

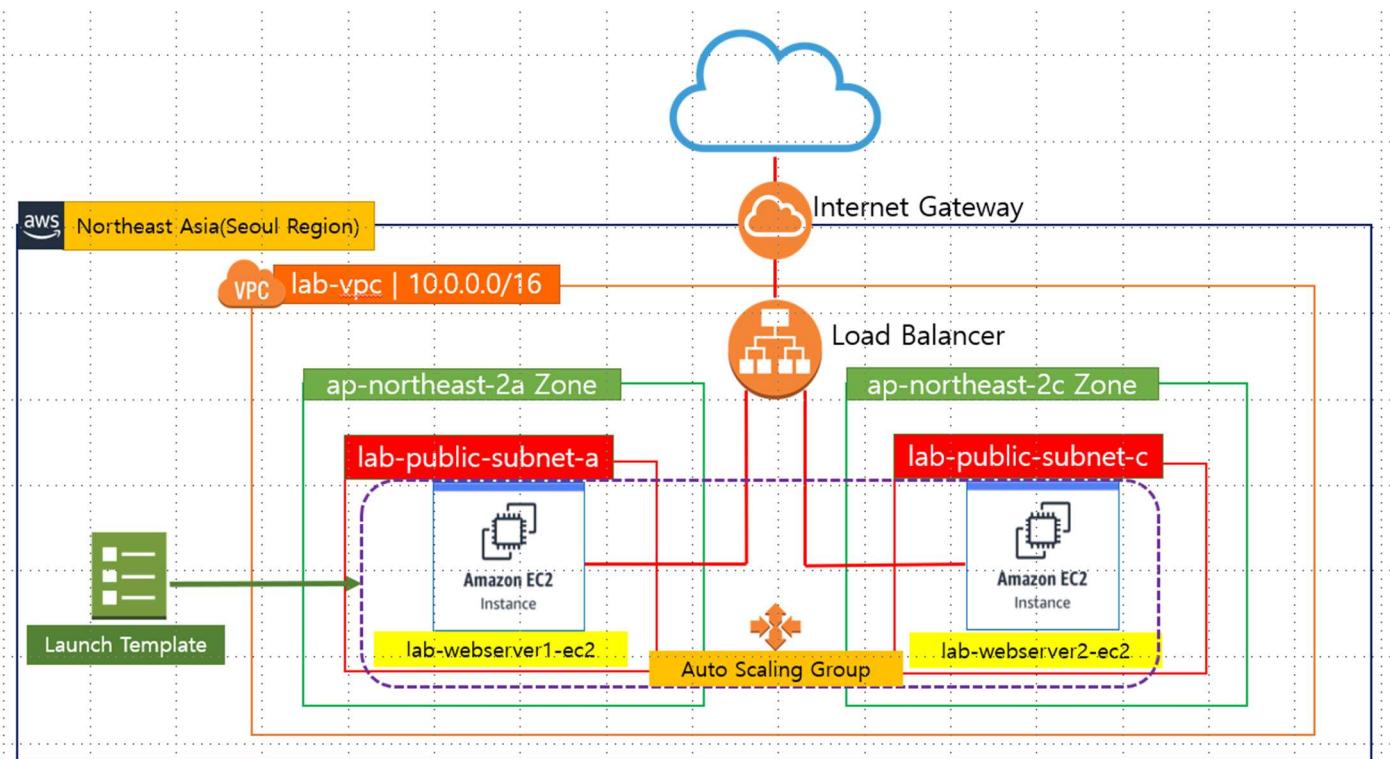
Lab. Understanding Amazon EC2 Auto Scaling

목적

Amazon EC2 Auto Scaling을 통해 애플리케이션의 로드를 처리할 수 있는 정확한 수의 Amazon EC2 인스턴스를 보유하도록 보장할 수 있다. Auto Scaling 그룹이라는 EC2 인스턴스 모음을 생성한다. 각 Auto Scaling 그룹의 최소/최대 인스턴스 수를 지정할 수 있으며, 그룹의 크기가 최소/최대 사이 값을 유지한다. 조정 정책을 지정했다면 Amazon EC2 Auto Scaling에서는 애플리케이션의 늘어나거나 줄어드는 수요에 따라 인스턴스를 시작하거나 종료할 수 있다.

사전 준비물

AWS Free-Tier 계정

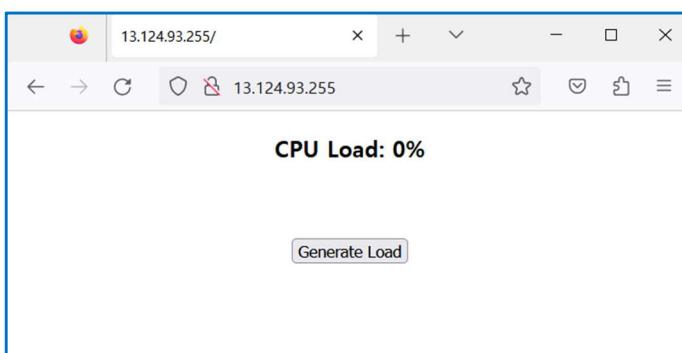


웹 서버 생성 및 AMI 생성하기

1. Auto Scaling 실습을 위한 EC2 인스턴스를 다음과 같이 새로 생성한다.
 - A. [이름] : lab-webserver1-ec2
 - B. [AMI] : Amazon Linux 2 AMI (HVM) – Kernel 5.10, SSD Volume Type, 64비트(x86)
 - C. [인스턴스 유형] : t2.micro
 - D. [키 페어(로그인)] > [새 키 페어 생성] : lab-webserver-ec2-key
 - E. [네트워크 설정] > [편집]
 - i. [VPC] : lab-vpc
 - ii. [서브넷] : lab-subnet-public1-ap-northeast-2a
 - iii. [퍼블릭 IP 자동 할당] : 활성화
 - iv. [방화벽(보안 그룹)] > [보안 그룹 생성]
 - v. [보안 그룹 이름] : lab-autoscaling-sg
 - vi. [인바운드 보안 그룹 규칙] : ssh:22 / 위치무관, http:80 / 위치 무관
 - F. [스토리지 구성] : 8 GiB, gp2
 - G. [고급 세부 정보] > [사용자 데이터] : userdata.txt

```
1 #!/bin/sh
2 yum -y install httpd php mysql php-mysql
3 chkconfig httpd on
4 systemctl enable httpd.service
5 systemctl start httpd.service
6 cd /tmp
7 wget https://henry-28day-lab-bucket.s3.ap-northeast-2.amazonaws.com/autoscaling-source.zip
8 unzip autoscaling-source.zip
9 sudo mv *.php /var/www/html
```

2. EC2 인스턴스 생성을 확인한다.
3. 방금 생성한 EC2 인스턴스의 요약 페이지로 이동하여 [퍼블릭 IPv4 주소]를 복사한다.
4. 웹 브라우저를 열고 방금 복사한 IP를 붙여넣기 하여 웹 페이지가 정상적으로 작동하는지 확인한다.



5. 인스턴스 생성 후 [작업] > [이미지 및 템플릿] > [이미지 생성]을 클릭한다.
6. [이미지 이름]을 **webserver-as-ami**로 설정한 후, 나머지 값은 기본값으로 놓고, [이미지 생성] 버튼을 클릭한다.
7. 좌측 메뉴의 [이미지] > [AMI]에서 확인할 수 있다.

The screenshot shows the AWS CloudFormation console with the following details:

- AMI ami-06aee060a78bcf4e 등록 취소를 완료했습니다.**
- Amazon Machine Images(AMI) (1) 정보**
- 내 소유** (선택)
AMI를 속성 또는 태그로 찾기
- AMI ID**: ami-03fa786da...
- AMI 이름**: webserver-...
- 원본**: 789534828...
- 소유자**: 789534828...
- 표시 여부**: 프라이빗
- 상태**: 사용 가능
- 생성 날짜**: 2023/10/17 09:05 GMT+9
- 플랫폼**: Linux/UNIX

시작 템플릿 생성

1. 오토스케일링 그룹을 생성하기 위해서는 시작 템플릿 또는 시작 구성이 준비되어야 한다. 이번 실습에서는 시작 템플릿을 사용한다. EC2 콘솔의 시작 템플릿 메뉴로 들어가 [시작 템플릿 생성] 버튼을 클릭한다.



