STACS

Designed for Finance

STACS Solutions Engineering -

Enhancing Security with AWS API Gateway and ECS Containers

Transformative Technology for the Financial Industry

PREPARED BY:
Tracy Thanda Aye
WEBSITE:
www.stacs.io
CONTACT:
info@stacs.io

Contents

01 | Introduction

In the previous article we discussed about the AWS Elastic Container Service (ECS) architecture system where the AWS Application Load Balancer (ALB) is used as the connectivity layer between AWS API Gateway and the ECS service. However, for the API Gateway to use HTTPs endpoint for integration with ALB, the security group of the ALB needs to allow for all inbound traffic. This would be a huge security concern and so ALB is not exactly a suitable connectivity layer for communicating requests from API Gateway to the ECS services.

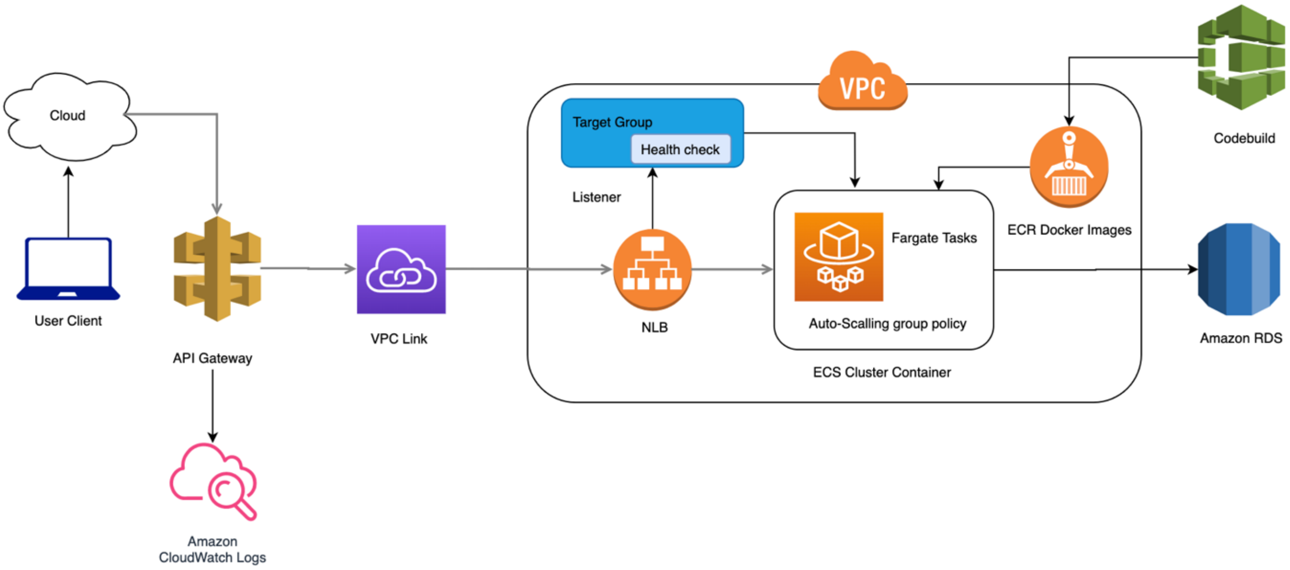
AWS Network Load Balancer (NLB) is a service distributes traffic across several servers by using TCP/IP network protocol, provides reliability and high performance for web servers. Once the API gateway receive the incoming request, the request data will be routed to the target integration service which is Virtual Private Cloud (VPC) link.

VPC Link provides private connectivity between API Gateway and NLB hosted in a VPC (VPC link can only target NLB). By providing a private endpoint integration that uses a VPC link to across the NLB, this ensures that the traffic is not exposed since network traffic is only allowed from the VPC Link.

The Application Load Balancer (ALB) supports both network and application layers, full 7 layers of the OSI model while the Network Load Balancer only supports 4 network layers. However, API gateway does not provide private IP or security groups that can be configured in the ALB security group's inbound rule. Thus, the best approach and solution is to connect ECS services to API Gateway with NLB instead since this setup uses the more secure VPC link.

