

Serviços de Computação em Nuvem: Escalabilidade com Auto-Scaling

Resumo

Referências Bibliográficas

- Criação de Instância EC2 Amazon Linux https://www.awsacademy.com/LMS_Login;
- [Endereços IP elásticos - Amazon Elastic Compute Cloud](#)
- Amazon EBS - [Armazenamento em blocos de alta performance – Amazon EBS; Amazon Elastic Block Store \(Amazon EBS\) - Amazon Elastic Compute Cloud](#)
- Snapshots - [Snapshots do Amazon EBS - Amazon Elastic Compute Cloud](#)

Auto Scaling - Criando instância Servidor Web APP para o Auto Scalling

Benefícios do Multi-AZ

Multi-AZ é importante para implementações de serviços na AWS pois com múltiplas Zonas de Disponibilidade (AZ) temos uma série de benefícios. Entre eles são:

- Resiliência;
- Failover automático;
- Maior disponibilidade;
- Durabilidade aperfeiçoada;
- Proteção do desempenho do seu banco de dados;

O que é Resiliência e Failover automático?

Resiliência é a capacidade de uma carga de trabalho se recuperar de interrupções de infraestrutura ou serviço, adquirir dinamicamente recursos computacionais para atender à demanda e mitigar interrupções, como configurações incorretas ou problemas temporários de rede.

Se um volume de armazenamento na instância primária falhar em uma implantação Multi-AZ do Amazon RDS, ele iniciará automaticamente um failover (instância que estava em standby) para a espera atualizada (ou para uma réplica, no caso do Amazon Aurora).

O que é maior disponibilidade?

Você se beneficia da disponibilidade otimizada do servidor ao executar implantações Multi-AZ. Se ocorrer uma falha em uma Zona de disponibilidade ou em uma instância de DB, por exemplo, o impacto na disponibilidade será limitado ao tempo necessário para a conclusão do failover automático, que geralmente é inferior a um minuto para o Amazon Aurora (chegando a 30 segundos com o uso do MariaDB Connector/J) e de um a dois minutos para outros mecanismos de banco de dados (consulte as perguntas frequentes do RDS para obter detalhes).

O que é Proteção do desempenho do seu banco de dados?

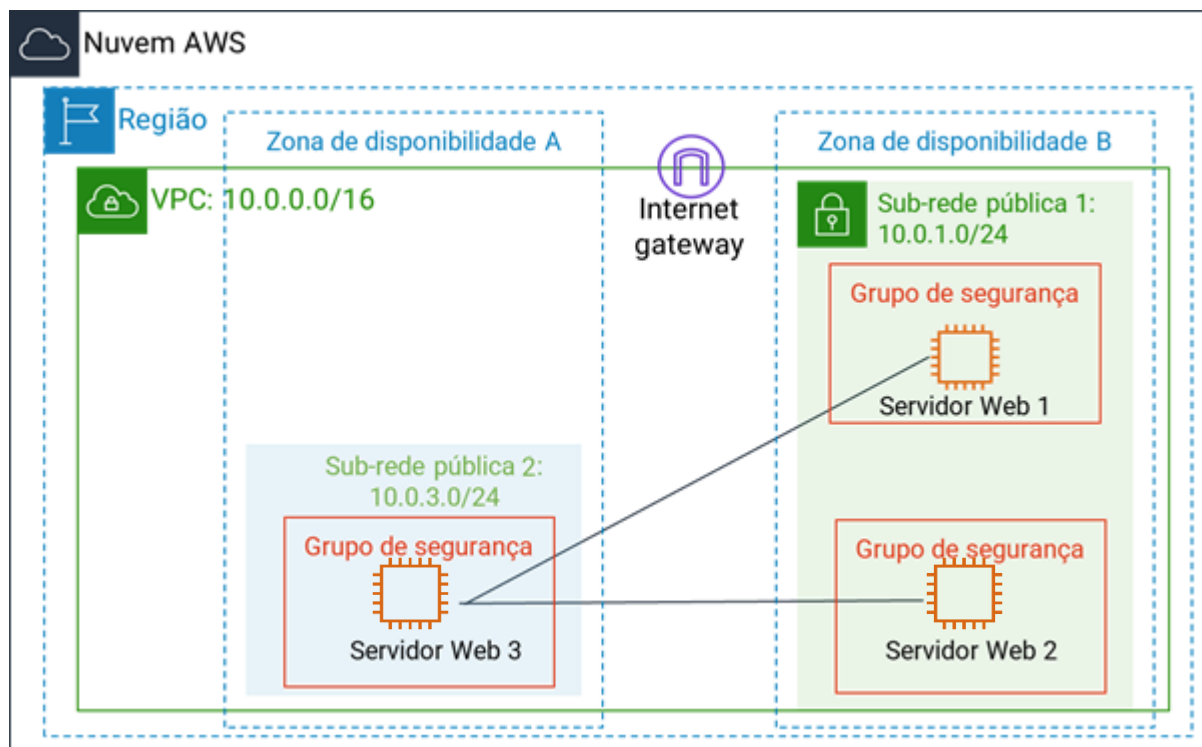
Diferentemente das implantações Single-AZ, a atividade de E/S não é suspensa em sua instância principal durante o backup nas implantações Multi-AZ de mecanismos MySQL, MariaDB, Oracle e PostgreSQL, pois o backup é feito por meio da instância de espera.

Mais sobre implantações Multi-AZ

Implantações Multi-AZ podem ter uma ou duas instâncias de banco de dados, por exemplo, em espera.

Uma implantação de instância de banco de dados Multi-AZ tem uma instância de banco de dados em espera (standby) que fornece suporte para failover, mas não serve tráfego de leitura.

Quando a implantação tem duas instâncias de banco de dados em espera, ela é chamada de implantação de cluster de banco de dados Multi-AZ. Uma implantação de cluster de banco de dados Multi-AZ tem instâncias de banco de dados em espera que fornecem suporte para failover e podem servir tráfego de leitura.

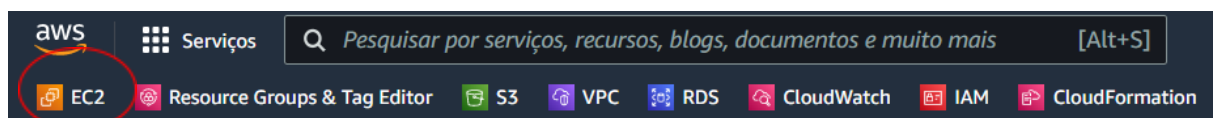
Exemplo de topologia MultiAZ:

Aqui temos um exemplo de uma topologia com multi-AZ com uma alta disponibilidade e resiliência.

Nesta topologia temos 3 instâncias web, estando uma em uma Zona de Disponibilidade A e duas em uma Zona de Disponibilidade B. Caso uma das zonas caia, perca conectividade ou tenha algum problema parecido.

Laboratório 10: Criar Instância com WebServer APP para o Auto Scalling

1. No Console de Gerenciamento da AWS no menu **Serviços**, clique em **EC2**.
2. No **Painel EC2**, clique em **Executar instância**.
3. **Etapa 1** - você irá se deparar com diversas **AMIs** encontradas na página de **AMIs**
4. Caso queira ver apenas as **AMIs gratuitas**, clique em **Somente nível gratuito**



1. Selecione a AMI 2. Escolher tipo de instância 3. Configurar instância 4. Adicionar armazenamento 5. Adicionar T

Etapa 1: Selecione uma Imagem de máquina da Amazon (AMI)

Uma AMI é um modelo que contém a configuração do software (sistema operacional, servidor de aplicativos e aplicativos) para usuários ou no AWS Marketplace, ou pode selecionar uma das suas próprias AMIs.

Procure uma AMI digitando um termo de pesquisa; por exemplo, "Windows"

Início rápido

Minhas AMIs

AWS Marketplace

AMIs da comunidade

☒ Somente nível gratuito



Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type

Amazon Linux
qualificado para o

Amazon Linux 2 comes with five years support. It provides Linux kernel 2.29.1, and the latest software packages through extras. This AMI is from this wizard.

Tipo de dispositivo raiz: ebs Tipo de virtualização: hvm ENA habilitado: S

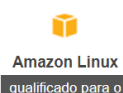


Amazon Linux 2 AMI (HVM) - Kernel 4.14, SSD Volume Type

Amazon Linux
qualificado para o

Amazon Linux 2 comes with five years support. It provides Linux kernel 2.29.1, and the latest software packages through extras. This AMI is

Selecione a AMI **Amazon Linux 2 AMI (HVM)** que é qualificada para o nível gratuito, mantenha **64 bits (x86)** selecionado e clique em **Selecionar**



Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type - ami-08e4e35cccc6189f4 (64 bits x86) / ami-0789681fae8b18e56 (64 bits Arm)

Amazon Linux 2 comes with five years support. It provides Linux kernel 5.10 tuned for optimal performance on Amazon EC2, systemd 219, GCC 7.3, Glibc 2.26, Binutils 2.29.1, and the latest software packages through extras. This AMI is the successor of the Amazon Linux AMI that is now under maintenance only mode and has been removed from this wizard.

Selecionar

- ☒ 64 bits (x86)
☐ 64 bits (Arm)

Etapa 2 – escolha o tipo da instância, **t2.micro** é neste momento que estamos escolhendo um processador e memória mais potentes conforme solicitação da empresa.

1. Selecione a AMI 2. Escolher tipo de instância 3. Configurar instância 4. Adicionar armazenamento 5. Adicionar Tags 6. Configure o security group 7. Análise

Etapa 2: Escolha um tipo de instância

O Amazon EC2 oferece uma ampla seleção de tipos de instâncias otimizadas para se adequarem a casos de uso diferentes. Instâncias são servidores virtuais que podem executar aplicativos. Possuem várias combinações de CPU, memória, armazenamento e capacidade de rede e oferecem flexibilidade de escolha da composição adequada de recursos para os seus aplicativos. Saiba mais sobre tipos de instância e como podem atender às suas necessidades de computação.

Filtrar por: Todas as famílias de instâncias Geração atual Mostrar/ocultar colunas

Selecionada atualmente: t2.micro (- ECUs, 1 vCPUs, 2.5 GHz, -, 1 GiB memória, Somente EBS)

	Família	Tipo	vCPUs	Memória (GiB)	Armazenamento da instância (GB)	Disponível otimizado para EBS	Desempenho de rede	Compatibilidade com IPv6
☐	t2	t2.nano	1	0.5	Somente EBS	-	Baixa a moderado	Sim
☒	t2	t2.micro qualificado para o nível gratuito	1	1	Somente EBS	-	Baixa a moderado	Sim
☐	t2	t2.small	1	2	Somente EBS	-	Baixa a moderado	Sim
☐	t2	t2.medium	2	4	Somente EBS	-	Baixa a moderado	Sim

Cancelar

Anterior

Verificar e ativar

Próximo: Configure os detalhes da instância

Clique em “Próximo: Configure os detalhes da instância”;

Segue relação e detalhamento de cada família de máquinas:

<https://aws.amazon.com/pt/ec2/instance-types/>

Etapa 3 – nesta etapa podemos fazer configurações detalhadas de nossa instância para este caso, usaremos **LAB-VPC** criada no **Laboratório 1**.

nossa instância **Amazon EC2** ficará na **Sub-rede-Pública-1**, onde será atribuído automaticamente um endereço IP público à nossa máquina virtual incluindo DNS.

