Sustentación: https://youtu.be/LNMfcF4AwGQ

Despliegue en Elastic Beanstalk (EB)

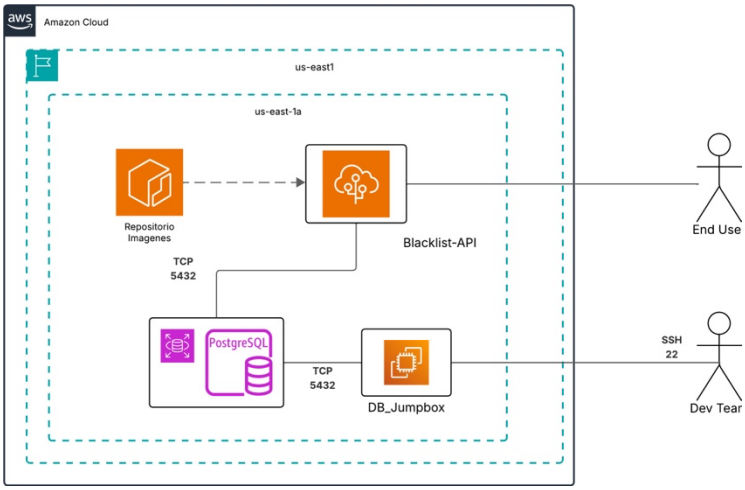
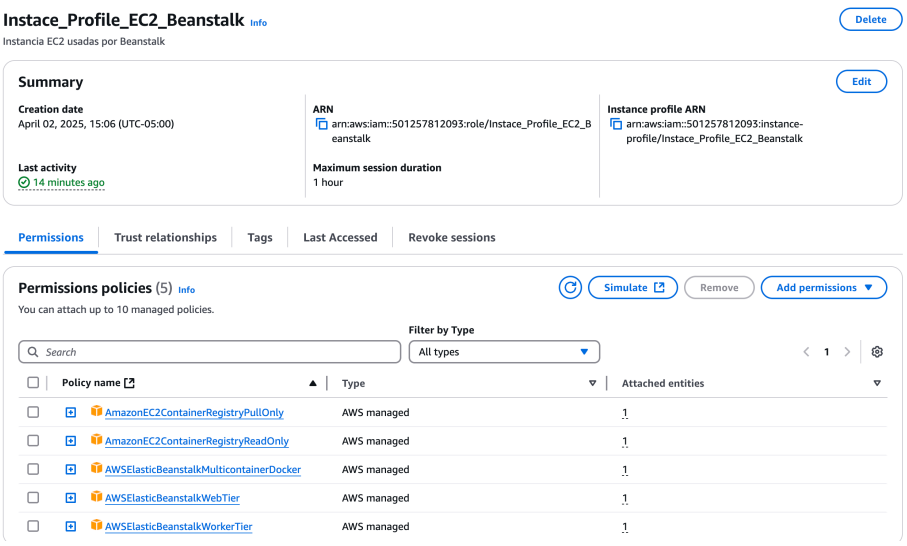


Ilustración 1. Diagrama de Contexto

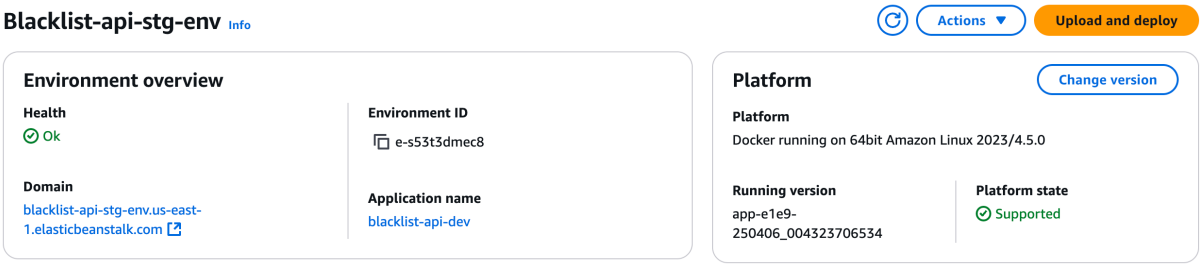
Para el despliegue de la aplicación en AWS Elastic Beanstalk (EB), el equipo optó por utilizar una estrategia basada en contenedores, debido a las ventajas que ofrece en términos de escalabilidad y mantenimiento futuro. Como parte de este proceso, se realizó el aprovisionamiento de un repositorio en Amazon Elastic Container Registry (ECR), en donde se almacenaron las imágenes Docker de la aplicación.

Posteriormente, se configuró un rol en AWS Identity and Access Management (IAM) con los permisos necesarios para permitir que EB pudiera acceder al repositorio, leer las imágenes y realizar operaciones de pull.

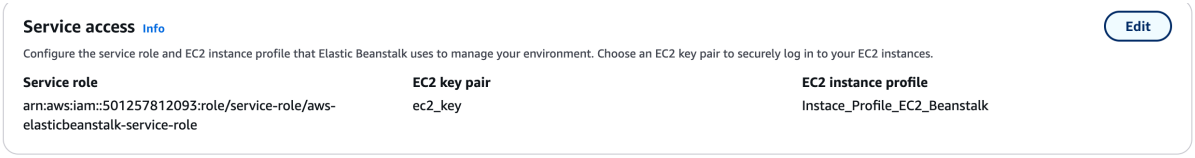


Una vez creado este rol, procedimos con la creación de la aplicación de EB, con la siguiente configuración:

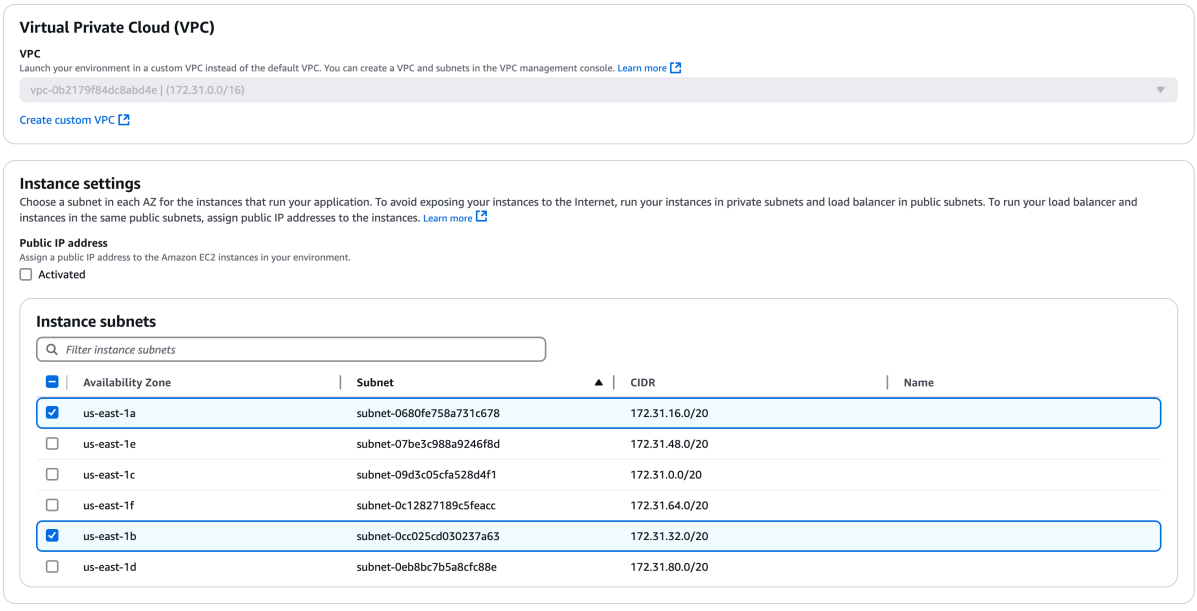
- Nombre de la aplicación: blacklist-api-dev
- Dominio: blacklist-api-stg-env
- Plataforma: Manejada, Docker corriendo en 64bit Amazon Linux