02 | Technical Architecture

The high-level technical architecture setup is shown as follows:

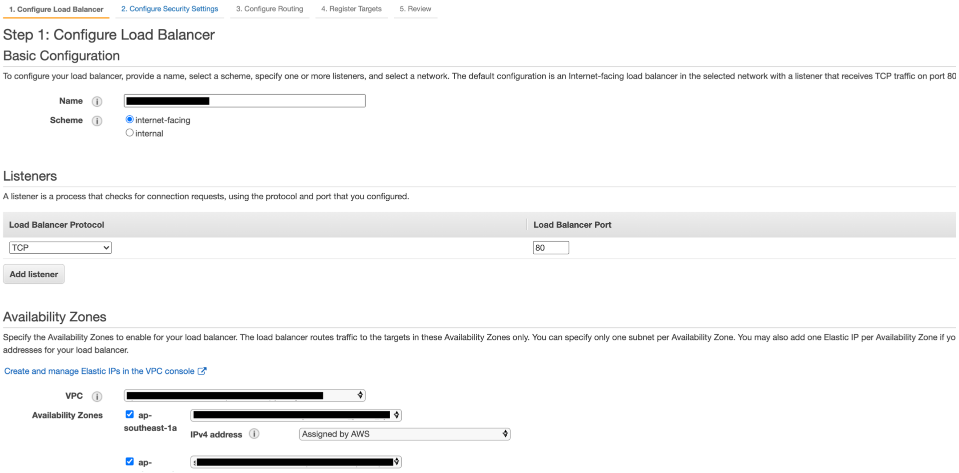


03 | Setting up Services for AWS API Gateway

This section provides a walkthrough that enables the full setup from AWS ECS to AWS API Gateway with the Network Load Balancer fully configured.

3.1 Network Load Balancer

To start off, let's setup a Network Load Balancer. Like the typical procedure, choose the default LB protocol TCP 80, VPC, and create new target group. Review the details and create.



Step 1: Configure Load Balancer

Basic Configuration

To configure your load balancer, provide a name, select a scheme, specify one or more listeners, and select a network. The default configuration is an Internet-facing load balancer in the selected network with a listener that receives TCP traffic on port 80

Name [redacted]

Scheme ● internet-facing ○ internal

Listeners

A listener is a process that checks for connection requests, using the protocol and port that you configured.

Load Balancer Protocol	Load Balancer Port
TCP	80

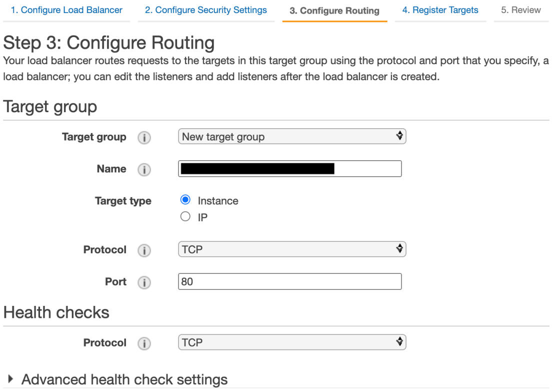
Availability Zones

Specify the Availability Zones to enable for your load balancer. The load balancer routes traffic to the targets in these Availability Zones only. You can specify only one subnet per Availability Zone. You may also add one Elastic IP per Availability Zone if yo addresses for your load balancer.

Create and manage Elastic IPs in the VPC console

VPC [redacted]

Availability Zones ☑ ap-southeast-1a [redacted]
IPv4 address Assigned by AWS
☑ ap- [redacted]



Step 3: Configure Routing

Your load balancer routes requests to the targets in this target group using the protocol and port that you specify, a load balancer; you can edit the listeners and add listeners after the load balancer is created.

