

Service Workbench Post Deployment Guide

Table of Contents

Account types.....	3
Enable Local Users.....	3
Preparing the organizational account	3
Creating a configuration file	4
Enabling AWS Organizations in the organizational account	4
Deploying the prepare organizational account SDC	4
Confirming the creation of organizational role	5
Prepare your Account for AppStream	5
Setting up an AppStream Image.....	6
Exporting data from AppStream-enabled workspaces.....	8
Create or add compute hosting accounts	11
Create AWS Account	13
Add AWS Account.....	14
Creating RStudio ALB workspace.....	17
Accessing RStudio workspace	18
Application load balancing for RStudio ALB workspace.....	18
Limitations	19
Create Indexes and Projects	19
Create an Administrator User.....	20
Import Service Catalog Products	21
Import a Product	22
Configuration for EC2 Linux.....	25
Configuration for Amazon EC2 Windows	27
Configuration for Amazon SageMaker	30
Configuration for Amazon EMR.....	32
Viewing logs.....	35

Viewing Service Workbench logs in CloudWatch.....	35
Metrics.....	36
Deploying Updates	36
Reference	36
Adding an AWS IAM role to an Amazon EC2 instance	37
Adding administrator’s role to a new Amazon EC2 instance	37
Adding a role to an existing instance	38
Usage of AWS Cloud Services	39
Amazon EC2	40
AWS IAM	40
AWS Organizations	41
Amazon S3	42
AWS Cost Explorer	43

Account types

Service Workbench uses *three* types of accounts. You will see these account names throughout the documentation.

- **Organizational:** Holds the AWS Organization which creates hosting accounts. Note that you may already have a method of obtaining AWS accounts supported by your organization. If this is the case, you will not create an organizational account or use the **Create Account** functionality within Service Workbench when onboarding a hosting account.
- **Main:** The account to which Service Workbench is deployed, serving as the central point of login and access to researcher workspaces.
- **Hosting:** Accounts that are established associated to the Service Workbench main account through the onboarding process to host the compute resources (Amazon SageMaker notebook instances, Amazon EC2 Windows and Linux instances, Amazon EMR clusters) associated to Service Workbench workspaces.

Read the following files in the source code documentation to learn more about the different types of AWS accounts within Service Workbench:

- `README.md`
- `main/solution/prepare-master-acc/README.md`

Enable Local Users

Local users are created only within the solution. Their credentials are stored in [Amazon DynamoDB](#). This is the easiest way to install. The alternative is to integrate with an Active Directory.

Preparing the organizational account

Note: This section is required only if Service Workbench is to be used to vend accounts in the AWS Organization by using Create Account. If Service Workbench uses only billing accounts onboarded using Add Account, this section can be omitted.

In this step, you deploy the **prepare_master_acc** Separately Deployable Component (SDC) in the `main/solution/prepare-master-acc` directory. This creates a role in the organizational account that allows **AssumeRole**, and has the **main** account as its trusted entity. If you have deployed Service Workbench in an **organizational** account, the **main** account is also the

organizational account, and the trusted entity is the same account ID as the **organizational** account.

The default settings in this step are:

- **Main account ID:** The current AWS account ID
- **External ID:** The string **workbench**

These default values are sufficient if you are deploying Service Workbench to an **organizational** account, and the default profile has permissions for the **organizational** account, since the **Main** account is also the **organizational** account. If deploying Service Workbench to the **main** AWS account that is not the same account as the **organizational** account, you must create a configuration file to specify the **main** account ID and the profile to use. This profile must have permissions for the **organizational** account.

Creating a configuration file

If deploying Service Workbench in an account other than that accessed by the current default profile, create a stage-named configuration file in the `main/solution/prepare-master-acc/config/settings` directory by copying `example.yml` to `<stage>.yml`. Edit the file as appropriate:

- **awsProfile:** The AWS credentials profile with permissions for the organizational account.
- **mainAccountID:** The 12-digit AWS account ID for the main AWS account, where the solution is deployed.
- **externalId:** As desired. The string **workbench** is often used. This string is required when creating an AWS account within Service Workbench.

Enabling AWS Organizations in the organizational account

In the [AWS Management Console](#), navigate to **AWS Organizations** to ensure that an organization exists for the **organizational** account. If it does not, then create a new one. There is no configuration setting. Service Workbench creates a new account in the AWS Organizations for this deployment, named after the **stage name** used at deployment.

Deploying the prepare organizational account SDC

To deploy the `prepare_master_acc` SDC:

- Read the `main/solution/prepare-master-acc/README.md` file.
- To deploy the **organizational** account SDC from the `main/solution/prepare-master-acc` directory, enter the following command:

```
pnpx sls deploy --stage <stage>
```

- To display the ARN of the **organizational role** from the same directory, enter the following command:

```
pnp sls info --verbose --stage <stage>
```

The **organizational role ARN** is needed when adding accounts within Service Workbench.

Note: Running the convenience script `scripts/master-account-deploy.sh <stage>` performs the same steps as `pnp sls deploy`.

Confirming the creation of organizational role

The newly-created role contains **MasterRole**. It has two policies and trusts the main account.

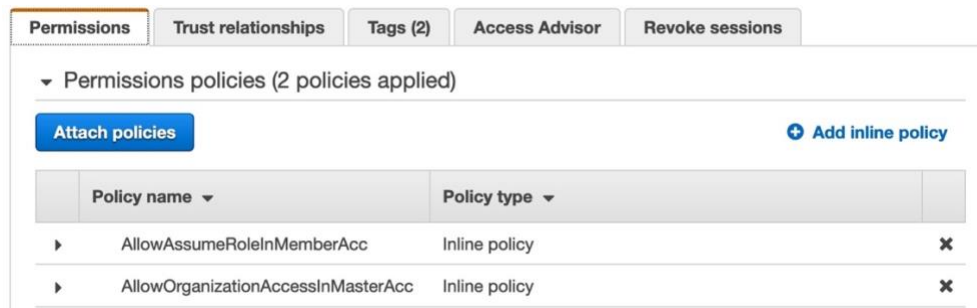


Figure 1: Permissions Policies

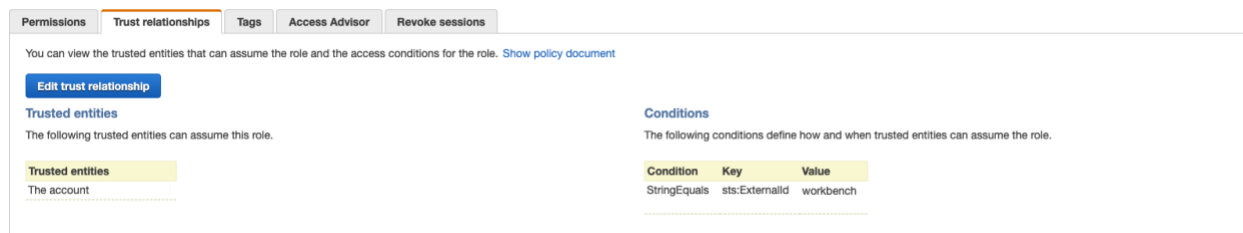


Figure 2: Edit Trust relationships

Prepare your Account for AppStream

Note: Read this section only if you want to enable AppStream.

AWS AppStream service limits may block resource creation in a new AWS account, or an account that has not yet hosted AppStream resources. The following actions will best prepare the hosting account.

1. **Required:** Go to AppStream 2.0 services (AWS Console), and choose **Get Started**. This will take you to a screen asking if you want to try out some templates. At this screen choose **Next**. This will only need to be done once for the account (it activates a role required for AppStream to be initiated).
2. **Recommended:** Launch and terminate at least one EC2 instance (size of instance and duration of run prior to termination do not matter), as this establishes the base compute service limits.
3. **Optional:** Open a support ticket within the hosting account to AWS AppStream service, noting your intention to create a fleet. This should only be necessary if you are creating a large fleet, but if the account is very new or has never had resources created prior to being onboarded as a hosting account, this is recommended.

Once prerequisites are complete and the CloudFormation stack has been created, go to AppStream on the AWS console of the hosting account. Go to **Fleet** and then choose the newly created fleet. Choose **Action>Start** to start the fleet. This step must be completed for any researcher workspace to be successfully launched within the hosting account.