En la sección de “Service Access”, se agregó el rol de IAM que se creó anteriormente, y además creamos un EC2 Key Pair para poder entrar a las instancias EC2 creadas por EB para propósitos de debugging

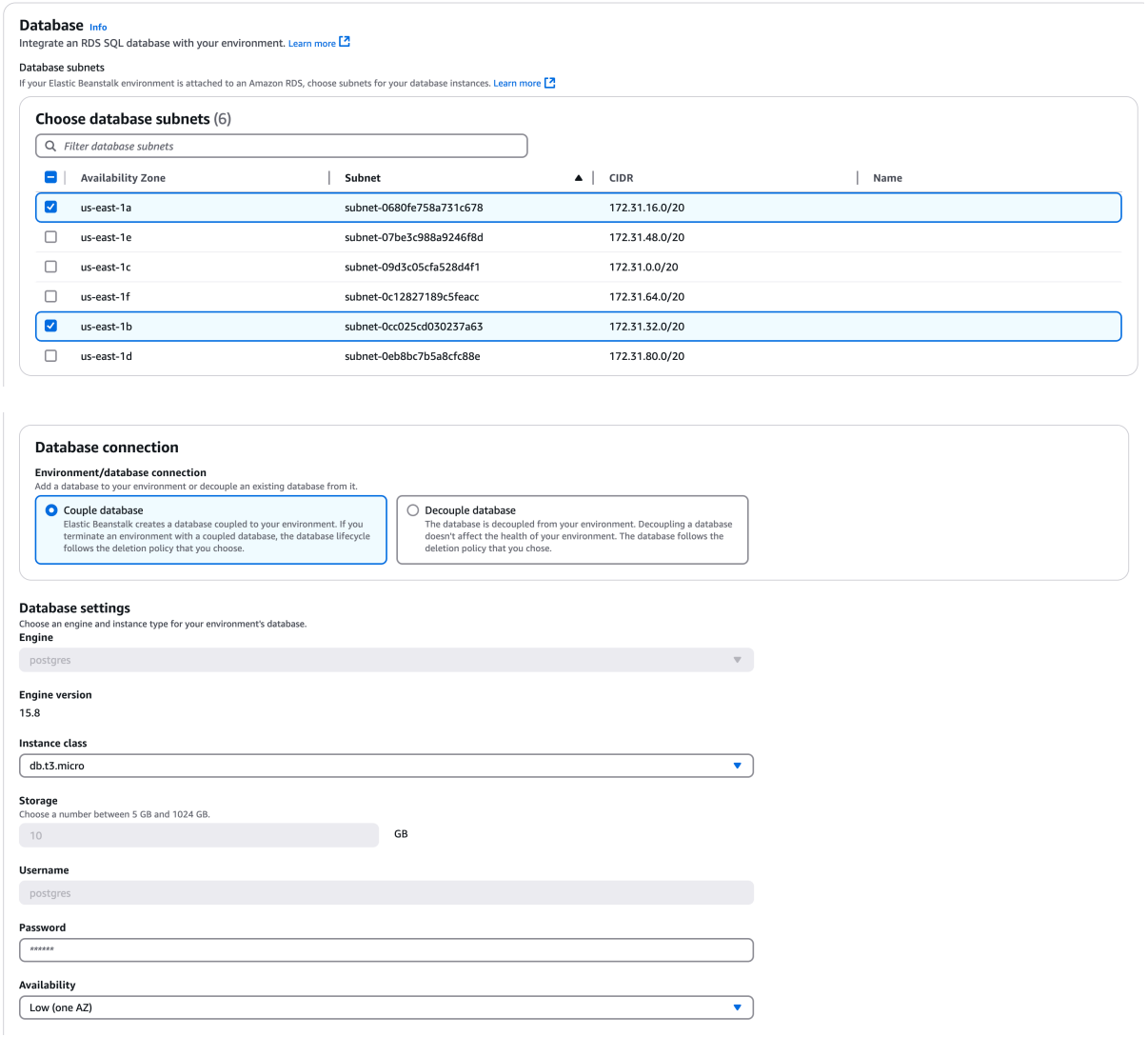


Para la configuración de red, se hizo uso de la VPC que se tiene por defecto, y se restringió las AZ a dos, us-east-1a y us-east1b.



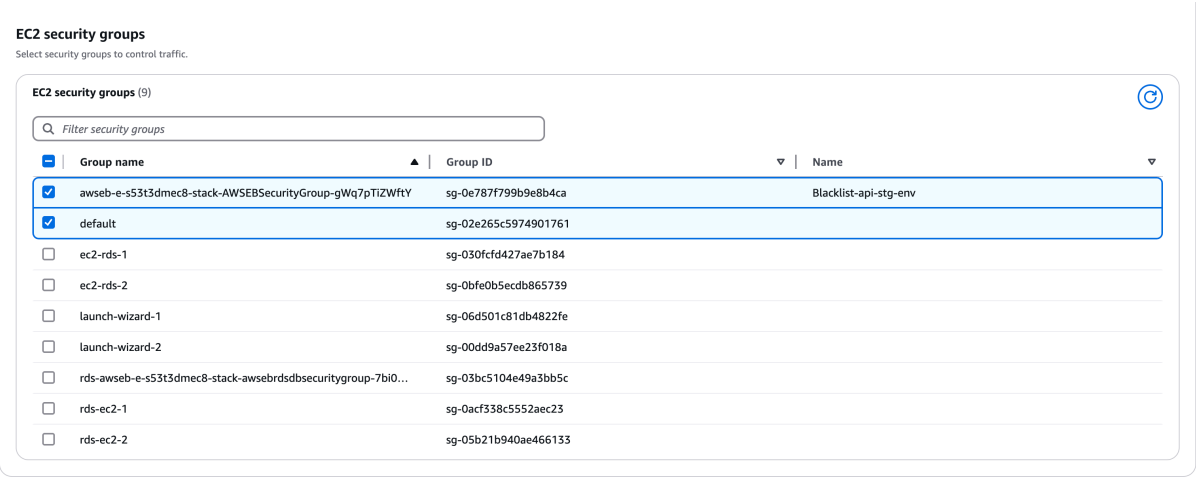
La BD fue creada directamente desde EB, configurando su disponibilidad en múltiples AZ y estableciendo su conexión a la aplicación. Para evitar costos adicionales, se

verificaron previamente los tipos de instancias incluidas en el Free Tier de Amazon RDS, seleccionando la opción más adecuada. Se optó por delegar la creación de la BD a EB debido a que este servicio genera automáticamente los grupos de seguridad necesarios para RDS y EC2, aplicándolos directamente sin intervención manual.

