROSA CLI to configure the cluster's OAuth provider to use Entra ID as its identity provider. Before configuring the identity provider, set the necessary variables for the identity provider configuration.

### Procedure

1. Create the variables by running the following command:

```
$ CLUSTER_NAME=example-cluster 1
$ IDP_NAME=AAD 2
$ APP_ID=yyyyyyyy-yyyy-yyyy-yyy-yyyyyyyyyyyyy 3
$ CLIENT_SECRET=xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxxx 4
$ TENANT_ID=zzzzzzzz-zzzz-zzzz-zzzz-zzzzzzzzzzzzz 5
```

1. Replace this with the name of your ROSA cluster.
2. Replace this value with the name you used in the OAuth callback URL that you generated earlier in this process.

- 3 Replace this with the Application (client) ID.
  - 4 Replace this with the Client Secret.
  - 5 Replace this with the Directory (tenant) ID.
2. Configure the cluster's OAuth provider by running the following command. If you enabled group claims, ensure that you use the **--group-claims groups** argument.

- If you enabled group claims, run the following command:

```
$ rosa create idp \
--cluster ${CLUSTER_NAME} \
--type openid \
--name ${IDP_NAME} \
--client-id ${APP_ID} \
--client-secret ${CLIENT_SECRET} \
--issuer-url https://login.microsoftonline.com/${TENANT_ID}/v2.0 \
--email-claims email \
--name-claims name \
--username-claims preferred_username \
--extra-scopes email,profile \
--groups-claims groups
```

- If you did not enable group claims, run the following command:

```
$ rosa create idp \
--cluster ${CLUSTER_NAME} \
--type openid \
--name ${IDP_NAME} \
--client-id ${APP_ID} \
--client-secret ${CLIENT_SECRET} \
--issuer-url https://login.microsoftonline.com/${TENANT_ID}/v2.0 \
--email-claims email \
--name-claims name \
--username-claims preferred_username \
--extra-scopes email,profile
```

After a few minutes, the cluster authentication Operator reconciles your changes, and you can log in to the cluster by using Entra ID.

## 8.5. GRANTING ADDITIONAL PERMISSIONS TO INDIVIDUAL USERS AND GROUPS

When your first log in, you might notice that you have very limited permissions. By default, Red Hat OpenShift Service on AWS only grants you the ability to create new projects, or namespaces, in the cluster. Other projects are restricted from view.

You must grant these additional abilities to individual users and groups.

### Granting additional permissions to individual users

Red Hat OpenShift Service on AWS includes a significant number of preconfigured roles, including the **cluster-admin** role that grants full access and control over the cluster.

## Procedure

- Grant a user access to the **cluster-admin** role by running the following command:

```
$ rosa grant user cluster-admin \  
  --user=<USERNAME> 1  
  --cluster=${CLUSTER_NAME}
```

- 1 Provide the Entra ID username that you want to have cluster admin permissions.

## Granting additional permissions to individual groups

If you opted to enable group claims, the cluster OAuth provider automatically creates or updates the user's group memberships by using the group ID. The cluster OAuth provider does not automatically create **RoleBindings** and **ClusterRoleBindings** for the groups that are created; you are responsible for creating those bindings by using your own processes.

To grant an automatically generated group access to the **cluster-admin** role, you must create a **ClusterRoleBinding** to the group ID.

## Procedure

- Create the **ClusterRoleBinding** by running the following command:

```
$ oc create clusterrolebinding cluster-admin-group \  
  --clusterrole=cluster-admin \  
  --group=<GROUP_ID> 1
```

- 1 Provide the Entra ID group ID that you want to have cluster admin permissions.

Now, any user in the specified group automatically receives **cluster-admin** access.

## 8.6. ADDITIONAL RESOURCES

For more information about how to use RBAC to define and apply permissions in Red Hat OpenShift Service on AWS, see [the Red Hat OpenShift Service on AWS documentation](#).

## CHAPTER 9. TUTORIAL: USING AWS SECRETS MANAGER CSI ON ROSA WITH STS

The AWS Secrets and Configuration Provider (ASCP) provides a way to expose AWS Secrets as Kubernetes storage volumes. With the ASCP, you can store and manage your secrets in Secrets Manager and then retrieve them through your workloads running on Red Hat OpenShift Service on AWS (ROSA).

### 9.1. PREREQUISITES

Ensure that you have the following resources and tools before starting this process:

- A ROSA cluster deployed with STS
- Helm 3
- **aws** CLI
- **oc** CLI
- **jq** CLI

#### Additional environment requirements

1. Log in to your ROSA cluster by running the following command:

```
$ oc login --token=<your-token> --server=<your-server-url>
```

You can find your login token by accessing your cluster in [pull secret from Red Hat OpenShift Cluster Manager](#).

2. Validate that your cluster has STS by running the following command:

```
$ oc get authentication.config.openshift.io cluster -o json \
| jq .spec.serviceAccountIssuer
```

#### Example output

```
"https://xxxxx.cloudfront.net/xxxxx"
```

If your output is different, do not proceed. See [Red Hat documentation on creating an STS cluster](#) before continuing this process.

3. Set the **SecurityContextConstraints** permission to allow the CSI driver to run by running the following command:

```
$ oc new-project csi-secrets-store
$ oc adm policy add-scc-to-user privileged \
system:serviceaccount:csi-secrets-store:secrets-store-csi-driver
$ oc adm policy add-scc-to-user privileged \
system:serviceaccount:csi-secrets-store:csi-secrets-store-provider-aws
```

4. Create environment variables to use later in this process by running the following command:

■

```
$ export REGION=$(oc get infrastructure cluster -o=jsonpath="{.status.platformStatus.aws.region}")
$ export OIDC_ENDPOINT=$(oc get authentication.config.openshift.io cluster \
  -o jsonpath='{.spec.serviceAccountIssuer}' | sed 's|^https://|')
$ export AWS_ACCOUNT_ID=`aws sts get-caller-identity --query Account --output text`
$ export AWS_PAGER=""
```

## 9.2. DEPLOYING THE AWS SECRETS AND CONFIGURATION PROVIDER

1. Use Helm to register the secrets store CSI driver by running the following command:

```
$ helm repo add secrets-store-csi-driver \
  https://kubernetes-sigs.github.io/secrets-store-csi-driver/charts
```

2. Update your Helm repositories by running the following command:

```
$ helm repo update
```

3. Install the secrets store CSI driver by running the following command:

```
$ helm upgrade --install -n csi-secrets-store \
  csi-secrets-store-driver secrets-store-csi-driver/secrets-store-csi-driver
```

4. Deploy the AWS provider by running the following command:

```
$ oc -n csi-secrets-store apply -f \
  https://raw.githubusercontent.com/rh-mobb/documentation/main/content/misc/secrets-
  store-csi/aws-provider-installer.yaml
```

5. Check that both Daemonsets are running by running the following command:

```
$ oc -n csi-secrets-store get ds \
  csi-secrets-store-provider-aws \
  csi-secrets-store-driver-secrets-store-csi-driver
```

6. Label the Secrets Store CSI Driver to allow use with the restricted pod security profile by running the following command:

```
$ oc label csidriver.storage.k8s.io/secrets-store.csi.k8s.io security.openshift.io/csi-ephemeral-
  volume-profile=restricted
```

## 9.3. CREATING A SECRET AND IAM ACCESS POLICIES

1. Create a secret in Secrets Manager by running the following command:

```
$ SECRET_ARN=$(aws --region "$REGION" secretsmanager create-secret \
  --name MySecret --secret-string \
  '{"username":"shadowman", "password":"hunter2"}' \
  --query ARN --output text)
$ echo $SECRET_ARN
```



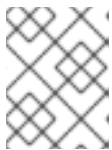
2. Create an IAM Access Policy document by running the following command:

```
$ cat << EOF > policy.json
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue",
      "secretsmanager:DescribeSecret"
    ],
    "Resource": ["$SECRET_ARN"]
  }]
}
EOF
```

3. Create an IAM Access Policy by running the following command:

```
$ POLICY_ARN=$(aws --region "$REGION" --query Policy.Arn \
--output text iam create-policy \
--policy-name openshift-access-to-mysecret-policy \
--policy-document file://policy.json)
$ echo $POLICY_ARN
```

4. Create an IAM Role trust policy document by running the following command:



#### NOTE

The trust policy is locked down to the default service account of a namespace you create later in this process.

```
$ cat <<EOF > trust-policy.json
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "${OIDC_ENDPOINT}.sub": ["system:serviceaccount:my-application:default"]
        }
      },
      "Principal": {
        "Federated": "arn:aws:iam:::$AWS_ACCOUNT_ID:oidc-provider:${OIDC_ENDPOINT}"
      },
      "Action": "sts:AssumeRoleWithWebIdentity"
    }
  ]
}
EOF
```

5. Create an IAM role by running the following command:

```
$ ROLE_ARN=$(aws iam create-role --role-name openshift-access-to-mysecret \
```

```
--assume-role-policy-document file://trust-policy.json \  
--query Role.Arn --output text)  
$ echo $ROLE_ARN
```

6. Attach the role to the policy by running the following command:

```
$ aws iam attach-role-policy --role-name openshift-access-to-mysecret \  
--policy-arn $POLICY_ARN
```

## 9.4. CREATE AN APPLICATION TO USE THIS SECRET

1. Create an OpenShift project by running the following command:

```
$ oc new-project my-application
```

2. Annotate the default service account to use the STS Role by running the following command:

```
$ oc annotate -n my-application serviceaccount default \  
eks.amazonaws.com/role-arn=$ROLE_ARN
```

3. Create a secret provider class to access our secret by running the following command:

```
$ cat << EOF | oc apply -f -  
apiVersion: secrets-store.csi.x-k8s.io/v1  
kind: SecretProviderClass  
metadata:  
  name: my-application-aws-secrets  
spec:  
  provider: aws  
  parameters:  
    objects: |  
    - objectName: "MySecret"  
      objectType: "secretsmanager"  
EOF
```

4. Create a Deployment by using our secret in the following command:

```
$ cat << EOF | oc apply -f -  
apiVersion: v1  
kind: Pod  
metadata:  
  name: my-application  
  labels:  
    app: my-application  
spec:  
  volumes:  
    - name: secrets-store-inline  
      csi:  
        driver: secrets-store.csi.k8s.io  
        readOnly: true  
        volumeAttributes:  
          secretProviderClass: "my-application-aws-secrets"  
  containers:
```

```

- name: my-application-deployment
  image: k8s.gcr.io/e2e-test-images/busybox:1.29
  command:
    - "/bin/sleep"
    - "10000"
  volumeMounts:
    - name: secrets-store-inline
      mountPath: "/mnt/secrets-store"
      readOnly: true
EOF

```

5. Verify the Pod has the secret mounted by running the following command:

```
$ oc exec -it my-application -- cat /mnt/secrets-store/MySecret
```

## 9.5. CLEAN UP

1. Delete the application by running the following command:

```
$ oc delete project my-application
```

2. Delete the secrets store csi driver by running the following command:

```
$ helm delete -n csi-secrets-store csi-secrets-store-driver
```

3. Delete Security Context Constraints by running the following command:

```

$ oc adm policy remove-scc-from-user privileged \
  system:serviceaccount:csi-secrets-store:secrets-store-csi-driver
$ oc adm policy remove-scc-from-user privileged \
  system:serviceaccount:csi-secrets-store:csi-secrets-store-provider-aws

```

4. Delete the AWS provider by running the following command:

```

$ oc -n csi-secrets-store delete -f \
  https://raw.githubusercontent.com/rh-mobb/documentation/main/content/misc/secrets-store-
  csi/aws-provider-installer.yaml

```

5. Delete AWS Roles and Policies by running the following command:

```

$ aws iam detach-role-policy --role-name openshift-access-to-mysecret \
  --policy-arn $POLICY_ARN
$ aws iam delete-role --role-name openshift-access-to-mysecret
$ aws iam delete-policy --policy-arn $POLICY_ARN

```

6. Delete the Secrets Manager secret by running the following command:

```
$ aws secretsmanager --region $REGION delete-secret --secret-id $SECRET_ARN
```

## CHAPTER 10. TUTORIAL: USING AWS CONTROLLERS FOR KUBERNETES ON ROSA

[AWS Controllers for Kubernetes](#) (ACK) lets you define and use AWS service resources directly from Red Hat OpenShift Service on AWS (ROSA). With ACK, you can take advantage of AWS-managed services for your applications without needing to define resources outside of the cluster or run services that provide supporting capabilities such as databases or message queues within the cluster.

You can install various ACK Operators directly from OperatorHub. This makes it easy to get started and use the Operators with your applications. This controller is a component of the AWS Controller for Kubernetes project, which is currently in developer preview.

Use this tutorial to deploy the ACK S3 Operator. You can also adapt it for any other ACK Operator in the OperatorHub of your cluster.

### 10.1. PREREQUISITES

- A ROSA cluster
- A user account with **cluster-admin** privileges
- The OpenShift CLI (**oc**)
- The Amazon Web Services (AWS) CLI (**aws**)

### 10.2. SETTING UP YOUR ENVIRONMENT

1. Configure the following environment variables, changing the cluster name to suit your cluster:

```
$ export CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}" | sed 's/-[a-z0-9]{5}$//')
$ export REGION=$(rosa describe cluster -c ${ROSA_CLUSTER_NAME} --output json | jq -r .region.id)
$ export OIDC_ENDPOINT=$(oc get authentication.config.openshift.io cluster -o json | jq -r .spec.serviceAccountIssuer | sed 's|^https://|')
$ export AWS_ACCOUNT_ID=`aws sts get-caller-identity --query Account --output text`
$ export ACK_SERVICE=s3
$ export ACK_SERVICE_ACCOUNT=ack-${ACK_SERVICE}-controller
$ export POLICY_ARN=arn:aws:iam::aws:policy/AmazonS3FullAccess
$ export AWS_PAGER=""
$ export SCRATCH="/tmp/${ROSA_CLUSTER_NAME}/ack"
$ mkdir -p ${SCRATCH}
```

2. Ensure all fields output correctly before moving to the next section:

```
$ echo "Cluster: ${ROSA_CLUSTER_NAME}, Region: ${REGION}, OIDC Endpoint:
${OIDC_ENDPOINT}, AWS Account ID: ${AWS_ACCOUNT_ID}"
```

### 10.3. PREPARING YOUR AWS ACCOUNT

1. Create an AWS Identity Access Management (IAM) trust policy for the ACK Operator:

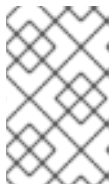
```
$ cat <<EOF > "${SCRATCH}/trust-policy.json"
```

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "${OIDC_ENDPOINT}:sub": "system:serviceaccount:ack-
system:${ACK_SERVICE_ACCOUNT}"
        }
      },
      "Principal": {
        "Federated": "arn:aws:iam::${AWS_ACCOUNT_ID}:oidc-provider/${OIDC_ENDPOINT}"
      },
      "Action": "sts:AssumeRoleWithWebIdentity"
    }
  ]
}
EOF

```

2. Create an AWS IAM role for the ACK Operator to assume with the **AmazonS3FullAccess** policy attached:



#### NOTE

You can find the recommended policy in each project's GitHub repository, for example <https://github.com/aws-controllers-k8s/s3-controller/blob/main/config/iam/recommended-policy-arn>.

```

$ ROLE_ARN=$(aws iam create-role --role-name "ack-${ACK_SERVICE}-controller" \
--assume-role-policy-document "file://${SCRATCH}/trust-policy.json" \
--query Role.Arn --output text)
$ echo $ROLE_ARN

$ aws iam attach-role-policy --role-name "ack-${ACK_SERVICE}-controller" \
--policy-arn ${POLICY_ARN}

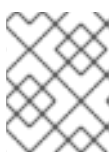
```

## 10.4. INSTALLING THE ACK S3 CONTROLLER

1. Create a project to install the ACK S3 Operator into:

```
$ oc new-project ack-system
```

2. Create a file with the ACK S3 Operator configuration:



#### NOTE

**ACK\_WATCH\_NAMESPACE** is purposefully left blank so the controller can properly watch all namespaces in the cluster.

```

$ cat <<EOF > "${SCRATCH}/config.txt"
ACK_ENABLE_DEVELOPMENT_LOGGING=true

```

```

ACK_LOG_LEVEL=debug
ACK_WATCH_NAMESPACE=
AWS_REGION=${REGION}
AWS_ENDPOINT_URL=
ACK_RESOURCE_TAGS=${CLUSTER_NAME}
ENABLE_LEADER_ELECTION=true
LEADER_ELECTION_NAMESPACE=
EOF

```

- Use the file from the previous step to create a ConfigMap:

```

$ oc -n ack-system create configmap \
  --from-env-file=${SCRATCH}/config.txt ack-${ACK_SERVICE}-user-config

```

- Install the ACK S3 Operator from OperatorHub:

```

$ cat << EOF | oc apply -f -
apiVersion: operators.coreos.com/v1
kind: OperatorGroup
metadata:
  name: ack-${ACK_SERVICE}-controller
  namespace: ack-system
spec:
  upgradeStrategy: Default
---
apiVersion: operators.coreos.com/v1alpha1
kind: Subscription
metadata:
  name: ack-${ACK_SERVICE}-controller
  namespace: ack-system
spec:
  channel: alpha
  installPlanApproval: Automatic
  name: ack-${ACK_SERVICE}-controller
  source: community-operators
  sourceNamespace: openshift-marketplace
EOF

```

- Annotate the ACK S3 Operator service account with the AWS IAM role to assume and restart the deployment:

```

$ oc -n ack-system annotate serviceaccount ${ACK_SERVICE_ACCOUNT} \
  eks.amazonaws.com/role-arn=${ROLE_ARN} && \
  oc -n ack-system rollout restart deployment ack-${ACK_SERVICE}-controller

```

- Verify that the ACK S3 Operator is running:

```

$ oc -n ack-system get pods

```

### Example output

```

NAME                                READY STATUS RESTARTS AGE
ack-s3-controller-585f6775db-s4lfz  1/1   Running  0      51s

```

## 10.5. VALIDATING THE DEPLOYMENT

1. Deploy an S3 bucket resource:

```
$ cat << EOF | oc apply -f -
apiVersion: s3.services.k8s.aws/v1alpha1
kind: Bucket
metadata:
  name: ${CLUSTER-NAME}-bucket
  namespace: ack-system
spec:
  name: ${CLUSTER-NAME}-bucket
EOF
```

2. Verify the S3 bucket was created in AWS:

```
$ aws s3 ls | grep ${CLUSTER_NAME}-bucket
```

### Example output

```
2023-10-04 14:51:45 mrmc-test-maz-bucket
```

## 10.6. CLEANING UP

1. Delete the S3 bucket resource:

```
$ oc -n ack-system delete bucket.s3.services.k8s.aws/${CLUSTER-NAME}-bucket
```

2. Delete the ACK S3 Operator and the AWS IAM roles:

```
$ oc -n ack-system delete subscription ack-${ACK_SERVICE}-controller
$ aws iam detach-role-policy \
  --role-name "ack-${ACK_SERVICE}-controller" \
  --policy-arn ${POLICY_ARN}
$ aws iam delete-role \
  --role-name "ack-${ACK_SERVICE}-controller"
```

3. Delete the **ack-system** project:

```
$ oc delete project ack-system
```

# CHAPTER 11. TUTORIAL: DEPLOYING THE EXTERNAL DNS OPERATOR ON ROSA

The External DNS Operator deploys and manages **ExternalDNS** to provide the name resolution for services and routes from the external DNS provider, like Amazon Route 53, to Red Hat OpenShift Service on AWS (ROSA) clusters. In this tutorial, we will deploy and configure the External DNS Operator with a secondary ingress controller to manage DNS records in Amazon Route 53.

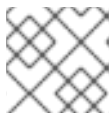


## IMPORTANT

The **External DNS** Operator does not support STS using IAM Roles for Service Accounts (IRSA) and uses long-lived Identity Access Management (IAM) credentials instead. This tutorial will be updated when the Operator supports STS.

## 11.1. PREREQUISITES

- A ROSA Classic cluster



## NOTE

ROSA with HCP is not supported at this time.

- A user account with **cluster-admin** privileges
- The OpenShift CLI (**oc**)
- The Amazon Web Services (AWS) CLI (**aws**)
- A unique domain, such as **apps.example.com**
- An Amazon Route 53 public hosted zone for the above domain

## 11.2. SETTING UP YOUR ENVIRONMENT

1. Configure the following environment variables:

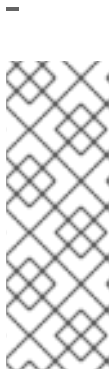
```
$ export DOMAIN=<apps.example.com> 1
$ export AWS_PAGER=""
$ export CLUSTER=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}"
| sed 's/-[a-z0-9]{5}$//')
$ export REGION=$(oc get infrastructure cluster -o=jsonpath="
{.status.platformStatus.aws.region}")
$ export AWS_ACCOUNT_ID=$(aws sts get-caller-identity --query Account --output text)
$ export SCRATCH="/tmp/${CLUSTER}/external-dns"
$ mkdir -p ${SCRATCH}
```

- 1** Replace with the custom domain you want to use for the **IngressController**.

2. Ensure all fields output correctly before moving to the next section:

```
$ echo "Cluster: ${CLUSTER}, Region: ${REGION}, AWS Account ID:
${AWS_ACCOUNT_ID}"
```



**NOTE**

The "Cluster" output from the previous command may be the name of your cluster, the internal ID of your cluster, or the cluster's domain prefix. If you prefer to use another identifier, you can manually set this value by running the following command:

```
$ export CLUSTER=my-custom-value
```

**11.3. SECONDARY INGRESS CONTROLLER SETUP**

Use the following procedure to deploy a secondary ingress controller using a custom domain.

**Prerequisites**

- A unique domain, such as **apps.example.com**
- A wildcard or SAN TLS certificate configured with the custom domain selected above (**CN=\*.apps.example.com**)

**Procedure**

1. Create a new TLS secret from a private key and a public certificate, where **fullchain.pem** is your full wildcard certificate chain (including any intermediaries) and **privkey.pem** is your wildcard certificate's private key:

```
$ oc -n openshift-ingress create secret tls external-dns-tls --cert=fullchain.pem --key=privkey.pem
```

2. Create a new **IngressController** resource:

```
$ cat << EOF | oc apply -f -
apiVersion: operator.openshift.io/v1
kind: IngressController
metadata:
  name: external-dns-ingress
  namespace: openshift-ingress-operator
spec:
  domain: ${DOMAIN}
  defaultCertificate:
    name: external-dns-tls
  endpointPublishingStrategy:
    loadBalancer:
      dnsManagementPolicy: Unmanaged
      providerParameters:
        aws:
          type: NLB
          type: AWS
        scope: External
      type: LoadBalancerService
EOF
```

**WARNING**

This **IngressController** example will create an internet accessible Network Load Balancer (NLB) in your AWS account. To provision an internal NLB instead, set the **.spec.endpointPublishingStrategy.loadBalancer.scope** parameter to **Internal** before creating the **IngressController** resource.

3. Verify that your custom domain IngressController has successfully created an external load balancer:

```
$ oc -n openshift-ingress get service/router-external-dns-ingress
```

**Example output**

```
NAME                                TYPE           CLUSTER-IP   EXTERNAL-IP
PORT(S)                            AGE
router-external-dns-ingress         LoadBalancer   172.30.71.250
a4838bb991c6748439134ab89f132a43-aeae124077b50c01.elb.us-east-1.amazonaws.com
80:32227/TCP,443:30310/TCP         43s
```

## 11.4. PREPARING YOUR AWS ACCOUNT

1. Retrieve the Amazon Route 53 public hosted zone ID:

```
$ export ZONE_ID=$(aws route53 list-hosted-zones-by-name --output json \
--dns-name "${DOMAIN}." --query 'HostedZones[0].Id --out text | sed 's/\//hostedzone\//')
```

2. Prepare a document with the necessary DNS changes to enable DNS resolution for the canonical domain of the Ingress Controller:

```
$ NLB_HOST=$(oc -n openshift-ingress get service/router-external-dns-ingress -ojsonpath="
{.status.loadBalancer.ingress[0].hostname}")
$ cat << EOF > "${SCRATCH}/create-cname.json"
{
  "Comment": "Add CNAME to ingress controller canonical domain",
  "Changes": [{
    "Action": "CREATE",
    "ResourceRecordSet": {
      "Name": "router-external-dns-ingress.${DOMAIN}",
      "Type": "CNAME",
      "TTL": 30,
      "ResourceRecords": [{
        "Value": "${NLB_HOST}"
      }]
    }
  ]
}
EOF
```

The External DNS Operator uses this canonical domain as the target for CNAME records.

3. Submit your changes to Amazon Route 53 for propagation:

```
aws route53 change-resource-record-sets \
  --hosted-zone-id ${ZONE_ID} \
  --change-batch file://${SCRATCH}/create-cname.json
```

4. Create an AWS IAM Policy document that allows the **External DNS** Operator to update *only* the custom domain public hosted zone:

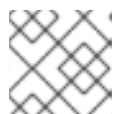
```
$ cat << EOF > "${SCRATCH}/external-dns-policy.json"
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "route53:ChangeResourceRecordSets"
      ],
      "Resource": [
        "arn:aws:route53:::hostedzone/${ZONE_ID}"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "route53:ListHostedZones",
        "route53:ListResourceRecordSets"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
EOF
```

5. Create an AWS IAM user:

```
$ aws iam create-user --user-name "${CLUSTER}-external-dns-operator"
```

6. Attach the policy:

```
$ aws iam attach-user-policy --user-name "${CLUSTER}-external-dns-operator" --policy-arn
$POLICY_ARN
```



#### NOTE

This will be changed to STS using IRSA in the future.

7. Create AWS keys for the IAM user:

```
$ SECRET_ACCESS_KEY=$(aws iam create-access-key --user-name "${CLUSTER}-external-dns-operator")
```

8. Create static credentials:

```
$ cat << EOF > "${SCRATCH}/credentials"
[default]
aws_access_key_id = $(echo $SECRET_ACCESS_KEY | jq -r '.AccessKey.AccessKeyId')
aws_secret_access_key = $(echo $SECRET_ACCESS_KEY | jq -r
'.AccessKey.SecretAccessKey')
EOF
```

## 11.5. INSTALLING THE EXTERNAL DNS OPERATOR

1. Create a new project:

```
$ oc new-project external-dns-operator
```

2. Install the **External DNS** Operator from OperatorHub:

```
$ cat << EOF | oc apply -f -
apiVersion: operators.coreos.com/v1
kind: OperatorGroup
metadata:
  name: external-dns-group
  namespace: external-dns-operator
spec:
  targetNamespaces:
  - external-dns-operator
---
apiVersion: operators.coreos.com/v1alpha1
kind: Subscription
metadata:
  name: external-dns-operator
  namespace: external-dns-operator
spec:
  channel: stable-v1.1
  installPlanApproval: Automatic
  name: external-dns-operator
  source: redhat-operators
  sourceNamespace: openshift-marketplace
EOF
```

3. Wait until the **External DNS** Operator is running:

```
$ oc rollout status deploy external-dns-operator --timeout=300s
```

4. Create a secret from the AWS IAM user credentials:

```
$ oc -n external-dns-operator create secret generic external-dns \
--from-file "${SCRATCH}/credentials"
```

5. Deploy the **ExternalDNS** controller:

```
$ cat << EOF | oc apply -f -
apiVersion: externaldns.olm.openshift.io/v1beta1
kind: ExternalDNS
metadata:
  name: ${DOMAIN}
spec:
  domains:
    - filterType: Include
      matchType: Exact
      name: ${DOMAIN}
  provider:
    aws:
      credentials:
        name: external-dns
      type: AWS
    source:
      openshiftRouteOptions:
        routerName: external-dns-ingress
        type: OpenShiftRoute
    zones:
      - ${ZONE_ID}
EOF
```

6. Wait until the controller is running:

```
$ oc rollout status deploy external-dns-${DOMAIN} --timeout=300s
```

## 11.6. DEPLOYING A SAMPLE APPLICATION

Now that the **ExternalDNS** controller is running, you can deploy a sample application to confirm that the custom domain is configured and trusted when you expose a new route.

1. Create a new project for your sample application:

```
$ oc new-project hello-world
```

2. Deploy a hello world application:

```
$ oc new-app -n hello-world --image=docker.io/openshift/hello-openshift
```

3. Create a route for the application specifying your custom domain name:

```
$ oc -n hello-world create route edge --service=hello-openshift hello-openshift-tls \
--hostname hello-openshift.${DOMAIN}
```

4. Check if the DNS record was created automatically by ExternalDNS:



### NOTE

It can take a few minutes for the record to appear in Amazon Route 53.