Target group

Target group New target group

Name [redacted]

Target type ● Instance ○ IP

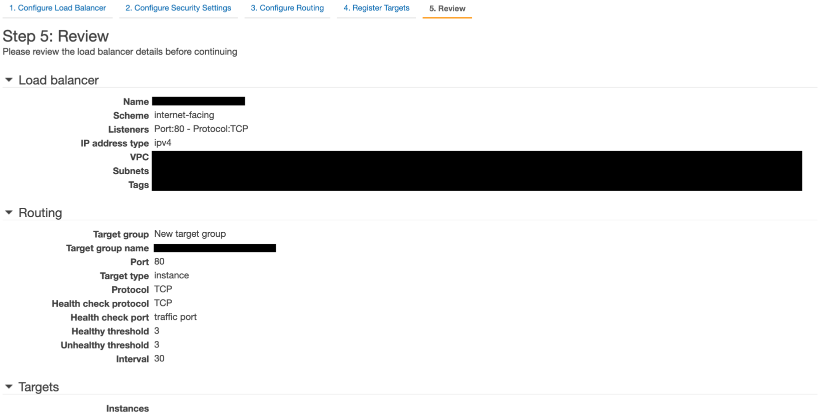
Protocol TCP

Port 80

Health checks

Protocol TCP

Advanced health check settings


1. Configure Load Balancer
2. Configure Security Settings
3. Configure Routing
4. Register Targets
5. Review

Step 5: Review

Please review the load balancer details before continuing

Load balancer

Name	[redacted]
Scheme	internet-facing
Listeners	Port:80 - Protocol:TCP
IP address type	ipv4
VPC	[redacted]
Subnets	
Tags	

Routing

Target group	New target group
Target group name	[redacted]
Port	80
Target type	instance
Protocol	TCP
Health check protocol	TCP
Health check port	traffic port
Healthy threshold	3
Unhealthy threshold	3
Interval	30

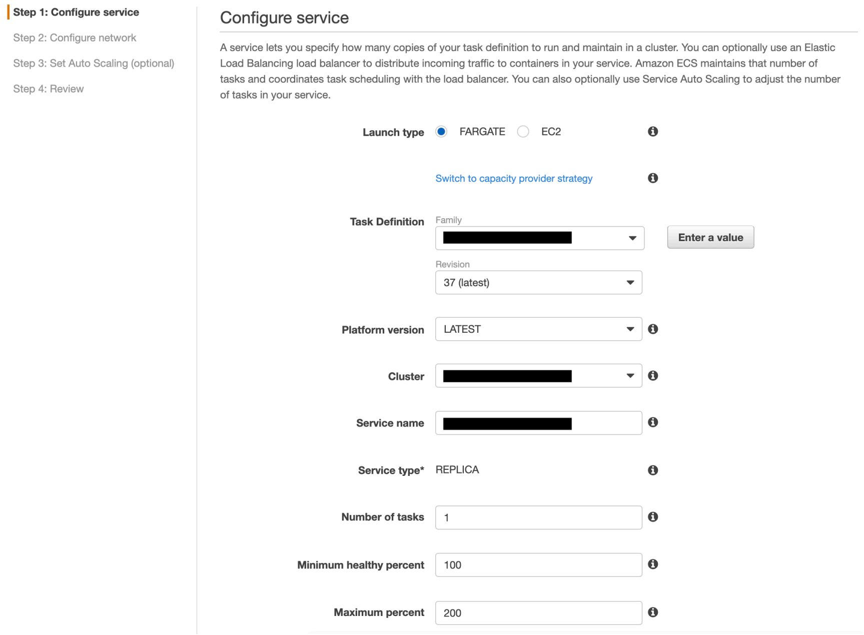
Targets

Instances

3.2 ECS Service

For our setup, we will be using an existing Task Definition for the new ECS service. If you have running ECS services, a new ECS service is still required because the load balancer can only be specified during the initial configuration phase of an ECS service.

In our setup, we use Fargate to create the ECS service with the latest Task Definition version