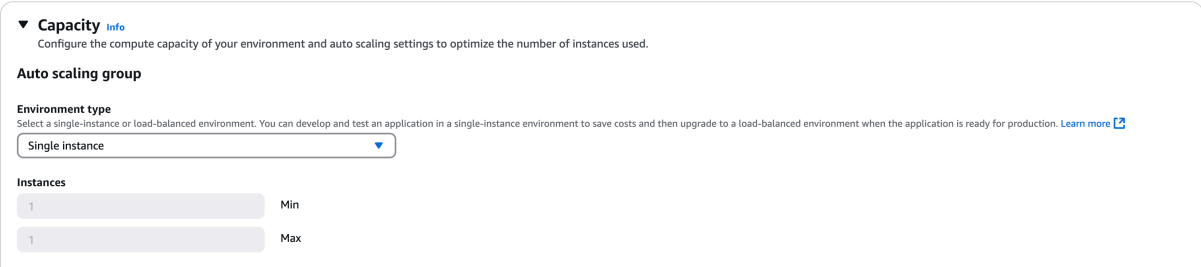


Por la parte de grupos de seguridad, al momento de la creación solo se seleccionó el “default”, pero como se menciona anteriormente, EB crea y añade uno para garantizar

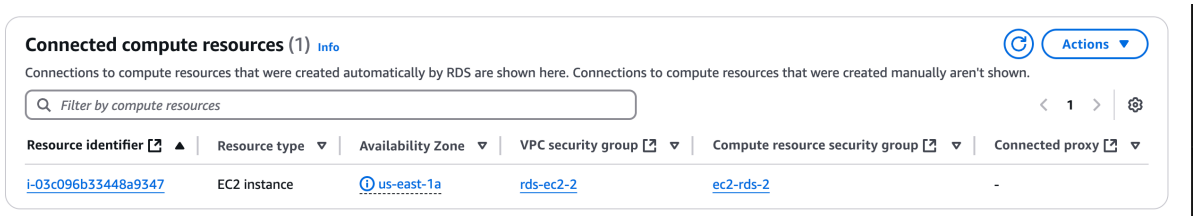
la conexión a la BD:



En la sección de auto-escalamiento, se define primero como un ambiente de tipo de instancia única. Esto para primero garantizar el correcto funcionamiento de la aplicación antes de realizar las comparativas del modelo de despliegue.

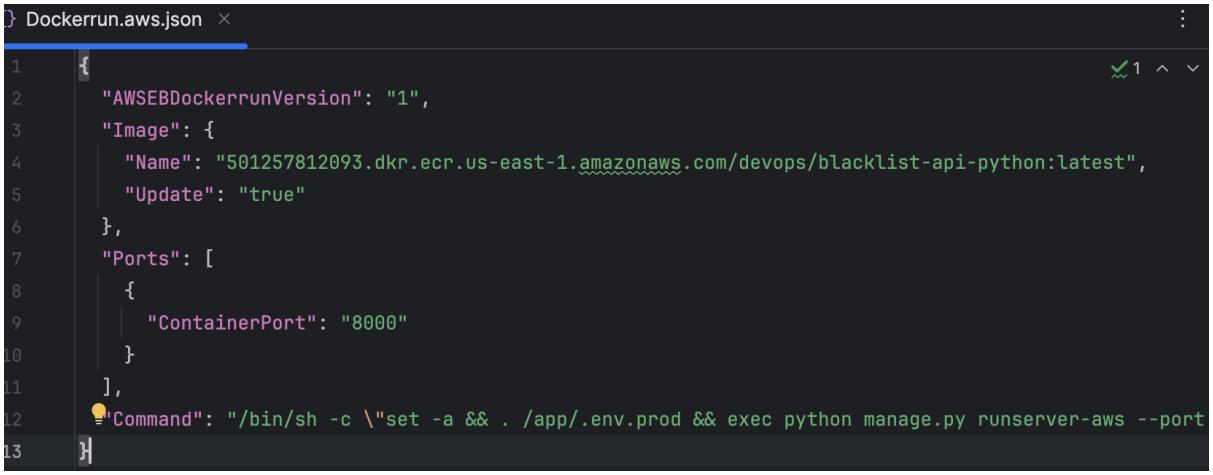


Finalmente, se creó una instancia EC2 que tuviera conexión a la BD creada. Esta la denominamos “Jumpbox”, y nos sirve de intermediario a la BD, y nos permite realizar consultas y conectarnos mediante SSH a esta ya que la BD está configurada para no ser accesible públicamente. Esta conexión se realizó mediante las herramientas de conexión RDS:



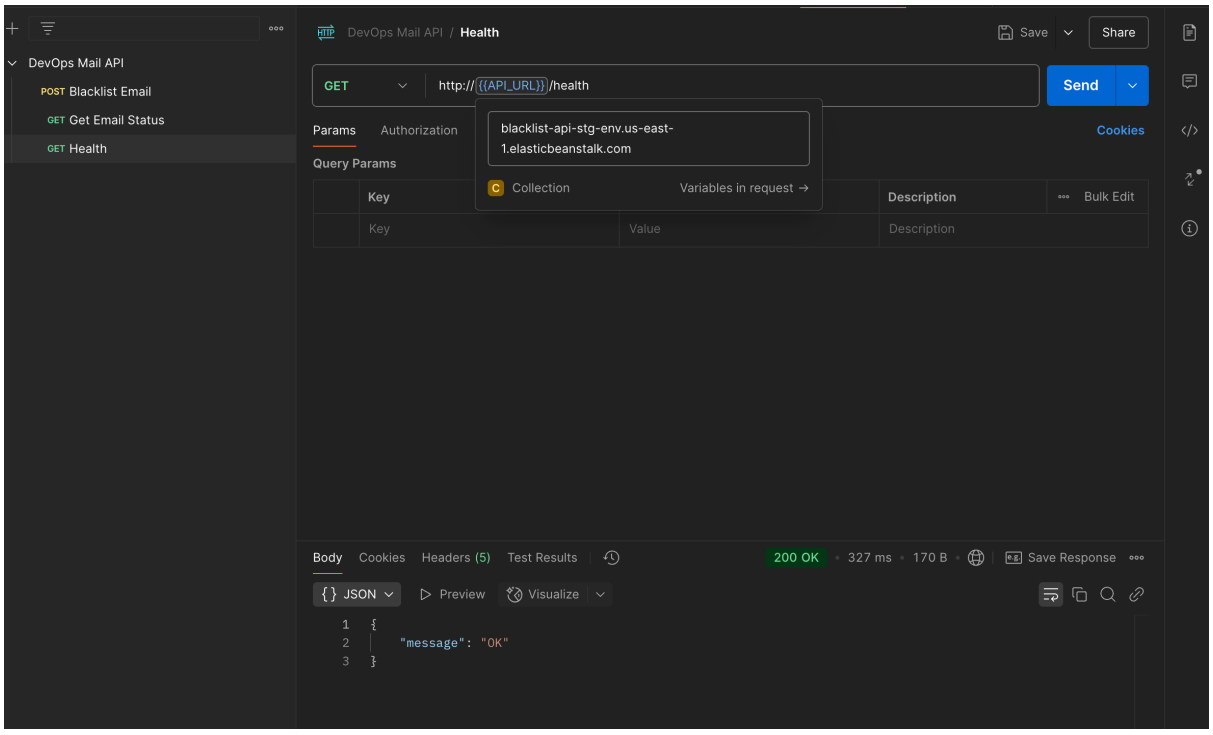
Finalmente, cuando el ambiente despliega correctamente la aplicación de prueba, creamos el archivo `Dockerrun.aws.json`, donde especificamos de que imagen debe realizar el pull al ECR, que puerto va a quedar expuesto y finalmente cual va a ser el entrypoint para la imagen de Docker.

