Project Documentation

BeatDrift-CLI Automation Tool Implementation

COSC595

Trimester 1 2022

Jason O. Aboh

Paldeep Kaur

Sumit Khokhar

Table of Contents

Project Description

The project is oriented towards creating a CLI tool which will be used to build a pipeline to address drift in an Infrastructure as Code (IaC) environment.

Client

Client: Mark Wallis, Director of Hunter Orbit Pty Ltd.

Email: mark@hunterorbit.com.au

1.2. Project Team

The team members are:

Name	**Email id (s)**
Jason Aboh	jaboh@myune.edu.au , bigbossw107@gmail.com
Paldeep Kaur	pkaur21@myune.edu.au , paldeepkaur1995@gmail.com
Sumit Khokhar	skhokha2@myune.edu.au , sumitkhokhar359@gmail.com

2.0. Stakeholders

The stakeholders of the project are as follows:

Client: **Mark Wallis**

Unit Coordinator (COSC 595): **Dr. Edmund Sadgrove**

Team Member (Project Lead): **Jason O. Aboh**

Team Member: **Sumit Khokhar**

Team Member: **Paldeep Kaur**

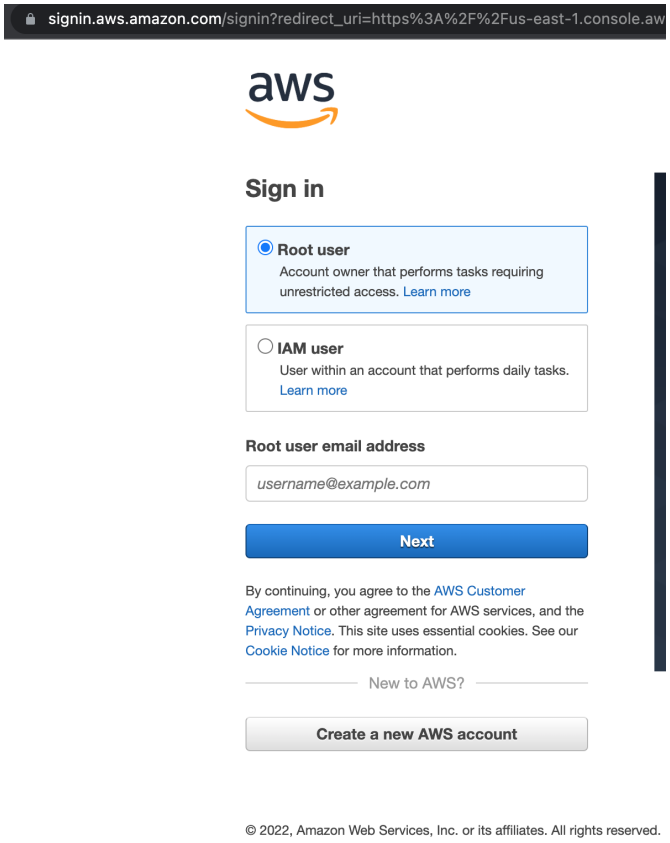
3.0. Installation Manual and User manual:

3.1 Creating Resources in AWS for testing (crawling)

After creation of an aws account and set up of credentials, to access the given aws; navigate to the aws Iam console website with the designated region at the beginning of the url link, the below link takes us to the "us-east-1" region:

https://us-east-1.console.aws.amazon.com/iamv2/home?region=us-east-1#/home

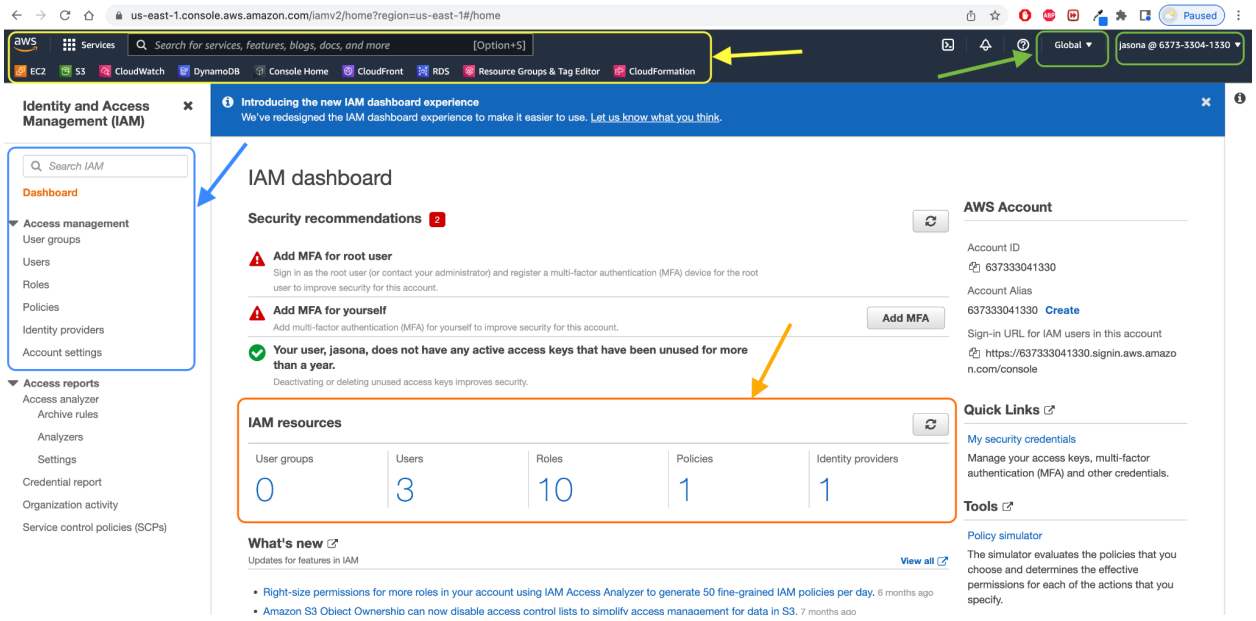
The user shall arrive at the page below; which they will be required to enter the credentials for the AWS account they have access to:



The user will need to input credentials for a “Root” user account (performs tasks with unrestricted access), or as a “IAM” user account (a user within an account that performs daily tasks).

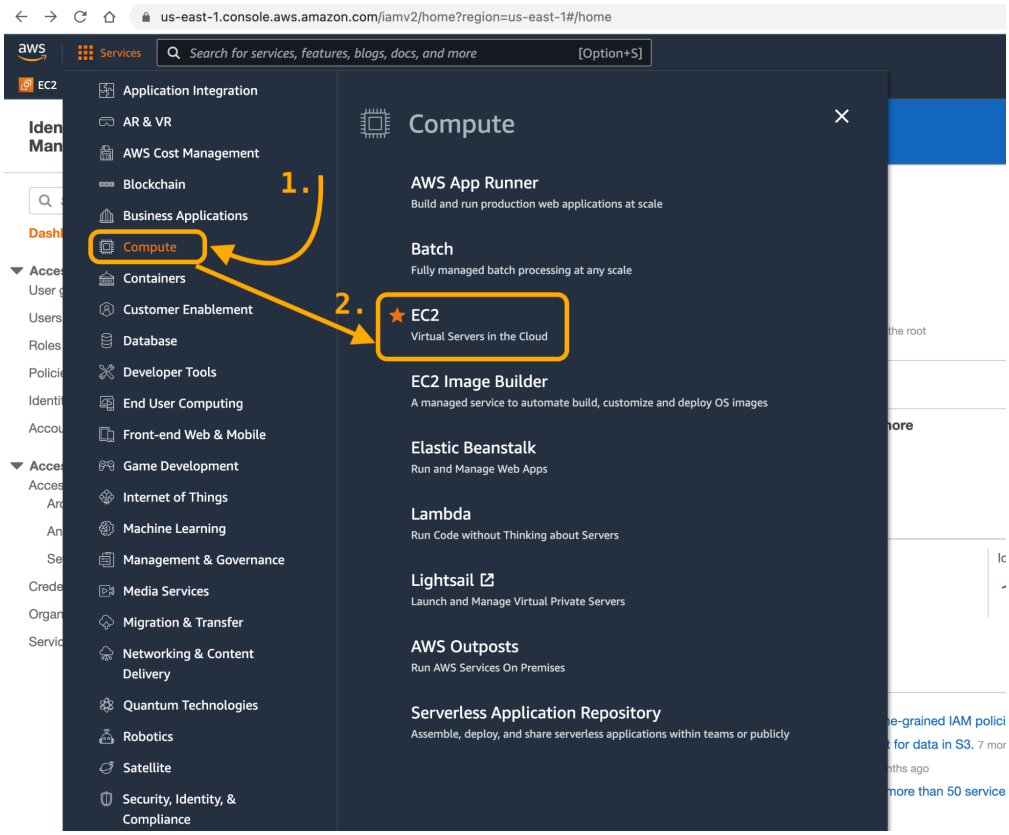
If the user supplies an alias (account name) or account ID number rather than the email address of a “Root” user account, they are taken to the IAM user page to sign in using an alias or account ID as shown below:

On successful sign in, the user arrives at the IAM dashboard as shown below:

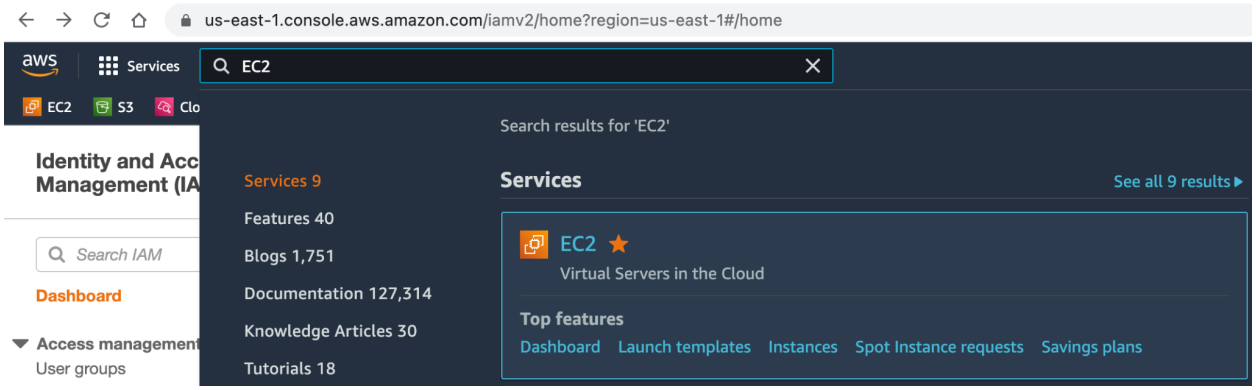


Please note that the IAM access management settings are located to the left within the blue rectangle, access to the AWS services is located to the top left within the yellow rectangle, access to IAM Users, Roles, Policies and Identity providers are located within the orange rectangle in the middle of the page; then finally, the current account username and ID , along with the region it is signed into, is located at the top right of the screen within the light green rectangle.

