Operationalizing an AWS ML Project

- Write up by Prafull Parmar

Note: *For each of the required steps all the snapshot are also present in the* ***snapshots*** *folder in the repository.*

1. Initial Setup

I have chosen the "**ml.t2.medium**" instance type for Notebook instance (Figure 1 - Sagemaker Notebook Instance). There are multiple reasons for selecting this instance type for my notebook.

- Firstly for completing the execution of this particular project's jupyter notebooks we **do not need** a very computationally **powerful CPU** and **high RAM.**
- We will need to keep this **notebook** instance in "**inService**" **status** for a **long time** while we are working on the project
- In order to **avoid high costs** we should **select a notebook** that is **low in per hour cost** and also **offers** reasonably **good CPU and RAM**.
- So looking at the instance type and their pricing : https://aws.amazon.com/sagemaker/pricing/ , we have two choices "ml.t2.medium" and "ml.t3.medium". Both have 2 vCPU and 4 GB Memory, and as per doc "ml.t3.medium" has a slightly higher cost as it has a fast boot time.
- Now given that we do have a critical requirement for a fast boot time, so we can go ahead with the "**ml.t2.medium**" as it offers the same computational power and is lower in per hour cost.

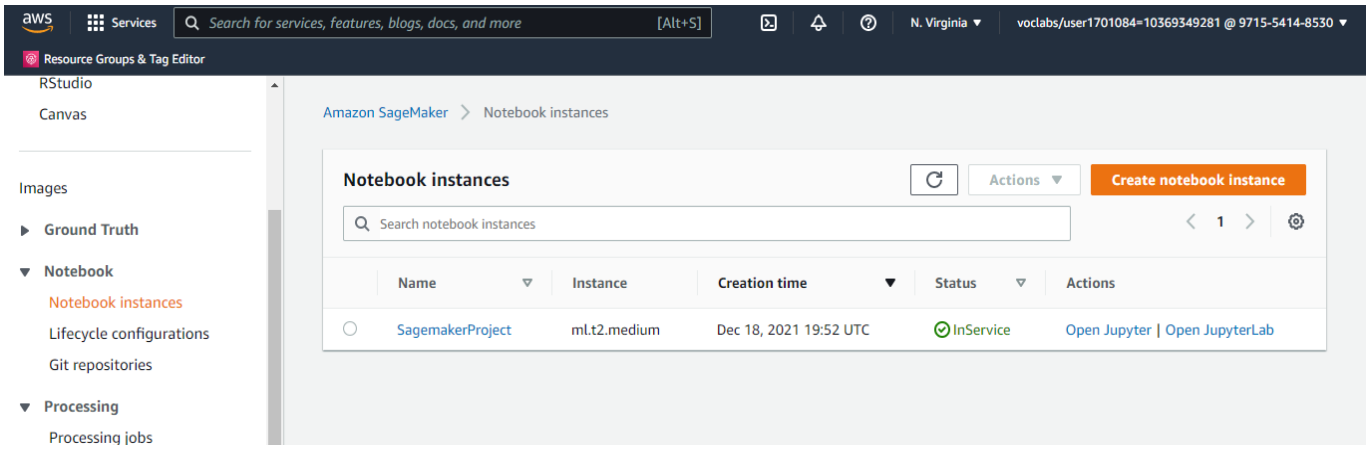


Figure 1 - Sagemaker Notebook Instance

The dog breed dataset was uploaded to a newly created S3 bucket, successfully.

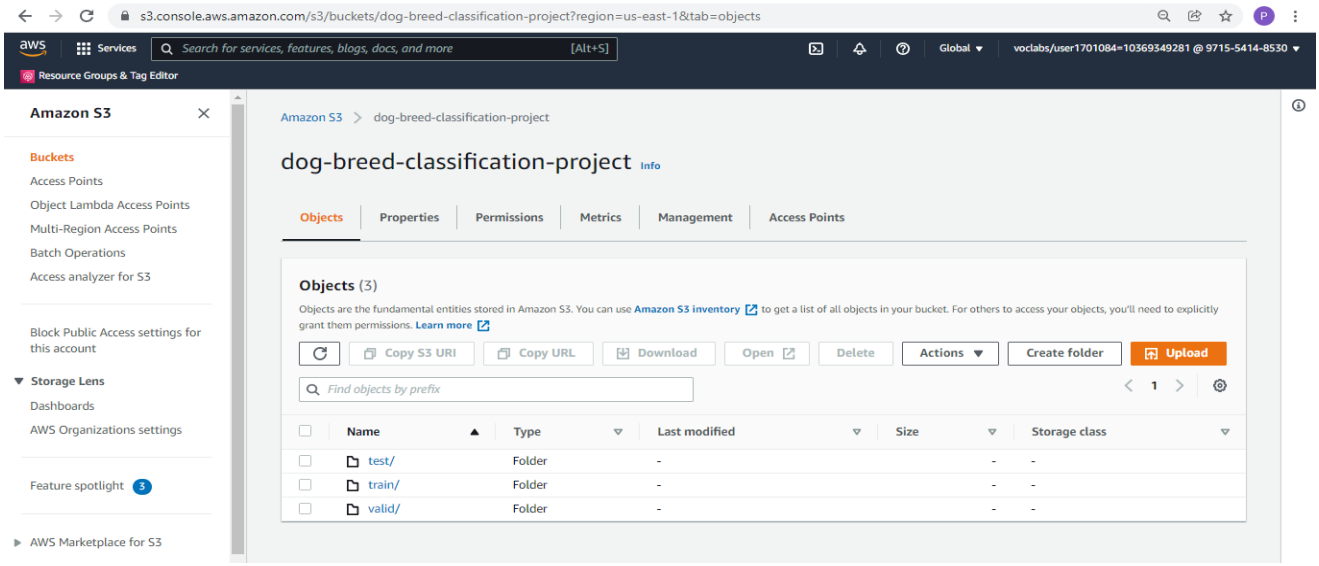


Figure 2 S3 Bucket snapshot

2. Sagemaker Training and Deployment

For hyperparameter tuning I used the same "ml.m5.xlarge" instance_type , however since it took too much time for the tuning, while the training process I tried to increase the processing power a bit by using the "ml.m5.2xlarge" for the single instance and multi instance training purposes.
Hyperparameter tuning job (max_jobs = 2, max_parallel_jobs = 2) completed successfully:

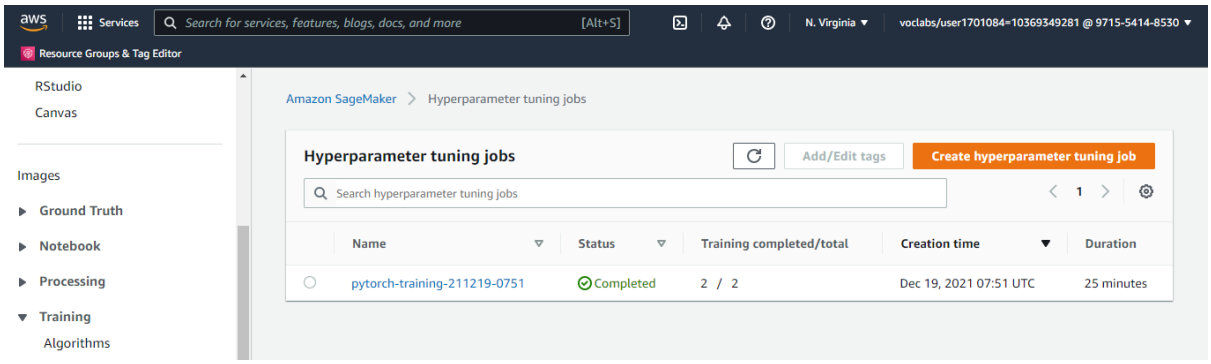


Figure 3 Hyperparameter Tuning Job

However, upon training the model with the best parameters from above tuning, the model gave a 0 test accuracy! So increased the max_jobs = 6, max_parallel_jobs = 3 and also changed its instance_type = "ml.m5.xlarge" to speed up the computations a bit.