Entrega 1



Seguido a esto, se ejecuta el comando “eb init”, se enlaza el ambiente de EB al directorio local, y procedemos con un “eb deploy”. Esto comprime los archivos necesarios en un .zip (En este caso, el Dockerrun.aws.json y el .env.prod), los envia a un bucket de S3, que después es descargado por la instancia de EC2 manejada por EB y ejecuta las instrucciones especificadas en el Dockerrun.aws.json.

Para validar el correcto funcionamiento de la aplicación, se cuenta con un endpoint de health que nos indica que la app está en funcionamiento:



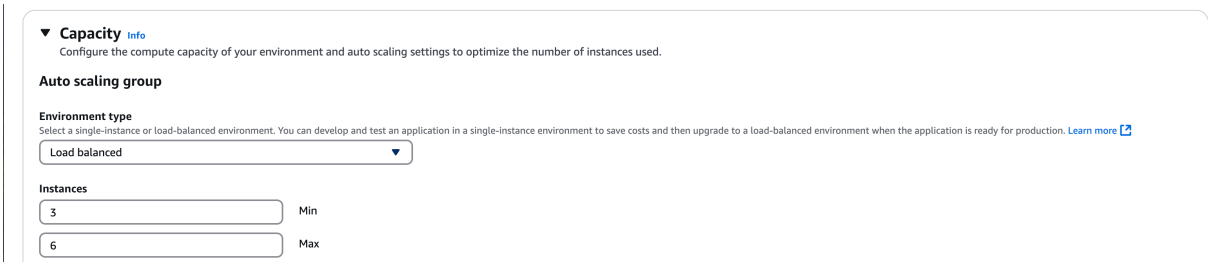
Entrega 1

Estrategias de Despliegue:

Configuración de health-checks y ambiente:

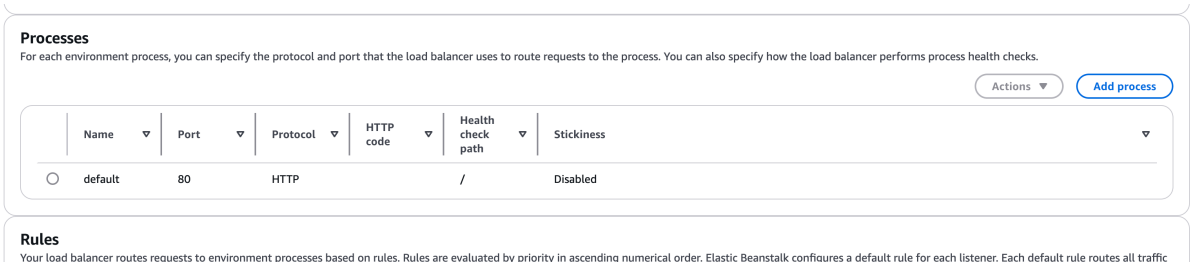
Con el objetivo de realizar una comparación adecuada, se establecieron las siguientes condiciones del entorno:

1. Todos los despliegues van a contar con 3 instancias



2. Para validar la corrección del despliegue, se verificará que haya un cambio en la aplicación y que ésta continúe cumpliendo con todos sus requerimientos funcionales, y que EB reporte un estado de salud OK.

3. Antes de comenzar cualquier prueba, se garantizará que la aplicación esté funcionando correctamente con la configuración predeterminada de un entorno con balanceo de carga ("Load Balanced"). Para esto fue necesario definir ciertas configuraciones para los health-checks. Nuestra API cuenta con un endpoint en la raíz que también sirve como health-check, coincidiendo así con la configuración predeterminada de Elastic Beanstalk (EB).

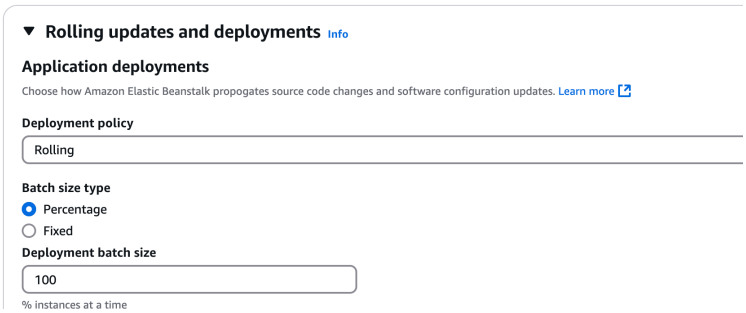


4. Para medir los tiempos de despliegue, primero se modificará la configuración y posteriormente se creará una nueva versión de la aplicación. En este caso particular, solo se cambió la variable APP_VERSION a 0.1.0 dentro del archivo settings.py.

Rolling:

Es una política de despliegue que intenta evitar en downtime y minimizar la disponibilidad reducida a costo de un tiempo de despliegue mayor. Este despliegue no

requiere que se eliminen las instancias anteriores y se generen instancias nuevas. La configuración que se dejó para la prueba es la siguiente:



Con esta configuración, el tiempo total para desplegar la aplicación y que los health-checks resultaran exitosos fue aproximadamente de 2 minutos y 50 segundos.

Time	Type	Details
April 6, 2025 17:03:17 (UTC-5)	INFO	Environment health has transitioned from Info to Ok. Application update completed 32 seconds ago and took 36 seconds.
April 6, 2025 17:02:17 (UTC-5)	INFO	Environment health has transitioned from Ok to Info. Application update in progress on 3 instances. 0 out of 3 instances completed (running for 29 seconds).
April 6, 2025 17:02:03 (UTC-5)	INFO	Environment update completed successfully.
April 6, 2025 17:02:03 (UTC-5)	INFO	New application version was deployed to running EC2 instances.
April 6, 2025 17:01:58 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:01:28 (UTC-5)	INFO	Deploying new version to instance(s).

Para mejorar la validación de resultados, se realizó adicionalmente un despliegue con un batch size del 50%, con el objetivo de observar las variaciones. Este despliegue tomó aproximadamente 4 minutos y 10 segundos.

