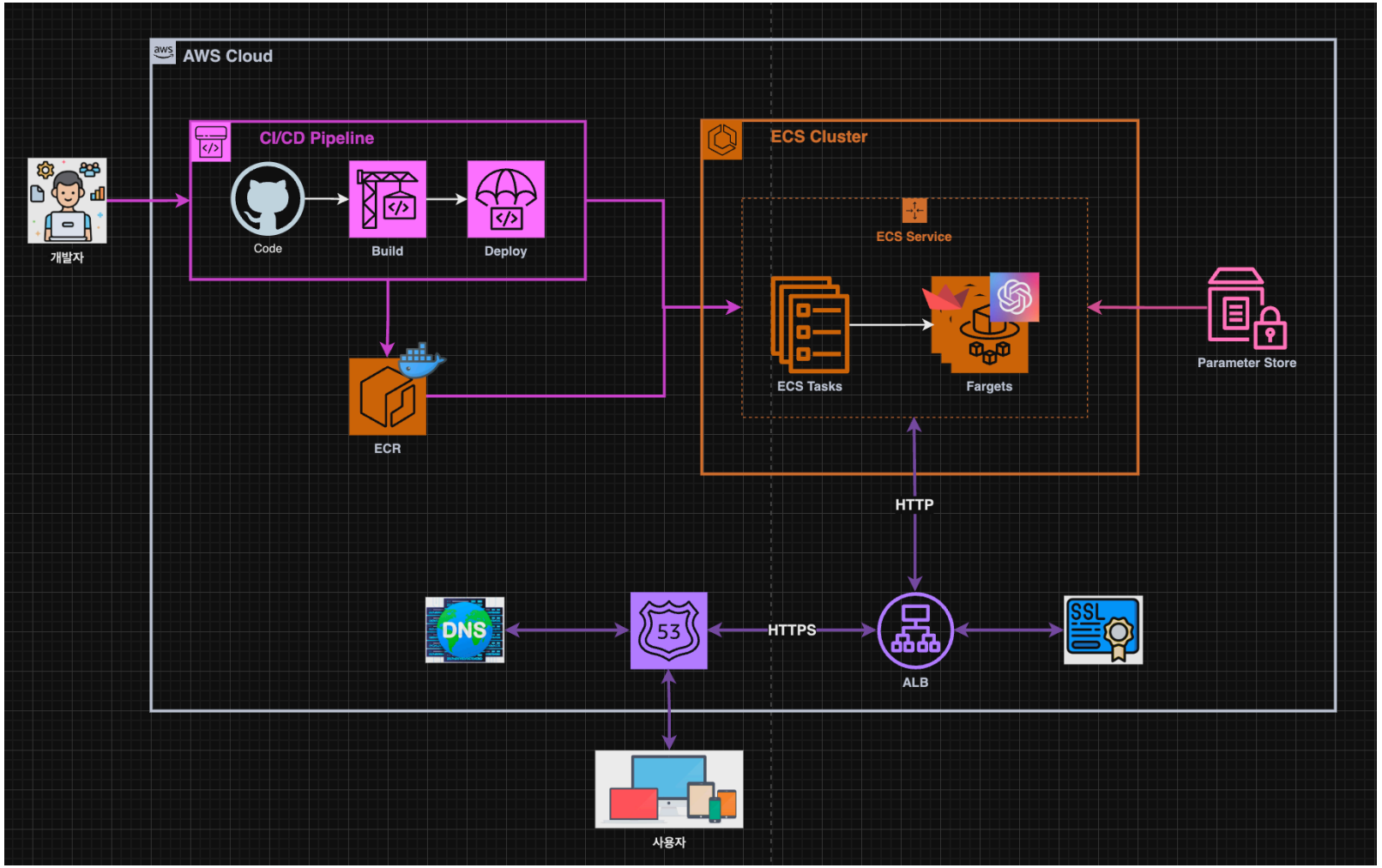
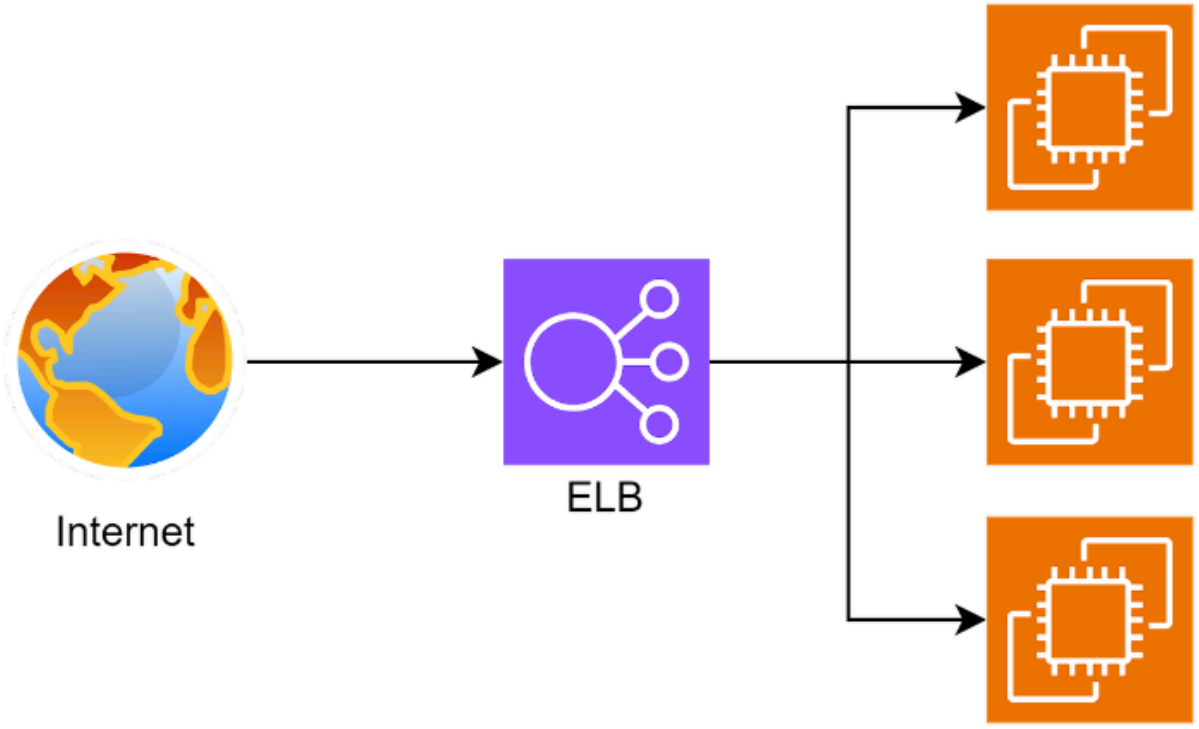
Architecture



ELB (Elastic Load Balancing)

- ELB는 하나 이상의 가용 영역(AZ)에 있는 애플리케이션에게 들어오는 트래픽을 자동으로 분산한다.
- ELB는 AWS의 로드 밸런스 종류 전체를 통틀어 칭하는 말이다.

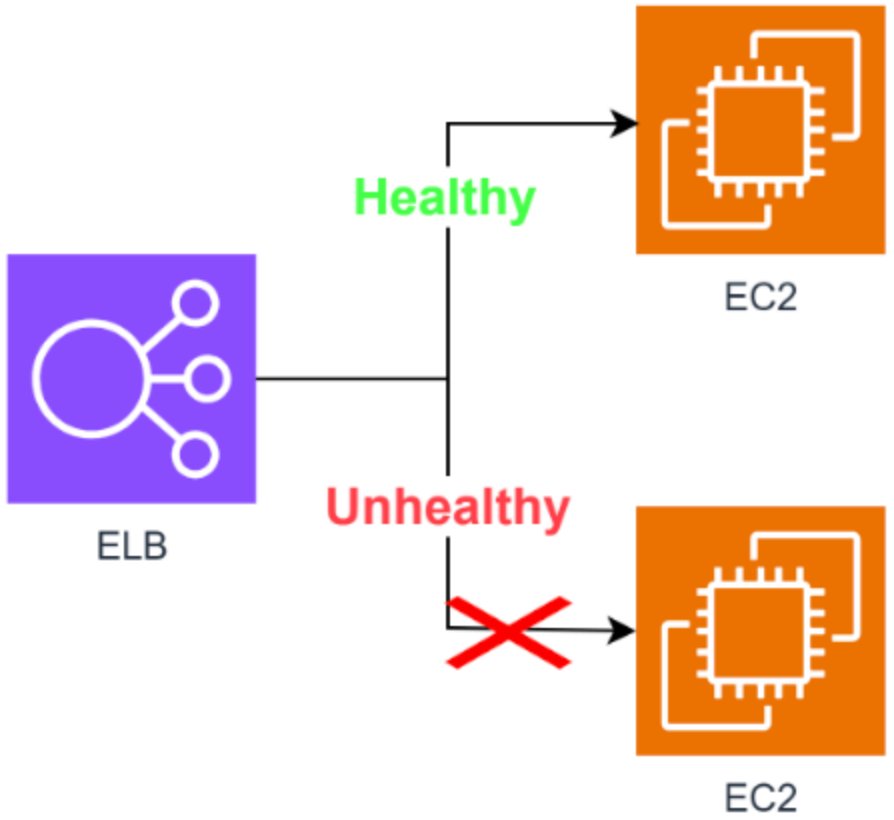


ELB 의 특징/기능

- 애플리케이션의 가용성(High Availability)과 확장성(Auto Scaling)을 향상시킨다.
- ELB는 EC2뿐만 아니라 ECS, Lambda 등 다양한 서비스와 연계하여 트레픽 부하를 분배할 수 있다.
- [Slow Start Configuration, 느린 시작 구성]: 새로 시작된 서버를 위해 천천히 트래픽을 부하
    - ALB와 NLB만 지원한다
    - Slow Start Configuration은 새로운 인스턴스가 부하가 걸리기 전에 천천히 트래픽을 수신하게 만드는 기능이다.
    - 이 방식은 서버가 처음부터 갑자기 모든 트래픽을 받는 것을 방지하고, 서버가 정상적으로 작동할 수 있도록 점진적으로 트래픽을 분배한다.

Health Checks

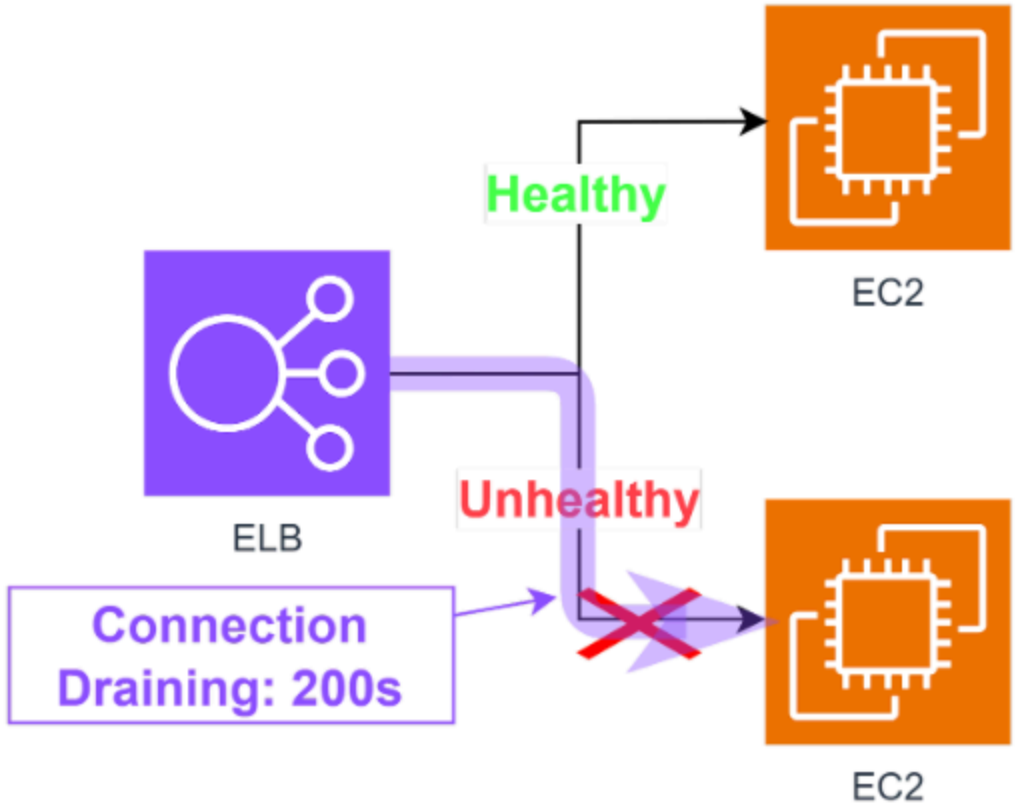
- 서버의 상태를 정기적으로 확인하고, 문제가 있을 경우 트래픽을 해당 서버로 전송하지 않는다.



Connection Draining

- Unhealthy 됐을 때 지정된 시간만큼 연결을 기다려준다
- 기본(Default)으로 켜져있다. ALB(와 CLB)만 지원한다
- 서버에 오류가 발생했을 때, 다른 서버에 트래픽을 이동시키기 전에 기존 연결을 기다려주는 시간

- 예시) Connection Draining을 200초로 설정하면, 서버에 오류 생긴 후 200초 동안 그나마 기존 연결을 기다려주고, 다른 서버에 트래픽을 이동한다.



ELB 로드 밸런서

Classic Load Balancer (CLB)

- ELB중에 가장 오래된 서비스. `No 추천!`
- AWS에서는 신규 생성으로 권장하지 않으며, 기존 사용자에게는 계속 지원중.

Application Load Balancer (ALB)

- ALB는 `L7단의 로드 밸런서`를 지원합니다.
- ALB는 `HTTP/HTTPS 프로토콜`의 헤더를 보고 적절한 패킷으로 전송합니다.
- ALB는 IP주소 + 포트번호 + 패킷 내용을 보고 스위칭합니다.
- ALB는 IP 주소가 변동되기 때문에 Client에서 Access 할 ELB의 DNS Name을 이용해야 합니다.
- ALB는 L7단을 지원하기 때문에 `SSL 적용`이 가능합니다.

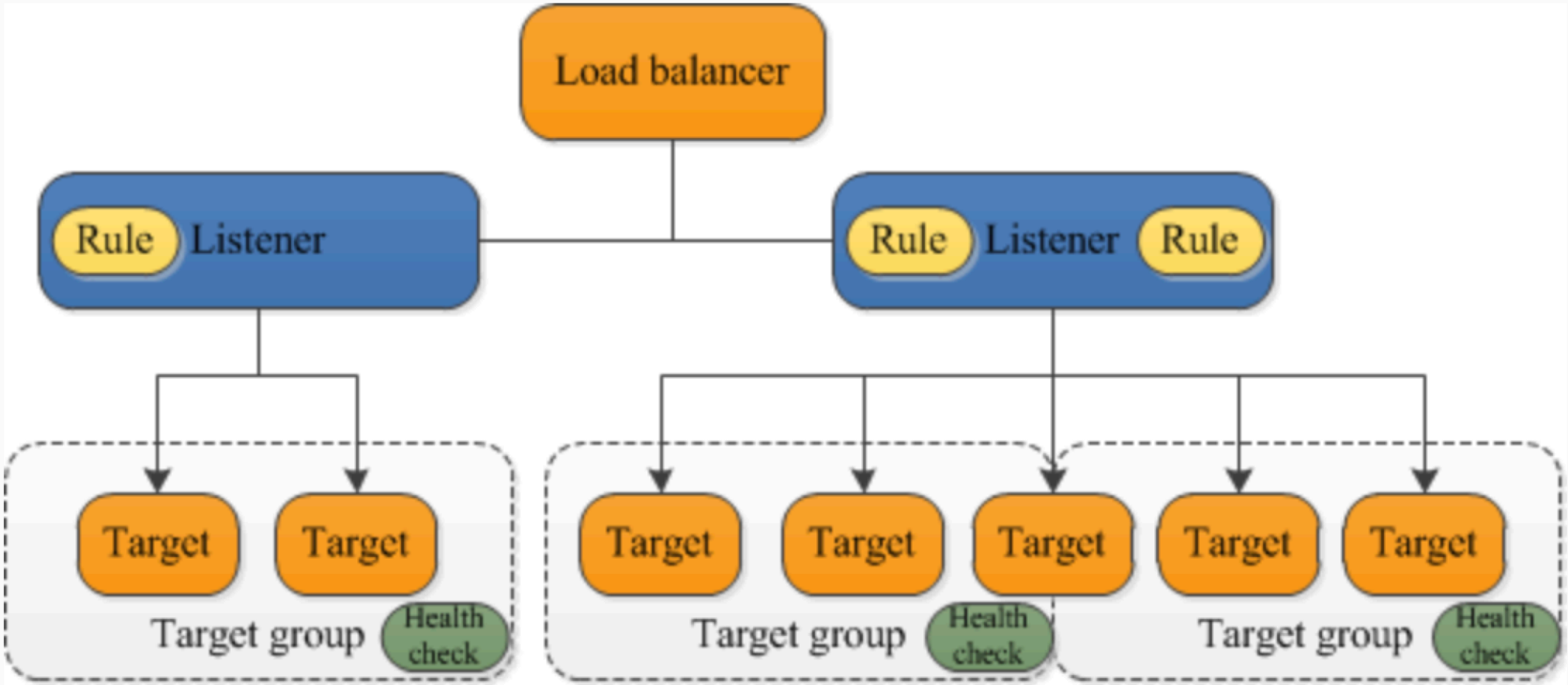
Network Load Balancer (NLB)

- NLB는 `L4단의 로드 밸런서`를 지원합니다.
- NLB는 `TCP/IP 프로토콜`의 헤더를 보고 적절한 패킷으로 전송합니다.
- NLB는 IP + 포트번호를 보고 스위칭합니다.
- NLB는 할당한 `Elastic IP를 Static IP로 사용이 가능`하여 DNS Name과 IP주소 모두 사용이 가능합니다.
- NLB는 `SSL 적용이 인프라 단에서 불가능`하여 애플리케이션에서 따로 적용해 주어야 합니다.