After proceeding from the **Onboard AWS Account** Screen, you will be returned to the **Account listing** page with a status.

- **Up to Date** : The account has successfully onboarded and the version of the main and hosting account code is in sync.
- **Needs Update** : This status means that the main account code is more up to date than the hosting account code. There is an update button available and it is advised that you use this to bring the hosting account to the latest version.
- **Needs Onboarding** : The account has been onboarded via the Service Workbench UI but the CloudFormation template has not been generated
- **Errored** : The onboarding of the hosting account failed. Check the CloudFormation stack in the hosting account.
- **Pending** : The account onboarding process is ongoing. If a hosting account remains in pending status beyond 30 minute, it is likely that there is an issue onboarding the account. Check the CloudFormation stack in the hosting account.

Setting up an AppStream Image

Overview

This section describes the procedure of setting up an AppStream image with the following applications installed: Firefox, Notepad, PuttyGen, and Putty. AppStream should be built in the **main account**.

Pre-requisites

1. Navigate to AWS AppStream in your main account using the AWS Management Console.

2. Choose **Get Started** and then choose **Next**. This activates AppStream in your main account, following which you could proceed with the commands in `SETUP.md` file.

The image build process relies on the presence of the default VPC and outbound internet connectivity. If this is not available (typical of AWS Control Tower environments), then [create a new default VPC](#).

Launching an AppStream Image Builder instance

1. Navigate to the `scripts/app-stream` directory.
2. Enter the following commands:

```
npm install
npm run start-image-builder -- <AWS Profile> <region> <Base image name> <instance size>
# Example: npm run start-image-builder -- default us-east-1 AppStream-WinServer2019-06-01-2021 stream.standard.medium
# If preferred you can choose the default base image name and instance size by running this command: npm run start-image-builder -- default us-east-1 default default
```

Note: Set up your AWS profile beforehand so that you have permission to launch an AppStream Image Builder instance in your AWS Account.

3. Once the Image Builder is launched and ready, choose the URL provided in the terminal console. This opens the AppStream page on the AWS Console.

Building AppStream image

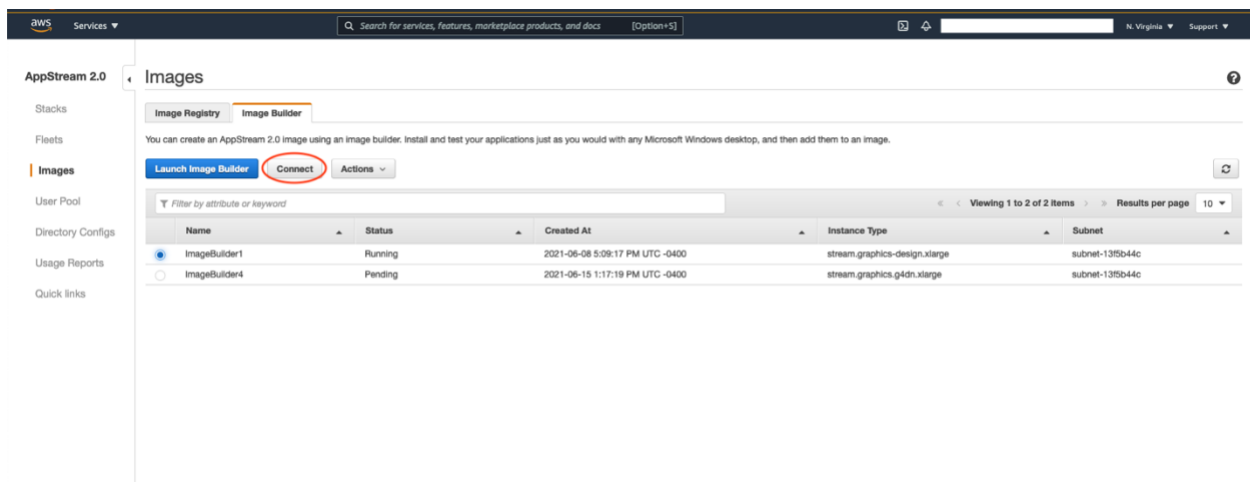


Figure: Create an AppStream image

1. On the AppStream page in the console, choose the AppStream image that was built in the previous step and choose **Connect**.
2. This opens a new tab in your browser. When prompted, log in as **Administrator**. This opens the Windows Desktop that you can interact with to create your AppStream image.
3. On the Windows Desktop, choose the **Start** button and enter `Windows Powershell`.
4. Right-click the application and choose `Run as administrator`.
5. Enter the following commands:

```
cd ~\Documents

# Pull the Image Builder script from Github
Invoke-WebRequest -Uri
https://raw.githubusercontent.com/aws-labs/service-workbench-on-aws/feat-secure-workspace-egress/scripts/app-stream/buildImage.ps1 -
OutFile buildImage.ps1

# Execute Image builder script
.\buildImage.ps1
```

6. At this point, the Image builder builds your image and the `Failed to reserve a session` message is displayed.
7. Log in to AppStream on the console again and wait till the AppStream image is built.

Exporting data from AppStream-enabled workspaces

As a researcher, if you have AppStream and Egress store enabled, AppStream provides limited access to the workspace that is being used. You can view and use data within that workspace. When you work on this data, it can be saved in the Egress folder, and then a request can be made to export that data. But if you want to export data from that workspace, it requires approval.

In the following example, you don't currently have access to the `hello.txt` file out of your workspace. Hence, you need permission to export this file. The **Submit Egress Request** button shown in the figure below is used to export data from your workspace.

Research Workspaces

Create Research Workspace

All

Available

Pending

Stopped

Errored

Terminated

AVAILABLE

rstudio-server

Created 1 day ago by Tim Nguyen ddf05a71

Connections

View Detail

Egress Store

Stop

Terminate

Name	ETag	Project	Workspace Id	LastModified	Size	StorageClass
hello.txt	*61ed55766...	sdharwad-a...	ddf05a79-6...	2021-12-22...	12 Bytes	STANDARD

Submit Egress Request

cdrrth

Owner	Tim Nguyen
Studies	No studies linked to this workspace
Project	sdharwad-appsec-hosting
Workspace Type	EC2-RStudio-Server-v1
Configuration Name	sdharwad-i
Instance Type	t3.large

\$0.00

YESTERDAY'S COST

When you choose **Submit Egress Request**, the file gets picked up by the SNS topic. The SNS topic has following format:

```
<env-name>-<AwsRegionName>-<SolutionName>-EgressTopic
```

For example, it looks like this:

```
prod-iad-sw-EgressTopic
```

Amazon SNS > Topics

Topics (1)

Edit

Delete

Publish message

Create topic

Search

< 1 >

Name	Type	ARN
thingut1-sel-sw-EgressTopic	Standard	arn:aws:sns:us-east-1:123456789012:thingut1-sel-sw-EgressTopic

The SNS topic allows an administrator to determine which files you would like to export from the workspace. The administrator can subscribe to the SNS topic to receive updates when a

user submits an Egress request. For example, the administrator can subscribe a Lambda function to process the Egress submission request or could also provide the email address.