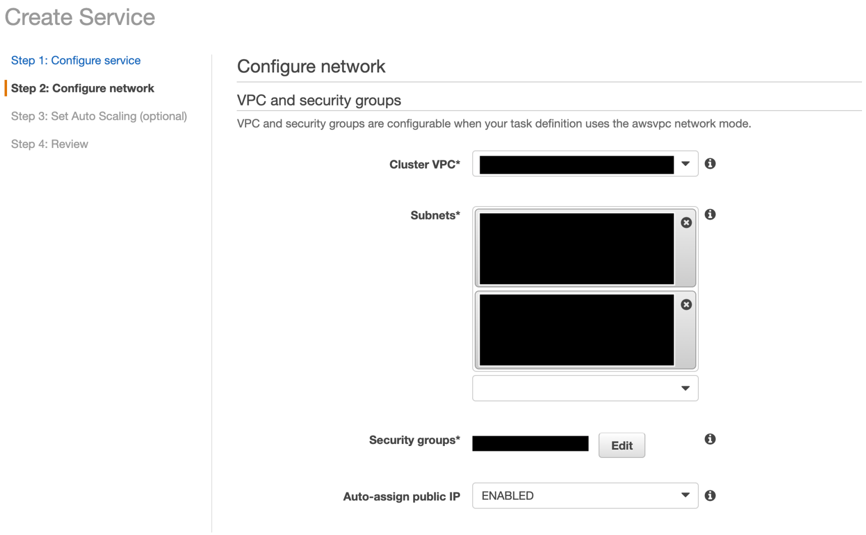
Step 1: Configure service
Step 2: Configure network
Step 3: Set Auto Scaling (optional)
Step 4: Review

Configure service

A service lets you specify how many copies of your task definition to run and maintain in a cluster. You can optionally use an Elastic Load Balancing load balancer to distribute incoming traffic to containers in your service. Amazon ECS maintains that number of tasks and coordinates task scheduling with the load balancer. You can also optionally use Service Auto Scaling to adjust the number of tasks in your service.

Launch type	● FARGATE ○ EC2
	Switch to capacity provider strategy
Task Definition	Family [redacted] Enter a value Revision 37 (latest)
Platform version	LATEST
Cluster	[redacted]
Service name	[redacted]
Service type*	REPLICA
Number of tasks	1
Minimum healthy percent	100
Maximum percent	200

Select VPC and public subnets.



Create Service

Step 1: Configure service
Step 2: Configure network
Step 3: Set Auto Scaling (optional)
Step 4: Review

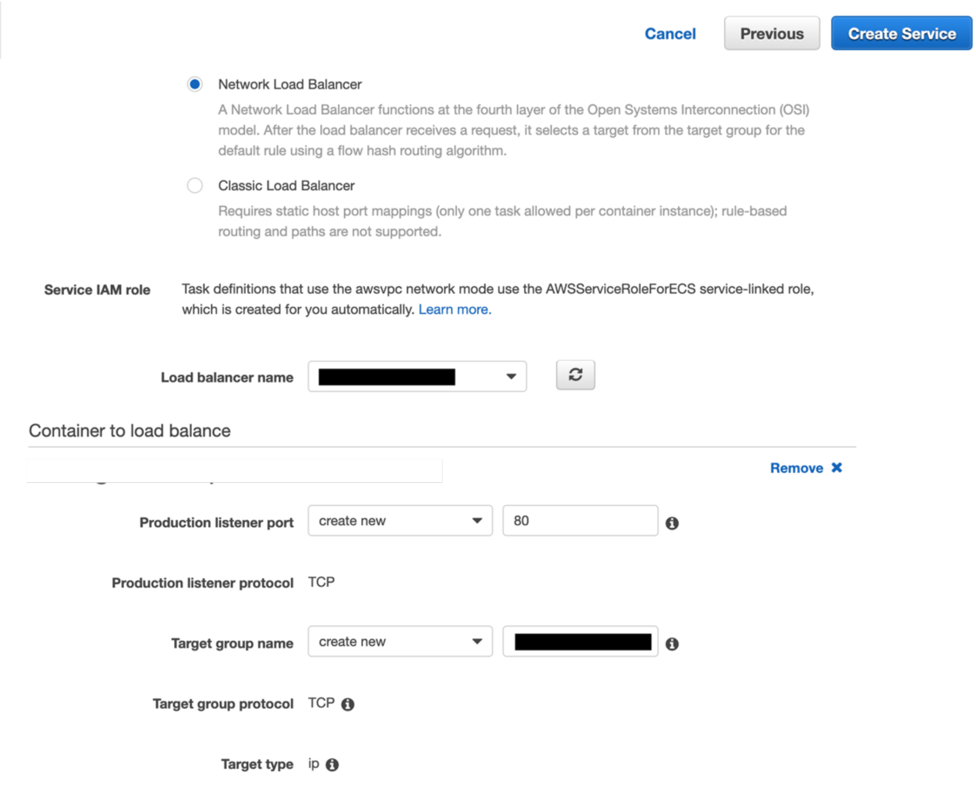
Configure network

VPC and security groups

VPC and security groups are configurable when your task definition uses the awsvpc network mode.