1. Selecione a AMI 2. Escolher tipo de instância 3. Configurar instância 4. Adicionar armazenamento 5. Adicionar Tags 6. Configure o security group 7. Análise

Etapa 3: Configure os detalhes da instância

Configure a instância para se adequar aos seus requisitos. Você pode executar várias instâncias na mesma AMI, solicitar instâncias spot para aproveitar a vantagem instância, e outros.

Número de instâncias Executar no grupo de Auto Scaling

Opção de compra ☐ Solicitar instâncias spot

Rede Criar nova VPC

Sub-rede Criar nova sub-rede

Auto-assign Public IP

Tipo de nome do host

DNS Hostname ☒ Enable IP name IPv4 (A record) DNS requests
☒ Habilitar solicitações de DNS IPv4 (registro A) com base em recursos
☐ Habilitar solicitações de DNS IPv6 (registro AAAA) com base em recursos

Grupo de posicionamento ☐ Adicione a instância a um grupo de posicionamento

Reserva de capacidade

a) Estamos usando **Lab-VPC** padrão da AWS “vpc-05c689....”;

- b) Alterar **Auto-assign Public IP** para: **Habilitar**
- c) Deixar habilitado a opção de DNS - "**Habilitar solicitações de DNS IPv4** (Registro A) com base em recursos".

Role para baixo para continuar nesta etapa 3, não vamos alterar nada neste campos.

Etapa 3: Configure os detalhes da instância

Grupo de posicionamento ⓘ ☐ Adicione a instância a um grupo de posicionamento

Reserva de capacidade ⓘ

Diretório de ingresso em domínio ⓘ [Criar novo diretório](#)

Função do IAM ⓘ [Criar nova função do IAM](#)

Comportamento de desligamento ⓘ

Interromper - Comportamento de hibernação ⓘ ☐ Habilitar hibernação como um comportamento de interrupção adicional

Habilitar a proteção contra encerramento ⓘ ☐ Proteger contra encerramento acidental

Monitoramento ⓘ ☐ Habilitar monitoramento detalhado do CloudWatch
[Aplicam-se cobranças adicionais.](#)

Localção ⓘ
[Serão aplicados encargos adicionais locação dedicada.](#)

Elastic Inference ⓘ ☐ Add an Elastic Inference accelerator
[Aplicam-se cobranças adicionais.](#)

Role para baixo para ver a seção **Detalhes avançados**.

- a) Expanda **detalhes avançados**. Aparecerá no último campo **Dados do Usuário**.
- b) Copie os comandos a seguir e cole-os no campo **Dados do Usuário**. Mantendo a opção "**Como texto**" selecionada.

eth0 [Adicionar IP](#) A sub-rede selecionada não oferece suporte a IPv6 porque não tem um CIDR IPv6.

[Adicionar dispositivo](#)

▼ Detalhes avançados

Enclave ⓘ ☐ Habilitar

Metadados acessíveis ⓘ

Versão de metadados ⓘ

Limite de salto de resposta do token de metadados ⓘ

Allow tags in metadata ⓘ

Dados do usuário ⓘ ☒ Como texto ☐ Como arquivo ☐ A entrada já está codificada como base64

```
# Habilitar Web-APP
cd /var/www/html
wget https://aws-logs-1-6.s3.amazonaws.com/CUR-TF-100-ACCLFO-2/lab5-rds/lab-app-php7.zip
unzip lab-app-php7.zip -d /var/www/html/
chown apache:root /var/www/html/rds.conf.php
```

[Cancelar](#) [Anterior](#) [Verificar e ativar](#) [Próximo: Adicionar armazenamento](#)

Clique em "**Próximo: Configure os detalhes da instância**";

```
#!/bin/bash -ex
# Updated Amazon Linux 2
yum -y update
# Instalar Apache2 - PHP - Banco de Dados MySQL
yum -y install httpd php mysql php-mysql
# Instalar LAMP para MariaDB 10.2 e PHP 7.2
amazon-linux-extras install -y lamp-mariadb10.2-php7.2 php7.2
# Instalar Apache2 - MariaDB
yum install -y httpd mariadb-server
# Habilitar Apache2
/usr/bin/systemctl enable httpd
# Iniciar Apache2
/usr/bin/systemctl start httpd
# Habilitar Web-APP
cd /var/www/html
wget
https://aws-tc-largeobjects.s3.amazonaws.com/CUR-TF-100-ACCLFO-2/lab5-rds/lab-app
-php7.zip
unzip lab-app-php7.zip -d /var/www/html/
chown apache:root /var/www/html/rds.conf.php
```

Bootstrapping, código fornecido que é executado quando um computador é inicializado.

Neste caso para o servidor se tornar um servidor Web com APP para Sobrecarregar a máquina virtual e fazer com que o Auto Scaling comece a funcionar.

Etapa 4 – nesta etapa de armazenamento vamos manter o padrão de 8GB.

1. Selecionar a AMI 2. Escolher tipo de instância 3. Configurar instância 4. Adicionar armazenamento 5. Adicionar Tags 6. Configure o security group 7. Análise

Etapa 4: Adicionar armazenamento

Sua instância será executada com as seguintes configurações de dispositivo de armazenamento. Você pode anexar volumes EBS adicionais e volumes de armazenamento de instâncias à sua instância ou editar as configurações do volume raiz. Você também pode anexar volumes EBS adicionais depois de executar uma instância, mas não volumes de armazenamento de instâncias. [Saiba mais](#) sobre opções de armazenamento no Amazon EC2.

Tipo de volume ⓘ	Dispositivo ⓘ	Snapshot ⓘ	Tamanho (GiB) ⓘ	Tipo de volume ⓘ	IOPS ⓘ	Transferência (MB/s) ⓘ	Excluir no encerramento ⓘ	Criptografia ⓘ
Root	/dev/xvda	snap-06705e15d64acb59a	8	Finalidade geral de SSD (gp2)	100 / 3000	N/D	☒	Não criptogra

Adicionar novo volume

Os clientes qualificados para o nível gratuito podem obter até 30 GB de armazenamento de uso geral de EBS (SSD) ou magnético. [Saiba mais](#) sobre a qualificação e restrições de utilização do nível de uso gratuito.

Shared file systems ⓘ

You currently don't have any file systems on this instance. Select "Add file system" button below to add a file system.

Cancelar Anterior Verificar e ativar Próximo: Adicionar Tags

Esta opção "Excluir no encerramento", precisa estar avida para quando excluirmos a instâncias o volume não fique consumindo nossos créditos. Se tiver em produção deixe desmarcado para evitar exclusões desnecessárias.

Clique em "Próximo: Adicionar Tags";

Etapa 5 – nesta etapa de Tags, vamos colocar:

Chave: **Name** e Valor: **Servidor Web APP Auto**

Esse nome será usado nas Instâncias, Volumes e Interfaces de Rede.

Etapa 5: Adicionar Tags

Uma tag consiste em um par chave-valor que diferencia maiúsculas de minúsculas. Por exemplo, você poderia definir uma tag com a chave : Uma cópia de uma tag pode ser aplicada a volumes, instâncias ou a ambos.

As tags serão aplicadas a todas as instâncias e volumes. [Saiba mais](#) sobre atribuição de tags aos seus recursos do Amazon EC2.

Chave (até 128 caracteres)	Valor (até 256 caracteres)
Name	Servidor Web APP Auto

Clique em "Próximo: Configure o security group";

Etapa 6 – nesta etapa são os grupos de segurança, já existe um grupo criando anteriormente.

Que liberar a porta de acesso SSH porta 22 e acesso Web HTTP porta 80 para qualquer IP vindo da internet.

Marque "Selecionar um grupo de segurança existente"

a) Seleciona o grupo com nome: GrupoSec-ServidoresLinux

Etapa 6: Configure o security group

Um grupo de segurança é um conjunto de regras de firewall que controla o tráfego da sua instância. Nesta página, você pode adicionar regras para permitir que tráfegos específicos cheguem até a sua instância. Por exemplo, se você quiser configurar um servidor Web e permitir que tráfego da Internet chegue até a sua instância, adicione regras que permitam acesso restrito às portas HTTP e HTTPS. Você pode criar um novo grupo de segurança ou selecionar um dos existentes abaixo. [Saiba mais](#) sobre grupo de segurança do Amazon EC2.

Atribuir um grupo de segurança: ☐ Criar um grupo de segurança novo
☒ Selecionar um grupo de segurança existente

ID do grupo de segurança	Nome	Descrição	Ações
☐ sg-01b4c694108cb0ad	default	default VPC security group	Copiar para novo
☒ sg-053e5bea90b942775	GrupoSec-ServidoresLinux	LiberaHTTP,SSH	Copiar para novo
☐ sg-0c920d44a2493c524	GrupoSecDB	Permitir acesso ao BD	Copiar para novo

Regras de entrada para sg-053e5bea90b942775 (Grupos de segurança selecionados: sg-053e5bea90b942775)

Tipo	Protocolo	Intervalo de Portas	Origem	Descrição
HTTP	TCP	80	0.0.0.0/0	
SSH	TCP	22	0.0.0.0/0	

[Cancelar](#) [Anterior](#) [Verificar e ativar](#)

Veja que as portas SSH 22/TCP e HTTP 80/TCP já estão cadastradas, clique em **Verificar e Ativar**

Etapa 7 – Esta é a última etapa de revisão, temos dois alertas, mas podemos ignorá-los.

Etapa 7: Review Instance Launch

Verifique os detalhes de execução da instância. Você pode voltar para editar alterações para cada seção. Clique em **Executar** para atribuir um par de chaves à sua instância e concluir o processo de execução.

A configuração da instância não está qualificada para o nível de uso gratuito
Para executar uma instância que seja qualificada para nível de uso gratuito, verifique sua seleção de AMI, tipo de instância, opções de configuração ou dispositivos de armazenamento. Saiba mais sobre a qualificação e restrições de uso de [nível de uso gratuito](#).

[Não mostrar me novamente](#)

Melhore a segurança da sua instância. Seu grupo de segurança, ServidorWEBGrupoSeg, está aberto para o mundo.
Sua instância pode ser acessada de qualquer endereço IP. Recomendamos atualizar as regras do seu grupo de segurança para permitir o acesso apenas de endereços IP conhecidos. Você também pode abrir portas adicionais no seu grupo de segurança para facilitar o acesso ao aplicativo ou serviço que está executando, por ex., HTTP (80) para servidores Web. [Editar grupos de segurança](#)

▼ Detalhes da AMI

[Editar AMI](#)



Amazon Linux 2 AMI (HVM) - Kernel 5.10, SSD Volume Type - ami-08e4e35cccc6189f4

qualificado para o

Amazon Linux 2 comes with five years support. It provides Linux kernel 5.10 tuned for optimal performance on Amazon EC2, systemd 219, GCC 7.3, Glibc 2.26, Binutils 2.29.1, and the latest software packages through extras. This AMI is the successor of the Amazon Linux AMI that is n...