AWS ALB 구현

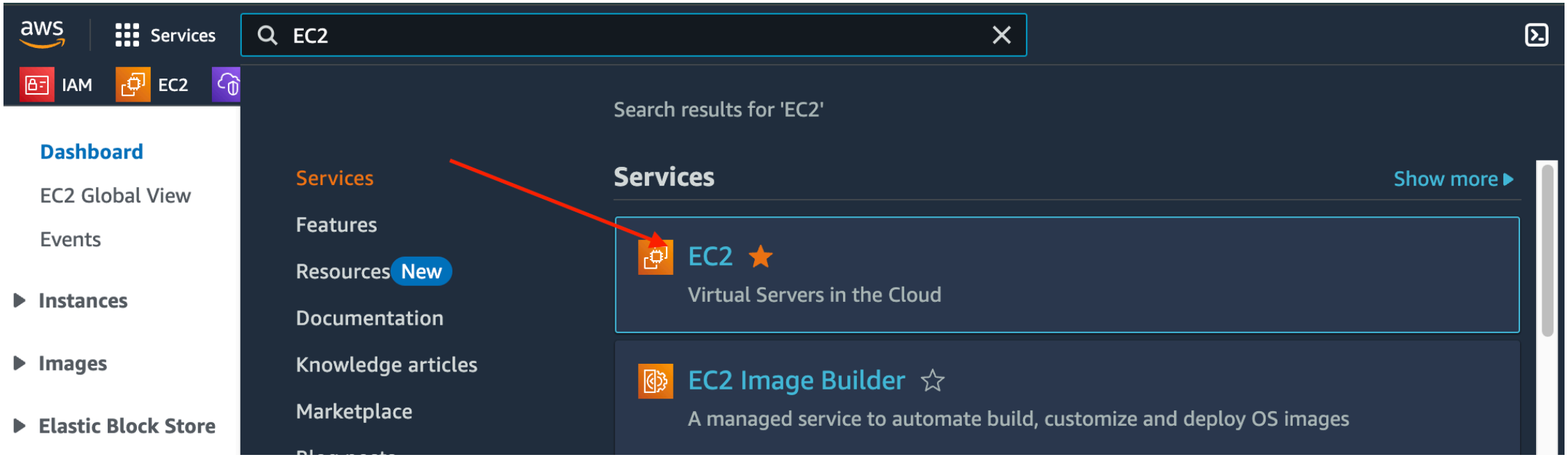
- (ALB에서만 사용 가능) Listener Rules는 수신한 요청을 어떻게 처리할지에 대한 규칙(Listener Rules)을 설정한 기능이다.
- 헤더 또는 쿼리 문자열의 내용에 따라 요청을 다르게 처리할 수도 있다.
- HTTP로 들어오는 요청을 HTTPS로 리다이렉트 할 수 있다.

다음 다이어그램은 ALB의 작동 방식을 보여줍니다.

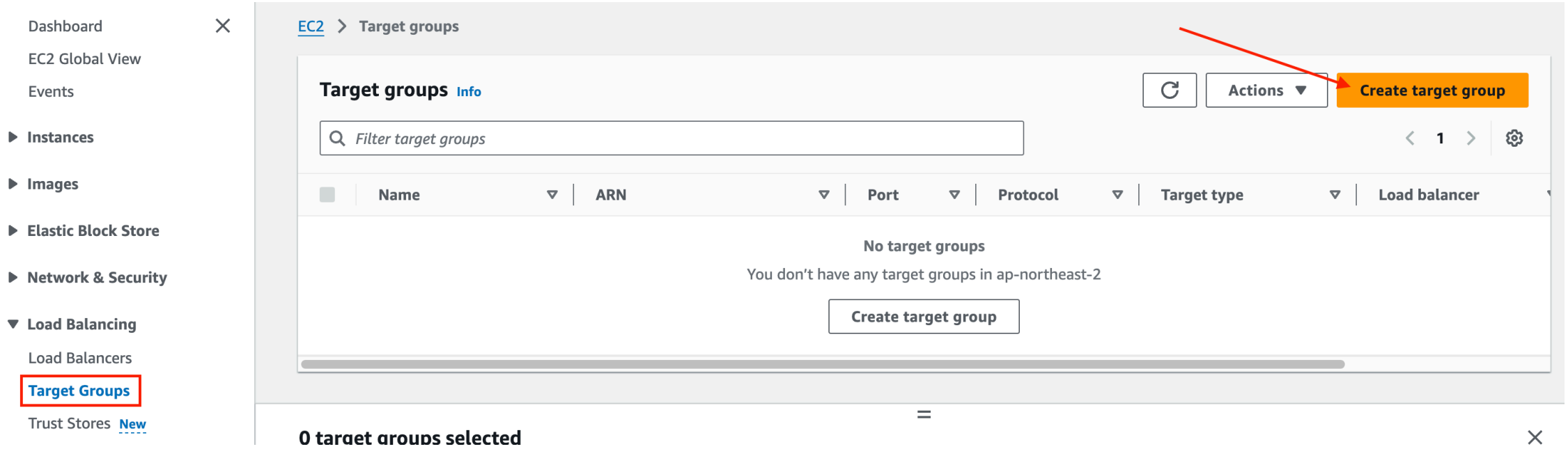


Target Group

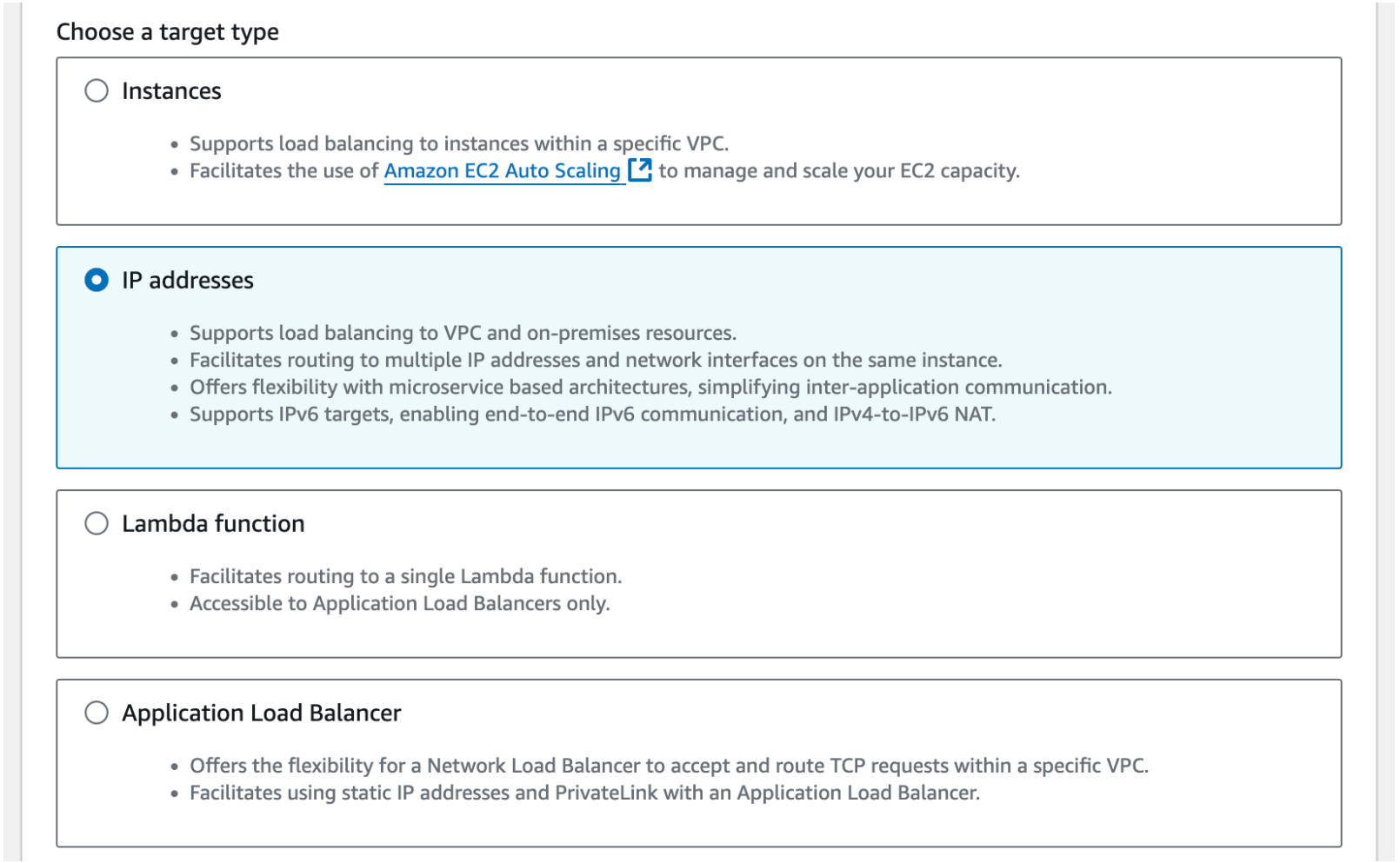
단계1: EC2 접속


aws
Services
EC2
IAM
EC2
Dashboard
EC2 Global View
Events
Instances
Images
Elastic Block Store
Search results for 'EC2'
Services
Features
Resources New
Documentation
Knowledge articles
Marketplace
Services
Show more
EC2
Virtual Servers in the Cloud
EC2 Image Builder
A managed service to automate build, customize and deploy OS images

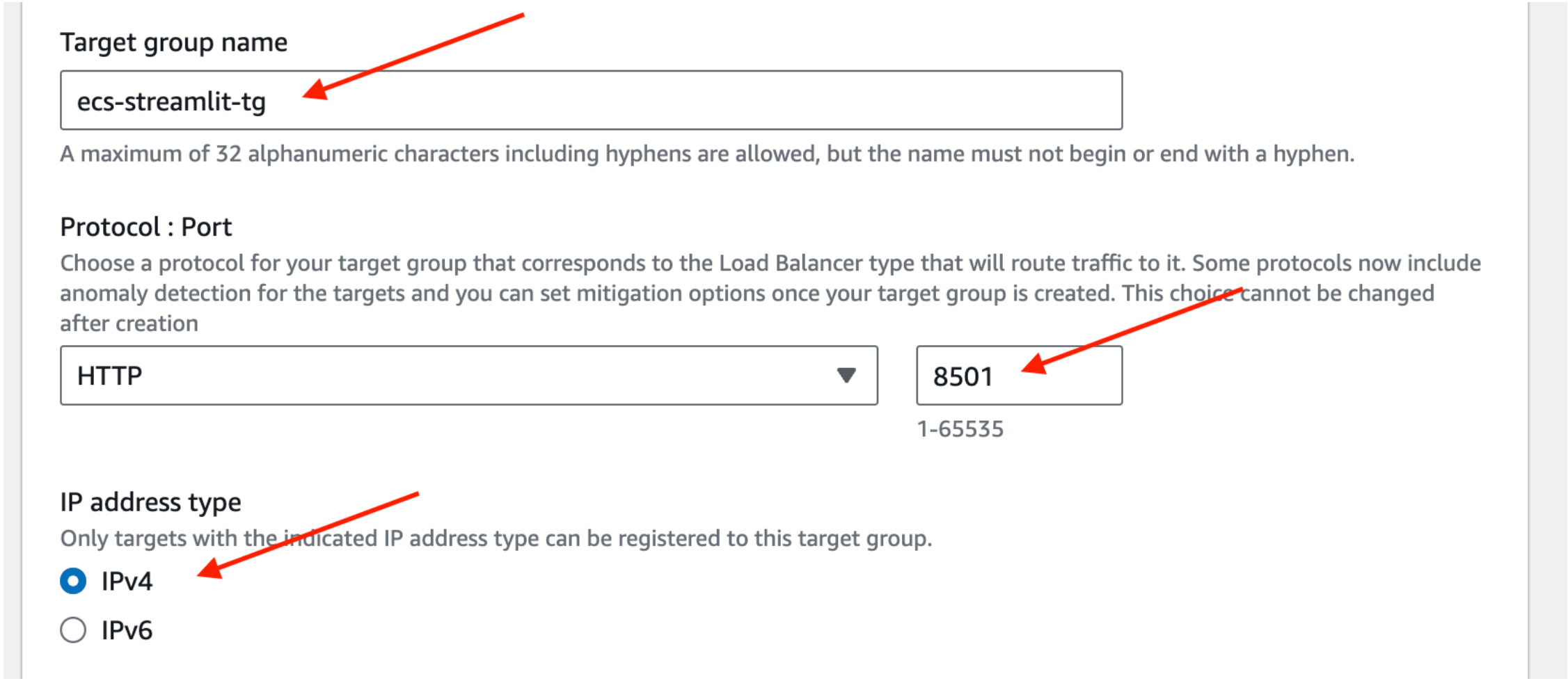
단계2: Create target group



단계3: Choose a target type



단계4: Target group name & port



단계5: VPC & Protocol version

VPC

Select the VPC that hosts the load balancer. Only VPCs that support the IP address type selected above are available in this list. On the **Register targets** page, you can register IP addresses from this VPC, or from private IP addresses located outside of this load balancer's VPC (such as a peered VPC, EC2-Classic, or on-premises targets that are reachable over Direct Connect or VPN).

-
vpc-0e092393ffbd671b9
IPv4 VPC CIDR: 172.31.0.0/16

Protocol version

- (●) HTTP1
  Send requests to targets using HTTP/1.1. Supported when the request protocol is HTTP/1.1 or HTTP/2.
- ( ) HTTP2
  Send requests to targets using HTTP/2. Supported when the request protocol is HTTP/2 or gRPC, but gRPC-specific features are not available.
- ( ) gRPC
  Send requests to targets using gRPC. Supported when the request protocol is gRPC.

단계6: Health checks

Health checks

The associated load balancer periodically sends requests, per the settings below, to the registered targets to test their status.

Health check protocol

HTTP

Health check path

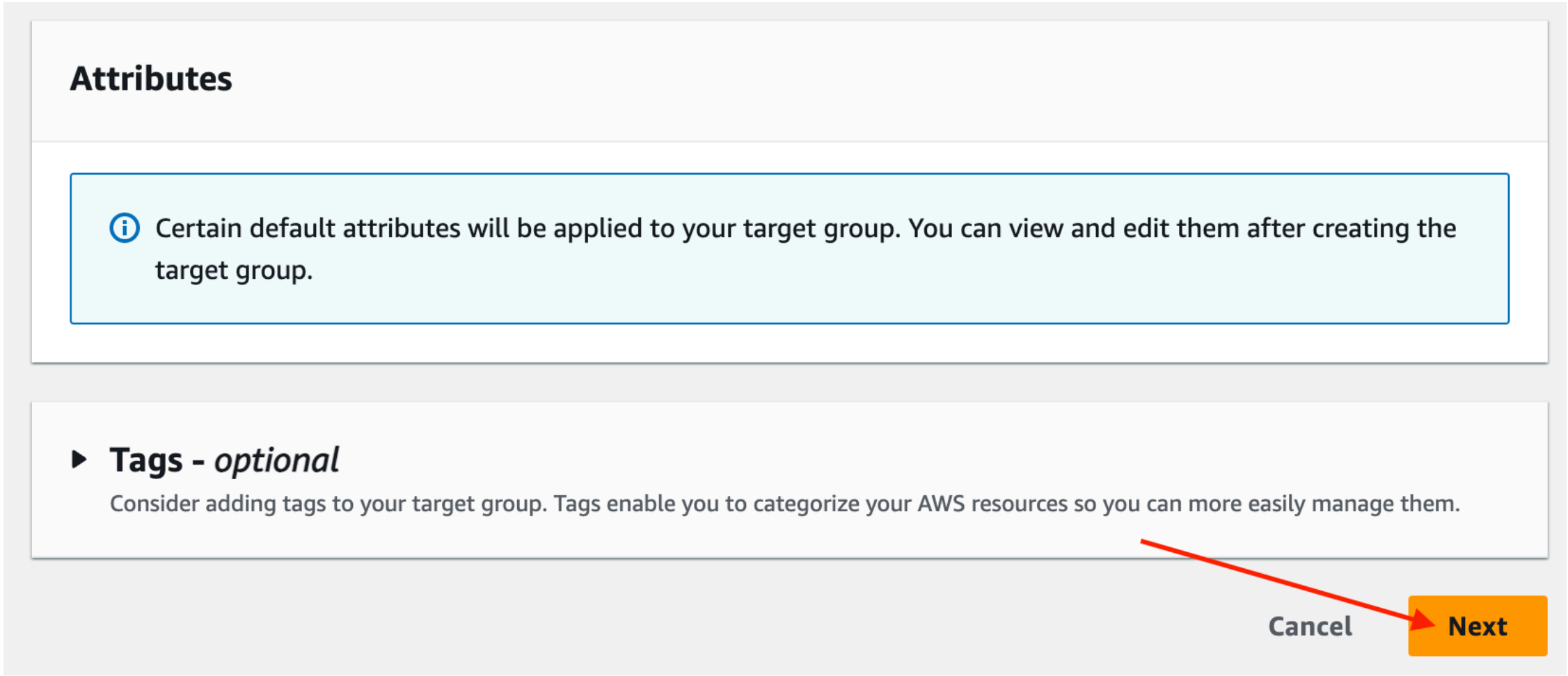
Use the default path of "/" to perform health checks on the root, or specify a custom path if preferred.

/

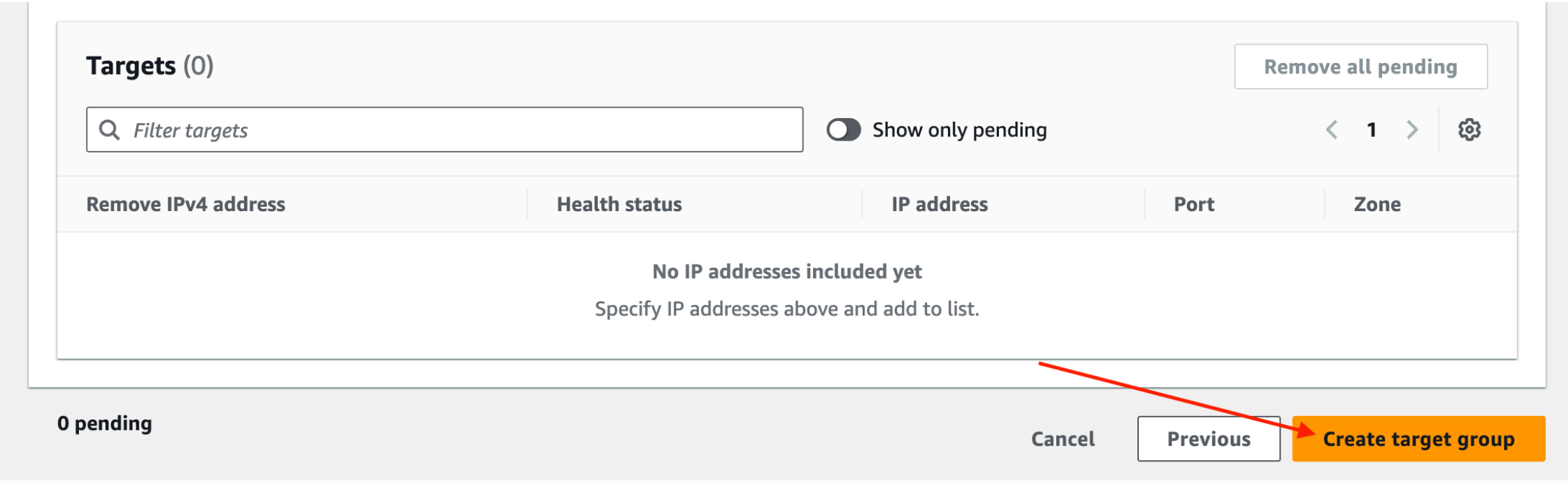
Up to 1024 characters allowed.