2. 시작 템플릿 생성 페이지에서 [시작 템플릿 이름]에 webserver-template, [템플릿 버전 설명]에 Template for Webserver Autoscaling를 입력하고, [Auto Scaling 지침] 항목에 체크한다.

시작 템플릿 생성

시작 템플릿을 생성하면 저장된 인스턴스 구성을 만들어 두었다가 나중에 이를 재사용하고, 공유하고, 시작할 수 있습니다. 여러 버전의 템플릿을 저장할 수 있습니다.

시작 템플릿 이름 및 설명

시작 템플릿 이름 - 필수
webserver-template

이 계정에 대해 고유해야 합니다. 최대 128자입니다. '&', '*', '@' 등의 특수 문자나 공백은 사용할 수 없습니다.

템플릿 버전 설명
Template for Webserver Autoscaling

최대 255자

Auto Scaling 지침 정보
EC2 Auto Scaling에 이 템플릿을 사용하려면 이 항목을 선택합니다.
 EC2 Auto Scaling에 사용할 수 있는 템플릿을 설정하는 데 도움이 되는 지침 제공

▶ 템플릿 태그
▶ 원본 템플릿

3. [시작 템플릿 콘텐츠] 섹션에서 [애플리케이션 및 OS 이미지(Amazon Machine Image) – 필수] 항목에서 [내 AMI] 탭을 선택하고, [내 소유]를 선택하여 위에서 만들어 놓은 webserver-as-ami를 선택한다.

시작 템플릿 콘텐츠

아래에서 시작 템플릿의 세부 정보를 지정합니다. 비워 둔 필드는 시작 템플릿에 포함되지 않습니다.

▼ 애플리케이션 및 OS 이미지(Amazon Machine Image) - 필수 정보

AMI는 인스턴스를 시작하는 데 필요한 소프트웨어 구성(운영 체제, 애플리케이션 서버 및 애플리케이션)이 포함된 템플릿입니다. 아래에서 찾고 있는 항목이 보이지 않으면 AMI를 검색하거나 찾아보십시오.

수천 개의 애플리케이션 및 OS 이미지를 포함하는 전체 카탈로그 검색

최근 사용 내 AMI Quick Start

내 소유 나와 공유 됨

더 많은 AMI 찾아보기

AWS, Marketplace 및 커뮤니티의 AMI 포함

Amazon Machine Image(AMI)

webserver-as-ami
ami-065ff25a1e9b3842c
2022-04-15T05:54:42.000Z 가상화: hvm ENA 활성화됨: true 루트 디바이스 유형: ebs

설명

-

아키텍처 AMI ID
x86_64 ami-065ff25a1e9b3842c

4. [인스턴스 유형] 섹션에서 [인스턴스 유형]에는 **t2.micro**를 선택한다. 이번 실습은 서비스용 웹 서버를 올리는 것으로 SSH 접근을 하지 않을 예정이어서 키 페어는 사용하지 않는다.

▼ 인스턴스 유형 정보 어드밴스드

인스턴스 유형

t2.micro 프리 티어 사용 가능

패밀리: t2 1 vCPU 1 GiB 메모리 온디맨드 Linux 요금: 0.0144 USD 시간당
온디맨드 Windows 요금: 0.019 USD 시간당

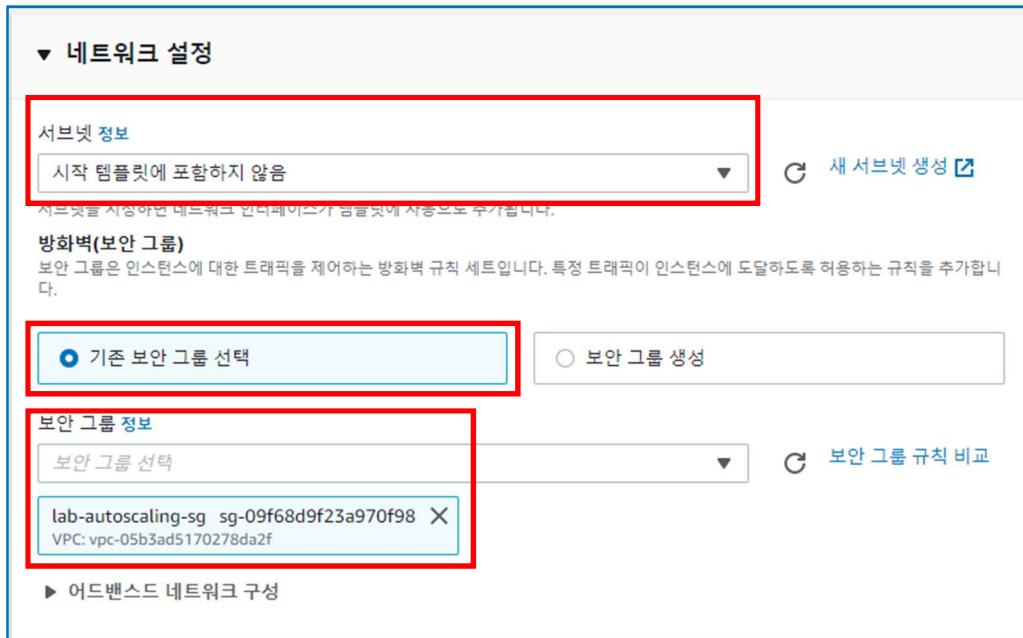
인스턴스 유형 비교

▼ 키 페어(로그인) 정보

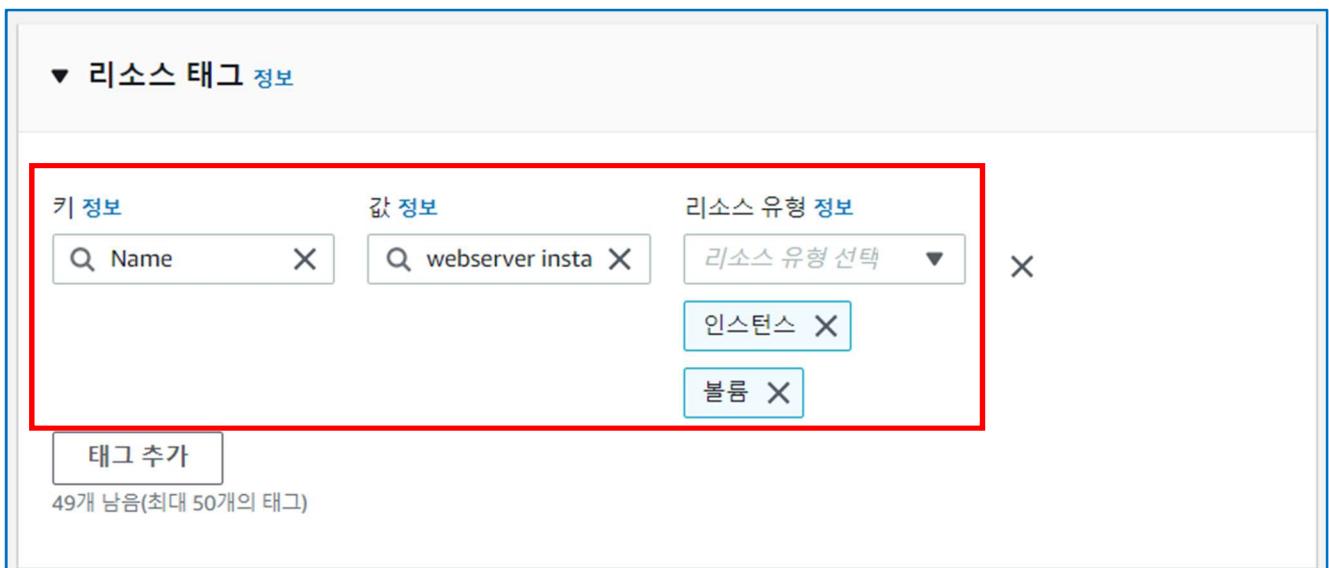
키 페어를 사용하여 인스턴스에 안전하게 연결할 수 있습니다. 인스턴스를 시작하기 전에 선택한 키 페어에 대한 액세스 권한이 있는지 확인하세요.