Cluster VPC*	[redacted]
Subnets*	[redacted] [redacted]
Security groups*	[redacted] Edit
Auto-assign public IP	ENABLED

Select the created Network Load Balancer and new target group.



● Network Load Balancer

A Network Load Balancer functions at the fourth layer of the Open Systems Interconnection (OSI) model. After the load balancer receives a request, it selects a target from the target group for the default rule using a flow hash routing algorithm.

○ Classic Load Balancer

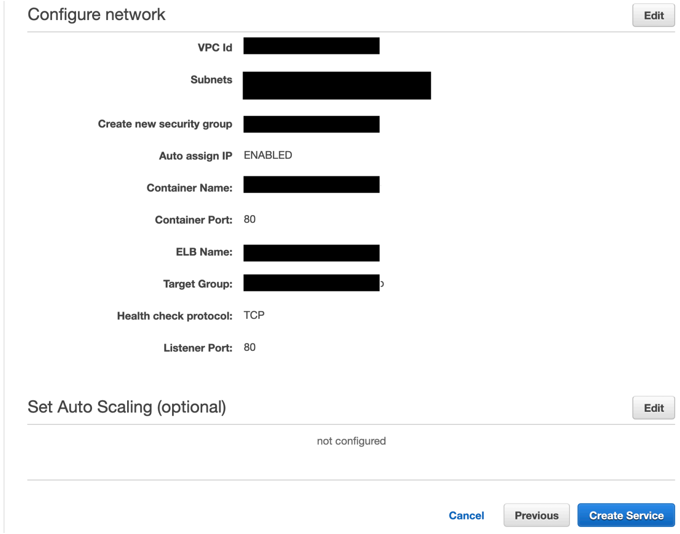
Requires static host port mappings (only one task allowed per container instance); rule-based routing and paths are not supported.

Service IAM role Task definitions that use the awsvpc network mode use the AWSServiceRoleForECS service-linked role, which is created for you automatically. Learn more.

Load balancer name [redacted]

Container to load balance

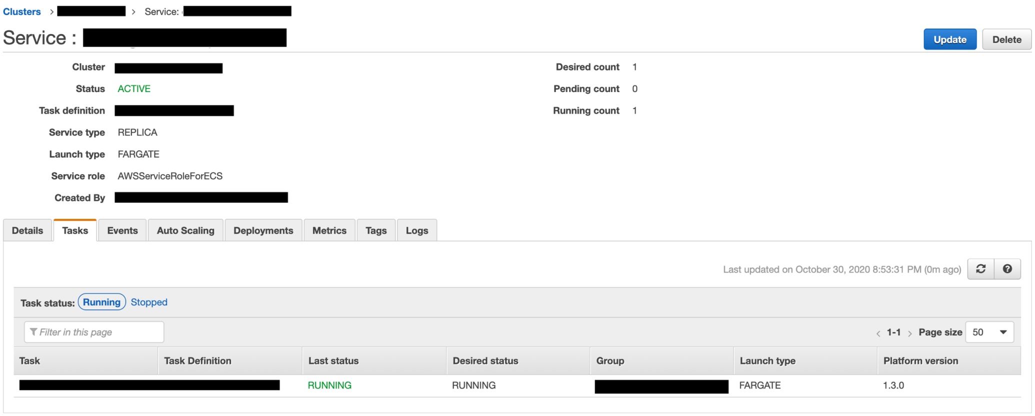
Production listener port	create new 80
Production listener protocol	TCP
Target group name	create new [redacted]
Target group protocol	TCP
Target type	ip