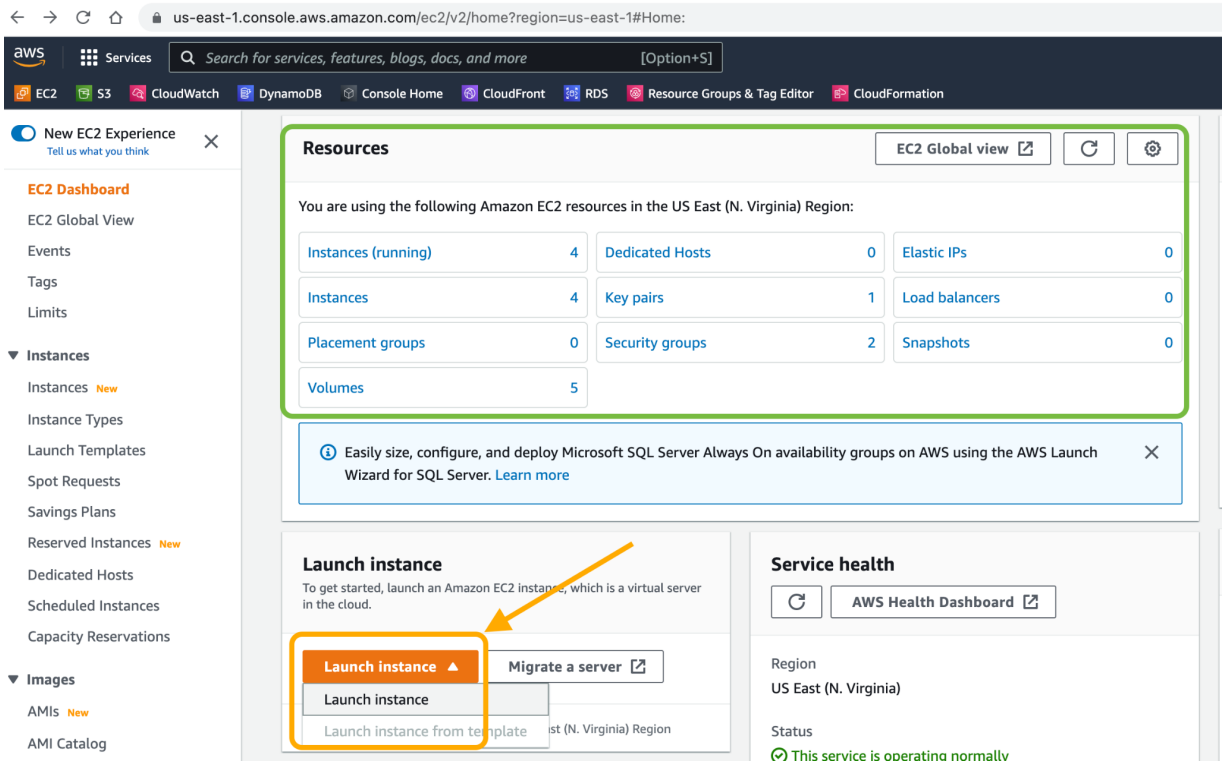
AWS resources can be created by either navigating to the services tab at the top right and clicking on the service sections within the scrollable options on the left, or searching for the service of interest to create:



Or by searching for “EC2” in the search bar located at the top left of the page:



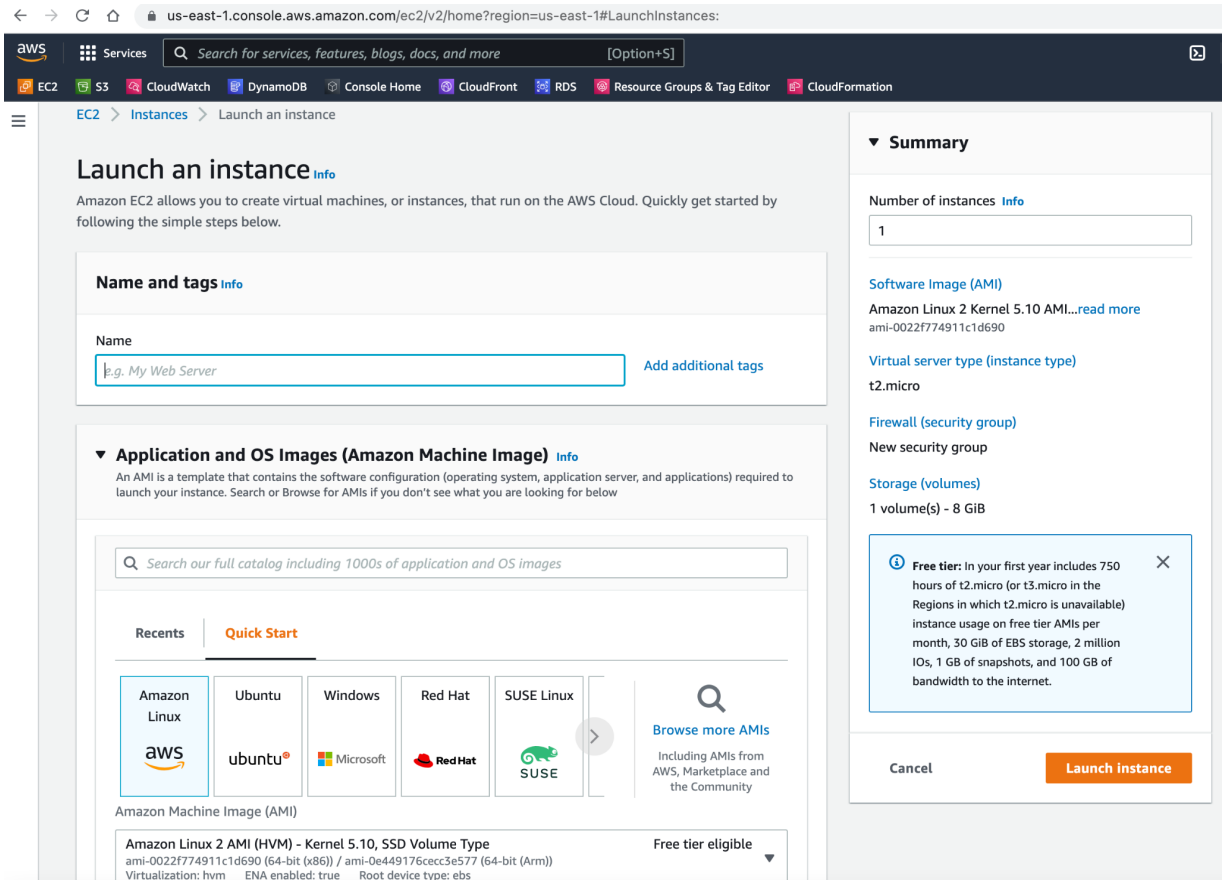
After selecting the "EC2" service (which is used to create virtual servers in the cloud); the user is then able to create an EC2 instance by clicking on the "Launch instance" button as shown below:



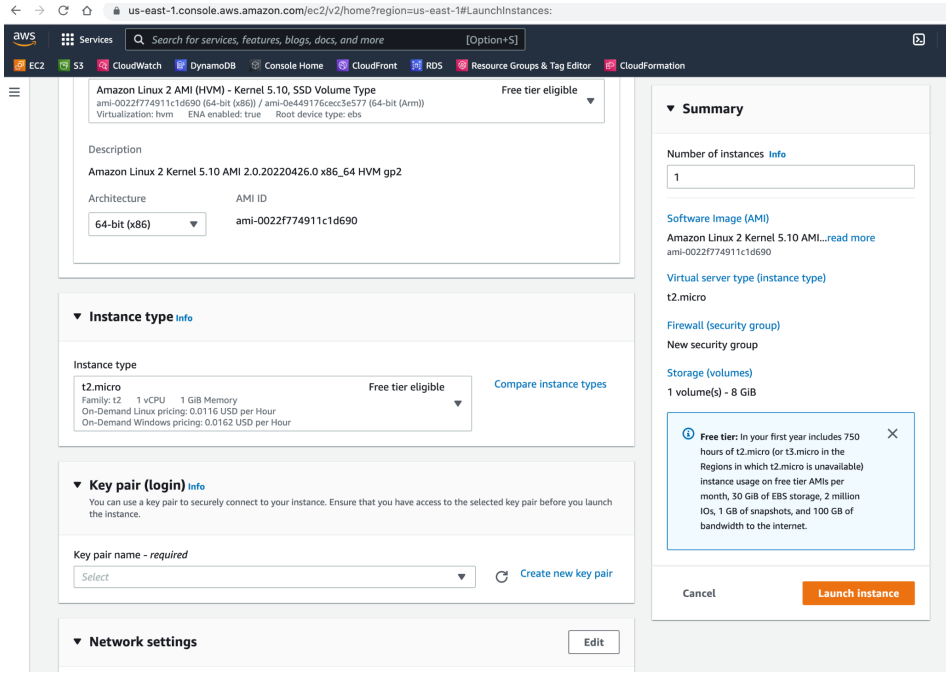
After clicking on the "Launch Instance" button, the user is directed to the form page below to configure the resource they intend to create, for EC2 instances the below is the form page the user must use to configure the instance:

At the top of the form, the user is able to assign a Name /Tag to the instance resource; A resource tag is a label that the user assigns to an AWS resource. Each tag consists of a key and an optional value, both of which the user defines - Tags allow the user to locate resources in the account via the use of the Tag Editor (which is used to list out the available tagged resources within a region - if specified).