Reran the hyperparmeter jobs and it executed successfully:

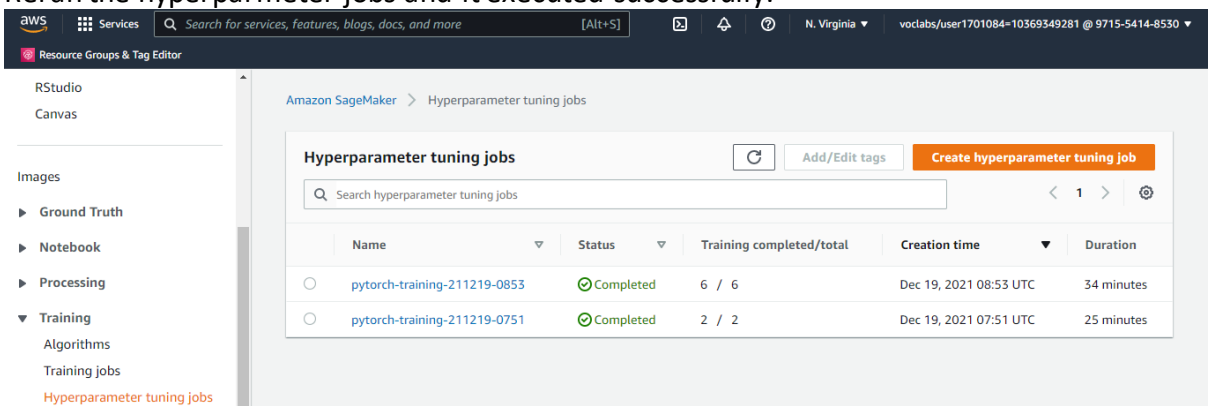


Figure 4 Hyperparameter jobs summary

Post which triggered the **single instance** and **multi instance training jobs**. Jobs completed successfully. For snapshot of the training jobs refer the images folder in the repository.

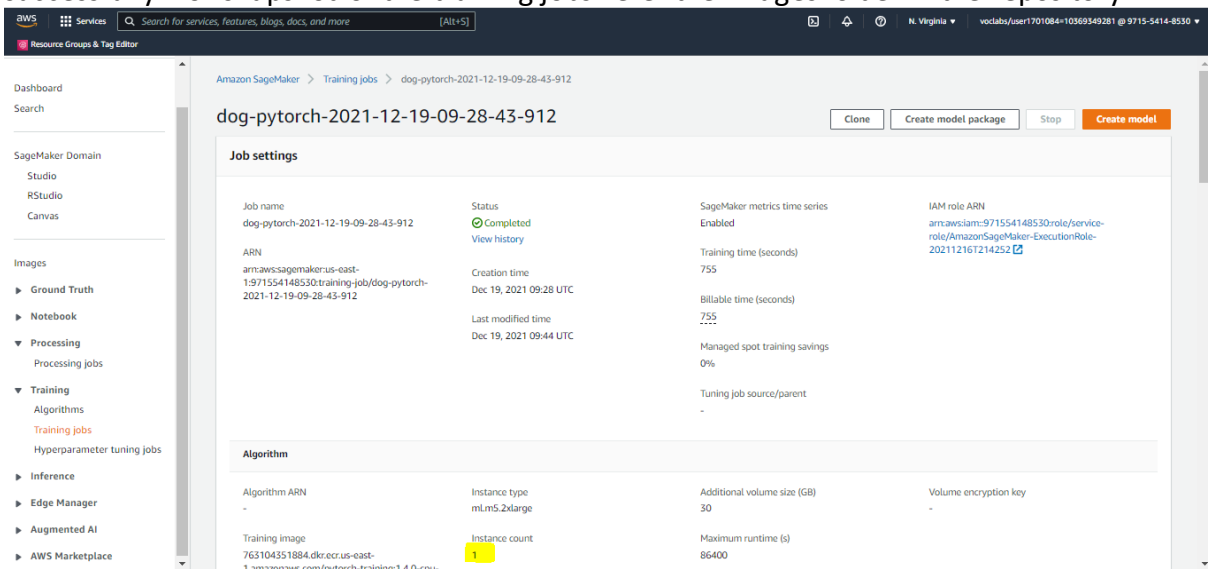


Figure 5 Single Instance Training Job

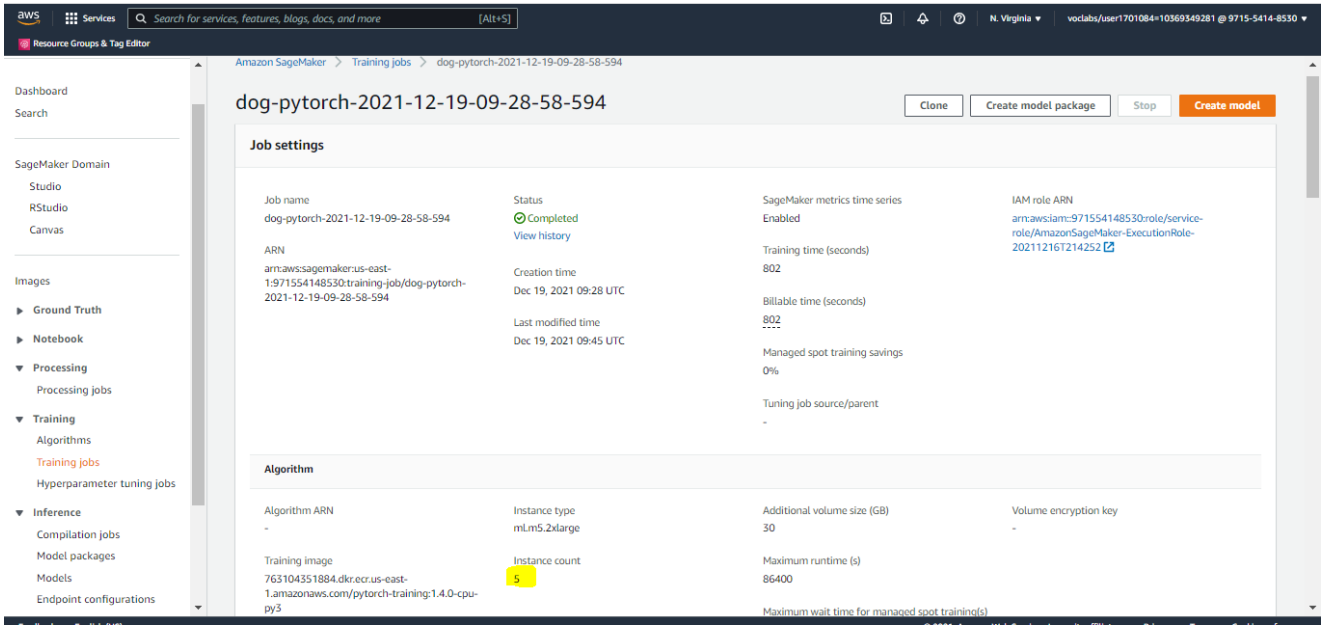


Figure 6 Multi Instance Training Job

Deployed Endpoints:

- Single instance deployed endpoint: "pytorch-inference-2021-12-19-09-46-12-148"
- Multi instance deployed endpoint: "pytorch-inference-2021-12-19-09-53-14-575"

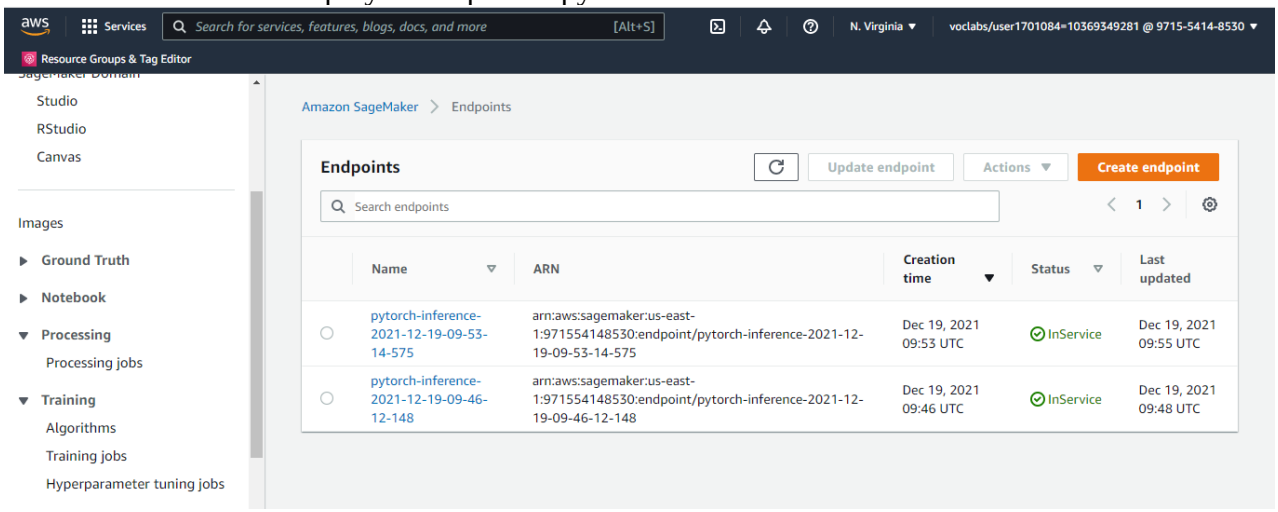


Figure 7 Sagemaker Endpoints

3. EC2 Training

- We have utilized the **t2.xlarge** instance and the **Deep Learning AMI (Amazon Linux 2) Version 55.0**. This seems like a reasonable balance of performance and affordability.
- As per the documentation, T2 instances can sustain high CPU performance for as long as a workload needs it.
- For most general-purpose workloads, T2 instances will provide ample performance without any additional charges.
- Similarly, because we don't know the duration for which we might need to keep this EC2 instance running for training, it's better to go with a medium size instance so we don't have to pay for a large instance while we're doing setup, debugging and other tasks.

Difference between **ec2train1.py (EC2 script)** and **train_and_deploy solution.ipynb + hpo.py (SageMaker scripts)**

- There is no logic for calling any Estimator or Tuner functions in the EC2 script. The code in the EC2 script is responsible for saving the model to the local path. While in the sagemaker scripts this was handled internally by sagemaker where the model data was stored to a S3 location.
- In the EC2 training script, all the variables like hyperparameters and output locations, etc are already mentioned in the script itself and so there is no need for **argparse.** Meaning while running the EC2 script we do not need to mention any arguments.
- In the EC2 script the training happens on the same server on which the script is invoked/executed, however in the sagemaker scripts the training job that is invoked, it runs on a separate container than the one on which the sagemaker notebook is running.
- Another difference is that ec2train1.py lacks the main function
- For the EC2 Training, given that the training data and model, all are stored on the EC2 instance host itself it would be difficult to deploy the saved model to an endpoint in sagemaker. If we wish to do that then we might need to manually upload the model first to sagemaker and then use that to deploy an endpoint. This is not the case in models trained via the sagemaker notebook instances, as the model can be easily deployed to an endpoint.

For snapshot refer below:

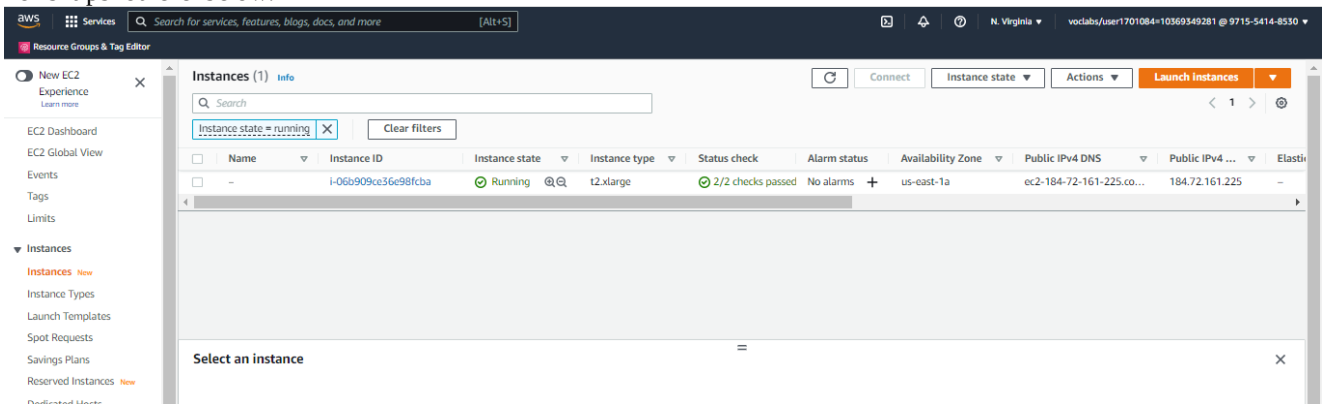


Figure 8 EC2 Instance snapshot

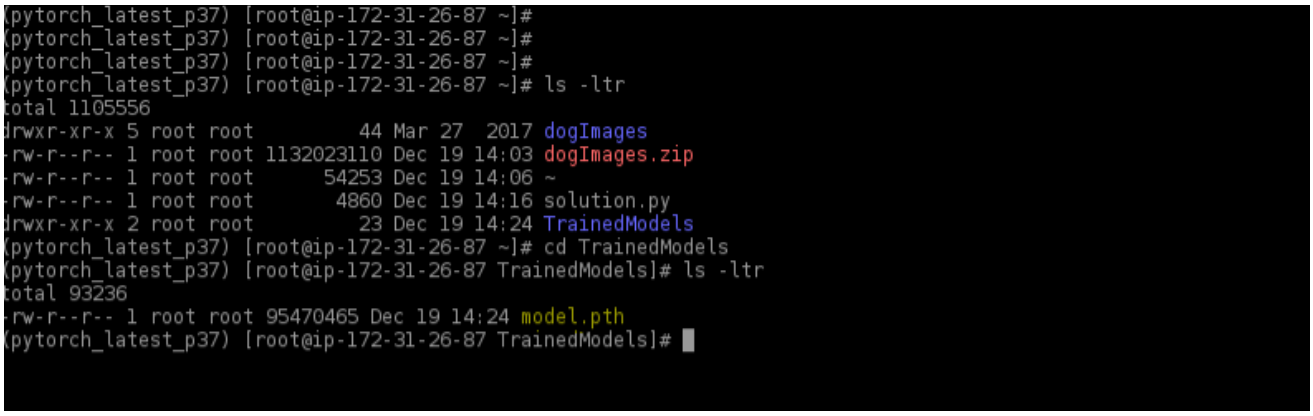


Figure 9 EC2 Training saved model.pth

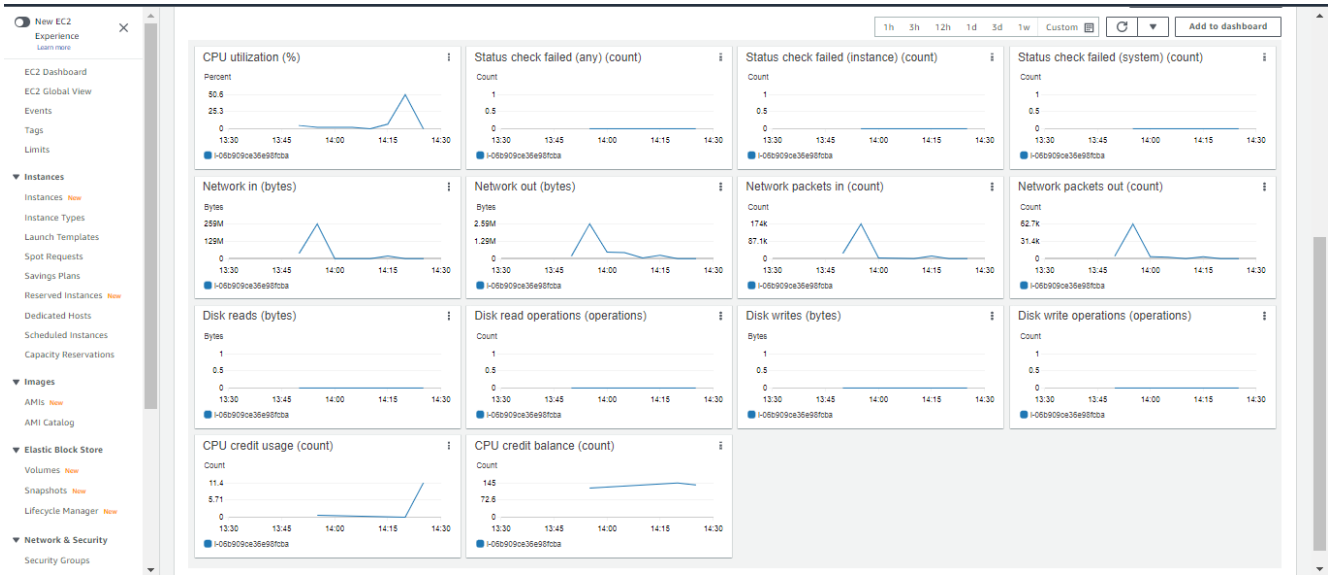


Figure 10 EC2 Instance Resource Metrics

4. Lambda functions

- The lambda functions will be used for invoking our deployed endpoints.

- The lambda function implemented in this project expects the image inputs in json format, which is used to invoke the model's deployed endpoint
- Given we have two endpoints deployed, one for the single instance training and the other for the multi-instance training, we will only use the multi-instance training jobs endpoint and create a lambda function for invoking that endpoint.
- Multi instance trained endpoint that we will be using: "**pytorch-inference-2021-12-19-09-53-14-575**"
- We created the lambda function with the corresponding changes to invoke the endpoint:

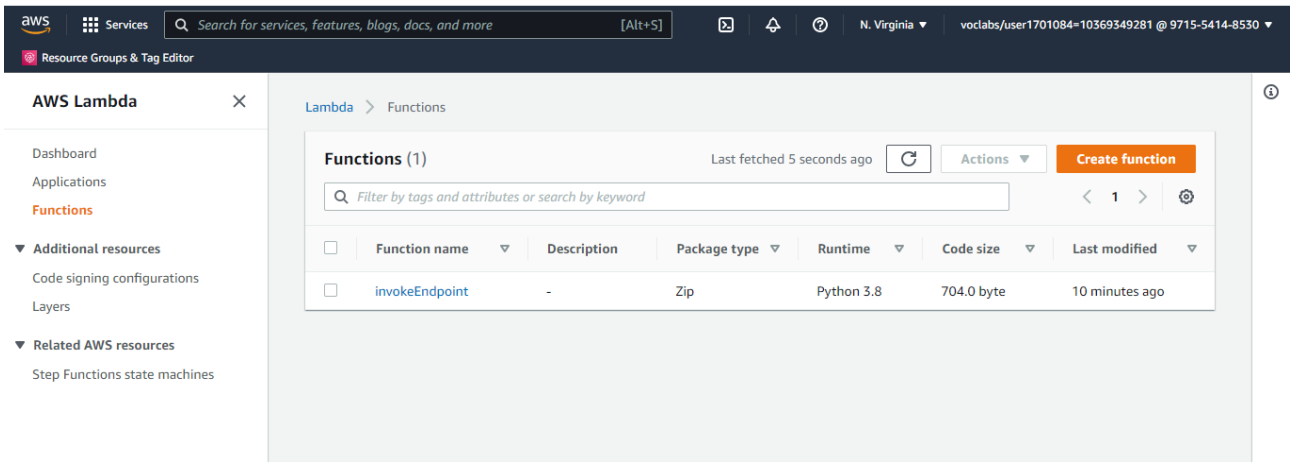
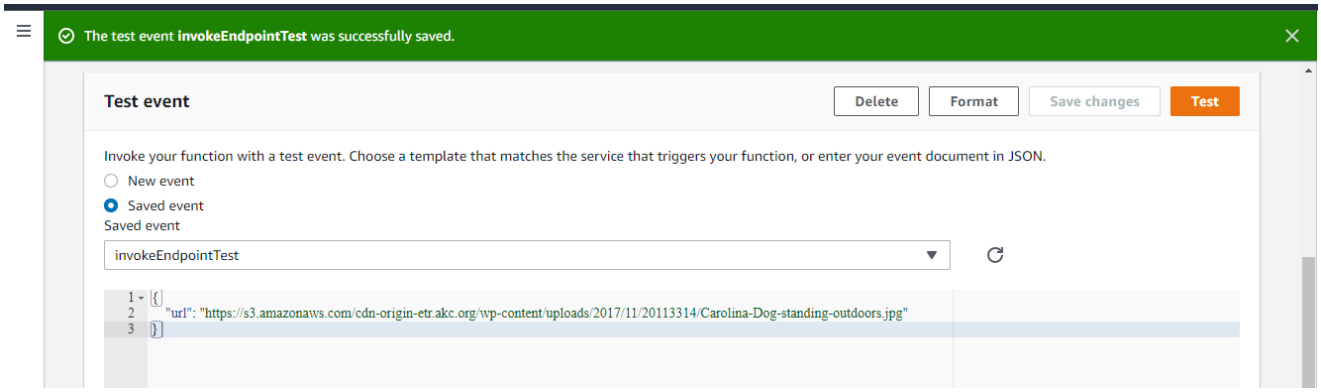


Figure 11 Lambda Function

5. Security and Testing

- We had created a new role for this lambda functions with basic accesses.
- We used the below test event to test our lambda function:



- Now when we tried to execute the test event we got and **AccessDeniedException.** This was expected as the lambda function did not have access to invoke the sagemaker endpoint.