Amazon SNS > Subscriptions > Create subscription

Create subscription

Details

Topic ARN
arn:aws:sns: [X]

Protocol
The type of endpoint to subscribe

Select protocol ▲

- Amazon Kinesis Data Firehose
- Amazon SQS
- AWS Lambda
- Email
- Email-JSON
- HTTP
- HTTPS
- Platform application endpoint

► **Redrive policy (dead-letter queue) - optional**
Send undeliverable messages to a dead-letter queue. [Info](#)

Cancel Create subscription

After subscribing to the SNS topic, the administrator receives information in the following JSON file format:

```
{
  "egress_store_object_list_location": "<REDACTED>/sagemaker-11-egress-
store-ver1.json",
  "id": "<REDACTED>",
  "egress_store_id": "<REDACTED>",
  "egress_store_name": "sagemaker-11-egress-store",
  "created_at": "2021-10-15T19:03:09.367Z",
  "created_by": "<REDACTED>",
  "created_by_email": "<REDACTED>@amazon.com",
  "workspace_id": "<REDACTED>",
  "project_id": "proj-new-account-1",
  "s3_bucketname": "<REDACTED>",
  "s3_bucketpath": "<REDACTED>/",
  "status": "PENDING",
  "updated_by": "<REDACTED>",
  "updated_by_email": "<REDACTED>@amazon.com",
  "updated_at": "2021-10-25T18:45:02.883Z",
  "ver": 1
}
```

The `egress_store_object_list_location` section of the code provides the S3 location of a JSON file that lists the files for which the export request has been made by the researcher. For example, the JSON file appears as:

```
{
```

```

"objects": [
  {
    "Key": "ddf05a79-646d-442c-9dd3-88454753a1f2/",
    "LastModified": "2021-12-21T18:49:15.000Z",
    "ETag": "\"ea2dea32611f8e65498d75036cdbf88a\"",
    "Size": 0,
    "StorageClass": "STANDARD",
    "VersionId": "ihxQnhu6DNw2OlKctZjAyj.XxBmEOk4B",
    "Owner": {
      "ID":
"c3bc599bfa8ef4ddd90576205045dd986d755dbf80330012efa9bbb00b95d453"
    }
  },
  {
    "Key": "ddf05a79-646d-442c-9dd3-88454753a1f2/hello.txt",
    "LastModified": "2021-12-22T19:39:58.000Z",
    "ETag": "\"61ed55766b67aa6bbad881c12cf1d922\"",
    "Size": 12,
    "StorageClass": "STANDARD",
    "VersionId": "2W0H0AMVor2Zxhw22xOrayf4a9frGxXY",
    "Owner": {
      "ID":
"c3bc599bfa8ef4ddd90576205045dd986d755dbf80330012efa9bbb00b95d453"
    }
  }
]
}

```

This means that the researcher requested permission for exporting the hello.txt file, which is stored in the ddf05a79-646d-442c-9dd3-88454753a1f2 folder.

The file is stored in an S3 bucket of the main account in the following format:

[<aws-account-id>-<env-name>-<AwsRegionName>-<SolutionName>-egress-store](#)

As an administrator, you can then use the information above to approve the file export.

Create or add compute hosting accounts

After logging in as the **root** user for the first time, go to the **Accounts** page.

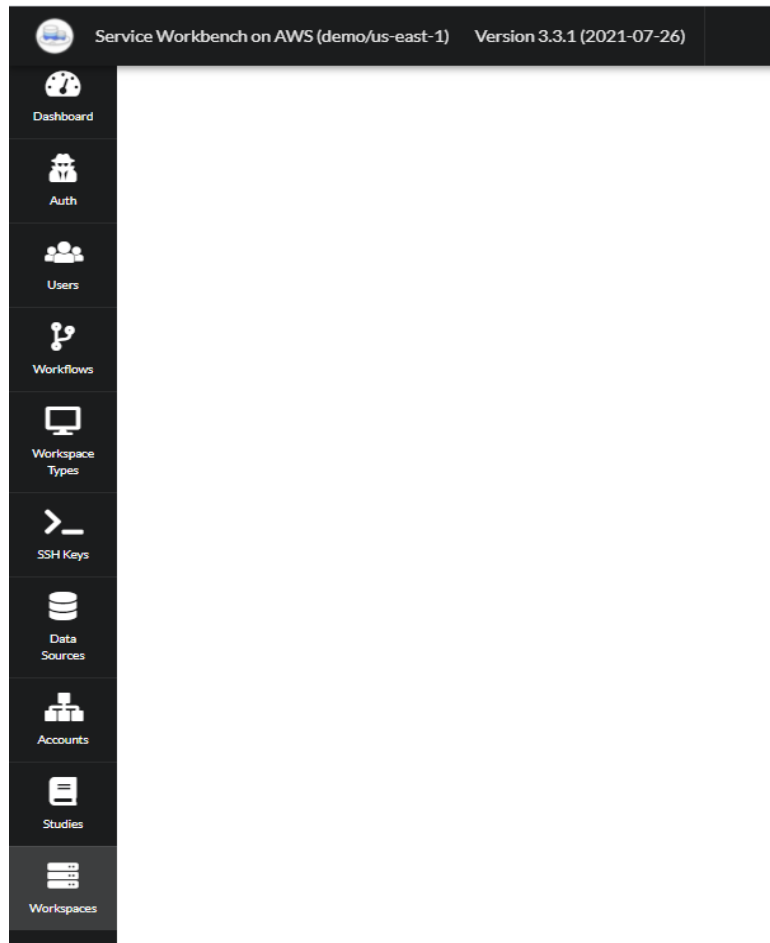


Figure: Service Workbench navigation bar

Service Workbench uses AWS accounts on this page for launching research workspaces. You can add existing AWS accounts or create new ones on the **Accounts** tab.

- **Create AWS Account:** Creates a new AWS account using AWS Organizations. Note that preparing organizational account steps outlined above are a pre-requisite should you require this capability.
- **Add AWS Account:** Associates an existing AWS account for purposes of hosting compute resources. This account can be responsible for its own billing.

Every user is linked to an account through a project and an index, so at least one account must be created or added before associating the first user to a project.

Note: If you do not need to create new AWS accounts from within Service Workbench, then skip to [Add AWS Account](#) section.

Create AWS Account

Prerequisites

Before creating an AWS account from Service Workbench, follow these guidelines:

- Configure an existing AWS account to be the organizational account for Service Workbench (steps outlined in *Preparing the organizational account* section). When Service Workbench creates new AWS accounts, billing for those accounts is applicable to the organizational account.

Creating a new Account

This creates a new **hosting** AWS account in the organization, whose billing goes to the **organizational** account.

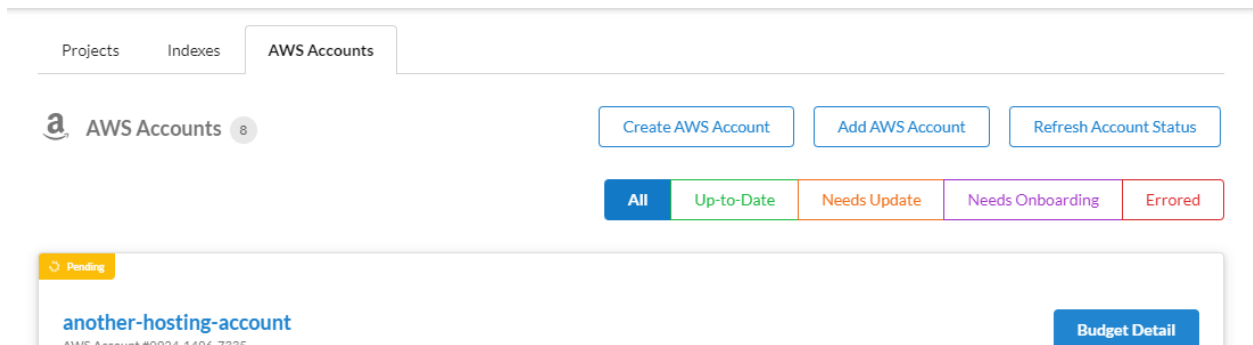


Figure: Create a new hosting account

To create an account:

1. In the Service Workbench console, choose **AWS Accounts** and then choose **Create AWS Account**.
 - In **Role ARN**, enter the **Master Role ARN** copied from the preparing the organizational account steps.
 - The email address that you specify here must be unique within the organization.
 - The **External ID** is **workbench** by default. See [IAM](#) for information on how to configure this to another value.
2. During processing, the following information displays in the **AWS Accounts** tab:
 - 'Trying to create accountID: xxx'
 - A workflow in progress in **Workflows** → **Provision Account** (see [Workflows](#))

Note: If instead you see an error message such as, 'Stop Internal State Account ID not found', check that there is an AWS Organization in the console of

your **organizational** account, if deploying Service Workbench in the **organizational** account. If you are deploying in a **Member** account, check and ensure that you followed the steps described in *Prepare Organizational Account*.

- Optionally, in the AWS console, you can inspect the following resources deployed by this script:
 - In AWS CloudFormation, a stack **prep-master** is running. It creates the **Master** role and its output is the **Master Role ARN**.
 - In the AWS Organization, in the **Organizational** account (see [IAM](#)), the new account displays.
 - In IAM, the new **Master** role is created.
- 3. Once the account is created it is listed in **AWS Accounts**.