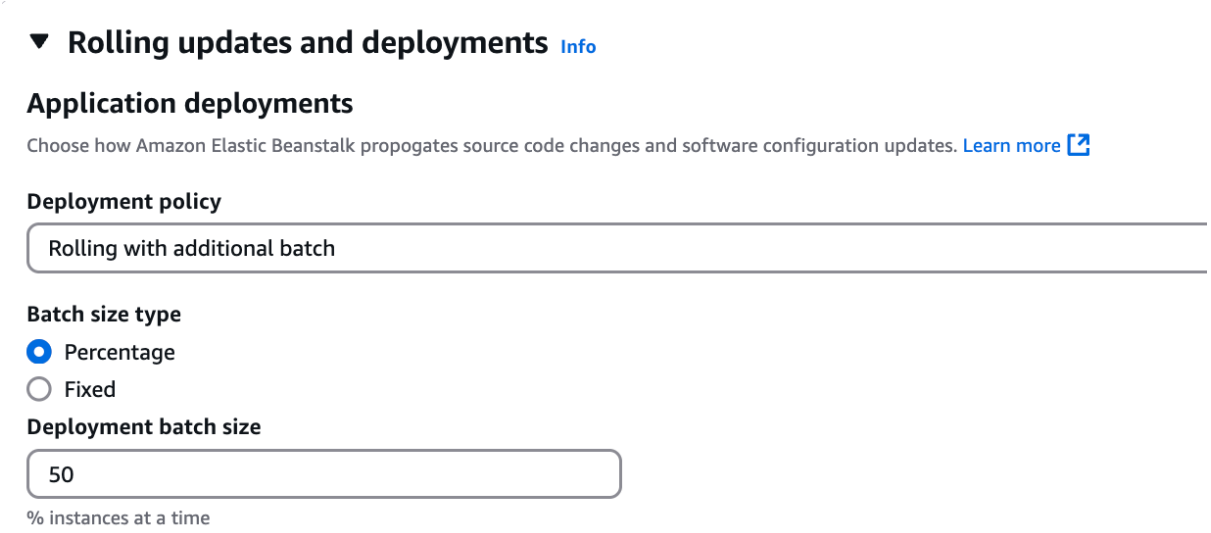
Time	Type	Details
April 6, 2025 17:18:16 (UTC-5)	INFO	Environment health has transitioned from Info to Ok. Application update completed 57 seconds ago and took 3 minutes.
April 6, 2025 17:16:50 (UTC-5)	INFO	Environment update completed successfully.
April 6, 2025 17:16:50 (UTC-5)	INFO	New application version was deployed to running EC2 instances.
April 6, 2025 17:16:47 (UTC-5)	INFO	Batch 3: Completed application deployment.
April 6, 2025 17:16:42 (UTC-5)	INFO	Batch 3: Registering instance(s) with the load balancer and waiting for them to be healthy.
April 6, 2025 17:16:42 (UTC-5)	INFO	Command execution completed on 3 of 3 instances in environment.
April 6, 2025 17:16:41 (UTC-5)	INFO	Batch 3: Completed application deployment command execution.
April 6, 2025 17:16:38 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:16:15 (UTC-5)	INFO	Batch 3: Starting application deployment command execution.
April 6, 2025 17:15:53 (UTC-5)	INFO	Batch 3: Starting application deployment on instance(s) [i-058f522a7a52f722b].
April 6, 2025 17:15:53 (UTC-5)	INFO	Batch 2: Completed application deployment.
April 6, 2025 17:15:48 (UTC-5)	INFO	Batch 2: Registering instance(s) with the load balancer and waiting for them to be healthy.
April 6, 2025 17:15:47 (UTC-5)	INFO	Command execution completed on 2 of 3 instances in environment.
April 6, 2025 17:15:47 (UTC-5)	INFO	Batch 2: Completed application deployment command execution.
April 6, 2025 17:15:43 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:15:21 (UTC-5)	INFO	Batch 2: Starting application deployment command execution.
April 6, 2025 17:14:58 (UTC-5)	INFO	Batch 2: Starting application deployment on instance(s) [i-0c69964b0fd0a257b].
April 6, 2025 17:14:58 (UTC-5)	INFO	Batch 1: Completed application deployment.
April 6, 2025 17:14:53 (UTC-5)	INFO	Batch 1: Registering instance(s) with the load balancer and waiting for them to be healthy.
April 6, 2025 17:14:52 (UTC-5)	INFO	Command execution completed on 1 of 3 instances in environment.

Time	Type	Details
April 6, 2025 17:14:52 (UTC-5)	INFO	Batch 1: Completed application deployment command execution.
April 6, 2025 17:14:51 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:14:30 (UTC-5)	INFO	Batch 1: Starting application deployment command execution.
April 6, 2025 17:14:16 (UTC-5)	INFO	Environment health has transitioned from Ok to Info. Application update in progress on 1 instance. 0 out of 3 instances completed (running for 3 seconds).
April 6, 2025 17:14:07 (UTC-5)	INFO	Batch 1: Starting application deployment on instance(s) [i-0a2f6d99640134e62].
April 6, 2025 17:14:06 (UTC-5)	INFO	Deploying new version to instance(s).

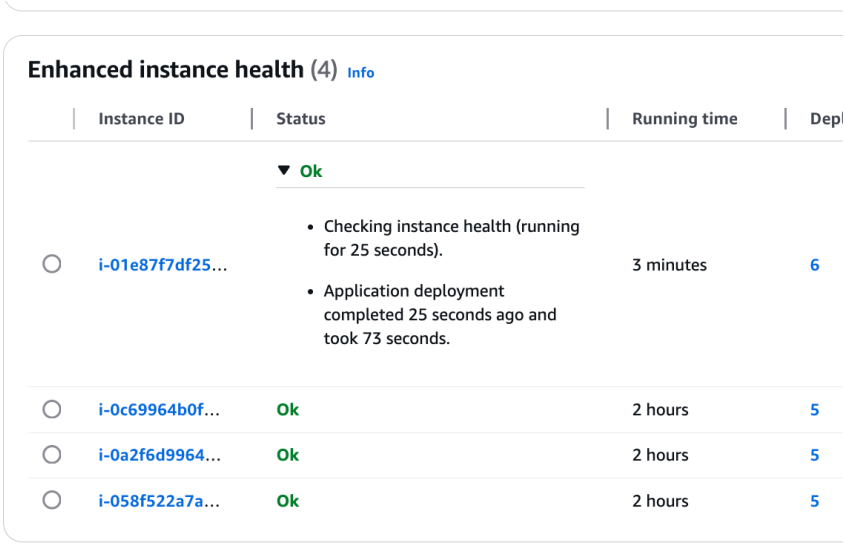
Estos resultados evidencian que un despliegue con batch size del 100% es muy similar a un despliegue del tipo "All at once", y que el tiempo de despliegue es inversamente proporcional al tamaño del batch; cuanto menor sea el tamaño del batch, mayor será el número de particiones y estas serán desplegadas secuencialmente.