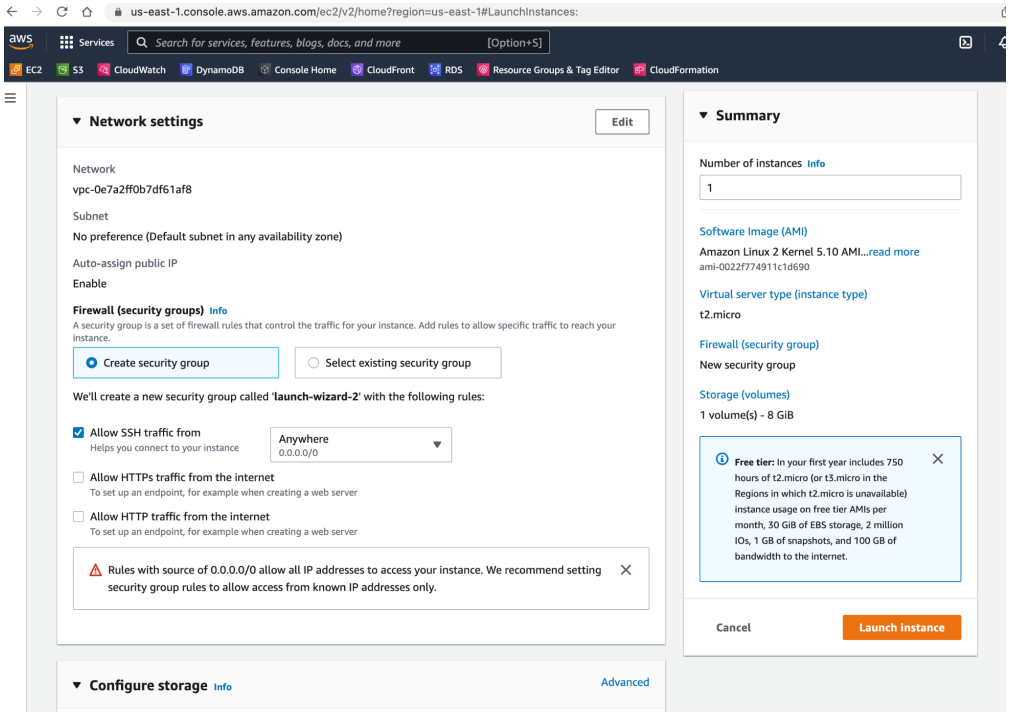
The user is also able to select an application and Operating system image, and on the right of the screen, the user can input the number of instances they intend to create:



The user has the ability to configure instance types and key pairs:



And also can configure Network settings for the EC2 instance being created:



And set other configurations such as the below:

Configure storage Info

Advanced

1x 8 GiB gp2 Root volume

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage

Add new volume

0 x File systems Edit

Instance type Info

Instance type

t2.micro Free tier eligible

Family: t2 1 vCPU 1 GiB Memory

On-Demand Linux pricing: 0.0116 USD per Hour

On-Demand Windows pricing: 0.0162 USD per Hour

Compare instance types

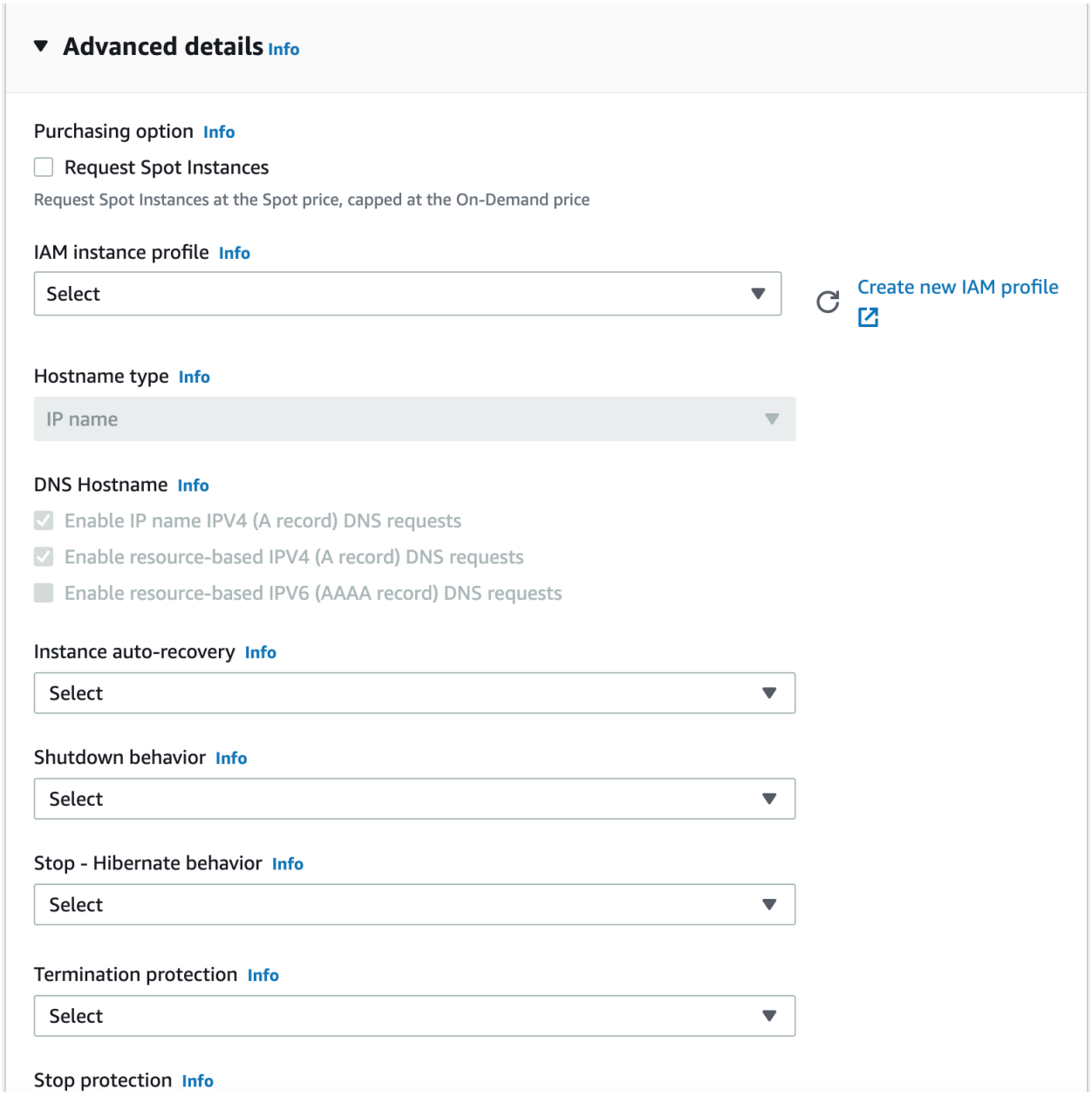
Key pair (login) Info

You can use a key pair to securely connect to your instance. Ensure that you have access to the selected key pair before you launch the instance.

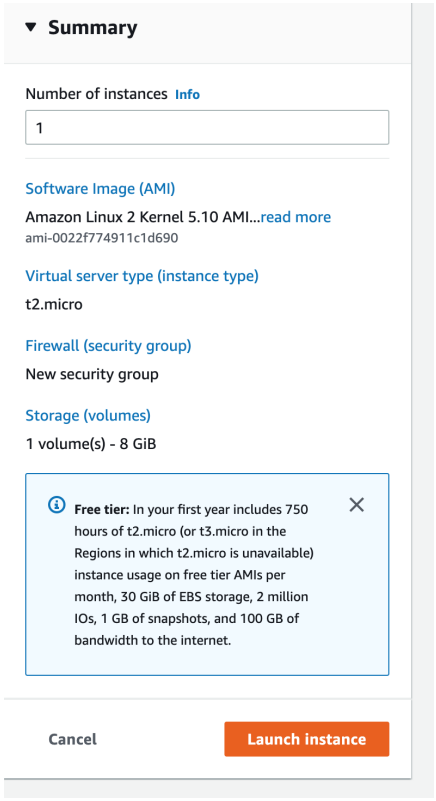
Key pair name - *required*

Select

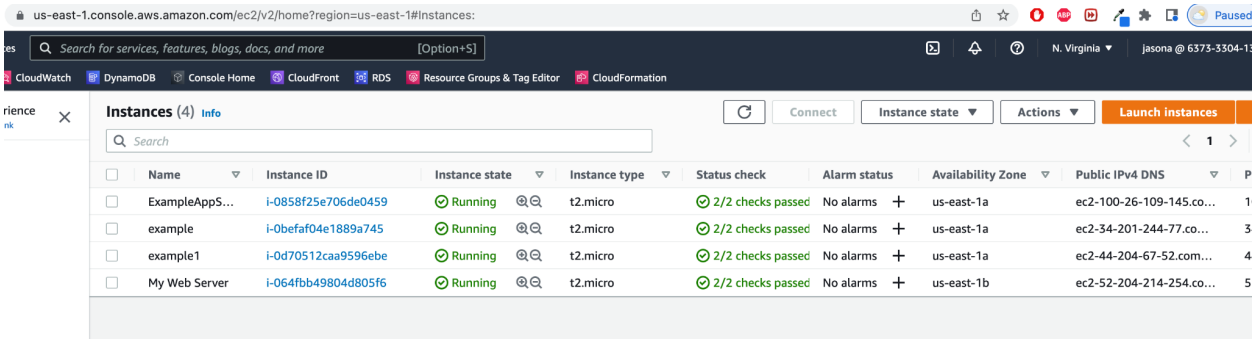
Create new key pair



When done with the settings; the user can click the “Launch Instance” button to create the resource:

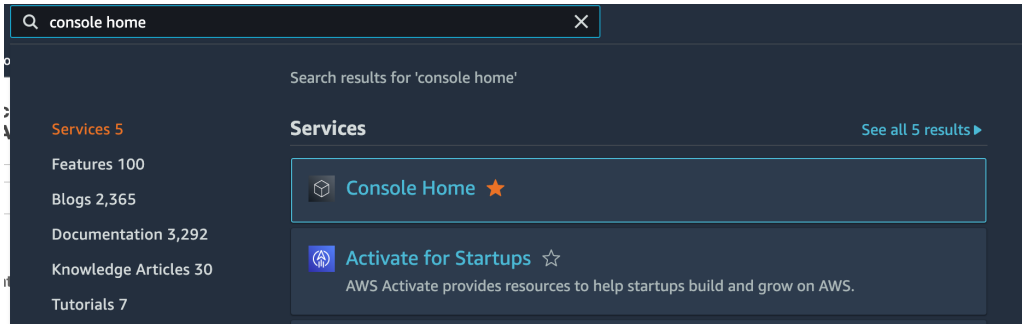


The new instance will then appear with its status in the list on the EC2 instance page, along with the other previously created instances:

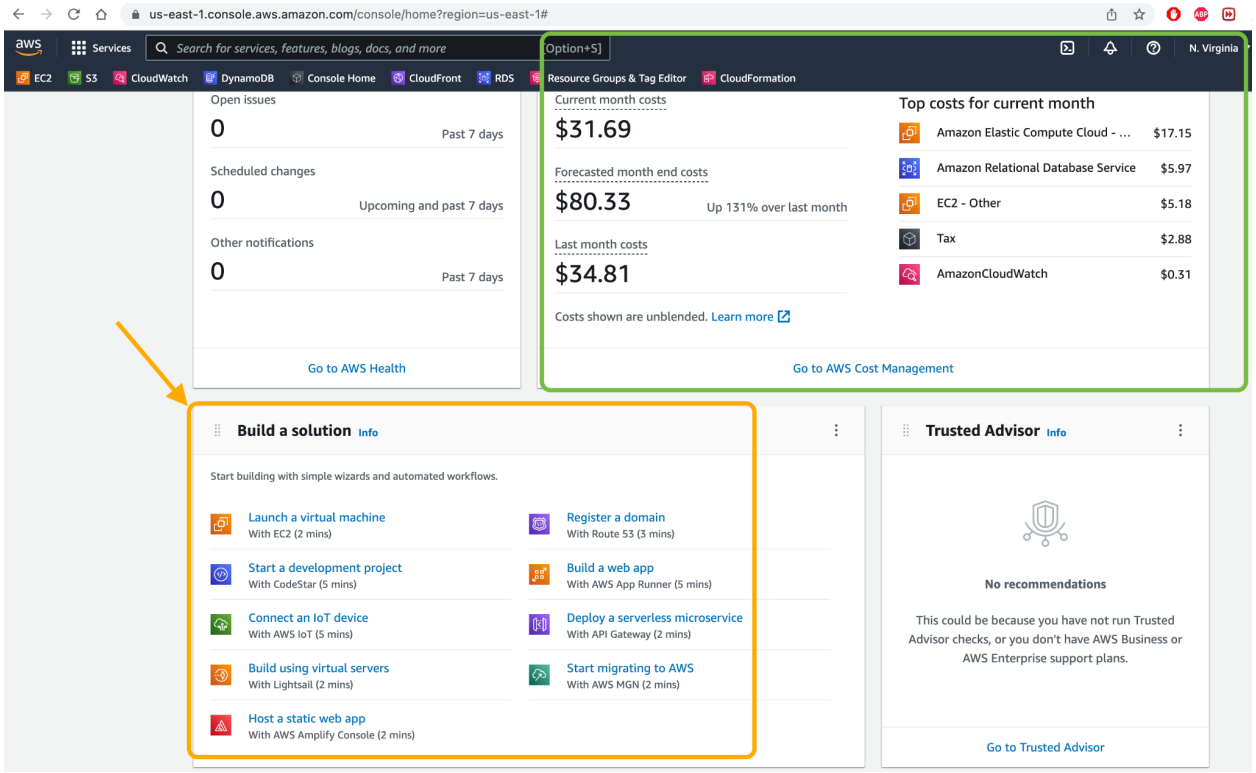


The user can also navigate to the AWS Management console where there is a dedicated section with links to build some common AWS resources (including the estimated time it would take to build on average).

This is accomplished by simply searching for “Console home” at the top left search bar of the page



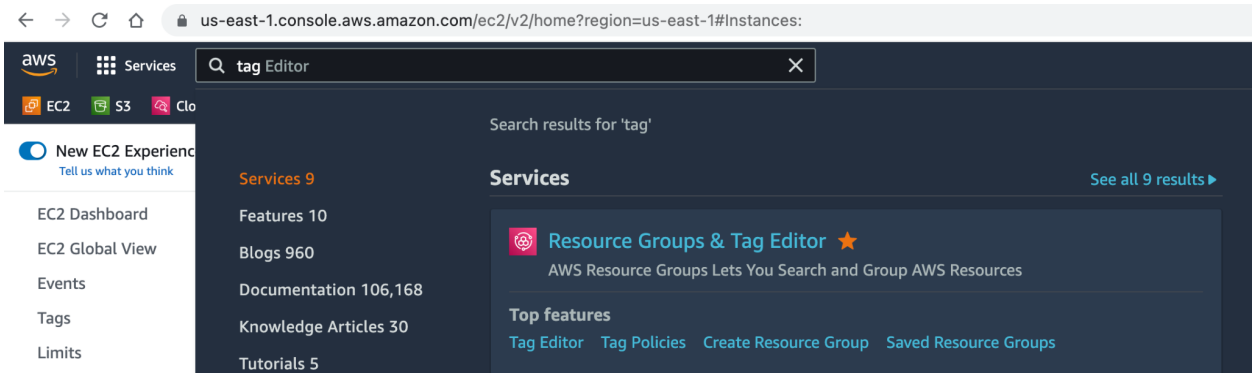
Then clicking on "Console Home" in the services results, to navigate to the "AWS Management Console". On the "AWS Management Console" hope page, the user can scroll down to utilize the provided links to begin the creation of some common AWS resources located at the bottom left of the screen in the orange rectangle highlight:



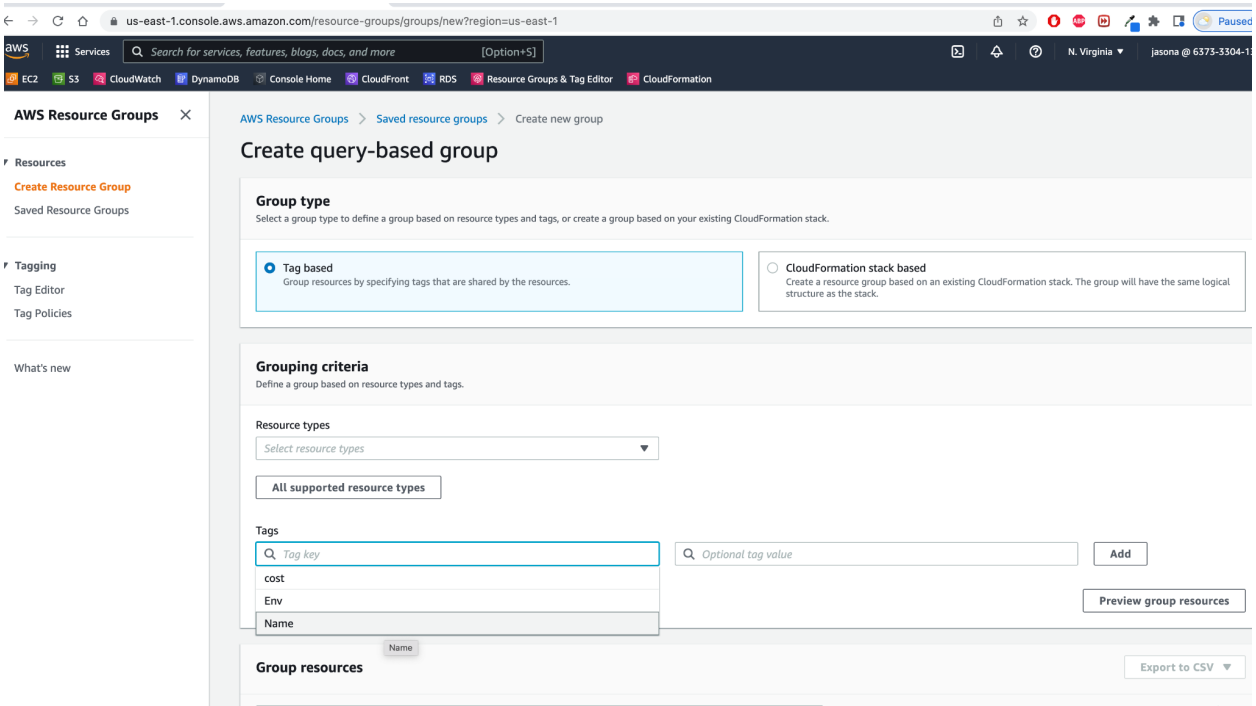
The user should note that the costs for the current month's use of AWS resources and services are displayed at the right of the page within the green rectangle highlight.

