

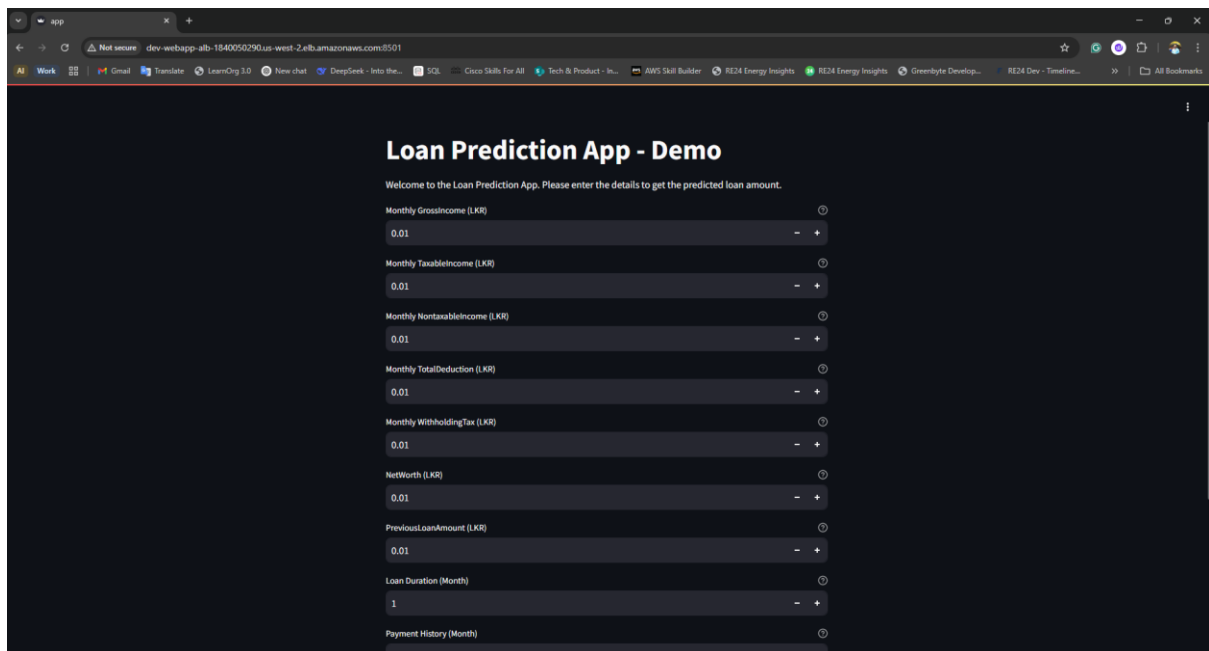
Platform Engineer Assignment

# Deploying a Containerised Application Using IaC

Akila Hettiarachchi

## Simple Web Application

URL: <http://dev-webapp-alb-1840050290.us-west-2.elb.amazonaws.com:8501>



## Step 1: Containerization

I developed a demo web application using Python and Streamlit. I then containerised it using Docker, based on the python:3.11-slim image. The Docker container exposes port 8501, the default port Streamlit uses.



Figure 01: Docker file

## Step 2: Infrastructure as Code (IaC) with Terraform

### 2.1 Define AWS ECS Cluster, Services, and Task Definitions, etc., using Terraform.

I set up a remote backend using AWS S3 and DynamoDB to store the Terraform state file and manage state locking. This setup ensures that Terraform works seamlessly both locally and in the GitLab pipeline, and it also facilitates team collaboration.

Using Terraform, I created an Amazon ECR repository to store the built web application image. I then defined an ECS cluster, service, and task definition to run the stored image. Additionally, I created a VPC and subnets to support the ECS infrastructure. A security group was configured to manage ECS network access, allowing public access on port 8501.

I assigned an IAM role for ECS task execution and configured CloudWatch Logs to view application logs. Finally, I set up an Application Load Balancer (ALB) and a target group to handle incoming web application traffic.