▶ **Advanced health check settings**

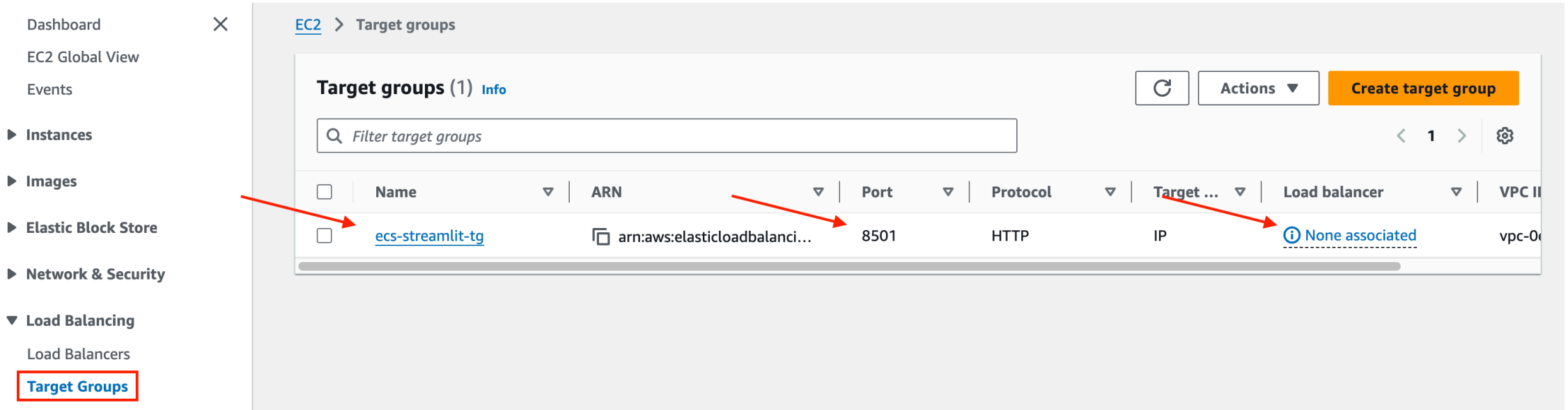
단계7: Next


Attributes
Certain default attributes will be applied to your target group. You can view and edit them after creating the target group.
Tags - optional
Consider adding tags to your target group. Tags enable you to categorize your AWS resources so you can more easily manage them.
Cancel
Next

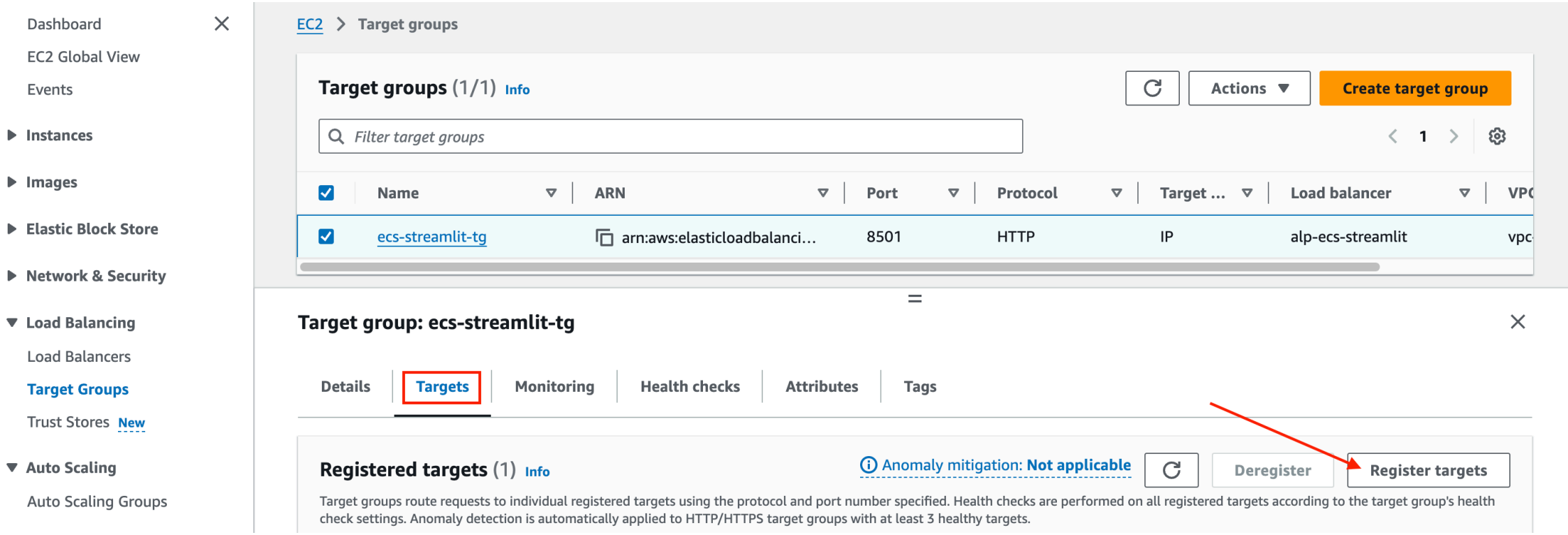
단계8: Create target group



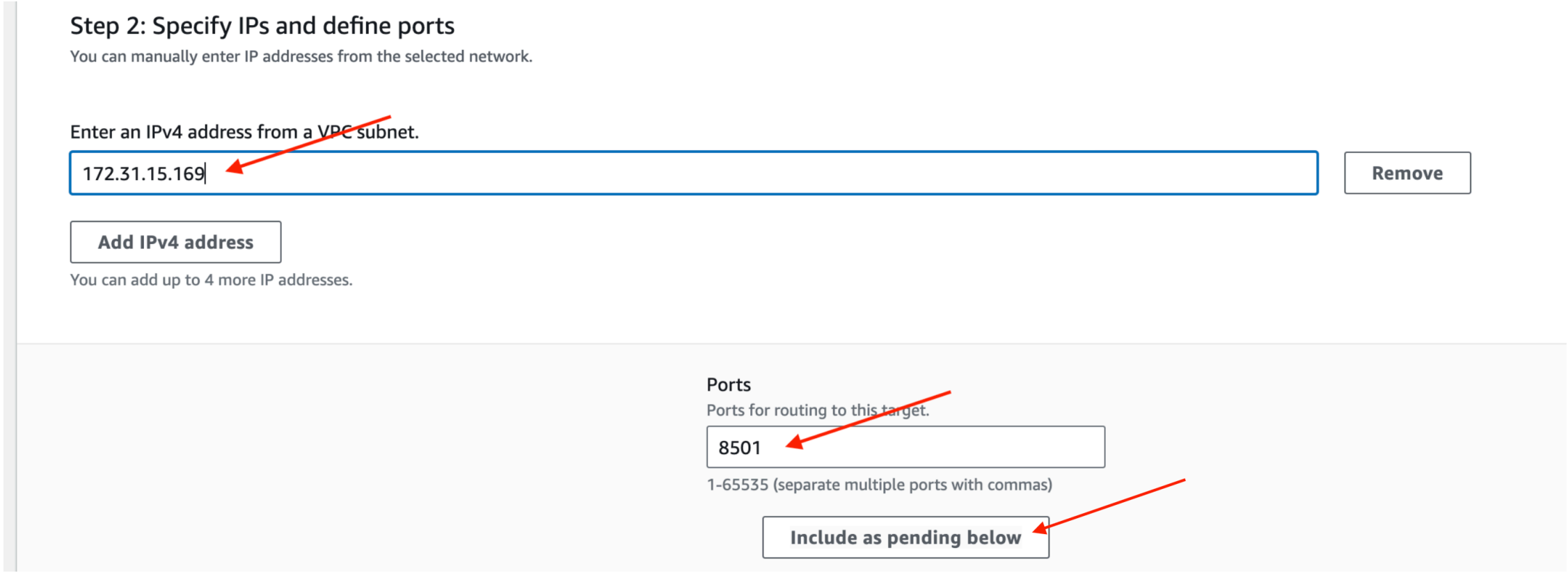
단계9: 결과 확인

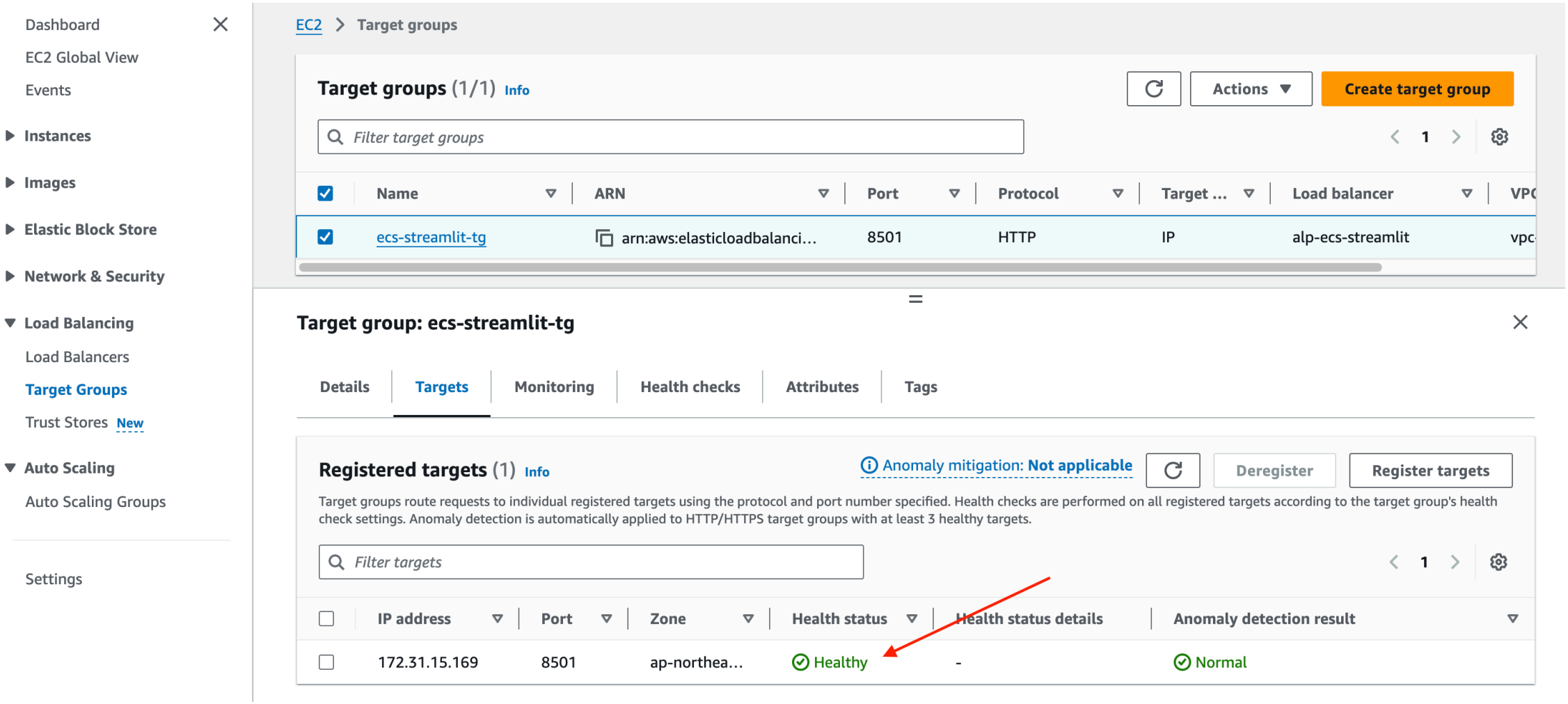

Dashboard
EC2 Global View
Events
Instances
Images
Elastic Block Store
Network & Security
Load Balancing
Load Balancers
Target Groups
EC2 > Target groups
Target groups (1) Info
Actions
Create target group
Filter target groups
1
Name
ARN
Port
Protocol
Target ...
Load balancer
VPC I
ecs-streamlit-tg
arn:aws:elasticloadbalanci...
8501
HTTP
IP
None associated
vpc-0

단계10: Registered targets



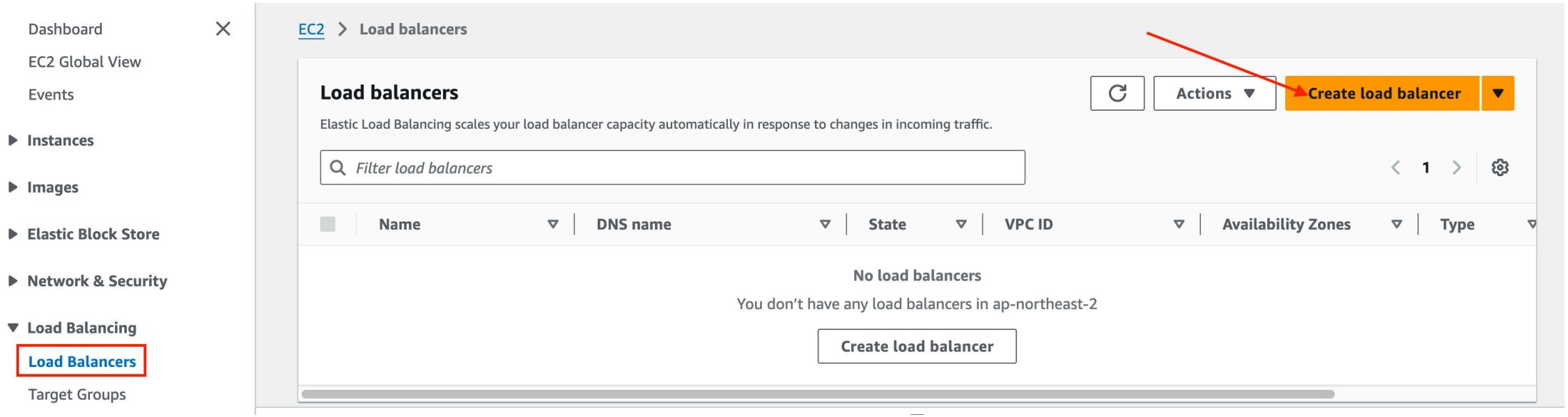
단계11: ECS의 target private ip & port 적용




Dashboard
EC2 Global View
Events
Instances
Images
Elastic Block Store
Network & Security
Load Balancing
Load Balancers
Target Groups
Trust Stores New
Auto Scaling
Auto Scaling Groups
Settings
EC2 > Target groups
Target groups (1/1) Info
Actions
Create target group
Filter target groups
1
Name
ARN
Port
Protocol
Target ...
Load balancer
ecs-streamlit-tg
arn:aws:elasticloadbalanci...
8501
HTTP
IP
alp-ecs-streamlit
Target group: ecs-streamlit-tg
Details
Targets
Monitoring
Health checks
Attributes
Tags
Registered targets (1) Info
Anomaly mitigation: Not applicable
Deregister
Register targets
Target groups route requests to individual registered targets using the protocol and port number specified. Health checks are performed on all registered targets according to the target group's health check settings. Anomaly detection is automatically applied to HTTP/HTTPS target groups with at least 3 healthy targets.
Filter targets
1
IP address
Port
Zone
Health status
Health status details
Anomaly detection result
172.31.15.169
8501
ap-northea...
Healthy
-
Normal

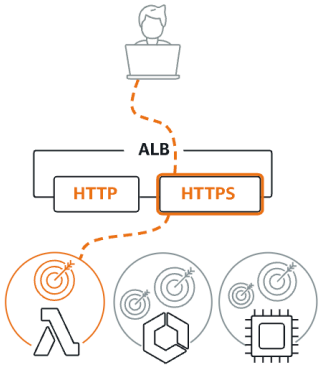
ALB

단계1: Create load balancer



단계2: Load balancer types

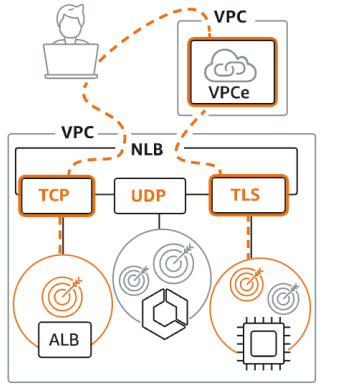
Application Load Balancer Info



Choose an Application Load Balancer when you need a flexible feature set for your applications with HTTP and HTTPS traffic. Operating at the request level, Application Load Balancers provide advanced routing and visibility features targeted at application architectures, including microservices and containers.

Create

Network Load Balancer Info



Choose a Network Load Balancer when you need ultra-high performance, TLS offloading at scale, centralized certificate deployment, support for UDP, and static IP addresses for your applications. Operating at the connection level, Network Load Balancers are capable of handling millions of requests per second securely while maintaining ultra-low latencies.

Create

Gateway Load Balancer Info

Choose a Gateway Load Balancer when you need to deploy and manage a fleet of third-party virtual appliances that support GENEVE. These appliances enable you to improve security, compliance, and policy controls.

Create

단계3: Basic configuration

Load balancer name

Name must be unique within your AWS account and can't be changed after the load balancer is created.

alp-ecs-streamlit

A maximum of 32 alphanumeric characters including hyphens are allowed, but the name must not begin or end with a hyphen.

Scheme | Info

Scheme can't be changed after the load balancer is created.

- (●) Internet-facing
  An internet-facing load balancer routes requests from clients over the internet to targets. Requires a public subnet. Learn more
- ( ) Internal
  An internal load balancer routes requests from clients to targets using private IP addresses. Compatible with the **IPv4** and **Dualstack** IP address types.