키 페어 이름 시작 템플릿에 포함하지 않음 새 키 페어 생성

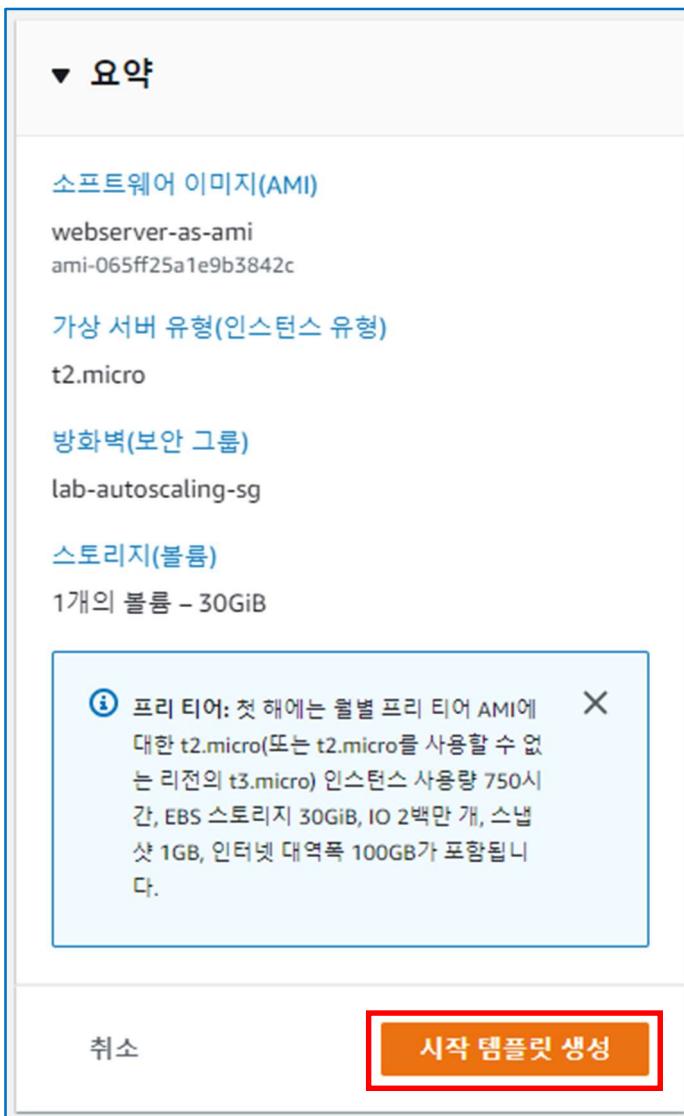
5. [네트워크 설정] 섹션에서는 기본값 그대로 놓고, [방화벽(보안 그룹)]은 [기존 보안 그룹 선택]을 선택하고, [보안 그룹] 목록에서 실습에서 생성한 webserver-autoscaling-sg를 선택한다.



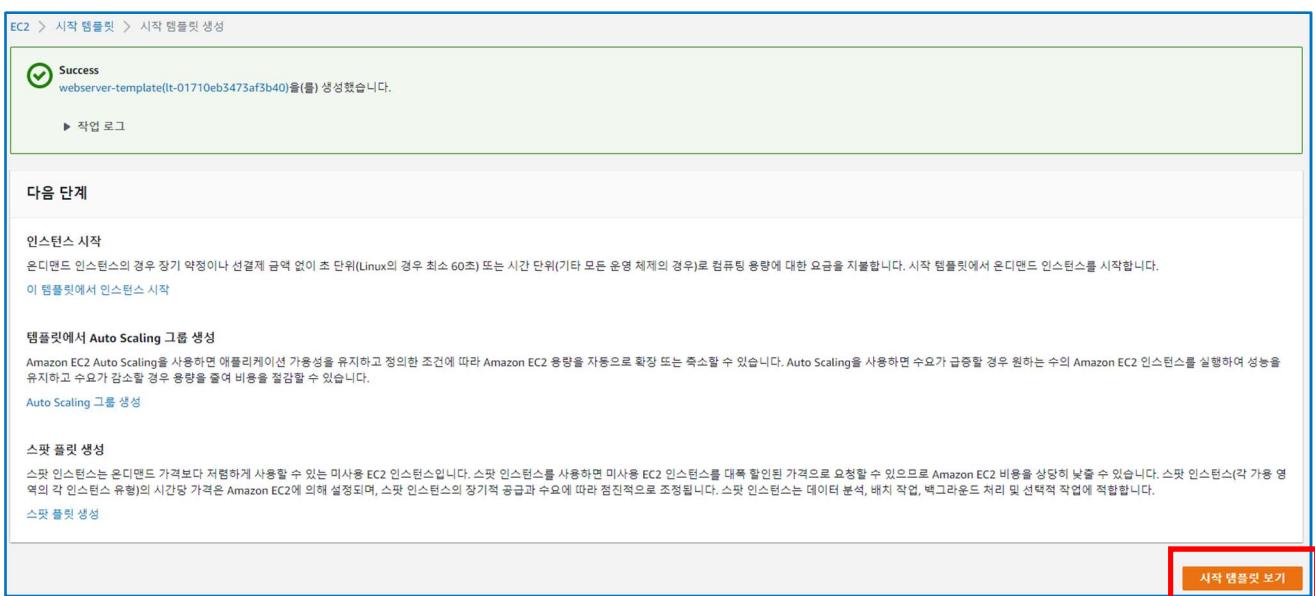
6. [리소스 태그] 섹션에서 [키]는 Name으로, [값]은 webserver instance로, [리소스 유형]은 인스턴스, 볼륨을 선택한다.



7. [요약] 섹션에서 이제까지 설정한 내용을 다시 한번 확인하고, [시작 템플릿 생성] 버튼을 클릭한다.



8. 다음 단계 지침을 숙지한 후, 페이지 우측 하단의 [시작 템플릿 보기] 버튼을 클릭한다.

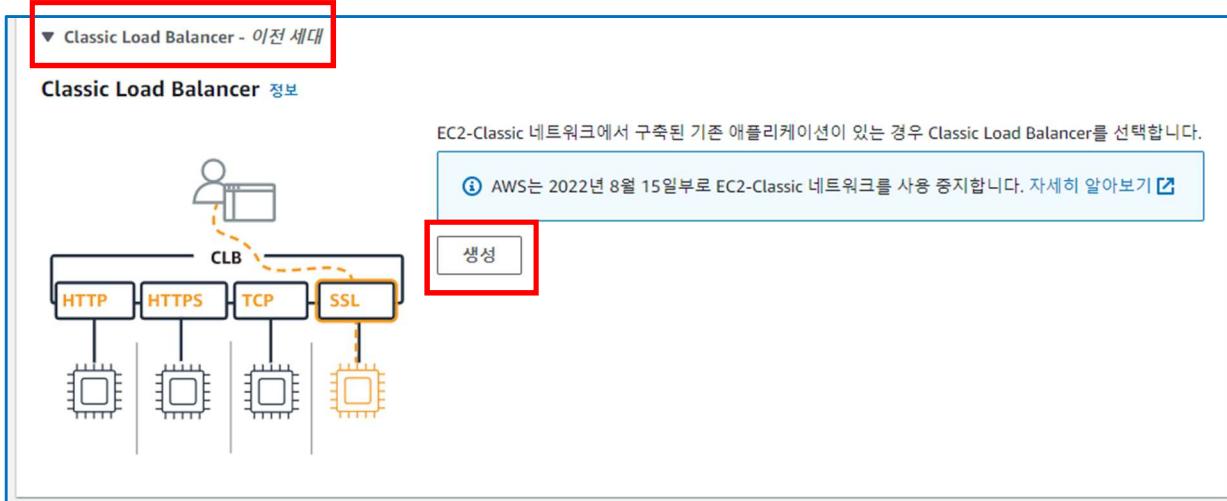


9. [시작 템플릿]이 잘 생성되었다.

시작 템플릿 (1) 정보					
시작 템플릿 ID	시작 템플릿 이름	기본 버전	최신 버전	생성 시간	
lt-0c05fa93c9e480c7f	webserver-template	1	1	2022-04	

Class Load Balancer 생성

1. 좌측 메뉴 중 [로드 밸런싱] > [로드밸런서]를 선택하여 [로드 밸런서 생성] 버튼을 클릭한다.
2. [로드 밸런서 유형 선택] 페이지에서 계속 아래로 페이지를 스크롤다운하여 [Classic Load Balancer – 이전 세대]를 클릭한다. [생성] 버튼을 클릭한다.