Rolling with additional batch

Este método sigue la misma metodología que un despliegue Rolling, pero además despliega un batch adicional de instancias para evitar completamente la reducción de disponibilidad. Para estas pruebas se configuró de la siguiente manera:



Durante el despliegue se verifico efectivamente que se tiene un numero de instancias adicionales equivalentes al tamaño del batch



En términos de tiempo de despliegue, se tiene un aproximado de 7min y 12 segundos. El aumento lo podemos ver gracias a que durante el despliegue se tienen que hacer actividades como el manejo del ciclo de vida de las instancias adicionales, además de

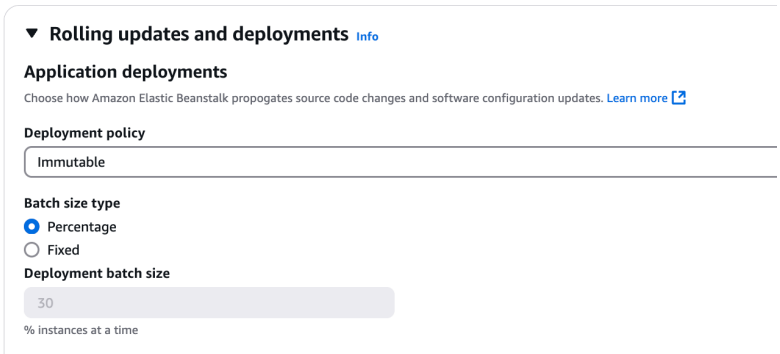
que solo se concluyó el despliegue hasta que el warning que se ven en los eventos desapareció y el ambiente paso de estado degradado a OK.

April 6, 2025 17:42:13 (UTC-5)	INFO	Removed instance [i-0c69964b0fd0a257b] from your environment.
April 6, 2025 17:41:14 (UTC-5)	WARN	Environment health has transitioned from Info to Degraded. Incorrect application version found on 1 out of 4 instances. Expected version "app-7ae2-250406_173459615916" (deployment 6). ELB processes are not healthy on 1 out of 4 instances. Application update completed 11 seconds ago and took 5 minutes. ELB health is failing or not available for 1 out of 4 instances.
April 6, 2025 17:40:24 (UTC-5)	INFO	Environment update completed successfully.
April 6, 2025 17:40:24 (UTC-5)	INFO	New application version was deployed to running EC2 instances.
April 6, 2025 17:40:22 (UTC-5)	INFO	Excess instance(s) terminated.
April 6, 2025 17:40:19 (UTC-5)	INFO	Terminating excess instance(s): [i-0c69964b0fd0a257b].
April 6, 2025 17:39:31 (UTC-5)	INFO	Batch 3: Completed application deployment.
April 6, 2025 17:39:26 (UTC-5)	INFO	Batch 3: Registering instance(s) with the load balancer and waiting for them to be healthy.
April 6, 2025 17:39:25 (UTC-5)	INFO	Command execution completed on 3 of 4 instances in environment.
April 6, 2025 17:39:25 (UTC-5)	INFO	Batch 3: Completed application deployment command execution.
April 6, 2025 17:39:25 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:39:03 (UTC-5)	INFO	Batch 3: Starting application deployment command execution.
April 6, 2025 17:38:41 (UTC-5)	INFO	Batch 3: Starting application deployment on instance(s) [i-058f522a7a52f722b].
April 6, 2025 17:38:40 (UTC-5)	INFO	Batch 2: Completed application deployment.
April 6, 2025 17:38:35 (UTC-5)	INFO	Batch 2: Registering instance(s) with the load balancer and waiting for them to be healthy.
April 6, 2025 17:38:35 (UTC-5)	INFO	Command execution completed on 2 of 4 instances in environment.
April 6, 2025 17:38:35 (UTC-5)	INFO	Batch 2: Completed application deployment command execution.
April 6, 2025 17:38:33 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:38:13 (UTC-5)	INFO	Batch 2: Starting application deployment command execution.
April 6, 2025 17:37:50 (UTC-5)	INFO	Batch 2: Starting application deployment on instance(s) [i-0a2f6d99640134e62].

Time	Type	Details
April 6, 2025 17:37:50 (UTC-5)	INFO	Batch 2: Starting application deployment on instance(s) [i-0a2f6d99640134e62].
April 6, 2025 17:37:49 (UTC-5)	INFO	Deploying new version to instance(s).
April 6, 2025 17:37:46 (UTC-5)	INFO	Command execution completed on 1 of 4 instances in environment.
April 6, 2025 17:37:46 (UTC-5)	INFO	Batch 1: Completed application deployment.
April 6, 2025 17:37:12 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:36:43 (UTC-5)	INFO	Batch 1: 1 EC2 instance(s) [i-01e87f7df254f55f4] launched. Deploying application version 'app-7ae2-250406_173459615916'.
April 6, 2025 17:36:14 (UTC-5)	INFO	Environment health has transitioned from Ok to Info. Application update in progress (running for 50 seconds).
April 6, 2025 17:36:14 (UTC-5)	INFO	Added instance [i-01e87f7df254f55f4] to your environment.
April 6, 2025 17:35:26 (UTC-5)	INFO	Rolling with Additional Batch deployment policy enabled. Launching a new batch of 1 additional instance(s).
April 6, 2025 17:35:01 (UTC-5)	INFO	Environment update is starting.

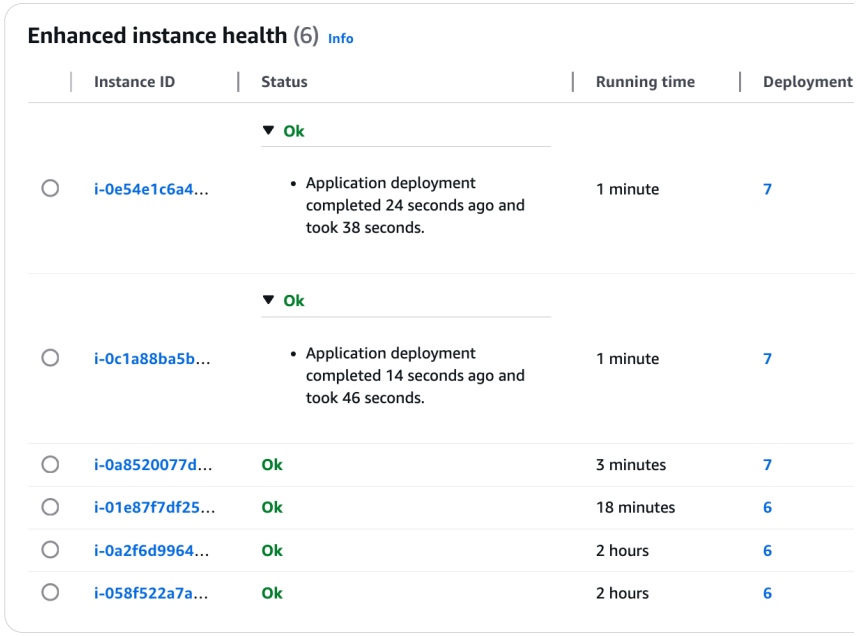
Inmutable

En esta política de despliegue, se requieren crear nuevas instancias. EB va a crear un segundo grupo de instancias con la versión nueva y ambas van a recibir tráfico hasta que todo el segundo grupo de instancias pase los health-checks. La política se configuro con los siguientes parámetros:



Entrega 1

Durante el despliegue se verifico que efectivamente se crea un segundo grupo de instancias. En este escenario se puede ver cómo se tienen 6 instancias EC2 corriendo, 3 del grupo nuevo y 3 del grupo original.