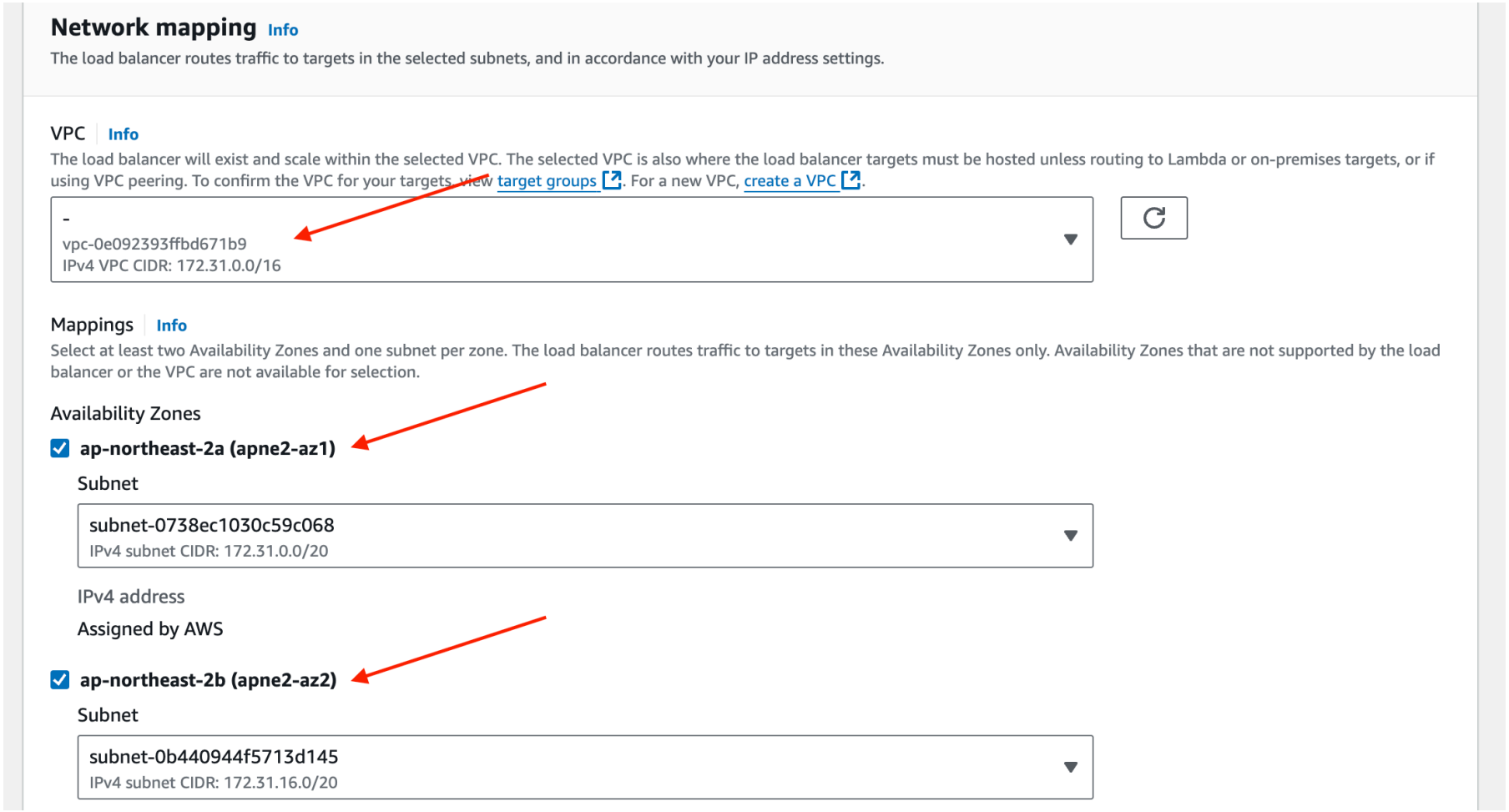
Load balancer IP address type | Info

Select the front-end IP address type to assign to the load balancer. The VPC and subnets mapped to this load balancer must include the selected IP address types. Public IPv4 addresses have an additional cost.

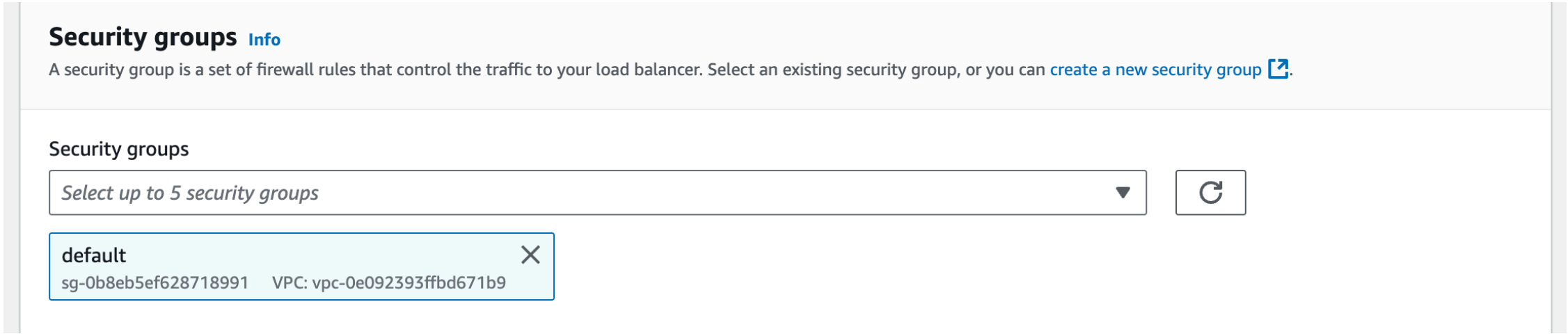
- (●) IPv4
  Includes only IPv4 addresses.
- ( ) Dualstack
  Includes IPv4 and IPv6 addresses.
- ( ) Dualstack without public IPv4
  Includes a public IPv6 address, and private IPv4 and IPv6 addresses. Compatible with **internet-facing** load balancers only.

단계4: Network mapping

> Mappings는 두개 이상 선택해줍니다.

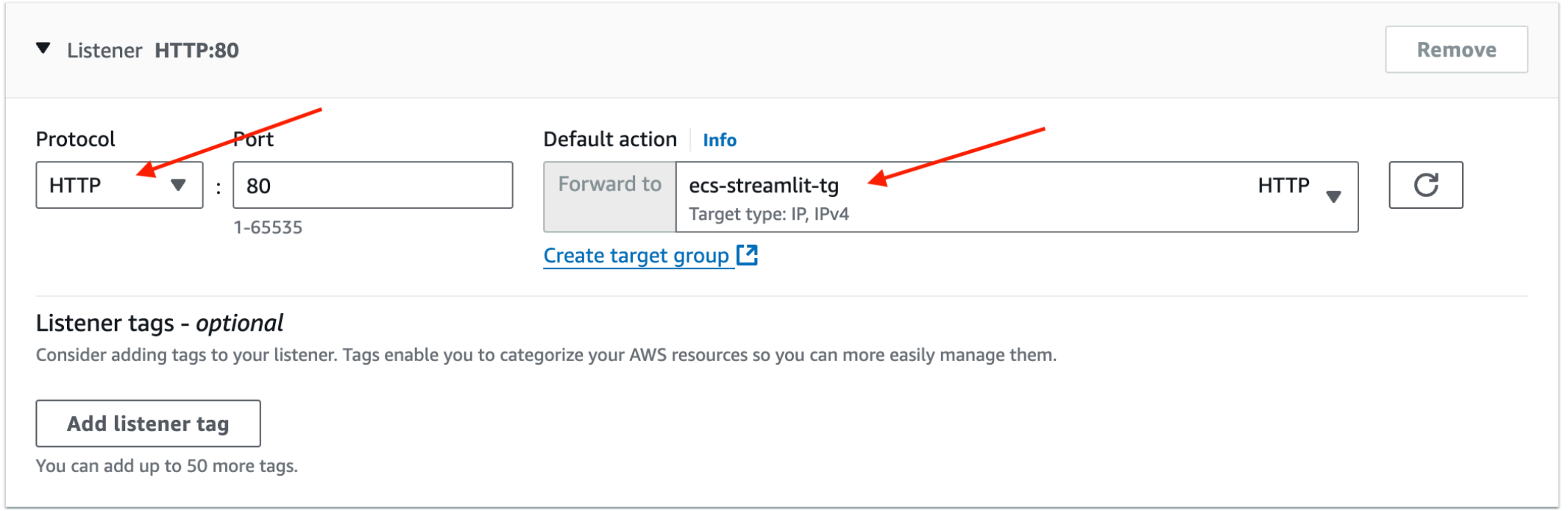


단계5: Security groups

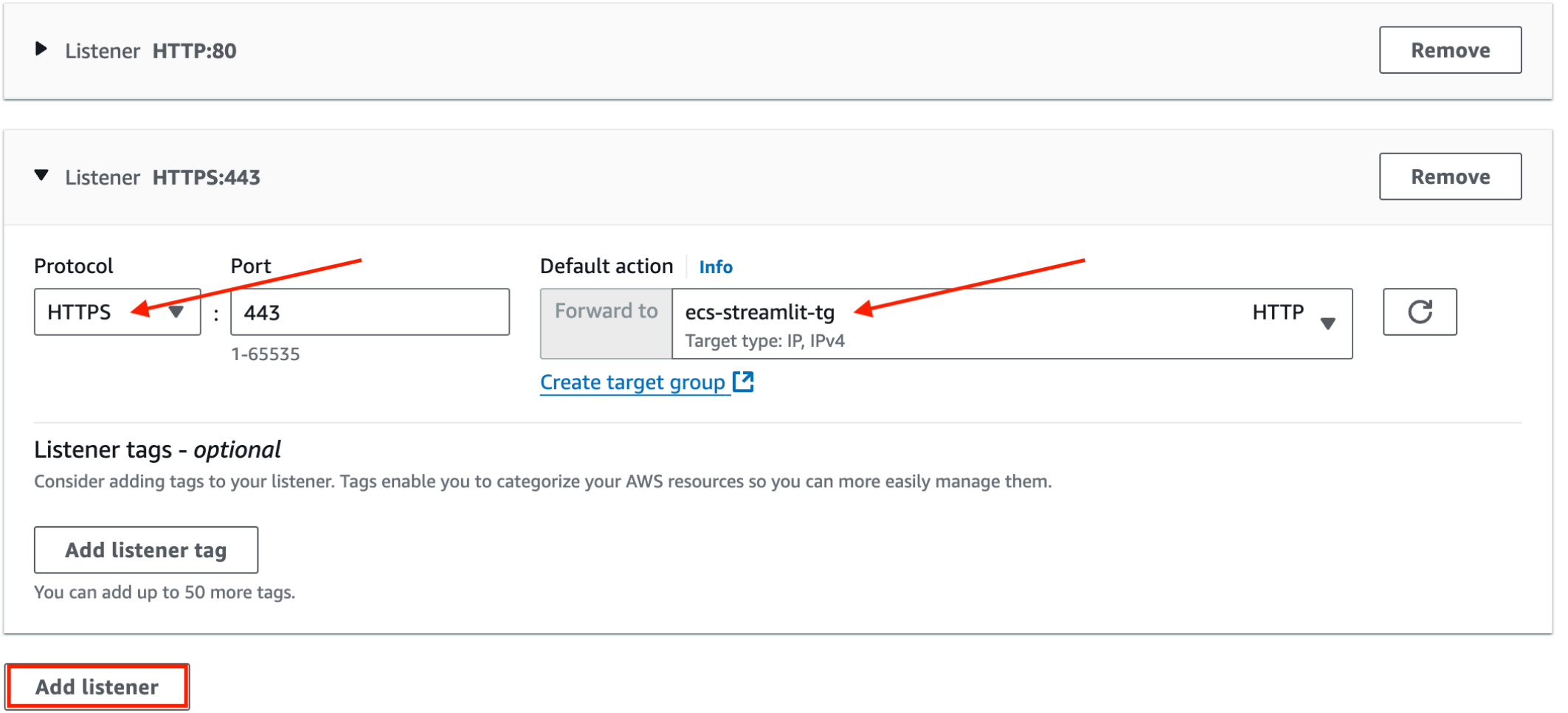


단계6: Listeners and routing

> 기존에 생성한 target group 선택

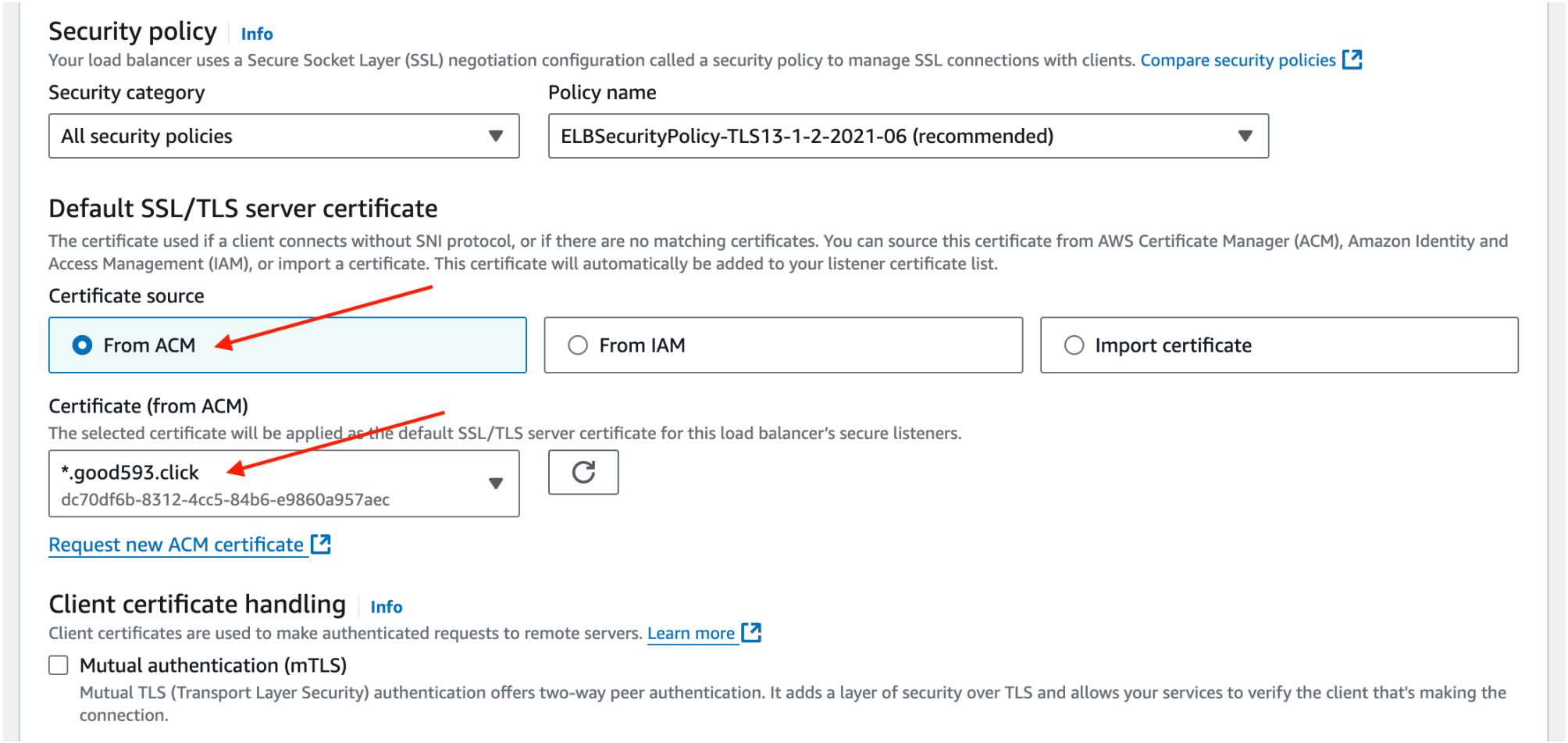


HTTPS도 추가


Listener HTTP:80
Remove
Listener HTTPS:443
Remove
Protocol
Port
HTTPS
:
443
1-65535
Default action
Info
Forward to
ecs-streamlit-tg
Target type: IP, IPv4
HTTP
Create target group
Listener tags - optional
Consider adding tags to your listener. Tags enable you to categorize your AWS resources so you can more easily manage them.
Add listener tag
You can add up to 50 more tags.
Add listener

단계7: Secure listener settings

HTTPS에 대한 생성한 ACM 적용



단계8: Optimize with service integrations - optional

AWS Web Application Firewall (WAF) Info

Optimizes: **Security**

☐ Apply application layer security protections - *in front of targets*

Your choice of either a pre-defined security configuration with the AWS-recommended security protections, or associate any of your existing WAF configurations for custom protections.
Additional charges apply

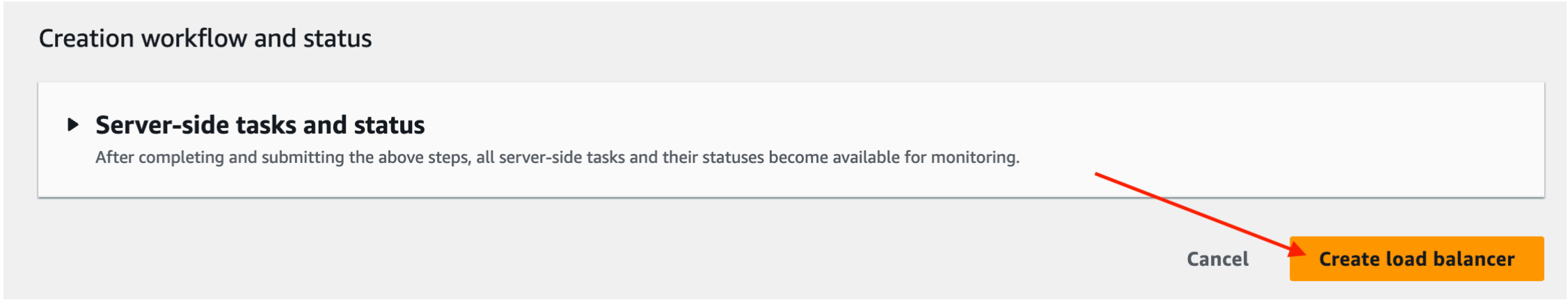
AWS Global Accelerator Info

Optimizes: **Performance, Availability**

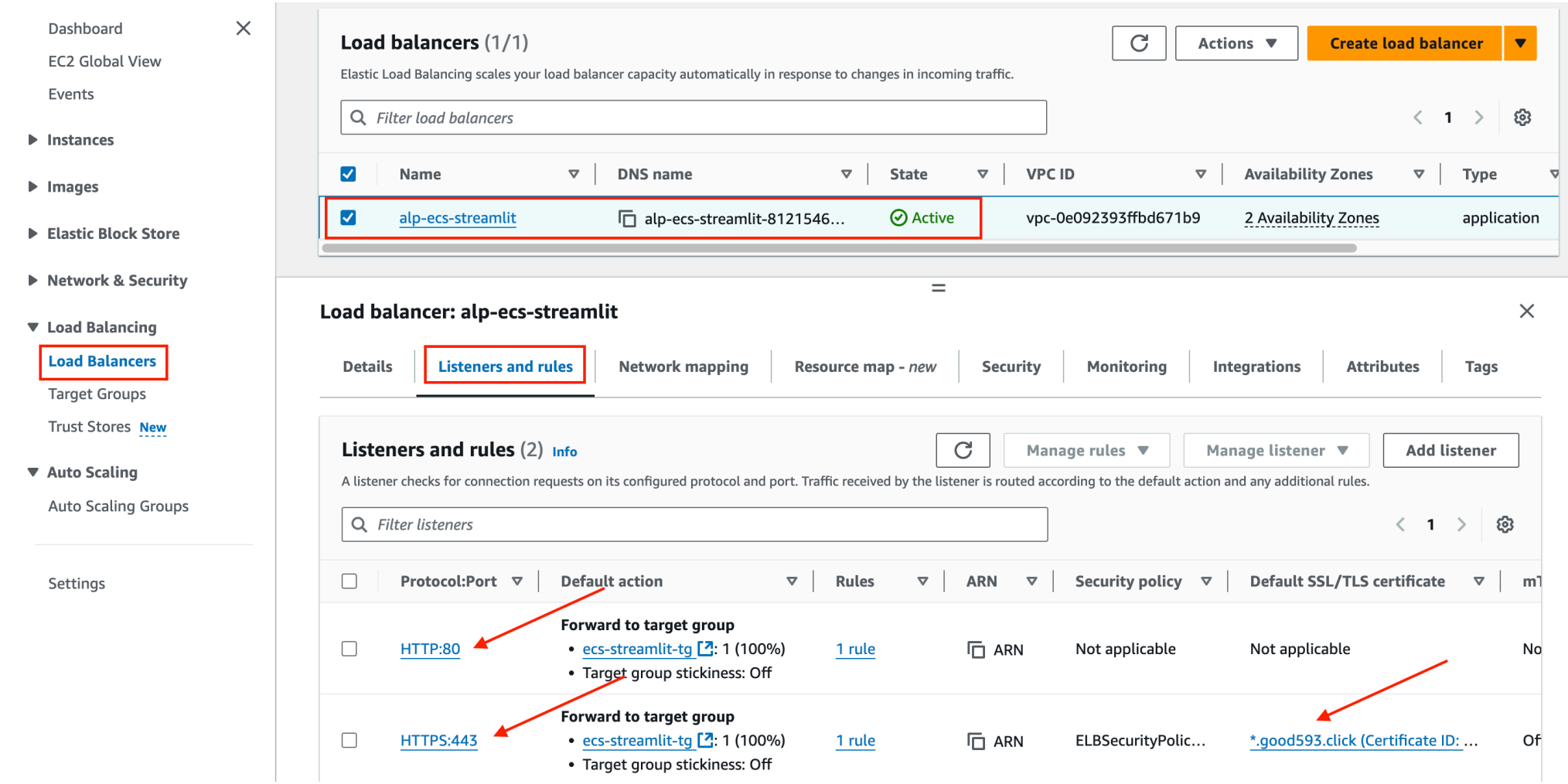
☐ Apply global load balancing across multiple regions