Add AWS Account

Adding an existing AWS account enables Service Workbench to launch research workspaces. The existing account alignment (standalone or associated to an organization) determines the billing responsibility.

- On the **Accounts Page**, choose **AWS Accounts**, and then choose **Add AWS Account**.

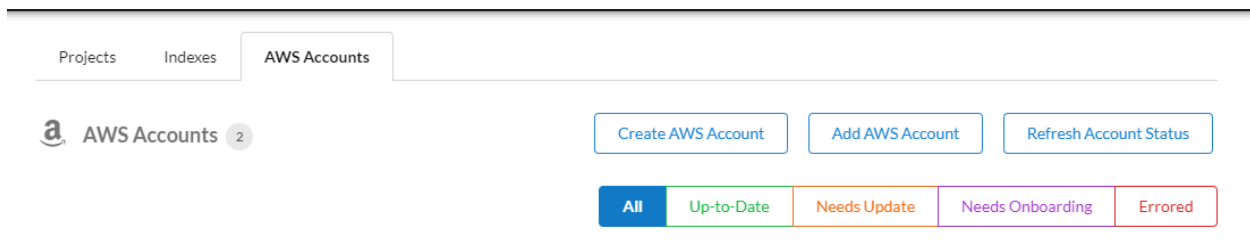


Figure: Add an existing account

- Create and run the AWS CloudFormation template:

If AppStream is not enabled

- Enter the **Account Name**, **Account ID** (12 digit AWS account ID for the account you wish to add), and **Description**.

Service Workbench on AWS (demo/us-east-1) Version 3.3.1 (2021-07-26)

Dashboard Auth Users Workflows Workspace Types SSH Keys Data Sources

Add AWS Account

Account Name
Type the name of this account

AWS Account ID
Type the 12-digit AWS account ID

Description
Type description for this AWS account

Figure: Specify account details

- The **Onboard AWS Account** screen is displayed, where you can select either of the following options:
 - **I have admin access** : You have administrator-level access to the hosting account that is being onboarded. Selecting this and then proceeding launches the CloudFormation template within the hosting account. Note that you need to be logged into the AWS account console for the hosting account when selecting this option and proceeding.
 - **I do not have admin access** : The CloudFormation template is generated and you can then share the template to be run by the party that does have administrator access in the AWS account that you are onboarding.

Onboard AWS Account

CloudFormation Stack Name

Initial-stack-1632323433789

AWSTemplateFormatVersion: 2010-09-09

Description: This stack provisions resources necessary to use this AWS account with Service Workbench.

Parameters:

EnableAppStream:

Type: String

AllowedValues: [true, false]

Description: Onboard this account to support AppStream

Namespace:

AWS Account # 888888888888

I have admin access I do not have admin access

Figure: Onboard AWS account

If AppStream is enabled

If you have chosen to enable AppStream for your installation, there are additional values required when onboarding a hosting account.

Add AWS Account

Account Name

AWS Account ID

Description

AppStream Fleet Desired Instance

AppStreamDisconnectTimeoutSeconds

AppStreamIdleDisconnectTimeoutSeconds

Figure: Add account when AppStream is enabled

- **AppStream Fleet Desired Instance:** The maximum number of concurrently running AppStream sessions allowed. If you set this to 5, that would mean that five workspaces can be viewed concurrently.
- **AppStreamDisconnectTimeoutSeconds:** With a minimum of 60 seconds, this is the amount between a researcher disconnection from a session (Manual Stop, Auto Stop, or Terminate) and the release of the AppStream instance that is supporting that session.
- **AppStreamIdleDisconnectTimeoutSeconds:** The amount of time that an AppStream session idle time (meaning no activity within the session) before the AppStream instance disconnects.

AppStreamIdleDisconnectTimeoutSeconds

The amount of time that users can be idle (inactive) before they are disconnected from their streaming session

AppStreamMaxUserDurationSeconds

The maximum amount of time that a streaming session can remain active, in seconds

AppStreamImageName

The name of the image used to create the fleet

AppStreamInstanceType

The instance type to use when launching fleet instances. List of images available at <https://aws.amazon.com/appstream2/pricing/>

AppStreamFleetType

The fleet type. Should be either ALWAYS_ON or ON_DEMAND

Cancel

Onboard AWS Account

Figure: Add account when AppStream is enabled (contd..)

- **AppStreamMaxUserDurationSeconds:** The maximum amount of time for an AppStream session.
- **AppStreamImageName:** The exact image name produced when you follow the instructions to [build your AppStream image](#)
- **AppStreamInstanceType:** The instance type for your AppStream fleet. Note that these instance types are unique to AppStream. A complete list and specifications for valid instance types is available at <https://aws.amazon.com/appstream2/pricing/>
- **AppStreamFleetType:** ALWAYS_ON ensures that the desired number of instances remain available at all times. ON_DEMAND only runs the instances when required. It is a cost versus convenience choice.

Note: If you needed to change these values later, you can do so through the AWS Console of the hosting account without negative impact to Service Workbench.

Once these options are specified, choosing **Onboard Account** displays the **Onboard AWS Account** screen. The same choice of Admin vs. No Admin access applies, but there are several important pre-requisites to complete before proceeding.

Creating RStudio ALB workspace

Accessing RStudio workspace

You can access the RStudio workspace type by using the template and AMI provided in AWS partner's [repository](#). RStudio legacy and RStudio ALB both use custom domain names. You can define custom domain name in `/main/config/settings/<stage>.yaml`.

Important: The legacy RStudio workspace type will be deprecated soon in a future release.

Application load balancing for RStudio ALB workspace

When you create RStudio ALB workspace, an ALB is created that can host upto 100 workspaces per hosting account. Per workspace, you can have upto 4 CIDR blocks (IP ranges) for port 443 only. This is updated in the hosting accounts. If there are no ALBs for that hosting account and you have created the very first workspace, then ALB will be created for the first time and it will be common for every workspace that you create subsequently. Once you terminate all the workspaces and you terminate the last workspace for that hosting account, the ALB is deleted. An ALB is created for minimum one active workspace.

The screenshot shows the 'Rules' tab for an ALB listener on port 443. It lists five rules:

Rule Number	ARN	IF Conditions	THEN Action
1	arn...e88cd	Host is rstudio-people.aws.dev Source IP is 54...	Forward to analysis-1634... (100%) Group-level stickiness: Off
2	arn...502f2	Source IP is 67... Host is rstudio-9057825b-95a3-people.aws.dev	Forward to analysis-163... (100%) Group-level stickiness: Off
3	arn...931a0	Host is rstudio-4d007f-people.aws.dev Source IP is 67.172...	Forward to analysis-1... (100%) Group-level stickiness: Off
4	arn...c8a48	Source IP is 67... Host is rstudio-f143fbaa...	Forward to analysis-... (100%) Group-level stickiness: Off
last	HTTPS 443: default action	Requests otherwise not routed	Return fixed response 403 (more...)

Each workspace type corresponds to one listener rule and there can be 100 such rules. In each rule, users can specify the IP ranges that user wants to allowlist. You can specify only upto 4 IP

ranges per listener rule. For more information about ALB, refer to [Application Load Balancer \(ALB\)](#).