```
$ aws route53 list-resource-record-sets --hosted-zone-id ${ZONE_ID} \  
  --query "ResourceRecordSets[?Type == 'CNAME']" | grep hello-openshift
```

- Optional: You can also view the TXT records that indicate they were created by ExternalDNS:

```
$ aws route53 list-resource-record-sets --hosted-zone-id ${ZONE_ID} \  
  --query "ResourceRecordSets[?Type == 'TXT']" | grep ${DOMAIN}
```

- Curl the newly created DNS record to your sample application to verify the hello world application is accessible:

```
$ curl https://hello-openshift.${DOMAIN}
```

### Example output

```
Hello OpenShift!
```

## CHAPTER 12. TUTORIAL: DYNAMICALLY ISSUING CERTIFICATES USING THE CERT-MANAGER OPERATOR ON ROSA

While wildcard certificates provide simplicity by securing all first-level subdomains of a given domain with a single certificate, other use cases can require the use of individual certificates per domain.

Learn how to use the [cert-manager Operator for Red Hat OpenShift](#) and [Let's Encrypt](#) to dynamically issue certificates for routes created using a custom domain.

### 12.1. PREREQUISITES

- A ROSA cluster (HCP or Classic)
- A user account with **cluster-admin** privileges
- The OpenShift CLI (**oc**)
- The Amazon Web Services (AWS) CLI (**aws**)
- A unique domain, such as **\*.apps.example.com**
- An Amazon Route 53 public hosted zone for the above domain

### 12.2. SETTING UP YOUR ENVIRONMENT

1. Configure the following environment variables:

```
$ export DOMAIN=apps.example.com 1
$ export EMAIL=email@example.com 2
$ export AWS_PAGER=""
$ export CLUSTER=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}"
| sed 's/-[a-z0-9]{5}$//')
$ export OIDC_ENDPOINT=$(oc get authentication.config.openshift.io cluster -o json | jq -r
.spec.serviceAccountIssuer | sed 's|^https://|')
$ export REGION=$(oc get infrastructure cluster -o=jsonpath="{
.status.platformStatus.aws.region}")
$ export AWS_ACCOUNT_ID=$(aws sts get-caller-identity --query Account --output text)
$ export SCRATCH="/tmp/${CLUSTER}/dynamic-certs"
$ mkdir -p ${SCRATCH}
```

- 1** Replace with the custom domain you want to use for the **IngressController**.
- 2** Replace with the e-mail you want Let's Encrypt to use to send notifications about your certificates.

2. Ensure all fields output correctly before moving to the next section:

```
$ echo "Cluster: ${CLUSTER}, Region: ${REGION}, OIDC Endpoint: ${OIDC_ENDPOINT},
AWS Account ID: ${AWS_ACCOUNT_ID}"
```

**NOTE**

The "Cluster" output from the previous command may be the name of your cluster, the internal ID of your cluster, or the cluster's domain prefix. If you prefer to use another identifier, you can manually set this value by running the following command:

```
$ export CLUSTER=my-custom-value
```

## 12.3. PREPARING YOUR AWS ACCOUNT

When cert-manager requests a certificate from Let's Encrypt (or another ACME certificate issuer), Let's Encrypt servers validate that you control the domain name in that certificate using *challenges*. For this tutorial, you are using a [DNS-01 challenge](#) that proves that you control the DNS for your domain name by putting a specific value in a TXT record under that domain name. This is all done automatically by cert-manager. To allow cert-manager permission to modify the Amazon Route 53 public hosted zone for your domain, you need to create an Identity Access Management (IAM) role with specific policy permissions and a trust relationship to allow access to the pod.

The public hosted zone that is used in this tutorial is in the same AWS account as the ROSA cluster. If your public hosted zone is in a different account, a few additional steps for [Cross Account Access](#) are required.

1. Retrieve the Amazon Route 53 public hosted zone ID:

**NOTE**

This command looks for a public hosted zone that matches the custom domain you specified earlier as the **DOMAIN** environment variable. You can manually specify the Amazon Route 53 public hosted zone by running **export ZONE\_ID=<zone\_ID>**, replacing **<zone\_ID>** with your specific Amazon Route 53 public hosted zone ID.

```
$ export ZONE_ID=$(aws route53 list-hosted-zones-by-name --output json \
--dns-name "${DOMAIN}." --query 'HostedZones[0].Id --out text | sed 's/\//hostedzone\//')
```

2. Create an AWS IAM policy document for the cert-manager Operator that provides the ability to update *only* the specified public hosted zone:

```
$ cat <<EOF > "${SCRATCH}/cert-manager-policy.json"
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "route53:GetChange",
      "Resource": "arn:aws:route53:::change/*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "route53:ChangeResourceRecordSets",
        "route53:ListResourceRecordSets"
      ]
    }
  ]
}
```



```

    "Resource": "arn:aws:route53:::hostedzone/${ZONE_ID}"
  },
  {
    "Effect": "Allow",
    "Action": "route53:ListHostedZonesByName",
    "Resource": "*"
  }
]
}
EOF

```

3. Create the IAM policy using the file you created in the previous step:

```

$ POLICY_ARN=$(aws iam create-policy --policy-name "${CLUSTER}-cert-manager-policy" \
--policy-document file://${SCRATCH}/cert-manager-policy.json \
--query 'Policy.Arn' --output text)

```

4. Create an AWS IAM trust policy for the cert-manager Operator:

```

$ cat <<EOF > "${SCRATCH}/trust-policy.json"
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "${OIDC_ENDPOINT}:sub": "system:serviceaccount:cert-manager:cert-manager"
        }
      },
      "Principal": {
        "Federated": "arn:aws:iam::${AWS_ACCOUNT_ID}:oidc-provider/${OIDC_ENDPOINT}"
      },
      "Action": "sts:AssumeRoleWithWebIdentity"
    }
  ]
}
EOF

```

5. Create an IAM role for the cert-manager Operator using the trust policy you created in the previous step:

```

$ ROLE_ARN=$(aws iam create-role --role-name "${CLUSTER}-cert-manager-operator" \
--assume-role-policy-document "file://${SCRATCH}/trust-policy.json" \
--query Role.Arn --output text)

```

6. Attach the permissions policy to the role:

```

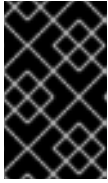
$ aws iam attach-role-policy --role-name "${CLUSTER}-cert-manager-operator" \
--policy-arn ${POLICY_ARN}

```

## 12.4. INSTALLING THE CERT-MANAGER OPERATOR

1. Create a project to install the cert-manager Operator into:

```
$ oc new-project cert-manager-operator
```



### IMPORTANT

Do not attempt to use more than one cert-manager Operator in your cluster. If you have a community cert-manager Operator installed in your cluster, you must uninstall it before installing the cert-manager Operator for Red Hat OpenShift.

2. Install the cert-manager Operator for Red Hat OpenShift:

```
$ cat << EOF | oc apply -f -
apiVersion: operators.coreos.com/v1
kind: OperatorGroup
metadata:
  name: openshift-cert-manager-operator-group
  namespace: cert-manager-operator
spec:
  targetNamespaces:
  - cert-manager-operator
---
apiVersion: operators.coreos.com/v1alpha1
kind: Subscription
metadata:
  name: openshift-cert-manager-operator
  namespace: cert-manager-operator
spec:
  channel: stable-v1
  installPlanApproval: Automatic
  name: openshift-cert-manager-operator
  source: redhat-operators
  sourceNamespace: openshift-marketplace
EOF
```



### NOTE

It takes a few minutes for this Operator to install and complete its set up.

3. Verify that the cert-manager Operator is running:

```
$ oc -n cert-manager-operator get pods
```

#### Example output

| NAME  | READY | STATUS  | RESTARTS | AGE |
|---|-------|---------|----------|-----|
| cert-manager-operator-controller-manager-84b8799db5-gv8mx | 2/2   | Running | 0        | 12s |

4. Annotate the service account used by the cert-manager pods with the AWS IAM role you created earlier:

```
$ oc -n cert-manager annotate serviceaccount cert-manager eks.amazonaws.com/role-arn=${ROLE_ARN}
```

- Restart the existing cert-manager controller pod by running the following command:

```
$ oc -n cert-manager delete pods -l app.kubernetes.io/name=cert-manager
```

- Patch the Operator's configuration to use external nameservers to prevent DNS-01 challenge resolution issues:

```
$ oc patch certmanager.operator.openshift.io/cluster --type merge \
  -p '{"spec":{"controllerConfig":{"overrideArgs":["--dns01-recursive-nameservers-only","--dns01-recursive-nameservers=1.1.1.1:53"]}}}'
```

- Create a **ClusterIssuer** resource to use Let's Encrypt by running the following command:

```
$ cat << EOF | oc apply -f -
apiVersion: cert-manager.io/v1
kind: ClusterIssuer
metadata:
  name: letsencrypt-production
spec:
  acme:
    server: https://acme-v02.api.letsencrypt.org/directory
    email: ${EMAIL}
    # This key doesn't exist, cert-manager creates it
    privateKeySecretRef:
      name: prod-letsencrypt-issuer-account-key
    solvers:
      - dns01:
          route53:
            hostedZoneID: ${ZONE_ID}
            region: ${REGION}
            secretAccessKeySecretRef:
              name: "
```

EOF

- Verify the **ClusterIssuer** resource is ready:

```
$ oc get clusterissuer.cert-manager.io/letsencrypt-production
```

### Example output

```
NAME                READY  AGE
letsencrypt-production  True  47s
```

## 12.5. CREATING A CUSTOM DOMAIN INGRESS CONTROLLER

- Create and configure a certificate resource to provision a certificate for the custom domain Ingress Controller:



### NOTE

The following example uses a single domain certificate. SAN and wildcard certificates are also supported.

```
$ cat << EOF | oc apply -f -
apiVersion: cert-manager.io/v1
kind: Certificate
metadata:
  name: custom-domain-ingress-cert
  namespace: openshift-ingress
spec:
  secretName: custom-domain-ingress-cert-tls
  issuerRef:
    name: letsencrypt-production
    kind: ClusterIssuer
  commonName: "${DOMAIN}"
  dnsNames:
  - "${DOMAIN}"
EOF
```

2. Verify the certificate has been issued:



### NOTE

It takes a few minutes for this certificate to be issued by Let's Encrypt. If it takes longer than 5 minutes, run **oc -n openshift-ingress describe certificate.cert-manager.io/custom-domain-ingress-cert** to see any issues reported by cert-manager.

```
$ oc -n openshift-ingress get certificate.cert-manager.io/custom-domain-ingress-cert
```

### Example output

```
NAME                READY SECRET                AGE
custom-domain-ingress-cert True  custom-domain-ingress-cert-tls  9m53s
```

3. Create a new **IngressController** resource:

```
$ cat << EOF | oc apply -f -
apiVersion: operator.openshift.io/v1
kind: IngressController
metadata:
  name: custom-domain-ingress
  namespace: openshift-ingress-operator
spec:
  domain: ${DOMAIN}
  defaultCertificate:
    name: custom-domain-ingress-cert-tls
  endpointPublishingStrategy:
    loadBalancer:
      dnsManagementPolicy: Unmanaged
      providerParameters:
        aws:
          type: NLB
          type: AWS
```

```
scope: External
type: LoadBalancerService
EOF
```



### WARNING

This **IngressController** example will create an internet accessible Network Load Balancer (NLB) in your AWS account. To provision an internal NLB instead, set the **.spec.endpointPublishingStrategy.loadBalancer.scope** parameter to **Internal** before creating the **IngressController** resource.

- Verify that your custom domain IngressController has successfully created an external load balancer:

```
$ oc -n openshift-ingress get service/router-custom-domain-ingress
```

### Example output

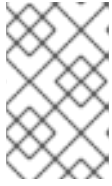
```
NAME                                TYPE          CLUSTER-IP      EXTERNAL-IP
PORT(S)                            AGE
router-custom-domain-ingress        LoadBalancer  172.30.174.34
a309962c3bd6e42c08cadb9202eca683-1f5bbb64a1f1ec65.elb.us-east-1.amazonaws.com
80:31342/TCP,443:31821/TCP          7m28s
```

- Prepare a document with the necessary DNS changes to enable DNS resolution for your custom domain Ingress Controller:

```
$ INGRESS=$(oc -n openshift-ingress get service/router-custom-domain-ingress -
ojsonpath="{.status.loadBalancer.ingress[0].hostname}")
$ cat << EOF > "${SCRATCH}/create-cname.json"
{
  "Comment": "Add CNAME to custom domain endpoint",
  "Changes": [{
    "Action": "CREATE",
    "ResourceRecordSet": {
      "Name": "*.${DOMAIN}",
      "Type": "CNAME",
      "TTL": 30,
      "ResourceRecords": [{
        "Value": "${INGRESS}"
      }]
    }
  ]
}
EOF
```

- Submit your changes to Amazon Route 53 for propagation:

```
$ aws route53 change-resource-record-sets \
  --hosted-zone-id ${ZONE_ID} \
  --change-batch file://${SCRATCH}/create-cname.json
```



## NOTE

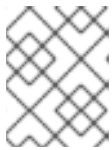
While the wildcard CNAME record avoids the need to create a new record for every new application you deploy using the custom domain Ingress Controller, the certificate that each of these applications use **is not** a wildcard certificate.

## 12.6. CONFIGURING DYNAMIC CERTIFICATES FOR CUSTOM DOMAIN ROUTES

Now you can expose cluster applications on any first-level subdomains of the specified domain, but the connection will not be secured with a TLS certificate that matches the domain of the application. To ensure these cluster applications have valid certificates for each domain name, configure cert-manager to dynamically issue a certificate to every new route created under this domain.

1. Create the necessary OpenShift resources cert-manager requires to manage certificates for OpenShift routes.

This step creates a new deployment (and therefore a pod) that specifically monitors annotated routes in the cluster. If the **issuer-kind** and **issuer-name** annotations are found in a new route, it requests the Issuer (ClusterIssuer in this case) for a new certificate that is unique to this route and which will honor the hostname that was specified while creating the route.



## NOTE

If the cluster does not have access to GitHub, you can save the raw contents locally and run **oc apply -f localfilename.yaml -n cert-manager**.

```
$ oc -n cert-manager apply -f https://github.com/cert-manager/openshift-
routes/releases/latest/download/cert-manager-openshift-routes.yaml
```

The following additional OpenShift resources are also created in this step:

- **ClusterRole** - grants permissions to watch and update the routes across the cluster
  - **ServiceAccount** - uses permissions to run the newly created pod
  - **ClusterRoleBinding** - binds these two resources
2. Ensure that the new **cert-manager-openshift-routes** pod is running successfully:

```
$ oc -n cert-manager get pods
```

### Example result

| NAME   | READY | STATUS  | RESTARTS | AGE   |
|--|-------|---------|----------|-------|
| cert-manager-866d8f788c-9kspc                  | 1/1   | Running | 0        | 4h21m |
| cert-manager-cainjector-6885c585bd-znws8       | 1/1   | Running | 0        | 4h41m |
| cert-manager-openshift-routes-75b6bb44cd-f8kd5 | 1/1   | Running | 0        | 6s    |
| cert-manager-webhook-8498785dd9-bvfd           | 1/1   | Running | 0        | 4h41m |

## 12.7. DEPLOYING A SAMPLE APPLICATION

Now that dynamic certificates are configured, you can deploy a sample application to confirm that certificates are provisioned and trusted when you expose a new route.

1. Create a new project for your sample application:

```
$ oc new-project hello-world
```

2. Deploy a hello world application:

```
$ oc -n hello-world new-app --image=docker.io/openshift/hello-openshift
```

3. Create a route to expose the application from outside the cluster:

```
$ oc -n hello-world create route edge --service=hello-openshift hello-openshift-tls --hostname
hello.${DOMAIN}
```

4. Verify the certificate for the route is untrusted:

```
$ curl -I https://hello.${DOMAIN}
```

### Example output

```
curl: (60) SSL: no alternative certificate subject name matches target host name
'hello.example.com'
More details here: https://curl.se/docs/sslcerts.html
```

```
curl failed to verify the legitimacy of the server and therefore could not
establish a secure connection to it. To learn more about this situation and
how to fix it, please visit the web page mentioned above.
```

5. Annotate the route to trigger cert-manager to provision a certificate for the custom domain:

```
$ oc -n hello-world annotate route hello-openshift-tls cert-manager.io/issuer-
kind=ClusterIssuer cert-manager.io/issuer-name=letsencrypt-production
```



### NOTE

It takes 2-3 minutes for the certificate to be created. The renewal of the certificate will automatically be managed by the cert-manager Operator as it approaches expiration.

6. Verify the certificate for the route is now trusted:

```
$ curl -I https://hello.${DOMAIN}
```

### Example output

```
HTTP/2 200
date: Thu, 05 Oct 2023 23:45:33 GMT
content-length: 17
```

```
content-type: text/plain; charset=utf-8
set-cookie: 52e4465485b6fb4f8a1b1bed128d0f3b=68676068bb32d24f0f558f094ed8e4d7;
path=/; HttpOnly; Secure; SameSite=None
cache-control: private
```

## 12.8. TROUBLESHOOTING DYNAMIC CERTIFICATE PROVISIONING



### NOTE

The validation process usually takes 2-3 minutes to complete while creating certificates.

If annotating your route does not trigger certificate creation during the certificate create step, run **oc describe** against each of the **certificate**, **certificaterequest**, **order**, and **challenge** resources to view the events or reasons that can help identify the cause of the issue.

```
$ oc get certificate,certificaterequest,order,challenge
```

For troubleshooting, you can refer to this [helpful guide in debugging certificates](#).

You can also use the **cmctl** CLI tool for various certificate management activities, such as checking the status of certificates and testing renewals.



## CHAPTER 13. TUTORIAL: ASSIGNING A CONSISTENT EGRESS IP FOR EXTERNAL TRAFFIC

You can assign a consistent IP address for traffic that leaves your cluster such as security groups which require an IP-based configuration to meet security standards.

By default, Red Hat OpenShift Service on AWS (ROSA) uses the OVN-Kubernetes container network interface (CNI) to assign random IP addresses from a pool. This can make configuring security lockdowns unpredictable or open.

See [Configuring an egress IP address](#) for more information.

### Objectives

- Learn how to configure a set of predictable IP addresses for egress cluster traffic.

### Prerequisites

- A ROSA cluster deployed with OVN-Kubernetes
- The [OpenShift CLI \(oc\)](#)
- The [ROSA CLI \(rosa\)](#)
- [jq](#)

## 13.1. SETTING YOUR ENVIRONMENT VARIABLES

- Set your environment variables by running the following command:



### NOTE

Replace the value of the **ROSA\_MACHINE\_POOL\_NAME** variable to target a different machine pool.

```
$ export ROSA_CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}" | sed 's/-[a-z0-9]{5}$//')
$ export ROSA_MACHINE_POOL_NAME=worker
```

## 13.2. ENSURING CAPACITY

The number of IP addresses assigned to each node is limited for each public cloud provider.

- Verify sufficient capacity by running the following command:

```
$ oc get node -o json | \
jq '.items[] | \
{
  "name": .metadata.name,
  "ips": (.status.addresses | map(select(.type == "InternalIP") | .address)),
```

```
"capacity": (.metadata.annotations."cloud.network.openshift.io/egress-ipconfig" |
fromjson[] | .capacity.ipv4)
}'
```

### Example output

```
---
{
  "name": "ip-10-10-145-88.ec2.internal",
  "ips": [
    "10.10.145.88"
  ],
  "capacity": 14
}
{
  "name": "ip-10-10-154-175.ec2.internal",
  "ips": [
    "10.10.154.175"
  ],
  "capacity": 14
}
---
```

## 13.3. CREATING THE EGRESS IP RULES

1. Before creating the egress IP rules, identify which egress IPs you will use.



### NOTE

The egress IPs that you select should exist as a part of the subnets in which the worker nodes are provisioned.

2. **Optional:** Reserve the egress IPs that you requested to avoid conflicts with the AWS Virtual Private Cloud (VPC) Dynamic Host Configuration Protocol (DHCP) service. Request explicit IP reservations on the [AWS documentation for CIDR reservations](#) page.

## 13.4. ASSIGNING AN EGRESS IP TO A NAMESPACE

1. Create a new project by running the following command:

```
$ oc new-project demo-egress-ns
```

2. Create the egress rule for all pods within the namespace by running the following command:

```
$ cat <<EOF | oc apply -f -
apiVersion: k8s.ovn.org/v1
kind: EgressIP
metadata:
  name: demo-egress-ns
spec:
  # NOTE: these egress IPs are within the subnet range(s) in which my worker nodes
  # are deployed.
  egressIPs:
```

```

- 10.10.100.253
- 10.10.150.253
- 10.10.200.253
namespaceSelector:
  matchLabels:
    kubernetes.io/metadata.name: demo-egress-ns
EOF

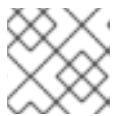
```

## 13.5. ASSIGNING AN EGRESS IP TO A POD

1. Create a new project by running the following command:

```
$ oc new-project demo-egress-pod
```

2. Create the egress rule for the pod by running the following command:



### NOTE

**spec.namespaceSelector** is a mandatory field.

```

$ cat <<EOF | oc apply -f -
apiVersion: k8s.ovn.org/v1
kind: EgressIP
metadata:
  name: demo-egress-pod
spec:
  # NOTE: these egress IPs are within the subnet range(s) in which my worker nodes
  #   are deployed.
  egressIPs:
    - 10.10.100.254
    - 10.10.150.254
    - 10.10.200.254
  namespaceSelector:
    matchLabels:
      kubernetes.io/metadata.name: demo-egress-pod
  podSelector:
    matchLabels:
      run: demo-egress-pod
EOF

```

### 13.5.1. Labeling the nodes

1. Obtain your pending egress IP assignments by running the following command:

```
$ oc get egressips
```

#### Example output

```

NAME           EGRESSIPS    ASSIGNED NODE  ASSIGNED EGRESSIPS
demo-egress-ns 10.10.100.253
demo-egress-pod 10.10.100.254

```

The egress IP rule that you created only applies to nodes with the **k8s.ovn.org/egress-assignable** label. Make sure that the label is only on a specific machine pool.

2. Assign the label to your machine pool using the following command:



### WARNING

If you rely on node labels for your machine pool, this command will replace those labels. Be sure to input your desired labels into the **--labels** field to ensure your node labels remain.

```
$ rosa update machinepool ${ROSA_MACHINE_POOL_NAME} \
  --cluster="${ROSA_CLUSTER_NAME}" \
  --labels "k8s.ovn.org/egress-assignable="
```

## 13.5.2. Reviewing the egress IPs

- Review the egress IP assignments by running the following command:

```
$ oc get egressips
```

### Example output

| NAME            | EGRESSIPS     | ASSIGNED NODE                 | ASSIGNED EGRESSIPS |
|-----------------|---------------|-------------------------------|--------------------|
| demo-egress-ns  | 10.10.100.253 | ip-10-10-156-122.ec2.internal | 10.10.150.253      |
| demo-egress-pod | 10.10.100.254 | ip-10-10-156-122.ec2.internal | 10.10.150.254      |

## 13.6. VERIFICATION

### 13.6.1. Deploying a sample application

To test the egress IP rule, create a service that is restricted to the egress IP addresses which we have specified. This simulates an external service that is expecting a small subset of IP addresses.

1. Run the **echoserver** command to replicate a request:

```
$ oc -n default run demo-service --image=gcr.io/google_containers/echoserver:1.4
```

2. Expose the pod as a service and limit the ingress to the egress IP addresses you specified by running the following command:

```
$ cat <<EOF | oc apply -f -
apiVersion: v1
kind: Service
metadata:
  name: demo-service
  namespace: default
```

```

annotations:
  service.beta.kubernetes.io/aws-load-balancer-scheme: "internal"
  service.beta.kubernetes.io/aws-load-balancer-internal: "true"
spec:
  selector:
    run: demo-service
  ports:
    - port: 80
      targetPort: 8080
  type: LoadBalancer
  externalTrafficPolicy: Local
  # NOTE: this limits the source IPs that are allowed to connect to our service. It
  # is being used as part of this demo, restricting connectivity to our egress
  # IP addresses only.
  # NOTE: these egress IPs are within the subnet range(s) in which my worker nodes
  # are deployed.
  loadBalancerSourceRanges:
    - 10.10.100.254/32
    - 10.10.150.254/32
    - 10.10.200.254/32
    - 10.10.100.253/32
    - 10.10.150.253/32
    - 10.10.200.253/32
EOF

```

- Retrieve the load balancer hostname and save it as an environment variable by running the following command:

```
$ export LOAD_BALANCER_HOSTNAME=$(oc get svc -n default demo-service -o json | jq -r '.status.loadBalancer.ingress[].hostname')
```

### 13.6.2. Testing the namespace egress

- Start an interactive shell to test the namespace egress rule:

```
$ oc run \
  demo-egress-ns \
  -it \
  --namespace=demo-egress-ns \
  --env=LOAD_BALANCER_HOSTNAME=$LOAD_BALANCER_HOSTNAME \
  --image=registry.access.redhat.com/ubi9/ubi -- \
  bash
```

- Send a request to the load balancer and ensure that you can successfully connect:

```
$ curl -s http://$LOAD_BALANCER_HOSTNAME
```

- Check the output for a successful connection:



#### NOTE

The **client\_address** is the internal IP address of the load balancer not your egress IP. You can verify that you have configured the client address correctly by connecting with your service limited to **.spec.loadBalancerSourceRanges**.

## Example output

```

CLIENT VALUES:
client_address=10.10.207.247
command=GET
real path=/
query=nil
request_version=1.1
request_uri=http://internal-a3e61de18bfca4a53a94a208752b7263-148284314.us-east-1.elb.amazonaws.com:8080/

SERVER VALUES:
server_version=nginx: 1.10.0 - lua: 10001

HEADERS RECEIVED:
accept=/*/*
host=internal-a3e61de18bfca4a53a94a208752b7263-148284314.us-east-1.elb.amazonaws.com
user-agent=curl/7.76.1
BODY:
-no body in request-

```

- Exit the pod by running the following command:

```
$ exit
```

### 13.6.3. Testing the pod egress

- Start an interactive shell to test the pod egress rule:

```

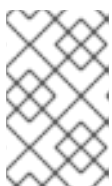
$ oc run \
  demo-egress-pod \
  -it \
  --namespace=demo-egress-pod \
  --env=LOAD_BALANCER_HOSTNAME=$LOAD_BALANCER_HOSTNAME \
  --image=registry.access.redhat.com/ubi9/ubi -- \
  bash

```

- Send a request to the load balancer by running the following command:

```
$ curl -s http://$LOAD_BALANCER_HOSTNAME
```

- Check the output for a successful connection:



#### NOTE

The **client\_address** is the internal IP address of the load balancer not your egress IP. You can verify that you have configured the client address correctly by connecting with your service limited to **.spec.loadBalancerSourceRanges**.

## Example output

```
CLIENT VALUES:
```

```

client_address=10.10.207.247
command=GET
real path=/
query=nil
request_version=1.1
request_uri=http://internal-a3e61de18bfca4a53a94a208752b7263-148284314.us-east-1.elb.amazonaws.com:8080/

SERVER VALUES:
server_version=nginx: 1.10.0 - lua: 10001

HEADERS RECEIVED:
accept=/*/*
host=internal-a3e61de18bfca4a53a94a208752b7263-148284314.us-east-1.elb.amazonaws.com
user-agent=curl/7.76.1
BODY:
-no body in request-

```

- Exit the pod by running the following command:

```
$ exit
```

#### 13.6.4. Optional: Testing blocked egress

- Optional:** Test that the traffic is successfully blocked when the egress rules do not apply by running the following command:

```

$ oc run \
  demo-egress-pod-fail \
  -it \
  --namespace=demo-egress-pod \
  --env=LOAD_BALANCER_HOSTNAME=$LOAD_BALANCER_HOSTNAME \
  --image=registry.access.redhat.com/ubi9/ubi -- \
  bash

```

- Send a request to the load balancer by running the following command:

```
$ curl -s http://$LOAD_BALANCER_HOSTNAME
```

- If the command is unsuccessful, egress is successfully blocked.
- Exit the pod by running the following command:

```
$ exit
```

## 13.7. CLEANING UP YOUR CLUSTER

- Clean up your cluster by running the following commands:

```

$ oc delete svc demo-service -n default; \
$ oc delete pod demo-service -n default; \
$ oc delete project demo-egress-ns; \

```

```
$ oc delete project demo-egress-pod; \  
$ oc delete egressip demo-egress-ns; \  
$ oc delete egressip demo-egress-pod
```

2. Clean up the assigned node labels by running the following command:



### WARNING

If you rely on node labels for your machine pool, this command replaces those labels. Input your desired labels into the **--labels** field to ensure your node labels remain.

```
$ rosa update machinepool ${ROSA_MACHINE_POOL_NAME} \  
--cluster="${ROSA_CLUSTER_NAME}" \  
--labels ""
```



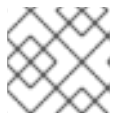
## CHAPTER 14. TUTORIAL: UPDATING COMPONENT ROUTES WITH CUSTOM DOMAINS AND TLS CERTIFICATES

This guide demonstrates how to modify the hostname and TLS certificate of the Web console, OAuth server, and Downloads component routes in Red Hat OpenShift Service on AWS (ROSA) version 4.14 and above.<sup>[1]</sup>

The changes that we make to the component routes<sup>[2]</sup> in this guide are described in greater detail in the customizing the [internal OAuth server URL](#), [console route](#), and [download route](#) OpenShift Container Platform documentation.

### 14.1. PREREQUISITES

- ROSA CLI (**rosa**) version 1.2.37 or higher
- AWS CLI (**aws**)
- A ROSA Classic cluster version 4.14 or higher



#### NOTE

ROSA with HCP is not supported at this time.

- OpenShift CLI (**oc**)
- **jq** CLI
- Access to the cluster as a user with the **cluster-admin** role.
- OpenSSL (for generating the demonstration SSL/TLS certificates)

### 14.2. SETTING UP YOUR ENVIRONMENT

1. Log in to your cluster using an account with **cluster-admin** privileges.
2. Configure an environment variable for your cluster name:

```
$ export CLUSTER_NAME=$(oc get infrastructure cluster -o=jsonpath="{.status.infrastructureName}" | sed 's/-[a-z0-9]{5}$//')
```

3. Ensure all fields output correctly before moving to the next section:

```
$ echo "Cluster: ${CLUSTER_NAME}"
```

#### Example output

```
Cluster: my-rosa-cluster
```

### 14.3. FIND THE CURRENT ROUTES

1. Verify that you can reach the component routes on their default hostnames.

You can find the hostnames by querying the lists of routes in the **openshift-console** and **openshift-authentication** projects.