Creates an accelerator in your account with two global static IPs that act as a fixed entry point to your load balancer. If you do not need global static IPs or traffic management across multiple regions, select Amazon CloudFront. **Additional charges apply**

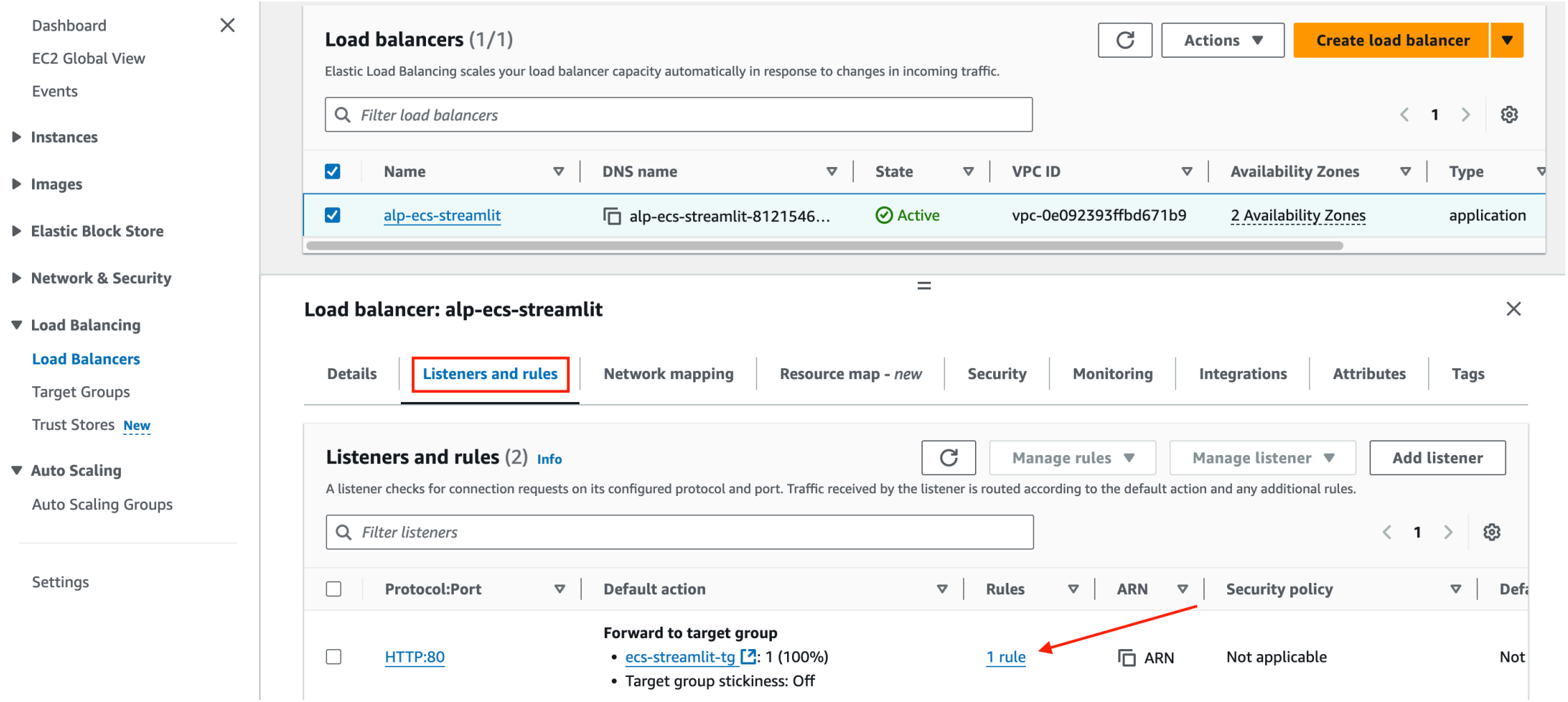
단계9: Create load balancer



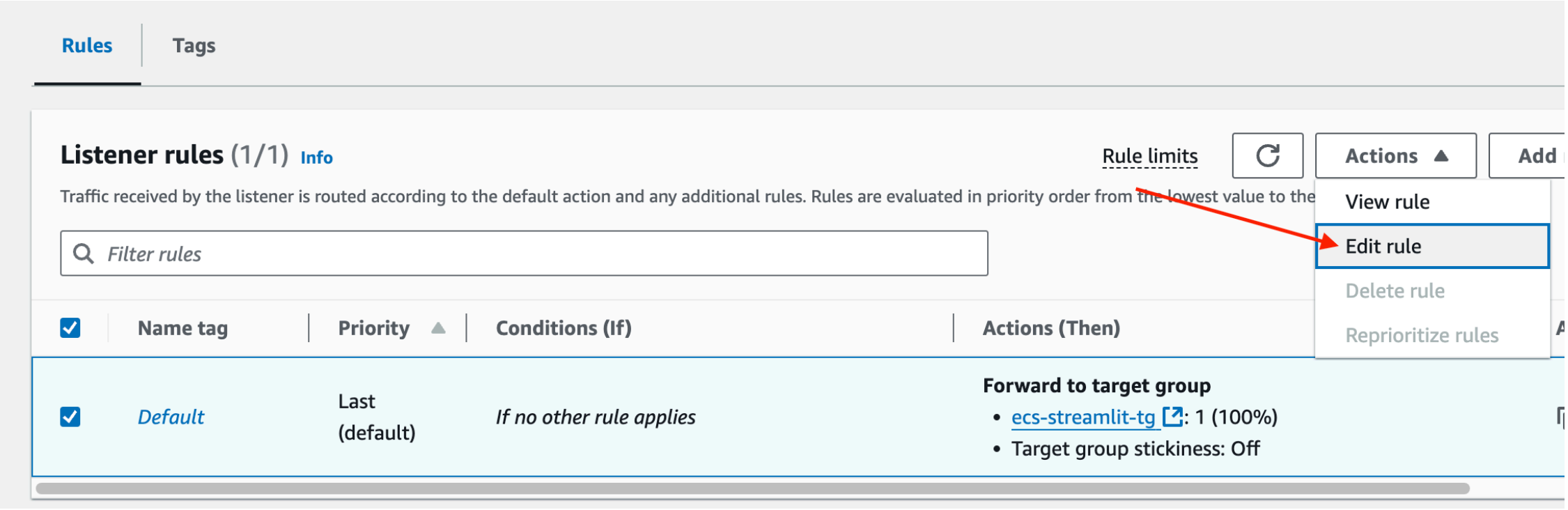
단계10: 결과 확인


Dashboard
EC2 Global View
Events
Instances
Images
Elastic Block Store
Network & Security
Load Balancing
Load Balancers
Target Groups
Trust Stores New
Auto Scaling
Auto Scaling Groups
Settings
Load balancers (1/1)
Elastic Load Balancing scales your load balancer capacity automatically in response to changes in incoming traffic.
Actions
Create load balancer
Filter load balancers
Name
DNS name
State
VPC ID
Availability Zones
Type
alp-ecs-streamlit
alp-ecs-streamlit-8121546...
Active
vpc-0e092393ffbd671b9
2 Availability Zones
application
Load balancer: alp-ecs-streamlit
Details
Listeners and rules
Network mapping
Resource map - new
Security
Monitoring
Integrations
Attributes
Tags
Listeners and rules (2) Info
A listener checks for connection requests on its configured protocol and port. Traffic received by the listener is routed according to the default action and any additional rules.
Manage rules
Manage listener
Add listener
Filter listeners
Protocol:Port
Default action
Rules
ARN
Security policy
Default SSL/TLS certificate
HTTP:80
Forward to target group
ecs-streamlit-tg: 1 (100%)
Target group stickiness: Off
1 rule
ARN
Not applicable
Not applicable
HTTPS:443
Forward to target group
ecs-streamlit-tg: 1 (100%)
Target group stickiness: Off
1 rule
ARN
ELBSecurityPolic...
*.good593.click (Certificate ID: ...

단계11: HTTP:80 > rule



단계12: Edit rule



단계13: Listener details

Listener details

A listener checks for connection requests using the protocol and port that you configure. The default action and any additional rules that you create determine how the Application Load Balancer routes requests to its registered targets.

Listener ARN

arn:aws:elasticloadbalancing:ap-northeast-2:426653742146:listener/app/alp-ecs-streamlit/403ec527635a07ff/7ad62826afef260c

Listener configuration

The listener will be identified by the protocol and port.

Protocol

Used for connections from clients to the load balancer.

HTTP

Port

The port on which the load balancer is listening for connections.

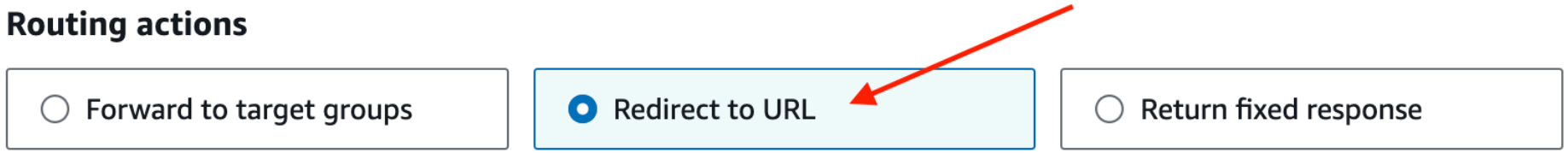
80

1-65535

Default actions | Info

The default action is used if no other rules apply. Choose the default action for traffic on this listener.

Routing actions



Redirect to URL | Info

Redirect client requests from one URL to another. You cannot redirect HTTPS to HTTP. To avoid a redirect loop, you must modify at least one of the following components: protocol, port, hostname or path. Components that you do not modify retain their original values.

Protocol

Used for connections from clients to the load balancer.

Port

The port on which the load balancer is listening for connections.

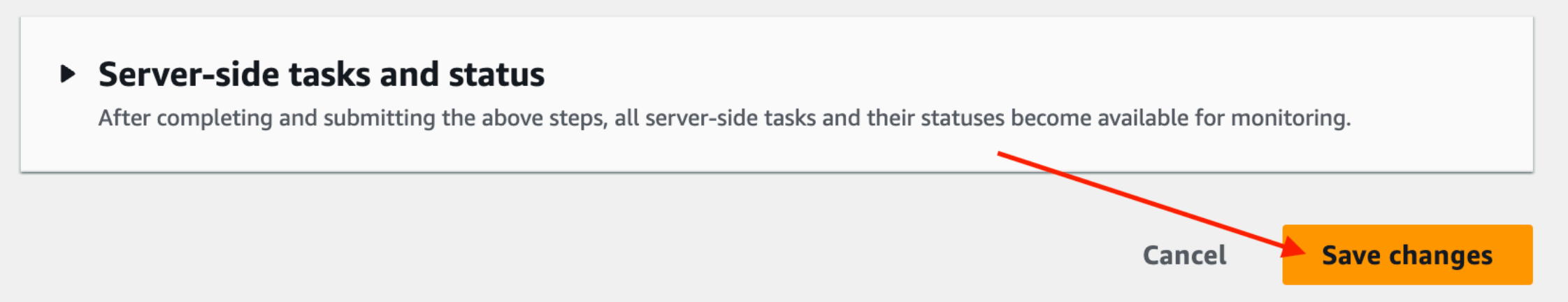
1-65535 or to retain the original port enter #{port}

Custom host, path, query

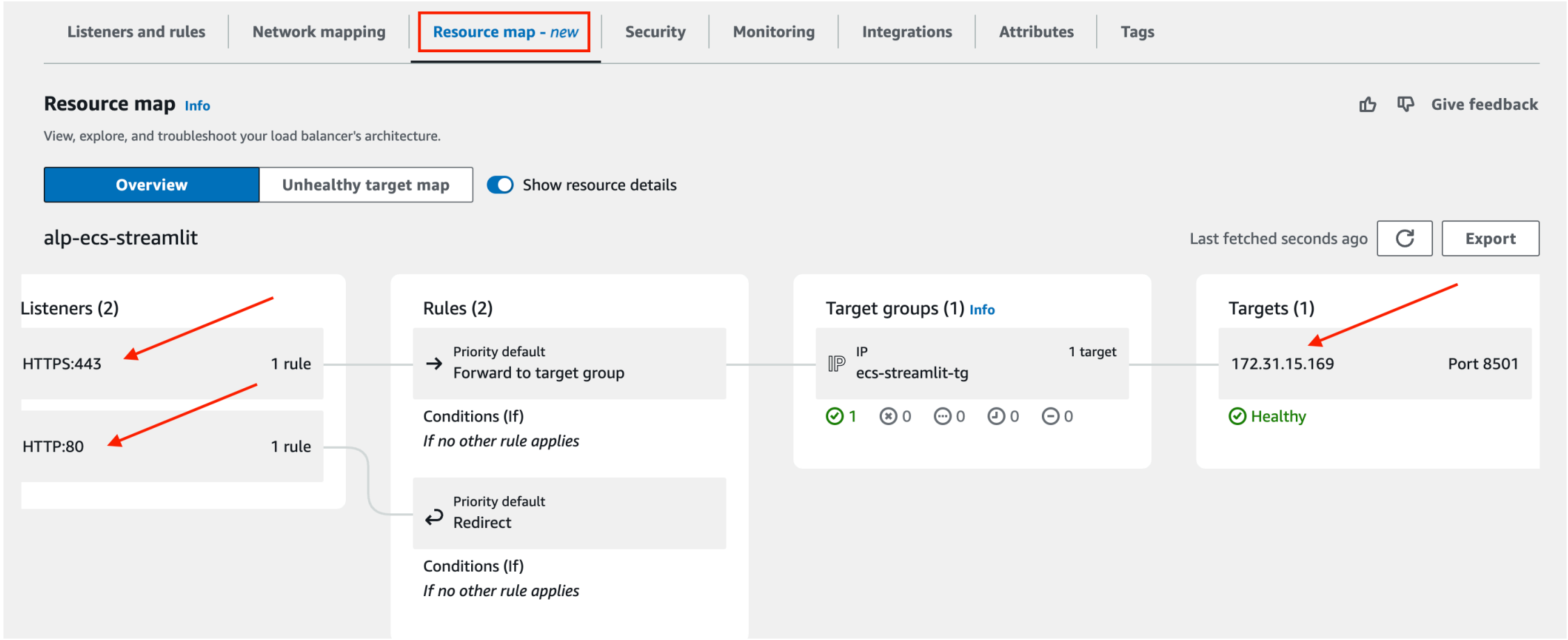
Select to modify host, path and query. If no changes are made, settings from the request URL are retained.