Configure network

VPC Id	[redacted]
Subnets	[redacted]
Create new security group	[redacted]
Auto assign IP	ENABLED
Container Name:	[redacted]
Container Port:	80
ELB Name:	[redacted]
Target Group:	[redacted]
Health check protocol:	TCP
Listener Port:	80

Set Auto Scaling (optional)

not configured


Clusters › [redacted] › Service: [redacted]

Service : [redacted]

Cluster	[redacted]	Desired count	1
Status	ACTIVE	Pending count	0
Task definition	[redacted]	Running count	1
Service type	REPLICA		
Launch type	FARGATE		
Service role	AWSServiceRoleForECS		
Created By	[redacted]		

Last updated on October 30, 2020 8:53:31 PM (0m ago)

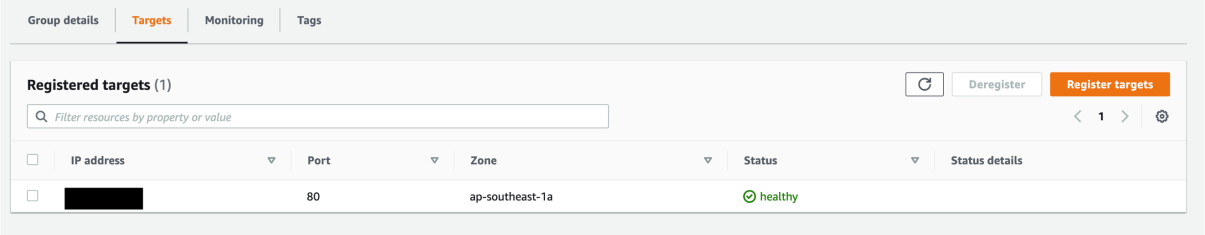
Task status: Running Stopped

Filter in this page

1-1 Page size 50

Task	Task Definition	Last status	Desired status	Group	Launch type	Platform version
[redacted]		RUNNING	RUNNING	[redacted]	FARGATE	1.3.0

After creating the ECS service, check the Target Group status, if the status is healthy, we can continue to create the VPC Link to connect with NLB.

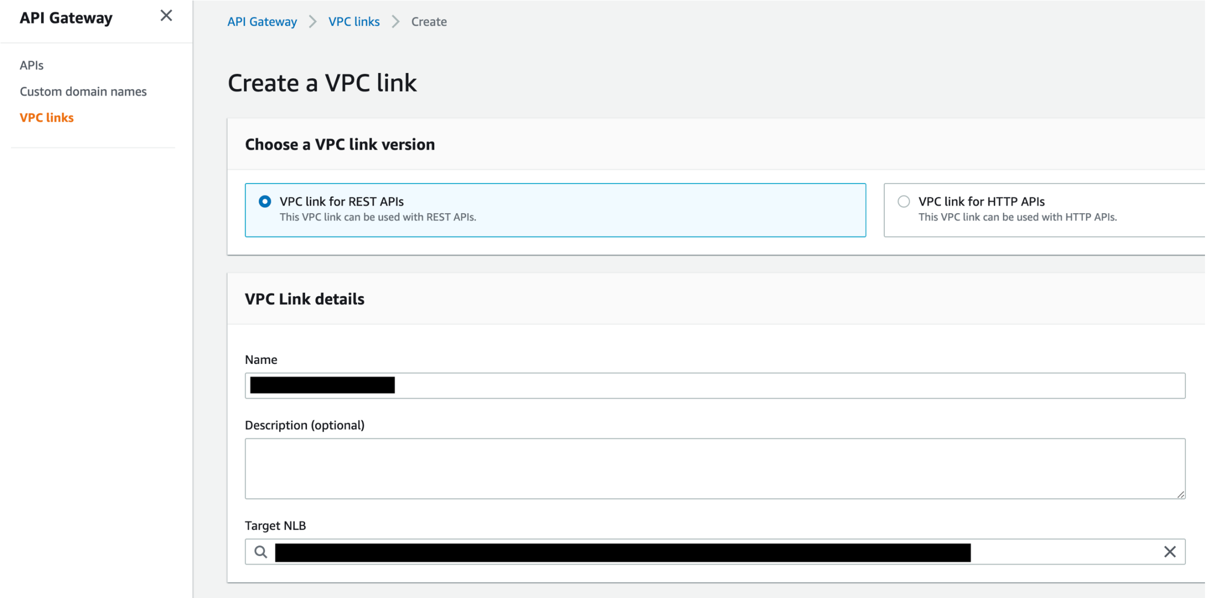


Registered targets (1)