```
$ oc get routes -n openshift-console
$ oc get routes -n openshift-authentication
```

### Example output

```
NAME      HOST/PORT                                PATH    SERVICES
PORT     TERMINATION      WILDCARD
console   console-openshift-console.apps.my-example-cluster-
aws.z9a9.p1.openshiftapps.com ... 1 more console https reencrypt/Redirect None
downloads downloads-openshift-console.apps.my-example-cluster-
aws.z9a9.p1.openshiftapps.com ... 1 more downloads http edge/Redirect None
NAME      HOST/PORT                                PATH    SERVICES
PORT     TERMINATION      WILDCARD
oauth-openshift oauth-openshift.apps.my-example-cluster-aws.z9a9.p1.openshiftapps.com
... 1 more oauth-openshift 6443 passthrough/Redirect None
```

From this output you can see that our base hostname is **z9a9.p1.openshiftapps.com**.

- Get the ID of the default ingress by running the following command:

```
$ export INGRESS_ID=$(rosa list ingress -c ${CLUSTER_NAME} -o json | jq -r '[] |
select(.default == true) | .id')
```

- Ensure all fields output correctly before moving to the next section:

```
$ echo "Ingress ID: ${INGRESS_ID}"
```

### Example output

```
Ingress ID: r3l6
```

By running these commands you can see that the default component routes for our cluster are:

- **console-openshift-console.apps.my-example-cluster-aws.z9a9.p1.openshiftapps.com** for Console
- **downloads-openshift-console.apps.my-example-cluster-aws.z9a9.p1.openshiftapps.com** for Downloads
- **oauth-openshift.apps.my-example-cluster-aws.z9a9.p1.openshiftapps.com** for OAuth

We can use the **rosa edit ingress** command to change the hostname of each service and add a TLS certificate for all of our component routes. The relevant parameters are shown in this excerpt of the command line help for the **rosa edit ingress** command:

```
$ rosa edit ingress -h
Edit a cluster ingress for a cluster. Usage:
rosa edit ingress ID [flags]
[...]
--component-routes string      Component routes settings. Available keys [oauth, console,
```

downloads]. For each key a pair of hostname and `tlsSecretRef` is expected to be supplied. Format should be a comma separate list `'oauth: hostname=example-hostname;tlsSecretRef=example-secret-ref,downloads:...'`

For this example, we'll use the following custom component routes:

- **console.my-new-domain.dev** for Console
- **downloads.my-new-domain.dev** for Downloads
- **oauth.my-new-domain.dev** for OAuth

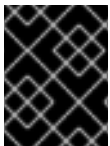
## 14.4. CREATE A VALID TLS CERTIFICATE FOR EACH COMPONENT ROUTE

In this section, we create three separate self-signed certificate key pairs and then trust them to verify that we can access our new component routes using a real web browser.



### WARNING

This is for demonstration purposes only, and is not recommended as a solution for production workloads. Consult your certificate authority to understand how to create certificates with similar attributes for your production workloads.



### IMPORTANT

To prevent issues with HTTP/2 connection coalescing, you must use a separate individual certificate for each endpoint. Using a wildcard or SAN certificate is not supported.

1. Generate a certificate for each component route, taking care to set our certificate's subject (**-subj**) to the custom domain of the component route we want to use:

#### Example

```
$ openssl req -newkey rsa:2048 -new -nodes -x509 -days 365 -keyout key-console.pem -out cert-console.pem -subj "/CN=console.my-new-domain.dev"
$ openssl req -newkey rsa:2048 -new -nodes -x509 -days 365 -keyout key-downloads.pem -out cert-downloads.pem -subj "/CN=downloads.my-new-domain.dev"
$ openssl req -newkey rsa:2048 -new -nodes -x509 -days 365 -keyout key-oauth.pem -out cert-oauth.pem -subj "/CN=oauth.my-new-domain.dev"
```

This generates three pairs of **.pem** files, **key-<component>.pem** and **cert-<component>.pem**.

## 14.5. ADD THE CERTIFICATES TO THE CLUSTER AS SECRETS

1. Create three TLS secrets in the **openshift-config** namespace.  
These become your secret reference when you update the component routes later in this guide.

```
$ oc create secret tls console-tls --cert=cert-console.pem --key=key-console.pem -n
openshift-config
$ oc create secret tls downloads-tls --cert=cert-downloads.pem --key=key-downloads.pem -n
openshift-config
$ oc create secret tls oauth-tls --cert=cert-oauth.pem --key=key-oauth.pem -n openshift-
config
```

## 14.6. FIND THE HOSTNAME OF THE LOAD BALANCER IN YOUR CLUSTER

When you create a cluster, the service creates a load balancer and generates a hostname for that load balancer. We need to know the load balancer hostname in order to create DNS records for our cluster.

You can find the hostname by running the **oc get svc** command against the **openshift-ingress** namespace. The hostname of the load balancer is the **EXTERNAL-IP** associated with the **router-default** service in the **openshift-ingress** namespace.

```
$ oc get svc -n openshift-ingress
NAME          TYPE          CLUSTER-IP   EXTERNAL-IP          PORT(S)
AGE
router-default LoadBalancer 172.30.237.88 a234gsr3242rsfsfs-1342r624.us-east-
1.elb.amazonaws.com 80:31175/TCP,443:31554/TCP 76d
```

In our case, the hostname is **a234gsr3242rsfsfs-1342r624.us-east-1.elb.amazonaws.com**.

Save this value for later, as we will need it to configure DNS records for our new component route hostnames.

## 14.7. ADD COMPONENT ROUTE DNS RECORDS TO YOUR HOSTING PROVIDER

In your hosting provider, add DNS records that map the **CNAME** of your new component route hostnames to the load balancer hostname we found in the previous step.

## 14.8. UPDATE THE COMPONENT ROUTES AND TLS SECRET USING THE ROSA CLI

When your DNS records have been updated, you can use the ROSA CLI to change the component routes.

1. Use the **rosa edit ingress** command to update your default ingress route with the new base domain and the secret reference associated with it, taking care to update the hostnames for each component route.

```
$ rosa edit ingress -c ${CLUSTER_NAME} ${INGRESS_ID} --component-routes 'console:
hostname=console.my-new-domain.dev;tlsSecretRef=console-tls,downloads:
hostname=downloads.my-new-domain.dev;tlsSecretRef=downloads-tls,oauth:
hostname=oauth.my-new-domain.dev;tlsSecretRef=oauth-tls'
```



## NOTE

You can also edit only a subset of the component routes by leaving the component routes you do not want to change set to an empty string. For example, if you only want to change the Console and OAuth server hostnames and TLS certificates, you would run the following command:

```
$ rosa edit ingress -c ${CLUSTER_NAME} ${INGRESS_ID} --component-
routes 'console: hostname=console.my-new-
domain.dev;tlsSecretRef=console-tls,downloads:
hostname="";tlsSecretRef="", oauth: hostname=oauth.my-new-
domain.dev;tlsSecretRef=oauth-tls'
```

2. Run the **rosa list ingress** command to verify that your changes were successfully made:

```
$ rosa list ingress -c ${CLUSTER_NAME} -ojson | jq ".[] | select(.id == \"${INGRESS_ID}\") |
.component_routes"
```

### Example output

```
{
  "console": {
    "kind": "ComponentRoute",
    "hostname": "console.my-new-domain.dev",
    "tls_secret_ref": "console-tls"
  },
  "downloads": {
    "kind": "ComponentRoute",
    "hostname": "downloads.my-new-domain.dev",
    "tls_secret_ref": "downloads-tls"
  },
  "oauth": {
    "kind": "ComponentRoute",
    "hostname": "oauth.my-new-domain.dev",
    "tls_secret_ref": "oauth-tls"
  }
}
```

3. Add your certificate to the truststore on your local system, then confirm that you can access your components at their new routes using your local web browser.

## 14.9. RESET THE COMPONENT ROUTES TO THE DEFAULT USING THE ROSA CLI

If you want to reset the component routes to the default configuration, run the following **rosa edit ingress** command:

```
$ rosa edit ingress -c ${CLUSTER_NAME} ${INGRESS_ID} --component-routes 'console:
hostname="";tlsSecretRef="",downloads: hostname="";tlsSecretRef="", oauth:
hostname="";tlsSecretRef=""'
```

[1] Modifying these routes on Red Hat OpenShift Service on AWS ROSA versions prior to 4.14 is not typically supported. However, if you have a cluster using version 4.13, you can request for Red Hat Support to enable support for this feature on your version 4.13 cluster by [opening a support case](#)

[2] We use the term "component routes" to refer to the OAuth, Console, and Downloads routes that are provided when ROSA are first installed.

## CHAPTER 15. GETTING STARTED WITH ROSA

### 15.1. TUTORIAL: WHAT IS ROSA

Red Hat OpenShift Service on AWS (ROSA) is a fully-managed turnkey application platform that allows you to focus on what matters most, delivering value to your customers by building and deploying applications. Red Hat and AWS SRE experts manage the underlying platform so you do not have to worry about infrastructure management. ROSA provides seamless integration with a wide range of AWS compute, database, analytics, machine learning, networking, mobile, and other services to further accelerate the building and delivering of differentiating experiences to your customers.

ROSA makes use of AWS Security Token Service (STS) to obtain credentials to manage infrastructure in your AWS account. AWS STS is a global web service that creates temporary credentials for IAM users or federated users. ROSA uses this to assign short-term, limited-privilege, security credentials. These credentials are associated with IAM roles that are specific to each component that makes AWS API calls. This method aligns with the principals of least privilege and secure practices in cloud service resource management. The ROSA command line interface (CLI) tool manages the STS credentials that are assigned for unique tasks and takes action on AWS resources as part of OpenShift functionality.

#### 15.1.1. Key features of ROSA

- **Native AWS service:** Access and use Red Hat OpenShift on-demand with a self-service onboarding experience through the AWS management console.
- **Flexible, consumption-based pricing:** Scale to your business needs and pay as you go with flexible pricing and an on-demand hourly or annual billing model.
- **Single bill for Red Hat OpenShift and AWS usage:** Customers will receive a single bill from AWS for both Red Hat OpenShift and AWS consumption.
- **Fully integrated support experience:** Installation, management, maintenance, and upgrades are performed by Red Hat site reliability engineers (SREs) with joint Red Hat and Amazon support and a 99.95% service-level agreement (SLA).
- **AWS service integration:** AWS has a robust portfolio of cloud services, such as compute, storage, networking, database, analytics, and machine learning. All of these services are directly accessible through ROSA. This makes it easier to build, operate, and scale globally and on-demand through a familiar management interface.
- **Maximum Availability:** Deploy clusters across multiple availability zones in supported regions to maximize availability and maintain high availability for your most demanding mission-critical applications and data.
- **Cluster node scaling:** Easily add or remove compute nodes to match resource demand.
- **Optimized clusters:** Choose from memory-optimized, compute-optimized, or general purpose EC2 instance types with clusters sized to meet your needs.
- **Global availability:** Refer to the [product regional availability page](#) to see where ROSA is available globally.

#### 15.1.2. ROSA and Kubernetes

In ROSA, everything you need to deploy and manage containers is bundled, including container management, Operators, networking, load balancing, service mesh, CI/CD, firewall, monitoring, registry,

authentication, and authorization capabilities. These components are tested together for unified operations as a complete platform. Automated cluster operations, including over-the-air platform upgrades, further enhance your Kubernetes experience.

### 15.1.3. Basic responsibilities

In general, cluster deployment and upkeep is Red Hat's or AWS's responsibility, while applications, users, and data is the customer's responsibility. For a more detailed breakdown of responsibilities, see the [responsibility matrix](#).

### 15.1.4. Roadmap and feature requests

Visit the [ROSA roadmap](#) to stay up-to-date with the status of features currently in development. Open a new issue if you have any suggestions for the product team.

### 15.1.5. AWS region availability

Refer to the [product regional availability](#) page for an up-to-date view of where ROSA is available.

### 15.1.6. Compliance certifications

ROSA is currently compliant with SOC-2 type 2, SOC 3, ISO-27001, ISO 27017, ISO 27018, HIPAA, GDPR, and PCI-DSS. We are also currently working towards FedRAMP High.

### 15.1.7. Nodes

#### 15.1.7.1. Worker nodes across multiple AWS regions

All nodes in a ROSA cluster must be located in the same AWS region. For clusters configured for multiple availability zones, control plane nodes and worker nodes will be distributed across the availability zones.

#### 15.1.7.2. Minimum number of worker nodes

For a ROSA cluster, the minimum is 2 worker nodes for single availability zone and 3 worker nodes for multiple availability zones.

#### 15.1.7.3. Underlying node operating system

As with all OpenShift v4.x offerings, the control plane, infra and worker nodes run Red Hat Enterprise Linux CoreOS (RHCOS).

#### 15.1.7.4. Node hibernation or shut-down

At this time, ROSA does not have a hibernation or shut-down feature for nodes. The shutdown and hibernation feature is an OpenShift platform feature that is not yet mature enough for widespread cloud services use.

#### 15.1.7.5. Supported instances for worker nodes

For a complete list of supported instances for worker nodes see [AWS instance types](#). Spot instances are also supported.



### 15.1.7.6. Node autoscaling

Autoscaling allows you to automatically adjust the size of the cluster based on the current workload. See [About autoscaling nodes on a cluster](#) for more details.

### 15.1.7.7. Maximum number of worker nodes

The maximum number of worker nodes is 180 worker nodes for each ROSA cluster. See [limits and scalability](#) for more details on node counts.

A list of the account-wide and per-cluster roles is provided in the [ROSA documentation](#).

### 15.1.8. Administrators

A ROSA customer's administrator can manage users and quotas in addition to accessing all user-created projects.

### 15.1.9. OpenShift versions and upgrades

ROSA is a managed service which is based on OpenShift Container Platform. You can view the current version and life cycle dates in the [ROSA documentation](#).

Customers can upgrade to the newest version of OpenShift and use the features from that version of OpenShift. For more information, see [life cycle dates](#). Not all OpenShift features are available on ROSA. Review the [Service Definition](#) for more information.

### 15.1.10. Support

You can open a ticket directly from the [OpenShift Cluster Manager](#). See the [ROSA support documentation](#) for more details about obtaining support.

You can also visit the [Red Hat Customer Portal](#) to search or browse through the Red Hat knowledge base of articles and solutions relating to Red Hat products or submit a support case to Red Hat Support.

#### 15.1.10.1. Limited support

If a ROSA cluster is not upgraded before the "end of life" date, the cluster continues to operate in a limited support status. The SLA for that cluster will no longer be applicable, but you can still get support for that cluster. See the [limited support status](#) documentation for more details.

#### Additional support resources

- [Red Hat Support](#)
- [AWS Support](#)  
AWS support customers must have a valid AWS support contract

### 15.1.11. Service-level agreement (SLA)

Refer to the [ROSA SLA](#) page for details.

### 15.1.12. Notifications and communication

Red Hat will provide notifications regarding new Red Hat and AWS features, updates, and scheduled maintenance through email and the Hybrid Cloud Console service log.

### 15.1.13. Open Service Broker for AWS (OSBA)

You can use OSBA with ROSA. However, the preferred method is the more recent [AWS Controller for Kubernetes](#). See [Open Service Broker for AWS](#) for more information on OSBA.

### 15.1.14. Offboarding

Customers can stop using ROSA at any time and move their applications to on-premise, a private cloud, or other cloud providers. Standard reserved instances (RI) policy applies for unused RI.

### 15.1.15. Authentication

ROSA supports the following authentication mechanisms: OpenID Connect (a profile of OAuth2), Google OAuth, GitHub OAuth, GitLab, and LDAP.

### 15.1.16. SRE cluster access

All SRE cluster access is secured by MFA. See [SRE access](#) for more details.

### 15.1.17. Encryption

#### 15.1.17.1. Encryption keys

ROSA uses a key stored in KMS to encrypt EBS volumes. Customers also have the option to provide their own KMS keys at cluster creation.

#### 15.1.17.2. KMS keys

If you specify a KMS key, the control plane, infrastructure and worker node root volumes and the persistent volumes are encrypted with the key.

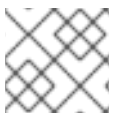
#### 15.1.17.3. Data encryption

By default, there is encryption at rest. The AWS Storage platform automatically encrypts your data before persisting it and decrypts the data before retrieval. See [AWS EBS Encryption](#) for more details.

You can also encrypt etcd in the cluster, combining it with AWS storage encryption. This results in double the encryption which adds up to a 20% performance hit. For more details see the [etcd encryption](#) documentation.

#### 15.1.17.4. etcd encryption

etcd encryption can only be enabled at cluster creation.



#### NOTE

etcd encryption incurs additional overhead with negligible security risk mitigation.

#### 15.1.17.5. etcd encryption configuration

etcd encryption is configured the same as in OpenShift Container Platform. The aescbc cypher is used and the setting is patched during cluster deployment. For more details, see the [Kubernetes documentation](#).

#### 15.1.17.6. Multi-region KMS keys for EBS encryption

Currently, the ROSA CLI does not accept multi-region KMS keys for EBS encryption. This feature is in our backlog for product updates. The ROSA CLI accepts single region KMS keys for EBS encryption if it is defined at cluster creation.

#### 15.1.18. Infrastructure

ROSA uses several different cloud services such as virtual machines, storage, and load balancers. You can see a defined list in the [AWS prerequisites](#).

#### 15.1.19. Credential methods

There are two credential methods to grant Red Hat the permissions needed to perform the required actions in your AWS account: AWS with STS or an IAM user with admin permissions. AWS with STS is the preferred method, and the IAM user method will eventually be deprecated. AWS with STS better aligns with the principles of least privilege and secure practices in cloud service resource management.

#### 15.1.20. Prerequisite permission or failure errors

Check for a newer version of the ROSA CLI. Every release of the ROSA CLI is located in two places: [Github](#) and the [Red Hat signed binary releases](#).

#### 15.1.21. Storage

Refer to the [storage](#) section of the service definition.

OpenShift includes the CSI driver for AWS EFS. For more information, see [Setting up AWS EFS for Red Hat OpenShift Service on AWS](#).

#### 15.1.22. Using a VPC

At installation you can select to deploy to an existing VPC or bring your own VPC. You can then select the required subnets and provide a valid CIDR range that encompasses the subnets for the installation program when using those subnets.

ROSA allows multiple clusters to share the same VPC. The number of clusters on one VPC is limited by the remaining AWS resource quota and CIDR ranges that cannot overlap. See [CIDR Range Definitions](#) for more information.

#### 15.1.23. Network plugin

ROSA uses the OpenShift OVN-Kubernetes default CNI network provider.

#### 15.1.24. Cross-namespace networking

Cluster admins can customize, and deny, cross-namespace on a project basis using NetworkPolicy objects. Refer to [Configuring multitenant isolation with network policy](#) for more information.

### 15.1.25. Using Prometheus and Grafana

You can use Prometheus and Grafana to monitor containers and manage capacity using OpenShift User Workload Monitoring. This is a check-box option in the [OpenShift Cluster Manager](#).

### 15.1.26. Audit logs output from the cluster control-plane

If the Cluster Logging Operator Add-on has been added to the cluster then audit logs are available through CloudWatch. If it has not, then a support request would allow you to request some audit logs. Small targeted and time-boxed logs can be requested for export and sent to a customer. The selection of audit logs available are at the discretion of SRE in the category of platform security and compliance. Requests for exports of a cluster's entirety of logs will be rejected.

### 15.1.27. AWS Permissions Boundary

You can use an AWS Permissions Boundary around the policies for your cluster.

### 15.1.28. AMI

ROSA worker nodes use a different AMI from OSD and OpenShift Container Platform. Control Plane and Infra node AMIs are common across products in the same version.

### 15.1.29. Cluster backups

ROSA STS clusters do not have backups. Users must have their own backup policies for applications and data. See our [backup policy](#) for more information.

### 15.1.30. Custom domain

You can define a custom domain for your applications. See [Configuring custom domains for applications](#) for more information.

### 15.1.31. ROSA domain certificates

Red Hat infrastructure (Hive) manages certificate rotation for default application ingress.

### 15.1.32. Disconnected environments

ROSA does not support an air-gapped, disconnected environment. The ROSA cluster must have egress to the internet to access our registry, S3, and send metrics. The service requires a number of egress endpoints. Ingress can be limited to a PrivateLink for Red Hat SREs and a VPN for customer access.

#### Additional Resources

- ROSA product pages:
  - [Red Hat product page](#)
  - [AWS product page](#)
  - [Red Hat Customer Portal](#)
- ROSA specific resources

- [AWS ROSA getting started guide](#)
- [ROSA documentation](#)
- [ROSA service definition](#)
- [ROSA responsibility assignment matrix](#)
- [Understanding Process and Security](#)
- [About Availability](#)
- [Updates Lifecycle](#)
- [Limits and Scalability](#)
- [ROSA roadmap](#)
- [Learn about OpenShift](#)
- [OpenShift Cluster Manager](#)
- [Red Hat Support](#)

## 15.2. TUTORIAL: ROSA WITH AWS STS EXPLAINED

This tutorial outlines the two options for allowing Red Hat OpenShift Service on AWS (ROSA) to interact with resources in a user's Amazon Web Service (AWS) account. It details the components and processes that ROSA with Security Token Service (STS) uses to obtain the necessary credentials. It also reviews why ROSA with STS is the more secure, preferred method.



### NOTE

This content currently covers ROSA Classic with AWS STS. For ROSA with hosted control planes (HCP) with AWS STS, see [AWS STS and ROSA with HCP explained](#).

This tutorial will:

- Enumerate two of the deployment options:
  - ROSA with IAM Users
  - ROSA with STS
- Explain the differences between the two options
- Explain why ROSA with STS is more secure and the preferred option
- Explain how ROSA with STS works

### 15.2.1. Different credential methods to deploy ROSA

As part of ROSA, Red Hat manages infrastructure resources in your AWS account and must be granted the necessary permissions. There are currently two supported methods for granting those permissions:

- Using static IAM user credentials with an **AdministratorAccess** policy

This is referred to as "ROSA with IAM Users" in this tutorial. It is not the preferred credential method.

- Using AWS STS with short-lived, dynamic tokens  
This is referred to as "ROSA with STS" in this tutorial. It is the preferred credential method.

### 15.2.1.1. Rosa with IAM Users

When ROSA was first released, the only credential method was ROSA with IAM Users. This method grants IAM users with an **AdministratorAccess** policy full access to create the necessary resources in the AWS account that uses ROSA. The cluster can then create and expand its credentials as needed.

### 15.2.1.2. ROSA with STS

ROSA with STS grants users limited, short-term access to resources in your AWS account. The STS method uses predefined roles and policies to grant temporary, least-privilege permissions to IAM users or authenticated federated users. The credentials typically expire an hour after being requested. Once expired, they are no longer recognized by AWS and no longer have account access from API requests made with them. For more information, see the [AWS documentation](#). While both ROSA with IAM Users and ROSA with STS are currently enabled, ROSA with STS is the preferred and recommended option.

## 15.2.2. ROSA with STS security

Several crucial components make ROSA with STS more secure than ROSA with IAM Users:

- An explicit and limited set of roles and policies that the user creates ahead of time. The user knows every requested permission and every role used.
- The service cannot do anything outside of those permissions.
- Whenever the service needs to perform an action, it obtains credentials that expire in one hour or less. This means that there is no need to rotate or revoke credentials. Additionally, credential expiration reduces the risks of credentials leaking and being reused.

### 15.2.3. AWS STS explained

ROSA uses AWS STS to grant least-privilege permissions with short-term security credentials to specific and segregated IAM roles. The credentials are associated with IAM roles specific to each component and cluster that makes AWS API calls. This method aligns with principles of least-privilege and secure practices in cloud service resource management. The ROSA command line interface (CLI) tool manages the STS roles and policies that are assigned for unique tasks and takes action upon AWS resources as part of OpenShift functionality.

STS roles and policies must be created for each ROSA cluster. To make this easier, the installation tools provide all the commands and files needed to create the roles as policies and an option to allow the CLI to automatically create the roles and policies. See [Creating a ROSA cluster with STS using customizations](#) for more information about the different **--mode** options.

### 15.2.4. Components specific to ROSA with STS

- **AWS infrastructure** - This provides the infrastructure required for the cluster. It contains the actual EC2 instances, storage, and networking components. See [AWS compute types](#) to see supported instance types for compute nodes and [provisioned AWS infrastructure](#) for control plane and infrastructure node configuration.

- **AWS STS** - See the credential method section above.
- **OpenID Connect (OIDC)** - This provides a mechanism for cluster Operators to authenticate with AWS, assume the cluster roles through a trust policy, and obtain temporary credentials from STS to make the required API calls.
- **Roles and policies** - The roles and policies are one of the main differences between ROSA with STS and ROSA with IAM Users. For ROSA with STS, the roles and policies used by ROSA are broken into account-wide roles and policies and Operator roles and policies. The policies determine the allowed actions for each of the roles. See [About IAM resources for ROSA clusters that use STS](#) for more details about the individual roles and policies.
  - The account-wide roles are:
    - ManagedOpenShift-Installer-Role
    - ManagedOpenShift-ControlPlane-Role
    - ManagedOpenShift-Worker-Role
    - ManagedOpenShift-Support-Role
  - The account-wide policies are:
    - ManagedOpenShift-Installer-Role-Policy
    - ManagedOpenShift-ControlPlane-Role-Policy
    - ManagedOpenShift-Worker-Role-Policy
    - ManagedOpenShift-Support-Role-Policy
    - ManagedOpenShift-openshift-ingress-operator-cloud-credentials <sup>[1]</sup>
    - ManagedOpenShift-openshift-cluster-csi-drivers-ebs-cloud-credential <sup>[1]</sup>
    - ManagedOpenShift-openshift-cloud-network-config-controller-cloud <sup>[1]</sup>
    - ManagedOpenShift-openshift-machine-api-aws-cloud-credentials <sup>[1]</sup>
    - ManagedOpenShift-openshift-cloud-credential-operator-cloud-credential <sup>[1]</sup>
    - ManagedOpenShift-openshift-image-registry-installer-cloud-credential <sup>[1]</sup>
      1. This policy is used by the cluster Operator roles, listed below. The Operator roles are created in a second step because they are dependent on an existing cluster name and cannot be created at the same time as the account-wide roles.
  - The Operator roles are:
    - <cluster-name>-xxxx-openshift-cluster-csi-drivers-ebs-cloud-credential
    - <cluster-name>-xxxx-openshift-cloud-network-config-controller-cloud
    - <cluster-name>-xxxx-openshift-machine-api-aws-cloud-credentials
    - <cluster-name>-xxxx-openshift-cloud-credential-operator-cloud-credential

- `<cluster-name>-xxxx-openshift-image-registry-installer-cloud-creden`
- `<cluster-name>-xxxx-openshift-ingress-operator-cloud-credentials`
- Trust policies are created for each account-wide and Operator role.

### 15.2.5. Deploying a ROSA STS cluster

You are not expected to create the resources listed in the below steps from scratch. The ROSA CLI creates the required JSON files for you and outputs the commands you need. The ROSA CLI can also take this a step further and run the commands for you, if desired.

#### Steps to deploy a ROSA with STS cluster

1. Create the account-wide roles and policies.
2. Assign the permissions policy to the corresponding account-wide role.
3. Create the cluster.
4. Create the Operator roles and policies.
5. Assign the permission policy to the corresponding Operator role.
6. Create the OIDC provider.

The roles and policies can be created automatically by the ROSA CLI, or they can be manually created by utilizing the **--mode manual** or **--mode auto** flags in the ROSA CLI. For further details about deployment, see [Creating a cluster with customizations](#) or the [Deploying the cluster tutorial](#).

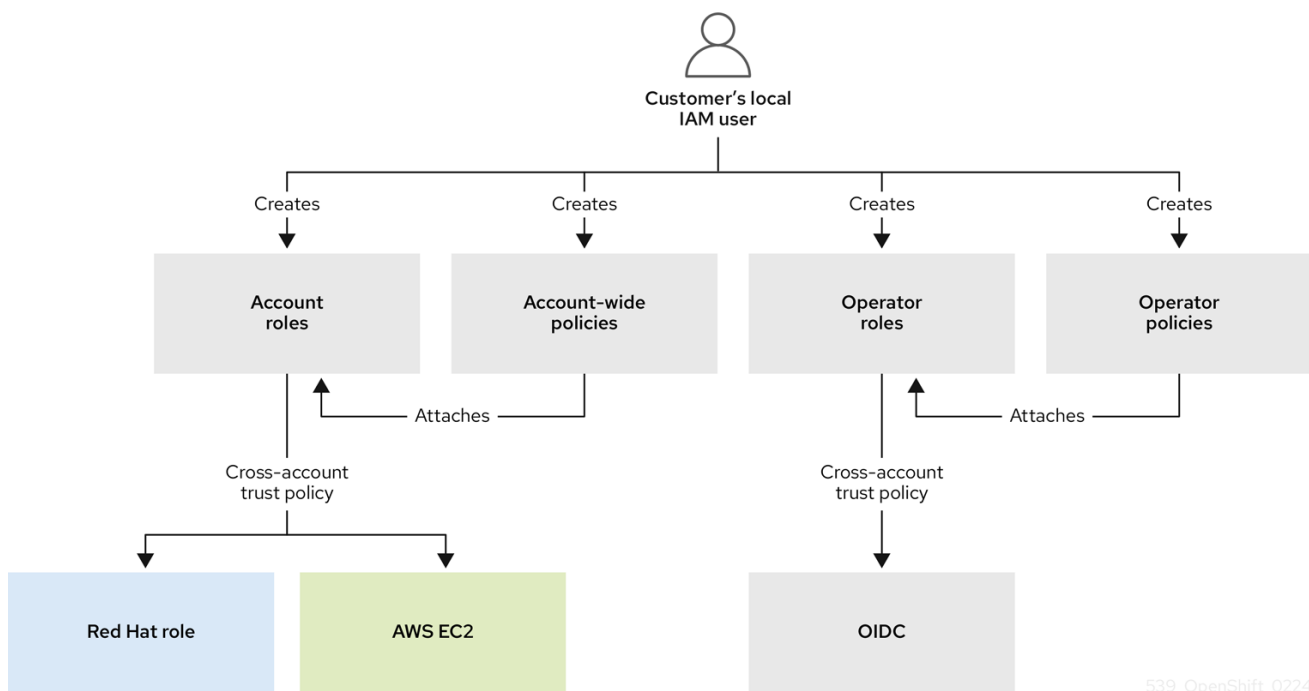
### 15.2.6. ROSA with STS workflow

The user creates the required account-wide roles and account-wide policies. For more information, see the components section in this tutorial. During role creation, a trust policy, known as a cross-account trust policy, is created which allows a Red Hat-owned role to assume the roles. Trust policies are also created for the EC2 service, which allows workloads on EC2 instances to assume roles and obtain credentials. The user can then assign a corresponding permissions policy to each role.

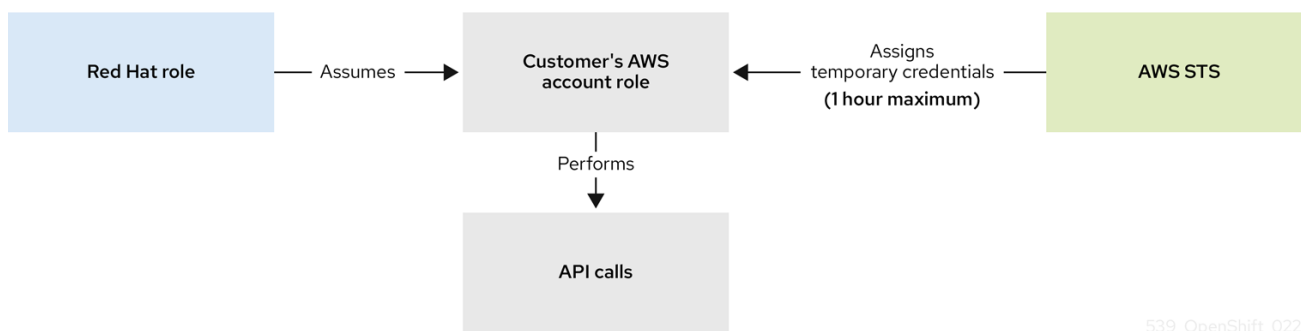
After the account-wide roles and policies are created, the user can create a cluster. Once cluster creation is initiated, the Operator roles are created so that cluster Operators can make AWS API calls. These roles are then assigned to the corresponding permission policies that were created earlier and a trust policy with an OIDC provider. The Operator roles differ from the account-wide roles in that they ultimately represent the pods that need access to AWS resources. Because a user cannot attach IAM roles to pods, they must create a trust policy with an OIDC provider so that the Operator, and therefore the pods, can access the roles they need.

Once the user assigns the roles to the corresponding policy permissions, the final step is creating the OIDC provider.



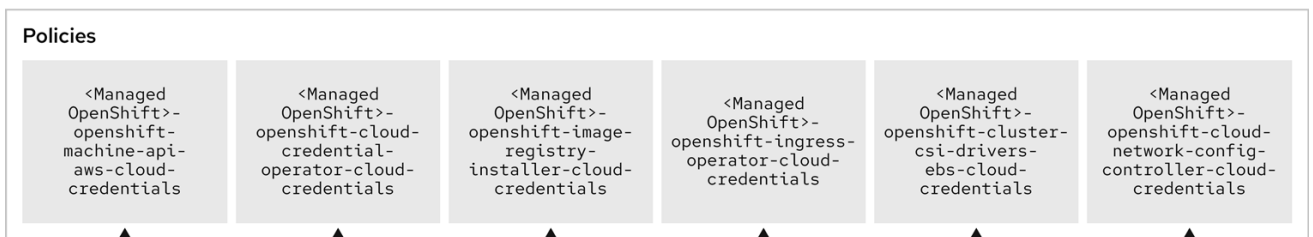
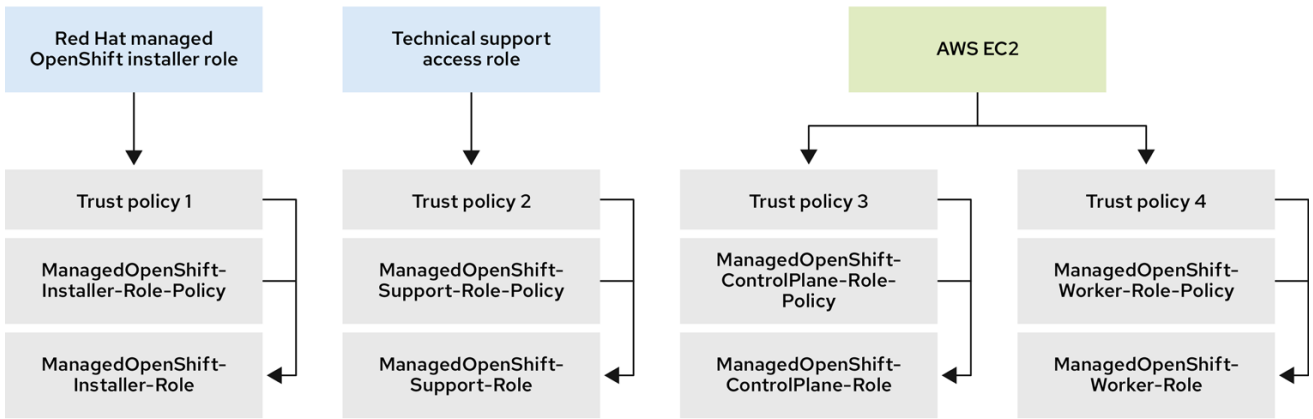


When a new role is needed, the workload currently using the Red Hat role will assume the role in the AWS account, obtain temporary credentials from AWS STS, and begin performing the actions using API calls within the customer's AWS account as permitted by the assumed role's permissions policy. The credentials are temporary and have a maximum duration of one hour.

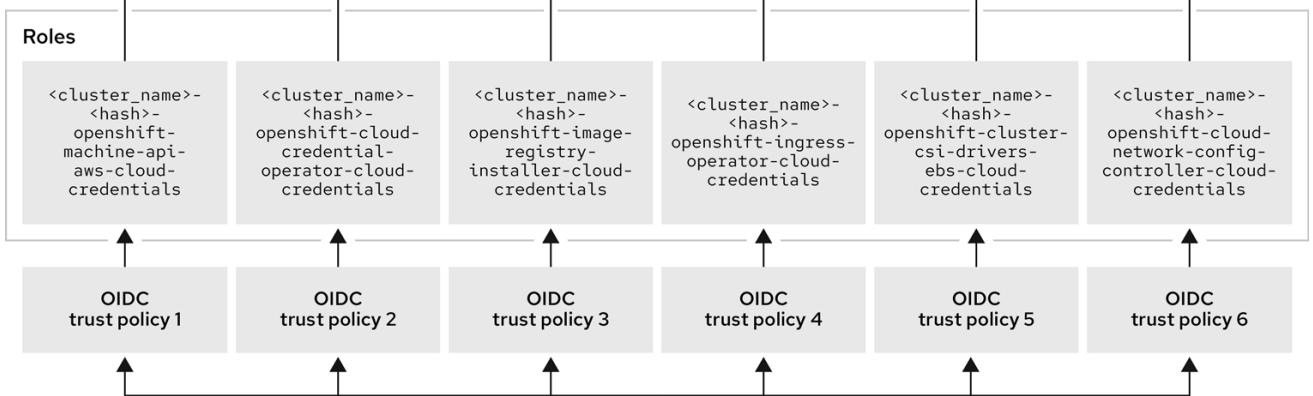


The entire workflow is depicted in the following graphic:

**1 Create account roles**



**2 Create Operator roles**

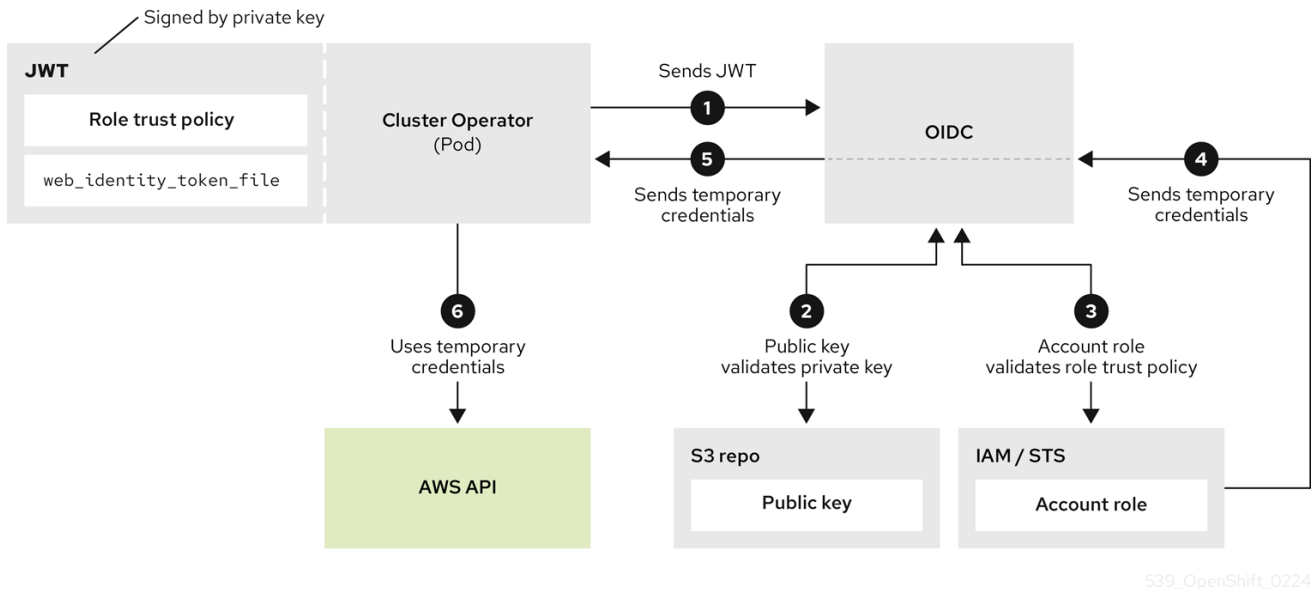


**3 Create OIDC provider**



539\_OpenShift\_0224

Operators use the following process to obtain the requisite credentials to perform their tasks. Each Operator is assigned an Operator role, a permissions policy, and a trust policy with an OIDC provider. The Operator will assume the role by passing a JSON web token that contains the role and a token file (**web\_identity\_token\_file**) to the OIDC provider, which then authenticates the signed key with a public key. The public key is created during cluster creation and stored in an S3 bucket. The Operator then confirms that the subject in the signed token file matches the role in the role trust policy which ensures that the OIDC provider can only obtain the allowed role. The OIDC provider then returns the temporary credentials to the Operator so that the Operator can make AWS API calls. For a visual representation, see below:



## 15.2.7. ROSA with STS use cases

### Creating nodes at cluster install

The Red Hat installation program uses the **RH-Managed-OpenShift-Installer** role and a trust policy to assume the **Managed-OpenShift-Installer-Role** role in the customer's account. This process returns temporary credentials from AWS STS. The installation program begins making the required API calls with the temporary credentials just received from STS. The installation program creates the required infrastructure in AWS. The credentials expire within an hour and the installation program no longer has access to the customer's account.

The same process also applies for support cases. In support cases, a Red Hat site reliability engineer (SRE) replaces the installation program.

### Scaling the cluster

The **machine-api-operator** uses [AssumeRoleWithWebIdentity](#) to assume the **machine-api-aws-cloud-credentials** role. This launches the sequence for the cluster Operators to receive the credentials. The **machine-api-operator** role can now make the relevant API calls to add more EC2 instances to the cluster.

## 15.3. DEPLOYING A CLUSTER

### 15.3.1. Tutorial: Choosing a deployment method

This tutorial outlines the different ways to deploy a cluster. Choose the deployment method that best fits your preferences and needs.

#### 15.3.1.1. Deployment options

If you want:

- Only the necessary CLI commands - [Simple CLI guide](#)
- A user interface - [Simple UI guide](#)

- The CLI commands with details - [Detailed CLI guide](#)
- A user interface with details - [Detailed UI guide](#)

All of the above deployment options work well for this tutorial. If you are doing this tutorial for the first time, the [Simple CLI guide](#) is the simplest and recommended method.

### 15.3.2. Tutorial: Simple CLI guide

This page outlines the minimum list of commands to deploy a Red Hat OpenShift Service on AWS (ROSA) cluster using the command line interface (CLI).



#### NOTE

While this simple deployment works well for a tutorial setting, clusters used in production should be deployed with a more detailed method.

#### 15.3.2.1. Prerequisites

- You have completed the prerequisites in the Setup tutorial.

#### 15.3.2.2. Creating account roles

Run the following command *once* for each AWS account and y-stream OpenShift version:

```
rosa create account-roles --mode auto --yes
```

#### 15.3.2.3. Deploying the cluster

1. Create the cluster with the default configuration by running the following command substituting your own cluster name:

```
rosa create cluster --cluster-name <cluster-name> --sts --mode auto --yes
```

2. Check the status of your cluster by running the following command:

```
rosa list clusters
```

### 15.3.3. Tutorial: Detailed CLI guide

This tutorial outlines the detailed steps to deploy a ROSA cluster using the ROSA CLI.

#### 15.3.3.1. CLI deployment modes

There are two modes with which to deploy a ROSA cluster. One is automatic, which is quicker and performs the manual work for you. The other is manual, requires you to run extra commands, and allows you to inspect the roles and policies being created. This tutorial documents both options.

If you want to create a cluster quickly, use the automatic option. If you prefer exploring the roles and policies being created, use the manual option.

Choose the deployment mode by using the **--mode** flag in the relevant commands.

Valid options for **--mode** are:

- **manual:** Role and policies are created and saved in the current directory. You must manually run the provided commands as the next step. This option allows you to review the policy and roles before creating them.
- **auto:** Roles and policies are created and applied automatically using the current AWS account.

## TIP

You can use either deployment method for this tutorial. The **auto** mode is faster and has less steps.

### 15.3.3.2. Deployment workflow

The overall deployment workflow follows these steps:

1. **rosa create account-roles** - This is executed only *once* for each account. Once created, the account roles do **not** need to be created again for more clusters of the same y-stream version.
2. **rosa create cluster**
3. **rosa create operator-roles** - For manual mode only.
4. **rosa create oidc-provider** - For manual mode only.

For each additional cluster in the same account for the same y-stream version, only step 2 is needed for automatic mode. Steps 2 through 4 are needed for manual mode.

### 15.3.3.3. Automatic mode

Use this method if you want the ROSA CLI to automate the creation of the roles and policies to create your cluster quickly.

#### 15.3.3.3.1. Creating account roles

If this is the *first time* you are deploying ROSA in this account and you have *not* yet created the account roles, then create the account-wide roles and policies, including Operator policies.

Run the following command to create the account-wide roles:

```
rosa create account-roles --mode auto --yes
```

#### Example output

```
I: Creating roles using 'arn:aws:iam::000000000000:user/rosa-user'
I: Created role 'ManagedOpenShift-ControlPlane-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-ControlPlane-Role'
I: Created role 'ManagedOpenShift-Worker-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-Worker-Role'
I: Created role 'ManagedOpenShift-Support-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-Support-Role'
I: Created role 'ManagedOpenShift-Installer-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-Installer-Role'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-
machine-api-aws-cloud-credentials'
```

```
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-cloud-credential-operator-cloud-crede'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-image-registry-installer-cloud-creden'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-ingress-operator-cloud-credentials'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-cluster-csi-drivers-ebs-cloud-credent'
I: To create a cluster with these roles, run the following command:
    rosa create cluster --sts
```

### 15.3.3.3.2. Creating a cluster

Run the following command to create a cluster with all the default options:

```
rosa create cluster --cluster-name <cluster-name> --sts --mode auto --yes
```



#### NOTE

This will also create the required Operator roles and OIDC provider. If you want to see all available options for your cluster use the **--help** flag or **--interactive** for interactive mode.

#### Example input

```
$ rosa create cluster --cluster-name my-rosa-cluster --sts --mode auto --yes
```

#### Example output

```
I: Creating cluster 'my-rosa-cluster'
I: To view a list of clusters and their status, run 'rosa list clusters'
I: Cluster 'my-rosa-cluster' has been created.
I: Once the cluster is installed you will need to add an Identity Provider before you can login into the cluster. See 'rosa create idp --help' for more information.
I: To determine when your cluster is Ready, run 'rosa describe cluster -c my-rosa-cluster'.
I: To watch your cluster installation logs, run 'rosa logs install -c my-rosa-cluster --watch'.
Name:                my-rosa-cluster
ID:                  1mlhulb3bo0l54ojd0ji000000000000
External ID:
OpenShift Version:
Channel Group:       stable
DNS:                 my-rosa-cluster.ibhp.p1.openshiftapps.com
AWS Account:         000000000000
API URL:
Console URL:
Region:              us-west-2
Multi-AZ:            false
Nodes:
- Master:            3
- Infra:              2
- Compute:           2
Network:
- Service CIDR:      172.30.0.0/16
- Machine CIDR:      10.0.0.0/16
```

```

- Pod CIDR:          10.128.0.0/14
- Host Prefix:       /23
STS Role ARN:        arn:aws:iam::000000000000:role/ManagedOpenShift-Installer-Role
Support Role ARN:    arn:aws:iam::000000000000:role/ManagedOpenShift-Support-Role
Instance IAM Roles:
- Master:            arn:aws:iam::000000000000:role/ManagedOpenShift-ControlPlane-Role
- Worker:            arn:aws:iam::000000000000:role/ManagedOpenShift-Worker-Role
Operator IAM Roles:
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-image-registry-installer-cloud-
credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-ingress-operator-cloud-credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-cluster-csi-drivers-ebs-cloud-credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-machine-api-aws-cloud-credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-cloud-credential-operator-cloud-
credential-oper
State:               waiting (Waiting for OIDC configuration)
Private:             No
Created:             Oct 28 2021 20:28:09 UTC
Details Page:
https://console.redhat.com/openshift/details/s/1wupmiQy45xr1nN0000000000000
OIDC Endpoint URL:  https://rh-oidc.s3.us-east-
1.amazonaws.com/1mlhulb3bo0l54ojd0ji0000000000000

```

#### 15.3.3.3.2.1. Default configuration

The default settings are as follows:

- Nodes:
  - 3 control plane nodes
  - 2 infrastructure nodes
  - 2 worker nodes
  - No autoscaling
  - See the documentation on [ec2 instances](#) for more details.
- Region: As configured for the **aws** CLI
- Networking IP ranges:
  - Machine CIDR: 10.0.0.0/16
  - Service CIDR: 172.30.0.0/16
  - Pod CIDR: 10.128.0.0/14
- New VPC
- Default AWS KMS key for encryption
- The most recent version of OpenShift available to **rosa**
- A single availability zone

- Public cluster

#### 15.3.3.3.3. Checking the installation status

1. Run one of the following commands to check the status of your cluster:

- For a detailed view of the status, run:

```
rosa describe cluster --cluster <cluster-name>
```

- For an abridged view of the status, run:

```
rosa list clusters
```

2. The cluster state will change from “waiting” to “installing” to “ready”. This will take about 40 minutes.
3. Once the state changes to “ready” your cluster is installed.

#### 15.3.3.4. Manual Mode

If you want to review the roles and policies before applying them to a cluster, use the manual method. This method requires running a few extra commands to create the roles and policies.

This section uses the **--interactive** mode. See the documentation on [interactive mode](#) for a description of the fields in this section.

##### 15.3.3.4.1. Creating account roles

1. If this is the *first time* you are deploying ROSA in this account and you have *not* yet created the account roles, create the account-wide roles and policies, including the Operator policies. The command creates the needed JSON files for the required roles and policies for your account in the current directory. It also outputs the **aws** CLI commands that you need to run to create these objects.

Run the following command to create the needed files and output the additional commands:

```
rosa create account-roles --mode manual
```

#### Example output

```
I: All policy files saved to the current directory
I: Run the following commands to create the account roles and policies:
aws iam create-role \
--role-name ManagedOpenShift-Worker-Role \
--assume-role-policy-document file://sts_instance_worker_trust_policy.json \
--tags Key=rosa_openshift_version,Value=4.8
Key=rosa_role_prefix,Value=ManagedOpenShift
Key=rosa_role_type,Value=instance_worker
aws iam put-role-policy \
--role-name ManagedOpenShift-Worker-Role \
--policy-name ManagedOpenShift-Worker-Role-Policy \
--policy-document file://sts_instance_worker_permission_policy.json
```



2. Check the contents of your current directory to see the new files. Use the **aws** CLI to create each of these objects.

### Example output

```
$ ls
openshift_cloud_credential_operator_cloud_credential_operator_iam_ro_creds_policy.json
sts_instance_controlplane_permission_policy.json
openshift_cluster_csi_drivers_ebs_cloud_credentials_policy.json
sts_instance_controlplane_trust_policy.json
openshift_image_registry_installer_cloud_credentials_policy.json
sts_instance_worker_permission_policy.json
openshift_ingress_operator_cloud_credentials_policy.json
sts_instance_worker_trust_policy.json
openshift_machine_api_aws_cloud_credentials_policy.json
sts_support_permission_policy.json
sts_installer_permission_policy.json          sts_support_trust_policy.json
sts_installer_trust_policy.json
```

3. **Optional:** Open the files to review what you will create. For example, opening the **sts\_installer\_permission\_policy.json** shows:

### Example output

```
$ cat sts_installer_permission_policy.json
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "autoscaling:DescribeAutoScalingGroups",
        "ec2:AllocateAddress",
        "ec2:AssociateAddress",
        "ec2:AssociateDhcpOptions",
        "ec2:AssociateRouteTable",
        "ec2:AttachInternetGateway",
        "ec2:AttachNetworkInterface",
        "ec2:AuthorizeSecurityGroupEgress",
        "ec2:AuthorizeSecurityGroupIngress",
        [...]
      ]
    }
  ]
}
```

You can also see the contents in the [About IAM resources for ROSA clusters](#) documentation.

4. Run the **aws** commands listed in step 1. You can copy and paste if you are in the same directory as the JSON files you created.

#### 15.3.3.4.2. Creating a cluster

1. After the **aws** commands are executed successfully, run the following command to begin ROSA cluster creation in interactive mode:

```
rosa create cluster --interactive --sts
```

See the [ROSA documentation](#) for a description of the fields.

2. For the purpose of this tutorial, copy and then input the following values:

```

Cluster name: my-rosa-cluster
OpenShift version: <choose version>
External ID (optional): <leave blank>
Operator roles prefix: <accept default>
Multiple availability zones: No
AWS region: <choose region>
PrivateLink cluster: No
Install into an existing VPC: No
Enable Customer Managed key: No
Compute nodes instance type: m5.xlarge
Enable autoscaling: No
Compute nodes: 2
Machine CIDR: <accept default>
Service CIDR: <accept default>
Pod CIDR: <accept default>
Host prefix: <accept default>
Encrypt etcd data (optional): No
Disable Workload monitoring: No

```

### Example output

```

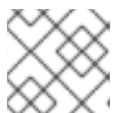
I: Creating cluster 'my-rosa-cluster'
I: To create this cluster again in the future, you can run:
rosa create cluster --cluster-name my-rosa-cluster --role-arn
arn:aws:iam::000000000000:role/ManagedOpenShift-Installer-Role --support-role-arn
arn:aws:iam::000000000000:role/ManagedOpenShift-Support-Role --master-iam-role
arn:aws:iam::000000000000:role/ManagedOpenShift-ControlPlane-Role --worker-iam-role
arn:aws:iam::000000000000:role/ManagedOpenShift-Worker-Role --operator-roles-prefix
my-rosa-cluster --region us-west-2 --version 4.8.13 --compute-nodes 2 --machine-cidr
10.0.0.0/16 --service-cidr 172.30.0.0/16 --pod-cidr 10.128.0.0/14 --host-prefix 23
I: To view a list of clusters and their status, run 'rosa list clusters'
I: Cluster 'my-rosa-cluster' has been created.
I: Once the cluster is installed you will need to add an Identity Provider before you can login
into the cluster. See 'rosa create idp --help' for more information.
Name:          my-rosa-cluster
ID:            1t6i760dbum4mqltqh6o000000000000
External ID:
OpenShift Version:
Channel Group:    stable
DNS:             my-rosa-cluster.abcd.p1.openshiftapps.com
AWS Account:     000000000000
API URL:
Console URL:
Region:         us-west-2
Multi-AZ:       false
Nodes:
- Control plane: 3
- Infra:        2
- Compute:     2
Network:
- Service CIDR: 172.30.0.0/16
- Machine CIDR: 10.0.0.0/16
- Pod CIDR:    10.128.0.0/14

```

```

- Host Prefix:          /23
STS Role ARN:          arn:aws:iam::000000000000:role/ManagedOpenShift-Installer-Role
Support Role ARN:      arn:aws:iam::000000000000:role/ManagedOpenShift-Support-Role
Instance IAM Roles:
- Control plane:       arn:aws:iam::000000000000:role/ManagedOpenShift-ControlPlane-
Role
- Worker:              arn:aws:iam::000000000000:role/ManagedOpenShift-Worker-Role
Operator IAM Roles:
- arn:aws:iam::000000000000:role/my-rosa-cluster-w7i6-openshift-ingress-operator-cloud-
credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-w7i6-openshift-cluster-csi-drivers-ebs-
cloud-credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-w7i6-openshift-cloud-network-config-
controller-cloud-cre
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-machine-api-aws-cloud-
credentials
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-cloud-credential-operator-cloud-
credentia
- arn:aws:iam::000000000000:role/my-rosa-cluster-openshift-image-registry-installer-cloud-
credential
State:                 waiting (Waiting for OIDC configuration)
Private:               No
Created:               Jul 1 2022 22:13:50 UTC
Details Page:
https://console.redhat.com/openshift/details/s/2BMQm8xz8Hq5yEN000000000000
OIDC Endpoint URL:    https://rh-oidc.s3.us-east-
1.amazonaws.com/1t6i760dbum4mqltqh6o000000000000
I: Run the following commands to continue the cluster creation:
rosa create operator-roles --cluster my-rosa-cluster
rosa create oidc-provider --cluster my-rosa-cluster
I: To determine when your cluster is Ready, run 'rosa describe cluster -c my-rosa-cluster'.
I: To watch your cluster installation logs, run 'rosa logs install -c my-rosa-cluster --watch'.

```

**NOTE**

The cluster state will remain as “waiting” until the next two steps are completed.

**15.3.3.4.3. Creating Operator roles**

1. The above step outputs the next commands to run. These roles need to be created *once* for *each* cluster. To create the roles run the following command:

```
rosa create operator-roles --mode manual --cluster <cluster-name>
```

**Example output**

```

I: Run the following commands to create the operator roles:
aws iam create-role \
  --role-name my-rosa-cluster-openshift-image-registry-installer-cloud-credentials \
  --assume-role-policy-document
file://operator_image_registry_installer_cloud_credentials_policy.json \
  --tags Key=rosa_cluster_id,Value=1mkesci269png3tck0000000000000000
Key=rosa_openshift_version,Value=4.8 Key=rosa_role_prefix,Value=
Key=operator_namespace,Value=openshift-image-registry

```

```
Key=operator_name,Value=installer-cloud-credentials
```

```
aws iam attach-role-policy \
  --role-name my-rosa-cluster-openshift-image-registry-installer-cloud-credentials \
  --policy-arn arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-image-
registry-installer-cloud-creden
[...]
```

2. Run each of the **aws** commands.

#### 15.3.3.4.4. Creating the OIDC provider

1. Run the following command to create the OIDC provider:

```
rosa create oidc-provider --mode manual --cluster <cluster-name>
```

2. This displays the **aws** commands that you need to run.

#### Example output

```
I: Run the following commands to create the OIDC provider:
$ aws iam create-open-id-connect-provider \
  --url https://rh-oidc.s3.us-east-1.amazonaws.com/1mkesci269png3tckknhh0rfs2da5fj9 \
  --client-id-list openshift sts.amazonaws.com \
  --thumbprint-list a9d53002e97e00e043244f3d170d000000000000

$ aws iam create-open-id-connect-provider \
  --url https://rh-oidc.s3.us-east-1.amazonaws.com/1mkesci269png3tckknhh0rfs2da5fj9 \
  --client-id-list openshift sts.amazonaws.com \
  --thumbprint-list a9d53002e97e00e043244f3d170d000000000000
```

3. Your cluster will now continue the installation process.

#### 15.3.3.4.5. Checking the installation status

1. Run one of the following commands to check the status of your cluster:

- For a detailed view of the status, run:

```
rosa describe cluster --cluster <cluster-name>
```

- For an abridged view of the status, run:

```
rosa list clusters
```

2. The cluster state will change from "waiting" to "installing" to "ready". This will take about 40 minutes.
3. Once the state changes to "ready" your cluster is installed.

#### 15.3.3.5. Obtaining the Red Hat Hybrid Cloud Console URL

- To obtain the Hybrid Cloud Console URL, run the following command:

```
■
```

```
rosa describe cluster -c <cluster-name> | grep Console
```

The cluster has now been successfully deployed. The next tutorial shows how to create an admin user to be able to use the cluster immediately.

### 15.3.4. Tutorial: Simple UI guide

This page outlines the minimum list of commands to deploy a ROSA cluster using the user interface (UI).



#### NOTE

While this simple deployment works well for a tutorial setting, clusters used in production should be deployed with a more detailed method.

#### 15.3.4.1. Prerequisites

- You have completed the prerequisites in the Setup tutorial.

#### 15.3.4.2. Creating account roles

Run the following command *once* for each AWS account and y-stream OpenShift version:

```
rosa create account-roles --mode auto --yes
```

#### 15.3.4.3. Creating Red Hat OpenShift Cluster Manager roles

1. Create one OpenShift Cluster Manager role for each AWS account by running the following command:

```
rosa create ocm-role --mode auto --admin --yes
```

2. Create one OpenShift Cluster Manager user role for each AWS account by running the following command:

```
rosa create user-role --mode auto --yes
```

3. Use the [OpenShift Cluster Manager](#) to select your AWS account, cluster options, and begin deployment.
4. OpenShift Cluster Manager UI displays cluster status.

Overview Access control Add-ons Cluster history Settings

Installing cluster [Download OC CLI](#)

Cluster creation usually takes 30 to 60 minutes to complete.

Account setup Completed

OIDC and operator roles Completed

Network settings Validating

DNS setup

Cluster installation

[View logs](#)

Details

Cluster ID  
N/A

Type  
ROSA

Region

Status  
Installing

Total vCPU  
0 vCPU

Total memory

### 15.3.5. Tutorial: Detailed UI guide

This tutorial outlines the detailed steps to deploy a Red Hat OpenShift Service on AWS (ROSA) cluster using the Red Hat OpenShift Cluster Manager user interface (UI).

#### 15.3.5.1. Deployment workflow

The overall deployment workflow follows these steps:

1. Create the account wide roles and policies.
2. Associate your AWS account with your Red Hat account.
  - a. Create and link the Red Hat OpenShift Cluster Manager role.
  - b. Create and link the user role.
3. Create the cluster.

Step 1 only needs to be performed the **first time** you are deploying into an AWS account. Step 2 only needs to be performed the **first time** you are using the UI. For successive clusters of the same y-stream version, you only need to create the cluster.

#### 15.3.5.2. Creating account wide roles



#### NOTE

If you already have account roles from an earlier deployment, skip this step. The UI will detect your existing roles after you select an associated AWS account.

If this is the *first time* you are deploying ROSA in this account and you have *not* yet created the account roles, create the account-wide roles and policies, including the Operator policies.

- In your terminal, run the following command to create the account-wide roles:

```
$ rosa create account-roles --mode auto --yes
```

## Example output

```

I: Creating roles using 'arn:aws:iam::000000000000:user/rosa-user'
I: Created role 'ManagedOpenShift-ControlPlane-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-ControlPlane-Role'
I: Created role 'ManagedOpenShift-Worker-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-Worker-Role'
I: Created role 'ManagedOpenShift-Support-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-Support-Role'
I: Created role 'ManagedOpenShift-Installer-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-Installer-Role'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-
machine-api-aws-cloud-credentials'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-
cloud-credential-operator-cloud-crede'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-
image-registry-installer-cloud-creden'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-
ingress-operator-cloud-credentials'
I: Created policy with ARN 'arn:aws:iam::000000000000:policy/ManagedOpenShift-openshift-
cluster-csi-drivers-ebs-cloud-credent'
I: To create a cluster with these roles, run the following command:
rosa create cluster --sts

```

### 15.3.5.3. Associating your AWS account with your Red Hat account

This step tells the OpenShift Cluster Manager what AWS account you want to use when deploying ROSA.



#### NOTE

If you have already associated your AWS accounts, skip this step.

1. Open the Red Hat Hybrid Cloud Console by visiting the [OpenShift Cluster Manager](#) and logging in to your Red Hat account.
2. Click **Create Cluster**.
3. Scroll down to the Red Hat OpenShift Service on AWS (ROSA) row and click **Create Cluster**.

Select an OpenShift cluster type to create

Cloud      Datacenter      Local

**Managed services**  
Create clusters in the cloud using a managed service.

| Offerings                                  | Purchased through   | Get started   |
|--|---------------------|---|
| Red Hat OpenShift Dedicated Trial          | Red Hat             | Available on AWS and GCP <a href="#">Create trial cluster</a>                           |
| >  Red Hat OpenShift Dedicated             | Red Hat             | Available on AWS and GCP <a href="#">Create cluster</a>                                 |
| >  Azure Red Hat OpenShift                 | Microsoft Azure     | Flexible hourly billing <a href="#">Try it on Azure</a>                                 |
| >  Red Hat OpenShift on IBM Cloud          | IBM                 | Flexible hourly billing <a href="#">Try it on IBM</a>                                   |
| >  Red Hat OpenShift Service on AWS (ROSA) | Amazon Web Services | Flexible hourly billing <a href="#">Create cluster</a><br><a href="#">Prerequisites</a> |

4. A dropdown menu appears. Click **With web interface**.

> Red Hat OpenShift Service on AWS (ROSA)      Amazon Web Services      Flexible hourly billing

[View your available AWS accounts and quota](#) →

[Create cluster](#) ▼

- With CLI
- With web interface

5. Under "Select an AWS control plane type," choose **Classic**. Then click **Next**.

Select an AWS control plane type  
Not sure what to choose? [Learn more about AWS control plane types](#)

**Hosted**  
Run an OpenShift cluster where the control plane is decoupled from the data plane, and is treated like a multi-tenant workload on a hosted service cluster. The data plane is on a separate network domain that allows segmentation between management and workload traffic.

- ✓ Control plane resources are hosted in a Red Hat-owned AWS account
- ✓ Better resource utilization with faster cluster creation
- ✓ Lower AWS infrastructure costs
- ✓ Red Hat SRE managed
- ⚠ Compliance certifications available soon
- ⚠ Virtual Private Cloud is required  
To create a ROSA cluster that is hosted by Red Hat, you must be able to create clusters on a VPC. [Learn more about Virtual Private Cloud](#)

**Classic**  
Run an OpenShift cluster where the control plane and data plane are coupled. The control plane is hosted by a dedicated group of physical or virtual nodes and the network stack is shared.

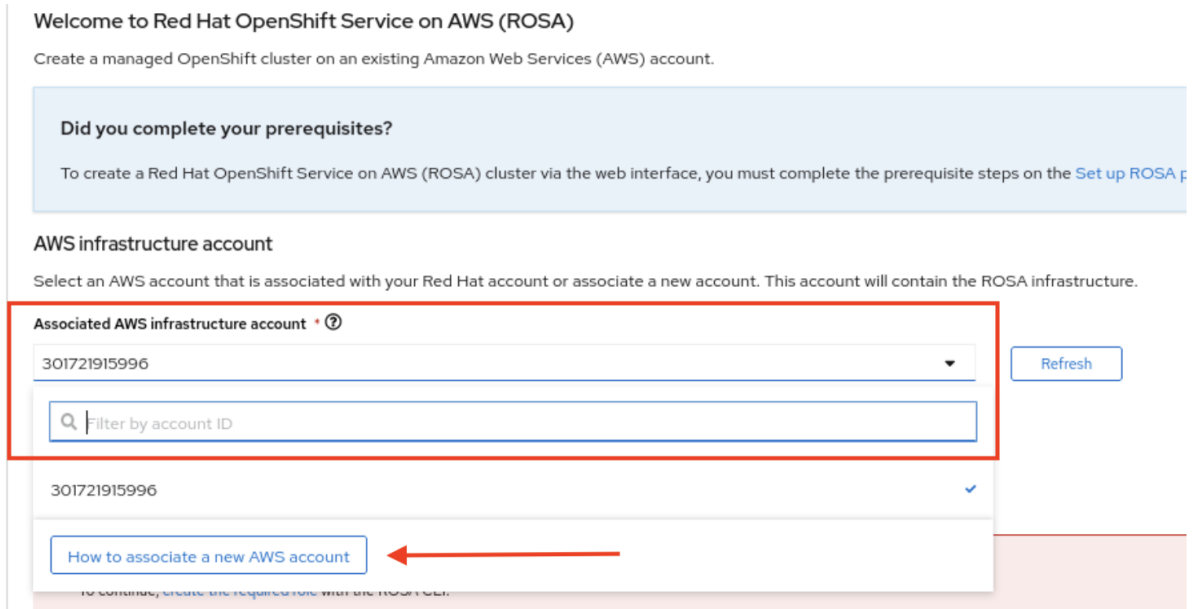
- ✓ Control plane resources are hosted in your own AWS account
- ✓ Full compliance certifications
- ✓ Red Hat SRE managed

[Next](#)   [Back](#)   [Cancel](#)

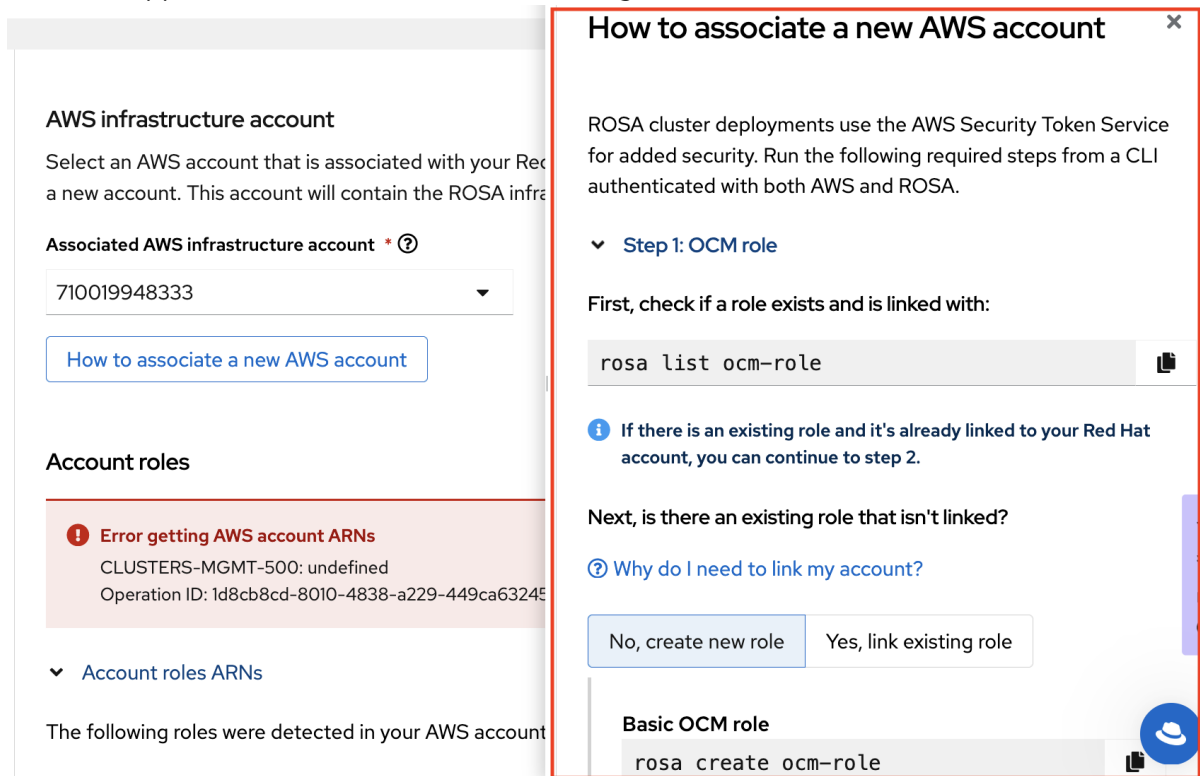
6. Click the dropbox under **Associated AWS infrastructure account** If you have not yet associated any AWS accounts, the dropbox may be empty.

7. Click **How to associate a new AWS account**





8. A sidebar appears with instructions for associating a new AWS account.



#### 15.3.5.4. Creating and associating an OpenShift Cluster Manager role

1. Run the following command to see if an OpenShift Cluster Manager role exists:

```
$ rosa list ocm-role
```

2. The UI displays the commands to create an OpenShift Cluster Manager role with two different levels of permissions:

- **Basic OpenShift Cluster Manager role:** Allows the OpenShift Cluster Manager to have read-only access to the account to check if the roles and policies that are required by ROSA are present before creating a cluster. You will need to manually create the required roles, policies, and OIDC provider using the CLI.

- **Admin OpenShift Cluster Manager role:** Grants the OpenShift Cluster Manager additional permissions to create the required roles, policies, and OIDC provider for ROSA. Using this makes the deployment of a ROSA cluster quicker since the OpenShift Cluster Manager will be able to create the required resources for you.  
To read more about these roles, see the [OpenShift Cluster Manager roles and permissions](#) section of the documentation.

For the purposes of this tutorial, use the **Admin OpenShift Cluster Manager role** for the simplest and quickest approach.

3. Copy the command to create the Admin OpenShift Cluster Manager role from the sidebar or switch to your terminal and enter the following command:

```
$ rosa create ocm-role --mode auto --admin --yes
```

This command creates the OpenShift Cluster Manager role and associates it with your Red Hat account.

### Example output

```
I: Creating ocm role
I: Creating role using 'arn:aws:iam::000000000000:user/rosa-user'
I: Created role 'ManagedOpenShift-OCM-Role-12561000' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-OCM-Role-12561000'
I: Linking OCM role
I: Successfully linked role-arn 'arn:aws:iam::000000000000:role/ManagedOpenShift-OCM-
Role-12561000' with organization account '1MpZfntsZeUdjWHg7XRgP000000'
```

4. Click **Step 2: User role**.

#### 15.3.5.4.1. Other OpenShift Cluster Manager role creation options

- **Manual mode:** If you prefer to run the AWS CLI commands yourself, you can define the mode as **manual** rather than **auto**. The CLI will output the AWS commands and the relevant JSON files are created in the current directory.  
Use the following command to create the OpenShift Cluster Manager role in manual mode:

```
$ rosa create ocm-role --mode manual --admin --yes
```

- **Basic OpenShift Cluster Manager role:** If you prefer that the OpenShift Cluster Manager has read only access to the account, create a basic OpenShift Cluster Manager role. You will then need to manually create the required roles, policies, and OIDC provider using the CLI.  
Use the following command to create a Basic OpenShift Cluster Manager role:

```
$ rosa create ocm-role --mode auto --yes
```

#### 15.3.5.5. Creating an OpenShift Cluster Manager user role

As defined in the [user role documentation](#), the user role needs to be created so that ROSA can verify your AWS identity. This role has no permissions, and it is only used to create a trust relationship between the installation program account and your OpenShift Cluster Manager role resources.

1. Check if a user role already exists by running the following command:

```
■
```

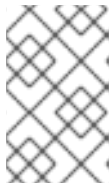
```
$ rosa list user-role
```

- Run the following command to create the user role and to link it to your Red Hat account:

```
$ rosa create user-role --mode auto --yes
```

### Example output

```
I: Creating User role
I: Creating ocm user role using 'arn:aws:iam::000000000000:user/rosa-user'
I: Created role 'ManagedOpenShift-User-rosa-user-Role' with ARN
'arn:aws:iam::000000000000:role/ManagedOpenShift-User-rosa-user-Role'
I: Linking User role
I: Successfully linked role ARN 'arn:aws:iam::000000000000:role/ManagedOpenShift-User-
rosa-user-Role' with account '1rbOQez0z5j1YollnhcXY000000'
```



### NOTE

As before, you can define **--mode manual** if you'd prefer to run the AWS CLI commands yourself. The CLI outputs the AWS commands and the relevant JSON files are created in the current directory. Make sure to link the role.

- Click **Step 3: Account roles**

#### 15.3.5.6. Creating account roles

- Create your account roles by running the following command:

```
$ rosa create account-roles --mode auto
```

- Click **OK** to close the sidebar.

#### 15.3.5.7. Confirming successful account association

- You should now see your AWS account in the **Associated AWS infrastructure account** dropdown menu. If you see your account, account association was successful.
- Select the account.
- You will see the account role ARNs populated below.

### Account roles

▼ [Account roles ARNs](#)

The following roles were detected in your AWS account. [Learn more about account roles](#)

[Refresh ARNs](#)

**Installer role** \*

arn:aws:iam:::role/ManagedOpenShift-Installer-Role

**Support role** \*

arn:aws:iam:::role/ManagedOpenShift-Support-Role

**Worker role** \*

arn:aws:iam:::role/ManagedOpenShift-Worker-Role

**Control plane role** \*

arn:aws:iam:::role/ManagedOpenShift-ControlPlane-Role

The selected account-wide roles are compatible with OpenShift version 4.10 and earlier.

[Next](#) [Back](#) [Cancel](#)

4. Click **Next**.

### 15.3.5.8. Creating the cluster

1. For the purposes of this tutorial make the following selections:

#### Cluster settings

- Cluster name: **<pick a name\>**
- Version: **<select latest version\>**
- Region: **<select region\>**
- Availability: **Single zone**
- Enable user workload monitoring: **leave checked**
- Enable additional etcd encryption: **leave unchecked**
- Encrypt persistent volumes with customer keys: **leave unchecked**

2. Click **Next**.

3. Leave the default settings on for the machine pool:

#### Default machine pool settings

- Compute node instance type: **m5.xlarge - 4 vCPU 16 GiB RAM**
- Enable autoscaling: **unchecked**
- Compute node count: **2**
- Leave node labels blank

4. Click **Next**.

#### 15.3.5.8.1. Networking

1. Leave all the default values for configuration.
2. Click **Next**.
3. Leave all the default values for CIDR ranges.
4. Click **Next**.

#### 15.3.5.8.2. Cluster roles and policies

For this tutorial, leave **Auto** selected. It will make the cluster deployment process simpler and quicker.



#### NOTE

If you selected a **Basic OpenShift Cluster Manager role** earlier, you can only use manual mode. You must manually create the operator roles and OIDC provider. See the "Basic OpenShift Cluster Manager role" section below after you have completed the "Cluster updates" section and started cluster creation.

#### 15.3.5.8.3. Cluster updates

- Leave all the options at default in this section.

#### 15.3.5.8.4. Reviewing and creating your cluster

1. Review the content for the cluster configuration.
2. Click **Create cluster**.

#### 15.3.5.8.5. Monitoring the installation progress

- Stay on the current page to monitor the installation progress. It should take about 40 minutes.

Overview Access control Settings

**Installing cluster** [Cancel cluster creation](#) [Download OC CLI](#)

[View logs](#)

**Details**

|                                    |   |
|------------------------------------|---|
| <b>Cluster ID</b><br>N/A           | <b>Status</b><br>Installing                         |
| <b>Type</b><br>ROSA                | <b>Total vCPU</b><br>0 vCPU                         |
| <b>Region</b><br>us-east-1         | <b>Total memory</b><br>0 B                          |
| <b>Availability</b><br>Single zone | <b>Nodes (actual/desired)</b><br>Control plane: 0/3 |

### 15.3.5.9. Basic OpenShift Cluster Manager Role



#### NOTE

If you created an **Admin OpenShift Cluster Manager role** as directed above **ignore** this entire section. The OpenShift Cluster Manager will create the resources for you.

If you created a **Basic OpenShift Cluster Manager role** earlier, you will need to manually create two more elements before cluster installation can continue:

- Operator roles
- OIDC provider

#### 15.3.5.9.1. Creating Operator roles

1. A pop up window will show you the commands to run.

### Action required to continue installation ✕

You must create the **operator roles** and **OIDC provider** to complete cluster installation.

Use one of the following methods:

ROSA CLI

AWS CLI

Copy and run the following commands:

```
rosa create operator-roles --interactive -c ssssssss
```

```
rosa create oidc-provider --interactive -c ssssssss
```

The options above will be available until the operator roles and OIDC provider are detected.

- Run the commands from the window in your terminal to launch interactive mode. Or, for simplicity, run the following command to create the Operator roles:

```
$ rosa create operator-roles --mode auto --cluster <cluster-name> --yes
```

#### Example output

```
I: Creating roles using 'arn:aws:iam::000000000000:user/rosauser'
I: Created role 'rosacluster-b736-openshift-ingress-operator-cloud-credentials' with ARN
'arn:aws:iam::000000000000:role/rosacluster-b736-openshift-ingress-operator-cloud-
credentials'
I: Created role 'rosacluster-b736-openshift-cluster-csi-drivers-ebs-cloud-credent' with ARN
'arn:aws:iam::000000000000:role/rosacluster-b736-openshift-cluster-csi-drivers-ebs-cloud-
credent'
I: Created role 'rosacluster-b736-openshift-cloud-network-config-controller-cloud' with ARN
'arn:aws:iam::000000000000:role/rosacluster-b736-openshift-cloud-network-config-controller-
cloud'
I: Created role 'rosacluster-b736-openshift-machine-api-aws-cloud-credentials' with ARN
'arn:aws:iam::000000000000:role/rosacluster-b736-openshift-machine-api-aws-cloud-
credentials'
I: Created role 'rosacluster-b736-openshift-cloud-credential-operator-cloud-crede' with ARN
'arn:aws:iam::000000000000:role/rosacluster-b736-openshift-cloud-credential-operator-
cloud-crede'
I: Created role 'rosacluster-b736-openshift-image-registry-installer-cloud-creden' with ARN
'arn:aws:iam::000000000000:role/rosacluster-b736-openshift-image-registry-installer-cloud-
creden'
```

#### 15.3.5.9.2. Creating the OIDC provider

- In your terminal, run the following command to create the OIDC provider:

```
$ rosa create oidc-provider --mode auto --cluster <cluster-name> --yes
```

#### Example output

```
I: Creating OIDC provider using 'arn:aws:iam::000000000000:user/rosauser'
I: Created OIDC provider with ARN 'arn:aws:iam::000000000000:oidc-provider/rh-oidc.s3.us-east-1.amazonaws.com/1tt4kvr2kha2rgs8gjfvf0000000000'
```

## 15.4. TUTORIAL: CREATING AN ADMIN USER

Creating an administration (admin) user allows you to access your cluster quickly. Follow these steps to create an admin user.



### NOTE

An admin user works well in this tutorial setting. For actual deployment, use a [formal identity provider](#) to access the cluster and grant the user admin privileges.

1. Run the following command to create the admin user:

```
rosa create admin --cluster=<cluster-name>
```

### Example output

```
W: It is recommended to add an identity provider to login to this cluster. See 'rosa create idp -
-help' for more information.
I: Admin account has been added to cluster 'my-rosa-cluster'. It may take up to a minute for
the account to become active.
I: To login, run the following command:
oc login https://api.my-rosa-cluster.abcd.p1.openshiftapps.com:6443 \
--username cluster-admin \
--password FWGYL-2mkJI-00000-00000
```

2. Copy the log in command returned to you in the previous step and paste it into your terminal. This will log you in to the cluster using the CLI so you can start using the cluster.

```
$ oc login https://api.my-rosa-cluster.abcd.p1.openshiftapps.com:6443 \
> --username cluster-admin \
> --password FWGYL-2mkJI-00000-00000
```

### Example output

```
Login successful.

You have access to 79 projects, the list has been suppressed. You can list all projects with '
projects'

Using project "default".
```

3. To check that you are logged in as the admin user, run one of the following commands:

- Option 1:

```
$ oc whoami
```

### Example output



```
cluster-admin
```

- Option 2:

```
oc get all -n openshift-apiserver
```

Only an admin user can run this command without errors.

4. You can now use the cluster as an admin user, which will suffice for this tutorial. For actual deployment, it is highly recommended to set up an identity provider, which is explained in the [next tutorial](#).

## 15.5. TUTORIAL: SETTING UP AN IDENTITY PROVIDER

To log in to your cluster, set up an identity provider (IDP). This tutorial uses GitHub as an example IDP. See the full list of [IDPs supported by ROSA](#).

- To view all IDP options, run the following command:

```
rosa create idp --help
```

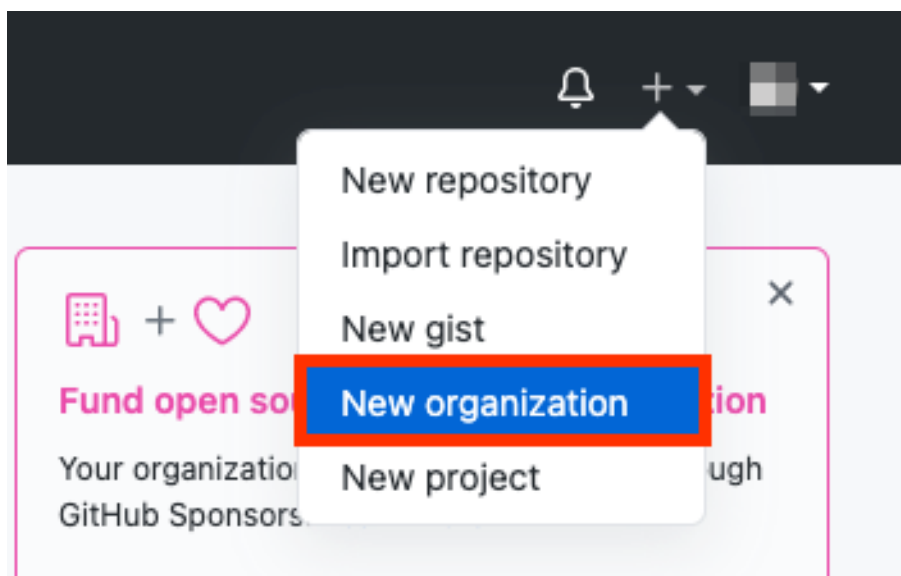
### 15.5.1. Setting up an IDP with GitHub

1. Log in to your GitHub account.
2. Create a new GitHub organization where you are an administrator.

#### TIP

If you are already an administrator in an existing organization and you want to use that organization, skip to step 9.

Click the + icon, then click **New Organization**.



3. Choose the most applicable plan for your situation or click **Join for free**.

4. Enter an organization account name, an email, and whether it is a personal or business account. Then, click **Next**.

Tell us about your organization

# Set up your team

## Organization account name \*



This will be the name of your account on GitHub.  
Your URL will be: <https://github.com/my-rosa-cluster>.

## Contact email \*



## This organization belongs to: \*

**My personal account**

i.e., ██████████

**A business or institution**

For example: GitHub, Inc., Example Institute, American Red Cross

**Next**

By creating an account, you agree to the [Terms of Service](#). For more information about GitHub's privacy practices, see the [GitHub Privacy Statement](#). We'll occasionally send you account-related emails.

5. **Optional:** Add the GitHub IDs of other users to grant additional access to your ROSA cluster. You can also add them later.
6. Click **Complete Setup**.
7. **Optional:** Enter the requested information on the following page.
8. Click **Submit**.
9. Go back to the terminal and enter the following command to set up the GitHub IDP:

```
rosa create idp --cluster=<cluster name> --interactive
```

10. Enter the following values:

```
Type of identity provider: github
Identity Provider Name: <IDP-name>
Restrict to members of: organizations
```

GitHub organizations: <organization-account-name>

- The CLI will provide you with a link. Copy and paste the link into a browser and press **Enter**. This will fill the required information to register this application for OAuth. You do not need to modify any of the information.