Status code

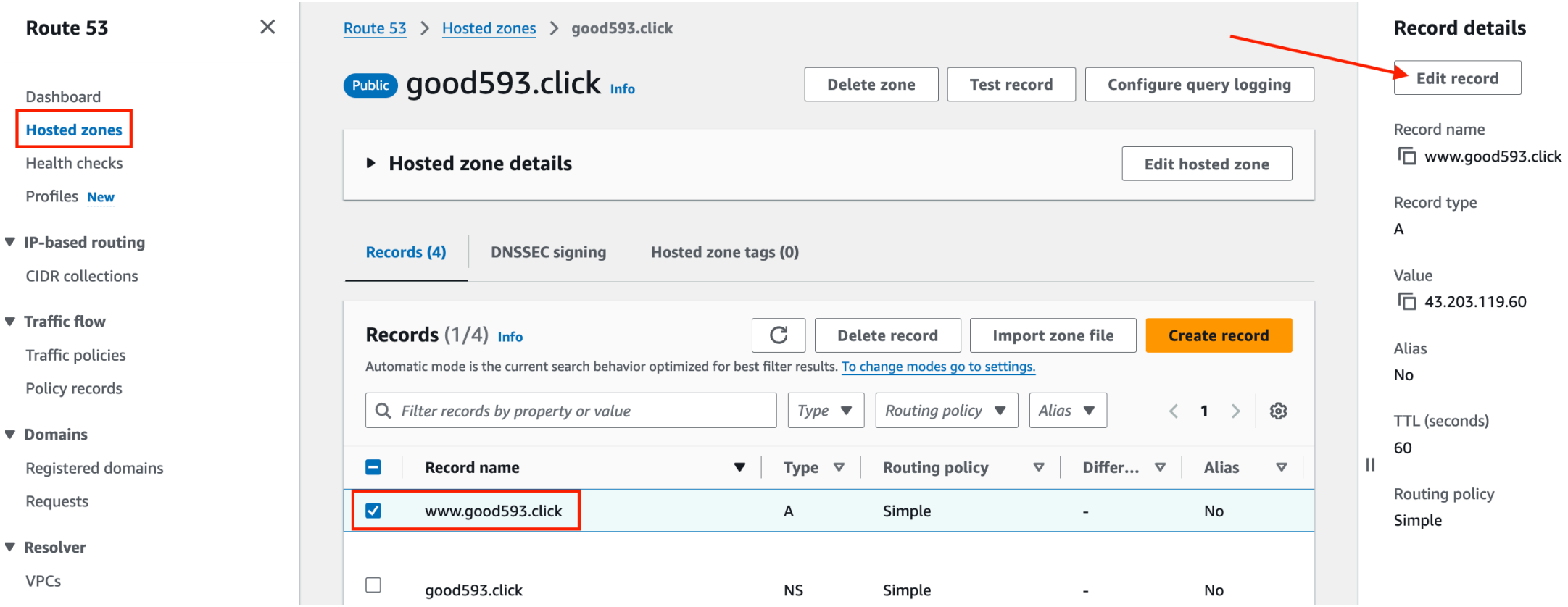
단계14: Save changes

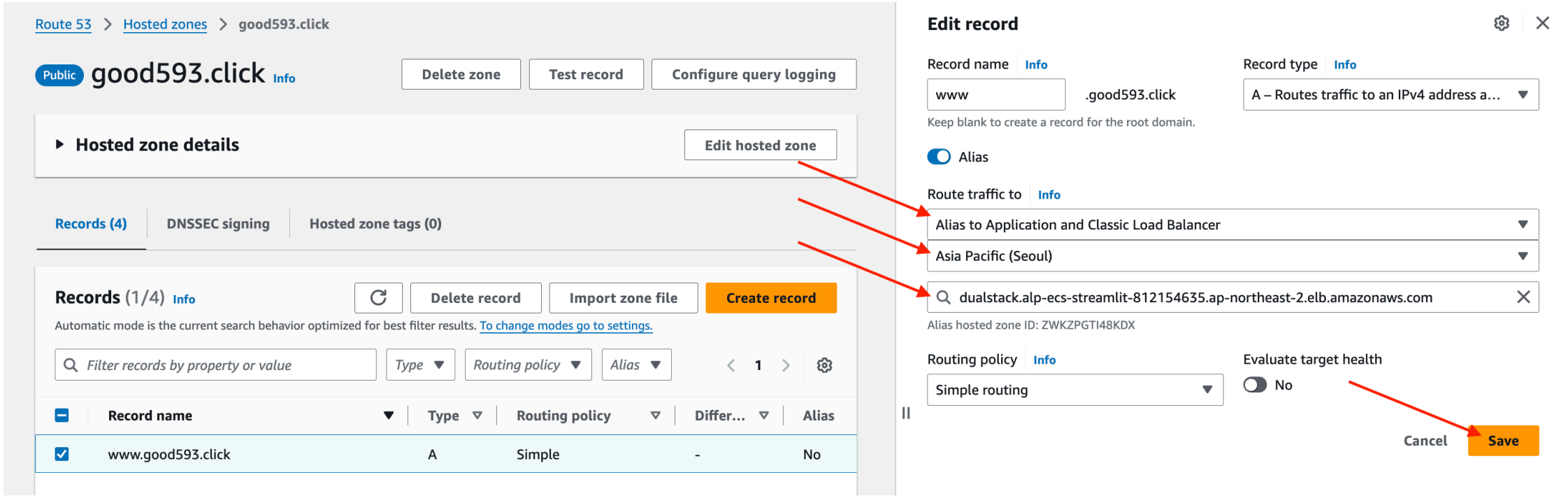


단계15: 결과 확인

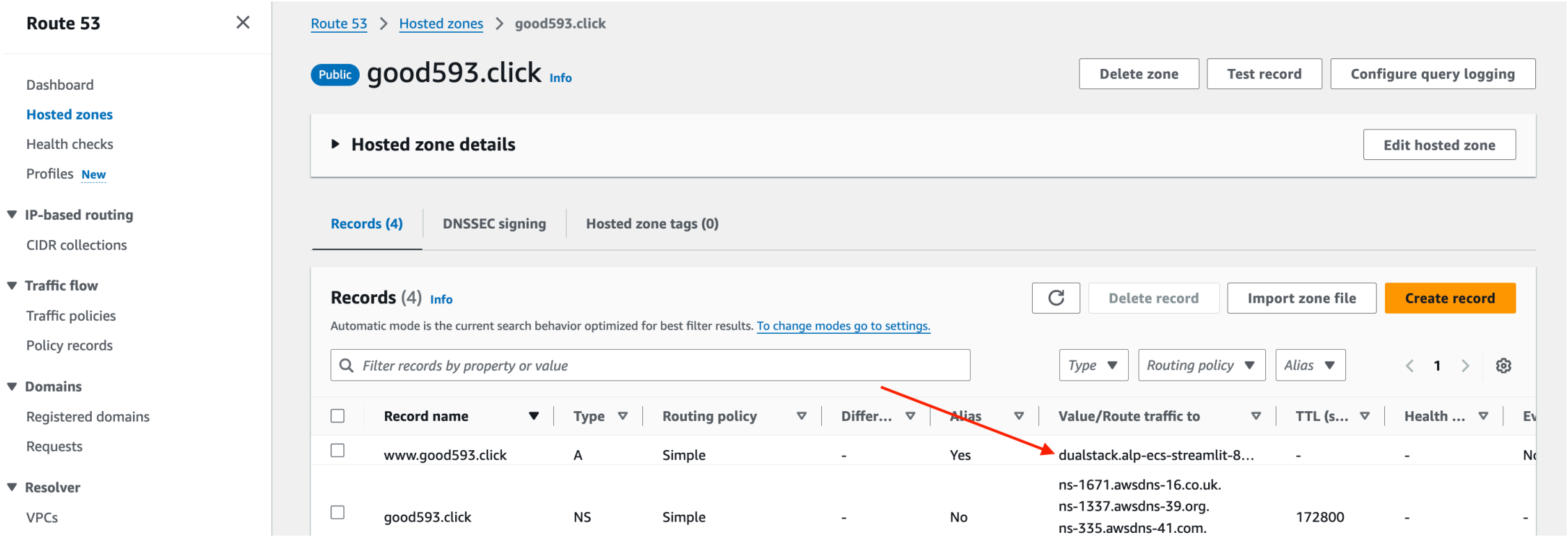

Listeners and rules
Network mapping
Resource map - new
Security
Monitoring
Integrations
Attributes
Tags
Resource map Info
Give feedback
View, explore, and troubleshoot your load balancer's architecture.
Overview
Unhealthy target map
Show resource details
alp-ecs-streamlit
Last fetched seconds ago
Export
Listeners (2)
HTTPS:443
1 rule
HTTP:80
1 rule
Rules (2)
Priority default
Forward to target group
Conditions (If)
If no other rule applies
Priority default
Redirect
Conditions (If)
If no other rule applies
Target groups (1) Info
IP
ecs-streamlit-tg
1 target
1
0
0
0
0
Targets (1)
172.31.15.169
Port 8501
Healthy

단계16: Route 53 > 레코드 등록/수정

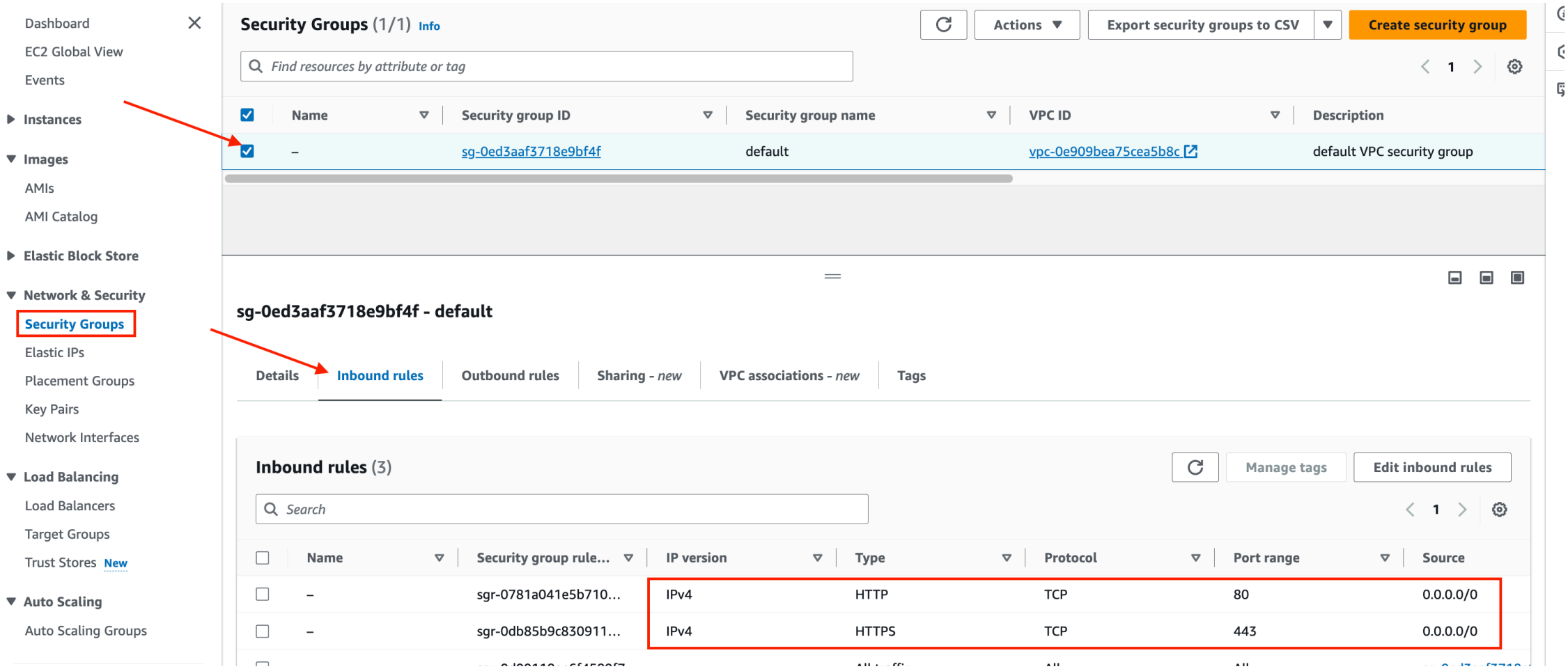



Route 53 > Hosted zones > good593.click
Public good593.click Info
Delete zone
Test record
Configure query logging
Hosted zone details
Edit hosted zone
Records (4)
DNSSEC signing
Hosted zone tags (0)
Records (1/4) Info
Delete record
Import zone file
Create record
Automatic mode is the current search behavior optimized for best filter results. To change modes go to settings.
Filter records by property or value
Type
Routing policy
Alias
1
Record name
Type
Routing policy
Differ...
Alias
www.good593.click
A
Simple
-
No
Edit record
Record name Info
www
.good593.click
Keep blank to create a record for the root domain.
Record type Info
A – Routes traffic to an IPv4 address a...
Alias
Route traffic to Info
Alias to Application and Classic Load Balancer
Asia Pacific (Seoul)
dualstack.alp-ecs-streamlit-812154635.ap-northeast-2.elb.amazonaws.com
Alias hosted zone ID: ZWKZPGTI48KDX
Routing policy Info
Simple routing
Evaluate target health
No
Cancel
Save

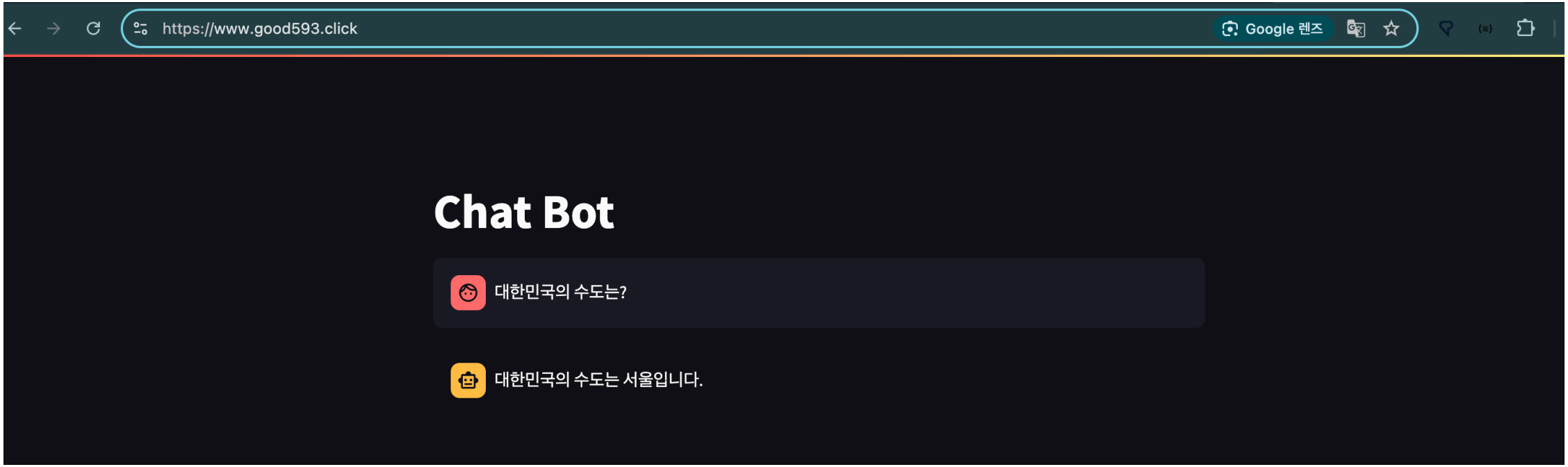
단계17: 결과 확인


Route 53
Dashboard
Hosted zones
Health checks
Profiles New
IP-based routing
CIDR collections
Traffic flow
Traffic policies
Policy records
Domains
Registered domains
Requests
Resolver
VPCs
Route 53 > Hosted zones > good593.click
Public good593.click Info
Delete zone
Test record
Configure query logging
Hosted zone details
Edit hosted zone
Records (4)
DNSSEC signing
Hosted zone tags (0)
Records (4) Info
Delete record
Import zone file
Create record
Automatic mode is the current search behavior optimized for best filter results. To change modes go to settings.
Filter records by property or value
Type
Routing policy
Alias
Record name
Type
Routing policy
Differ...
Alias
Value/Route traffic to
TTL (s...
Health ...
www.good593.click
A
Simple
-
Yes
dualstack.alp-ecs-streamlit-8...
-
-
good593.click
NS
Simple
-
No
ns-1671.awsdns-16.co.uk.
ns-1337.awsdns-39.org.
ns-335.awsdns-41.com.
172800
-
-