3. [Classic Load Balancer 생성] 페이지에서 다음과 같이 설정한다.
 - A. [Load Balancer 이름] : lab-autoscaling-clb
 - B. [VPC] : lab-vpc
 - C. [매핑] : ap-northeast-2a/lab-subnet-public1-ap-northeast-2a, ap-northeast-2c/lab-subnet-public2-ap-northeast-2c
 - D. [보안 그룹] : lab-autoscaling-sg
 - E. [인스턴스 추가] > [인스턴스 추가] 버튼 클릭 : 위에서 생성한 lab-webserver1-ec2 선택 후 [확인] 버튼 클릭



- F. [로드 밸런서 생성] 버튼 클릭

4. 방금 생성한 **Classic Load Balancer**의 요약 페이지를 볼 수 있다.

The screenshot shows the AWS EC2 Load Balancers page. At the top, there is a search bar labeled "Filter 로드 밸런서". To the right of the search bar are buttons for "작업" (Work) and "로드 밸런서 생성" (Create Load Balancer). Below the search bar is a table header with columns: 이름 (Name), DNS 이름 (DNS Name), 상태 (Status), VPC ID, 가용 영역 (Available Regions), 유형 (Type), and 생성된 날짜 (Created Date). A single row is selected in the table, corresponding to the load balancer "lab-autoscaling-clb". The table row includes the name, DNS name (lab-autoscaling-clb-11943...), status (-), VPC ID (vpc-02ff6d53ce26bb1c6), 2 가용 영역 (2 Available Regions), classic type, and creation date (2023년 10월 17일, 09:00 (UTC+09:00)).

Below the table, a section titled "로드 밸런서: lab-autoscaling-clb" is displayed. It contains tabs for "세부 정보" (Detailed Information), "리스너" (Listeners), "네트워크 매핑" (Network Mapping), "보안" (Security), "상태 검사" (Health Check), "대상 인스턴스" (Target Instances), "모니터링" (Monitoring), "속성" (Attributes), and "태그" (Tags). The "세부 정보" tab is currently selected.

In the "세부 정보" section, there are two main sections: "로드 밸런서 유형" (Load Balancer Type) and "체계" (System). Under "로드 밸런서 유형", it shows "Classic" and "서비스 중인 인스턴스 0/1". Under "체계", it shows "Internet-facing" and "호스팅 영역 ZWKZPGTI48KDX". To the right of these sections, there are fields for "VPC" (vpc-02ff6d53ce26bb1c6), "가용 영역" (subnet-00741cb2ca85f0f48), and "생성된 날짜" (2023년 10월 17일, 09:21 (UTC+09:00)).

오토스케일링 그룹 생성

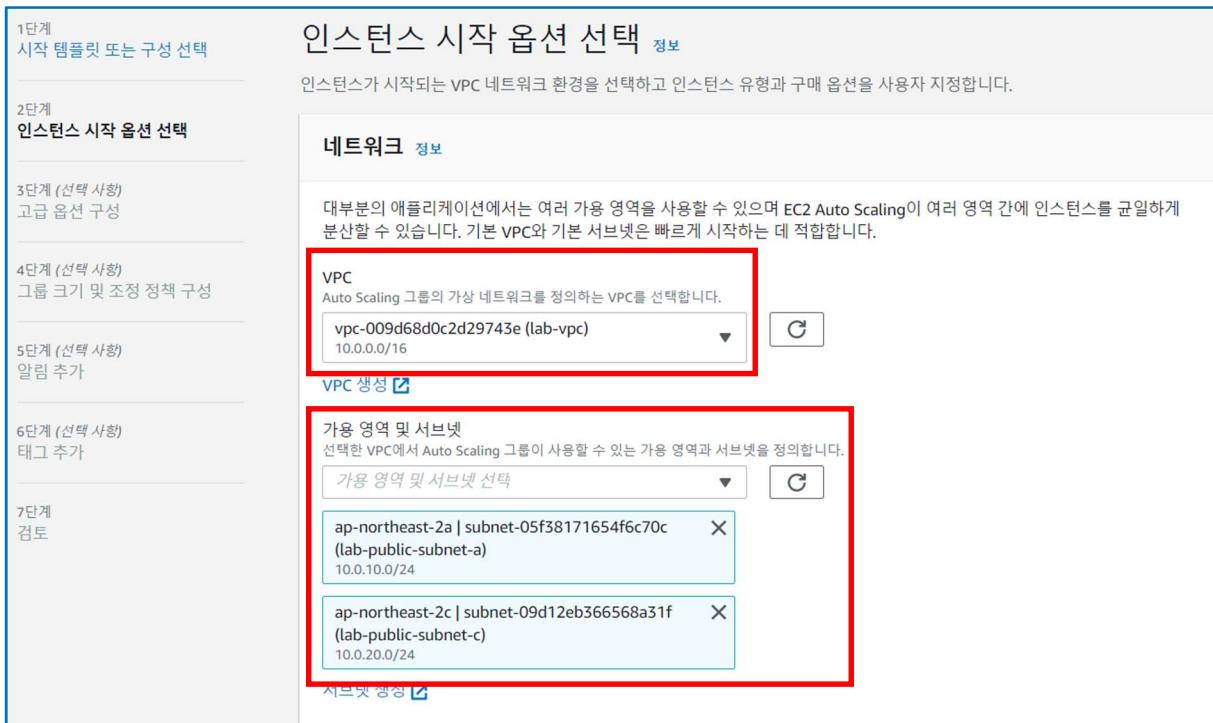
- 좌측 메뉴 중 [Auto Scaling] > [Auto Scaling 그룹]을 선택한 후, [Auto Scaling 그룹 생성] 버튼을 클릭한다.

The screenshot shows the 'Auto Scaling Group Creation' wizard. On the left sidebar, there are several sections: AMI (새로 만들기), Elastic Block Store (볼륨 새로 만들기, 스냅샷 새로 만들기), Auto Scaling (대상 그룹 새로 만들기), Auto Scaling (시작 구성), and Auto Scaling 그룹 (selected). The main content area has a title 'Amazon EC2 Auto Scaling' with a sub-section '애플리케이션의 가용성을 유지하는 데 유용함'. It includes a diagram titled '작동 방식' showing an 'Auto Scaling group' with four EC2 instances. A callout box points to the 'Auto Scaling 그룹 생성' button at the top right of the main panel. To the right, there's a '요금' section with information about costs.

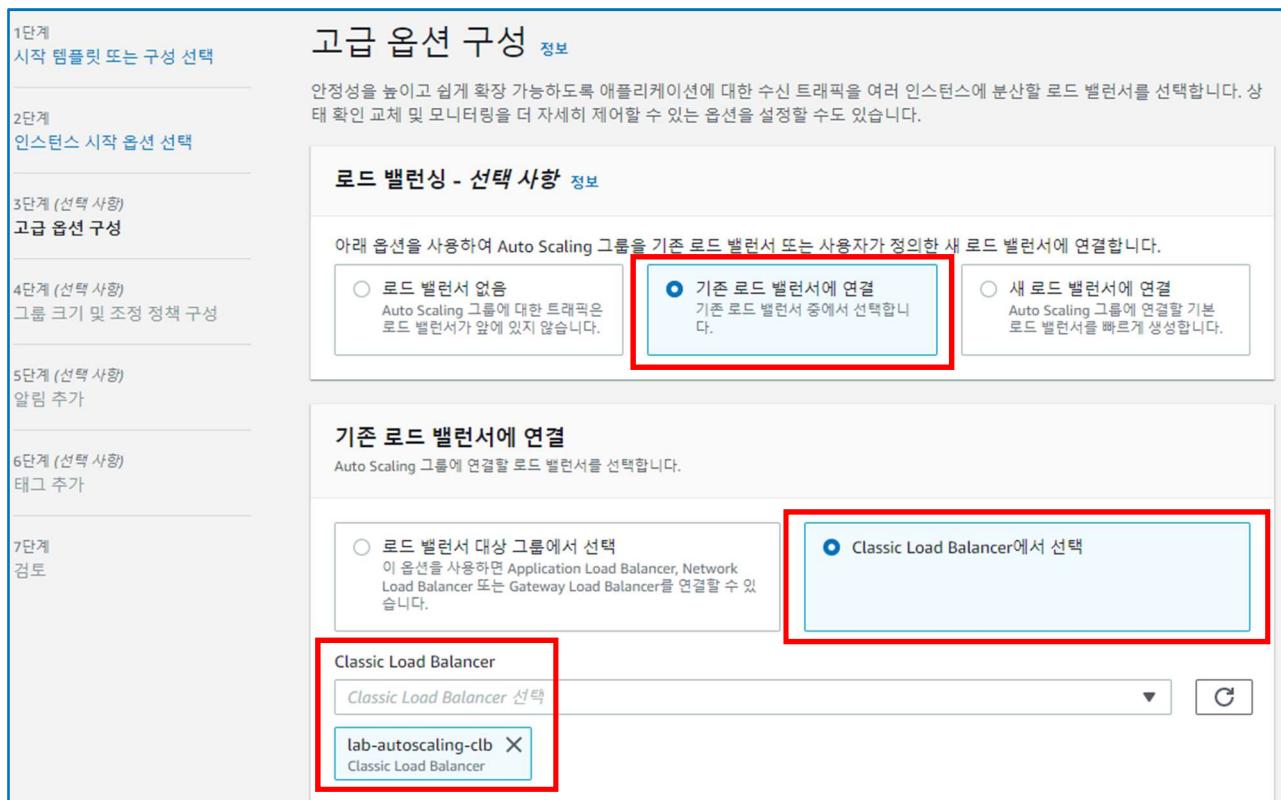
- [1단계: 시작 템플릿 또는 구성 선택]에서 [Auto Scaling 그룹 이름]에 **webserver-asg**를 입력하고, 아래의 [시작 템플릿]에서 위에서 생성한 **시작 템플릿(webserver-template)**을 선택하고 [다음] 버튼을 클릭한다.