```
I: Interactive mode enabled.
Any optional fields can be left empty and a default will be selected.
? Type of identity provider: github
[?] Identity provider name: rosa-github
? Restrict to members of: organizations
[?] GitHub organizations: my-rosa-cluster
? To use GitHub as an identity provider, you must first register the application:
  - Open the following URL:
    https://github.com/organizations/my-rosa-cluster/settings/applications/new?oauth_application%5Bcallback_url%5D=https%3A%2F%2Fconsole-openshift-console.apps.x.p1
  - Click on 'Register application'
? Client ID: [?] [? for help]
```

- Click **Register application**.

## Register a new OAuth application

### Application name \*

Something users will recognize and trust.

### Homepage URL \*

The full URL to your application homepage.

### Application description

This is displayed to all users of your application.

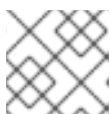
### Authorization callback URL \*

Your application's callback URL. Read our [OAuth documentation](#) for more information.

**Register application**

Cancel


- The next page displays a **Client ID**. Copy the ID and paste it in the terminal where it asks for **Client ID**.



### NOTE

Do not close the tab.

- The CLI will ask for a **Client Secret**. Go back in your browser and click **Generate a new client secret**.

 **my-rosa-cluster** owns this application. [Transfer ownership](#)

---

You can list your application in the [GitHub Marketplace](#) so that other users can discover it. [List this application in the Marketplace](#)

---

**0 users** [Revoke all user tokens](#)

---

**Client ID**  
caa31


---

**Client secrets** [Generate a new client secret](#)

You need a client secret to authenticate as the application to the API.

---

**Application logo**



Drag & drop

[Upload new logo](#)

You can also drag and drop a picture from your computer.

**Application name \***

15. A secret is generated for you. Copy your secret because it will never be visible again.
16. Paste your secret into the terminal and press **Enter**.
17. Leave **GitHub Enterprise Hostname** blank.
18. Select **claim**.
19. Wait approximately 1 minute for the IDP to be created and the configuration to land on your cluster.

```
I: Interactive mode enabled.
Any optional fields can be left empty and a default will be selected.
? Type of identity provider: github
? Identity provider name: rosa-github
? Restrict to members of: organizations
? GitHub organizations: my-rosa-cluster
? To use GitHub as an identity provider, you must first register the application:
  - Open the following URL:
    https://github.com/organizations/my-rosa-cluster/settings/applications/new?oauth_application%5Bcallback_url%5D=
https%3A%2F%2Foauth-openshift.apps.█.p1.openshiftapps.com%2Foauth2callback%2Frosa-github&oauth_ap
plication%5Bname%5D=█&oauth_application%5Burl%5D=https%3A%2F%2Fconsole-openshift-console.apps.█.
█.p1.openshiftapps.com
  - Click on 'Register application'
? Client ID: caa31
? Client Secret: [? for help] *****
? GitHub Enterprise Hostname (optional):
? Mapping method: claim
I: Configuring IDP for cluster '█'
I: Identity Provider 'rosa-github' has been created.
It will take up to 1 minute for this configuration to be enabled.
To add cluster administrators, see 'rosa create user --help'.
To login into the console, open https://console-openshift-console.apps.█.p1.openshiftapps.com
and click on rosa-github.
```

20. Copy the returned link and paste it into your browser. The new IDP should be available under your chosen name. Click your IDP and use your GitHub credentials to access the cluster.

Log in with...

Cluster-Admin

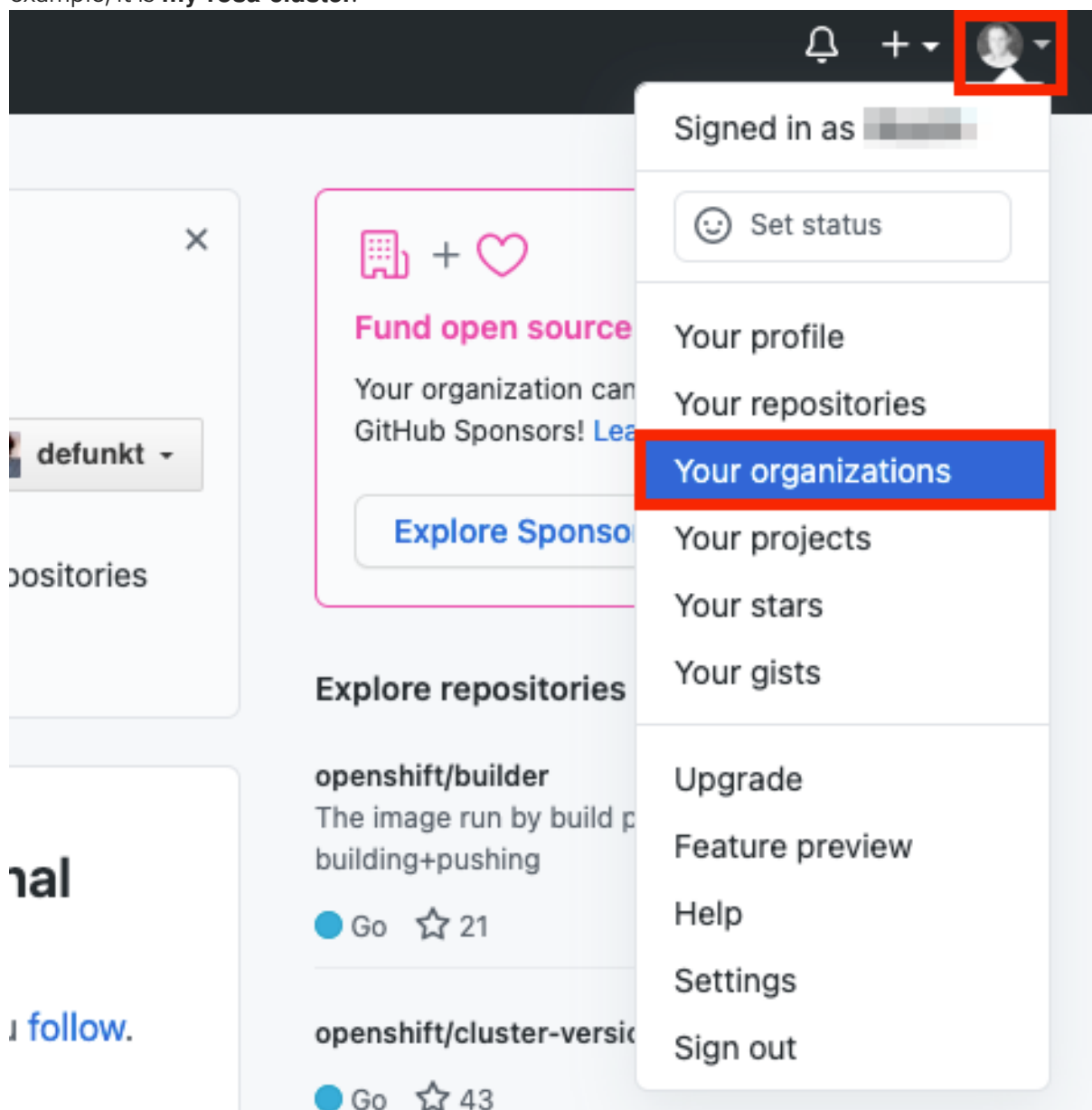
rosa-github



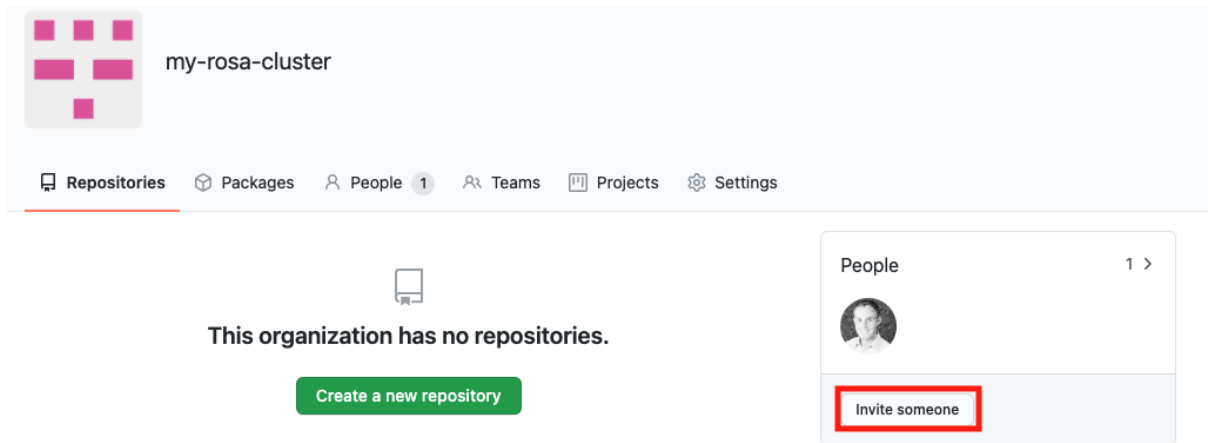
### 15.5.2. Granting other users access to the cluster

To grant access to other cluster user you will need to add their GitHub user ID to the GitHub organization used for this cluster.

1. In GitHub, go to the **Your organizations** page.
2. Click your **profile icon**, then **Your organizations**. Then click `<your-organization-name>`. In our example, it is **my-rosa-cluster**.



3. Click **Invite someone**.



4. Enter the GitHub ID of the new user, select the correct user, and click **Invite**.
5. Once the new user accepts the invitation, they will be able to log in to the ROSA cluster using the Hybrid Cloud Console link and their GitHub credentials.

## 15.6. TUTORIAL: GRANTING ADMIN PRIVILEGES

Administration (admin) privileges are not automatically granted to users that you add to your cluster. If you want to grant admin-level privileges to certain users, you will need to manually grant them to each user. You can grant admin privileges from either the ROSA command line interface (CLI) or the Red Hat OpenShift Cluster Manager web user interface (UI).

Red Hat offers two types of admin privileges:

- **cluster-admin:** **cluster-admin** privileges give the admin user full privileges within the cluster.
- **dedicated-admin:** **dedicated-admin** privileges allow the admin user to complete most administrative tasks with certain limitations to prevent cluster damage. It is best practice to use **dedicated-admin** when elevated privileges are needed.

For more information on admin privileges, see the [administering a cluster](#) documentation.

### 15.6.1. Using the ROSA CLI

1. Assuming you are the user who created the cluster, run one of the following commands to grant admin privileges:

- For **cluster-admin**:

```
$ rosa grant user cluster-admin --user <idp_user_name> --cluster=<cluster-name>
```

- For **dedicated-admin**:

```
$ rosa grant user dedicated-admin --user <idp_user_name> --cluster=<cluster-name>
```

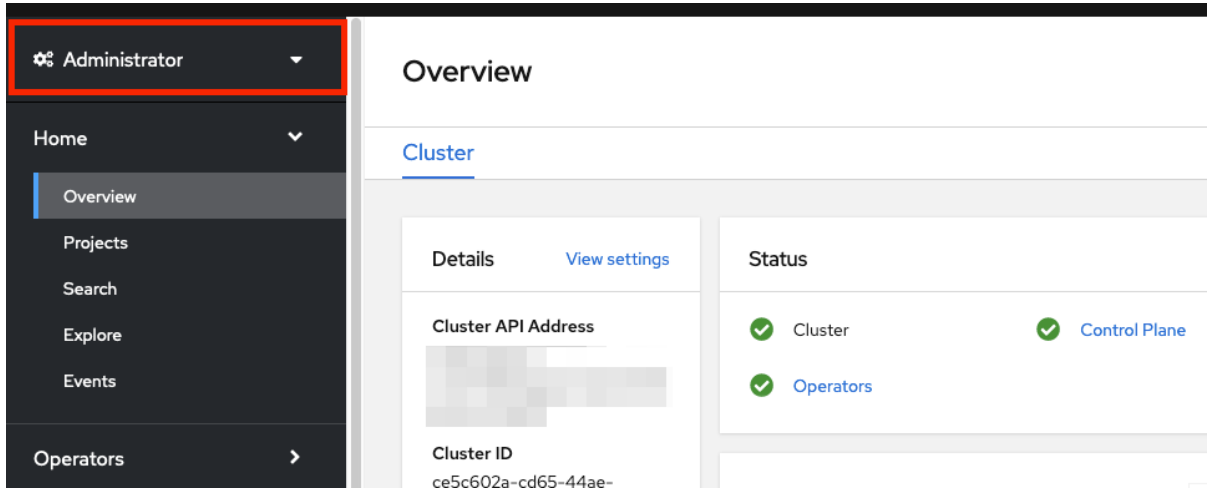
2. Verify that the admin privileges were added by running the following command:

```
$ rosa list users --cluster=<cluster-name>
```

**Example output**

```
$ rosa list users --cluster=my-rosa-cluster
ID          GROUPS
<idp_user_name> cluster-admins
```

- If you are currently logged into the Red Hat Hybrid Cloud Console, log out of the console and log back in to the cluster to see a new perspective with the "Administrator Panel". You might need an incognito or private window.

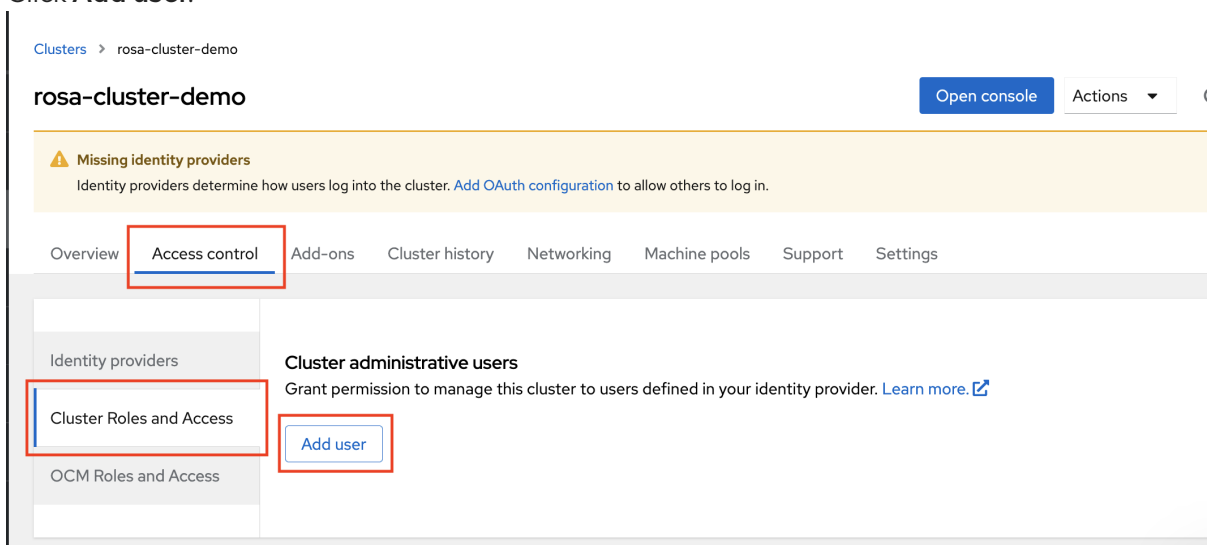


- You can also test that admin privileges were added to your account by running the following command. Only a **cluster-admin** users can run this command without errors.

```
$ oc get all -n openshift-apiserver
```

## 15.6.2. Using the Red Hat OpenShift Cluster Manager UI

- Log in to the [OpenShift Cluster Manager](#).
- Select your cluster.
- Click the **Access Control** tab.
- Click the **Cluster roles and Access** tab in the sidebar.
- Click **Add user**.



- On the pop-up screen, enter the user ID.

7. Select whether you want to grant the user **cluster-admins** or **dedicated-admins** privileges.

**Add cluster user** ×

User ID \*  
sample-user ←

Group

dedicated-admins  
Grants standard administrative privileges for OpenShift Dedicated. Users can perform administrative actions listed in the [documentation](#).

cluster-admins  
Gives users full administrative access to the cluster. This is the highest level of privilege available to users. It should be granted with extreme care, because it is possible with this level of access to get the cluster into an unsupported state.

**Add user** Cancel

## 15.7. TUTORIAL: ACCESSING YOUR CLUSTER

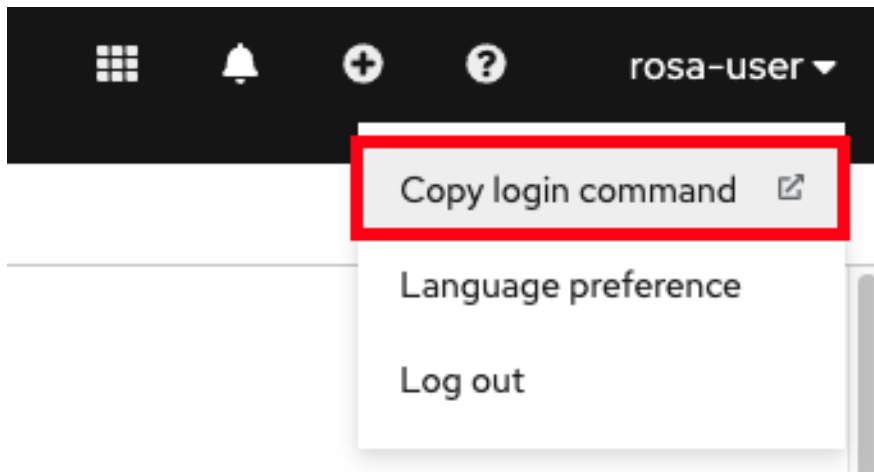
You can connect to your cluster using the command line interface (CLI) or the Red Hat Hybrid Cloud Console user interface (UI).

### 15.7.1. Accessing your cluster using the CLI

To access the cluster using the CLI, you must have the **oc** CLI installed. If you are following the tutorials, you already installed the **oc** CLI.

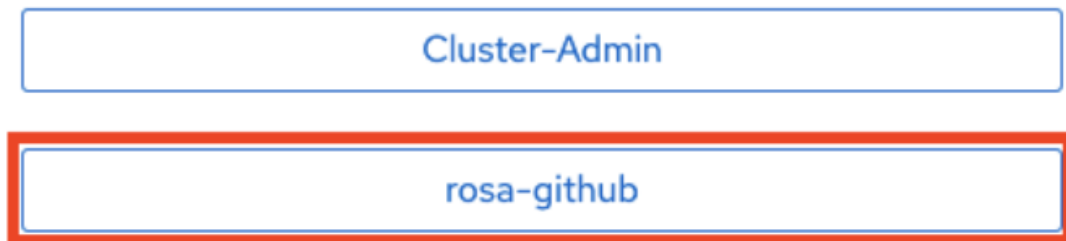
1. Log in to the [OpenShift Cluster Manager](#).
2. Click your username in the top right corner.
3. Click **Copy Login Command**.





- This opens a new tab with a choice of identity providers (IDPs). Click the IDP you want to use. For example, "rosa-github".

## Log in with...



- A new tab opens. Click **Display token**.
- Run the following command in your terminal:

```
$ oc login --token=sha256~GBAfS4JQ0t1UTKYHbWAK6OUWGUkdMGz000000000000 --
server=https://api.my-rosa-cluster.abcd.p1.openshiftapps.com:6443
```

### Example output

```
Logged into "https://api.my-rosa-cluster.abcd.p1.openshiftapps.com:6443" as "rosa-user"
using the token provided.
```

```
You have access to 79 projects, the list has been suppressed. You can list all projects with '
projects'
```

```
Using project "default".
```

- Confirm that you are logged in by running the following command:

```
$ oc whoami
```

### Example output

```
rosa-user
```

- You can now access your cluster.

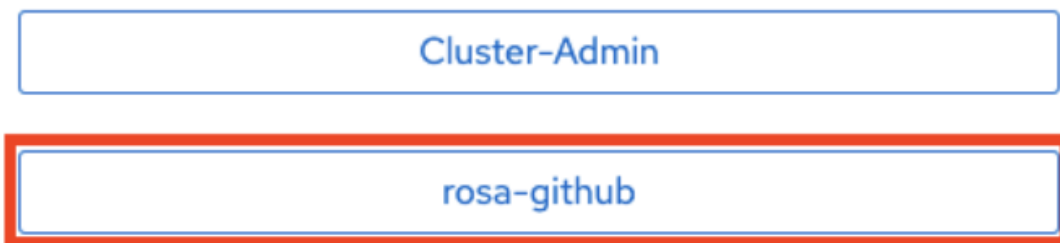
## 15.7.2. Accessing the cluster via the Hybrid Cloud Console

- Log in to the [OpenShift Cluster Manager](#).
  - To retrieve the Hybrid Cloud Console URL run:

```
rosa describe cluster -c <cluster-name> | grep Console
```

- Click your IDP. For example, "rosa-github".

# Log in with...



- Enter your user credentials.
- You should be logged in. If you are following the tutorials, you will be a cluster-admin and should see the Hybrid Cloud Console webpage with the **Administrator** panel visible.

## 15.8. TUTORIAL: MANAGING WORKER NODES

In Red Hat OpenShift Service on AWS (ROSA), changing aspects of your worker nodes is performed through the use of machine pools. A machine pool allows users to manage many machines as a single entity. Every ROSA cluster has a default machine pool that is created when the cluster is created. For more information, see the [machine pool](#) documentation.

## 15.8.1. Creating a machine pool

You can create a machine pool with either the command line interface (CLI) or the user interface (UI).

### 15.8.1.1. Creating a machine pool with the CLI

1. Run the following command:

```
rosa create machinepool --cluster=<cluster-name> --name=<machinepool-name> --replicas=
<number-nodes>
```

#### Example input

```
$ rosa create machinepool --cluster=my-rosa-cluster --name=new-mp
--replicas=2
```

#### Example output

```
I: Machine pool 'new-mp' created successfully on cluster 'my-rosa-cluster'
I: To view all machine pools, run 'rosa list machinepools -c my-rosa-cluster'
```

2. **Optional:** Add node labels or taints to specific nodes in a new machine pool by running the following command:

```
rosa create machinepool --cluster=<cluster-name> --name=<machinepool-name> --replicas=
<number-nodes> --labels='<key=pair>'
```

#### Example input

```
$ rosa create machinepool --cluster=my-rosa-cluster --name=db-nodes-mp --replicas=2 --
labels='app=db','tier=backend'
```

#### Example output

```
I: Machine pool 'db-nodes-mp' created successfully on cluster 'my-rosa-cluster'
```

This creates an additional 2 nodes that can be managed as a unit and also assigns them the labels shown.

3. Run the following command to confirm machine pool creation and the assigned labels:

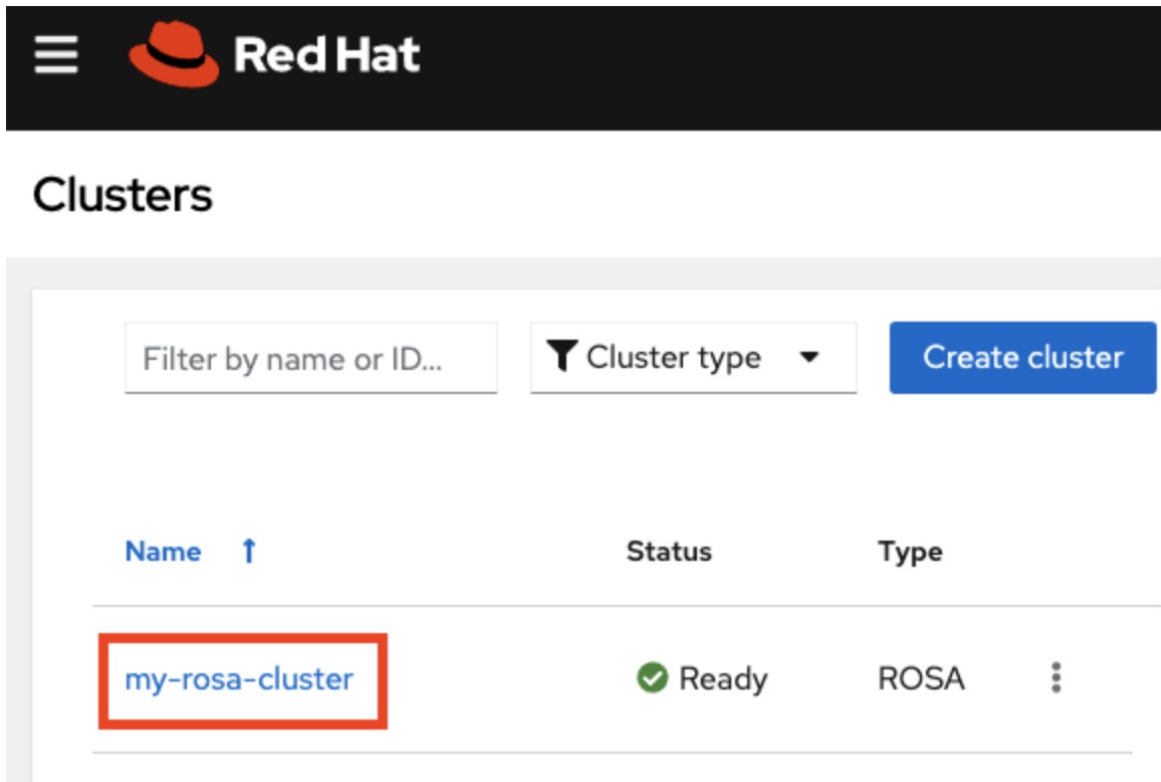
```
rosa list machinepools --cluster=<cluster-name>
```

#### Example output

| ID      | AUTOSCALING | REPLICAS | INSTANCE TYPE | LABELS     | TAINTS |
|---------|-------------|----------|---------------|------------|--------|
| Default | No          | 2        | m5.xlarge     | us-east-1a |        |

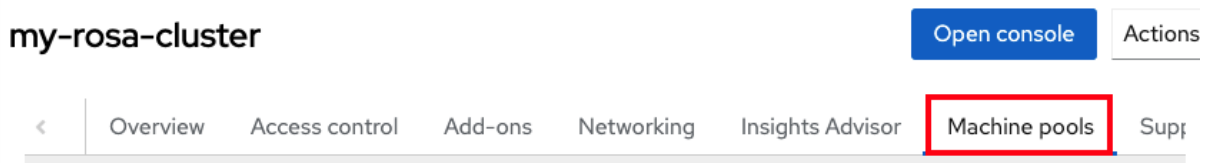
### 15.8.1.2. Creating a machine pool with the UI

1. Log in to the [OpenShift Cluster Manager](#) and click your cluster.



The screenshot shows the Red Hat OpenShift Cluster Manager interface. At the top, there is a navigation bar with the Red Hat logo and a hamburger menu icon. Below the navigation bar, the word "Clusters" is displayed. The main content area contains a search bar labeled "Filter by name or ID...", a dropdown menu for "Cluster type", and a blue "Create cluster" button. Below these controls is a table with columns for "Name", "Status", and "Type". The table contains one row with the cluster name "my-rosa-cluster", a status of "Ready" with a green checkmark, and a type of "ROSA". The "my-rosa-cluster" text in the table is highlighted with a red rectangular box.

2. Click **Machine pools**.



The screenshot shows the OpenShift Cluster Manager interface for a specific cluster named "my-rosa-cluster". At the top right, there is a blue "Open console" button and an "Actions" button. Below the cluster name, there is a horizontal navigation bar with several tabs: "Overview", "Access control", "Add-ons", "Networking", "Insights Advisor", "Machine pools", and "Support". The "Machine pools" tab is highlighted with a red rectangular box.

3. Click **Add machine pool**.
4. Enter the desired configuration.

### TIP

You can also expand the **Edit node labels and taints** section to add node labels and taints to the nodes in the machine pool.

## Add machine pool ×

A machine pool is a group of machines that are all clones of the same configuration, that can be used on demand by an application running on a pod.

**Machine pool name \***

**Compute node instance type \* ?**

### Scaling

**Enable autoscaling ?**

Autoscaling automatically adds and removes worker (compute) nodes from the cluster based on resource requirements.

**Compute node count \* ?**

**Root disk size \* ?**

GiB

5. You will see the new machine pool you created.

Add machine pool

| Machine pool                         | Instance type | Availability zones | Node co... | Autoscaling |
|--------------------------------------|---------------|--------------------|------------|-------------|
| Default                              | m5.xlarge     | us-west-2a         | 2          | Disabled    |
| new-mp                               | m5.xlarge     | us-west-2a         | 2          | Disabled    |
| <input type="checkbox"/> db-nodes-mp | m5.xlarge     | us-west-2a         | 2          | Disabled    |

Labels

app = db tier = backend

## 15.8.2. Scaling worker nodes

Edit a machine pool to scale the number of worker nodes in that specific machine pool. You can use either the CLI or the UI to scale worker nodes.

### 15.8.2.1. Scaling worker nodes using the CLI

1. Run the following command to see the default machine pool that is created with each cluster:

```
rosa list machinepools --cluster=<cluster-name>
```

#### Example output

```
ID      AUTOSCALING REPLICAS INSTANCE TYPE LABELS      TAINTS
AVAILABILITY ZONES
Default No        2      m5.xlarge           us-east-1a
```

2. To scale the default machine pool out to a different number of nodes, run the following command:

```
rosa edit machinepool --cluster=<cluster-name> --replicas=<number-nodes> <machinepool-name>
```

#### Example input

```
rosa edit machinepool --cluster=my-rosa-cluster --replicas 3 Default
```

3. Run the following command to confirm that the machine pool has scaled:

```
rosa describe cluster --cluster=<cluster-name> | grep Compute
```

#### Example input

■

```
$ rosa describe cluster --cluster=my-rosa-cluster | grep Compute
```

### Example output

