

AUTOR:	Josely Castro
DATA CRIAÇÃO:	15/01/2025
ÚLTIMA REVISÃO:	12/02/2025
REVISADO POR:	Josely Castro
STATUS:	Ativo
TÍTULO:	Laboratório 6. Este laboratório demonstra Auto Scaling e Load Balancing na AWS, usando Launch Templates.

Objetivos: Este laboratório ensina como configurar um ambiente web básico, altamente disponível e escalável na AWS, utilizando Auto Scaling Groups, Launch Templates e Application Load Balancer. Você aprenderá a criar e integrar esses serviços essenciais para garantir que sua aplicação possa lidar com variações de tráfego e se recuperar automaticamente de falhas.

Cenário: Você é responsável por implantar e gerenciar uma aplicação web que precisa estar sempre disponível para os usuários, mesmo durante picos de acesso ou falhas de instâncias individuais. Para garantir a alta disponibilidade e a escalabilidade da aplicação, você utilizará os seguintes serviços da AWS:

- **Auto Scaling Groups:** Para gerenciar automaticamente o número de instâncias EC2 executando sua aplicação, com base na demanda.
- **Launch Templates:** Para definir as configurações das instâncias EC2 (sistema operacional, software, configurações de rede, etc.).
- **Application Load Balancer:** Para distribuir o tráfego de entrada entre as instâncias EC2 de forma inteligente, garantindo que nenhuma instância fique sobrecarregada.
- Seu desafio é configurar esses serviços na AWS, integrá-los corretamente e testar o funcionamento do ambiente, garantindo que ele seja resiliente a falhas e possa se ajustar automaticamente a mudanças na demanda.
 1. Usar a VPC *padrão* em us-east-1.
 2. Criar um *Launch Template*.
 3. Criar um *Auto Scaling Group*.
 4. Criar um *Application Load Balancer*.
 5. Integrar e testar.

Pré-requisitos

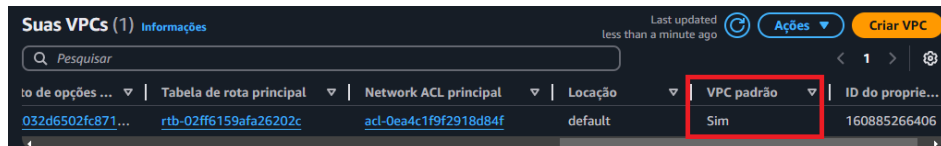
- Conta AWS Ativa.
- Permissões IAM: AmazonEC2FullAccess, AmazonEC2AutoScalingFullAccess, ElasticLoadBalancingFullAccess.
- Navegador Web.

Tarefas:

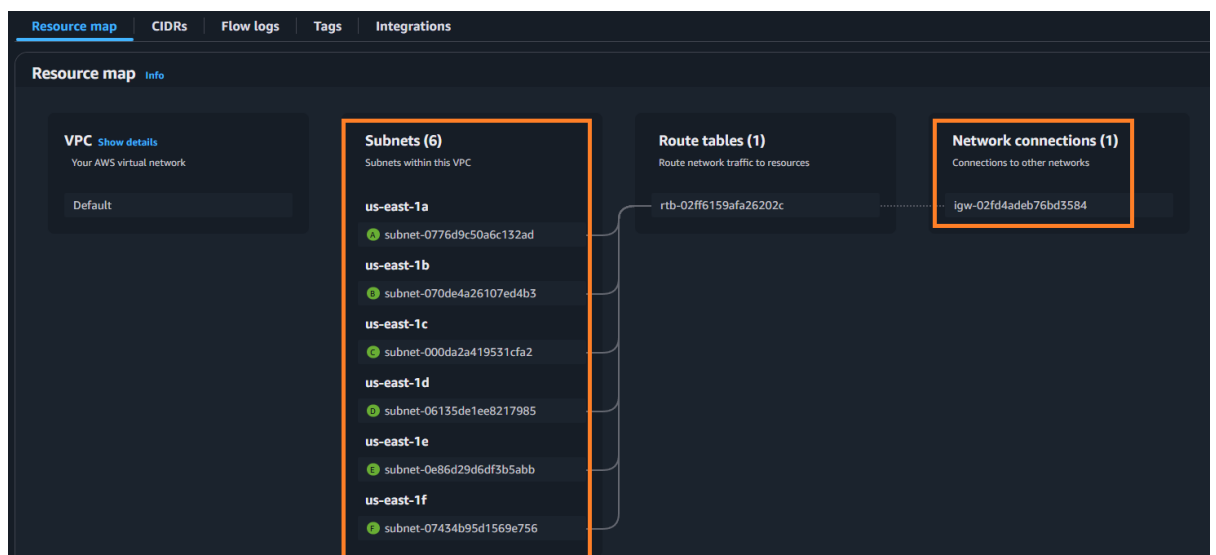
Passo 1:

1. **Console AWS - Região:**

- Acesse o console da AWS.
 - **Crucial:** No canto superior direito, selecione a região **Norte da Virgínia (us-east-1)**.
2. **Console VPC:**
- Na barra de pesquisa, digite "VPC" e selecione o serviço **VPC**.
3. **Verificar VPC Padrão:**
- No menu lateral, clique em **Your VPCs** (Suas VPCs).
 - Você deve ver uma VPC listada com "Default VPC" = "Yes".



4. **Anote** o ID da VPC (**vpc-xxxxxxxxxxxxxxxxxx**).
5. No menu lateral, clique em **Subnets**.
6. **Verifique** se há Subnets associadas à **VPC padrão**, distribuídas em **diferentes zonas de disponibilidade (AZs)**.
7. **Anote** os IDs das Subnets (**subnet-xxxxxxxxxxxxxxxxxx**) – você precisará de pelo menos duas. **Registre** as **AZs** que serão utilizadas neste laboratório, pois elas serão as mesmas nos próximos passos.
8. **Confirme** se as Subnets são **públicas**, verificando a **tabela de rotas** e garantindo que há uma **rota para um Internet Gateway (igw-xxxxxxx)**. Em VPC > Your VPCs > Selecione a VPC Padrão. Você terá exatamente essa visão em “**Resource map**”:



Passo 2:

1. **Console EC2:** Acesse o console do EC2 (us-east-1).
2. **Security Groups:** Clique em "Security Groups".
3. **Create Security Group:**
 - **Security group name:** SG-Lab-SeuNomeSobrenome
 - **Description:** HTTP, HTTPS, SSH.
 - **VPC:** Sua *VPC padrão*.
 - **Inbound rules:**
 - HTTP, Source: Anywhere-IPv4

- HTTPS, Source: Anywhere-IPv4
 - **Outbound rules:** Padrão.
- 4. **Create security group.**

Passo 3:

1. **Console EC2:** No console do EC2 (us-east-1).
2. **Launch Templates:** No menu lateral, em "Instances", clique em **Launch Templates**.
3. **Create launch template:**
 - **Launch template name:** LaunchTemplates_NomeSobrenome
 - **Template version description:** (Opcional)
 - **Auto Scaling guidance:** *Marque a caixa.*
 - **Application and OS Images (AMI):**
 - **Quick Start -> Amazon Linux.**
 - **Amazon Machine Image (AMI):** Escolha **Amazon Linux 2** (x86_64, Free tier eligible).
 - **Instance type:** t2.micro
 - **Key pair (login):** Selecione "Proceed without a key pair (Not recommended)".
 - **Network settings:**
 - **Subnet:** Não selecione uma sub-rede aqui. Faremos isso no Auto Scaling Group.
 - **Firewall (security groups):** Select an existing security group: SG-Lab-SeuNomeSobrenome.
 - **Advanced network configuration:** Não precisa mexer.
 - **Storage (volumes):** Padrão (8 GiB).
 - **Advanced details > User data:**

```
#!/bin/bash
yum update -y
yum install -y httpd
systemctl start httpd
systemctl enable httpd
echo "<h1>Servidor Web - Instância: $(hostname -f)</h1>" >
/var/www/html/index.html
```

- **Análise do Script:**

- Define que o script será interpretado pelo Bash:


```
#!/bin/bash
```
- Atualiza os pacotes da instância: `yum update -y`
- Instala o servidor Apache (httpd): `yum install -y httpd`
- Inicia o serviço Apache: `systemctl start httpd`

- **Configura o serviço para iniciar automaticamente no boot:**
`systemctl enable httpd`
- **Cria uma página HTML simples no diretório padrão do Apache (`/var/www/html/`), incluindo o nome da instância:**

```
echo "<h1>Servidor Web (Ultra Simples) -  
Instância: $(hostname -f)</h1>" >  
/var/www/html/index.html
```

- Deixe o restante como padrão.