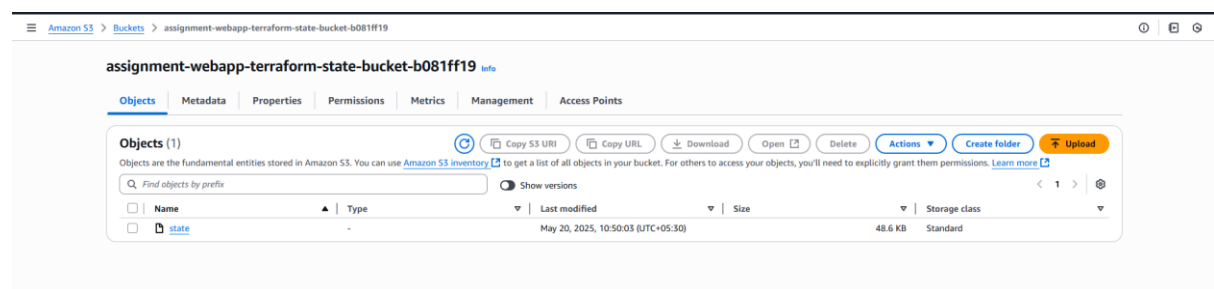


Figure 02: S3 bucket for state store

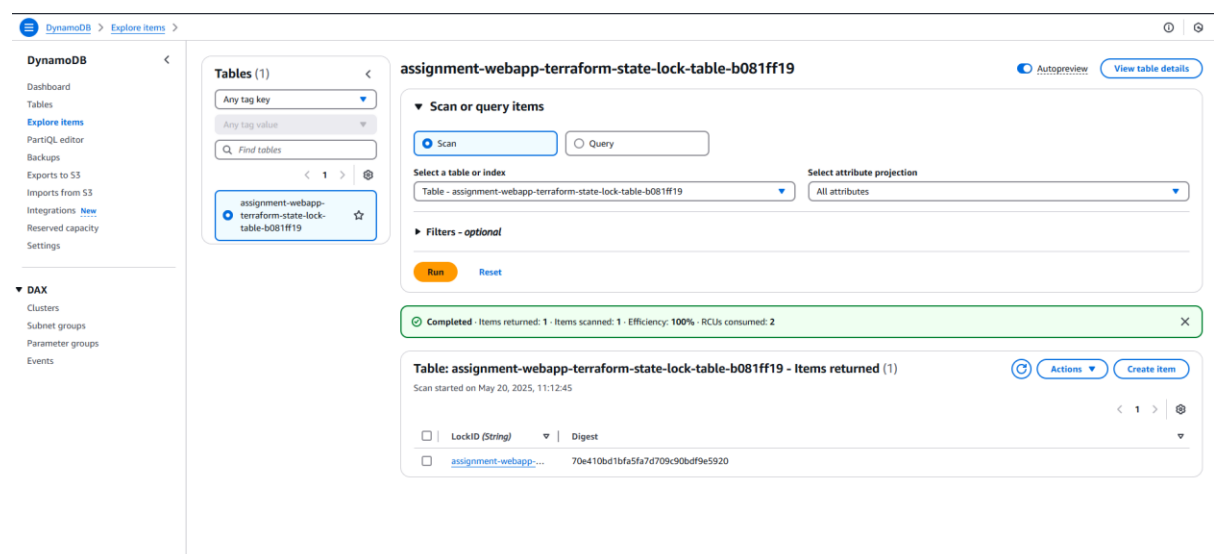


Figure 03: DynamoDB for state locking

Amazon ECR > Private registry > Repositories > dev-webapp-ecr

Amazon Elastic Container Registry

▼ Private registry

- Repositories
- Summary
- Images**
- Permissions
- Lifecycle Policy
- Repository tags
- Features & Settings

▼ Public registry

- Repositories
- Settings

ECR public gallery

Amazon ECS

Amazon EKS

Getting started

Documentation

Images (18)

Search artifacts

<input type="checkbox"/>	Image tag	Artifact type	Pushed at	Size (MB)	Image URI	Digest	Last recorded pull time
<input type="checkbox"/>	webapp-1.0	Image	May 20, 2025, 10:49:22 (UTC+05.5)	401.37	Copy URI	sha256:ef185297d5e24e254d12989009b19...	-
<input type="checkbox"/>	-	Image	May 20, 2025, 03:02:52 (UTC+05.5)	397.62	Copy URI	sha256:4cc3db472669c8b6e314a3cab4d83...	May 20, 2025, 03:04:04 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 18, 2025, 23:45:42 (UTC+05.5)	397.60	Copy URI	sha256:c1134ffec7e87b3507deebc7f0e55c...	May 20, 2025, 00:43:22 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 18, 2025, 00:56:41 (UTC+05.5)	397.60	Copy URI	sha256:96bfa8bb03cfc5e17b610ee9b5ba2b...	May 18, 2025, 00:56:58 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 23:01:16 (UTC+05.5)	397.60	Copy URI	sha256:305b388c7436f113fb2a4fe7c3a8d1...	May 17, 2025, 23:01:32 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 22:53:21 (UTC+05.5)	397.60	Copy URI	sha256:de7a7443cb43f2d361d6e46049cf89...	May 17, 2025, 22:53:36 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 22:49:00 (UTC+05.5)	397.60	Copy URI	sha256:5aa5da3bf7f1d10cda393c1d4f301...	May 17, 2025, 22:49:17 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 22:09:36 (UTC+05.5)	397.60	Copy URI	sha256:ff7c7579a7e716c248dee09d1a1028...	May 17, 2025, 22:09:53 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 21:53:47 (UTC+05.5)	397.60	Copy URI	sha256:143a373efb1f5ea931a7093f244ad...	May 17, 2025, 21:54:05 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 21:45:22 (UTC+05.5)	397.60	Copy URI	sha256:7f075454902fda87b9e15ce25ae617...	May 17, 2025, 21:45:38 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 21:35:05 (UTC+05.5)	397.60	Copy URI	sha256:75432d56af2146a2b855924fe5db5...	May 17, 2025, 21:35:21 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 21:13:20 (UTC+05.5)	397.60	Copy URI	sha256:6c7ca794f580c7dc3552abfde53bc...	May 17, 2025, 21:13:36 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 20:27:37 (UTC+05.5)	397.60	Copy URI	sha256:9d03acbb3c444f0faec921ae05d5c9...	May 17, 2025, 20:27:52 (UTC+05.5)
<input type="checkbox"/>	-	Image	May 17, 2025, 20:19:37 (UTC+05.5)	397.60	Copy URI	sha256:590b98be7f67b8fc49c248d1c2542a...	May 17, 2025, 20:19:53 (UTC+05.5)