Limitations

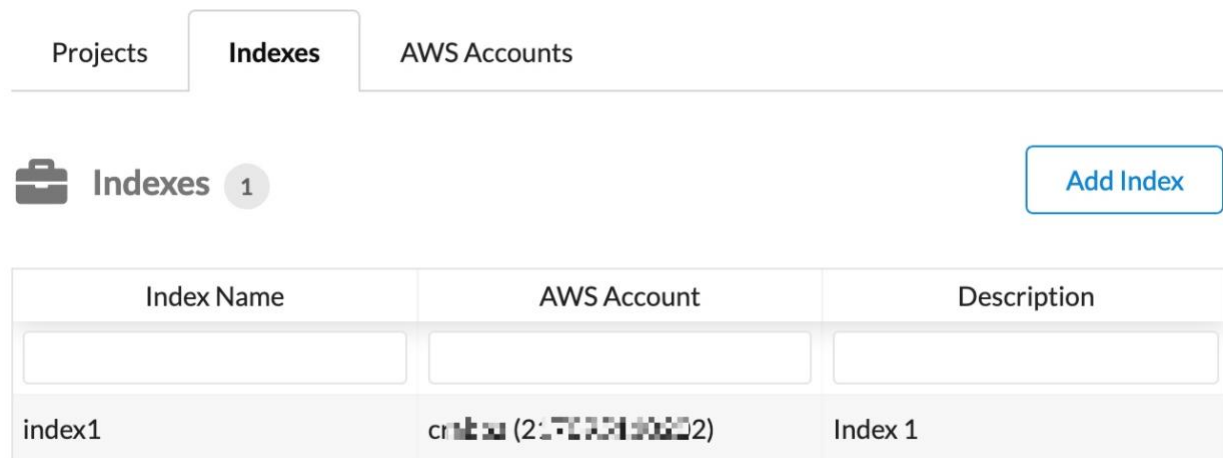
1. With RStudio ALB implementation, you can create only upto 100 workspaces per hosting account.
2. Per workspace, you can specify only upto 4 CIDR blocks for port 443.

Create Indexes and Projects


Projects and **Indexes** form a hierarchy under **Accounts**. Each Account can have multiple **Indexes**, each **Index** can have multiple **Projects**. **Projects** are attached to **Users**, so you must create the **Projects** first.

After you create an [account](#) in the **Accounts** tab of the administrative interface, create an **Index** that links to the Account, by selecting the **Account ID** from the drop-down list.

1. On the **Indexes** tab, click **Add Index**. See the figure below.



Projects	Indexes	AWS Accounts
----------	---------	--------------

 **Indexes** 1 Add Index

Index Name	AWS Account	Description
<input type="text"/>	<input type="text"/>	<input type="text"/>
index1	crlbx (2:713219022)	Index 1


Figure : Create an Index

2. Create a **Project** that links to the new Index. In the **Projects** tab, choose **Add Project**.

Projects

Indexes

AWS Accounts

 Projects 1

Add Project

Project Name	Index Id	Description
<input type="text"/>	<input type="text"/>	<input type="text"/>
project1	index1	Project 1

Figure: Create a Project

Create an Administrator User

Once you create an [account](#) and an [index and project](#), you must create an administrator user in the **Users** tab.

Dashboard


Auth

Users

API Keys

Users

Roles

 Users 6

Add Local UserAdd Federated User






Name	Email	Identity Provid...	Type	Role	Project	Status	
root	 @amazon.com	internal	Internal	admin	<<none>>	 Active	Detail
 @amazon.com	 @amazon.com	internal	Internal	admin	project1	 Active	Detail

Figure: Create an Administrator

Note: A root user account will already be created, however, you must not routinely use the root user account.

For testing purposes, you can create a local user by clicking ‘**Add Local User**’. Assign the user the administrator’s role, and associate the user with the **Project** you created, and set the status to ‘**Active**’. See **Figure 8**.

Add Local User



Not for production usage

Creating local users is not meant to be used in production environments.

Username

Username in email format

First Name

Last Name

Password

UserRole

Select user's role

Projects

Select projects that this user are associated with

Status

☒ Active

or

☐ Inactive

Active users can log into the Research Portal

Figure: Add Local User

In prod environments we highly recommend using an IDP. For more details, refer to the **Service Workbench Configuration Guide**.

Import Service Catalog Products

Service Workbench uses [AWS Service Catalog](#) to manage different types of computation resources available for researchers to use through the platform.

With AWS Service Catalog integration, Service Workbench allows Admin users to create and manage catalogs of IT services that are approved for use on AWS. These IT services can include everything from virtual machine images, servers, software, and databases to complete multi-tier application architectures.

With this integration, Service Workbench helps organization to centrally manage commonly deployed IT services, and helps achieve consistent governance and meet compliance requirements, while enabling users to quickly deploy only the approved IT services they need.

When Service Workbench is deployed, an AWS Service Catalog portfolio is created by default with four commonly used products: Amazon SageMaker, Amazon EC2 for Windows, Amazon EC2 for Linux and Amazon EMR. The **administrator** needs to import and configure these products using Service Workbench user interface before they can be deployed. If you want to include additional custom products in the AWS Service Catalog portfolio, complete these steps:

1. Add the AWS CloudFormation template in the following directory:

```
addons/addon-base-raas/packages/base-raas-cfn-templates/src/templates/service-catalog
```

2. Add the AWS CloudFormation template file name in the productsToCreate list in the following location:

```
addons/addon-base-raas/packages/base-raas-post-deployment/lib/steps/create-service-catalog-portfolio.js
```

3. Deploy Service Workbench.

Import a Product

In this step, you import a pre-defined product, configure parameters to be used for product launch, and approve the configured product to be used. The following sections use Amazon EC2 Linux as an example, followed by setting different configuration required for Amazon EC2 Windows, Amazon SageMaker and Amazon EMR.

Prerequisites

Ensure the following prerequisites are met in order to import a product.

Creating AMI

Make sure you completed the step, deploy the Machine Images SDC as part of the deployment process.

To check if AMIs were created successfully:

1. Navigate to Amazon EC2.
2. Select the **AMI** tab.
3. Note down the four AMIs created for Amazon EC2 Linux, Amazon EC2 Windows, Amazon EMR, and Amazon EC2 RStudio.
4. Copy the AMI IDs and use for workspace import and configuration. Alternatively, you can also copy these AMI IDs from the terminal when the machine-images SDC is deployed.

Note: *If you run the machine images SDC multiple times, duplicated AMIs are created. This is okay and will not affect any Service Workbench functionalities. You can choose to remove the duplicates to avoid confusion or leave them.*

Viewing AWS Service Catalog portfolio

1. Log in to Service Workbench UI as an **administrator**.
2. Navigate to **Workspace Types** tab. Four AWS Service Catalog Products display as shown in **Figure 9**.



Available AWS Service Catalog Product Versions not imported yet

EC2 Windows-V1.0.0

V1.0.0

- An EC2 Windows instance with RDP access
- Secure compute in the cloud

Import ↓

SageMaker Notebook-V1.0.0

V1.0.0

An Amazon SageMaker Jupyter Notebook that comes with:

- TensorFlow
- Apache MXNet
- Scikit-learn

Import ↓

EC2 Linux-V1.0.0

V1.0.0

- An EC2 Linux instance with SSH access
- Secure compute in the cloud

Import ↓

EMR-V1.0.0

V1.0.0

An Amazon EMR research workspace that comes with:

- Hail 0.2
- Jupyter Lab
- Spark 2.4.4
- Hadoop 2.8.5

Import ↓

Figure: AWS Service Catalog Products

These four products come from the AWS Service Catalog portfolio created by the system during deployment. And they'll be ready for use once imported and configured.

If you wish to include other AWS computation resources in the future:

1. Add a new product to the existing Service Workbench portfolio in AWS Service Catalog
2. Update the role `ServiceCatalogLaunchConstraintRole` in [cloudformation.yml](#) to include permission needed to launch and terminate the product

Importing a workspace

In this section, the Amazon EC2 Linux is used as an example.

1. Choose the **Import** button under `ec2-linux-instance`.
2. Enter **Name** and **Description** so you can easily identify the workspace.

Configuring the workspace

Once you import a workspace type, perform the following actions:

1. Choose **Add Configuration**
2. Enter **ID**, **Name**, **Description**, and **Estimated Costs** for the configuration. A common naming convention here is to attach the instance size after the product name. For example, use ec2-linux-instance-V1-small for a small Linux Amazon EC2 instance.
3. Choose **Next**.
4. Add access control for the workspace configuration.
5. Choose **Next**.

The input parameters are parameters used for the product, AWS CloudFormation template. The number and type of parameters are different for different products. Most of the parameters used for the four system created products can be evaluated automatically at launch time. These parameters are available for selection in the drop-down when filling the input parameters page.

Configuration for EC2 Linux

For Amazon EC2 Linux, the only two fields that are not available in the drop-down are **InstanceType** and **Amild**.

The following figures display screenshot images that exemplify Amazon EC2 Linux configurations.