Tipo de dispositivo raíz: ebs Tipo de virtualização: hvm

▼ Tipo de instância

[Editar tipo de instância](#)

[Cancelar](#) [Anterior](#) [Executar](#)

Clique em **Executar**

Essa janela será aberta, é o momento para gerar o par de chaves:

Não vamos usar ele agora

×

Selecione um par de chaves existente ou crie um novo par de chaves

chaves

Um par de chaves consiste em uma **chave pública** armazenada pela AWS e um **arquivo de chave privada** que você armazena. Juntos, eles permitem que você se conecte à sua instância com segurança. Em AMIs do Windows, o arquivo de chave privada é necessário para obter a senha usada para fazer login na sua instância. Para AMIs do Linux, o arquivo de chave privada permite fazer SSH com segurança na sua instância. O Amazon EC2 oferece suporte aos tipos de par de chaves ED25519 e RSA.

Observação: O par de chaves selecionado será adicionado ao conjunto de chaves autorizado para essa instância. Saiba mais sobre [Como remover pares de chaves existentes de uma AMI pública](#).

Continuar sem par de chaves

▼

☒ Reconheço que, sem um par de chaves, só posso me conectar a esta instância usando o EC2 Instance Connect ou se eu souber a senha integrada na AMI. Observe que o EC2 Instance Connect é compatível somente com o Amazon Linux 2 e o Ubuntu. [Saiba mais](#).

Cancelar

Executar instâncias

No menu suspenso, selecione **Continuar sem par de chaves**, marque a caixa a com a instrução **"Reconheço que, sem um par de chaves, só posso me conectar usando o EC2 Instance Connect..."** e clique em **Executar Instância**.

Você será direcionado para essa tela:

Launch Status

✓

Sua instância está sendo iniciada

A seguinte execução de instância foi iniciada: i-03e2d0f1576673701 [Exibir log de execução](#)

ℹ

Obter notificação de cobranças estimadas

[Criar alertas de pagamento](#) obter notificação por e-mail quando as cobranças estimadas na sua fatura da AWS ultrapassarem a quantia definida por você (por exemplo, se tiver excedido o nível de uso gratuito).

Como conectar-se à sua instância

Sua instância está sendo iniciada e pode demorar alguns minutos até que esteja no status de **em execução**, quando estará pronta para ser usada. As horas de utilização da sua nova instância iniciarão imediatamente e serão acumuladas até que você interrompa ou encerre sua instância.

Clique em **Exibir instâncias** para monitorar o status da sua instância. Assim que sua instância estiver no status de **executando**, você pode se **conectar** a ela na tela Instâncias. [Saiba como](#) como se conectar à sua instância.

▼ Veja alguns recursos úteis para ajudá-lo a começar

- Como se conectar à sua instância do Linux
- Saiba mais sobre o nível de uso gratuito da AWS

- Amazon EC2: Guia do usuário
- Amazon EC2: Fórum de discussão

Enquanto suas instâncias estão executando, você também pode

- Criar alarmes de verificação de status ser notificado quando essas instâncias forem reprovadas em verificações de status. (Podem se aplicar cobranças adicionais)
- Criar e anexar volumes adicionais do EBS (Podem se aplicar cobranças adicionais)
- Gerenciar grupos de segurança

Exibir instâncias

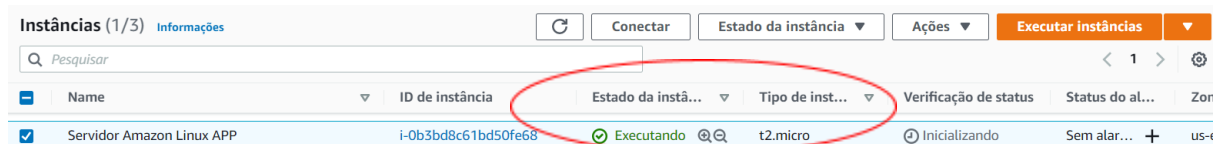
Clique em **Exibir Instâncias**

Etapa 8 – Acesso APP do Servidor Web.

10

Temos as colunas:

- Name "Servidor WEB Descomplica";
- Estado da Instância (**Executando**);
- Tipo da Instância: ex.: **t2.micro**;
- Verificação de Status (Inicializando, **2/2 verificações aprovadas**).



Name	ID de instância	Estado da instância	Tipo de instância	Verificação de status	Status do alarme	Zona de disponibilidade
Servidor Amazon Linux APP	i-0b3bd8c61bd50fe68	Executando	t2.micro	Inicializando	Sem alarme	us-east-1a

Assim que iniciar, faça o teste de acesso WEB.

Instância: i-0b3bd8c61bd50fe68 (Servidor Amazon Linux APP)

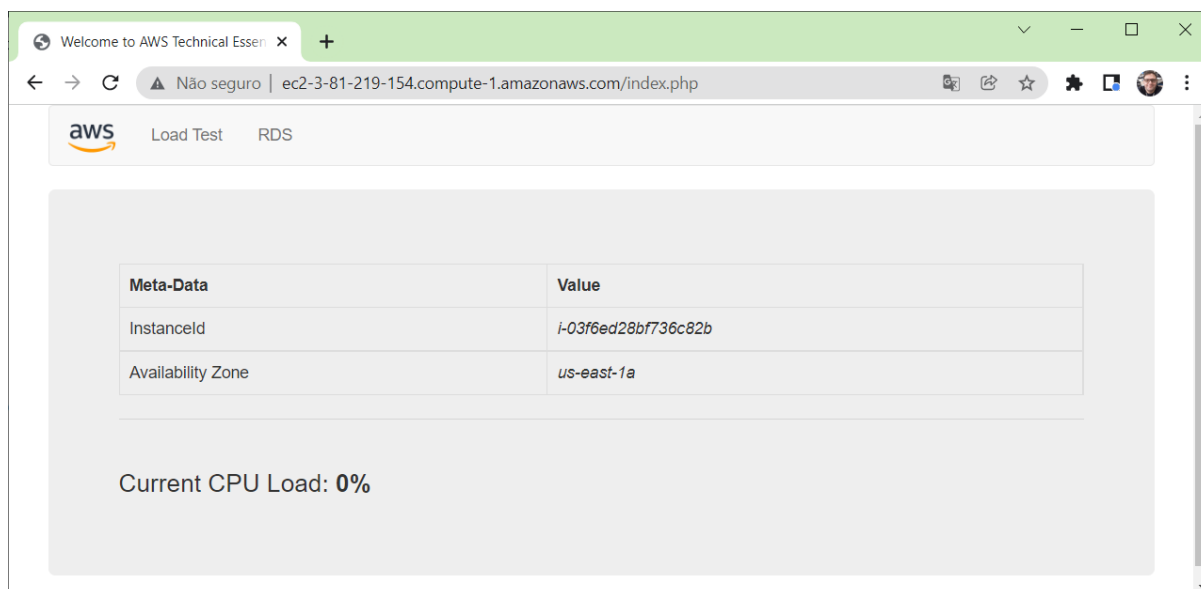


Detalhes	Segurança	Redes	Armazenamento	Verificações de status	Monitoramento	Tags
Resumo da instância						
ID de instância	i-0b3bd8c61bd50fe68 (Servidor Amazon Linux APP)		Endereço IPv4 público	35.175.207.67 endereço aberto		
Endereço IPv6	-		Estado da instância	Executando		
				DNS IPv4 público		
				ec2-35-175-207-67.compute-1.amazonaws.com endereço aberto		

Nesta mesma aba que já está marcada colete a informação: "Nome IPv4 público": ex.: ec2-35-175-207-67.compute-1.amazonaws.com

Complete com: ec2-35-175-207-67.compute-1.amazonaws.com/index.php

Será exibida essa tela:



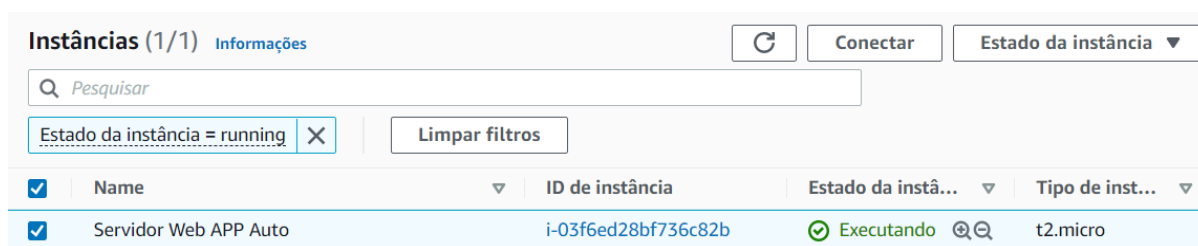
Laboratório 10 concluído, vamos para o próximo laboratório sequencial.

Laboratório 11 - Criando Instância modelo AMI para o Auto Scaling

Nesta tarefa, você criará uma AMI pelo **Servidor Web APP Auto** existente.

Isso salvará o conteúdo do disco de inicialização para que novas instâncias possam ser executadas com conteúdo idêntico.

5. No **Console de Gerenciamento da AWS**, no menu **Serviços**, clique em **EC2**.
6. No painel de navegação esquerdo, clique em **Instâncias**
 - 1° - Confirme se a instância está em execução.
7. Aguarde até que as etapas de **Verificação de status** do **Servidor Web APP Auto**, estão **2/2 verificações aprovadas**.

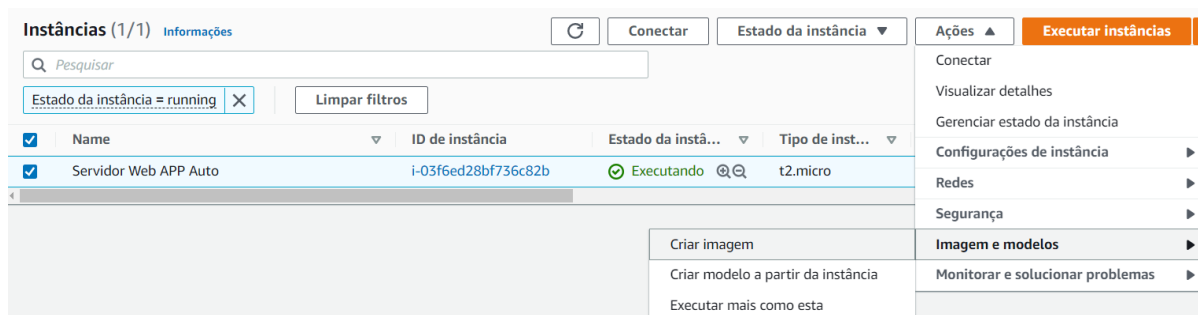


Instâncias (1/1) Informações		Atualizar	Conectar	Estado da instância ▼
Pesquisar				
Estado da instância = running ✕		Limpar filtros		
☒	Name	ID de instância	Estado da instância	Tipo de instância
☒	Servidor Web APP Auto	i-03f6ed28bf736c82b	Executando	t2.micro

8. Clique em Refresh  (atualizar) até chegar ao estado desejado.

Agora, você criará uma AMI com base nessa instância.