The user can utilise the tag manager to list out created resources which has been tagged and add it to a group of resources, which they can monitor:

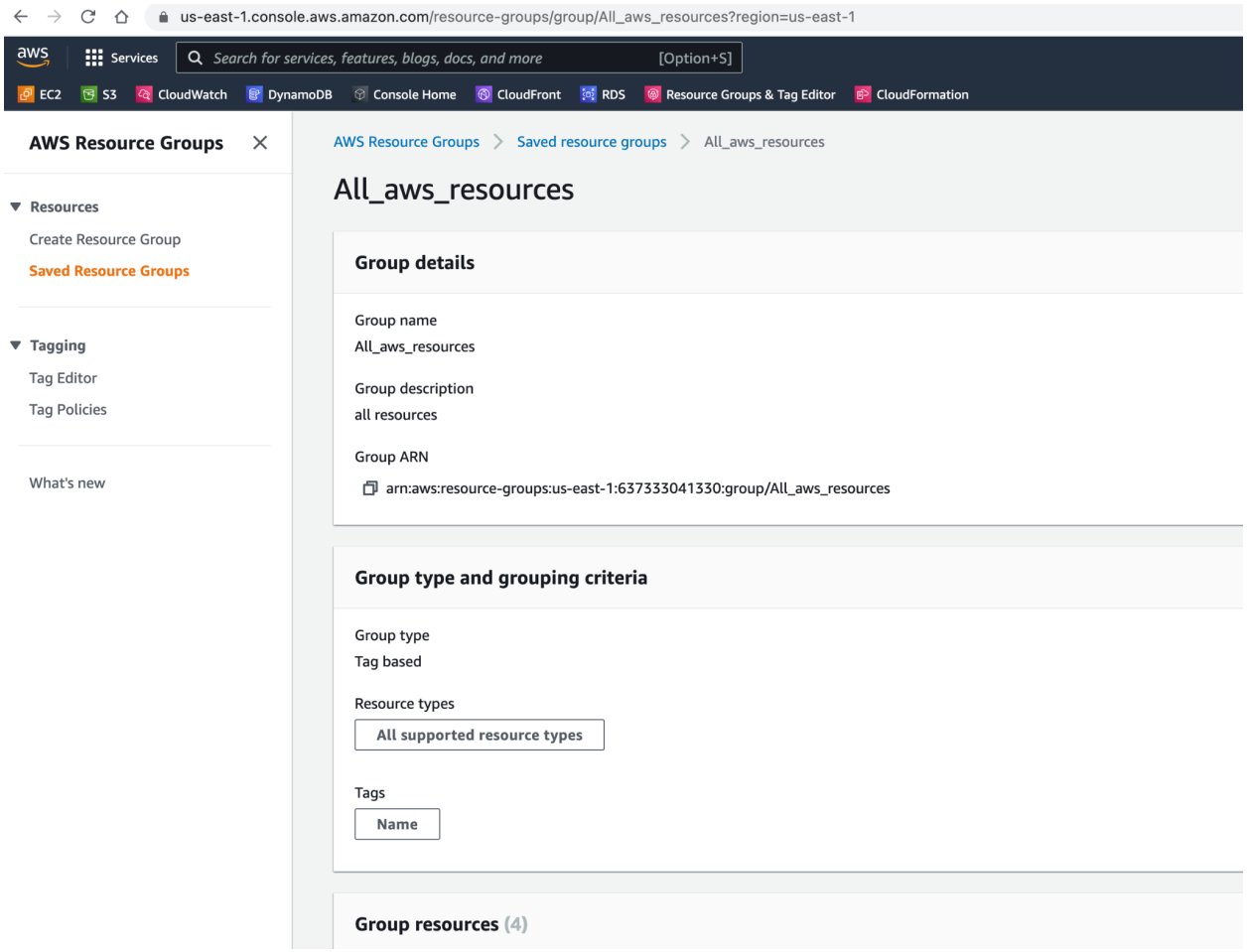
Search for the tag manager in the top bar:

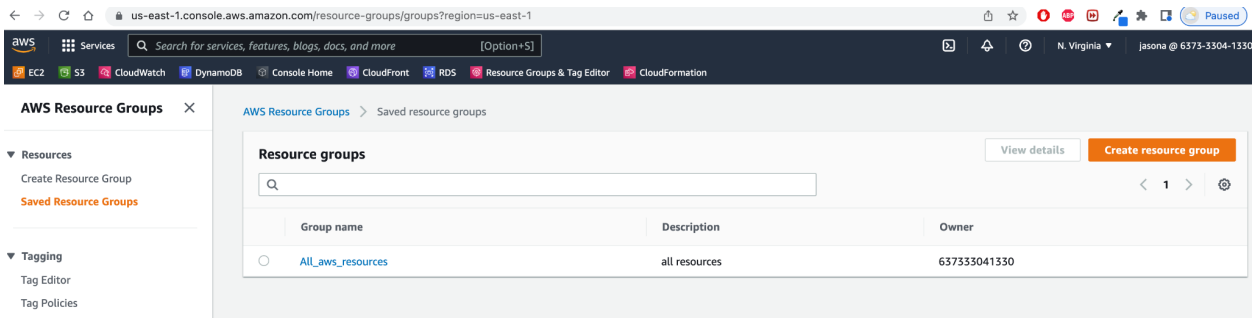


When the user clicks into this option, they are able to create groups of tagged resources:

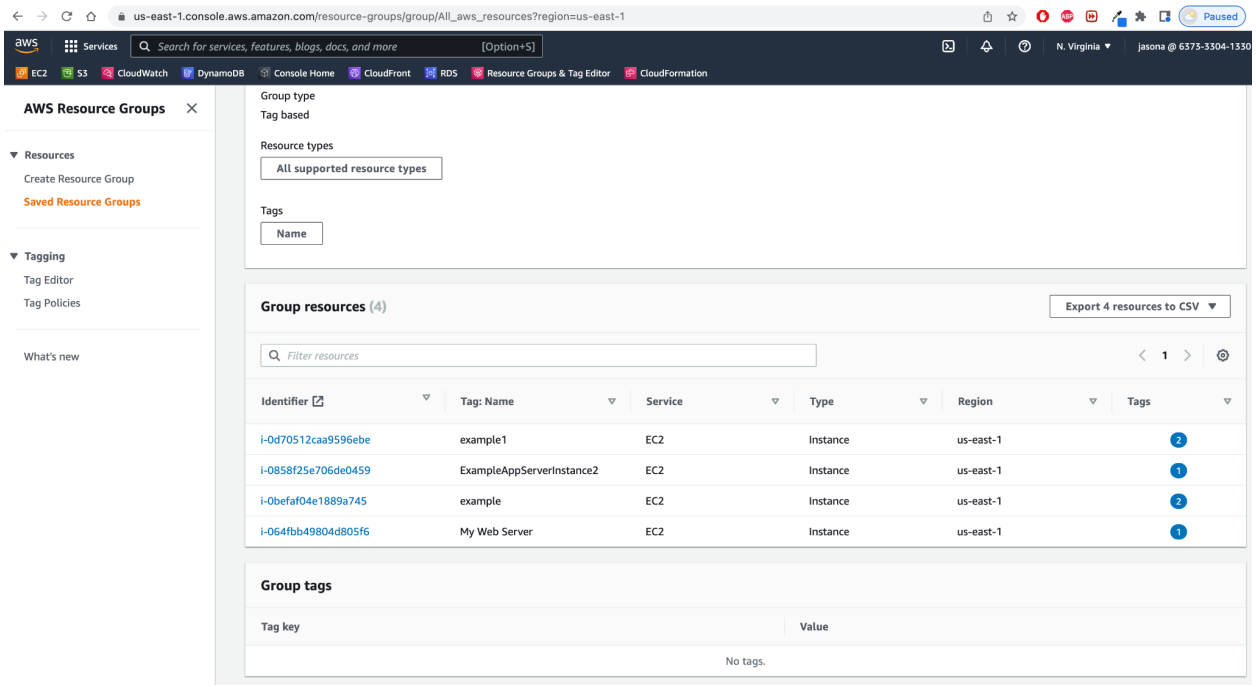


After filling in the fields and saving the group, the users will now have a saved group list of tagged resources:


us-east-1.console.aws.amazon.com/resource-groups/group/All_aws_resources?region=us-east-1
Services
Search for services, features, blogs, docs, and more
[Option+S]
EC2
S3
CloudWatch
DynamoDB
Console Home
CloudFront
RDS
Resource Groups & Tag Editor
CloudFormation
AWS Resource Groups
Resources
Create Resource Group
Saved Resource Groups
Tagging
Tag Editor
Tag Policies
What's new
AWS Resource Groups > Saved resource groups > All_aws_resources
All_aws_resources
Group details
Group name
All_aws_resources
Group description
all resources
Group ARN
arn:aws:resource-groups:us-east-1:637333041330:group/All_aws_resources
Group type and grouping criteria
Group type
Tag based
Resource types
All supported resource types
Tags
Name
Group resources (4)


us-east-1.console.aws.amazon.com/resource-groups/groups?region=us-east-1
Paused
Services
Search for services, features, blogs, docs, and more
[Option+S]
N. Virginia
jasona @ 6373-3304-1330
EC2
S3
CloudWatch
DynamoDB
Console Home
CloudFront
RDS
Resource Groups & Tag Editor
CloudFormation
AWS Resource Groups
Resources
Create Resource Group
Saved Resource Groups
Tagging
Tag Editor
Tag Policies
AWS Resource Groups > Saved resource groups
Resource groups
View details
Create resource group
1
Group name
Description
Owner
All_aws_resources
all resources
637333041330

which they can click into and monitor on the aws web app while cross-checking with the aws cli crawler:



3.2 Configure AWS Environment in CLI

The configuration of the AWS environment in the Command Line Interface is quite simple. It involves the use of amazon's "aws cli", which needs to be installed on the user's machine and configured with the user's account details and credentials.