Edit Configuration

first

Basic Information

Access Control

Input Parameters

Tags

EncryptionKeyArn

The ARN of the KMS encryption Key used to encrypt data in the instance

`${encryptionKeyArn}`



IamPolicyDocument

The IAM policy to be associated with the launched workstation

`${iamPolicyDocument}`



AccessFromCIDRBlock

The CIDR used to access the ec2 instances.

`${cidr}`



VPC

The VPC in which the EC2 instance will reside

`${vpclid}`



Figure: Configurations for Amazon EC2 Linux

The screenshot displays a series of configuration fields for an Amazon Linux EC2 instance. Each field has a title, a description, and a text input box with a clear button (X).

- EnvironmentInstanceFiles**: An S3 URI (starting with "s3://") that specifies the location of files to be copied to the environment instance, including any bootstrap scripts. The input field contains `${environmentInstanceFiles}`.
- InstanceType**: EC2 instance type to launch. The input field contains `t3.large`.
- Subnet**: The VPC subnet in which the EC2 instance will reside. The input field contains `${subnetId}`.
- S3Mounts**: A JSON array of objects with name, bucket, and prefix properties used to mount data. The input field contains `${s3Mounts}`.
- Namespace**: An environment name that will be prefixed to resource names. The input field contains `${namespace}`.
- Amild**: Amazon Machine Image for the EC2 instance. The input field contains `ami-0d1b1e3a0e3d9021c`.

Figure: Configurations for Amazon Linux EC2

Configuration for Amazon EC2 Windows

For Amazon EC2 Windows, the only two fields that are not available in the drop-down are **InstanceType** and **Amild**. (Use the AMI ID you copied in Prerequisites - AMI)

The following figures display screenshot images that exemplify Amazon EC2 Windows configurations.

Add Configuration

1 Basic Information
Enter basic information

2 Access Control
Define who can access

3 Input Parameters
Provide AWS CloudFormation Inputs

4 Tags
Specify Resource Tags

DownloadInterval

An interval in seconds to wait between two downloads in case of recurring downloads. This is only applicable when RecurringDownloads is set to "true". Note that this does not include the download time. This specifies the duration in seconds to wait before initiating the next download after the previous one completes.

20

RecurringDownloads

A flag indicating whether to keep syncing studies data to local EBS volumes on recurring basis. Setting this to false will download studies data only once at the instance bootstrap time. When this flag is set to true the instance will periodically sync changes from S3 to local EBS i.e., it will download any new files added to S3, re-download any files changed in S3 (will use object ETag value to determine if file changed in S3), delete files from local EBS if they are deleted from S3.

true

EncryptionKeyArn

The ARN of the KMS encryption Key used to encrypt data in the instance

\${encryptionKeyArn}

AccessFromCIDRBlock

The CIDR used to access the ec2 instances.

\${cidr}

VPC

The VPC in which the EC2 instance will reside

\${vpcid}

Figure: Configurations for EC2 Windows

S3Mounts	
A JSON array of objects with name, bucket, and prefix properties used to mount data	
<input type="text" value="{s3Mounts}"/>	✕
Namespace	
An environment name that will be prefixed to resource names	
<input type="text" value="{namespace}"/>	✕
KeyName	
Keypair name for admin password encryption/decryption	
<input type="text" value="{adminKeyPairName}"/>	✕
RaidDataVolumeSize	
The size of each volume in the RAID array used to hold studies data, in GiB. The template creates a striped volume (RAID 0) by joining 8 volumes. The total size of the data volume would be roughly 8 times the size specified here.	
<input type="text" value="10"/>	✕
StopRecurringDownloadsAfter	
Duration in seconds after which to stop the recurring downloads. Value of -1 means keep doing the recurring downloads (sync) indefinitely.	
<input type="text" value="-1"/>	✕
IamPolicyDocument	
The IAM policy to be associated with the launched workstation	
<input type="text" value="{iamPolicyDocument}"/>	✕
EnvironmentInstanceFiles	
An S3 URI (starting with "s3://") that specifies the location of files to be copied to the environment instance, including any bootstrap scripts	
<input type="text" value="{environmentInstanceFiles}"/>	✕

Figure: Configurations for EC2 Windows

The screenshot displays a configuration window for an Amazon SageMaker EC2 instance. It contains three input fields, each with a title, a description, and a value. The 'Instance Type' field is set to 'r5.2xlarge'. The 'Subnet' field is set to '\$[subnetId]'. The 'Amild' field is set to 'ami-0#####'. At the bottom, there are three buttons: 'Cancel', 'Previous', and 'Next'.

Field	Description	Value
Instance Type	EC2 instance type to launch	r5.2xlarge
Subnet	The VPC subnet in which the EC2 instance will reside	\$(subnetId)
Amild	Amazon Machine Image for the EC2 instance	ami-0#####

Figure: Configurations for EC2 Windows

Configuration for Amazon SageMaker

For Amazon SageMaker, the only field that's not available in the drop-down is '**InstanceType**'.

The following figures display screenshot images that exemplify Amazon SageMaker configurations.

Add Configuration

1 Basic Information
Enter basic information

2 Access Control
Define who can access

3 [Input Parameters](#)
Provide AWS CloudFormation Inputs

4 Tags
Specify Resource Tags

EncryptionKeyArn

The ARN of the KMS encryption Key used to encrypt data in the notebook

`${encryptionKeyArn}`

✕

IamPolicyDocument

The IAM policy to be associated with the launched workstation

`${iamPolicyDocument}`

✕

VPC

VPC for EMR nodes.

`${vpcId}`

✕

AccessFromCIDRBlock

The CIDR used to access sagemaker.

`${cidr}`

✕

EnvironmentInstanceFiles

An S3 URI (starting with "s3://") that specifies the location of files to be copied to the environment instance, including any bootstrap scripts

`${environmentInstanceFiles}`

✕

InstanceType

EC2 instance type to launch

`ml.t3.medium`

✕

Subnet

Subnet for EMR nodes, from the VPC selected above

`${subnetId}`

✕

Figure: Configurations for Amazon SageMaker

S3Mounts
A JSON array of objects with name, bucket and prefix properties used to mount data

`$(s3Mounts)` ✕

Namespace
An environment name that will be prefixed to resource names

`$(namespace)` ✕

AutoStopIdleTimeInMinutes
Number of idle minutes for auto stop to shutdown the instance (0 to disable auto-stop)

15 ✕

Cancel Previous Next

Figure: Configurations for Amazon SageMaker

Configuration for Amazon EMR

Amazon EMR requires a few more fields that are not available in the drop-down menu, including the following:

- DiskSizeGB (≥ 10)
- CoreNodeCount (1-80)
- MasterInstanceType
- Market (ON_DEMAND / SPOT)
- WorkerBidPrice (only applicable when Market = SPOT. Specify 0 for Market = ON_DEMAND)
- WorkerInstanceType
- Amild (Use the AMI id we copied in prerequisites - AMI)

The following figures display screenshot images that exemplify Amazon EMR configurations.

DiskSizeGB	EBS Volume size (GB) for each node	10	✕
CoreNodeCount	Number of core nodes to provision (1-80)	1	✕
EncryptionKeyArn	The ARN of the KMS encryption Key used to encrypt data in the cluster	`\${encryptionKeyArn}`	✕
VPC	VPC for EMR nodes.	`\${vpcId}`	✕
AccessFromCIDRBlock	Restrict WebUI access to specified address or range	`\${cidr}`	✕
MasterInstanceType	EMR node ec2 instance type.	c5.xlarge	✕

Figure: Configurations for Amazon EMR

Market
Which market to purchase workers on - ON_DEMAND or SPOT.
ON_DEMAND
S3Mounts
A JSON array of objects with name, bucket and prefix properties used to mount data
`\${s3Mounts}`
Namespace
An environment name that will be prefixed to resource names
`\${namespace}`
KeyName
SSH key pair to use for EMR node login
`\${adminKeyPairName}`
IamPolicyDocument
The IAM policy to be associated with the launched workstation
`\${iamPolicyDocument}`
WorkerBidPrice
Bid price for the worker spot nodes. This is only applicable when Market = SPOT. Specify 0 for Market = ON_DEMAND.
0

Figure: Configurations for Amazon EMR

EnvironmentInstanceFiles

An S3 URI (starting with "s3://") that specifies the location of files to be copied to the environment instance, including any bootstrap scripts

\$(environmentInstanceFiles)

✕

Subnet

Subnet for EMR nodes, from the VPC selected above

\$(subnetId)

✕

WorkerInstanceType

EMR node ec2 instance type.

c5.xlarge

✕

AmiId

Ami Id to use for the cluster

ami-0a2b3c4d5e6f7g8h9i0j

✕

Cancel

Save

Figure: Configurations for Amazon EMR

Approving the workspace

Once the configuration completes, choose the **Approve** button; the newly created workspace type will be available for launch in the **Study and Workspace** tab.