9. Selecione o **Servidor Web APP Auto**.
10. No menu **Ações**, clique em **Imagem e Modelos > Criar Imagem**



Instâncias (1/1) Informações		Atualizar	Conectar	Estado da instância ▼	Ações ▲	Executar instâncias
Pesquisar						
Estado da instância = running ✕		Limpar filtros				
☒	Name	ID de instância	Estado da instância	Tipo de instância		
☒	Servidor Web APP Auto	i-03f6ed28bf736c82b	Executando	t2.micro		
					Criar imagem	Imagem e modelos ▶
					Criar modelo a partir da instância	Monitorar e solucionar problemas ▶
					Executar mais como esta	

Configure:

EC2 > Instâncias > i-03f6ed28bf736c82b > Criar imagem

Criar imagem Informações

Uma imagem (também conhecida como AMI) define os programas e as configurações que são aplicados quando você executa uma instância do EC2. É possível criar uma imagem a partir de:

ID de instância
i-03f6ed28bf736c82b (Servidor Web APP Auto)

Nome da imagem
AMI do Servidor Web APP
Máximo de 127 caracteres. Não é possível modificar após a criação.

Descrição da imagem - *opcional*
AMI do Servidor Web para o LAB
Máximo de 255 caracteres

Não reinicializar
☐ Habilitar

- o Nome da imagem: AMI do Servidor Web APP
- o Descrição da imagem: AMI do Servidor Web para o LAB
- o Deixe desmarcado: Não reiniciar

Volumes de instâncias

Tipo de volume	Dispositivo	Snapshot	Tamanho	Tipo de volume	IOPS	Taxa de transferência	Excluir no encerramento	Criptografado
EBS	/dev/x...	Criar novo snapshot a p...	8	SSD de uso geral do EB...	100		☒ Habilitar	☐ Habilitar

Adicionar volume

ⓘ Durante o processo de criação de imagens, o Amazon EC2 cria um snapshot de cada um dos volumes acima.

Tags — *opcional*
Uma tag é um rótulo que você atribui a um recurso da AWS. Cada tag consiste em uma chave e um valor opcional. É possível usar as tags para pesquisar e filtrar seus recursos ou para controlar seus custos da AWS.

☒ Marcar imagem e snapshots juntos
Marque a imagem e os snapshots com a mesma tag.

☐ Marcar imagem e snapshots separadamente
Marque a imagem e os snapshots com tags diferentes.

Nenhuma tag associada ao recurso.

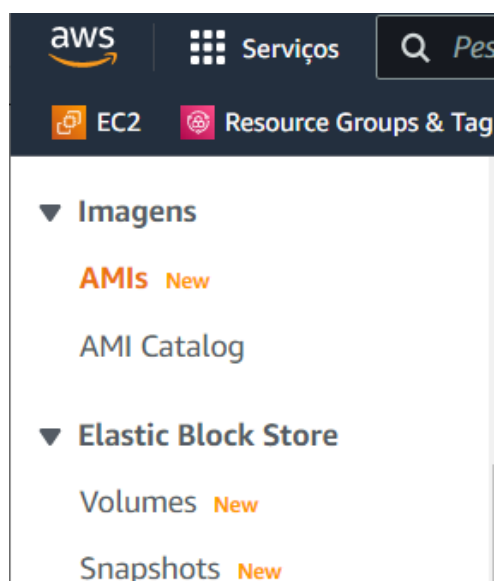
Adicionar tag
Você pode adicionar mais 50 tags

Cancelar Criar imagem

11. Clique em **Criar imagem**

No **Console de Gerenciamento da AWS**, no menu **Serviços**, clique em **EC2**.

No painel de navegação esquerdo, clique em **AMIs**



Imagens de máquina da Amazon (AMIs) (1/1) Informações					EC2 Image Builder		Ações ▾	Executar instância
De minha propriedade ▾					Q Pesquisar			
Nome da AMI	Origem	Proprietário	Visibilidade	Status				
AMI do Servidor Web APP	324815927342/AMI do Servidor Web APP	324815927342	Privada	Pendente				

Aguarde o Status mudar de **Pendente** para **Disponível**

Demora até 5 minutos para criar.

Máquina modelo criada com sucesso:

Imagens de máquina da Amazon (AMIs) (1)

Informações

EC2 Image Builder

Ações ▾

Executar ins

De minha propriedade ▾

Nome da AMI ▾	Origem ▾	Proprietário ▾	Visibilidade ▾	Status ▾
AMI do Servidor Web APP	324815927342/AMI do Servidor Web APP	324815927342	Privada	Disponível

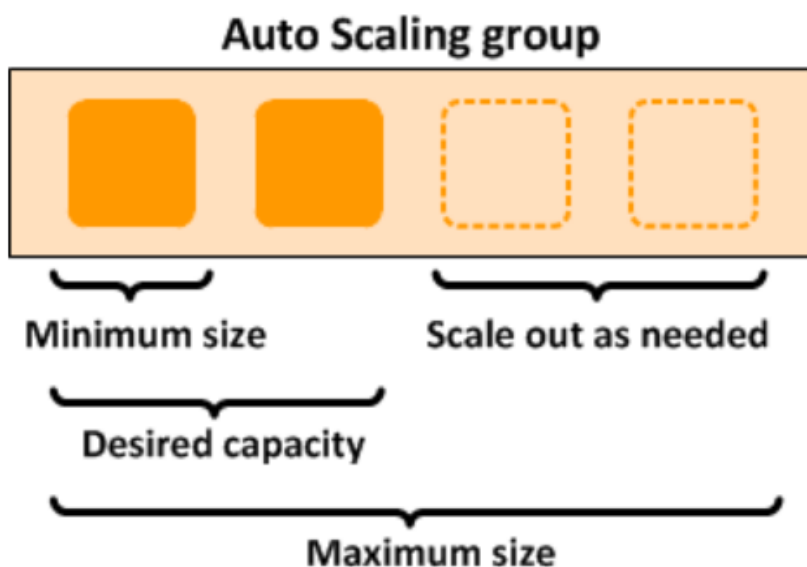
Você usará essa AMI ao iniciar o grupo de Auto Scaling posteriormente, no laboratório 12.

Auto Scaling Group

É um conjunto de instâncias do Amazon EC2 que são tratadas como um agrupamento lógico para fins de escalabilidade e gerenciamento automáticos. O tamanho de um grupo de Auto Scaling depende do número de instâncias definidas como a capacidade desejada.

Você pode ajustar seu tamanho para atender à demanda manualmente ou usando a escalabilidade automática.

A seguir, você verá uma imagem ilustrativa contendo as informações necessárias ao criar um Auto Scaling Group e poder entender um pouco melhor como ele funciona na prática logo em seguida.



Fonte:

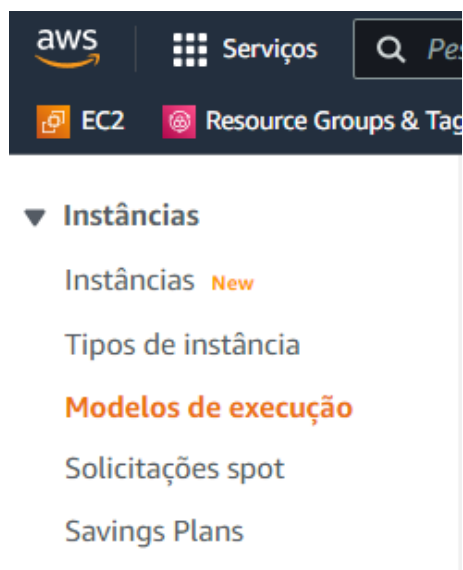
https://docs.aws.amazon.com/pt_br/autoscaling/ec2/userguide/what-is-amazon-ec2-auto-scaling.html

- Minimum size (tamanho mínimo): O menor número de instâncias ativas que seu Auto Scaling pode ter.
- Desired capacity (capacidade desejada): O número de instâncias recomendado.
- Scale out as needed (dimensionar conforme necessário): Espaço disponível para expansão do número de instâncias.
- Maximum size (tamanho máximo): O número máximo de instâncias que podem ser executadas ao mesmo tempo.

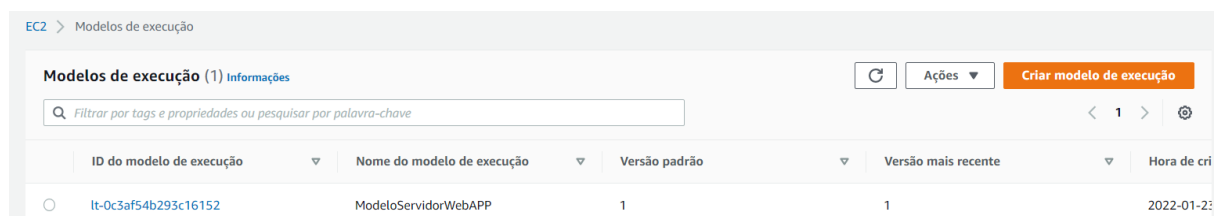
Laboratório 12 – Auto Scaling Group

Etapa 1 – Criar Modelo de Execução

Entre no console AWS, **EC2** no menu à esquerda clique em **Modelos de execução**



Dentro de **Modelos de execução**



Clique em **Criar modelo de execução**

EC2 > Modelos de execução > Criar modelo de execução

Criar modelo de execução

A criação de um modelo de execução permite criar uma configuração de instância salva que pode ser reutilizada, compartilhada e executada posteriormente. Os modelos podem ter várias versões.

Nome e descrição do modelo de execução

Nome do modelo de execução - *obrigatório*

Deve ser exclusivo para essa conta. Máximo de 128 caracteres. Nenhum espaço ou caracteres especiais, como "&", "*", "@".

Descrição da versão do modelo

Máximo de 255 caracteres

Orientação sobre o Auto Scaling [Informações](#)

Selecione essa opção se você pretende usar esse modelo com o Auto Scaling do EC2

- ☒ Fornecer orientação para me ajudar a configurar um modelo que eu possa usar com o Auto Scaling do EC2

► Tags de modelo

► Modelo de origem

Preencha o nome: **ModeloServidorWebAPP**

Descrição o nome: **Modelo para ser usado no Grupo de Auto Scaling**

Marque: **"Fornecer orientação para me ajudar a configurar um modelo..."**

Não vamos preencher **Tags de modelo**

Não vamos mudar **Modelo de origem**

Dentro de conteúdo do modelo, marque **"Minhas AMIs"**

Conteúdo do modelo de execução

Especifique os detalhes do modelo de execução abaixo. Se um campo for deixado em branco, isso fará com que ele não seja incluído no modelo de execução.

▼ Imagens de aplicativo e sistema operacional (Imagem de máquina da Amazon) - obrigatório [Informações](#)

Uma AMI é um modelo que contém a configuração do software (sistema operacional, servidor de aplicativos e aplicativos) necessária para executar a instância. Pesquise ou navegue pelas AMIs se você não estiver vendo o que está buscando abaixo

Recentes

Minhas AMIs

Início rápido

☒ De minha propriedade

☐ Compartilhado comigo


Procurar mais AMIs
Incluindo AMIs da AWS, do Marketplace e da comunidade

Imagem de máquina da Amazon (AMI)

AMI do Servidor Web APP
ami-0fdb69d8671ad9c4a
2022-01-23T16:28:15.000Z Virtualização: hvm ENA habilitado: true Tipo de dispositivo raiz: ebs

Descrição

AMI do Servidor Web para o LAB

Arquitetura	ID da AMI
x86_64	ami-0fdb69d8671ad9c4a

Selecione a **AMI do Servidor Web APP**

Dentro de “**Tipo de instância**”, selecione: **t2.micro**

▼ Tipo de instância [Informações](#)

Avançado

Tipo de instância

t2.micro
Família: t2 1 vCPU 1 GiB Memória
Sob demanda Linux definição de preço: 0.0116 USD por hora
Sob demanda Windows definição de preço: 0.0162 USD por hora

Qualificado para o nível gratuito

▼

[Comparar tipos de instância](#)

▼ Par de chaves (login) [Informações](#)

Você pode usar um par de chaves para se conectar com segurança à sua instância. Certifique-se de ter acesso ao par de chaves selecionado antes de executar a instância.