4. **Create launch template.**

Passo 4:

1. **Console EC2:** No console do EC2.
2. **Auto Scaling Groups:** Clique em "Auto Scaling Groups".
3. **Create Auto Scaling group:**
 - **Step 1:**
 - **Name:** ASG-SeuNomeSobrenome
 - **Launch template:** LT-SeuNomeSobrenome
 - **Version:** Default.
 - **Next.**
 - **Step 2:**
 - **VPC:** Sua *VPC padrão*.
 - **Subnets:** Selecione as duas subnets públicas que você anotou anteriormente no passo "*Verificar VPC Padrão*".
 - **Next.**
 - **Step 3:**
 - **Load balancing:** *Selecione Attach to a new load balancer.*
 - **Load balancer type:** Application Load Balancer.
 - **Load balancer scheme:** Internet-facing.
 - **Listeners and routing:** Listeners = HTTP:80(Padrão).
 - **Default routing (forward to):** Selecione "Create a target group"
 - **Health checks:** Selecione "Turn on Elastic Load Balancing health checks".
 - **Next.**
 - **Step 4:**
 - **Desired capacity, Min desired capacity, Max desired capacity:** 2
 - **Scaling policies:** None
 - **Next.**
 - **Step 5, 6:** (Next, Next).
 - **Step 7:** Add Tag = Key=Name, Value=WebServer-SeuNomeSobrenome.
 - **Step 8:** Review.
 - **Create Auto Scaling group.**

Passo 5:

OBS: Tenha paciência nesta etapa, pois o processo pode levar cerca de 5 minutos para você ver a mágica acontecer.

1. **DNS do Load Balancer:** Console EC2 -> Load Balancers -> copie o "DNS name".
2. **Acessar:** Cole no navegador.
3. **Resultado:** A página "Servidor Web ..." será exibida.
4. Ao atualizar (com nomes de host diferentes), você perceberá que o número do IP do seu servidor Web muda. Esse é exatamente o propósito! Aqui, o load balancer está funcionando corretamente, distribuindo o tráfego entre as instâncias Web 1 e Web 2. Você verá algo semelhante às imagens abaixo:

Server 1

Servidor Web - Instância: ip-172-31-47-55.ec2.internal

Server 2

Servidor Web - Instância: ip-172-31-6-17.ec2.internal

Importante: Tire um print dos seus testes e anexe as imagens ao entregar a atividade no Classroom.

Informações adicionais

Se a página não for exibida ou houver falha na resposta, verifique os seguintes pontos:

Sem Resposta:

- Grupo de segurança (verifique se as regras de entrada permitem tráfego HTTP/HTTPS).
- Instâncias com IP público (confirme se possuem IP público atribuído e se a conectividade está correta).
- Listener do Load Balancer (garanta que a porta e as regras de roteamento estão configuradas corretamente).

Instâncias marcadas como "Unhealthy" no Load Balancer:

- User Data (verifique se o script de inicialização configurou corretamente o servidor Web).
- Grupo de segurança (certifique-se de que as regras permitem tráfego interno entre o Load Balancer e as instâncias).

Passo 6:

Após o término do laboratório, exclua todos os recursos para evitar cobranças desnecessárias.

1. **Exclua o Auto Scaling Group.**

2. Exclua o Load Balancer.
3. Exclua o Target Group.
4. Exclua o Launch Template.
5. Exclua o Security Group.

Parabéns! Você concluiu com sucesso o laboratório de Auto Scaling e Load Balancing! Você agora tem uma compreensão prática de como criar uma infraestrutura web básica, mas altamente disponível e escalável na AWS. Você aprendeu a:

- Usar a VPC padrão para simplificar a configuração de rede.
- Criar um Launch Template para definir as configurações das suas instâncias EC2.
- Configurar um Auto Scaling Group para gerenciar automaticamente o número de instâncias com base na demanda (mesmo que, neste laboratório simplificado, tenhamos mantido um número fixo de instâncias).
- Criar um Application Load Balancer para distribuir o tráfego de forma inteligente entre as instâncias.
- Integrar o Auto Scaling Group e o Load Balancer para garantir que novas instâncias sejam automaticamente adicionadas ao balanceamento de carga e instâncias com falha sejam removidas.
- Verificar o funcionamento do ambiente e solucionar problemas comuns.

Esses são conceitos fundamentais para a construção de aplicações robustas e escaláveis na nuvem. Você deu um passo importante no seu aprendizado de AWS!