□	IP address	Port	Zone	Status	Status details
□	[redacted]	80	ap-southeast-1a	✓ healthy	

3.3 VPC Link

A VPC Link provides private connectivity between API gateway and the Network Load Balancer. Create the VPC Link from API Gateway console. Fill in the name and choose the newly created NLB from the dropdown list 'Target NLB'.



Create a VPC link

Choose a VPC link version

● VPC link for REST APIs
This VPC link can be used with REST APIs.

○ VPC link for HTTP APIs
This VPC link can be used with HTTP APIs.

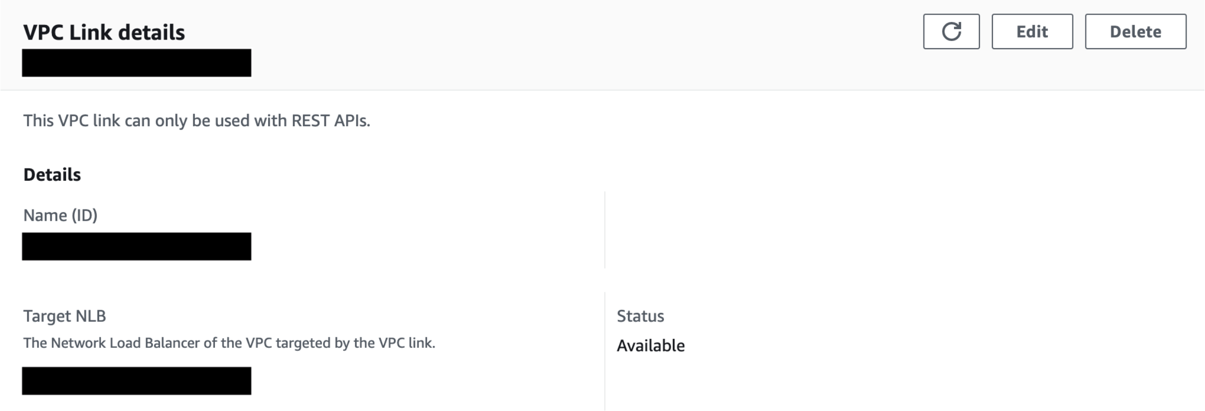
VPC Link details

Name
[redacted]

Description (optional)

Target NLB
[redacted]

It will take a few minutes for the VPC Link to be ready to use, its status will transition from Pending to Available.



VPC Link details
[redacted]

This VPC link can only be used with REST APIs.

Details

Name (ID)
[redacted]

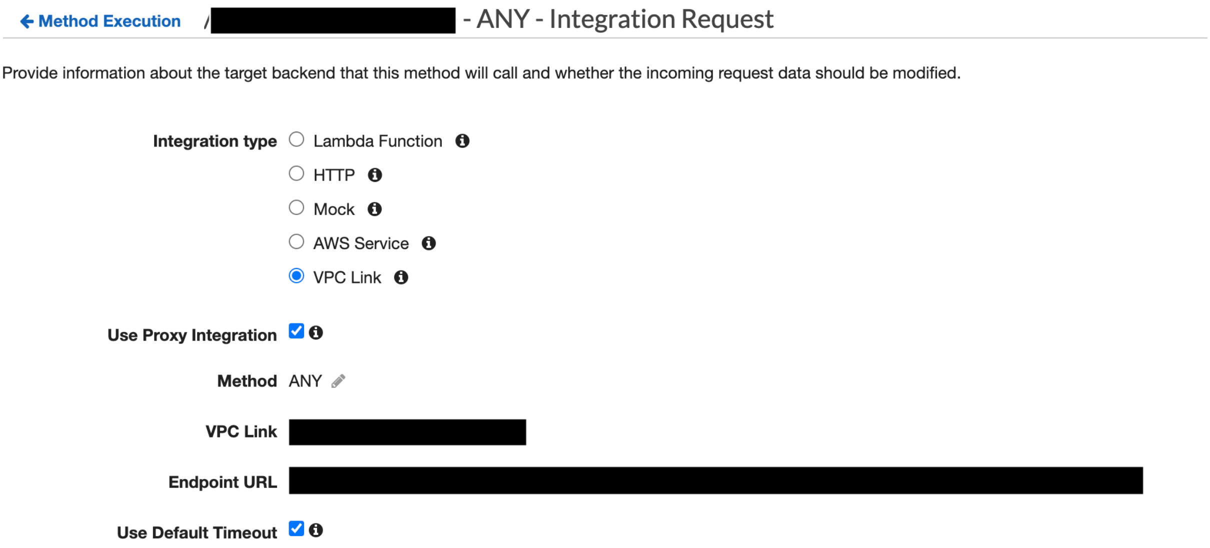
Target NLB
The Network Load Balancer of the VPC targeted by the VPC link.
[redacted]