Nome do par de chaves

Não incluir no modelo de execução

▼

 [Criar novo par de chaves](#)

Dentro de “Par de chaves”, não selecione nada, vamos aumentar nossa segurança.

Dentro de “Configurações de rede”, selecione: **Sub-Rede-Publica-1**

▼ Configurações de rede

Sub-rede [Informações](#)

subnet-06e2c1bb9df5c1ab9
VPC: vpc-091dbf50d1445f94c Proprietário: 324815927342
Zona de disponibilidade: us-east-1a Endereços IP disponíveis: 251

Sub-Rede-Privada-1

▼

 [Criar nova sub-rede](#)

Quando você especifica uma sub-rede, uma interface de rede é adicionada automaticamente ao modelo.

Firewall (grupos de segurança)

Um grupo de segurança é um conjunto de regras de firewall que controlam o tráfego para sua instância. Adicione regras para permitir que o tráfego específico alcance sua instância.

☒ Selecionar grupo de segurança existente

☐ Criar grupo de segurança

Grupos de segurança comuns [Informações](#)

Selecionar grupos de segurança

▼

GrupoSec-ServidoresLinux sg-053e5bea90b942775 ✕
VPC: vpc-091dbf50d1445f94c

 [Comparar regras de grupo de segurança](#)

Os grupos de segurança que você adicionar ou remover aqui serão adicionados ou removidos em todas as suas interfaces de rede.

► Configuração avançada de rede

Dentro de “Firewall”, “Selecionar grupo de segurança existente”, selecione: **GrupoSec-ServidoresLinux**.

20

Dentro de “**Detalhes avançados**”, role até “**Nome do host DNS**”, selecione: **Habilitar solicitações de DNS IPv4 (registro A)**.

▼ Detalhes avançados [Informações](#)

Opção de compra [Informações](#)

☐ Solicitar instâncias spot

Se Spot for selecionado, você não poderá criar um grupo de Auto Scaling que abranja várias opções de definição de preço e tipos de instância

Perfil de instância do IAM [Informações](#)

Não incluir no modelo de execução ▼

 [Criar novo perfil do IAM](#) 

Tipo de nome do host [Informações](#)

Não incluir no modelo de execução ▼

Nome do host DNS [Informações](#)

☒ Habilitar solicitações de DNS IPv4 (registro A) com base em recursos

☐ Habilitar solicitações de DNS IPv6 (registro AAAA) com base em recursos

No menu do lado direito “**Resumo**”

▼ Resumo

Imagem do software (AMI)

AMI do Servidor Web para o LAB...[read more](#)
ami-0fdb69d8671ad9c4a

Tipo de servidor virtual (tipo de instância)

t2.micro

Firewall (grupo de segurança)

GrupoSec-ServidoresLinux

Armazenamento (volumes)

1 volume(s) - 8 GiB

i **Nível gratuito:** In your first year includes 750 ✕
hours of t2.micro (or t3.micro in the Regions
in which t2.micro is unavailable) instance
usage on free tier AMIs per month, 30 GiB of
EBS storage, 2 million IOs, 1 GB of
snapshots, and 100 GB of bandwidth to the
internet

Cancelar

Criar modelo de execução

Clique em “**Criar um modelo de execução**”

Etapa 2 – Criar Grupo do Auto Scaling

Entre no console AWS, **EC2** no menu à esquerda clique em **Grupos do Auto Scaling**

Você será redirecionado para essa página:

Amazon EC2 Auto Scaling

ajuda a manter a
disponibilidade de seus
aplicativos

Os grupos de Auto Scaling são coleções de instâncias do Amazon EC2 que permitem recursos de escalabilidade automática e gerenciamento de frotas. Esses recursos ajudam você a manter a integridade e a disponibilidade dos seus aplicativos.

Criar o grupo do Auto Scaling

Comece a usar o EC2 Auto Scaling criando um grupo de Auto Scaling.

Criar o grupo do Auto Scaling

Clique em “**Criar o grupo de Auto Scaling**”:

Etapa 1: Escolha o nome e o modelo de execução:

Etapa 1
Escolher o modelo ou a configuração de execução

Etapa 2
Escolha as opções de execução da instância

Etapa 3 (opcional)
Configurar opções avançadas

Etapa 4 (opcional)
Configurar políticas de escalabilidade e tamanho do

Escolher o modelo ou a configuração de execução [Informações](#)

Especifique um modelo de execução que contenha configurações comuns a todas as instâncias do EC2 executadas por esse grupo de Auto Scaling. Se você usa configurações de execução no momento, pode considerar a migração para modelos de execução.

Nome

Grupo do Auto Scaling
Insira um nome para identificar o grupo.

Deve ser exclusivo para essa conta na região atual e não ter mais de 255 caracteres.

- Informe o nome: **GrupoAutoScaling**
- Escolha em “Modelo de execução”: **ModeloServidorWebAPP**

grupo

Etapa 5 (opcional)
Adicionar notificações

Etapa 6 (opcional)
Adicionar tags

Etapa 7
Análise

Modelo de execução [Informações](#) [Alternar para configuração de execução](#)

Modelo de execução
Escolha um modelo de execução que contenha as configurações no nível da instância, como a Imagem de máquina da Amazon (AMI), o tipo de instância, o par de chaves e os grupos de segurança.

[Criar um modelo de execução](#)

Versão
 [Criar uma versão de modelo de execução](#)

Descrição Modelo para ser usado no Grupo de Auto Scalling	Modelo de execução ModeloServidorWebAPP lt-0c3af54b293c16152	Tipo de instância t2.micro
AMI ID ami-0fdb69d8671ad9c4a	Grupos de segurança -	Solicitar instâncias Spot Não
Nome do par de chaves -	IDs de grupo de segurança -	

Detalhes adicionais

Armazenamento (volumes) -	Data de criação Sun Jan 23 2022 14:10:52 GMT-0300 (Horário Padrão de Brasília)
------------------------------	--

Cancelar

Próximo

Manter versão “Default(1)” e Clique em **Próximo**.

Etapas 2: Selecione sua VPC e AZ's:

Etapa 1
Escolher o modelo ou a configuração de execução

Etapa 2
Escolha as opções de execução da instância

Etapa 3 (opcional)
Configurar opções avançadas

Etapa 4 (opcional)
Configurar políticas de escalabilidade e tamanho do grupo

Etapa 5 (opcional)
Adicionar notificações

Etapa 6 (opcional)
Adicionar tags

Etapa 7
Análise

Escolha as opções de execução da instância [Informações](#)

Escolha o ambiente de rede da VPC no qual as instâncias são executadas e personalize os tipos de instância e as opções de compra.

Rede [Informações](#)

Para a maioria dos aplicativos, você pode usar várias zonas de disponibilidade e permitir que o EC2 Auto Scaling equilibre suas instâncias entre as zonas. A VPC padrão e as sub-redes padrão são adequadas para o início de uso rápido.

VPC
Escolha a VPC que define a rede virtual do grupo do Auto Scaling.

vpc-091dbf50d1445f94c (Lab-VPC)
10.0.0.0/16

Criar uma VPC [🔗](#)

Zonas de disponibilidade e sub-redes
Defina quais zonas de disponibilidade e sub-redes seu grupo do Auto Scaling pode usar na VPC escolhida.

Selecionar zonas de disponibilidade e sub-redes

us-east-1a | subnet-058d66561031d3e28 (Sub-Rede-Publica-1)
10.0.0.0/24

us-east-1b | subnet-07c6207930706f3a4 (Sub-Rede-Publica-2)
10.0.2.0/24

Criar uma sub-rede [🔗](#)

Usar a VPC **Lab-VPC** e as AZs **Sub-Rede-Publica-1** e **Sub-Rede-Publica-2**.

Requisitos de tipo de instância [Informações](#)

Você pode manter os mesmos atributos da instância ou tipo de instância do modelo de execução ou modificar o modelo de execução especificando atributos de instância diferentes ou adicionando manualmente tipos de instância.

Modelo de execução
[ModeloServidorWebAPP](#)
lt-0c3af54b293c16152

Versão
Default

Descrição
Modelo para ser usado no Grupo de Auto Scaling

Tipo de instância
t2.micro

Cancelar

Anterior

Pular para revisão

Próximo

Clique em **Próximo**.

Etapa 3:

Configurar opções avançadas [Info](#)

Escolha um load balancer para distribuir o tráfego de entrada do aplicativo entre instâncias para torná-lo mais confiável e facilmente escalável. Você também pode definir opções que oferecem mais controle sobre as substituições e o monitoramento de verificação de integridade.

Balanceamento de carga - *opcional* [Info](#)

Use as opções abaixo para anexar seu grupo de Auto Scaling a um load balancer existente ou a um novo load balancer definido por você.

☒ Nenhum load balancer
O tráfego para seu grupo de Auto Scaling não será liderado por um load balancer.

☐ Anexar a um load balancer existente
Escolha entre seus balanceadores de carga existentes.

☐ Anexar a um novo load balancer
Crie rapidamente um load balancer básico para anexar ao seu grupo de Auto Scaling.

Verificações de integridade - *opcional*Tipo de verificação de integridade [Info](#)

O EC2 Auto Scaling substitui automaticamente as instâncias que falham nas verificações de integridade. Se você habilitar o balanceamento de carga, poderá habilitar as verificações de integridade do ELB além das verificações de integridade do EC2 que estão sempre habilitadas.

☒ EC2 ☐ ELB

Período de tolerância da verificação de integridade

A quantidade de tempo até que o EC2 Auto Scaling execute a primeira verificação de integridade em novas instâncias depois que elas são colocadas em serviço.

segundos

- Aqui, nas configurações avançadas, você pode anexar diretamente com um load balancer (balanceador de carga) existente.
- Porém, como esse conteúdo não contempla o ELB, não estaremos fazendo por enquanto.

Configurações adicionais - *opcional*Monitoramento [Informações](#)

☒ Habilitar coleta de métricas de grupo no CloudWatch

[Cancelar](#)[Anterior](#)[Pular para revisão](#)[Próximo](#)

Habilitar coleta de métricas de grupo no CloudWatch e clique em **Próximo**.

Etapa 4:

Configurar políticas de escalabilidade e tamanho do grupo Info

Defina a capacidade desejada, mínima e máxima de seu grupo do Auto Scaling. Opcionalmente, você pode adicionar uma política de escalabilidade para escalar dinamicamente o número de instâncias no grupo.