- Error Message:


```
{
  "errorMessage": "An error occurred (AccessDeniedException) when calling the InvokeEndpoint operation: User: arn:aws:sts::971554148530:assumed-role/invokeEndpoint-role-gzgryuzu/invokeEndpoint is not authorized to perform: sagemaker:InvokeEndpoint on resource: arn:aws:sagemaker:us-east-1:971554148530:endpoint/pytorch-inference-2021-12-19-09-53-14-575 because no identity-based policy allows the sagemaker:InvokeEndpoint action",
  "errorType": "ClientError",
  "stackTrace": [
```

```
    "  File \"/var/task/lambda_function.py\", line 25, in lambda_handler\n
response=runtime.invoke_endpoint(EndpointName=endpoint_Name,\n",
    "  File \"/var/runtime/botocore/client.py\", line 386, in _api_call\n    return
self._make_api_call(operation_name, kwargs)\n",
    "  File \"/var/runtime/botocore/client.py\", line 705, in _make_api_call\n    raise
error_class(parsed_response, operation_name)\n"
  ]
}
```

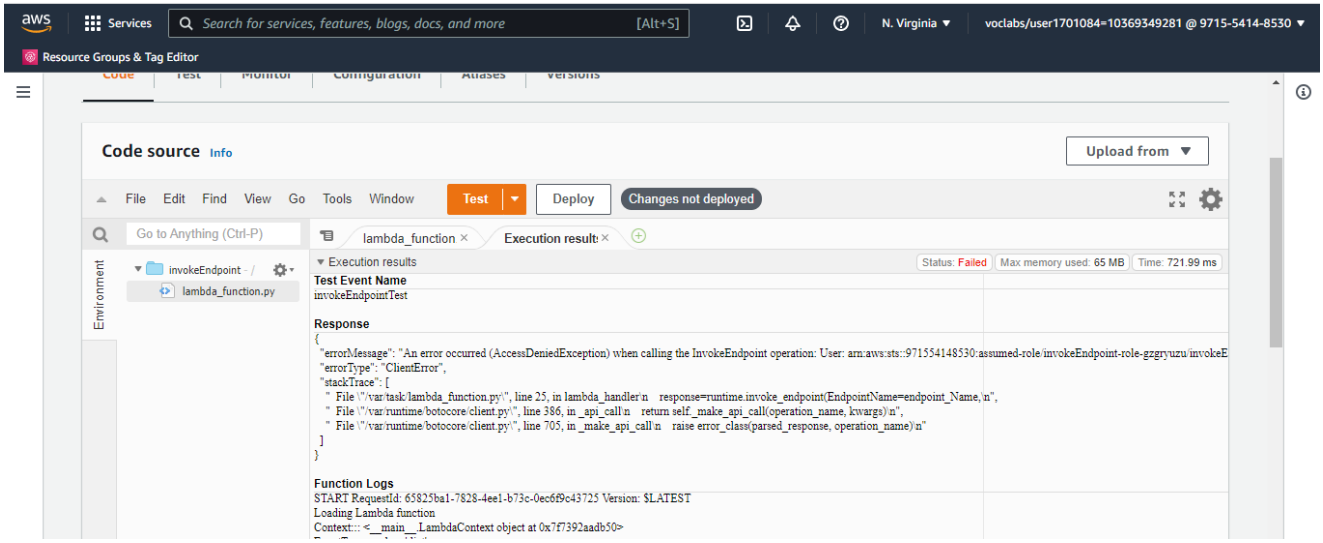



Figure 12 Lambda function test event failure response

- So went ahead and added the "**SagemakerFullAccess"** policy role to the lambda functions role.