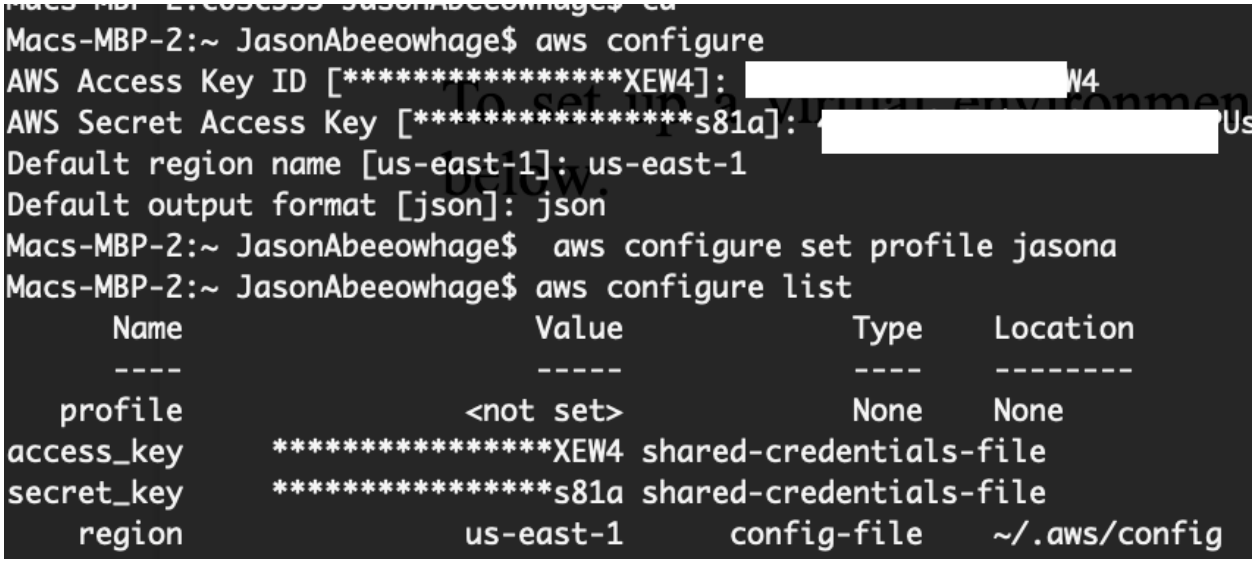
After the user accesses the directory of choice with

cd /directoryOfChoice

Input the following command and hit the return/enter button in the command line interface:

aws configure

You will be prompted and enter the corresponding credentials supplied from the AWS account you are using :



After entering the correct credentials for AWS Access Key ID, AWS Secret Access Key, Default region name, and output format; the user's aws cli environment is configured.

The user is also able to list out the aws accounts that have been configured in the environment by entering :

aws configure list

As shown in the screenshot above.

3.3 Crawling AWS for resources with BeatDrift-CLI tool

Crawling AWS for specific or all resources can be done with the use of the “BeatDrift-CLI” tool.

The user simply needs to download the source code zip file and then unzip this file,

Then go into the “beatdrift-cli-master_copy” directory from the home directory it was downloaded into by typing in:

cd ./beatdrift-cli-master_copy”

within the directory of their choice (This can also be a virtual environment). The environment in which the python script is run, needs to have the required installed system software environment tools and versions as specified in the ‘setup.py’ file:

Programming language version: Python 2’, Python 2.7’, Python 3’, or Python 3.6’

Boto3 version: 1.16.57

App_json_file_cache version: 0.2.2

To install the latest version of a package enter:

pip install 'PackageName'

To install a specific version, type the package name followed by the required version enter:

pip install 'PackageName ==1.4'

The user should note that there are compatibility issues with the latest version of boto3, due to several changes made that would affect the ability to run BeatDrift-cli as intended; until updates are made, users are advised to stick to the requirements stated.

pip CLI python packaging - version (or equivalent python package manager)

Can be used to install the specific versions of these packages, if they are still available; the user can also search the web to install the specific versions of the required libraries, to ensure the script runs as intended.

Tip:

Care must be taken to ensure the user's machine is set to the current date and time to avoid denial of access to list the available resources, due to invalid AWS credentials authentication caused by a non- automatically set time and date on the machine/computer; an example of this authentication error is shown below:

Once the environment is set up correctly; cd into the "BeatDrift-CLI" folder via the command line to execute the python scripts;

Please enter :

chmod +x __main__.py

Then :

cp __main__.py ~/bin/beatdrift-cli-master

The user would then be able to run the program using commands such as:

beatdrift-cli query --region us-east-1 --service ec2 --directory ./listdataCheck2/

However, since full integration was not acheived in the project, if you are having problems with the execution or have prompts about missing modules, you can utilize the 'aws-list-all' package incorporated by BeatDrift-cli (with credit given in the notice to the author) to achieve the same results of crawling AWS resources and storing them as json files within the designated directory/folder within the project's folder:

AWS resources can be crawled by installing and using the "aws-list-all" package

If you are using a virtual environment simply enter the below code:

mkvirtualenv -p $(which python3) aws

pip install aws-list-all

After installation, The user is able to query specific AWS resources by typing in the following.

aws-list-all query --region us-east-1 --service ec2 --directory ./listdataCheck2/demo

This command line argument crawls the provided aws account for all resources contained within the specified parameters:

(1) The main command is “query”