Tamanho do grupo - *opcional* Info

Especifique o tamanho do grupo do Auto Scaling alterando a capacidade desejada. Você também pode especificar os limites de capacidade mínima e máxima. Sua capacidade desejada deve estar dentro do intervalo limite.

Capacidade desejada

Capacidade mínima

Capacidade máxima

Políticas de escalabilidade - *opcional*

Escolha se deseja usar uma política de escalabilidade para redimensionar dinamicamente seu grupo de Auto Scaling para atender às alterações na demanda. Info

☐ Política de escalabilidade de rastreamento de destino
Escolha um resultado desejado e deixe-o com a política de escalabilidade para adicionar e remover capacidade conforme necessário para alcançar esse resultado.

☒ Nenhum

Proteção contra redução de instâncias - *opcional*

Proteção contra redução de instâncias
Se a opção Proteção contra redução estiver habilitada, as instâncias recém-iniciadas serão protegidas contra redução por padrão.

☐ Habilitar a proteção contra redução de instâncias

- Aqui se encontra a tela de configuração que foi representada pela imagem ilustrativa anteriormente, com as mesmas informações:
 - Capacidade mínima: 1
 - Capacidade desejada: 2
 - Capacidade máxima: 4
- Política de escalabilidade – opcional, manter marcado “Nenhum”.

Proteção contra redução de instâncias - *opcional*

Proteção contra redução de instâncias

Se a opção Proteção contra redução estiver habilitada, as instâncias recém-iniciadas serão protegidas contra redução por padrão.

☐ Habilitar a proteção contra redução de instâncias

Cancelar

Anterior

Pular para revisão

Próximo

Desmarcar “Habilitar a proteção contra redução de instâncias” e clique em **Próximo**.

Etapa 5:

Etapa 1
Escolher o modelo ou a configuração de execução

Etapa 2
Escolha as opções de execução da instância

Etapa 3 (opcional)
Configurar opções avançadas

Etapa 4 (opcional)
Configurar políticas de escalabilidade e tamanho do grupo

**Etapa 5 (opcional)
Adicionar notificações**

Etapa 6 (opcional)
Adicionar tags

Etapa 7
Análise

Adicionar notificações Info

Envie notificações para tópicos do SNS sempre que o Auto Scaling do Amazon EC2 iniciar ou encerrar as instâncias do EC2 em seu grupo de Auto Scaling.

Adicionar notificação

Cancelar

Anterior

Pular para revisão

Próximo

- Você também pode enviar notificações para tópicos do SNS (Simple Notification Service) mas neste caso não vamos usar.

Etapa 6 - Tags:

Etapa 1

Escolher o modelo ou a configuração de execução

Etapa 2

Escolha as opções de execução da instância

Etapa 3 (opcional)

Configurar opções avançadas

Etapa 4 (opcional)

Configurar políticas de escalabilidade e tamanho do grupo

Etapa 5 (opcional)

Adicionar notificações

Etapa 6 (opcional)

Adicionar tags

Etapa 7

Análise

Adicionar tags Informações

Adicione tags para ajudá-lo a pesquisar, filtrar e rastrear seu grupo do Auto Scaling na AWS. Você também pode optar por adicionar automaticamente essas tags às instâncias quando elas forem executadas.

❗ Opcionalmente, você poderá adicionar tags a instâncias (e seus volumes de EBS anexados) especificando tags em seu modelo de execução. No entanto, recomendamos cautela, pois os valores de tag das instâncias de seu modelo de execução serão substituídos se houver chaves duplicadas especificadas para o grupo do Auto Scaling.

Tags (1)

Chave	Valor - opcional	Adicionar tags em novas instâncias	
Name	WebAutoScaling	☒	Remover
Adicionar tag			

49 restantes

Cancelar

Anterior

Próximo

Adicionar Tag: **WebAutoScaling**

Adicionar tags às suas configurações na AWS são de extrema ajuda em casos de grandes empresas, podendo te poupar muito tempo na procura por erros e mais.

Etapa 7 - Revisão:

Aqui, você irá revisar suas configurações, mostrando as 6 etapas anteriores.

Análise Informações

Etapa 1: escolher o modelo ou a configuração de execução

Detalhes do grupo

Grupo do Auto Scaling

GrupoAutoScaling

Modelo de execução

Modelo de execução

[ModeloServidorWebAPP](#)

lt-0c3af54b293c16152

Versão

Default

Descrição

Modelo para ser usado no Grupo de Auto Scaling

Etapa 2: Escolher as opções de execução de instâncias

Editar

Rede

Rede

VPC

vpc-091dbf50d1445f94c [🔗](#)

Zona de disponibilidade

Sub-rede

us-east-1a

subnet-058d66561031d3e28 [🔗](#)

10.0.0.0/24

us-east-1b

subnet-07c6207930706f3a4 [🔗](#)

10.0.2.0/24

Requisitos de tipo de instância

Esse grupo do Auto Scaling aderirá ao modelo de execução.

Etapa 3: Configurar opções avançadas

Editar

Balanceamento de carga

Verificações de integridade

Tipo de verificação de integridade
EC2Período de tolerância da verificação
de integridade
300 segundos

Configurações adicionais

Monitoramento
Habilitado

Etapa 4: configurar políticas de escalabilidade e tamanho do grupo

Editar

Tamanho do grupo

Capacidade desejada

2

Capacidade mínima

1

Capacidade máxima

6

Política de escalabilidade

Nenhuma política de escalabilidade

Proteção contra redução de instâncias

Proteção contra redução de instâncias

☐ Habilitar a proteção de instância contra redução

Etapa 5: Adicionar notificações

Editar

Notificações

Nenhuma notificação

Etapa 6: Adicionar tags

Editar

Tags (1)

Chave

Valor

Adicionar tags em novas instâncias

Name

TagAutoScaling

Sim

Cancelar

Criar o grupo do Auto Scaling

Clique em “**Criar o grupo do Auto Scaling**”.Veja em **EC2** e **Instâncias** que as máquinas começaram a ser criadas:

Instâncias (3) Informações						
Pesquisar						
☐	Name	ID de instância	Estado da instâ...	Tipo de inst...	Verificação de status	
☐	Servidor Web APP Auto	i-03f6ed28bf736c82b	Executando	t2.micro	2/2 verificações apro	
☐	TagAutoScaling	i-0b1846e3b060d9d2d	Executando	t2.micro	Iniciando	
☐	TagAutoScaling	i-0770c74e7e55e7b91	Executando	t2.micro	Iniciando	

Laboratório 13 - Aplicando Auto-Scaling e ambiente com muitos acessos

Este laboratório orienta você a usar os serviços Elastic Load Balancing (ELB) e Auto Scaling para balancear cargas e dimensionar automaticamente a infraestrutura.

O **Auto Scaling** ajuda a manter a disponibilidade da aplicação e permite aumentar ou reduzir a capacidade do Amazon EC2 de forma automática, de acordo com condições definidas por você. Você pode usar o Auto Scaling para ajudar a garantir que está executando o número desejado de instâncias do Amazon EC2. O Auto Scaling também pode aumentar automaticamente o número de instâncias do Amazon EC2 durante picos de demanda para manter a performance, e diminuir a capacidade durante períodos ociosos para reduzir os custos. O Auto Scaling é ideal para aplicações com padrões de demanda estáveis ou que passam por variações de utilização horárias, diárias ou semanais.

Objetivos

Depois de concluir este laboratório, você será capaz de:

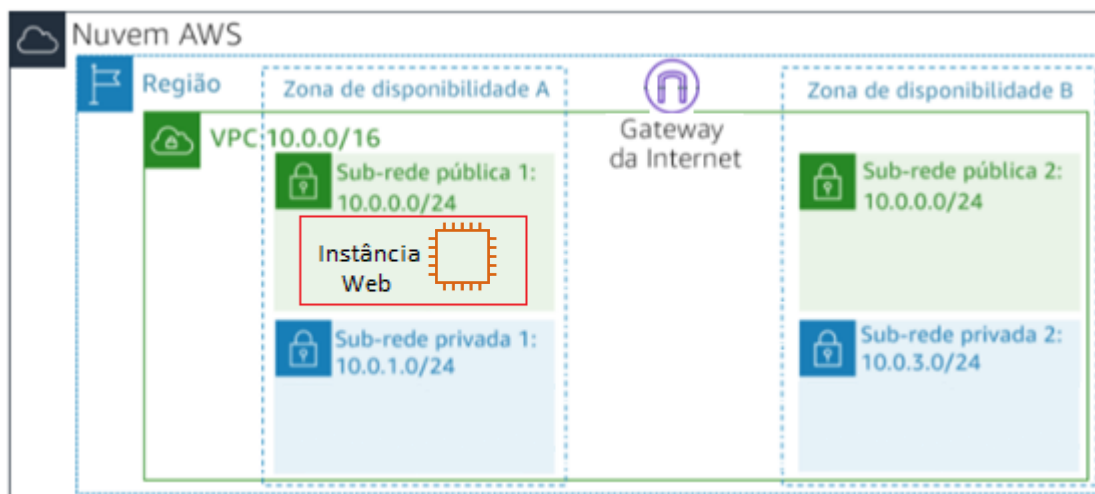
- Criar uma imagem de máquina da Amazon (AMI) por meio de uma instância em execução (**Servidor Web APP Auto**);
- Criar uma configuração de execução e um grupo de Auto Scaling;
- Ajustar a escala automaticamente de novas instâncias em uma sub-rede privada;
- Criar alarmes do Amazon CloudWatch e monitorar a performance da sua infraestrutura.

Duração

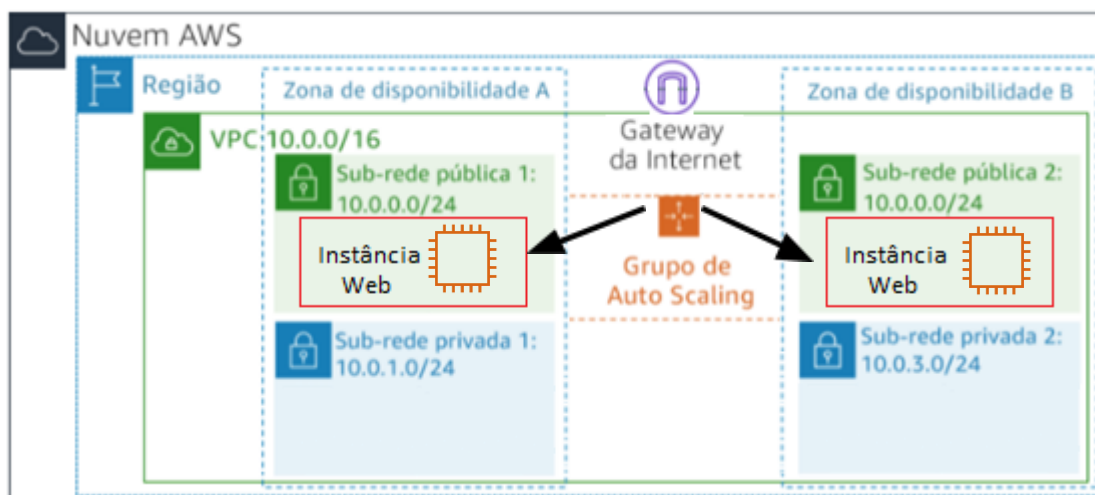
Este laboratório tem duração aproximada de **10 minutos**.