The screenshot shows the 'Start Template or Configuration Selection' step. On the left sidebar, there are seven numbered steps. Step 1 is '1단계 시작 템플릿 또는 구성 선택' (selected). Step 2 is '2단계 인스턴스 시작 옵션 선택'. Step 3 is '3단계 (선택 사항) 고급 옵션 구성'. Step 4 is '4단계 (선택 사항) 그룹 크기 및 조정 정책 구성'. Step 5 is '5단계 (선택 사항) 알림 추가'. Step 6 is '6단계 (선택 사항) 태그 추가'. Step 7 is '7단계 검토'. The main content area has a title '시작 템플릿 또는 구성 선택' with a sub-section '정보'. It shows the 'Auto Scaling 그룹 이름' input field containing 'webserver-asg' (highlighted with a red box). Below it is a note: '현재 디자인에서 이 경계에 대해서 고유해야 하며 255자를 넘지 않아야 합니다.' In the bottom section, there's a '시작 템플릿 정보' panel with a sub-section '시작 템플릿' containing the text 'Amazon Machine Image(AMI), 인스턴스 유형, 키 페어 및 보안 그룹과 같은 인스턴스 수준 설정이 포함된 시작 템플릿을 선택합니다.' and a dropdown menu set to 'webserver-template' (highlighted with a red box). There are also '시작 템플릿 생성' and '버전' buttons.

3. [2단계: 인스턴스 시작 옵션 선택]에서 [네트워크] 섹션에서, [VPC]는 lab-vpc로, [가용 영역 및 서브넷]은 lab-subnet-public1-ap-northeast-2a와 lab-subnet-public2-ap-northeast-2c를 선택한 후, [다음] 버튼을 클릭합니다.



4. [3단계: 고급 옵션 구성]에서 [로드 밸런싱] 섹션은 [기존 로드 밸런서에 연결]을 선택하고, 자동으로 로딩된 [기존 로드 밸런서에 연결] 섹션에서 **Classic Load Balancer**에서 선택을 선택하면, [Classic Load Balancer] 목록에서 선택할 수 있는데, 위에서 생성한 lab-autoscaling-clb를 선택한다.



5. [상태 확인] 섹션에서 [상태 확인 유예 기간] 300초를 확인하고, [추가 설정 – 선택 사항] 섹션에서 추가 설정의 모니터링 활성화를 통해 오토스케일링 그룹에서 추적할 수 있는 여러 지표를 받을 수 있도록 체크하고 [다음] 버튼을 클릭한다.

상태 확인 - 선택 사항

상태 확인 유형 정보
EC2 Auto Scaling은 상태 확인에 실패한 인스턴스를 자동으로 교체합니다. 로드 밸런싱을 활성화한 경우 항상 활성화된 EC2 상태 확인 외에도 ELB 상태 확인을 활성화할 수 있습니다.

EC2 ELB

상태 확인 유예 기간
EC2 Auto Scaling이 새 인스턴스가 서비스 상태가 된 후 새 인스턴스에 대한 첫 번째 상태 확인을 수행할 때까지의 시간입니다.

300 초

추가 설정 - 선택 사항

모니터링 정보

CloudWatch 내에서 그룹 지표 수집 활성화

취소 이전 검토로 이동 **다음**

6. [4단계: 그룹 크기 및 조정 정책 구성]에서 [그룹 크기]를 다음과 같이 용량 한도를 지정한다.

- [원하는 용량] : 2
- [최소 용량] : 2
- [최대 용량] : 4

EC2 > Auto Scaling 그룹 > Auto Scaling 그룹 생성

1단계 시작 템플릿 또는 구성 선택

2단계 인스턴스 시작 옵션 선택

3단계 (선택 사항)
고급 옵션 구성

4단계 (선택 사항)
그룹 크기 및 조정 정책 구성

5단계 (선택 사항)
알림 추가

6단계 (선택 사항)
태그 추가

7단계
검토

그룹 크기 및 조정 정책 구성 정보

Auto Scaling 그룹의 원하는 용량, 최소 용량 및 최대 용량을 설정합니다. 선택적으로 조정 정책을 추가하여 그룹의 인스턴스 수를 동적으로 조정할 수 있습니다.

그룹 크기 - 선택 사항 정보

원하는 용량을 변경하여 Auto Scaling 그룹의 크기를 지정합니다. 최소 및 최대 용량 한도를 지정할 수도 있습니다. 원하는 용량은 한도 범위 내에 있어야 합니다.

원하는 용량	2
최소 용량	2
최대 용량	4

7. [조정 정책] 섹션은 조정 정책을 사용하여 트래픽에 따라 오토스케일링 그룹의 크기를 동적으로 조정하도록 설정하는 부분이다. 아래 그림과 같이 [대상 추적 조정 정책]을 선택하고 [조정 정책 이름]은 기본값 그대로 놓고, [지표 유형]은 평균 CPU 사용률을 선택하고, [대상 값]은 30으로 설정하고 [다음] 버튼을 클릭한다.

조정 정책 - 선택 사항

조정 정책을 사용하여 수요의 변화를 충족하도록 Auto Scaling 그룹의 크기를 동적으로 조정할지 여부를 선택합니다. [정보](#)

대상 추적 조정 정책
원하는 결과를 선택하고 조정 정책에 따라 필요한 경우 용량을 추가 및 제거하여 해당 결과를 달성합니다.

없음

조정 정책 이름

지표 유형

대상 값

인스턴스 요구 사항
 지표에 포함하기 전 워밍업 시간(초)

확대 정책만 생성하려면 축소 비활성화

8. [5단계: 알림 추가]에서는 오토스케일링 이벤트가 발생할 때마다 알림을 받을 수 있도록 설정할 수 있다. 필요에 따라 설정을 하거나 바로 [다음] 버튼을 클릭하여 단계로 넘어간다.

EC2 > Auto Scaling 그룹 > Auto Scaling 그룹 생성

1단계
시작 템플릿 또는 구성 선택

2단계
인스턴스 시작 옵션 선택

3단계 (선택 사항)
고급 옵션 구성

4단계 (선택 사항)
그룹 크기 및 조정 정책 구성

5단계 (선택 사항)
알림 추가

6단계 (선택 사항)
태그 추가

알림 추가 [정보](#)
Amazon EC2 Auto Scaling | Auto Scaling 그룹의 EC2 인스턴스를 시작하거나 종료할 때마다 SNS 주제로 알림을 전송합니다.

취소

9. [6단계: 태그 추가]에서 아래와 같이 [키]를 Name, [값]은 webserver-asg로 입력하고 [다음]을 클릭하여 다음 단계로 넘어간다.

EC2 > Auto Scaling 그룹 > Auto Scaling 그룹 생성

태그 추가 정보

태그를 추가하면 AWS에서 Auto Scaling 그룹의 검색, 필터링 및 추적에 도움이 됩니다. 인스턴스가 시작될 때 이러한 태그를 인스턴스에 자동으로 추가하도록 선택할 수도 있습니다.