(2) Parameter 1 - region : us-east-1: this parameter

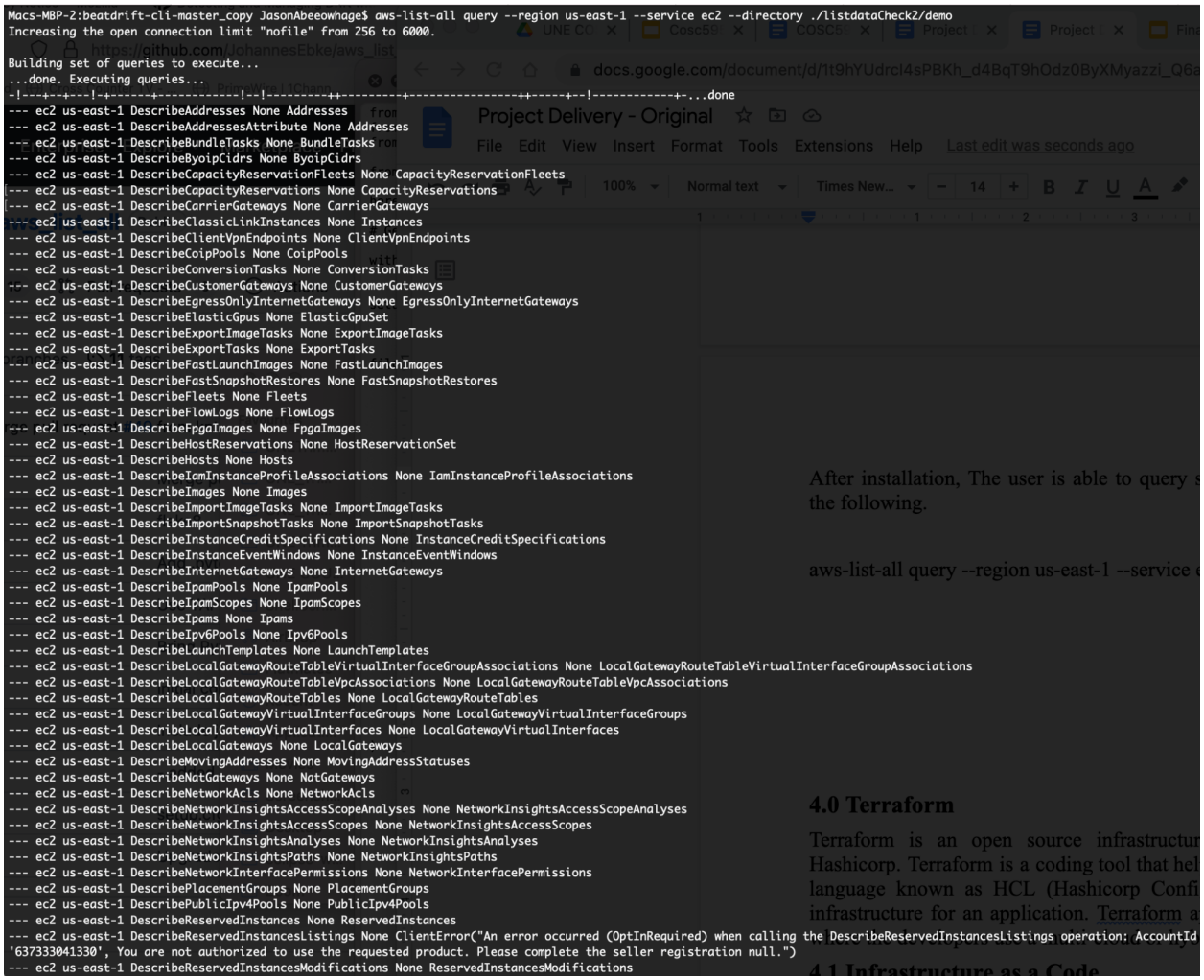
(3) Parameter 2 - service : ec2

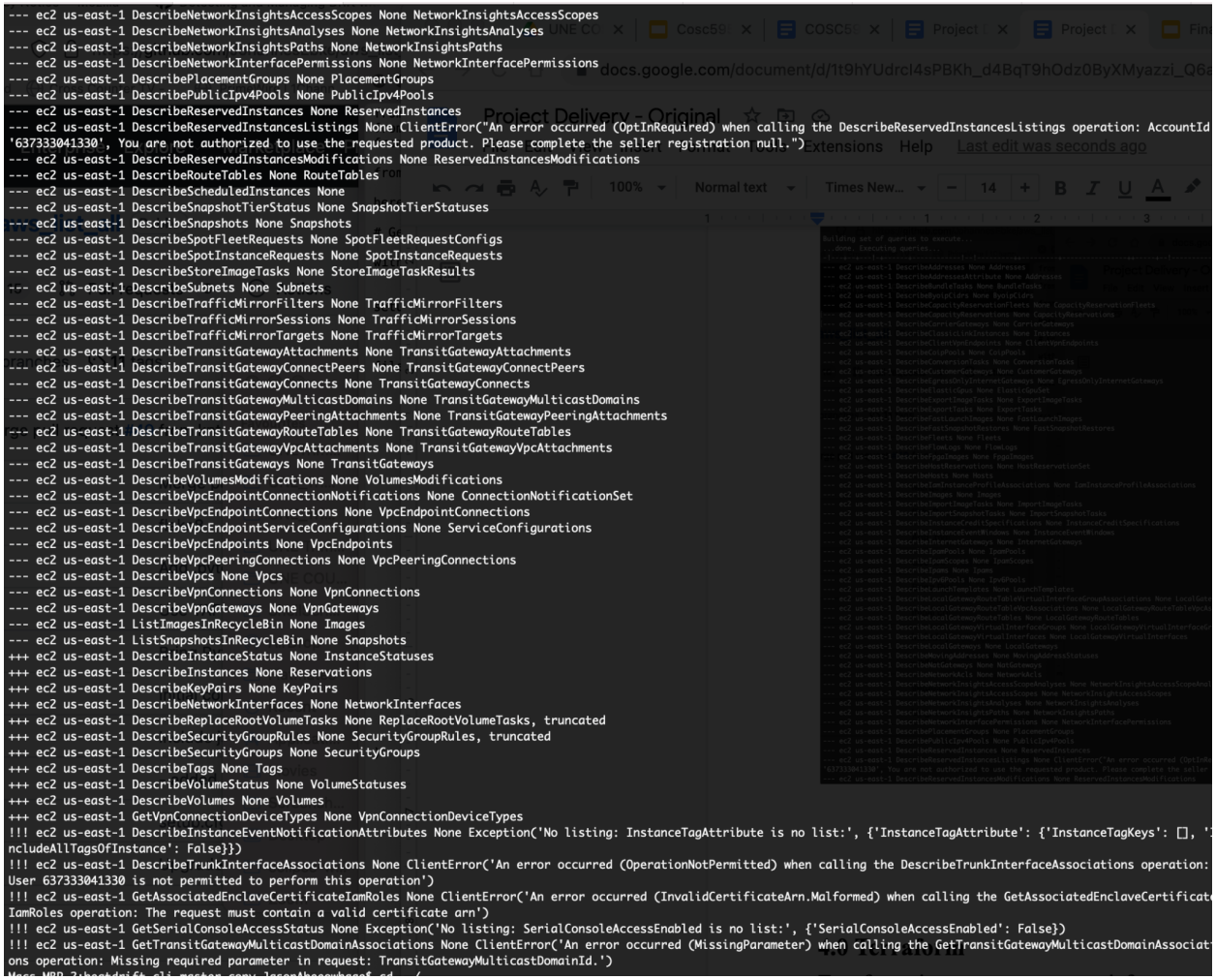
- This parameter tells the script to only crawl ec2 resources.

(4) Parameter 3 - Directory: ./listdataCheck2/demo

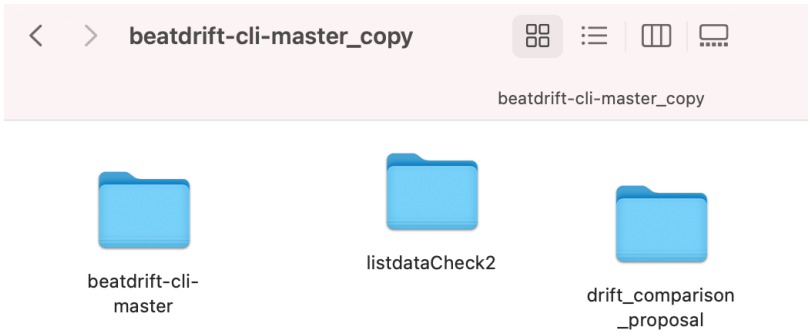
- This parameter tells the script where to store the json files created to store the information about the AWS resources it has detected and listed.

More parameters like “-verbose” can be added and repeated for emphasis on expected outcome.


Macs-MBP-2:beatdrift-cli-master_copy JasonAbeeowhage$ aws-list-all query --region us-east-1 --service ec2 --directory ./listdataCheck2/demo
Increasing the open connection limit "nofile" from 256 to 6000.
Building set of queries to execute...
...done. Executing queries...
--- ec2 us-east-1 DescribeAddresses None Addresses
--- ec2 us-east-1 DescribeAddressesAttribute None Addresses
--- ec2 us-east-1 DescribeBundleTasks None BundleTasks
--- ec2 us-east-1 DescribeByoipCidrs None ByoipCidrs
--- ec2 us-east-1 DescribeCapacityReservationFleets None CapacityReservationFleets
--- ec2 us-east-1 DescribeCapacityReservations None CapacityReservations
--- ec2 us-east-1 DescribeCarrierGateways None CarrierGateways
--- ec2 us-east-1 DescribeClassicLinkInstances None Instances
--- ec2 us-east-1 DescribeClientVpnEndpoints None ClientVpnEndpoints
--- ec2 us-east-1 DescribeCoipPools None CoipPools
--- ec2 us-east-1 DescribeConversionTasks None ConversionTasks
--- ec2 us-east-1 DescribeCustomerGateways None CustomerGateways
--- ec2 us-east-1 DescribeEgressOnlyInternetGateways None EgressOnlyInternetGateways
--- ec2 us-east-1 DescribeElasticGpus None ElasticGpuSet
--- ec2 us-east-1 DescribeExportImageTasks None ExportImageTasks
--- ec2 us-east-1 DescribeExportTasks None ExportTasks
--- ec2 us-east-1 DescribeFastLaunchImages None FastLaunchImages
--- ec2 us-east-1 DescribeFastSnapshotRestores None FastSnapshotRestores
--- ec2 us-east-1 DescribeFleets None Fleets
--- ec2 us-east-1 DescribeFlowLogs None FlowLogs
--- ec2 us-east-1 DescribeFpgaImages None FpgaImages
--- ec2 us-east-1 DescribeHostReservations None HostReservationSet
--- ec2 us-east-1 DescribeHosts None Hosts
--- ec2 us-east-1 DescribeIamInstanceProfileAssociations None IamInstanceProfileAssociations
--- ec2 us-east-1 DescribeImages None Images
--- ec2 us-east-1 DescribeImportImageTasks None ImportImageTasks
--- ec2 us-east-1 DescribeImportSnapshotTasks None ImportSnapshotTasks
--- ec2 us-east-1 DescribeInstanceCreditSpecifications None InstanceCreditSpecifications
--- ec2 us-east-1 DescribeInstanceEventWindows None InstanceEventWindows
--- ec2 us-east-1 DescribeInternetGateways None InternetGateways
--- ec2 us-east-1 DescribeIpamPools None IpamPools
--- ec2 us-east-1 DescribeIpamScopes None IpamScopes
--- ec2 us-east-1 DescribeIpams None Ipams
--- ec2 us-east-1 DescribeIpv6Pools None Ipv6Pools
--- ec2 us-east-1 DescribeLaunchTemplates None LaunchTemplates
--- ec2 us-east-1 DescribeLocalGatewayRouteTableVirtualInterfaceGroupAssociations None LocalGatewayRouteTableVirtualInterfaceGroupAssociations
--- ec2 us-east-1 DescribeLocalGatewayRouteTableVpcAssociations None LocalGatewayRouteTableVpcAssociations
--- ec2 us-east-1 DescribeLocalGatewayRouteTables None LocalGatewayRouteTables
--- ec2 us-east-1 DescribeLocalGatewayVirtualInterfaceGroups None LocalGatewayVirtualInterfaceGroups
--- ec2 us-east-1 DescribeLocalGatewayVirtualInterfaces None LocalGatewayVirtualInterfaces
--- ec2 us-east-1 DescribeLocalGateways None LocalGateways
--- ec2 us-east-1 DescribeMovingAddresses None MovingAddressStatuses
--- ec2 us-east-1 DescribeNatGateways None NatGateways
--- ec2 us-east-1 DescribeNetworkAcls None NetworkAcls
--- ec2 us-east-1 DescribeNetworkInsightsAccessScopeAnalyses None NetworkInsightsAccessScopeAnalyses
--- ec2 us-east-1 DescribeNetworkInsightsAccessScopes None NetworkInsightsAccessScopes
--- ec2 us-east-1 DescribeNetworkInsightsAnalyses None NetworkInsightsAnalyses
--- ec2 us-east-1 DescribeNetworkInsightsPaths None NetworkInsightsPaths
--- ec2 us-east-1 DescribeNetworkInterfacePermissions None NetworkInterfacePermissions
--- ec2 us-east-1 DescribePlacementGroups None PlacementGroups
--- ec2 us-east-1 DescribePublicIpv4Pools None PublicIpv4Pools
--- ec2 us-east-1 DescribeReservedInstances None ReservedInstances
--- ec2 us-east-1 DescribeReservedInstancesListings None ClientError("An error occurred (OptInRequired) when calling the DescribeReservedInstancesListings operation: AccountId
'637333041330', You are not authorized to use the requested product. Please complete the seller registration null.")
--- ec2 us-east-1 DescribeReservedInstancesModifications None ReservedInstancesModifications
Project Delivery - Original
File Edit View Insert Format Tools Extensions Help Last edit was seconds ago
After installation, The user is able to query
the following.
aws-list-all query --region us-east-1 --service
4.0 Terraform
Terraform is an open source infrastructur
Hashicorp. Terraform is a coding tool that hel
language known as HCL (Hashicorp Confi
infrastructure for an application. Terraform a