Cenário

Você começa com a seguinte infraestrutura:



O estado final da infraestrutura é:



Acessar o Console de Gerenciamento da AWS

1. Na parte superior destas instruções, clique em **Start Lab** (Iniciar laboratório) para iniciar o laboratório.

Um painel Start Lab (Iniciar laboratório) é aberto exibindo o status do laboratório.

2. Aguarde até ver a mensagem "**Lab status: in creation** (Status do laboratório: em criação)" e clique no **X** para fechar o painel Start Lab (Iniciar laboratório).

Observação: pode levar aproximadamente 20 minutos ou mais para que o status do laboratório mude para "ready" (pronto).

3. Na parte superior destas instruções, clique em **AWS**

Isso abrirá o Console de Gerenciamento da AWS em uma nova guia do navegador. O sistema fará o seu login automaticamente.

Dica: se uma nova guia do navegador não for aberta normalmente, um banner ou um ícone na parte superior do navegador indicará que o navegador está impedindo que o site abra janelas pop-up. Clique no banner ou ícone e escolha "Allow pop ups" (Permitir pop-ups).

4. Organize a guia do Console de Gerenciamento da AWS para que ela seja exibida com estas instruções. Em um cenário ideal, você poderá ver as duas guias do navegador ao mesmo tempo, para facilitar o acompanhamento das etapas do laboratório.

Já passamos pelos Laboratórios

Laboratório 10: Criar uma Instância EC2 Amazon Linux com APP

Laboratório 11: Criar um Modelo AMI (Amazon Machine Image) para o Auto Scaling

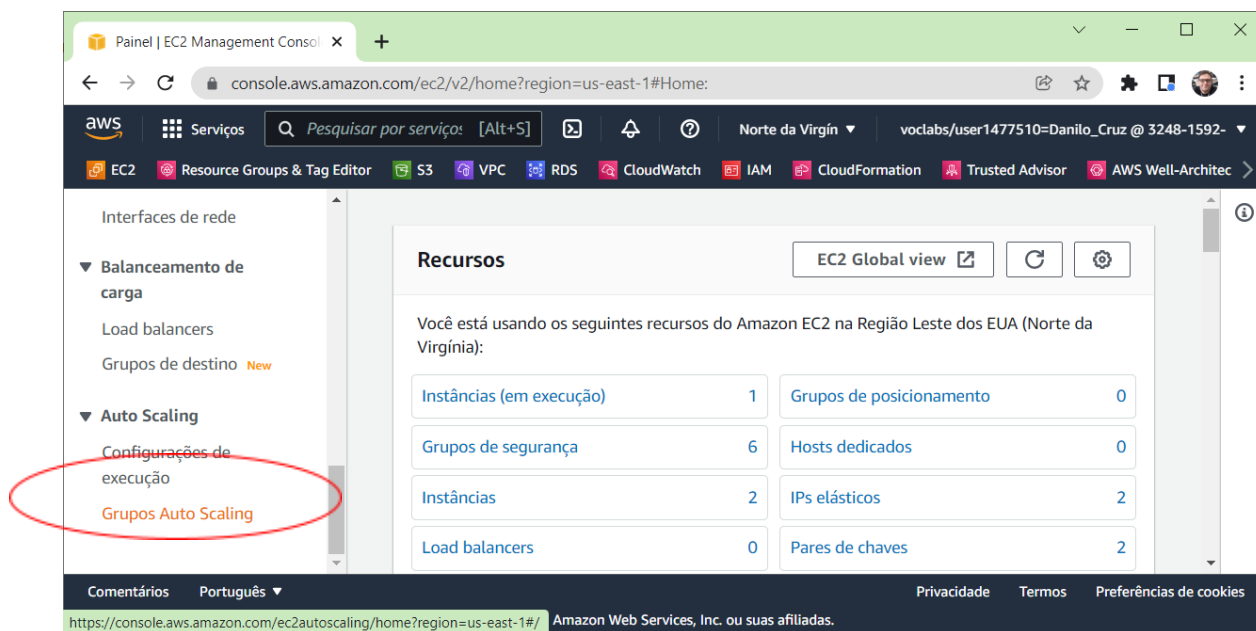
Laboratório 12: Criamos o Auto Scalling

Nesta tarefa, você criou regras de funcionamento do Auto Scalling onde identificou que 2 instâncias EC2 em zonas de disponibilidade diferente foram criadas.

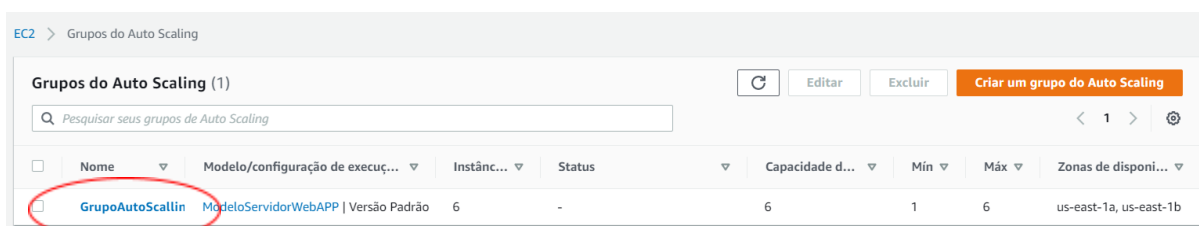
Laboratório 13 - Aplicando Auto-Scaling e ambiente com muitos acessos

Etapa 1 - Vamos criar uma política de escalabilidade automática

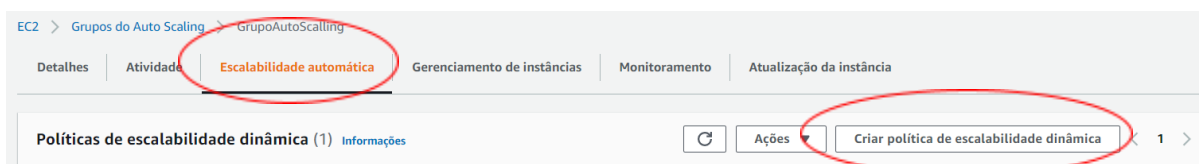
No Console **AWS**, em **EC2** clique em **Grupos do Auto Scaling:**



Clique em **GrupoAutoScaling**



Dentro do **GrupoAutoScaling** escolha a aba “**Escalabilidade automática**”



Clique em “**Criar política de escalabilidade dinâmica**”

EC2 > Grupos do Auto Scaling > GrupoAutoScaling

Criar política de escalabilidade dinâmica

Tipo de política

Escalabilidade de rastreamento de destino ▼

Nome da política de escalabilidade

Politica-Rastreio-de-Consumo-CPU

Tipo de métrica

Média de utilização da CPU ▼

Valor de destino

50

As instâncias precisam

300 segundos de aquecimento antes de incluir na métrica

☐ Desabilitar a redução para criar apenas uma política de expansão

Cancelar

Criar

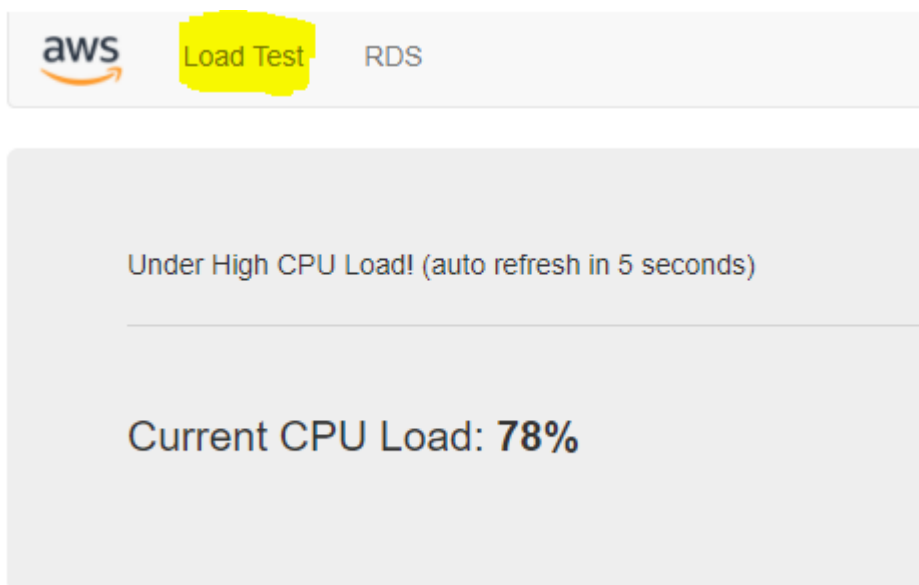
Dentro de “Criar política de escalabilidade dinâmica

- Escolher o “Tipo de política” escolha: **Política de escalabilidade de rastreamento de destino**;
- Nome da política: **Politica-Rastreio-de-Consumo-CPU**
- Tipo de métrica: **Média de Utilização média da CPU**;
- Valor de destino: **50**

Isso informa ao Auto Scaling para manter uma utilização de CPU *média* em *todas as instâncias* de 50%.

O Auto Scaling adicionará ou removerá automaticamente a capacidade conforme necessário para manter a métrica no valor de destino especificado ou próximo a ele.

Acesse a guia do navegador com a aplicação Web APP de **cada instância**.



Clique em **Load Test** (Carregar teste) ao lado do logotipo da AWS.

A máquina irá começar a consumir muito processamento chegando em 100%.

Veja se a instância esta com IP Válido e DNS e faça acesso em cada uma pelo endereço:

<http://ec2-107-22-24-147.compute-1.amazonaws.com/index.php>

Veja que as instancias estão sendo criadas em média de 5 minutos.

Instâncias (6) Informações								
Pesquisar				Atualizar	Conectar	Estado da instância	Ações	Executar instâncias
☐	Name	ID de instância	Estado da instância	Tipo de instância	Verificação de status	Status do alarme	Zona de disponibilidade	
☐	WebAutoScalling	i-0cf08a14a506fe810	Executando	t2.micro	2/2 verificações aprovadas	Sem alarme	+	us-east-1a
☐	WebAutoScalling	i-0332c8381f644c047	Pendente	t2.micro	-	Sem alarme	+	us-east-1a
☐	WebAutoScalling	i-0ccdee6666d4eabe1	Encerrado	t2.micro	-	Sem alarme	+	us-east-1b
☐	WebAutoScalling	i-0909b320ce93c47cc	Encerrado	t2.micro	-	Sem alarme	+	us-east-1b
☐	WebAutoScalling	i-05b8f670b8bbe4269	Executando	t2.micro	2/2 verificações aprovadas	Sem alarme	+	us-east-1b
☐	WebAutoScalling	i-0e012ca8133b13da3	Executando	t2.micro	-	Sem alarme	+	us-east-1b

Você deve ver duas novas instâncias chamadas de **WebAutoScalling**

Elas foram iniciadas pelo **Auto Scaling**.

Se as instâncias ou nomes não forem exibidos, aguarde 30 segundos e clique em atualizar no canto superior direito.

Primeiro, você confirmará que as novas instâncias foram aprovadas na verificação de integridade.