태그(1)

키	값 - 선택 사항	새 인스턴스에 태그 지정
Name	webserver-asg	<input checked="" type="checkbox"/>

태그 추가

49 남은 수

다음

10. [7단계: 검토]에서 앞서 설정한 정보를 확인한 후, [Auto Scaling 그룹 생성] 버튼을 클릭한다.

EC2 > Auto Scaling 그룹 > Auto Scaling 그룹 생성

검토 정보

1단계: 시작 템플릿 또는 구성 선택

그룹 세부 정보

Auto Scaling 그룹 이름: webserver-asg

시작 템플릿: webserver-template [?] 버전: Default 설명: Template for Webserver Autoscaling

다음

6단계: 태그 추가

태그 (1)

키	값	새 인스턴스에 태그 지정
Name	webserver-asg	예

다음

11. [Auto Scaling 그룹]이 잘 생성되었다.

The screenshot shows the AWS Auto Scaling Groups page. At the top, there's a search bar labeled 'Auto Scaling 그룹 검색'. Below it is a table with the following columns: '이름' (Name), '시작 템플릿/구성' (Start Template/Configuration), '인스턴스' (Instances), '상태' (Status), '원하는 용량' (Desired Capacity), '최소' (Min), '최대' (Max), and '가용 영역' (Available Regions). One row is visible for 'webserver-asg' with a status of '실행 중' (Running), a desired capacity of 2, and two instances listed. A yellow button at the top right says 'Auto Scaling 그룹 생성' (Create Auto Scaling Group).

12. [인스턴스] 메뉴를 클릭하여 원하는 용량 2개가 생성되었는지 확인한다.

The screenshot shows the AWS Instances page. At the top, there's a search bar labeled '인스턴스를 속성 또는 (case-sensitive) 태그로 찾기'. Below it is a table with columns: 'Name', '인스턴스 ID', '인스턴스 상태', '인스턴스 유형', '상태 검사', '경보 상태', '가용 영역', and '피블릭 IPv4 DNS'. Three instances are listed: 'webserver-asg' (running, t2.micro, 2/2 healthy, ap-northeast-2a), 'lab-webserver1-ec2' (running, t2.micro, 2/2 healthy, ap-northeast-2a), and 'webserver-asg' (running, t2.micro, 2/2 healthy, ap-northeast-2c). A yellow button at the top right says '인스턴스 시작' (Start Instances).

13. 또한 위에서 생성한 Auto Scaling 그룹인 webserver-asg 페이지의 [활동] 탭을 확인하면 성공적으로 2개의 EC2 인스턴스의 생성을 확인할 수 있다.

The screenshot shows the AWS Auto Scaling Groups - 'webserver-asg' page. At the top, there are tabs: '세부 정보', '활동' (Activities), 'Auto Scaling', '인스턴스 관리', '모니터링', and '인스턴스 새로 고침'. The '활동' tab is selected. It displays two activity logs under '활동 알림 (0)'. Below that is a section for '작업 기록 (2)'. Two entries are listed, both marked as '성공' (Success) and timestamped on October 17, 2023, at 09:26:55 AM +09:00. The log details show the creation of new EC2 instances based on a user request to change the desired capacity from 0 to 2. A red box highlights the second log entry.

부하 발생 테스트 후, 오토 스케일링 동작 여부 확인

- 좌측 메뉴 중 [로드 밸런싱] > [로드밸런서]를 클릭하여 이미 생성한 로드 밸런서인 lab-autoscaling-clb 페이지의 [상태 검사] 탭을 클릭한다. 그리고 [편집] 버튼을 클릭한다.

리스너 네트워크 매핑 보안 상태 검사 대상 인스턴스 모니터링 속성 태그

상태 검사 설정

로드 밸런서가 자동으로 상태 확인을 수행하여 등록된 모든 인스턴스의 가용성을 테스트합니다. 상태 확인은 지정된 프로토콜, 포트 및 경로를 사용하여 ping 요청을 전송합니다. 트래픽은 상태 확인 설정에서 지정한 상태 기준을 만족하여 정상으로 확인된 인스턴스로만 라우팅됩니다. 상태 확인에 응답하지 않은 인스턴스는 비정상으로 간주되어 정상 임계값을 충족할 때까지 트래픽 수신을 중지합니다.

Ping 대상	정상 임계 값	비정상 임계값	제한 시간
HTTP:80/index.html	연속 상태 확인 성공 10회	연속 상태 확인 실패 2회	2초
간격			
5초			

- [상태 검사 구성] 창에서 [Ping 경로]의 값은 /로 변경하고 [저장] 버튼을 클릭한다.

EC2 > 로드 밸런서 > lab-autoscaling-clb > 상태 검사 설정 편집

▶ 로드 밸런서 세부 정보: lab-autoscaling-clb

상태 검사 정보

로드 밸런서가 자동으로 상태 확인을 수행하여 등록된 모든 인스턴스의 가용성을 테스트합니다. 상태 확인은 지정된 프로토콜, 포트 및 경로를 사용하여 ping 요청을 전송합니다. 트래픽은 상태 확인 설정에서 지정한 상태 기준을 만족하여 정상으로 확인된 인스턴스로만 라우팅됩니다. 상태 확인에 응답하지 않은 인스턴스는 비정상으로 간주되어 정상 임계값을 충족할 때까지 트래픽 수신을 중지합니다.

Ping 대상
지정한 프로토콜 및 포트를 사용하여 상태 확인 ping이 전송됩니다. HTTP/HTTPS 프로토콜을 사용하는 경우 대상 경로도 제공해야 합니다.

Ping 프로토콜	Ping 포트	Ping 경로
HTTP	: 80	/
1-65535		

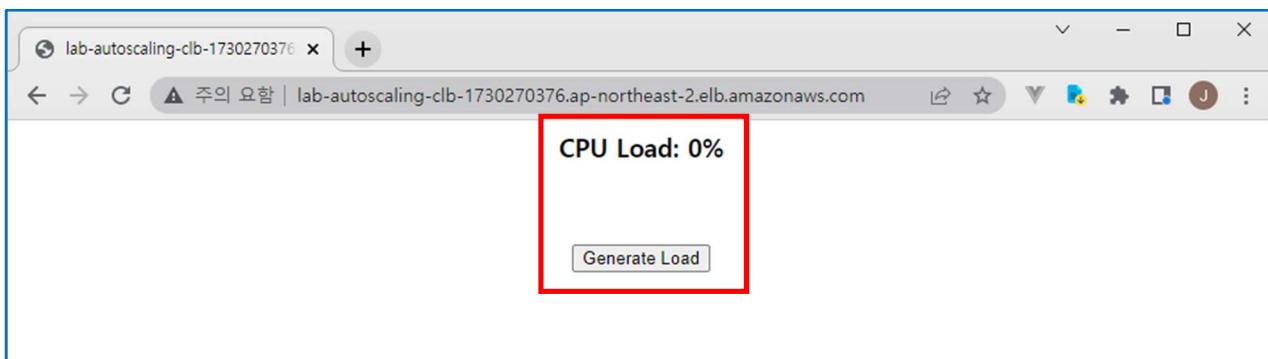
▶ 고급 상태 검사 설정

취소 변경 내용 저장

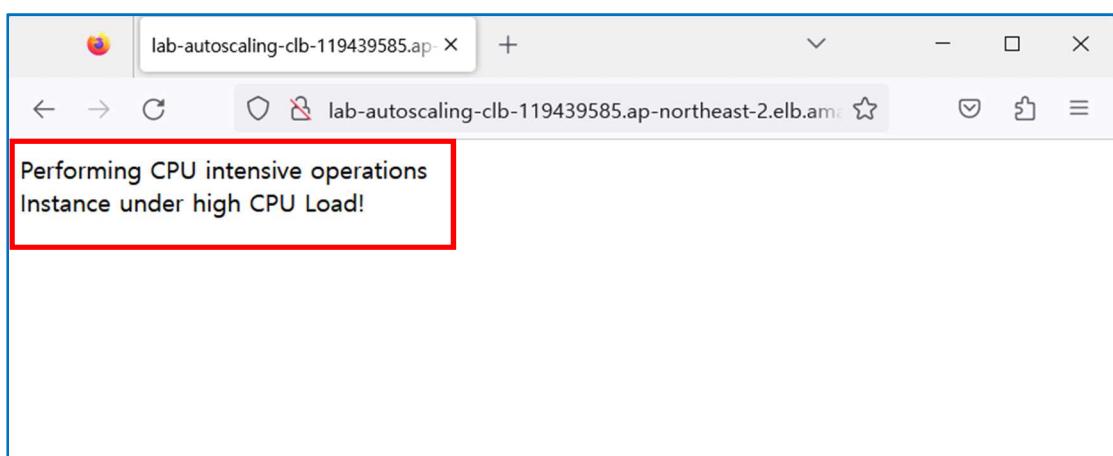
3. 다시 [설명] 탭 페이지로 이동하여 [DNS 이름]의 A 레코드 값을 복사한다.