Enhanced instance health (6) Info

Instance ID	Status	Running time	Deployment
	▼ Ok		
i-0e54e1c6a4…	• Application deployment completed 24 seconds ago and took 38 seconds.	1 minute	7
	▼ Ok		
i-0c1a88ba5b…	• Application deployment completed 14 seconds ago and took 46 seconds.	1 minute	7
i-0a8520077d…	Ok	3 minutes	7
i-01e87f7df25…	Ok	18 minutes	6
i-0a2f6d9964…	Ok	2 hours	6
i-058f522a7a…	Ok	2 hours	6

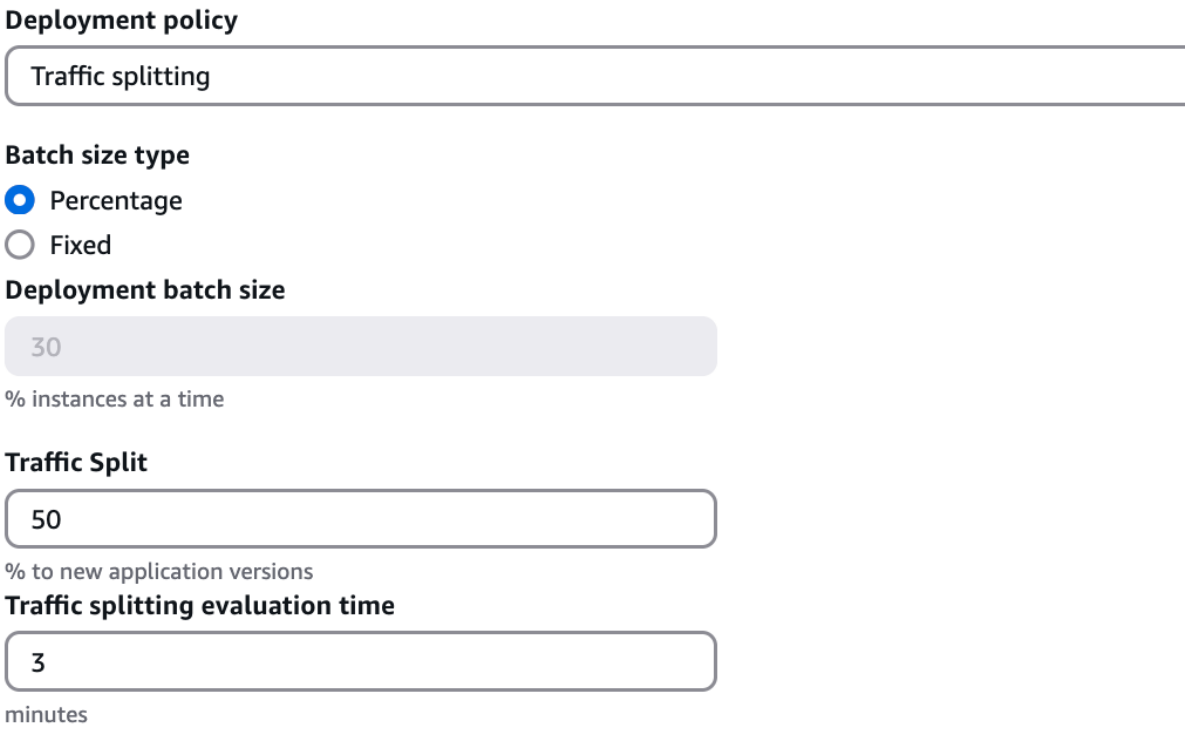
En términos del tiempo de despliegue, se tuvo un aproximado de 18min. El tiempo de este despliegue es totalmente dependiente de la cantidad de instancias que se deban desplegar nuevas y el tiempo que toman estas en configurarse y finalmente eliminar las instancias anteriores.

April 6, 2025 18:07:11 (UTC-5)	INFO	Environment health has transitioned from Info to Ok. Application update completed 46 seconds ago and took 17 minutes.
April 6, 2025 18:06:49 (UTC-5)	INFO	Environment update completed successfully.
April 6, 2025 18:06:49 (UTC-5)	INFO	New application version was deployed to running EC2 instances.
April 6, 2025 18:05:11 (UTC-5)	INFO	Removed instance [i-01e87f7df254f55f4] from your environment.
April 6, 2025 18:02:12 (UTC-5)	INFO	Removed instance [i-058f522a7a52f722b] from your environment.
April 6, 2025 17:59:12 (UTC-5)	INFO	Removed instance [i-0a2f6d99640134e62] from your environment.
April 6, 2025 17:57:20 (UTC-5)	INFO	Deployment succeeded. Terminating old instances and temporary Auto Scaling group.
April 6, 2025 17:55:47 (UTC-5)	INFO	Waiting for post-deployment configuration to complete.
April 6, 2025 17:55:44 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:55:33 (UTC-5)	INFO	Starting post-deployment configuration on new instances.
April 6, 2025 17:55:06 (UTC-5)	INFO	Attached new instance(s) to the permanent auto scaling group awseb-e-7p5kfdrbyb-stack-AWSEBAutoScalingGroup-4H9x6Bo3JezW.
April 6, 2025 17:55:02 (UTC-5)	INFO	Detached new instance(s) from temporary auto scaling group awseb-e-7p5kfdrbyb-immutable-stack-AWSEBAutoScalingGroup-OuYatyOlCocW.
April 6, 2025 17:54:57 (UTC-5)	INFO	Waiting for instance(s) (i-0a8520077da1a8405,i-0c1a88ba5ba9c5f82,i-0e54e1c6a4e264f40) to pass health checks.
April 6, 2025 17:54:56 (UTC-5)	INFO	Adding new instance(s) (i-0c1a88ba5ba9c5f82,i-0e54e1c6a4e264f40) to the load balancer.
April 6, 2025 17:54:55 (UTC-5)	INFO	Successfully launched all instances.
April 6, 2025 17:54:12 (UTC-5)	INFO	Added instances [i-0c1a88ba5ba9c5f82, i-0e54e1c6a4e264f40] to your environment.
April 6, 2025 17:53:13 (UTC-5)	INFO	Environment health has transitioned from Ok to Info. Application update in progress on 1 instance. 0 out of 1 instance completed (running for 3 minutes).
April 6, 2025 17:53:13 (UTC-5)	INFO	Added instance [i-0a8520077da1a8405] to your environment.
April 6, 2025 17:52:54 (UTC-5)	INFO	Test instance passed health checks. Launching remaining new instances.
April 6, 2025 17:52:33 (UTC-5)	INFO	Waiting for instance(s) (i-0a8520077da1a8405) to pass health checks.
April 6, 2025 17:52:33 (UTC-5)	INFO	Adding new instance(s) (i-0a8520077da1a8405) to the load balancer.
April 6, 2025 17:52:28 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 17:50:40 (UTC-5)	INFO	Created temporary auto scaling group awseb-e-7p5kfdrbyb-immutable-stack-AWSEBAutoScalingGroup-OuYatyOlCocW.
April 6, 2025 17:50:24 (UTC-5)	INFO	Immutable deployment policy enabled. Launching one instance with the new settings to verify health.
April 6, 2025 17:49:26 (UTC-5)	INFO	Environment update is starting.