Figure 04: ECR Repo

Services | Tasks | Infrastructure | Metrics | Scheduled tasks | Configuration | Tags

Services (1) Info

Filter services by value

Filter launch type: Any launch type

Filter service type: Any service type

Manage tags | Update | Delete service | Create

<input type="checkbox"/>	Service name	ARN	Status	Service type	Created at	Deployments and tasks	Last deployment	Task definiti...
<input type="checkbox"/>	dev-webapp-service	arn:aws:ecs:us-v...	Active	REPLICA	11 hours ago	1/1 Tasks running	Completed	View dev-webapp-ta...

Figure 05: ECS Service

Amazon Elastic Container Service > Clusters > dev-webapp-cluster > Tasks > 2fb20adaba4641d7b58b48f1216f7906 > Configuration

Amazon Elastic Container Service

Clusters

- Namespaces
- Task definitions
- Account settings

Install AWS Copilot

Amazon ECR

Repositories

AWS Batch

Documentation

Discover products

Subscriptions

Tell us what you think

Configuration

Operating system/Architecture: Linux/X86\_64

CPU | Memory: .25 vCPU | 5 GB

Platform version: 1.4.0

Fault injection: -

ECS Exec: info

Turned off

Capacity provider: -

Launch type: FARGATE

Container instance ID: -

Task definition: revision: dev-webapp-task-7

Task group: service:dev-webapp-service

Service: dev-webapp-service

ENI ID: eni-0030a86b678a73c33

Network mode: awsvpc

Subnet ID: subnet-0a8db35937dac61e9

Public IP: 18.256.249.228 | open address

Private IP: 10.0.1.126

MAC address: 06:bc:f8:5d:42:01

Containers (1)

Filter containers

<input type="checkbox"/>	Container name	Container runtime ID	Image URI	Image Digest	Status	Health status	CPU	Memory hard/soft limit
<input checked="" type="checkbox"/>	dev-webapp-container	2fb20adaba4641d...	686255942291.dkr.ecr.us...	sha256:4c...	Running	Unknown	0	- / -