The screenshot shows the AWS Classic Load Balancer configuration page. It displays various details about the load balancer, including its type (Classic), state (Service 0/3 instances), VPC (vpc-02ff6d53ce26bb1c6), and creation date (2023-10-17). A red box highlights the 'DNS 이름 정보' section, which contains the copied DNS value: lab-autoscaling-clb-119439585.ap-northeast-2.elb.amazonaws.com (A 레코드). Below this, a note suggests migrating to an Application Load Balancer. A 'Generate Load' button is also visible.

4. 웹 브라우저를 열고 방금 복사한 주소를 붙여넣는다. 다음과 같은 페이지로 로딩된다. 여기서 [Generate Load] 버튼을 클릭한다.



5. CPU Load의 값이 100%가 되는 것을 확인한다.



6. 300초 즉 5분이 지난 뒤 EC2 인스턴스 목록에서 보면 4개의 인스턴스가 실행되고 있음을 확인할 수 있다.

인스턴스 (5) 정보		인스턴스 ID	인스턴스 상태	인스턴스 유형	상태 검사	경보 상태	가용 영역
<input type="checkbox"/>	Name	i-03805076fbcf915c4	실행 중	t2.micro	2/2개 검사 통과	경보 없음	ap-northeast-2a
<input type="checkbox"/>	webserver-asg	i-07c5c0f65b77e8774	실행 중	t2.micro	2/2개 검사 통과	경보 없음	ap-northeast-2a
<input type="checkbox"/>	lab-webserver1-ec2	i-0a6cb2cdaf34c1335	실행 중	t2.micro	-	경보 없음	ap-northeast-2a
<input type="checkbox"/>	webserver-asg	i-0c47c0df2eaefcd39	실행 중	t2.micro	2/2개 검사 통과	경보 없음	ap-northeast-2c
<input type="checkbox"/>	webserver-asg	i-0b2d87aa1da34f17d	실행 중	t2.micro	초기화	경보 없음	ap-northeast-2c

7. webserver-asg Auto Scaling 그룹에서도 인스턴스 개수가 최대의 개수와 일치하는 것을 확인할 수 있다.

Auto Scaling 그룹 (1) Info							
Auto Scaling 그룹 검색							
이름	시작 템플릿/구성	인스턴스	상태	원하는 ...	최소	최대	가
webserver-asg	webserver-template 버전 기본값	4	-	4	2	4	ap

8. webserver-asg 그룹의 [활동] 탭에서도 확인할 수 있다.

작업 기록 (4)			
상태	설명	원인	시작 시간
① 인스턴스 워밍업 대기 중	Launching a new EC2 instance: i-0b2d87aa1da34f17d	At 2023-10-17T02:31:53Z a monitor alarm TargetTracking-webserver-asg-AlarmHigh-55199366-1a92-45c6-9df1-65e251cc6de4 in state ALARM triggered policy Target Tracking Policy changing the desired capacity from 2 to 4. At 2023-10-17T02:31:58Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 2 to 4.	2023 October 17, 11:32:00 AM +09:00
① 인스턴스 워밍업 대기 중	Launching a new EC2 instance: i-0a6cb2cdaf34c1335	At 2023-10-17T02:31:53Z a monitor alarm TargetTracking-webserver-asg-AlarmHigh-55199366-1a92-45c6-9df1-65e251cc6de4 in state ALARM triggered policy Target Tracking Policy changing the desired capacity from 2 to 4. At 2023-10-17T02:31:58Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 2 to 4.	2023 October 17, 11:32:00 AM +09:00
성공	Launching a new EC2 instance: i-03805076fbcf915c4	At 2023-10-17T00:26:48Z a user request created an AutoScalingGroup changing the desired capacity from 0 to 2. At 2023-10-17T00:26:53Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 2.	2023 October 17, 09:26:55 AM +09:00
성공	Launching a new EC2 instance: i-0c47c0df2eaefcd39	At 2023-10-17T00:26:48Z a user request created an AutoScalingGroup changing the desired capacity from 0 to 2. At 2023-10-17T00:26:53Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 2.	2023 October 17, 09:26:55 AM +09:00

9. [자동 조정] 탭을 클릭하여 해당 페이지로 이동한다. 현재는 **Target Tracking Policy** 즉 [동적 크기 조정 정책]만 확인할 수 있다.

The screenshot shows the AWS Auto Scaling console for the 'webserver-asg' group. The 'Auto Scaling' tab is selected. A red box highlights the '동적 크기 조정 정책 (1) Info' section, which contains the configuration for the 'Target Tracking Policy'. The policy details include:

- 대상 주적 크기 조정
- 활성
- 평균 CPU 사용률(률) 30(으)로 유지하는 데 필요한 경우
- 필요에 따라 충당 단위 추가 또는 제거
- 300초(지표에 포함하기 전 워밍업 시간)
- 활성

At the top right, there are buttons for '작업 ▾' and '동적 크기 조정 정책 생성'.

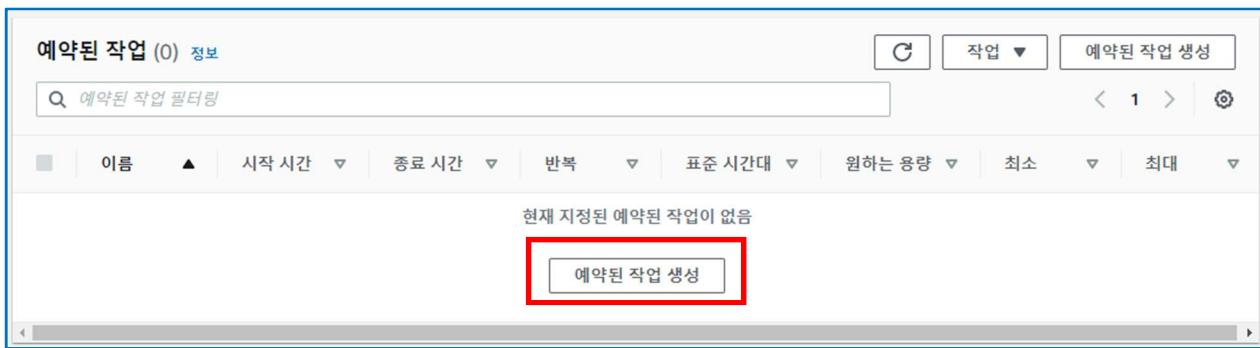
10. 페이지를 스크롤다운하여 [예측 크기 조정 정책]에서 [예측 크기 조정 정책 생성] 버튼을 클릭한다.

The screenshot shows the '예측 크기 조정 정책' creation page. It displays a message stating '생성된 예측 크기 조정 정책이 없음' and '예측 크기 조정 정책은 기록 데이터를 사용하여 예측 시간당 로드에 앞서 그룹을 확장합니다.' A red box highlights the '예측 크기 조정 정책 생성' button at the bottom.

11. [예측 크기 조정 정책 생성] 페이지에서도 CPU 사용률 기준으로 설정할 수 있다. [취소]를 클릭한다.

The screenshot shows the '예측 크기 조정 정책 생성' configuration page. The '지표 및 목표 사용률' section is highlighted with a red box. It shows the '지표' dropdown set to 'CPU 사용률' and the '목표 사용률' input field set to '50 인스턴스당 %'. Below this, the '추가 크기 조정 설정 - 선택 사항' section is partially visible, showing '인스턴스 사전 시작' and '최대 용량 동작' options. At the bottom right, there are '취소' and '생성' buttons.

12. 이번에는 그 아래 섹션인 [예약된 작업]을 생성하기 위해 [예약된 작업 생성]을 클릭한다.



13. [예약된 작업 생성] 창에서 다음과 같이 설정한 후, [생성] 버튼을 클릭한다.