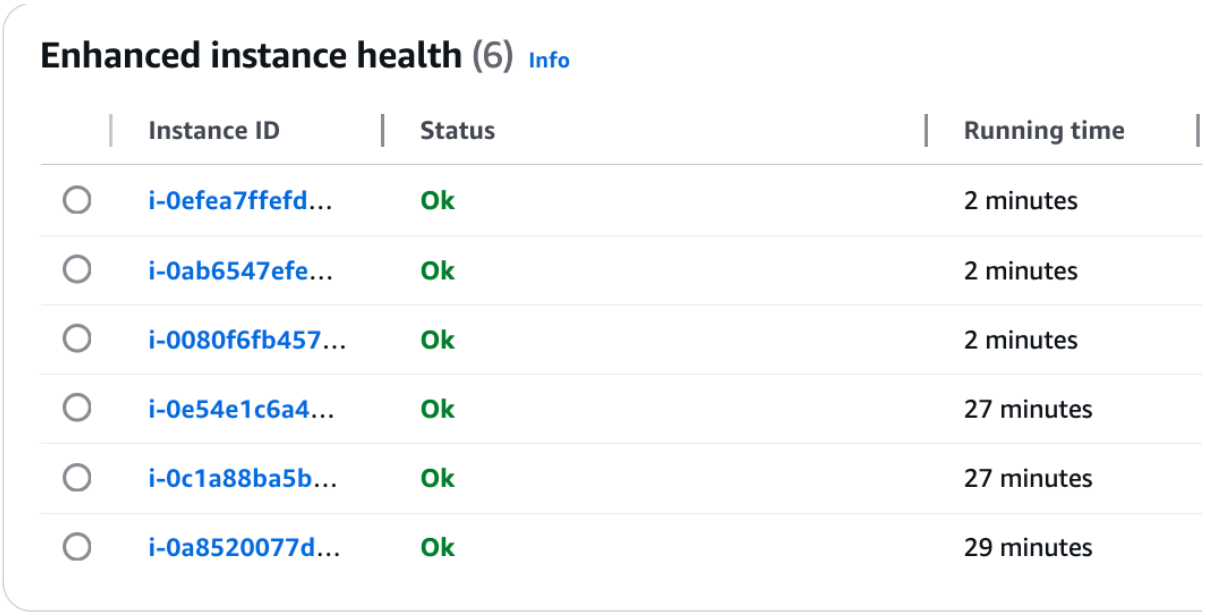
Entrega 1

Traffic Splitting

Esta política de despliegue sigue un canary deployment. Para esta prueba fue configurada de la siguiente manera.



En este despliegue, también se verifico que efectivamente se crea un segundo grupo de instancias que reciben el tráfico nuevo.



Entrega 1

En términos de tiempo de despliegue, se tomó un aproximado de 11min, cabe resaltar que este tiempo pudo haber sido mayor si en la configuración se hubiera dado un mayor tiempo de muestreo.

April 6, 2025 18:29:09 (UTC-5)	INFO	Removed instances [i-0a8520077da1a8405, i-0c1a88ba5ba9c5f82, i-0e54e1c6a4e264f40] from your environment.
April 6, 2025 18:27:09 (UTC-5)	INFO	Environment health has transitioned from Info to Ok. Application update completed 46 seconds ago and took 9 minutes. 3 instances online which meets Auto Scaling group desired capacity of 3.
April 6, 2025 18:26:38 (UTC-5)	INFO	Environment update completed successfully.
April 6, 2025 18:26:38 (UTC-5)	INFO	New application version was deployed to running EC2 instances.
April 6, 2025 18:24:30 (UTC-5)	INFO	Waiting for post-deployment configuration to complete.
April 6, 2025 18:24:26 (UTC-5)	INFO	Instance deployment completed successfully.
April 6, 2025 18:24:15 (UTC-5)	INFO	Starting post-deployment configuration on new instances.
April 6, 2025 18:23:48 (UTC-5)	INFO	Attached new instance(s) to the permanent auto scaling group awseb-e-7p5kfdrbyb-stack-AWSEBAutoScalingGroup-4H9x6Bo3JezW.
April 6, 2025 18:23:45 (UTC-5)	INFO	Detached new instance(s) from temporary auto scaling group awseb-e-7p5kfdrbyb-immutable-stack-AWSEBAutoScalingGroup-Ol0Jlbo0F20R.
April 6, 2025 18:23:15 (UTC-5)	INFO	Detached old instance(s) from environment auto scaling group awseb-e-7p5kfdrbyb-stack-AWSEBAutoScalingGroup-4H9x6Bo3JezW.
April 6, 2025 18:23:14 (UTC-5)	INFO	Traffic-splitting deployment: routed all client traffic to the new application version.
April 6, 2025 18:23:14 (UTC-5)	INFO	Traffic-splitting deployment: starting to shift all client traffic to the new application version.
April 6, 2025 18:23:14 (UTC-5)	INFO	Traffic-splitting deployment: evaluation is complete.
April 6, 2025 18:20:13 (UTC-5)	INFO	Traffic-splitting deployment: all instances passed enhanced health check.
April 6, 2025 18:20:12 (UTC-5)	INFO	Waiting for instance(s) (i-0ab6547efeea8bb82,i-0efea7ffefd949902,i-0080f6fb457f1f9d2) to pass health checks.
April 6, 2025 18:20:12 (UTC-5)	INFO	Traffic-splitting deployment: starting evaluation time for 3 minute(s).
April 6, 2025 18:20:12 (UTC-5)	INFO	Traffic-splitting deployment: all instances that have the new application version passed load balancer health checks.
April 6, 2025 18:20:10 (UTC-5)	INFO	Environment health has transitioned from Ok to Info. Application update in progress on 3 instances. 0 out of 3 instances completed (running for 2 minutes).
April 6, 2025 18:20:10 (UTC-5)	INFO	Added instance [i-0efea7ffefd949902] to your environment.
April 6, 2025 18:19:51 (UTC-5)	INFO	Traffic-splitting deployment: routed 50% of client traffic to the new application version.

Time	Type	Details
April 6, 2025 18:19:51 (UTC-5)	INFO	Traffic-splitting deployment: starting to shift 50% of client traffic to the new application version.
April 6, 2025 18:19:49 (UTC-5)	INFO	Traffic-splitting deployment: successfully launched 3 test instance(s).
April 6, 2025 18:19:10 (UTC-5)	INFO	Added instances [i-0080f6fb457f1f9d2, i-0ab6547efeea8bb82] to your environment.
April 6, 2025 18:18:03 (UTC-5)	INFO	Created temporary auto scaling group awseb-e-7p5kfdrbyb-immutable-stack-AWSEBAutoScalingGroup-Ol0Jlbo0F20R.
April 6, 2025 18:17:47 (UTC-5)	INFO	Traffic Splitting deployment policy enabled. Preparing a new fleet for the new application version.
April 6, 2025 18:17:07 (UTC-5)	INFO	Environment update is starting.