Figure 06: ECS Task

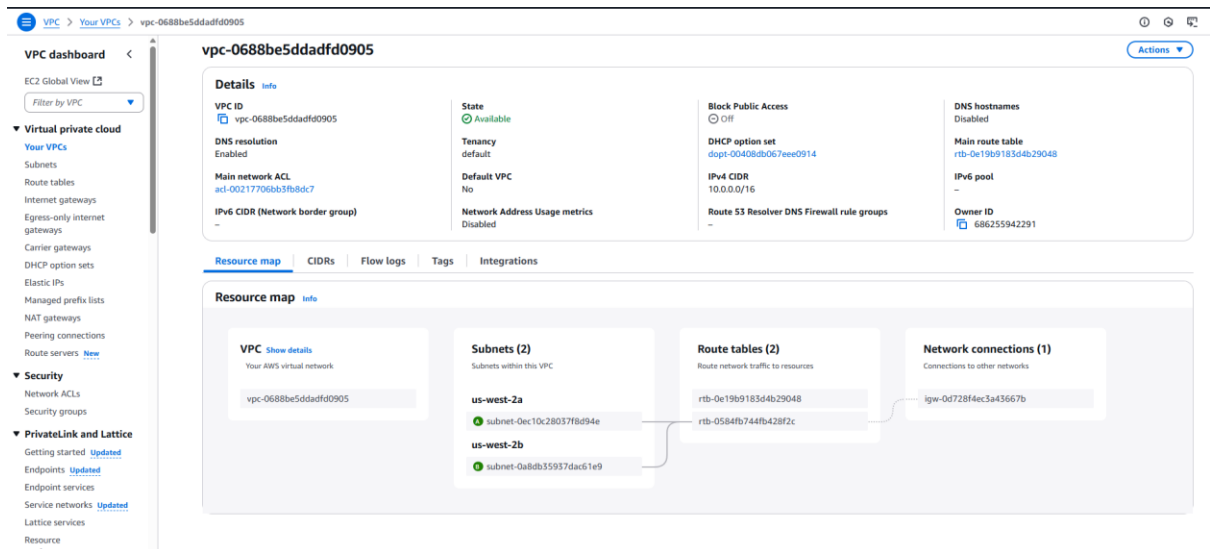


Figure 07: VPC

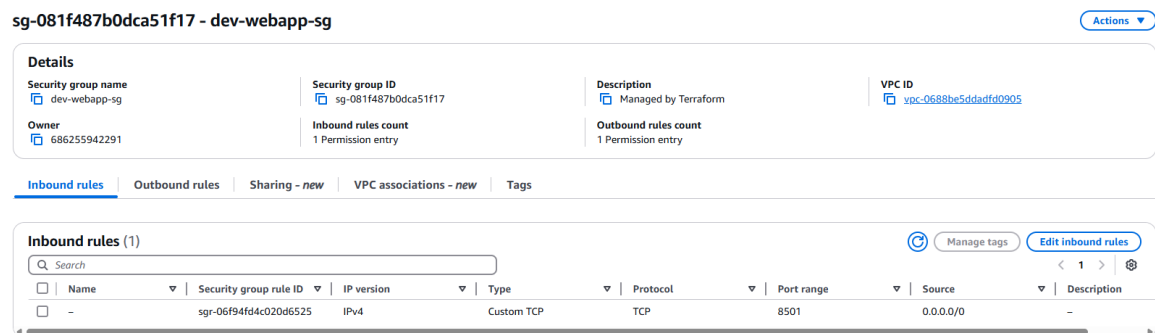


Figure 08: Security group

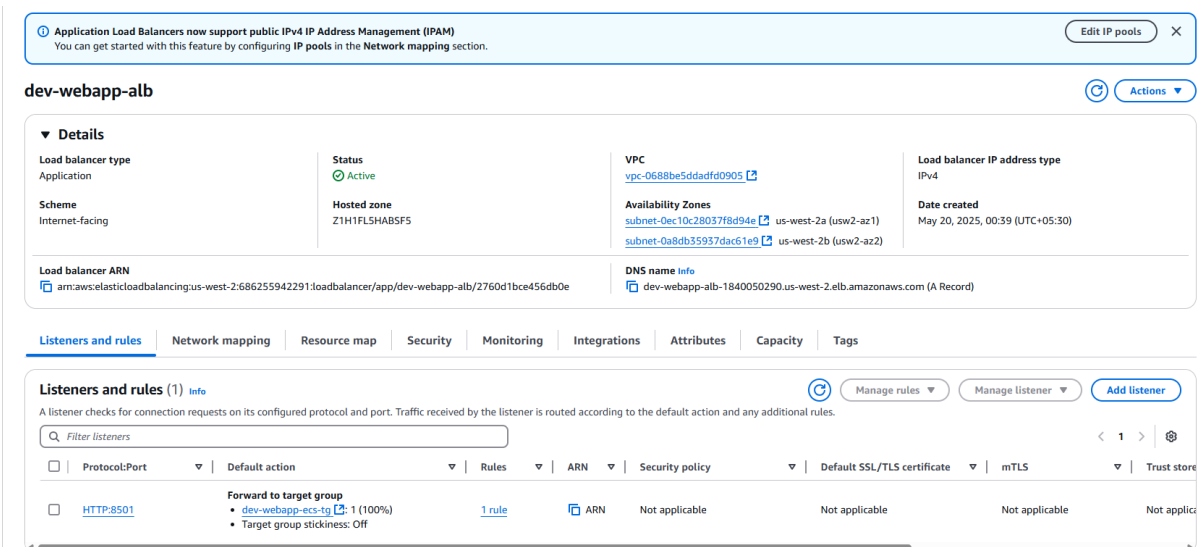


Figure 09: Load Balancer

## 2.2 Implement vertical scaling.

AWS does not provide a native service for vertical auto scaling. Therefore, I implemented a custom solution using CloudWatch Alarms and Lambda functions. CloudWatch metrics monitor ECS service CPU usage and ALB target group activity.

If CPU usage exceeds 80%, a CloudWatch alarm triggers the scale-up Lambda function, which vertically scales the ECS task by increasing vCPU and memory. Conversely, if CPU usage drops below 20% and there is at least one active request, a scale-down Lambda function is triggered to reduce vCPU and memory allocation. It helps efficiently manage resources while maintaining application performance.