Viewing logs

Viewing Service Workbench logs in CloudWatch

Service Workbench has API Gateway access logging enabled. The logs are available in CloudWatch at the `/aws/api-gateway/<name of your API>` log group:

Following is the format of the access logs:

```
{ "authorizer.principalId": "u-000000000000", "error.message": "-", "extendedRequestId":  
"ZuT4rGDNoAMFxXw=", "httpMethod": "GET", "identity.sourceIp": "22.22.22.22", "integration.error": "-",  
"integration.integrationStatus": "200", "integration.latency": "79", "integration.requestId": "67394741-90ae-
```

```
4c6c-94fb-df8bf7be33ec", "integration.status": "200", "path": "/dev/api/user-roles", "requestId": "468a1b4d-3015-4901-b749-37e4e0551029", "responseLatency": "83", "responseLength": "819", "stage": "dev", "status": "200"}
```

Lambda logs are also available in CloudWatch with the default log group names `/aws/lambda/<lambda function name>`.

Metrics

The default metrics for Lambda and API Gateway are available in CloudWatch. For the full list of available metrics, see:

- [Working with AWS Lambda function metrics - AWS Lambda](#)
- [Amazon API Gateway dimensions and metrics - Amazon API Gateway](#)

Service Workbench does not emit any custom metrics.

Deploying Updates

After a successful initial deployment, you can deploy individually to the five serverless projects that are a part of this solution.

Deploying Updates to the Infrastructure Serverless Project

```
$ cd solution/infrastructure$ pnpm sls deploy -s <stage>
```

Deploying Updates to the Backend Serverless Project

```
$ cd solution/backend$ pnpm sls deploy -s <stage>
```

Deploying Updates to the Machine-Images Serverless Project

```
$ cd solution/machine-images$ pnpm sls deploy -s <stage>
```

Deploying Updates to the Post-Deployment Serverless Project

```
$ cd solution/post-deployment$ pnpm sls invoke local -f postDeployment --env WEBPACK_ON=true -s <stage>
```

Deploying Updates to the UI Serverless Project

```
$ cd solution/ui$ pnpm sls package-ui --stage <stage> --local$ pnpm sls package-ui --stage <stage>$ pnpm sls deploy-ui --stage <stage> --invalidate-cache
```

Reference

Service Workbench on AWS interacts with multiple AWS resources, including Amazon EC2, AWS IAM, AWS Organizations, and more. You can easily add an AWS IAM role to an Amazon EC2 instance, leverage our AWS Organizations with your account structure, and create multiple Amazon S3 buckets.

Adding an AWS IAM role to an Amazon EC2 instance

An Amazon EC2 instance can be assigned an instance profile that contains an AWS IAM role. The AWS IAM role gives the Amazon EC2 instance a set of permissions. The Amazon EC2 instance only performs the actions defined by its AWS IAM role. Adding an AWS IAM role to the Amazon EC2 instance allows your application to make API calls securely—eliminating the need to manage security credentials.

The Service Workbench deployment application must be able to create AWS resources. The easiest way to meet this requirement is to give the Amazon EC2 instance an administrator's role.

Adding administrator's role to a new Amazon EC2 instance

When creating a new Amazon EC2 instance for the Service Workbench deployment, an instance profile may be assigned to the Amazon EC2 instance. Choose **Create a new IAM role** located next to the **IAM role** drop-down.



Figure: Create a new AWS IAM role

To continue the process, highlight Amazon EC2 and proceed to permissions.

1. Under **Permissions**, filter for **AdministratorAccess**.

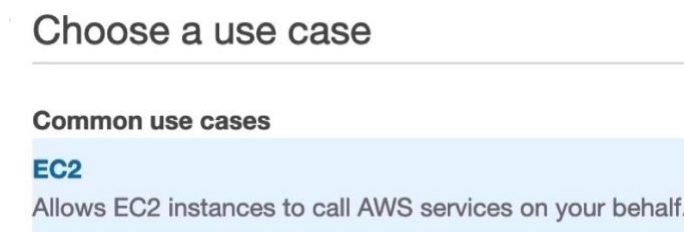


Figure: Permissions in Amazon EC2

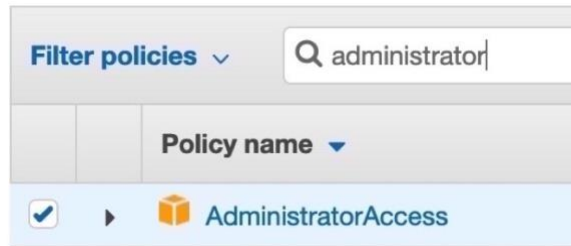


Figure: Filtering for AdministratorAccess

2. Proceed through **Tags**.
3. In the **Review** page, enter the role name.

Role name*

Figure: Choosing a role name for an Amazon EC2 instance

4. In the **Amazon EC2** tab, refresh the AWS IAM role drop-down, and choose the administrator role that is to be attached to the new Amazon EC2 instance.


IAM role ⓘ ✓ None
ec2-admin 

Figure: Selecting the administrator role of the Amazon EC2 instance

5. Now, proceed through the process to create an Amazon EC2 instance.

Adding a role to an existing instance

To add a role to an Amazon EC2 instance that is already running, select the Amazon EC2 instance in the EC2 Console. Open the **Action > Instance Settings** menu, and select **Attach/Replace IAM Role**.

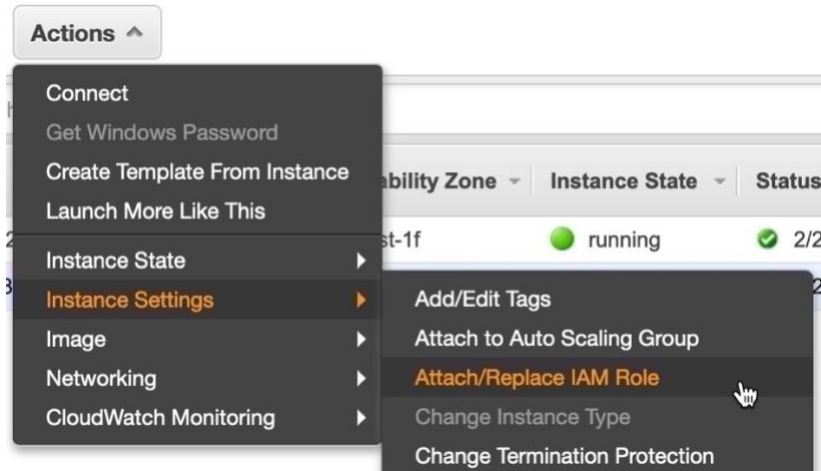


Figure: Attach/Replace an AWS IAM Role in the EC2 Console

In the **Attach/Replace IAM Role** screen, search for the role you created, select it, and choose **Apply**. **Figure 30** shows the screen where you can **Attach/Replace IAM Role**.