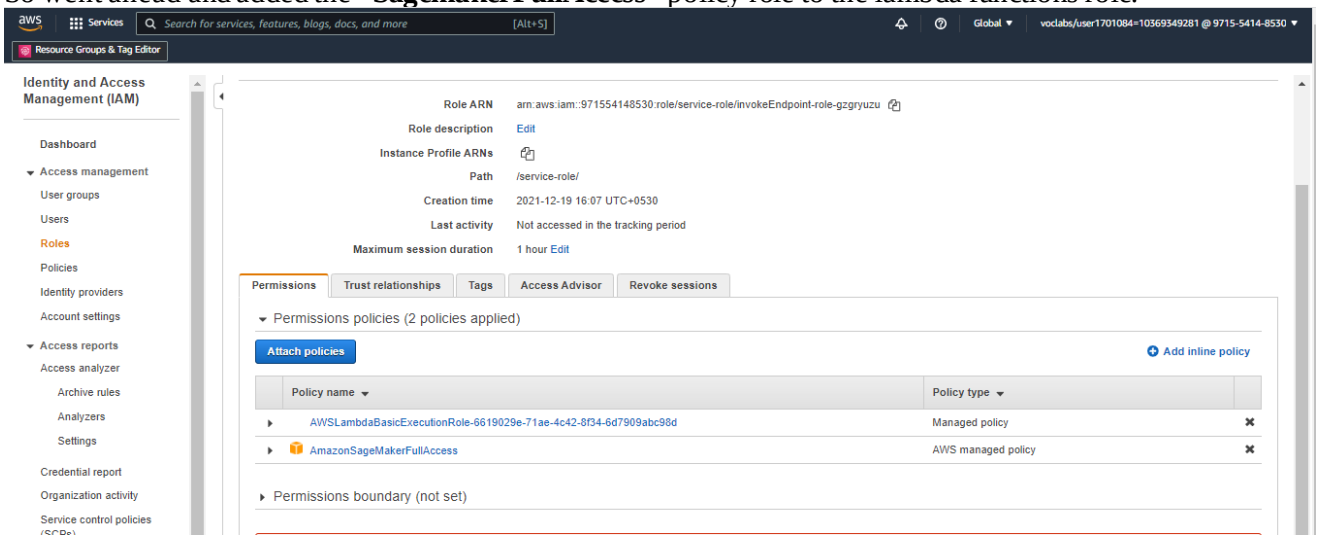


Figure 13 Lambda function role IAM permissions

- Post which we got the following output from the test event:
- (Please note that there are 133 dog breed and not 33 as mentioned in the project instructions. So we do expect there to be around 133 elements in the response object.)
- Response:


```
{
  "statusCode": 200,
  "headers": {
    "Content-Type": "text/plain",
    "Access-Control-Allow-Origin": "*"
```

```
  },
  "type-result": "<class 'str'>",
  "COntent-Type-In": "<__main__.LambdaContext object at 0x7fd68e644b20>",
  "body": "[[-8.698302268981934, -4.179177761077881, -4.179353713989258, -0.2588956952095032, -
1.487127661705017, -2.1415610313415527, -3.4534683227539062, -1.7365180253982544, -
6.828235626220703, 0.2051820158958435, -2.6021368503570557, -2.8923962116241455, -
6.523495197296143, -2.6120450496673584, -5.095528602600098, -2.9676473140716553, -
3.553805351257324, -3.8343188762664795, -7.553815841674805, 0.12643051147460938, -
7.497677803039551, -3.4655911922454834, -5.898045063018799, -5.160151958465576, -
3.6235978603363037, -8.570965766906738, -1.8589730262756348, -4.407120704650879, -
3.5737500190734863, -5.470921039581299, -6.042724132537842, -3.959641933441162, -
10.052687644958496, -2.8568434715270996, -5.760976791381836, -6.494334697723389, -
4.099149703979492, -5.4974493980407715, -2.2055768966674805, -2.012829065322876, -
1.8965739011764526, -6.571795463562012, -0.27010539174079895, -3.2459044456481934, -
1.2880767583847046, -5.312222003936768, -3.113386631011963, -1.4697904586791992, -
1.6495234966278076, -1.0736843347549438, -2.064608573913574, -5.6230669021606445, -
5.238499641418457, -2.8898324966430664, -7.891391277313232, -2.679391622543335, -
5.404878616333008, -6.20999002456665, -2.989795684814453, -0.5135252475738525, -
5.893244743347168, -5.591003894805908, -5.265442371368408, -4.828556060791016, -
1.2217851877212524, -5.282491683959961, -2.0269908905029297, -7.6696085929870605, -
3.052612543106079, -2.5704259872436523, -0.8108125925064087, -5.741940975189209, -
6.503178596496582, -7.224191188812256, -6.6316328048706055, -2.221468925476074, -
6.28223991394043, -4.524191856384277, -2.2097902297973633, -2.223500967025757, -
1.4676966667175293, -5.632909774780273, -3.556790590286255, -1.4694610834121704, -
4.029181480407715, -5.289389610290527, -3.64164400100708, -7.0696120262146, -6.260553359985352,
-3.436521291732788, -3.2722976207733154, -4.284804344177246, -5.729438781738281, -
5.0031046867370605, -3.415367603302002, -3.6307361125946045, -6.322794437408447, -
3.6496636867523193, -5.761388301849365, -5.25753116607666, -3.751535654067993, -
3.4450197219848633, -2.7776808738708496, -6.916290760040283, -1.4568134546279907, -
7.214130878448486, -3.1477553844451904, -0.7226721048355103, -1.8855265378952026,
0.5120564699172974, -3.6698319911956787, -1.5091665983200073, -4.080456733703613, -
5.634585380554199, -3.0056800842285156, -0.5162562131881714, -3.0180211067199707, -
0.9867972135543823, -5.96330451965332, -2.6183390617370605, -4.35217809677124, -
3.350148916244507, -4.202278137207031, -5.385364055633545, -7.7063307762146, -
2.8624188899993896, -4.602965831756592, -2.9397473335266113, -5.298011302947998, -
5.27507209777832, -7.07054328918457, -3.5364830493927, -5.396554470062256]]"
}
```

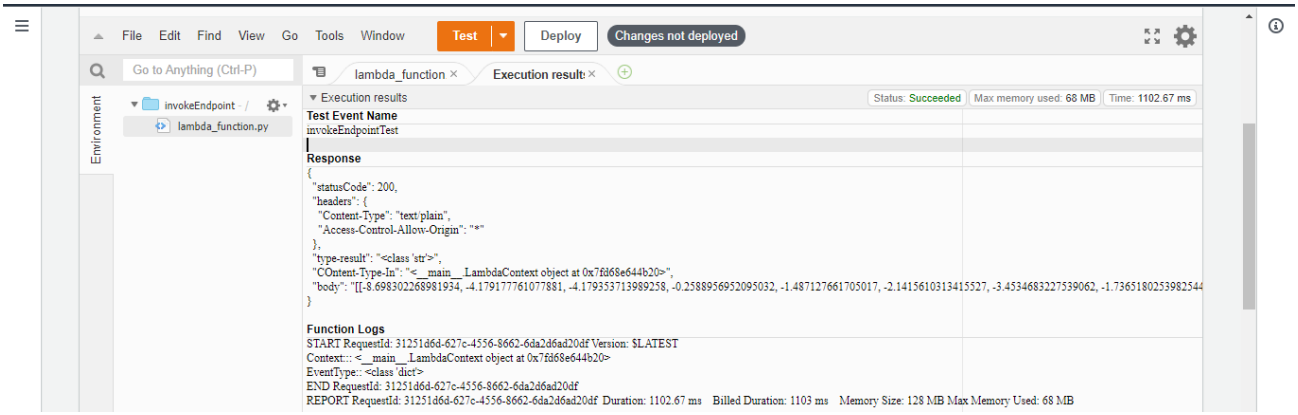



Figure 14 lambda function test event success response

- I'm concerned about the "Full Access" type permission policies that are available.

- For example, for this lambda functions we have provided the lambda function a Full Access to Sagemaker resources, but this does not seem to follow the concept of least privilege access.
- Ideally, one should only allow these lambda functions to query the endpoints that they're intended and allowed to query.
- We will have to do some more analysis to figure out if there's anything we could do about it.
- Furthermore, another concern is that the account's root user does not employ MFA
- Looking at the IAM roles that are currently active, all the roles seems to be necessary and also most of the roles have been added on a per need basis.
- However, we need to keep an eye on the roles dashboard to make sure only relevant roles absolutely necessary for currently active projects , are the only roles that are in active state to prevent unauthorized accesses.