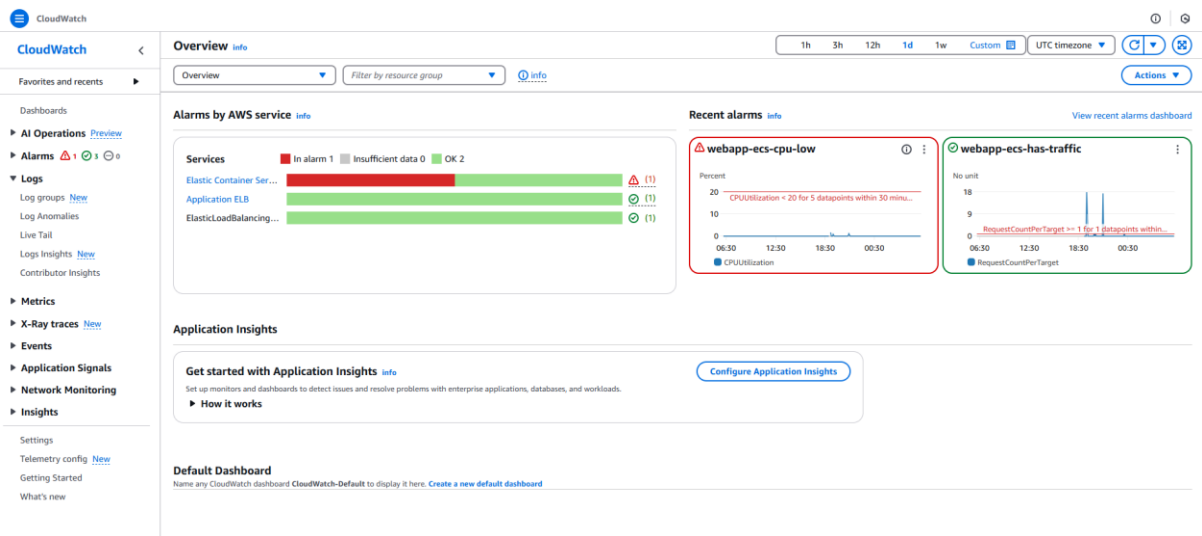


Figure 10: CloudWatch alarm overview

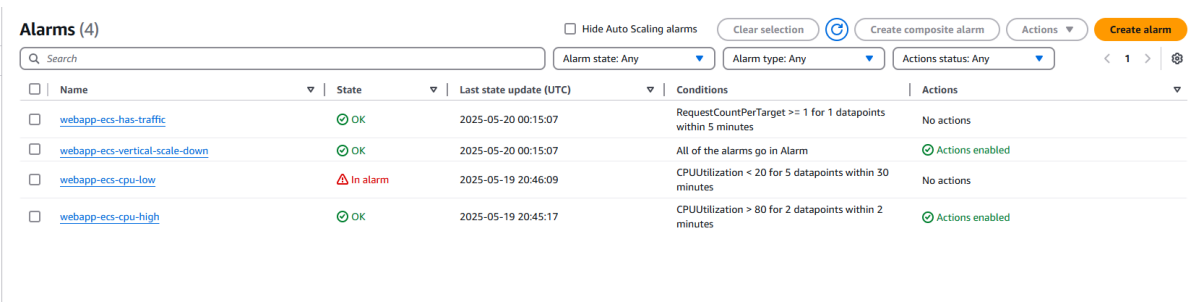


Figure 11: CloudWatch alarms list

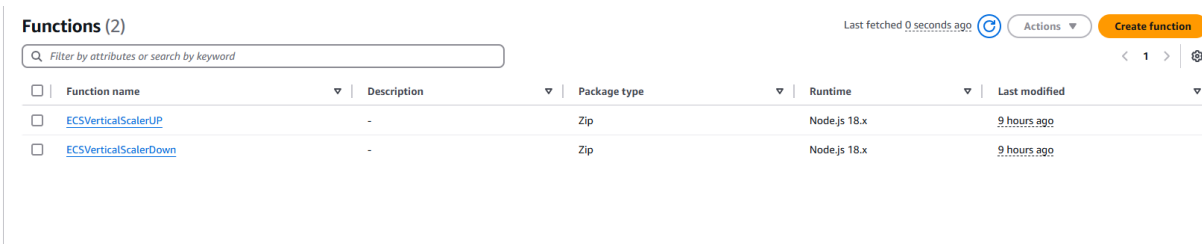


Figure 12: Lambda Functions

## Step 3: GitLab CI/CD Pipeline

I defined a GitLab CI/CD pipeline with three stages:

### 1. Run Tests

Used the “python:3.10-slim” image to execute web application tests. The tests were run using “pytest”.

### 2. Build Image

Used the “docker:28.1.1” image and “docker:28.1.1-dind” as the service image. In this stage, the pipeline logs into Amazon ECR using the AWS CLI, builds the Docker container, and pushes it to the ECR repository.

### 3. Deploy Image

Used the “alpine:3.20” image. A Python virtual environment is created, and the AWS CLI is installed and configured. Then, Terraform is installed using the official binary. Finally, the Terraform configuration is applied to provision infrastructure and deploy the newly built image to ECS.