```
- Compute:          3 (m5.xlarge)
```

#### 15.8.2.2. Scaling worker nodes using the UI

1. Click the three dots to the right of the machine pool you want to edit.
2. Click **Edit**.
3. Enter the desired number of nodes, and click **Save**.
4. Confirm that the cluster has scaled by selecting the cluster, clicking the **Overview** tab, and scrolling to **Compute listing**. The compute listing should equal the scaled nodes. For example, 3/3.

The screenshot displays the Red Hat OpenShift Cluster Manager interface. On the left is a dark sidebar with navigation options: Clusters, Subscriptions, Overview, Support Cases, Cluster Manager Feedback, Red Hat Marketplace, and Documentation. The main content area has a top navigation bar with tabs: Overview (selected), Access control, Add-ons, Networking, Machine pools, Support, and Update set. Below the tabs, there are two circular gauges for 'Resource usage': 'vCPU' at 14.59% of 28 Cores used, and 'Memory' at 23.83% of 107.26 GiB used. Below these is a 'Details' section with the following information:

|   |   |
|---|---|
| <b>Cluster ID</b><br>ce5c602a-cd65-44ae-a888-4e8d641bb3b3 | <b>Status</b><br>✔ Ready  |
| <b>Type</b><br>ROSA                                       | <b>Total vCPU</b><br>28 vCPU  |
| <b>Location</b><br>US East, N. Virginia                   | <b>Total memory</b><br>107.26 GiB   |
| <b>Provider</b><br>AWS                                    | <b>Nodes (actual/desired)</b><br>Master: 3/3<br>Infra: 2/2<br><b>Compute: 3/3</b> |
| <b>Availability</b><br>Single                             |   |

#### 15.8.2.3. Adding node labels

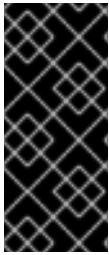
1. Use the following command to add node labels:

```
rosa edit machinepool --cluster=<cluster-name> --replicas=<number-nodes> --  
labels='key=value' <machinepool-name>
```

### Example input

```
rosa edit machinepool --cluster=my-rosa-cluster --replicas=2 --labels 'foo=bar','baz=one'  
new-mp
```

This adds 2 labels to the new machine pool.



### IMPORTANT

This command replaces all machine pool configurations with the newly defined configuration. If you want to add another label **and** keep the old label, you must state both the new and preexisting the label. Otherwise the command will replace all preexisting labels with the one you wanted to add. Similarly, if you want to delete a label, run the command and state the ones you want, excluding the one you want to delete.

### 15.8.3. Mixing node types

You can also mix different worker node machine types in the same cluster by using new machine pools. You cannot change the node type of a machine pool once it is created, but you can create a new machine pool with different nodes by adding the **--instance-type** flag.

1. For example, to change the database nodes to a different node type, run the following command:

```
rosa create machinepool --cluster=<cluster-name> --name=<mp-name> --replicas=<number-nodes> --labels='<key=pair>' --instance-type=<type>
```

#### Example input

```
rosa create machinepool --cluster=my-rosa-cluster --name=db-nodes-large-mp --replicas=2 --labels='app=db','tier=backend' --instance-type=m5.2xlarge
```

2. To see all the [instance types available](#), run the following command:

```
rosa list instance-types
```

3. To make step-by-step changes, use the **--interactive** flag:

```
rosa create machinepool -c <cluster-name> --interactive
```

```
[? Machine pool name: large-nodes-pool
[? Enable autoscaling (optional): No
[? Replicas: 3
? Instance type: [Use arrows to move, type to filter, ? for more help]
> m5.xlarge
  r5.xlarge
  r5.2xlarge
  m5.2xlarge
  c5.2xlarge
  r5.4xlarge
  m5.4xlarge
```

4. Run the following command to list the machine pools and see the new, larger instance type:

```
rosa list machinepools -c <cluster-name>
```



| ID                 | AUTOSCALING | REPLICAS | INSTANCE TYPE | TAINTS               | AVAILABILITY ZONES |
|--------------------|-------------|----------|---------------|----------------------|--------------------|
| default            | No          | 3        | m5.xlarge     |                      | us-east-1a         |
| db-nodes-large--mp | No          | 2        | m5.2xlarge    | app=db, tier=backend | us-east-1a         |
| db-nodes-mp        | No          | 2        | m5.xlarge     | app=db, tier=backend | us-east-1a         |

## 15.9. TUTORIAL: AUTOSCALING

The [cluster autoscaler](#) adds or removes worker nodes from a cluster based on pod resources.

The cluster autoscaler increases the size of the cluster when:

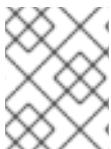
- Pods fail to schedule on the current nodes due to insufficient resources.
- Another node is necessary to meet deployment needs.

The cluster autoscaler does not increase the cluster resources beyond the limits that you specify.

The cluster autoscaler decreases the size of the cluster when:

- Some nodes are consistently not needed for a significant period. For example, when a node has low resource use and all of its important pods can fit on other nodes.

### 15.9.1. Enabling autoscaling for an existing machine pool using the CLI



#### NOTE

Cluster autoscaling can be enabled at cluster creation and when creating a new machine pool by using the **--enable-autoscaling** option.

1. Autoscaling is set based on machine pool availability. To find out which machine pools are available for autoscaling, run the following command:

```
$ rosa list machinepools -c <cluster-name>
```

#### Example output

```
ID      AUTOSCALING  REPLICAS  INSTANCE TYPE  LABELS  TAINTS
AVAILABILITY ZONES
Default No          2         m5.xlarge      us-east-1a
```

2. Run the following command to add autoscaling to an available machine pool:

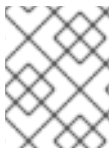
```
$ rosa edit machinepool -c <cluster-name> --enable-autoscaling <machinepool-name> --min-replicas=<num> --max-replicas=<num>
```

#### Example input

```
$ rosa edit machinepool -c my-rosa-cluster --enable-autoscaling Default --min-replicas=2 --max-replicas=4
```

The above command creates an autoscaler for the worker nodes that scales between 2 and 4 nodes depending on the resources.

## 15.9.2. Enabling autoscaling for an existing machine pool using the UI



### NOTE

Cluster autoscaling can be enabled at cluster creation by checking the **Enable autoscaling** checkbox when creating machine pools.

1. Go to the **Machine pools** tab and click the three dots in the right..
2. Click **Scale**, then **Enable autoscaling**.
3. Run the following command to confirm that autoscaling was added:

```
$ rosa list machinepools -c <cluster-name>
```

### Example output

```
ID      AUTOSCALING REPLICAS INSTANCE TYPE LABELS  TAINTS
AVAILABILITY ZONES
Default Yes      2-4    m5.xlarge          us-east-1a
```

## 15.10. TUTORIAL: UPGRADING YOUR CLUSTER

Red Hat OpenShift Service on AWS (ROSA) executes all cluster upgrades as part of the managed service. You do not need to run any commands or make changes to the cluster. You can schedule the upgrades at a convenient time.

Ways to schedule a cluster upgrade include:

- **Manually using the command line interface (CLI)** Start a one-time immediate upgrade or schedule a one-time upgrade for a future date and time.
- **Manually using the Red Hat OpenShift Cluster Manager user interface (UI)** Start a one-time immediate upgrade or schedule a one-time upgrade for a future date and time.
- **Automated upgrades** Set an upgrade window for recurring y-stream upgrades whenever a new version is available without needing to manually schedule it. Minor versions have to be manually scheduled.

For more details about cluster upgrades, run the following command:

```
$ rosa upgrade cluster --help
```

### 15.10.1. Manually upgrading your cluster using the CLI

1. Check if there is an upgrade available by running the following command:

```
$ rosa list upgrade -c <cluster-name>
```

### Example output

```
$ rosa list upgrade -c <cluster-name>
```

**VERSION NOTES**

4.14.7 recommended

4.14.6

...

In the above example, versions 4.14.7 and 4.14.6 are both available.

- Schedule the cluster to upgrade within the hour by running the following command:

```
$ rosa upgrade cluster -c <cluster-name> --version <desired-version>
```

- Optional:** Schedule the cluster to upgrade at a later date and time by running the following command:

```
$ rosa upgrade cluster -c <cluster-name> --version <desired-version> --schedule-date <future-date-for-update> --schedule-time <future-time-for-update>
```

### 15.10.2. Manually upgrading your cluster using the UI

- Log in to the OpenShift Cluster Manager, and select the cluster you want to upgrade.
- Click **Settings**.
- If an upgrade is available, click **Update**.

**Monitoring**

Enable user workload monitoring ⓘ  
Monitor your own projects in isolation from Red Hat Site Reliability Engineering (SRE) platform metrics.

---

**Update strategy**

Note: In the event of [Critical security concerns](#) (CVEs) that significantly impact the security or stability of the cluster, updates may be automatically scheduled by Red Hat SRE to the latest z-stream version not impacted by the CVE within 2 business days after customer notifications.

Individual updates  
Schedule each update individually. Take into consideration end of life dates from the [lifecycle policy](#) when planning updates.

Recurring updates  
The cluster will be automatically updated based on your preferred day and start time when new patch updates ([z-stream](#)) are available. When a new minor version is available, you'll be notified and must manually allow the cluster to update to the next minor version.

---

**Node draining**

You may set a grace period for how long pod disruption budget-protected workloads will be respected during updates. After this grace period, any workloads protected by pod disruption budgets that have not been successfully drained from a node will be forcibly evicted.

**Update status**

Update available

4.12.13 ————— 4.12.45

Additional versions available between 4.12.13 and 4.12.45

[Update](#)

Feedback

- Select the version to which you want to upgrade in the new window.
- Schedule a time for the upgrade or begin it immediately.

### 15.10.3. Setting up automatic recurring upgrades

- Log in to the OpenShift Cluster Manager, and select the cluster you want to upgrade.
- Click **Settings**.
  - Under **Update Strategy**, click **Recurring updates**.

3. Set the day and time for the upgrade to occur.
4. Under **Node draining**, select a grace period to allow the nodes to drain before pod eviction.
5. Click **Save**.

## 15.11. TUTORIAL: DELETING YOUR CLUSTER

You can delete your Red Hat OpenShift Service on AWS (ROSA) cluster using either the command line interface (CLI) or the user interface (UI).

### 15.11.1. Deleting a ROSA cluster using the CLI

1. **Optional:** List your clusters to make sure you are deleting the correct one by running the following command:

```
$ rosa list clusters
```

2. Delete a cluster by running the following command:

```
$ rosa delete cluster --cluster <cluster-name>
```



#### WARNING

This command is non-recoverable.

3. The CLI prompts you to confirm that you want to delete the cluster. Press **y** and then **Enter**. The cluster and all its associated infrastructure will be deleted.



#### NOTE

All AWS STS and IAM roles and policies will remain and must be deleted manually once the cluster deletion is complete by following the steps below.

4. The CLI outputs the commands to delete the OpenID Connect (OIDC) provider and Operator IAM roles resources that were created. Wait until the cluster finishes deleting before deleting these resources. Perform a quick status check by running the following command:

```
$ rosa list clusters
```

5. Once the cluster is deleted, delete the OIDC provider by running the following command:

```
$ rosa delete oidc-provider -c <clusterID> --mode auto --yes
```

6. Delete the Operator IAM roles by running the following command:

```
$ rosa delete operator-roles -c <clusterID> --mode auto --yes
```

**NOTE**

This command requires the cluster ID and not the cluster name.

- Only remove the remaining account roles if they are no longer needed by other clusters in the same account. If you want to create other ROSA clusters in this account, do not perform this step.  
To delete the account roles, you need to know the prefix used when creating them. The default is "ManagedOpenShift" unless you specified otherwise.

Delete the account roles by running the following command:

```
$ rosa delete account-roles --prefix <prefix> --mode auto --yes
```

### 15.11.2. Deleting a ROSA cluster using the UI

- Log in to the [OpenShift Cluster Manager](#), and locate the cluster you want to delete.
- Click the three dots to the right of the cluster.

| Name      | Status | Type | Created     | Version | Provider (Region) |
|-----------|--------|------|-------------|---------|-------------------|
| rosa-test | Ready  | ROSA | 12 Jan 2024 | 4.14.7  | AWS (us-east-1)   |

- In the dropdown menu, click **Delete cluster**.

| Name      | Status | Type | Created     | Version | Provider (Region) |
|-----------|--------|------|-------------|---------|-------------------|
| rosa-test | Ready  | ROSA | 12 Jan 2024 | 4.14.7  | AWS (us-east-1)   |

- Open console
- Edit display name
- Edit machine pool
- Delete cluster

- Enter the name of the cluster to confirm deletion, and click **Delete**.

## 15.12. TUTORIAL: OBTAINING SUPPORT

Finding the right help when you need it is important. These are some of the resources at your disposal when you need assistance.

### 15.12.1. Adding support contacts

You can add additional email addresses for communications about your cluster.

- On the Red Hat OpenShift Cluster Manager user interface (UI), click **select cluster**.

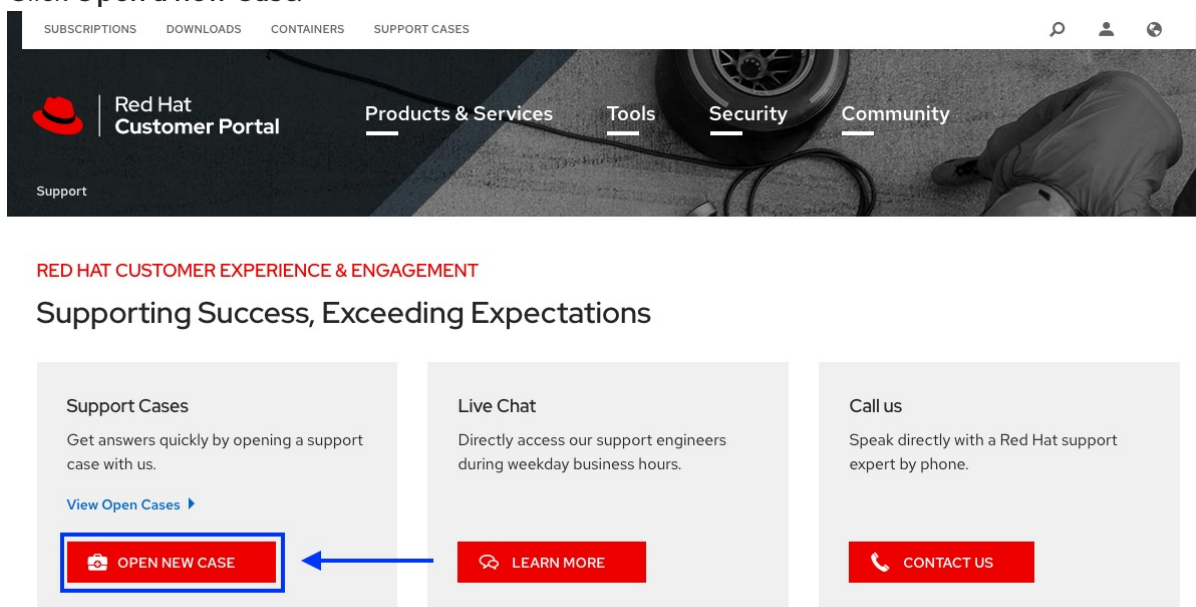
2. Click the **Support** tab.
3. Click **Add notification contact**, and enter the additional email addresses.

### 15.12.2. Contacting Red Hat for support using the UI

1. On the OpenShift Cluster Manager UI, click the **Support** tab.
2. Click **Open support case**.

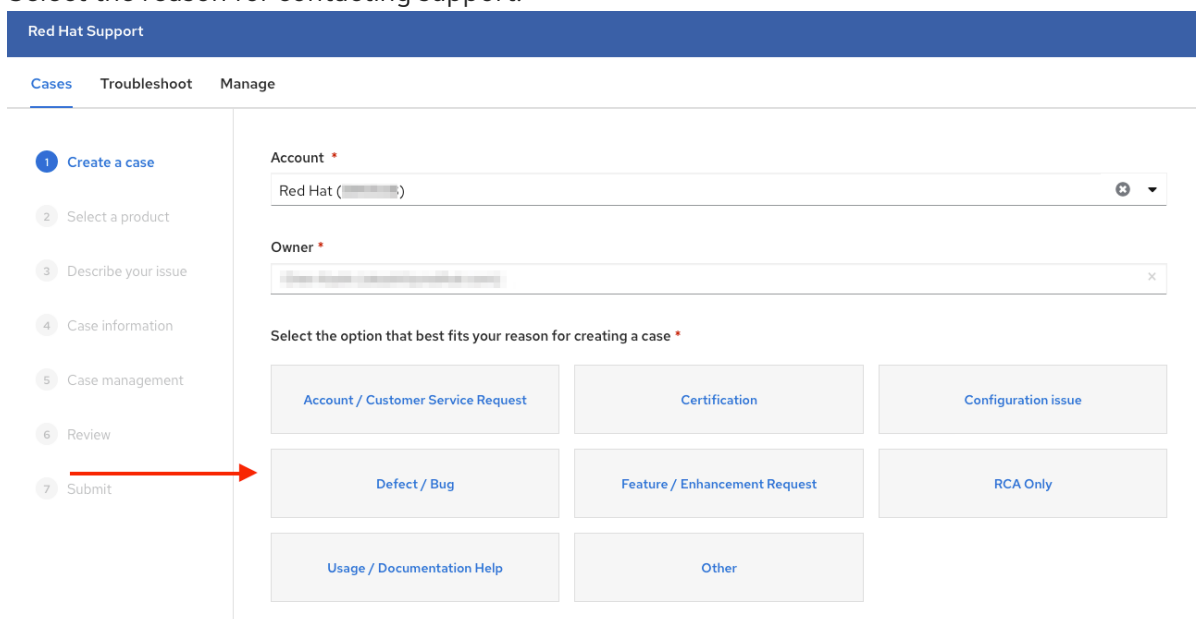
### 15.12.3. Contacting Red Hat for support using the support page

1. Go to the [Red Hat support page](#).
2. Click **Open a new Case**.

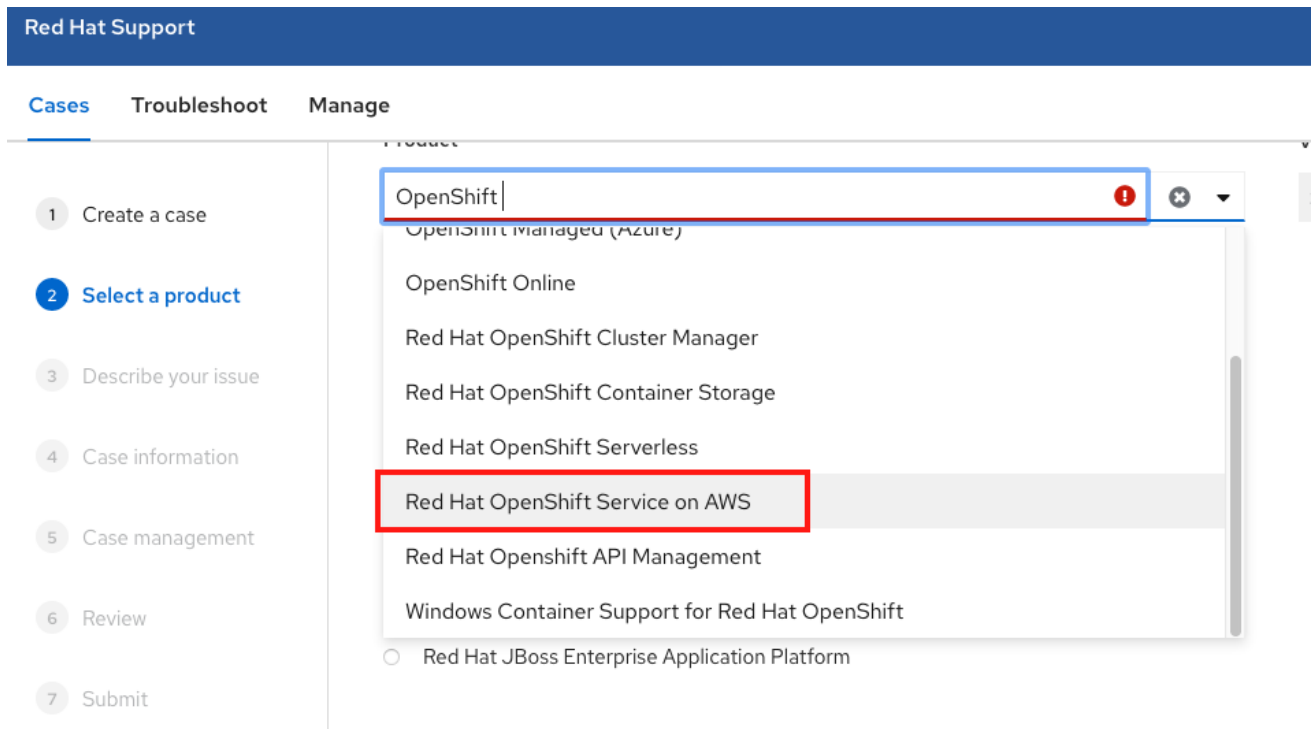


Note: If needed, our engineers can initiate a [Remote Support Session](#) to view and/or access your system.

3. Log in to your Red Hat account.
4. Select the reason for contacting support.



## 5. Select Red Hat OpenShift Service on AWS.

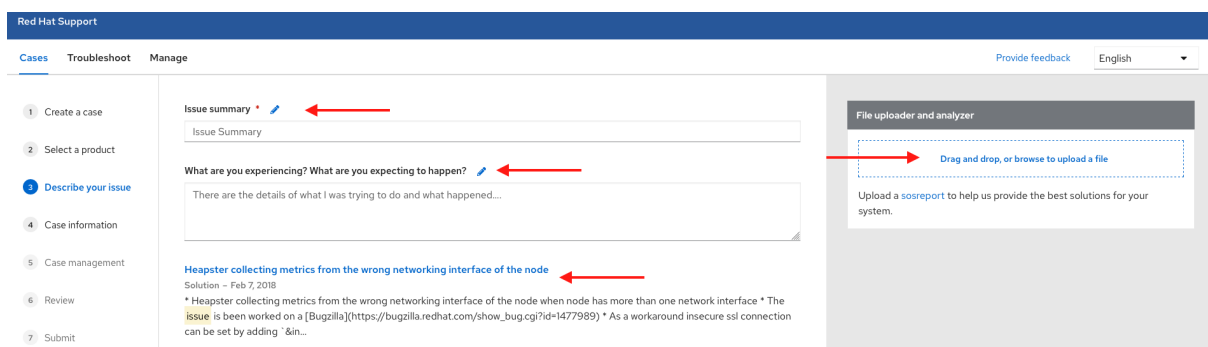


The screenshot shows the Red Hat Support portal interface. On the left, there is a navigation menu with steps: 1. Create a case, 2. Select a product (highlighted in blue), 3. Describe your issue, 4. Case information, 5. Case management, 6. Review, and 7. Submit. The main content area shows a search bar with 'OpenShift' entered. A dropdown menu is open, listing several products. The product 'Red Hat OpenShift Service on AWS' is highlighted with a red rectangular box. Other products listed include OpenShift managed (Azure), OpenShift Online, Red Hat OpenShift Cluster Manager, Red Hat OpenShift Container Storage, Red Hat OpenShift Serverless, Red Hat OpenShift API Management, Windows Container Support for Red Hat OpenShift, and Red Hat JBoss Enterprise Application Platform.

1. Click **continue**.
2. Enter a summary of the issue and the details of your request. Upload any files, logs, and screenshots. The more details you provide, the better Red Hat support can help your case.

**NOTE**

Relevant suggestions that might help with your issue will appear at the bottom of this page.



The screenshot shows the 'Describe your issue' step in the Red Hat Support portal. The navigation menu on the left has 'Describe your issue' highlighted in blue. The main content area has three input fields: 'Issue summary', 'What are you experiencing? What are you expecting to happen?', and a text area for details. Red arrows point to each of these fields. Below the text area, a suggested solution is displayed: 'Heapster collecting metrics from the wrong networking interface of the node'. To the right, there is a 'File uploader and analyzer' section with a 'Drag and drop, or browse to upload a file' button and a note about uploading a sosreport.

3. Click **Continue**.
4. Answer the questions in the new fields.
5. Click **Continue**.
6. Enter the following information about your case:
  - a. **Support level:** Premium
  - b. **Severity:** Review the Red Hat Support Severity Level Definitions to choose the correct one.

- c. **Group:** If this is related to a few other cases you can select the corresponding group.
  - d. **Language**
  - e. **Send notifications:** Add any additional email addresses to keep notified of activity.
  - f. **Red Hat associates:** If you are working with anyone from Red Hat and want to keep them in the loop you can enter their email address here.
  - g. **Alternate Case ID:** If you want to attach your own ID to it you can enter it here.
7. Click **Continue**.
  8. On the review screen make sure you select the correct cluster ID that you are contacting support about.

Red Hat Support

Cases Troubleshoot Manage

- 1 Create a case
- 2 Select a product
- 3 Describe your issue
- 4 Case information
- 5 Case management
- 6 **Review**
- 7 Submit

**Account \***  
Red Hat ( )

**Owner \***  
( )@redhat.com

**Product \*** Red Hat OpenShift Service on AWS **Version \*** Red Hat OpenShift Service on AWS

**OpenShift Cluster ID \*** ←  
Select a Cluster

Cluster is not listed ▶

9. Click **Submit**.
10. You will be contacted based on the response time committed to for the [indicated severity level](#).



## CHAPTER 16. DEPLOYING AN APPLICATION

### 16.1. TUTORIAL: DEPLOYING AN APPLICATION

#### 16.1.1. Introduction

After successfully provisioning your cluster, you can deploy an application on it. This application allows you to become more familiar with some of the features of Red Hat OpenShift Service on AWS (ROSA) and Kubernetes.

##### 16.1.1.1. Lab overview

In this lab, you will complete the following set of tasks designed to help you understand the concepts of deploying and operating container-based applications:

- Deploy a Node.js based app by using S2I and Kubernetes Deployment objects.
- Set up a continuous delivery (CD) pipeline to automatically push source code changes.
- Explore logging.
- Experience self healing of applications.
- Explore configuration management through configmaps, secrets, and environment variables.
- Use persistent storage to share data across pod restarts.
- Explore networking within Kubernetes and applications.
- Familiarize yourself with ROSA and Kubernetes functionality.
- Automatically scale pods based on loads from the Horizontal Pod Autoscaler.
- Use AWS Controllers for Kubernetes (ACK) to deploy and use an S3 bucket.

This lab uses either the ROSA CLI or ROSA web user interface (UI).

### 16.2. TUTORIAL: DEPLOYING AN APPLICATION

#### 16.2.1. Prerequisites

1. A Provisioned ROSA cluster  
This lab assumes you have access to a successfully provisioned a ROSA cluster. If you have not yet created a ROSA cluster, see [Red Hat OpenShift Service on AWS quick start guide](#) for more information.
2. The OpenShift Command Line Interface (CLI)  
For more information, see [Getting started with the OpenShift CLI](#).
3. A GitHub Account  
Use your existing GitHub account or register at <https://github.com/signup>.

### 16.3. TUTORIAL: DEPLOYING AN APPLICATION

## 16.3.1. Lab overview

### 16.3.1.1. Lab resources

- [Source code for the OSToy application](#)
- [OSToy front-end container image](#)
- [OSToy microservice container image](#)
- Deployment Definition YAML files:

#### **ostoy-frontend-deployment.yaml**

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
  name: ostoy-pvc
spec:
  accessModes:
    - ReadWriteOnce
  resources:
    requests:
      storage: 1Gi
---
apiVersion: apps/v1
kind: Deployment
metadata:
  name: ostoy-frontend
  labels:
    app: ostoy
spec:
  selector:
    matchLabels:
      app: ostoy-frontend
  strategy:
    type: Recreate
  replicas: 1
  template:
    metadata:
      labels:
        app: ostoy-frontend
    spec:
      # Uncomment to use with ACK portion of the workshop
      # If you chose a different service account name please replace it.
      # serviceAccount: ostoy-sa
      containers:
        - name: ostoy-frontend
          securityContext:
            allowPrivilegeEscalation: false
            runAsNonRoot: true
          seccompProfile:
            type: RuntimeDefault
          capabilities:
            drop:
              - ALL
```

```

image: quay.io/ostoylab/ostoy-frontend:1.6.0
imagePullPolicy: IfNotPresent
ports:
- name: ostoy-port
  containerPort: 8080
resources:
  requests:
    memory: "256Mi"
    cpu: "100m"
  limits:
    memory: "512Mi"
    cpu: "200m"
volumeMounts:
- name: configvol
  mountPath: /var/config
- name: secretvol
  mountPath: /var/secret
- name: datavol
  mountPath: /var/demo_files
livenessProbe:
  httpGet:
    path: /health
    port: 8080
  initialDelaySeconds: 10
  periodSeconds: 5
env:
- name: ENV_TOY_SECRET
  valueFrom:
    secretKeyRef:
      name: ostoy-secret-env
      key: ENV_TOY_SECRET
- name: MICROSERVICE_NAME
  value: OSTOY_MICROSERVICE_SVC
- name: NAMESPACE
  valueFrom:
    fieldRef:
      fieldPath: metadata.namespace
volumes:
- name: configvol
  configMap:
    name: ostoy-configmap-files
- name: secretvol
  secret:
    defaultMode: 420
    secretName: ostoy-secret
- name: datavol
  persistentVolumeClaim:
    claimName: ostoy-pvc
---
apiVersion: v1
kind: Service
metadata:
  name: ostoy-frontend-svc
  labels:
    app: ostoy-frontend
spec:

```

```

type: ClusterIP
ports:
  - port: 8080
    targetPort: ostoy-port
    protocol: TCP
    name: ostoy
selector:
  app: ostoy-frontend
---
apiVersion: route.openshift.io/v1
kind: Route
metadata:
  name: ostoy-route
spec:
  to:
    kind: Service
    name: ostoy-frontend-svc
---
apiVersion: v1
kind: Secret
metadata:
  name: ostoy-secret-env
type: Opaque
data:
  ENV_TOY_SECRET: VGhpcyBpcyBhIHRIc3Q=
---
kind: ConfigMap
apiVersion: v1
metadata:
  name: ostoy-configmap-files
data:
  config.json: '{ "default": "123" }'
---
apiVersion: v1
kind: Secret
metadata:
  name: ostoy-secret
data:
  secret.txt:
VVNFUk5BTUU9bXlfdXNlcgpQQVNTV09SRD1AT3RCbCVYQXAhIzYzMIk1RndDQE1UUWsK
U01UUD1sb2NhbGhvc3QKU01UUF9QT1JUPT11
type: Opaque

```

### ostoy-microservice-deployment.yaml

```

apiVersion: apps/v1
kind: Deployment
metadata:
  name: ostoy-microservice
  labels:
    app: ostoy
spec:
  selector:
    matchLabels:
      app: ostoy-microservice

```

```

replicas: 1
template:
  metadata:
    labels:
      app: ostoy-microservice
  spec:
    containers:
      - name: ostoy-microservice
        securityContext:
          allowPrivilegeEscalation: false
          runAsNonRoot: true
          seccompProfile:
            type: RuntimeDefault
        capabilities:
          drop:
            - ALL
        image: quay.io/ostoylab/ostoy-microservice:1.5.0
        imagePullPolicy: IfNotPresent
        ports:
          - containerPort: 8080
            protocol: TCP
        resources:
          requests:
            memory: "128Mi"
            cpu: "50m"
          limits:
            memory: "256Mi"
            cpu: "100m"
    ---
  apiVersion: v1
  kind: Service
  metadata:
    name: ostoy-microservice-svc
    labels:
      app: ostoy-microservice
  spec:
    type: ClusterIP
    ports:
      - port: 8080
        targetPort: 8080
        protocol: TCP
    selector:
      app: ostoy-microservice

```

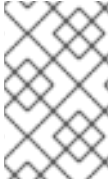
- S3 bucket manifest for ACK S3

### s3-bucket.yaml

```

apiVersion: s3.services.k8s.aws/v1alpha1
kind: Bucket
metadata:
  name: ostoy-bucket
  namespace: ostoy
spec:
  name: ostoy-bucket

```



## NOTE

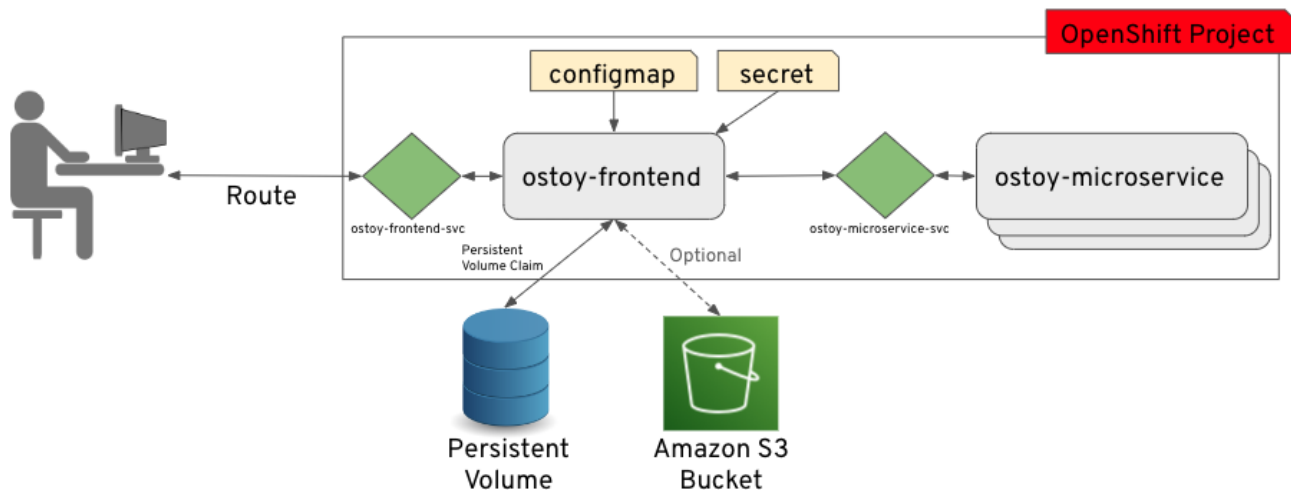
To simplify deployment of the OSToy application, all of the objects required in the above deployment manifests are grouped together. For a typical enterprise deployment, a separate manifest file for each Kubernetes object is recommended.

### 16.3.1.2. About the OSToy application

OSToy is a simple Node.js application that you will deploy to a ROSA cluster to help explore the functionality of Kubernetes. This application has a user interface where you can:

- Write messages to the log (stdout / stderr).
- Intentionally crash the application to view self-healing.
- Toggle a liveness probe and monitor OpenShift behavior.
- Read config maps, secrets, and env variables.
- If connected to shared storage, read and write files.
- Check network connectivity, intra-cluster DNS, and intra-communication with the included microservice.
- Increase the load to view automatic scaling of the pods to handle the load using the Horizontal Pod Autoscaler.
- Optional: Connect to an AWS S3 bucket to read and write objects.

### 16.3.1.3. OSToy Application Diagram



### 16.3.1.4. Understanding the OSToy UI

1. Shows the pod name that served your browser the page.
2. **Home:** The main page of the application where you can perform some of the functions listed which we will explore.
3. **Persistent Storage:** Allows you to write data to the persistent volume bound to this application.
4. **Config Maps:** Shows the contents of configmaps available to the application and the key:value pairs.
5. **Secrets:** Shows the contents of secrets available to the application and the key:value pairs.
6. **ENV Variables:** Shows the environment variables available to the application.
7. **Networking:** Tools to illustrate networking within the application.
8. **Pod Auto Scaling:** Tool to increase the load of the pods and test the HPA.
9. **ACK S3:** Optional: Integrate with AWS S3 to read and write objects to a bucket.



## NOTE

In order see the "ACK S3" section of OSToy, you must complete the ACK section of this workshop. If you decide not to complete that section, the OSToy application will still function.

10. **About:** Displays more information about the application.

## 16.4. TUTORIAL: DEPLOYING AN APPLICATION

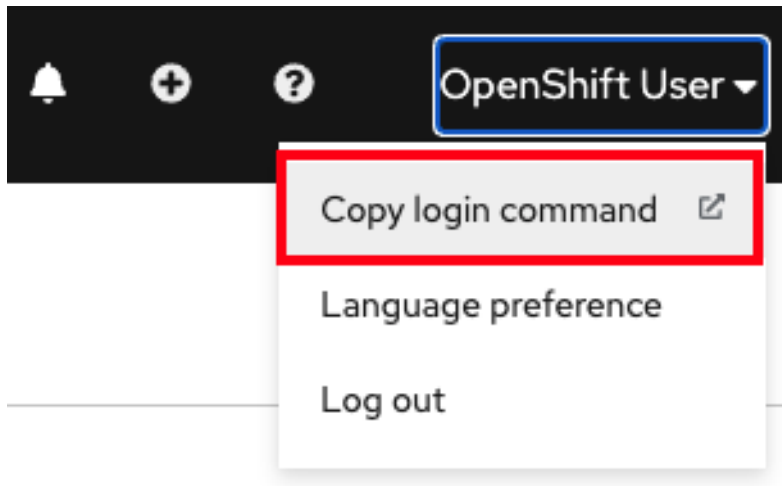
### 16.4.1. Deploying the OSToy application with Kubernetes

You can deploy the OSToy application by creating and storing the images for the front-end and back-end microservice containers in an image repository. You can then create Kubernetes deployments to deploy the application.

#### 16.4.1.1. Retrieving the login command

1. If you are not logged in to the CLI, access your cluster with the web console.

- Click the dropdown arrow next to your login name in the upper right, and select **Copy Login Command**.



A new tab opens.

- Select your authentication method.
- Click **Display Token**.
- Copy the command under **Log in with this token**
- From your terminal, paste and run the copied command. If the login is successful, you will see the following confirmation message:

```
$ oc login --token=<your_token> --server=https://api.osd4-
demo.abc1.p1.openshiftapps.com:6443
Logged into "https://api.myrosacluster.abcd.p1.openshiftapps.com:6443" as "rosa-user"
using the token provided.
```

You don't have any projects. You can try to create a new project, by running

```
oc new-project <project name>
```

## 16.4.1.2. Creating a new project

### 16.4.1.2.1. Using the CLI

- Create a new project named **ostoy** in your cluster by running following command:

```
$ oc new-project ostoy
```

#### Example output

```
Now using project "ostoy" on server
"https://api.myrosacluster.abcd.p1.openshiftapps.com:6443".
```

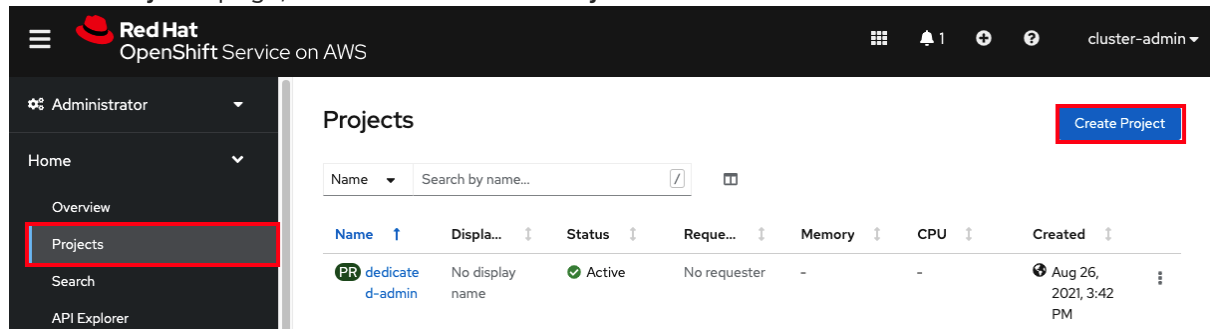
- Optional:** Alternatively, create a unique project name by running the following command:

```
$ oc new-project ostoy-$(uuidgen | cut -d - -f 2 | tr '[:upper:]' '[:lower:]')
```



### 16.4.1.2.2. Using the web console

1. From the web console, click **Home** → **Projects**.
2. On the **Projects** page, click create **Create Project**.



### 16.4.1.3. Deploying the back-end microservice

The microservice serves internal web requests and returns a JSON object containing the current hostname and a randomly generated color string.

- Deploy the microservice by running the following command from your terminal:

```
$ oc apply -f https://raw.githubusercontent.com/openshift-cs/rosaworkshop/master/rosaworkshop/ostoy/yaml/ostoy-microservice-deployment.yaml
```

#### Example output

```
$ oc apply -f https://raw.githubusercontent.com/openshift-cs/rosaworkshop/master/rosaworkshop/ostoy/yaml/ostoy-microservice-deployment.yaml
deployment.apps/ostoy-microservice created
service/ostoy-microservice-svc created
```

### 16.4.1.4. Deploying the front-end service

The front-end deployment uses the Node.js front-end for the application and additional Kubernetes objects.

The **ostoy-frontend-deployment.yaml** file shows that front-end deployment defines the following features:

- Persistent volume claim
- Deployment object
- Service
- Route
- Configmaps
- Secrets
  - Deploy the application front-end and create all of the objects by entering the following command:

```
$ oc apply -f https://raw.githubusercontent.com/openshift-cs/rosaworkshop/master/rosa-workshop/ostoy/yaml/ostoy-frontend-deployment.yaml
```

### Example output

```
persistentvolumeclaim/ostoy-pvc created
deployment.apps/ostoy-frontend created
service/ostoy-frontend-svc created
route.route.openshift.io/ostoy-route created
configmap/ostoy-configmap-env created
secret/ostoy-secret-env created
configmap/ostoy-configmap-files created
secret/ostoy-secret created
```

You should see all objects created successfully.

#### 16.4.1.5. Getting the route

You must get the route to access the application.

- Get the route to your application by running the following command:

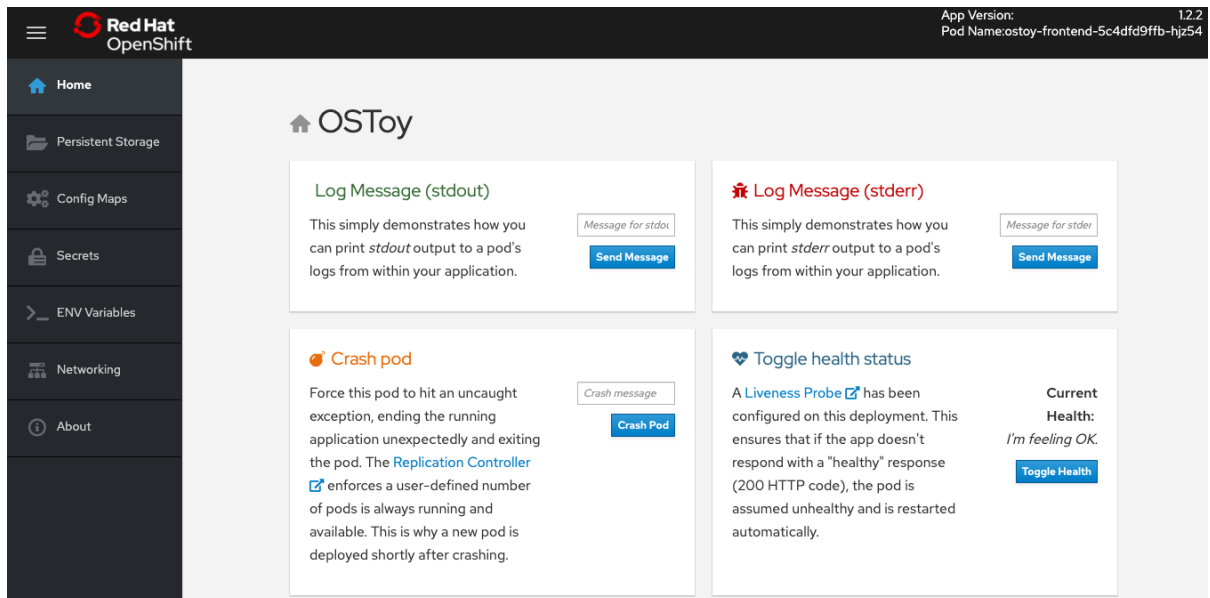
```
$ oc get route
```

### Example output

| NAME               | HOST/PORT  | PATH | SERVICES | PORT |
|--------------------|--|------|----------|------|
| ostoy-route        | ostoy-route-ostoy.apps.<your-rosa-cluster>.abcd.p1.openshiftapps.com |      |          |      |
| ostoy-frontend-svc | <all>  |      | None     |      |

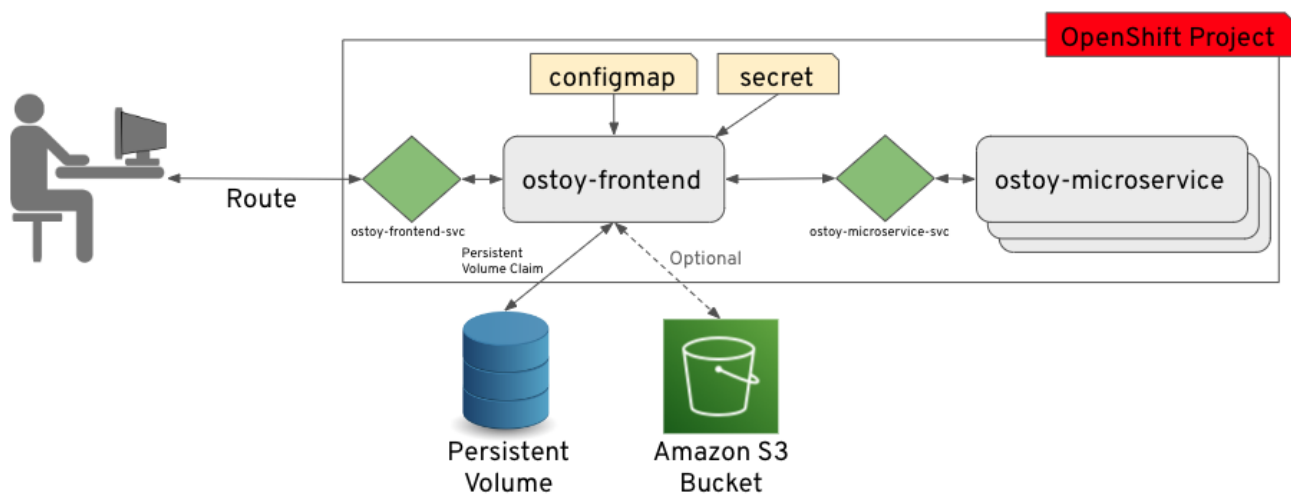
#### 16.4.1.6. Viewing the application

1. Copy the **ostoy-route-ostoy.apps.<your-rosa-cluster>.abcd.p1.openshiftapps.com** URL output from the previous step.
2. Paste the copied URL into your web browser and press enter. You should see the homepage of your application. If the page does not load, make sure you use **http** and not **https**.



## 16.5. TUTORIAL: NETWORKING

This tutorial shows how the OSToy app uses intra-cluster networking to separate functions by using microservices and visualize the scaling of pods.



The diagram shows there are at least two separate pods, each with its own service.

One pod functions as the front end web application with a service and a publicly accessible route. The other pod functions as the backend microservice with a service object so that the front end pod can communicate with the microservice. This communication occurs across the pods if more than one. Because of these communication limits, this microservice is not accessible from outside this cluster or from other namespaces or projects if these are configured. The sole purpose of this microservice is to serve internal web requests and return a JSON object containing the current hostname, which is the pod's name, and a randomly generated color string. This color string is used to display a box with that color displayed in the tile titled "Intra-cluster Communication".

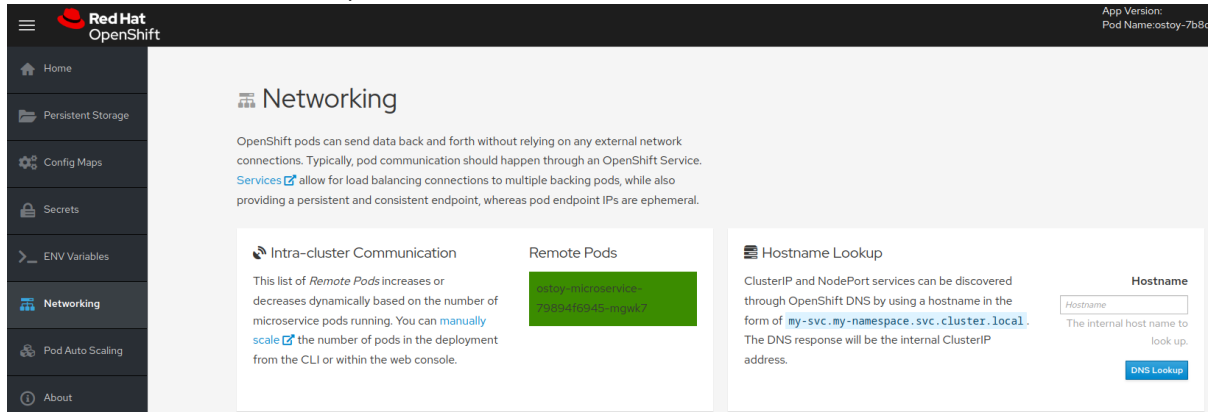
For more information about the networking limitations, see [About network policy](#).

### 16.5.1. Intra-cluster networking

You can view your networking configurations in your OSToy application.

## Procedure

1. In the OSToy application, click **Networking** in the left menu.
2. Review the networking configuration. The right tile titled "Hostname Lookup" illustrates how the service name created for a pod can be used to translate into an internal ClusterIP address.



3. Enter the name of the microservice created in the right tile ("Hostname Lookup") following the format of **<service\_name>.<namespace>.svc.cluster.local**. You can find this service name in the service definition of **ostoy-microservice.yaml** by running the following command:

```
$ oc get service <name_of_service> -o yaml
```

### Example output

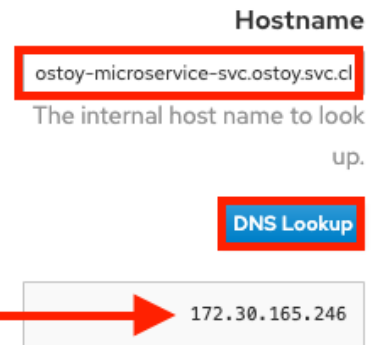
```
apiVersion: v1
kind: Service
metadata:
  name: ostoy-microservice-svc
  labels:
    app: ostoy-microservice
spec:
  type: ClusterIP
  ports:
    - port: 8080
      targetPort: 8080
      protocol: TCP
  selector:
    app: ostoy-microservice
```

In this example, the full hostname is **ostoy-microservice-svc.ostoy.svc.cluster.local**.

4. You see an IP address returned. In this example it is **172.30.165.246**. This is the intra-cluster IP address, which is only accessible from within the cluster.

## ☰ Hostname Lookup

ClusterIP and NodePort services can be discovered through [OpenShift DNS](#) by using a hostname in the form of `my-svc.my-namespace.svc.cluster.local`. The DNS response will be the internal ClusterIP address.



## 16.6. TUTORIAL: PERSISTENT VOLUMES FOR CLUSTER STORAGE

Red Hat OpenShift Service on AWS (ROSA) (classic architecture) and Red Hat OpenShift Service on AWS (ROSA) support storing persistent volumes with either [Amazon Web Services \(AWS\) Elastic Block Store \(EBS\)](#) or [AWS Elastic File System \(EFS\)](#).

### 16.6.1. Using persistent volumes

Use the following procedures to create a file, store it on a persistent volume in your cluster, and confirm that it still exists after pod failure and re-creation.

#### 16.6.1.1. Viewing a persistent volume claim

1. Navigate to the cluster's OpenShift web console.
2. Click **Storage** in the left menu, then click **PersistentVolumeClaims** to see a list of all the persistent volume claims.
3. Click a persistence volume claim to see the size, access mode, storage class, and other additional claim details.

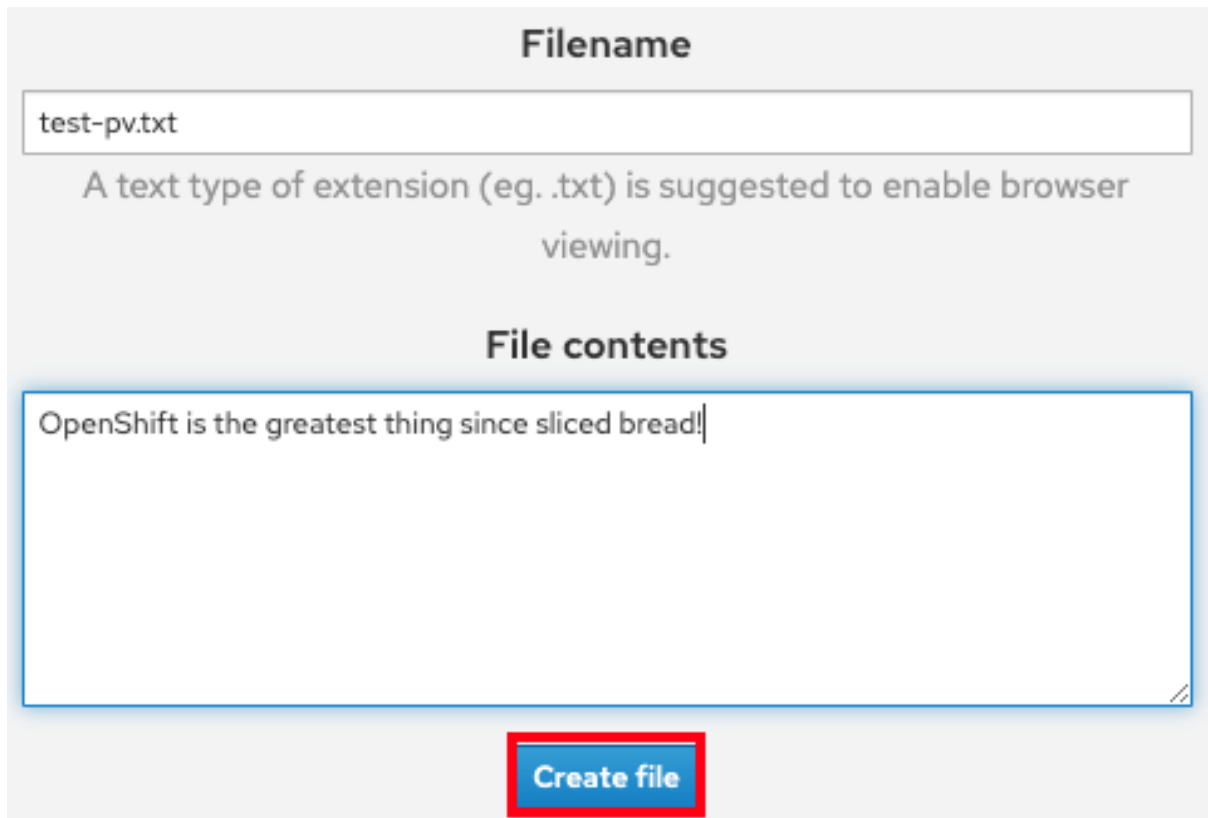


#### NOTE

The access mode is **ReadWriteOnce** (RWO). This means that the volume can only be mounted to one node and the pod or pods can read and write to the volume.

#### 16.6.1.2. Storing your file

1. In the OSToy app console, click **Persistent Storage** in the left menu.
2. In the **Filename** box, enter a file name with a **.txt** extension, for example **test-pv.txt**.
3. In the **File contents** box, enter a sentence of text, for example **OpenShift is the greatest thing since sliced bread!**
4. Click **Create file**.



The screenshot shows a form for creating a new file. At the top, the heading "Filename" is centered. Below it is a text input field containing "test-pv.txt". Underneath the input field is a greyed-out instruction: "A text type of extension (eg. .txt) is suggested to enable browser viewing." Below this is another heading "File contents" centered above a large text area. The text area contains the text "OpenShift is the greatest thing since sliced bread!". At the bottom center of the form is a blue button with the text "Create file", which is highlighted with a red rectangular border.

5. Scroll to **Existing files** on the OSToy app console.
6. Click the file you created to see the file name and contents.

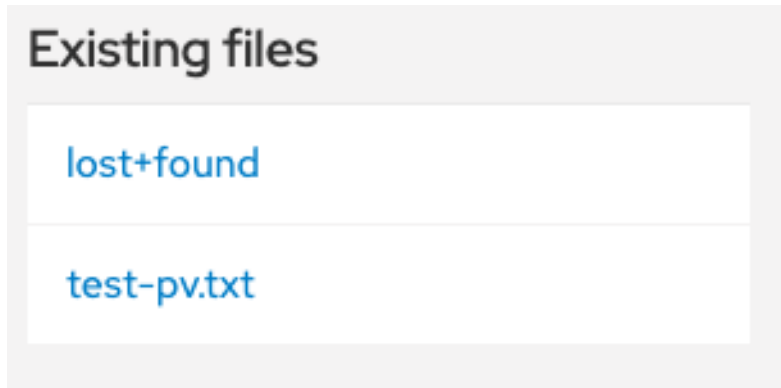


### 16.6.1.3. Crashing the pod

1. On the OSToy app console, click **Home** in the left menu.
2. Click **Crash pod**.

### 16.6.1.4. Confirming persistent storage

1. Wait for the pod to re-create.
2. On the OSToy app console, click **Persistent Storage** in the left menu.
3. Find the file you created, and open it to view and confirm the contents.



## Verification

The deployment YAML file shows that we mounted [the directory /var/demo\\_files](#) to our persistent volume claim.

1. Retrieve the name of your front-end pod by running the following command:

```
$ oc get pods
```

2. Start a secure shell (SSH) session in your container by running the following command:

```
$ oc rsh <pod_name>
```

3. Go to the directory by running the following command:

```
$ cd /var/demo_files
```

4. **Optional:** See all the files you created by running the following command:

```
$ ls
```

5. Open the file to view the contents by running the following command:

```
$ cat test-pv.txt
```

6. Verify that the output is the text you entered in the OSToy app console.

### Example terminal

```
$ oc get pods
NAME                                READY  STATUS   RESTARTS  AGE
ostoy-frontend-5fc8d486dc-wsw24    1/1    Running  0         18m
ostoy-microservice-6cf764974f-hx4qm 1/1    Running  0         18m

$ oc rsh ostoy-frontend-5fc8d486dc-wsw24

$ cd /var/demo_files/

$ ls
lost+found  test-pv.txt
```

```
$ cat test-pv.txt  
OpenShift is the greatest thing since sliced bread!
```

### 16.6.1.5. Ending the session

- Type **exit** in your terminal to quit the session and return to the CLI.

### 16.6.2. Additional resources

- For more information about persistent volume storage, see [Understanding persistent storage](#).
- For more information about ROSA storage options, see [Storage overview](#).