단계18: EC2 > Security Group > Inbound rules 추가

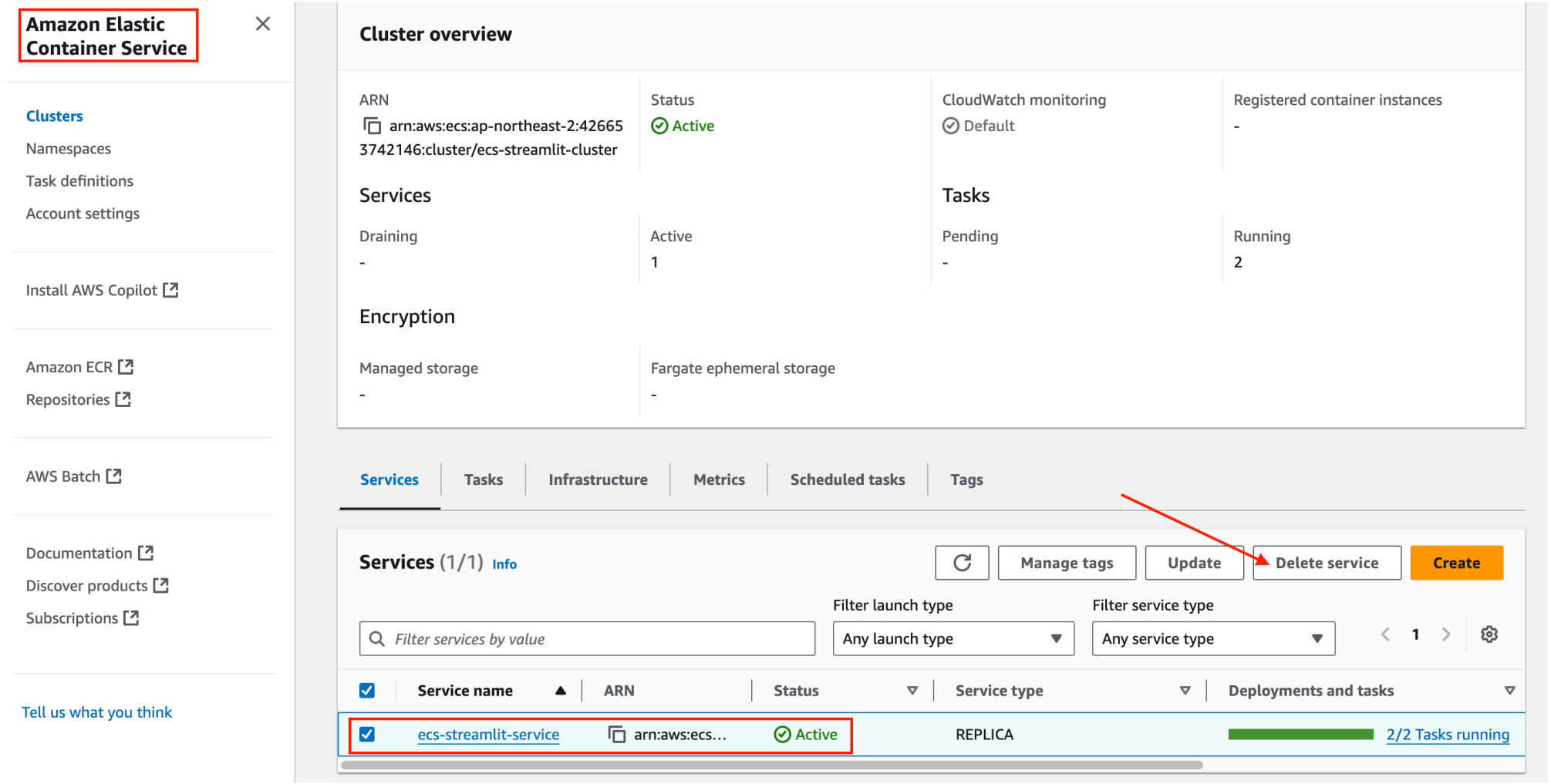


https://도메인

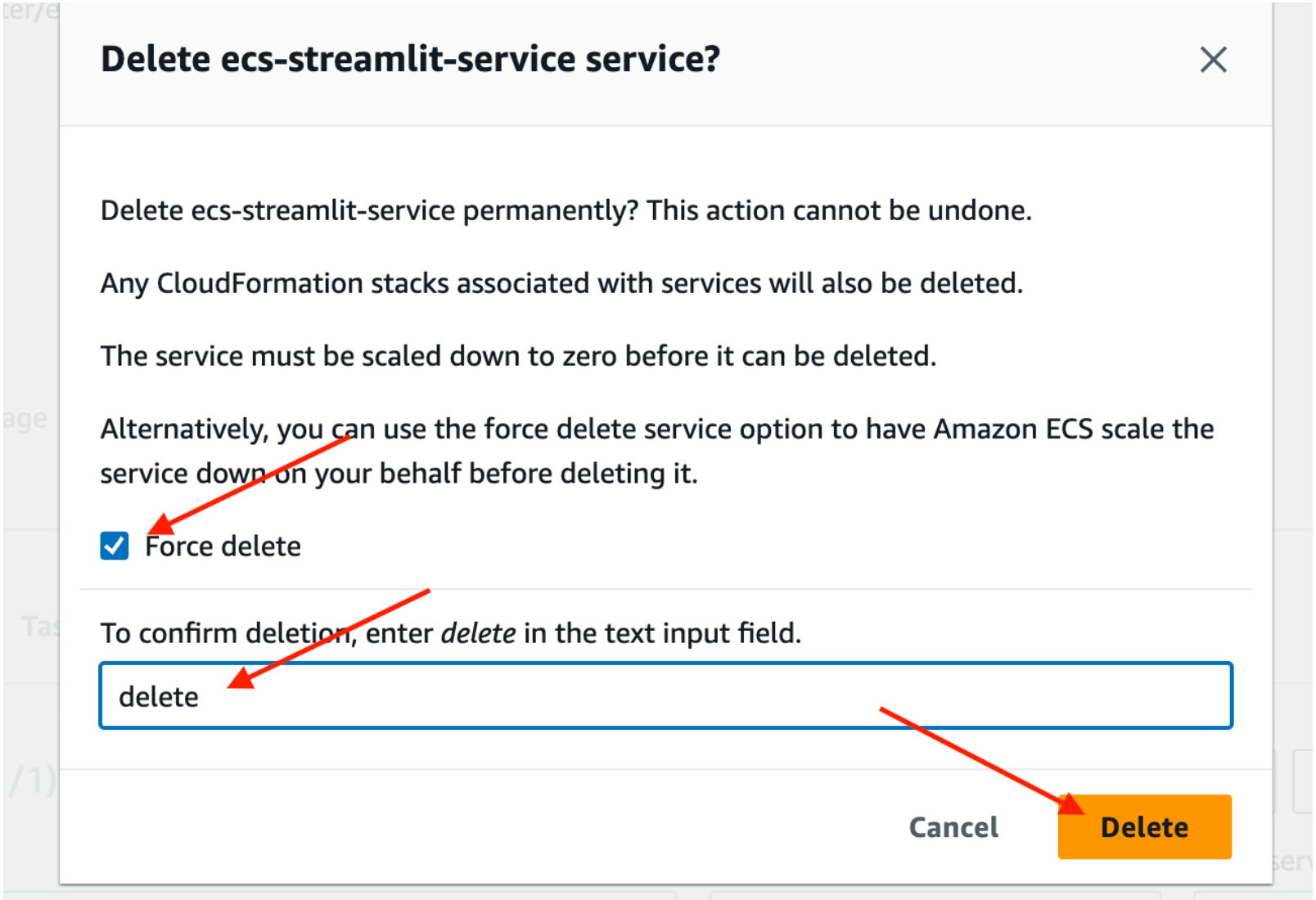

https://www.good593.click
Google 렌즈
Chat Bot
대한민국의 수도는?
대한민국의 수도는 서울입니다.

ECS Service with ALB

단계1: ECS > 기존 Service 선택

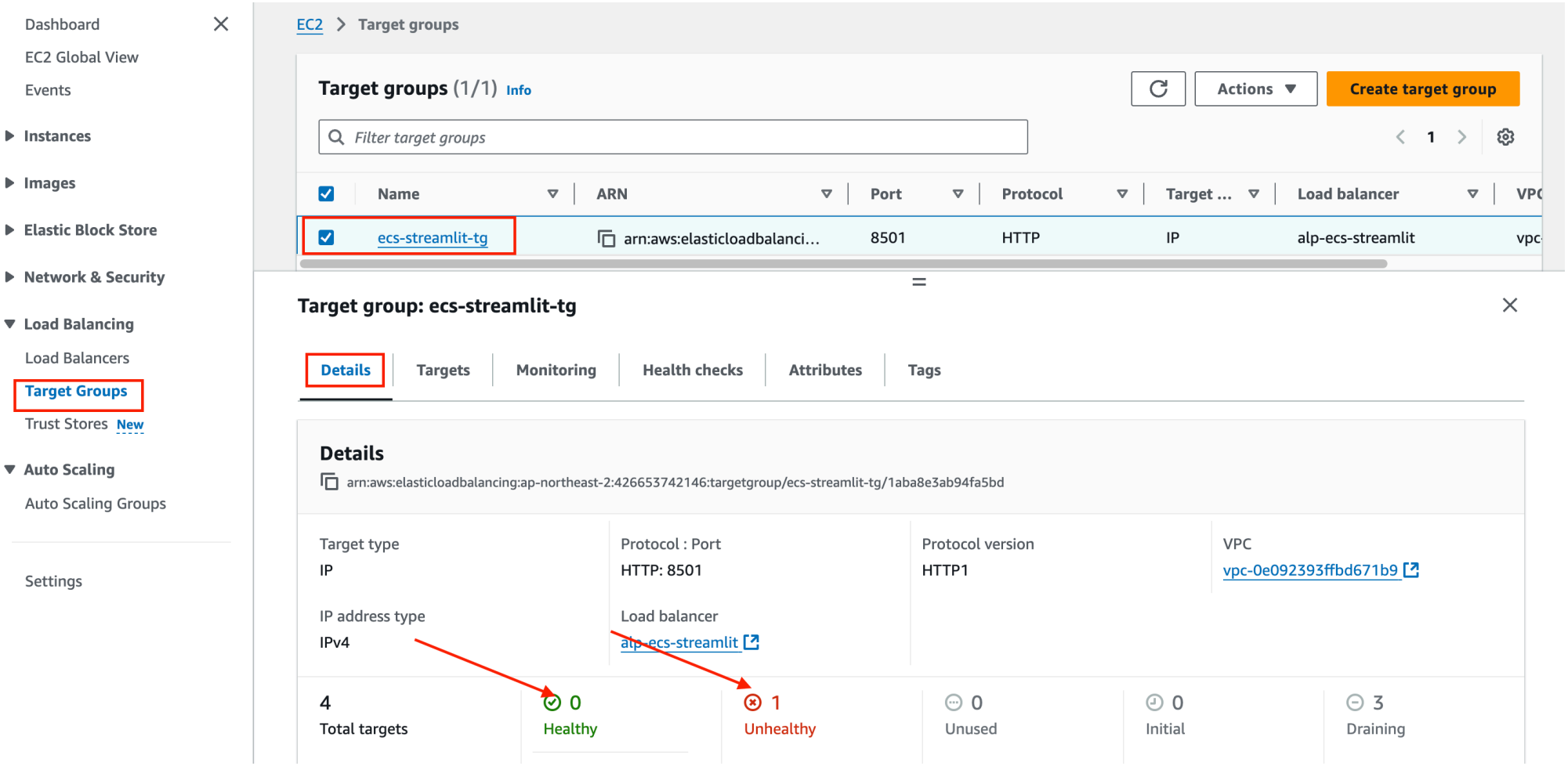


단계2: ECS > Service Delete

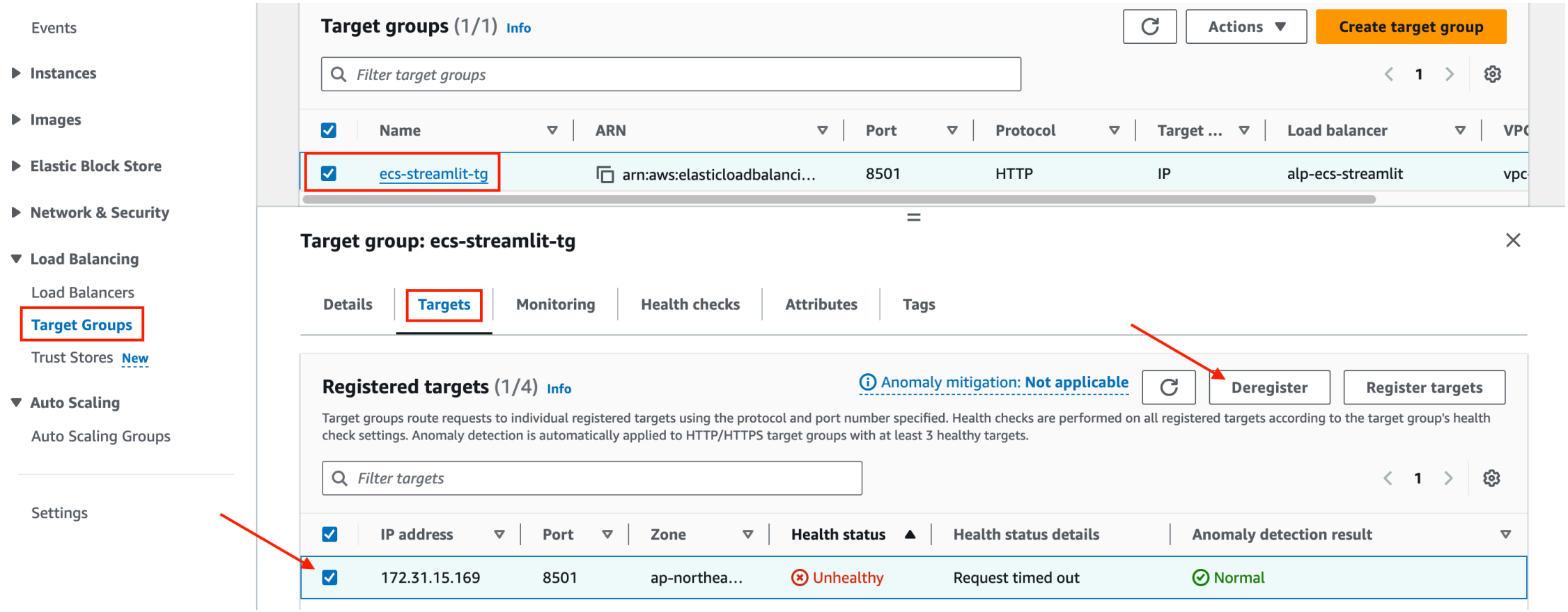


단계3: EC2 > Target Groups 확인

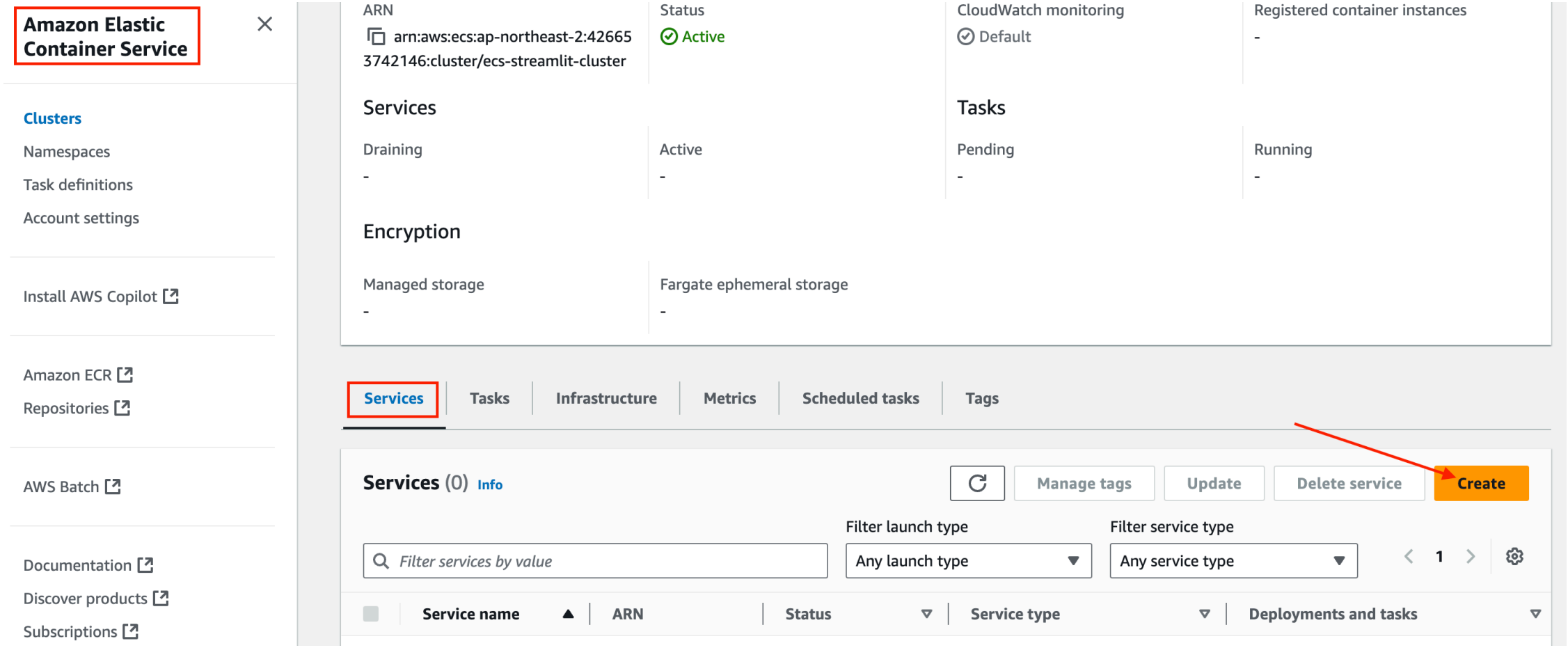
> 시간이 좀 걸림

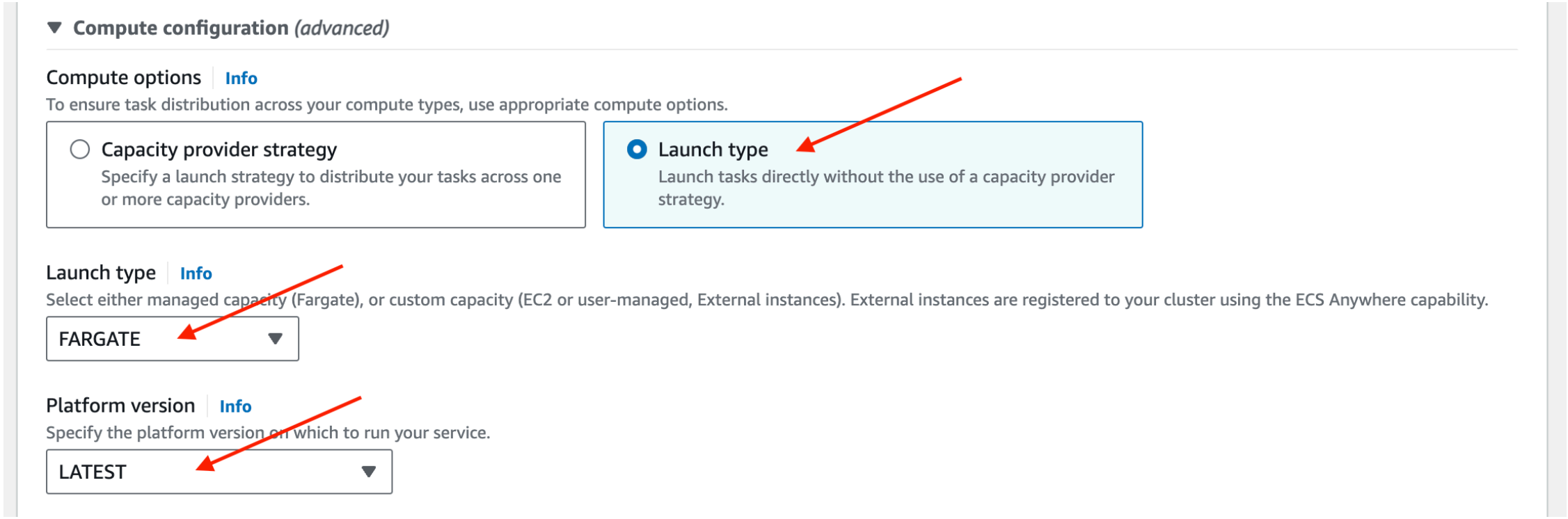


단계4: EC2 > 등록된 Target 삭제



단계5: ECS > Create Service




Compute configuration (advanced)
Compute options Info
To ensure task distribution across your compute types, use appropriate compute options.
Capacity provider strategy
Specify a launch strategy to distribute your tasks across one or more capacity providers.
Launch type
Launch tasks directly without the use of a capacity provider strategy.
Launch type Info
Select either managed capacity (Fargate), or custom capacity (EC2 or user-managed, External instances). External instances are registered to your cluster using the ECS Anywhere capability.
FARGATE
Platform version Info
Specify the platform version on which to run your service.
LATEST

Deployment configuration

Application type | Info

Specify what type of application you want to run.

- (•) **Service**
  Launch a group of tasks handling a long-running computing work that can be stopped and restarted. For example, a web application.
- ( ) **Task**
  Launch a standalone task that runs and terminates. For example, a batch job.

Task definition

Select an existing task definition. To create a new task definition, go to Task definitions.

- [ ] Specify the revision manually
  Manually input the revision instead of choosing from the 100 most recent revisions for the selected task definition family.

Family: ecs-streamlit-task

Revision: 2 (LATEST)

Service name

Assign a service name that is unique for this cluster.

ecs-streamlit-service

Up to 255 letters (uppercase and lowercase), numbers, underscores, and hyphens are allowed. Service names must be unique within a cluster.

Service type | Info

Specify the service type that the service scheduler will follow.

Replica

Place and maintain a desired number of tasks across your cluster.

Daemon

Place and maintain one copy of your task on each container instance.