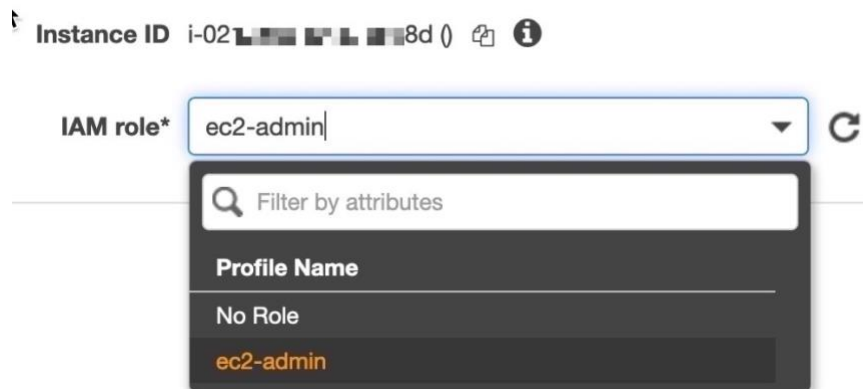


Figure : AWS IAM role search

Usage of AWS Cloud Services

This section describes some of the AWS Cloud services used by Service Workbench. The resource names usually include the **Namespace**, including the [Stage Name](#) used at deployment. You can deploy multiple instances of Service Workbench from the same account if you use a different Stage Name for each deployment.

Amazon EC2

Amazon EC2 is used only as a platform from which to deploy Service Workbench. For more details see the [Deployment Instance](#) section.

AWS IAM

Service Workbench creates several roles in your account. The role `<namespace>-prep-raas-master-MasterRole-XXX` is created when you run the [Post Deployment](#) SDC. This role possesses a trust relationship with the Main account from which you deployed Service Workbench. There are two policies that allow the Main account to assume a role in this Management account. The [Account Structure](#) defines each type of account. The following figure shows the AWS IAM **‘Trust Relationships’** tab.

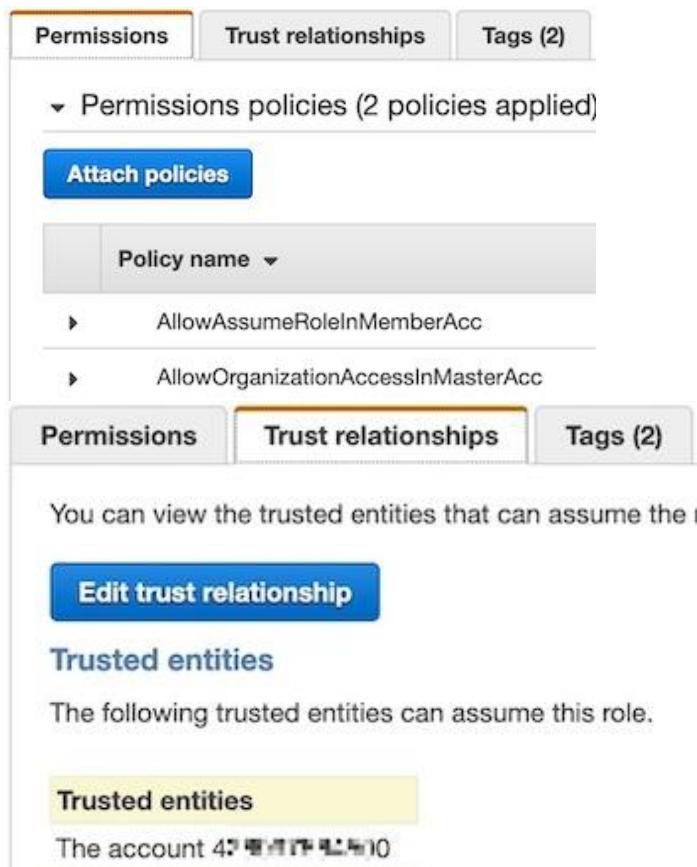


Figure: AWS IAM Trust Relationships Tab

An [External ID](#) is associated with the role. The External ID is an identifying string that is provided once a role is created. In order for the Trusted Entity (your Main account) to assume its role in the management account, it must supply this External ID. Providing the External ID of establishes a revocable relationship between the Trusted Entity and the Management account.

In the current Service Workbench deployment, the External ID is configured as a default value in the following string workbench:

```
main/solution/prepare-master-acc/config/settings/.defaults.yml
```

To change this value, create a stage-named configuration file (`mystagename.yml`) in the same directory. For more information, see the [Configuration](#) section. The following figure displays a screenshot image of the conditions that define how **Trusted Entities** assume a role.

Conditions

The following conditions define how and when trusted entities can assume the role.

Condition	Key	Value
StringEquals	sts:ExternalId	...

Figure: Defining Conditions for Trusted Entities

AWS Organizations

An AWS Organization is created in the **Management account**. The **Management account** is discussed in the Account Structure section of the Service Workbench User Guide . The AWS Organization use the **Management account** to create a separate account for each deployment. The account's name is the **Stage Name** used. The following figure shows a screenshot image of the AWS Organizations 'Accounts' tab.

AWS Organizations				
Accounts Organize accounts Policies				
<div>Add account Remove account <input type="checkbox"/> Hide Failed account creation requests <input type="text" value="Filter"/></div>				
<input type="checkbox"/>	Account name	Email	Account ID	Status
<input type="checkbox"/> ★@amazon.com	4-...	Joined on 5/25/20
<input type="checkbox"/>@amazon.com	100-...	Created on 6/18/20

Figure: AWS Organizations Account Page

Amazon S3

Multiple Amazon S3 buckets are created by Service Workbench. Filtering by **Stage Name** shows the Amazon S3 buckets for a deployment. The following figure shows the Amazon S3 buckets for the Service Workbench deployment.

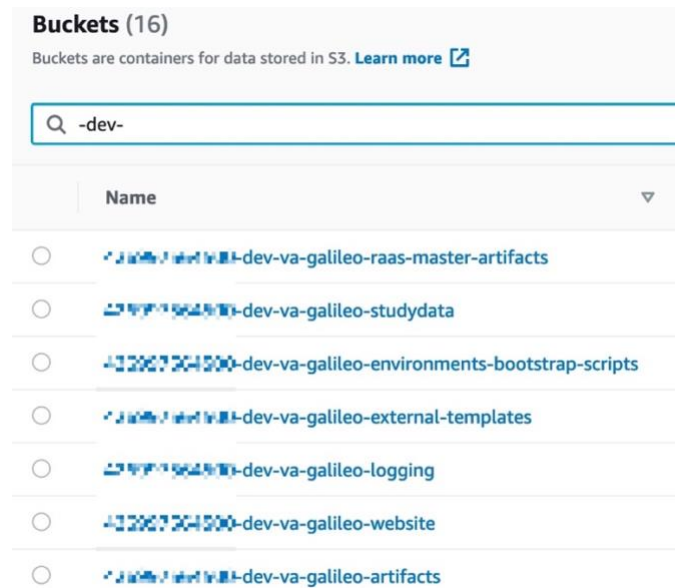


Figure: Amazon S3 Buckets for a Service Workbench Deployment

The 'studydata' bucket contains all the data for the various studies in this deployment at the individual and organization level. The following figure displays an image of the contents within the studydata bucket.

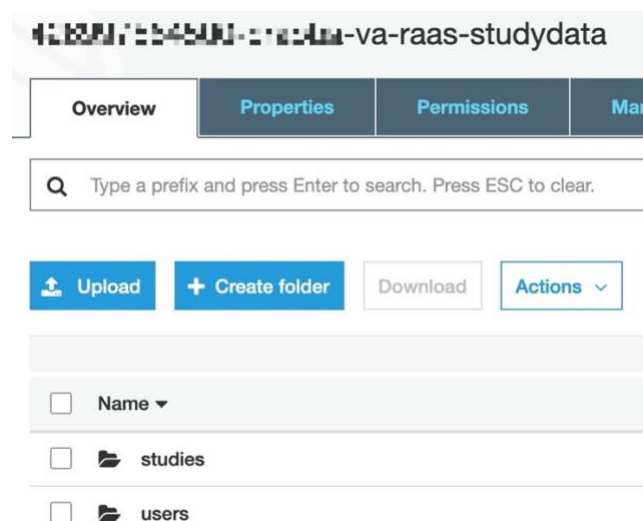


Figure: Amazon S3 StudyData Bucket

AWS Cost Explorer

Service Workbench has the ability to show actual cost incurred by workspaces running under the Management account. This is using the AWS Cost Explorer service in the AWS Management Console. AWS Cost Explorer must be manually set up for each Management account once in order to allow requests for cost data to process. Setting this up requires background processes to complete in the Management account, which can take up to 24 hours.