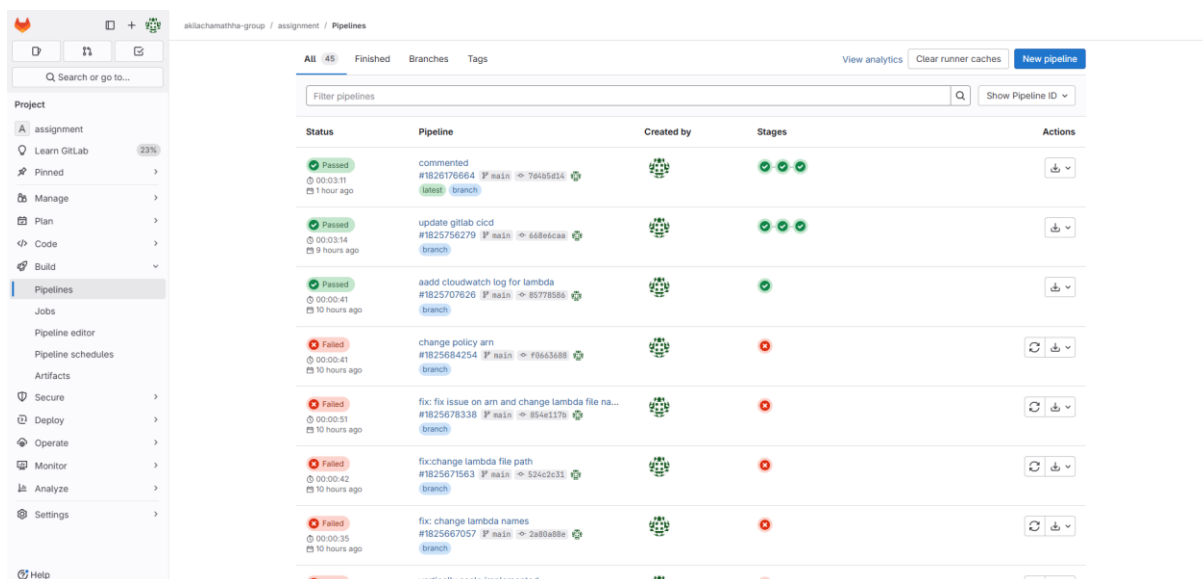


Figure 13: GitLab Pipeline runs

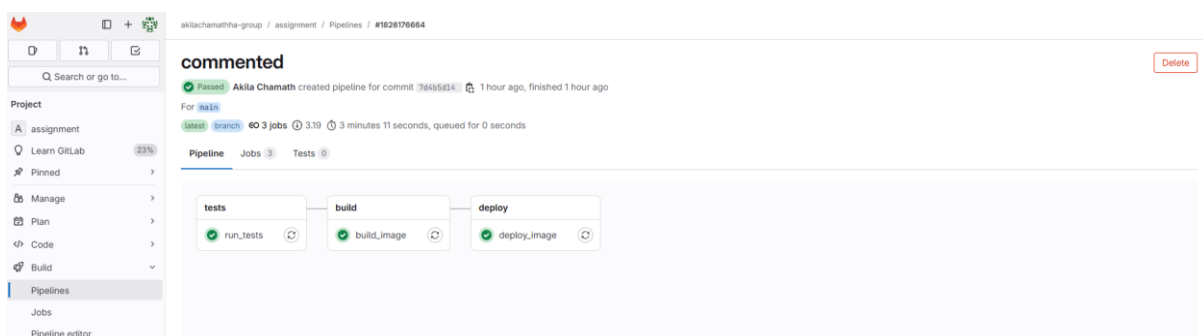


Figure 14: GitLab Pipeline Stages

```

197 Downloading tzdata-2025.2-py2.py3-none-any.whl (347 kB)
198 Installing collected packages: pytz, watchdog, urllib3, tzdata, tornado, toml, threadpoolctl, tenacity, smmap, six, rpds-py, pyarrow, protobuf, pillow, pac
kaging, numpy, narwhals, MarkupSafe, joblib, idna, click, charset-normalizer, certifi, cachetools, blinker, attrs, scipy, requests, referencing, python-dat
eutil, Jinja2, gitdb, scikit-learn, pydeck, pandas, jsonschema-specifications, gitpython, jsonschema, altair, streamlit
199 Attempting uninstall: packaging
200 Found existing installation: packaging 25.0
201 Uninstalling packaging-25.0:
202 Successfully uninstalled packaging-25.0
203 Successfully installed MarkupSafe-3.0.2 altair-5.5.0 attrs-25.3.0 blinker-1.9.0 cachetools-5.5.2 certifi-2025.4.26 charset-normalizer-3.4.2 click-8.2.0 git
db-4.0.12 gitpython-3.1.44 idna-3.10 Jinja2-3.1.6 joblib-1.5.0 jsonschema-4.23.0 jsonschema-specifications-2025.4.1 narwhals-1.40.0 numpy-2.0.2 packaging-2
4.2 pandas-2.2.3 pillow-11.2.1 protobuf-5.29.4 pyarrow-2025.0.0 pydeck-0.9.1 python-dateutil-2.9.0.post0 pytz-2025.2 referencing-0.36.2 requests-2.32.3 rpds-
py-0.25.0 scikit-learn-1.6.1 scipy-1.15.3 six-1.17.0 smmap-5.0.2 streamlit-1.43.0 tenacity-9.1.2 threadpoolctl-3.6.0 toml-0.10.2 tornado-6.5 tzdata-2025.2
urllib3-2.4.0 watchdog-6.0.0
204 WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager, possibly rendering your
system unusable. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv. Use the --root-user-action option if you know w
hat you are doing and want to suppress this warning.
205 $ echo "Running Test..."
206 Running Test...
207 $ pytest test_app.py
208 ===== test session starts =====
209 platform linux -- Python 3.10.17, pytest-8.3.5, pluggy-1.6.0
210 rootdir: /builds/[MASKED]-group/assignment/webapp
211 collected 1 item
212 test_app.py . [100%]
213 ===== warnings summary =====
214 test_app.py::test_loan_prediction_ui_valid_inputs
215 /usr/local/lib/python3.10/site-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but RandomForestRegressor was
fitted with feature names
216 warnings.warn(
217 -- Docs: https://docs.pytest.org/en/stable/how-to/capture-warnings.html
218 ===== 1 passed, 1 warning in 1.30s =====
219 $ echo "Test completed successfully."
220 Test completed successfully.
221 Cleaning up project directory and file based variables
222 Job succeeded