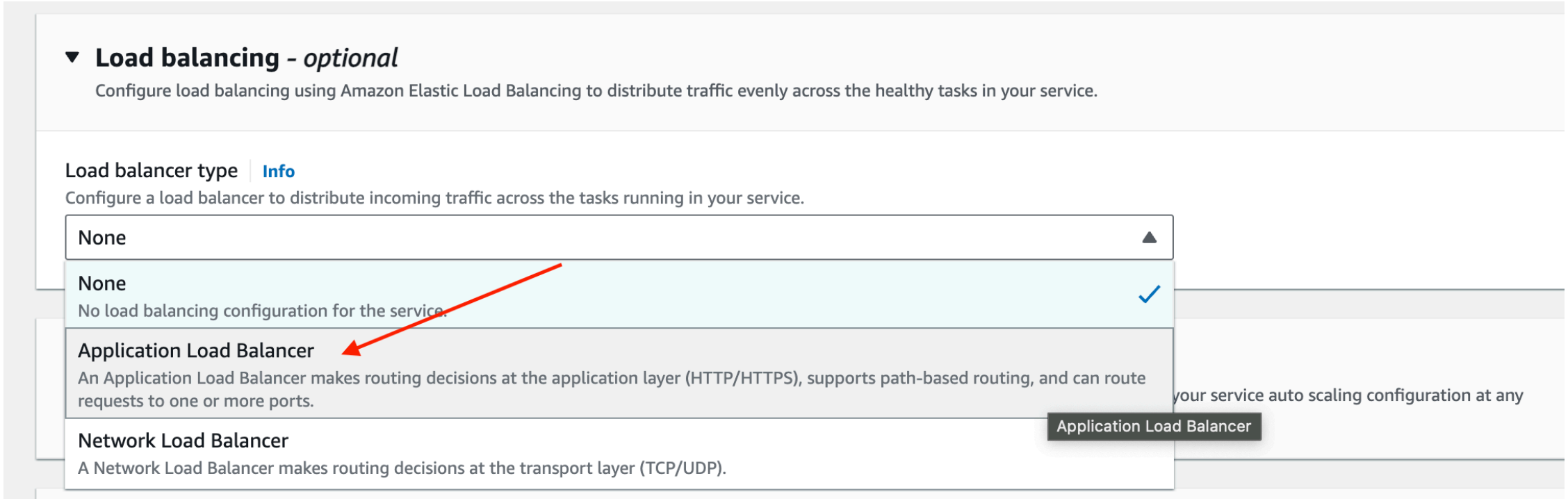
Desired tasks

Specify the number of tasks to launch.

2

▶ Deployment options

▶ Deployment failure detection Info


Load balancing - optional
Configure load balancing using Amazon Elastic Load Balancing to distribute traffic evenly across the healthy tasks in your service.
Load balancer type
Info
Configure a load balancer to distribute incoming traffic across the tasks running in your service.
None
None
No load balancing configuration for the service.
Application Load Balancer
An Application Load Balancer makes routing decisions at the application layer (HTTP/HTTPS), supports path-based routing, and can route requests to one or more ports.
Network Load Balancer
A Network Load Balancer makes routing decisions at the transport layer (TCP/UDP).
your service auto scaling configuration at any
Application Load Balancer

Load balancer type | Info

Configure a load balancer to distribute incoming traffic across the tasks running in your service.

Application Load Balancer

Container

The container and port to load balance the incoming traffic to

ecs-streamlit-container 8501:8501

Host port:Container port

Application Load Balancer

Specify whether to create a new load balancer or choose an existing one.

- ○ Create a new load balancer
- ◉ Use an existing load balancer

Load balancer

Select the load balancer you wish to use to distribute incoming traffic across the tasks running in your service.

alp-ecs-streamlit

Health check grace period | Info

0

seconds

Listener | Info

Specify the port and protocol that the load balancer will listen for connection requests on.

- ○ Create new listener
- ◉ Use an existing listener

Listener

443:HTTPS

Listener rules for 443:HTTPS (1)

Traffic received by the listener is routed according to its rules. Rules are evaluated in priority order, from the lowest value to the highest value. The default rule is evaluated last.

1

Evaluation order	Rule path	Target group
default	/	ecs-streamlit-tg

Target group | Info

Specify whether to create a new target group or choose an existing one that the load balancer will use to route requests to the tasks in your service.

- ○ Create new target group
- ◉ Use an existing target group

Target group name

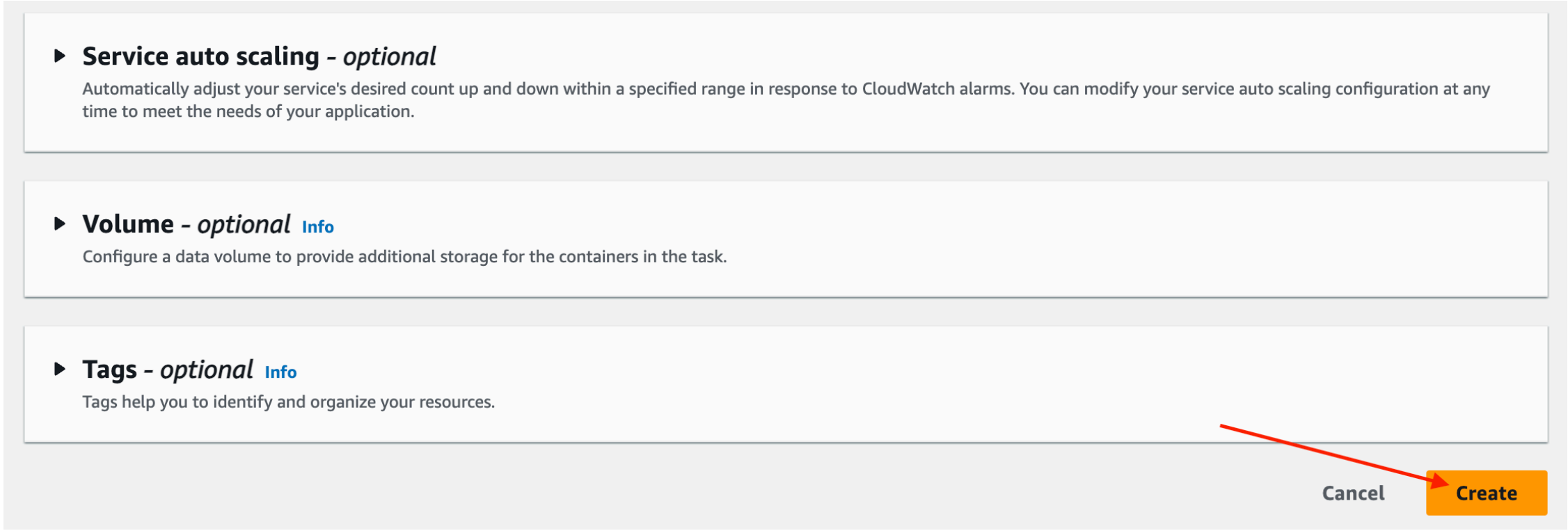
ecs-streamlit-tg

Health check path

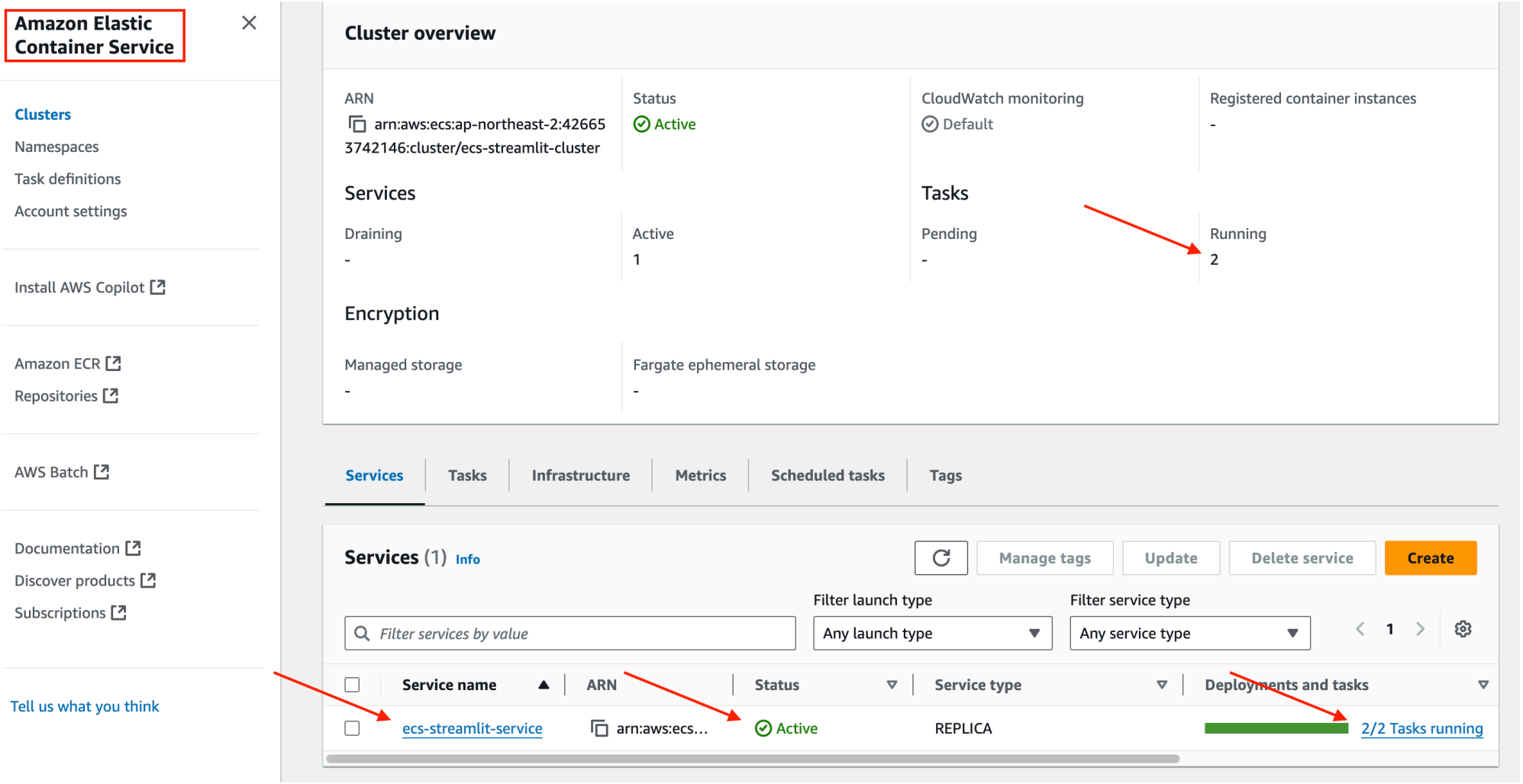
/

Health check protocol

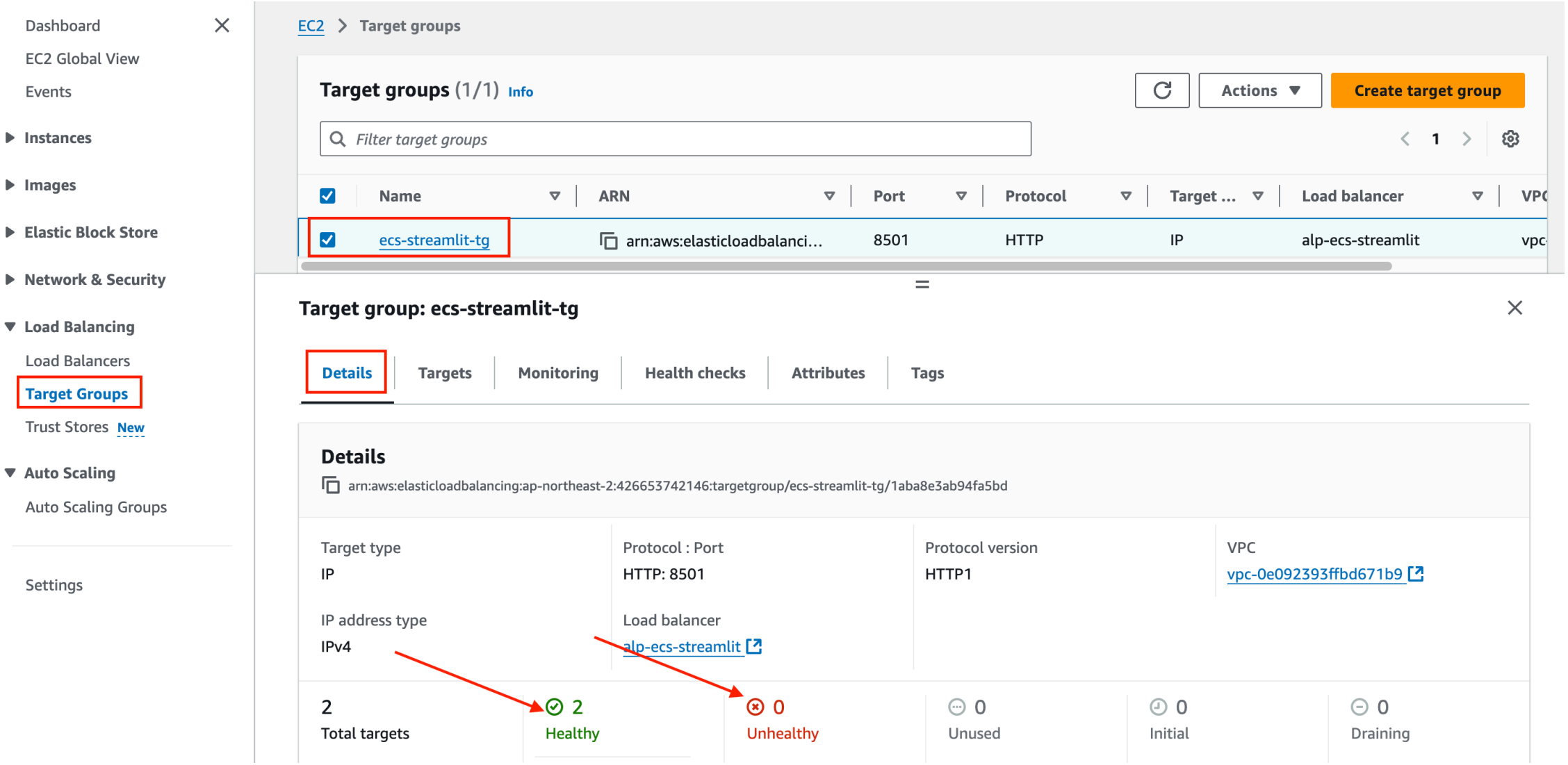
HTTP

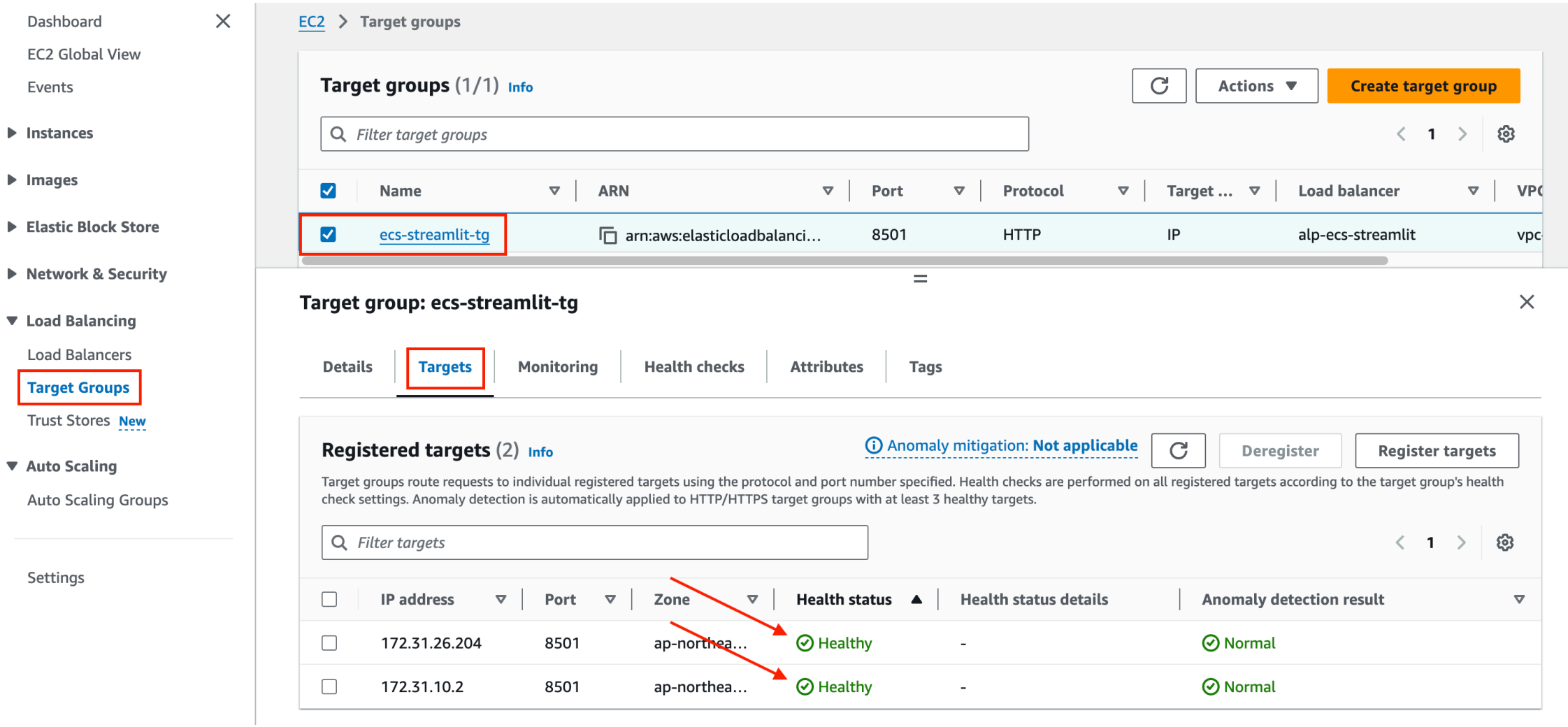

Service auto scaling - optional
Automatically adjust your service's desired count up and down within a specified range in response to CloudWatch alarms. You can modify your service auto scaling configuration at any time to meet the needs of your application.
Volume - optional Info
Configure a data volume to provide additional storage for the containers in the task.
Tags - optional Info
Tags help you to identify and organize your resources.
Cancel
Create

단계6: ECS > Service 결과 확인



단계7: EC2 > Target Groups 확인




Dashboard
EC2 Global View
Events
Instances
Images
Elastic Block Store
Network & Security
Load Balancing
Load Balancers
Target Groups
Trust Stores New
Auto Scaling
Auto Scaling Groups
Settings
EC2 > Target groups
Target groups (1/1) Info
Actions
Create target group
Filter target groups
1
Name
ARN
Port
Protocol
Target ...
Load balancer
ecs-streamlit-tg
arn:aws:elasticloadbalanci...
8501
HTTP
IP
alp-ecs-streamlit
Target group: ecs-streamlit-tg
Details
Targets
Monitoring
Health checks
Attributes
Tags
Registered targets (2) Info
Anomaly mitigation: Not applicable
Deregister
Register targets
Target groups route requests to individual registered targets using the protocol and port number specified. Health checks are performed on all registered targets according to the target group's health check settings. Anomaly detection is automatically applied to HTTP/HTTPS target groups with at least 3 healthy targets.
Filter targets
1
IP address
Port
Zone
Health status
Health status details
Anomaly detection result
172.31.26.204
8501
ap-northea...
Healthy
-
Normal
172.31.10.2
8501
ap-northea...
Healthy
-
Normal

단계8: https 접속

> https://도메인