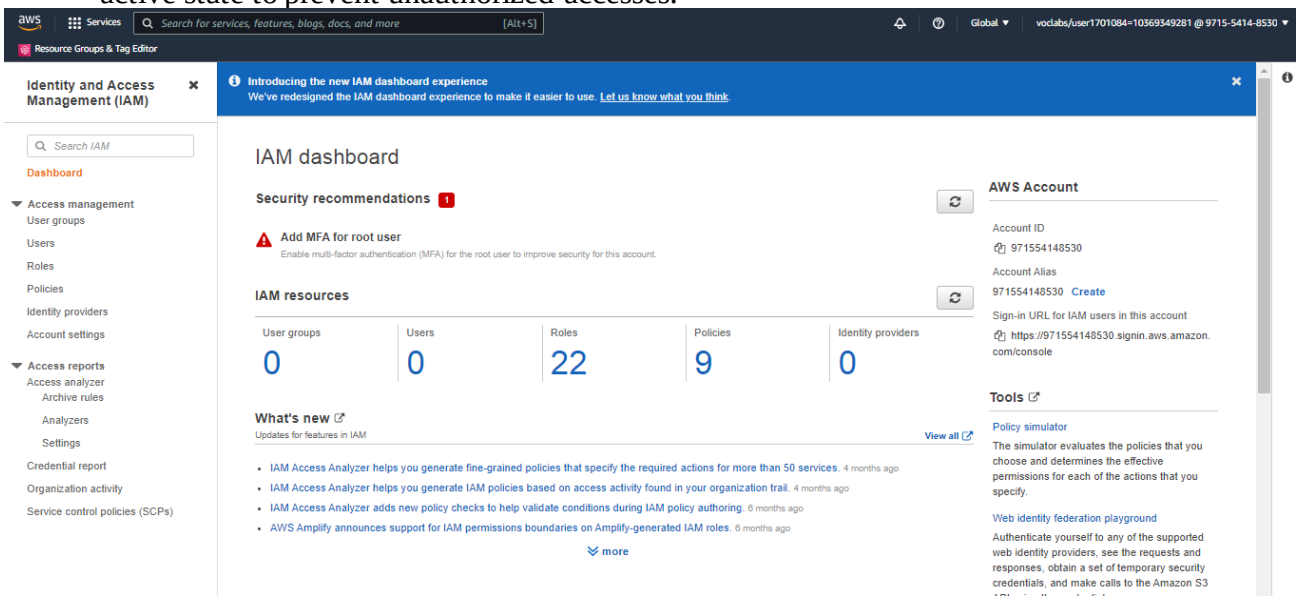


Figure 15 IAM Dashboard

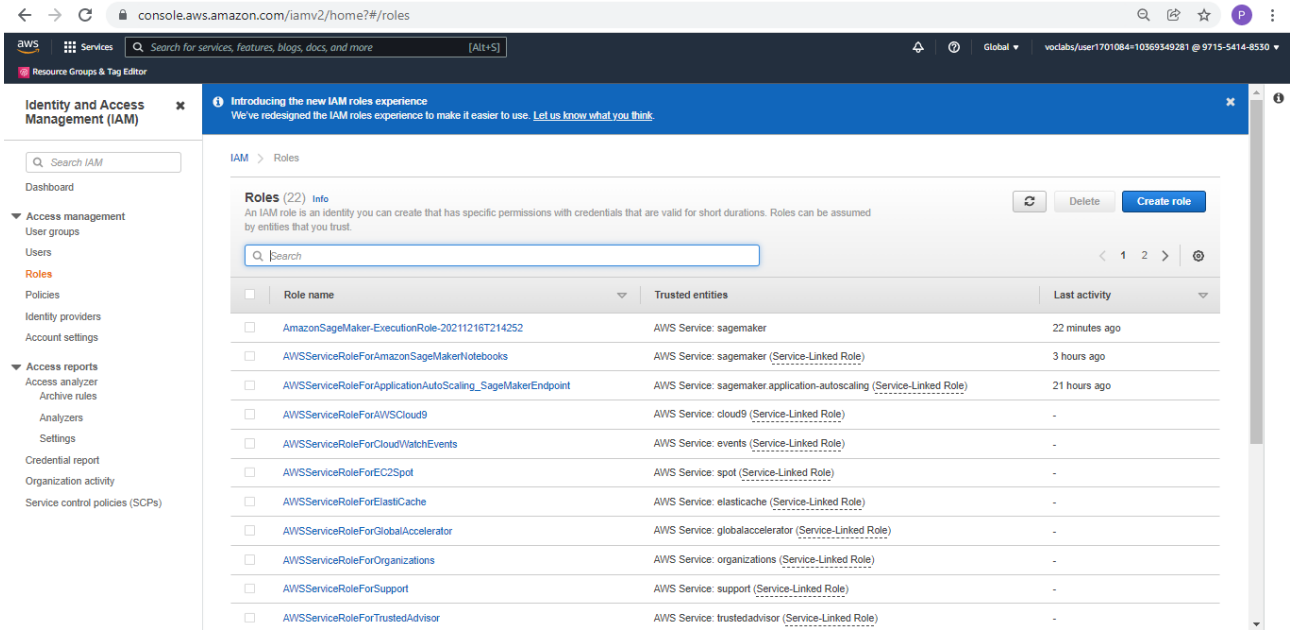


Figure 16 IAM Roles

6. Concurrency and Auto-scaling

- Before adding in configs for Concurrency and Auto-scaling for our lambda functions we will first create a version config for our lambda function.

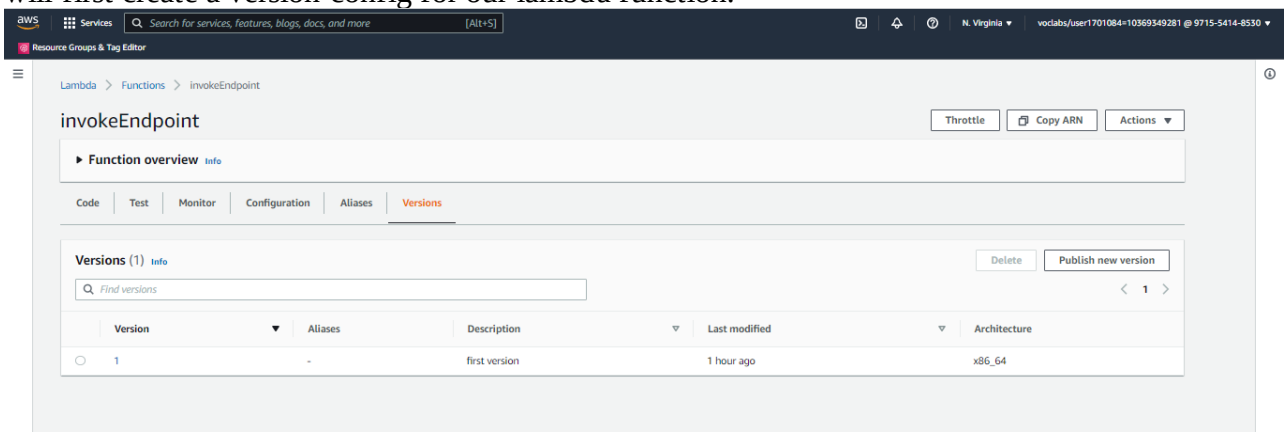


Figure 17 Lambda function version config

- For the lambda function we have set the **reserved concurrency** to be 5. This implies that the lambda function would be able to handle upto 5 requests concurrently at the same time. This would definitely help lower latency issues in situations when there is higher traffic than usual.

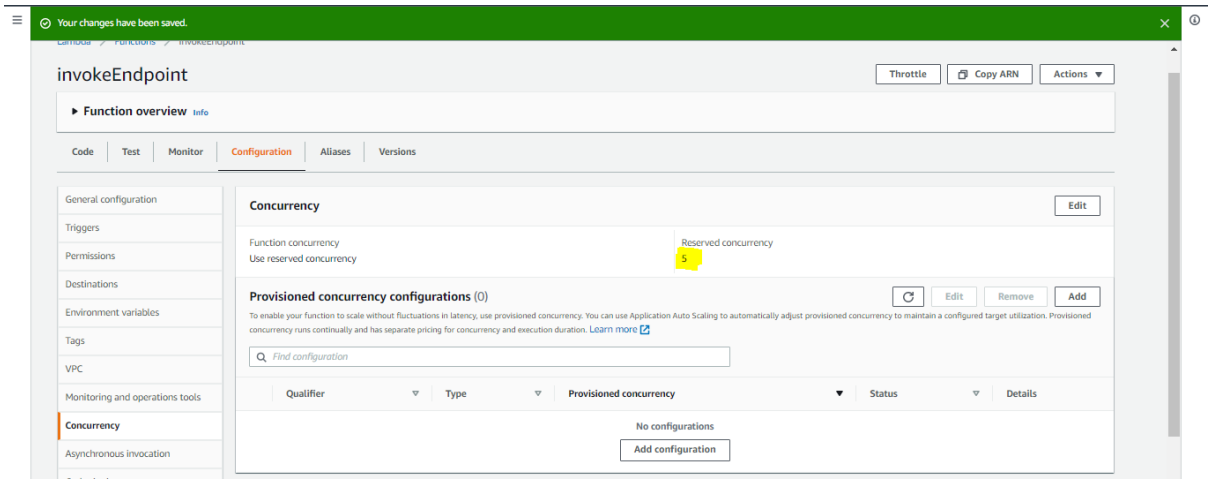


Figure 18 lambda function reversed concurrency

- Given the current use case, ideally using only reserved concurrency should have sufficed for our use case and we might not need to consider using the provisioned concurrency configs. However, for the sake of completion, I tried to add in the config for the provisioned concurrency as well. We set the provision concurrency to be 2.