```

Figure 15: GitLab Pipeline Run\_Test Log

```

294 #11 exporting to image
295 #11 exporting layers
296 #11 exporting layers 2.5s done
297 #11 writing image sha256:96c5b105f8ea992e7378f3d7792546c71bbdf587f3a5737b13415557ebd97ef done
298 #11 naming to docker.io/library/dev-webapp-ecr:webapp-1.0 done
299 #11 DONE 2.5s
300 $ docker tag $ECR_REPO_NAME:$IMAGE_TAG $ECR_REGISTRY/$ECR_REPO_NAME:$IMAGE_TAG
301 $ docker push $ECR_REGISTRY/$ECR_REPO_NAME:$IMAGE_TAG
302 The push refers to repository [[MASKED].dkr.ecr.us-west-2.amazonaws.com/dev-webapp-ecr]
303 c0494ce0b0f8: Preparing
304 9503192ecfa1: Preparing
305 d4dee230217e: Preparing
306 f1ba4063e60f: Preparing
307 6ba140ab1e68: Preparing
308 23aa89a8a424: Preparing
309 91bd78b864ed: Preparing
310 adb057d02f88: Preparing
311 6c4c763d22d0: Preparing
312 23aa89a8a424: Waiting
313 91bd78b864ed: Waiting
314 adb057d02f88: Waiting
315 6c4c763d22d0: Waiting
316 6ba140ab1e68: Pushed
317 f1ba4063e60f: Pushed
318 23aa89a8a424: Layer already exists
319 91bd78b864ed: Layer already exists
320 adb057d02f88: Layer already exists
321 c0494ce0b0f8: Pushed
322 d4dee230217e: Pushed
323 6c4c763d22d0: Pushed
324 9503192ecfa1: Pushed
325 webapp-1.0: digest: sha256:ef185297d5e24e254d12989009b191b11aedbeb501c4ec9e6c72aeadd83138659 size: 2207
326 Cleaning up project directory and file based variables
327 Job succeeded

```

Figure 16: GitLab Pipeline Build\_Image Log



```
175 aws_s3_bucket.public[1]: Refreshing state... [id=s3bucket-6a8d03537d8c0e7]
176 aws_route_table.public: Refreshing state... [id=rtb-9584fb744fb428f2c]
177 aws_lb_alb: Refreshing state... [id=arn:aws:elasticloadbalancing:us-west-2:[MASKED]:loadbalancer/app/dev-webapp-alb/2760d1bce456db0e]
178 aws_route_table_association.public[0]: Refreshing state... [id=rtbassoc-04bfa84124a68cc7e]
179 aws_route_table_association.public[1]: Refreshing state... [id=rtbassoc-0a1fec8d91eb09c71]
180 aws_cloudwatch_metric_alarm.has_traffic: Refreshing state... [id=webapp-ecs-has-traffic]
181 aws_lb_listener.http: Refreshing state... [id=arn:aws:elasticloadbalancing:us-west-2:[MASKED]:listener/app/dev-webapp-alb/2760d1bce456db0e/a137e3aa2a75f72f]
182 aws_ecs_service.webapp: Refreshing state... [id=arn:aws:ecs:us-west-2:[MASKED]:service/dev-webapp-cluster/dev-webapp-service]
183 aws_lambda_function.ecs_vertical_scaler_up: Refreshing state... [id=ECSVerticalScalerUP]
184 aws_cloudwatch_metric_alarm.cpu_low: Refreshing state... [id=webapp-ecs-cpu-low]
185 aws_lambda_function.ecs_vertical_scaler_down: Refreshing state... [id=ECSVerticalScalerDown]
186 aws_cloudwatch_composite_alarm.scale_down: Refreshing state... [id=webapp-ecs-vertical-scale-down]
187 aws_cloudwatch_metric_alarm.cpu_alarm: Refreshing state... [id=webapp-ecs-cpu-high]
188 Terraform used the selected providers to generate the following execution
189 plan. Resource actions are indicated with the following symbols:
190   ~ update in-place
191 Terraform will perform the following actions:
192   # aws_cloudwatch_metric_alarm.cpu_alarm will be updated in-place
193   ~ resource "aws_cloudwatch_metric_alarm" "cpu_alarm" {
194     ~ alarm_description      = "This metric monitors high CPU" -> "Scale Up when CPU > 88%"
195     id                      = "webapp-ecs-cpu-high"
196     tags                    = {}
197     # (21 unchanged attributes hidden)
198   }
199 Plan: 0 to add, 1 to change, 0 to destroy.
200 aws_cloudwatch_metric_alarm.cpu_alarm: Modifying... [id=webapp-ecs-cpu-high]
201 aws_cloudwatch_metric_alarm.cpu_alarm: Modifications complete after 0s [id=webapp-ecs-cpu-high]
202 Apply complete! Resources: 0 added, 1 changed, 0 destroyed.
203 Outputs:
204   ecs_cluster_name = "dev-webapp-cluster"
205   ecs_service_name = "dev-webapp-service"
206   load_balancer_dns = "dev-webapp-alb-1840650290.us-west-2.elb.amazonaws.com"
207 Cleaning up project directory and file based variables
208 Job succeeded
```