- A. [이름] : lab-predict-job
- B. [원하는 용량] : 3
- C. [최소] : 2
- D. [최대] : 5
- E. [반복] : 한 번
- F. [표준 시간대] : Asia/Seoul
- G. [특정 시작 시간] : 원하는 시간(실습에서는 현재 시간에서 5분 추가로 맞춤)

The screenshot shows a modal dialog titled '예약된 작업 생성'.
- **이름**: lab-predict-job
- **원하는 용량**: 3
- **최소**: 2
- **최대**: 5
- **반복**: 한 번
- **표준 시간대**: Asia/Seoul
- **특정 시작 시간**: 2023/10/17 22:50 (선택한 표준 시간대의 현재 시간은 2023-10-17/22:42 KST)
- **Note**: 원하는 용량, 최소 용량 또는 최대 용량 중 하나 이상의 값 제공
At the bottom right, there are '취소' and '생성' buttons.

14. 아래와 같이 잘 생성되었다.

예약된 작업 (1) Info	
□ 이름	▲ 시작 시간 종료 시간 반복 표준 시간대 월하는 용량 최소 최대
□ lab-predict-job	2023 October 17, 1... Asia/Seoul 3 2 5

15. 예약된 시간이 되면 아래 그림과 같이 [상태]가 [용량 업데이트 중]으로 변경된다.

Auto Scaling 그룹 (1) Info	
□ 이름	▼ 시작 템플릿/구성 인스턴스 상태 원하는 ... 최소 최대 가용 영역
□ webserver-asg	webserver-template 버전 기본값 2 3 2 5 ap-northeast-2a, ap-northeast-2c

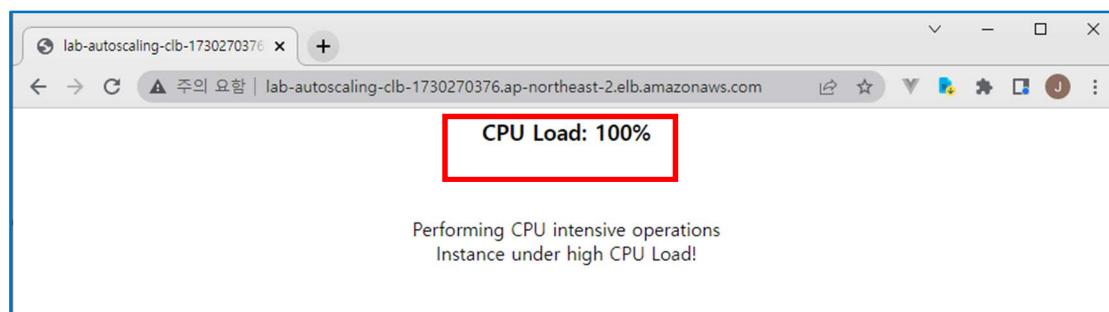
16. 잠시 기다리면 아래의 그림과 같이 [인스턴스] 개수는 3, [원하는 용량]은 2, [최소] 2, [최대] 5로 설정됨을 확인할 수 있다.

Auto Scaling 그룹 (1) Info	
□ 이름	▼ 시작 템플릿/구성 인스턴스 상태 원하는 ... 최소 최대 가용 영역
□ webserver-asg	webserver-template 버전 기본값 3 - 3 2 5 ap-northeast-2a, ap-northeast-2c

17. 인스턴스 목록에서 확인해 보면 인스턴스 개수가 변경되어 자동으로 줄어든 것을 확인할 수 있다.

인스턴스 (4) 정보									
□ Name	▼ 인스턴스 ID 인스턴스 상태 인스턴스 유형 상태 검사 경보 상태 가용 영역 ▼ 퍼블릭 IPv4 DNS 퍼블릭 IPv4 ... ▼ 단역적 IP								
□ lab-webserver1-ec2	i-0eb74b996f6910012	● 실행 중 2/2개 검사 통과	t2.micro	● 2/2개 검사 통과	경보 없음 +	ap-northeast-2a	ec2-54-180-157-235.ap...	54.180.157.235	-
□ webserver-asg	i-0fb436885b91b95c4	● 실행 중 2/2개 검사 통과	t2.micro	● 2/2개 검사 통과	경보 없음 +	ap-northeast-2a	-	-	-
□ webserver-asg	i-096d2909594e7b776	● 종료됨 0/0개 검사 통과	t2.micro	-	경보 없음 +	ap-northeast-2a	-	-	-
□ webserver-asg	i-0b4acc58a3c5f017e	● 실행 중 2/2개 검사 통과	t2.micro	● 2/2개 검사 통과	경보 없음 +	ap-northeast-2c	-	-	-

18. 다시 로드밸런스에 설정되어 있는 [DNS 이름] 값을 이용해서 CPU 이용율을 100%로 올리고 다시 확인해보자.



EC2 > Auto Scaling 그룹

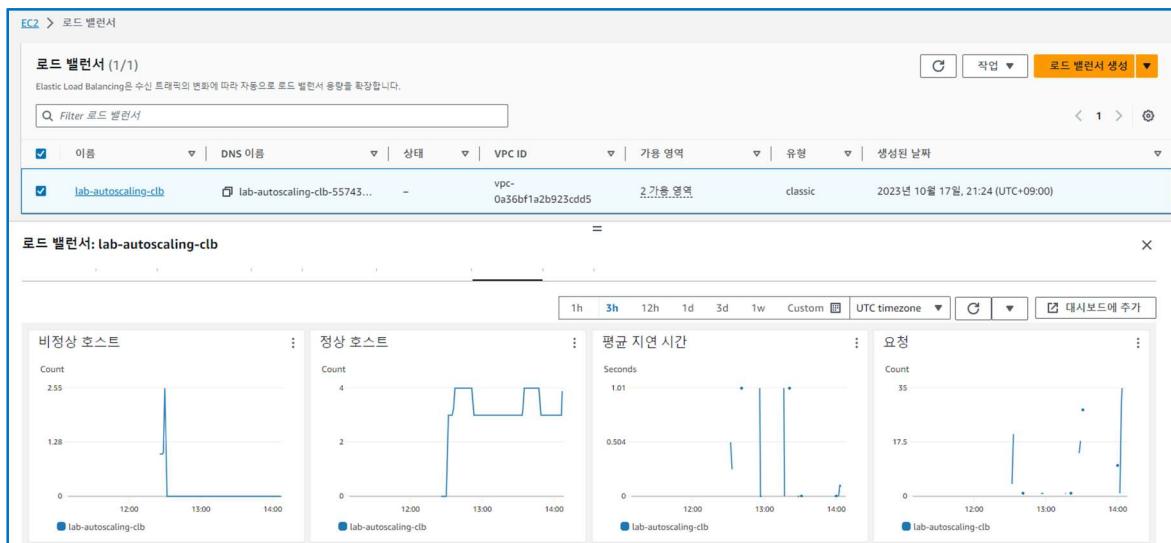
Auto Scaling 그룹 (1)

이름	시작 템플릿/구성	인스턴스	상태	원하는 용량	최소	최대
webserver-asg	webserver-template 버전 기본값	3	-	3	2	5

인스턴스 (7) 정보

Name	인스턴스 ID	인스턴스 상태	인스턴스 유형	상태 검사	경보 상태	가용 영역	퍼블릭 IPv4 DNS	퍼블릭 IPv4 ...	단락적 IP
lab-webserver1-ec2	i-0eb74b996f6910012	실행 중	t2.micro	2/2개 검사 통과	경보 없음	+ ap-northeast-2a	ec2-54-180-157-235.ap...	54.180.157.235	-
webserver-asg	i-0fb436885b91b95c4	실행 중	t2.micro	2/2개 검사 통과	경보 없음	+ ap-northeast-2a	-	-	-
webserver-asg	i-0d058579ef4730f2b	실행 중	t2.micro	2/2개 검사 통과	경보 없음	+ ap-northeast-2a	-	-	-
webserver-asg	i-096d2909594e7b776	종료됨	t2.micro	-	경보 없음	+ ap-northeast-2a	-	-	-
webserver-asg	i-0b4acc58a3c5f017e	실행 중	t2.micro	2/2개 검사 통과	경보 없음	+ ap-northeast-2c	-	-	-
webserver-asg	i-09f51bf0676b1ef35	실행 중	t2.micro	① 초기화	경보 없음	+ ap-northeast-2c	-	-	-
webserver-asg	i-08eb104a426ce652c	실행 중	t2.micro	① 초기화	경보 없음	+ ap-northeast-2c	-	-	-

19. 다음 그림은 로드 밸런서의 [모니터링] 탭이다.



20. 당연히 Auto Scaling 그룹의 [모니터링]에서도 관련 지표 확인 가능하다.