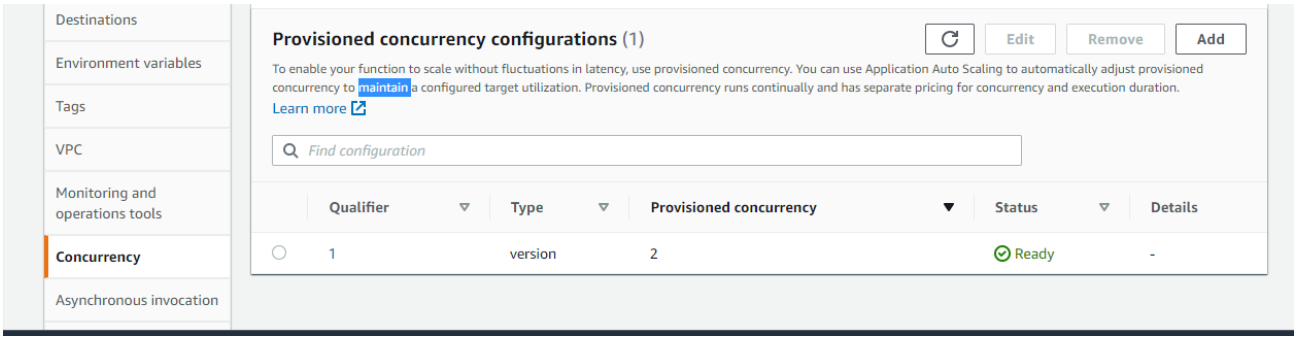


Figure 19 lambda function provisioned concurrency

- For adding config for auto-scaling we have added the below configs:

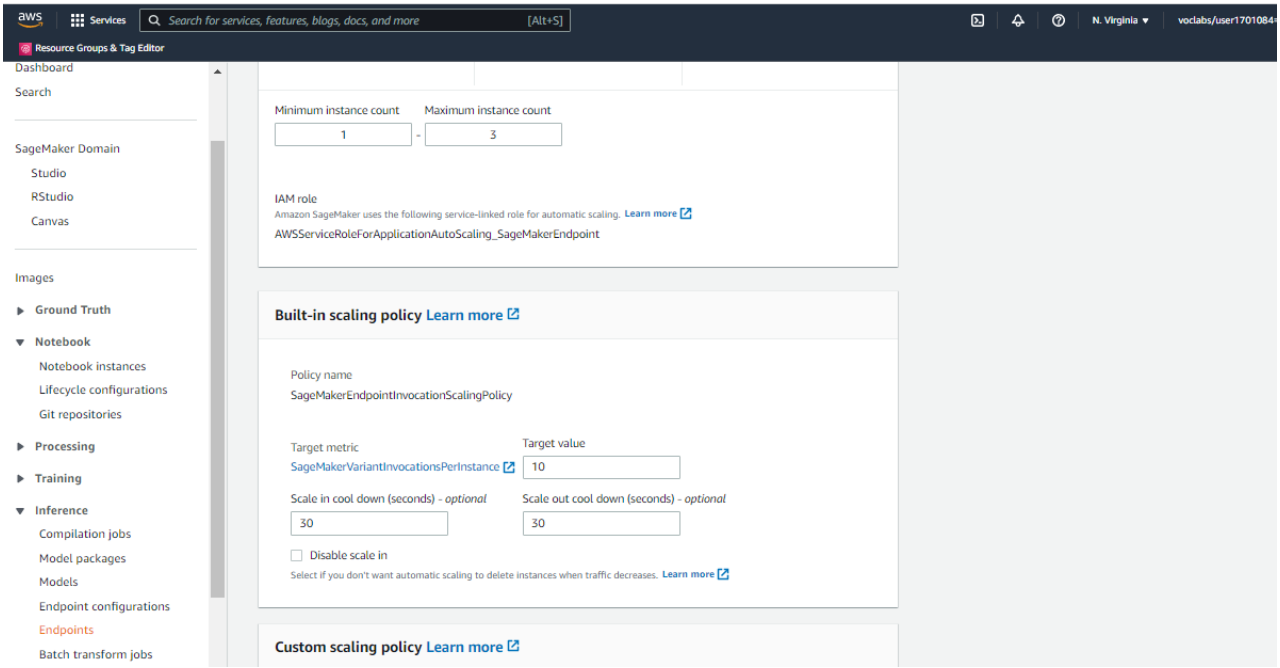


Figure 20 Endpoint Auto-scaling Config

- We have set the max instance count to 3 for Auto-scaling, as considering the current requirement, auto scaling on 3 instances with a scale-in and scale-out cool down time of 30 seconds should be a reasonably good config.

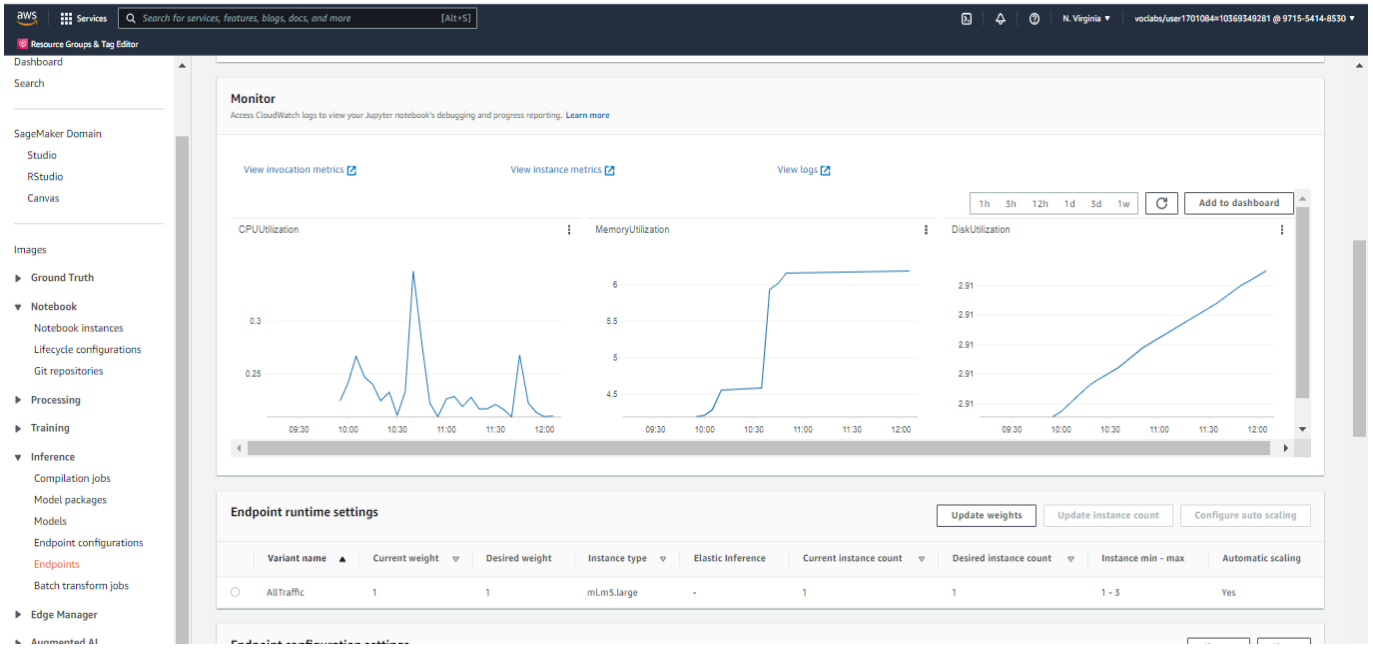


Figure 21 Endpoint Auto-scaling Metrics