Figure 17: GitLab Pipeline Deploy\_Image Log

## Step 4: Monitoring & Logging

### 4.1 Enable ECS service logs for vertical scaling.

I created a CloudWatch Log Group to monitor and view the ECS service auto scaling logs generated by the Lambda functions.

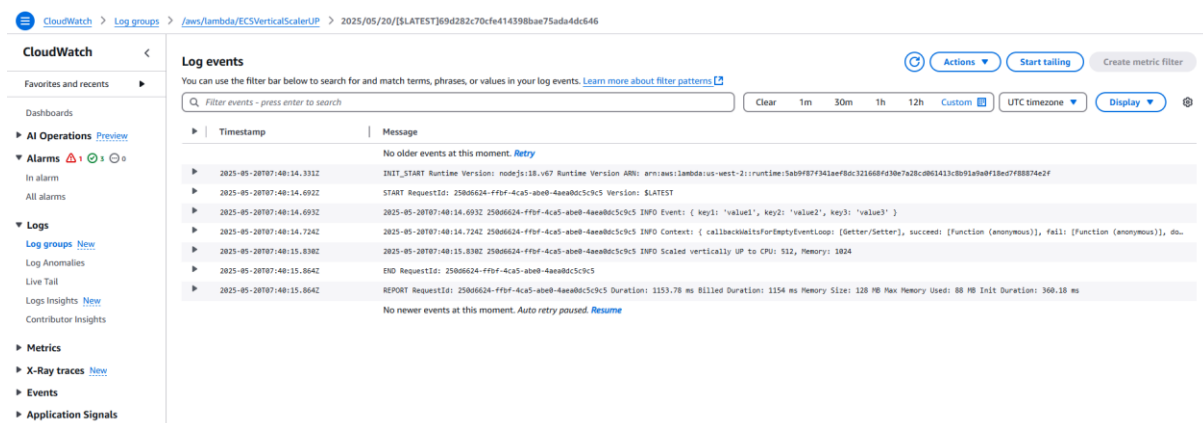


Figure 18: Lambda Log for auto scale up

The screenshot shows the AWS CloudWatch console interface. On the left is a navigation sidebar with sections like 'AI Operations', 'Alarms', 'Logs', 'Metrics', 'X-Ray traces', 'Events', 'Application Signals', and 'Network Monitoring'. The main panel is titled 'Log events' and displays a list of log entries for a specific Lambda function. The entries include timestamps and messages such as 'INIT\_START Runtime Version: nodejs18.v67', 'START RequestId: c1c91341-f714-4f3a-aa55-c0638e5fa937 Version: \$LATEST', and 'INFO Scaled vertically Down to CPU: 256, Memory: 512'. A message at the bottom states 'No newer events at this moment. Auto retry paused. Resume'.

Figure 19: Lambda Log for auto scale down

## 4.2 Capture ECS service logs using AWS CloudWatch.

I created an AWS CloudWatch Log Group for the ECS service to monitor and view its logs

This screenshot shows the AWS CloudWatch console for a specific ECS service log group. The breadcrumb navigation at the top indicates the path: CloudWatch > Log groups > /ecs/dev-webapp > ecs/dev-webapp-container/8b3bfda1f9f645be9802f663ef88ba73. The 'Log events' section displays a list of log entries with timestamps and messages. The messages include 'Collecting usage statistics. To deactivate, set browser.gatherUsageStats to false.', 'You can now view your Streamlit app in your browser.', and 'URL: http://0.0.0.0:8501'. A final message at the bottom says 'No newer events at this moment. Auto retry paused. Resume'.

Figure 20: ECS Service Log