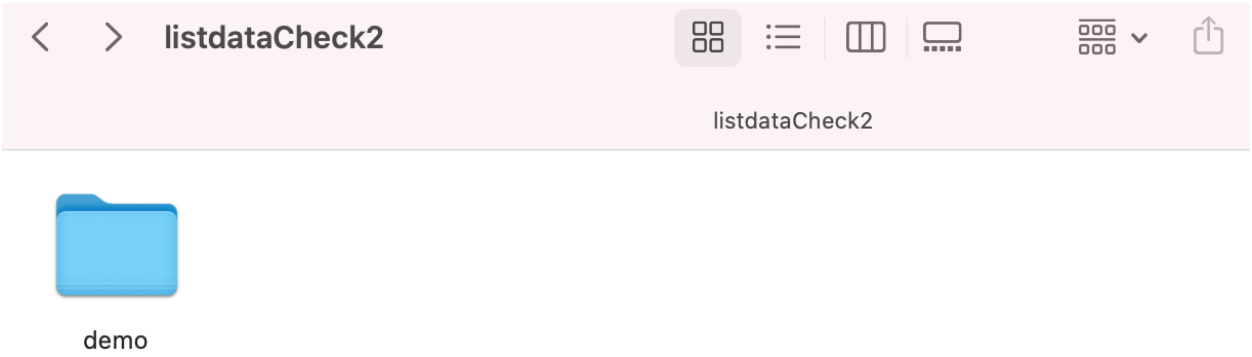


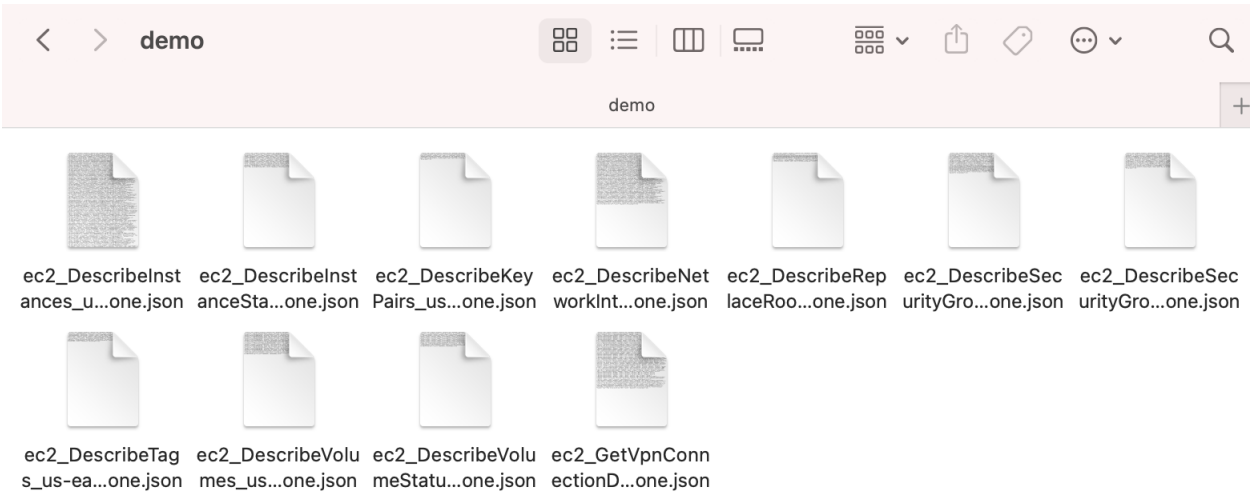
The data is listed with the “+” symbols, indicating resources that where found in the AWS account, the “-” symbol indicates resources not found in the AWS account and the “!!!” symbol indicates a warning.



The json files then appear within the created “listdataCheck2” folder

In the second created folder directory named “demo”:





The contents of the files created are of “JSON” format and contain the details of each crawled AWS resource; including properties such as “ARN” (Amazon Resource Names) which uniquely identify AWS resources. We require an ARN when you need to specify a resource unambiguously across all of AWS, such as in IAM policies, Amazon Relational Database Service (Amazon RDS) tags, and API calls.

When the query is run without a filter on services:

aws-list-all query --region us-east-1 --directory ./listdata2/demo2

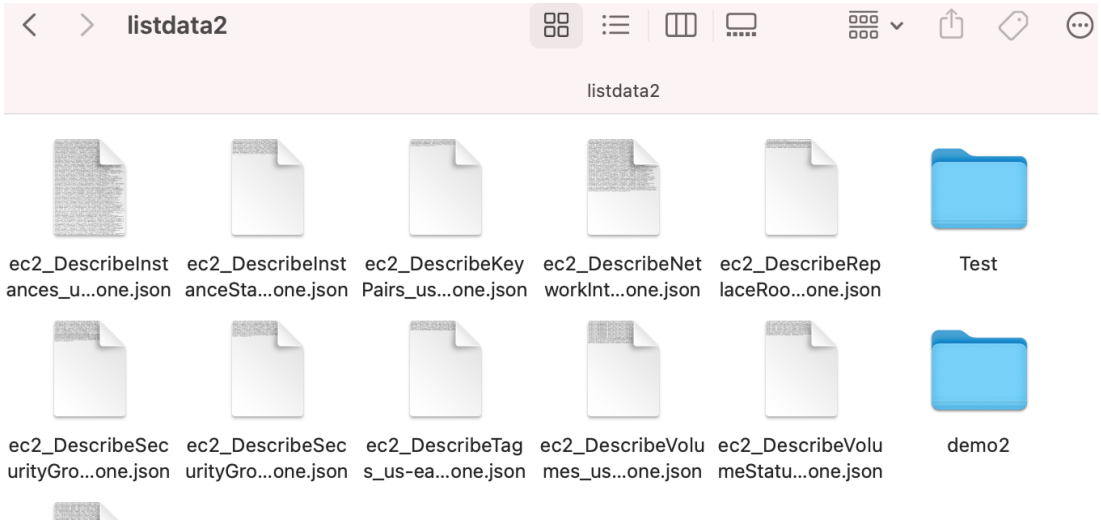
it takes a much longer time to crawl the AWS account for all resources in the specified region the difference in speed is

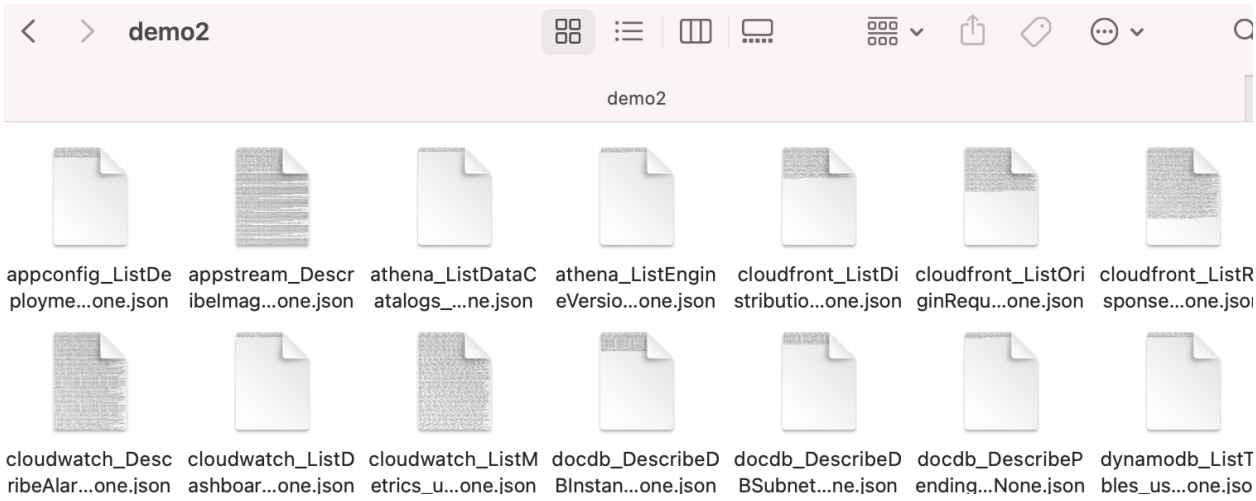
Benchmarked at

~ 4-6 minutes as opposed to the usual 45 seconds to 1 minute that it usually takes to crawl the AWS account for specific services.

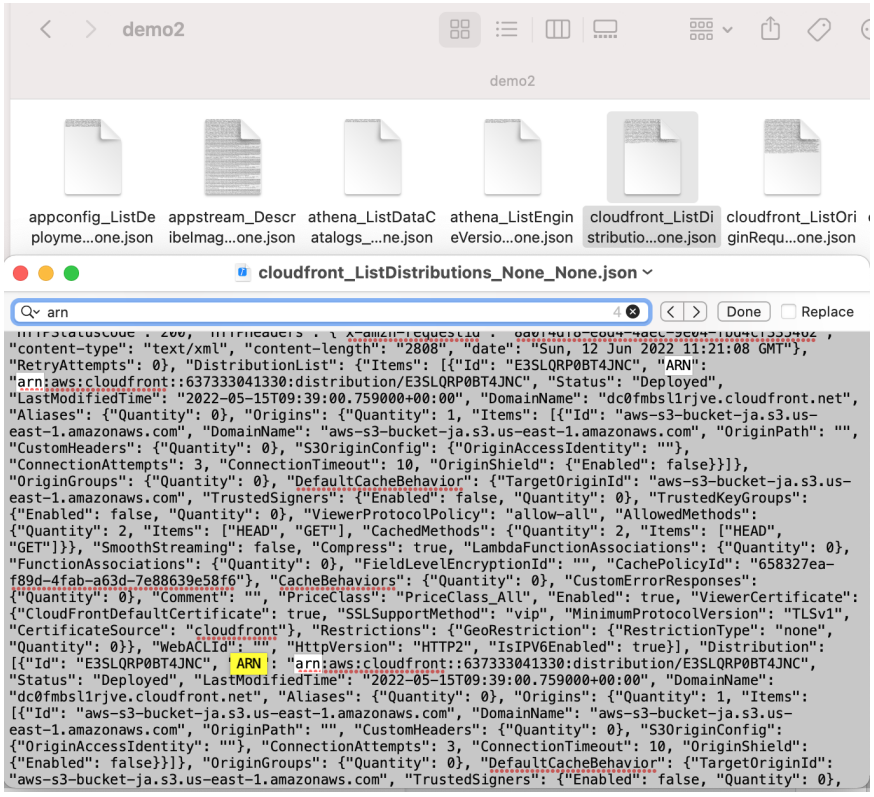
This is due to the scripts utilisation of a list of "service regions" stored in 'service_regions.json' file and "endpoint hosts" stored in the 'endpoints_hosts.json' file for the script to search through, when specific services are not specified in the command line argument.

With the second command , we get more resources listed and stored in json file formats within the demo2 folder located in the "**./listdata2/demo2**" directory:





And within these json files are the “ARNs” of the crawled AWS resources which can be used to compare against the ARN’s within terraform’s tfstate file to detect infrastructural drift.



Project github link: https://github.com/Jase-The-Ace/beatdrift-cli