Status
Available

3.4 API Gateway

Finally, since in our setup we already have API Gateway setup and running, we will simply update the integration type from HTTP to VPC Link. If you have not setup API Gateway, the VPC Link option will show up during the creation process.

API Gateway has a proxy configuration which will forward all request to our Network Load Balancer directly through VPC Link.



← Method Execution /[redacted] - ANY - Integration Request

Provide information about the target backend that this method will call and whether the incoming request data should be modified.

Integration type ○ Lambda Function ○ HTTP ○ Mock ○ AWS Service ● VPC Link

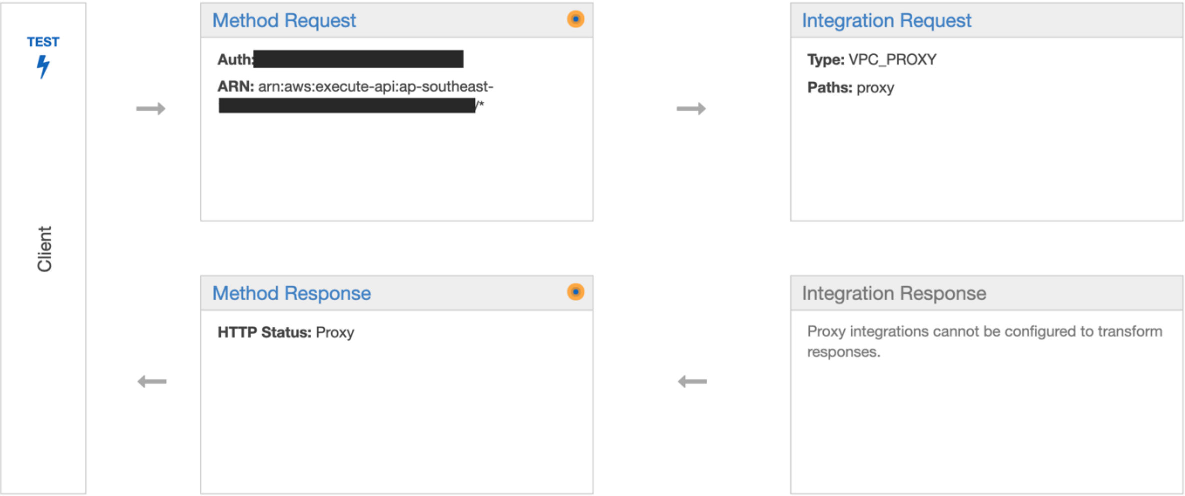
Use Proxy Integration ☑

Method ANY

VPC Link [redacted]

Endpoint URL [redacted]

Use Default Timeout ☑



Method Request

Auth:[redacted]

ARN: arn:aws:execute-api:ap-southeast-[redacted]/*

Integration Request

Type: VPC_PROXY

Paths: proxy

Method Response

HTTP Status: Proxy

Integration Response

Proxy integrations cannot be configured to transform responses.

Once everything is configured correctly, deploy the API on the API Gateway dashboard. You should now be able to make endpoint request and have them directly forward to ECS service which is behind the Network Load Balancer.