Tarefa 6: Testar o Auto Scaling

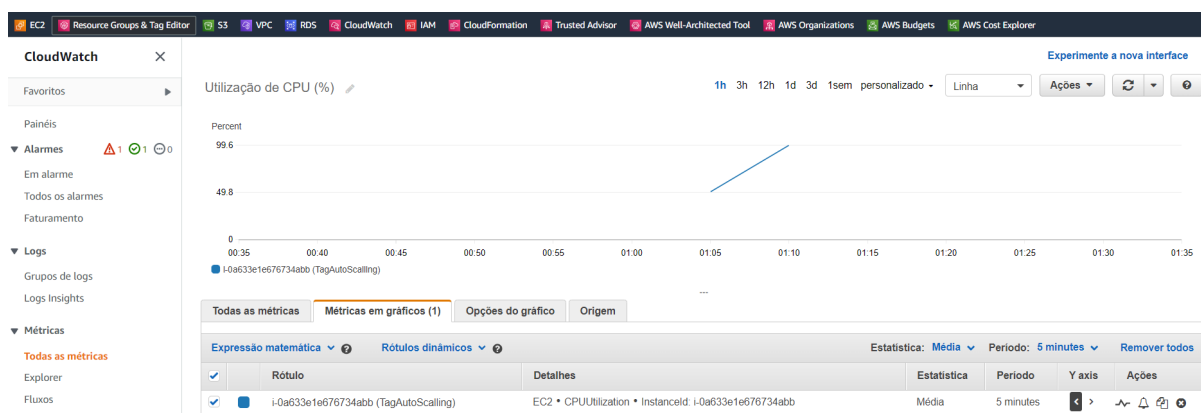
Você criou um grupo de Auto Scaling com um mínimo de duas instâncias e um máximo de seis instâncias. Atualmente, duas instâncias estão em execução porque o tamanho mínimo é duas e o grupo não está atualmente sob nenhuma carga. Agora, você aumentará a carga para fazer com que o Auto Scaling acrescente instâncias adicionais.

52. Volte para o Console de Gerenciamento da AWS, mas não feche a guia da aplicação. Você retornará a ele em breve.

53. No menu **Services** (Serviços), clique em **CloudWatch**.

54. No painel de navegação à esquerda, clique em **Alarms** (Alarmes) (não em **ALARM** (ALARME)).

Verifique no serviço do CloudWatch que aletas foram gerados.



Clique na guia **Alarmes**.

CloudWatch > Alarms

Alarms (2) ☐ Hide Auto Scaling alarms

< 1 >

☐	Name	State	Last state update	Conditions	Actions
☐	TargetTracking-LabGroupAutoS-AlarmLow-5e952431-5661-4a53-84e9-7df72d568d22	⚠ In alarm	2021-10-03 11:11:23	CPUUtilization < 35 for 15 datapoints within 15 minutes	☒ Actions enabled
☐	TargetTracking-LabGroupAutoS-AlarmHigh-6a0f4733-3f72-4102-b4b1-93756ccc1ef1	✅ OK	2021-10-03 11:10:53	CPUUtilization > 50 for 3 datapoints within 3 minutes	☒ Actions enabled

48. Aguarde até que o **Status** de ambas as instâncias mude para **Healthy** (*íntegro*).

49. Clique em atualizar no canto superior direito para verificar se há atualizações.

O status **Healthy** (*íntegro*) indica que a instância passou na verificação de integridade.

Alarmes serão exibidos. Eles foram criados automaticamente pelo grupo de Auto Scaling.

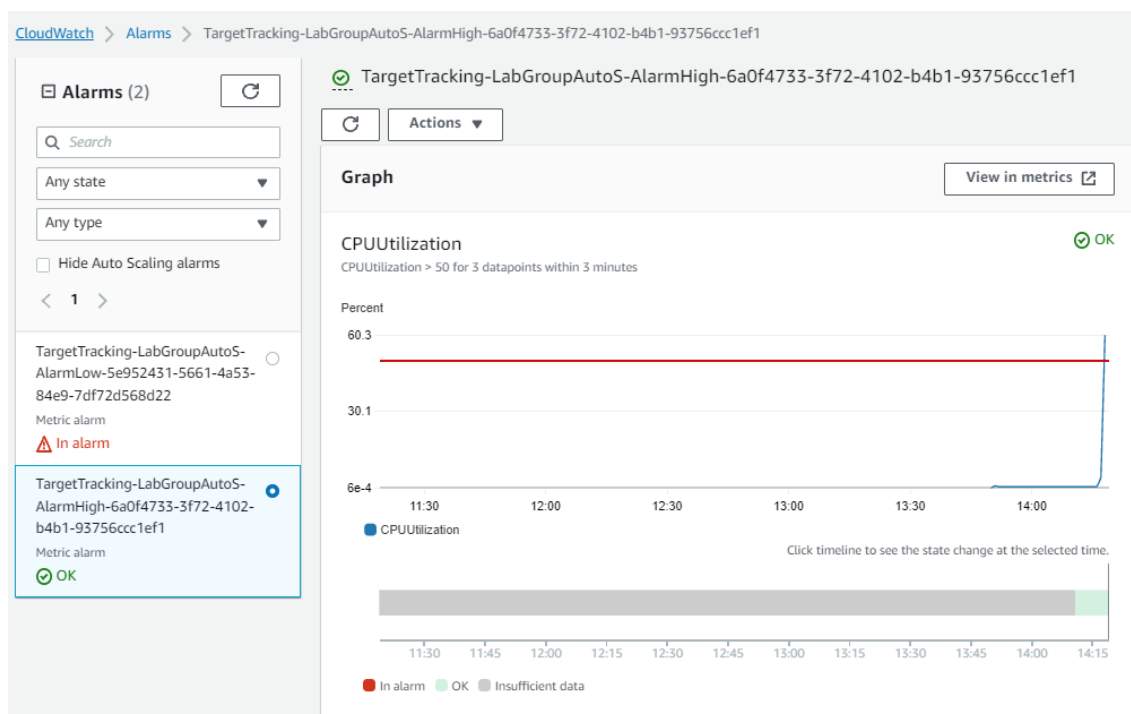
Eles manterão automaticamente a carga média da CPU próxima a 50%, permanecendo também dentro da limitação de ter duas a seis instâncias.

- No menu **Services** (Serviços), clique em **CloudWatch**.
- No painel de navegação à esquerda, clique em **All Alarms** (Todos os Alarmes) (*não em ALARM (ALARME)*) e confirme se vê dois alarmes.

55. Clique no alarme **OK**, que tem *AlarmHigh (Alarme alto)* no nome.

Se nenhum alarme estiver mostrando **OK**, aguarde um minuto e clique em Refresh (Atualizar) no canto superior direito até que o status do alarme mude.

O **OK** indica que o alarme *não* foi acionado.



Em menos de 5 minutos, o alarme **AlarmLow** (Alarme baixo) deverá mudar para **OK** e o status do alarme **AlarmHigh** (Alarme alto) deverá mudar para **ALARM** (ALARME).

Você pode clicar em Refresh (Atualizar) no canto superior direito a cada 60 segundos para atualizar a exibição.

Você deve ver o gráfico **AlarmHigh**(Alarme alto) indicando uma porcentagem crescente de CPU. Depois de passar da linha de 60% por mais de 3 minutos, o Auto Scaling acionará a adição de novas instâncias.

59. Aguarde até que o alarme **AlarmHigh** (Alarme alto) entre no estado **ALARM** (ALARME).

Agora você pode visualizar as instâncias adicionais que foram executadas.

60. No menu **Services** (Serviços), clique em **EC2**.

61. No painel de navegação esquerdo, clique em **Instances** (Instâncias).

Agora deve haver mais de duas instâncias rotuladas **Lab Instance** (Instância do laboratório) em execução. As novas instâncias foram criadas pelo Auto Scaling em resposta ao alarme.

Remover Grupos de Auto Scaling, AMIs e Instâncias

Exercícios do Módulo

1. Você deseja redefinir a configuração de instâncias EC2 que você planeja iniciar manualmente e usando o Auto Scaling. Que recurso você deve usar??

- a) Launch configuration (Configuração de lançamento).
- b) Modelo CloudFormation.
- c) Papel de instância.
- d) Launch template (Modelo de lançamento).
- e) Modelo Lambda.

2. Qual o parâmetro de grupo do Auto Scaling define o limite para o número de instâncias que o Auto Scaling cria?

- a) Tamanho do grupo.
- b) Capacidade desejada.
- c) Capacidade mínima.
- d) Capacidade limitante.
- e) Capacidade máxima.

3. Um grupo de Autoscaling tem uma capacidade desejada de 7 e um tamanho máximo de 7. O que o Auto Scaling fará se alguém encerrar manualmente uma dessas instâncias?

- a) Ele iniciará uma nova instância.
- b) Isso mudará a capacidade desejada para 6.
- c) Ele encerra uma instância.
- d) Não iniciarão nenhuma nova instância.
- e) Isso mudará a capacidade máxima para 6.

4. Como uma empresa pode habilitar elasticidade para um aplicativo em execução no Amazon EC2?

- a) Usando balanceamento de carga elástica.
- b) Usando o Amazon EC2 Auto Scaling.
- c) Configurando multi-AZ.
- d) Ao habilitar o failover no Amazon EC2.
- e) Configurando um bucket no S3.

5. Como o Amazon EC2 Auto Scaling ajuda no custo-benefício?

- a) Escolhendo o tipo de instância com melhor custo-benefício.
- b) Equilibrando a carga entre instâncias uniformemente.
- c) Ao iniciar e encerrar instâncias à medida que a demanda muda.

- d) Por falha de aplicativo automaticamente.
- e) Ao iniciar as instâncias à medida que a demanda diminui e encerrar instâncias à medida que a demanda cresce.

6. Dentre as afirmações a seguir, qual NÃO é verdade?

- a) O Auto Scaling pode ajudar a diminuir os custos.
- b) O Auto Scaling pode distribuir carga igualmente para as instâncias EC2.
- c) O Auto Scaling te ajuda a ter menos problemas com instâncias ociosas e sem necessidade na sua topologia.
- d) O Auto Scaling pode tanto criar novas instâncias quanto encerrar de acordo com a sua demanda.
- e) O Auto Scaling tem número de instâncias mínimo, desejado e máximo totalmente customizáveis.

Gabarito

1. Letra D. É no Launch Template (Modelo de Lançamento) que você define essas informações, como a AMI usada, o tipo da instância, o par de chaves e mais.
2. Letra E. Ao definir a capacidade máxima, você está definindo o limite de instâncias que podem ser executadas.
3. Letra A. A capacidade desejada não mudaria nesse caso, sendo assim, o Auto Scaling iniciaria uma nova instância para satisfazer as 6 desejadas.
4. Letra B. O serviço responsável por aumentar a elasticidade dentre os citados é o Auto Scaling.
5. Letra C. Ele ajuda no custo-benefício por conseguir iniciar as instâncias à medida que a demanda cresce e encerrar à medida que a demanda diminui, por isso não se encaixa na letra e.
6. Letra B. Essa alternativa está incorreta, visto que o papel de balancear carga é do Load Balancer, e não do Auto Scaling.