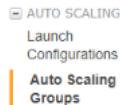


07

Mão à obra: Criando políticas de escalonamento

Devemos definir agora quais serão as políticas utilizadas para poder fazer esse escalonamento. No painel de console da Amazon, pesquisar pelo serviço EC2, feito isso, no menu lateral esquerdo pesquise pela aba Auto Scaling e clique em Auto Scaling Groups.



Na sequência, clique no grupo de auto-escalonamento criado e vá até a aba Details. Vamos primeiro indicar que os demais servidores deverão utilizar a configuração customizada que criamos. Clique no botão Edit e coloque na opção Launch Configuration a configuração customizada, imagem-customizada-casadocodigo. No projeto, vamos configurar para que o mínimo de servidores ativos seja 1 e nesses casos, que teremos uma grande demanda de acessos que seja configurado um segundo servidor, Min: 1 e Max:2 servidores.

A screenshot of the 'Auto Scaling Group: Escalonando-servidor-casadocodigo' details page. The 'Edit' button is highlighted in red. The 'Launch Configuration' dropdown shows 'Escalonando-servidor-casadocodigo'. The 'Desired' field is set to 1. The 'Availability Zone(s)' field shows 'us-east-1b'. Other tabs like Activity History, Scaling Policies, Instances, Monitoring, Notifications, Tags, Scheduled Actions, and Lifecycle Hooks are visible at the top.

Para finalizar, especifique as localidades (zonas de disponibilidade) nas quais iremos configurar esses servidores. O primeiro servidor configuramos na zona de disponibilidade us-east-1b, a fim de garantir sempre uma disponibilidade de nossos serviços, vamos configurar uma segunda zona de disponibilidade para o segundo servidor do grupo de auto-escalonamento, por exemplo a zona de disponibilidade us-east-1c. No final, salve essas alterações da aba Detalhes, clicando no botão Save

A screenshot of the 'Edit Launch Configuration' dialog. It shows the 'Launch Configuration' dropdown set to 'imagem-customizada-casadocodigo'. The 'Desired' field is set to 1, and the 'Max' field is set to 2. The 'Health Check Type' is set to 'EC2'. On the right, under 'Availability Zone(s)', there are two selected subnets: 'subnet-6cd4b808(172.31.16.0/20)' and 'subnet-9928f6b6(172.31.32.0/20)', both marked as 'Default in us-east-1b'.

Agora devemos configurar as políticas de escalonamento para que a Amazon "suba" esse segundo servidor. Vá até a aba Scaling Policies e crie uma política com o nome Aumentando servidores e clique na opção Create a scaling policy with steps.

Name: Aumentando-servidores
Metric type: Average CPU Utilization
Target value:
Instances need: 300 seconds to warm up after scaling
Disable scale-in:

Create a simple scaling policy ⓘ
Create a scaling policy with steps ⓘ

Feito isso, clique na opção **Create new alarm** para que configuremos assim a política de escalonamento. A política que iremos criar será para analisar o consumo de CPU do servidor que temos no momento, se esse valor for ultrapassado, a Amazon irá subir o segundo servidor com as configurações que fizemos. Para evitar um longo processo de espera, coloque um consumo baixo, por exemplo 0.5% por um período de 1 minuto:

Send a notification to: NotifyMe (rafanercessario@gmail.com)
Whenever: Average of CPU Utilization
Is: >= 0.5 Percent
For at least: 1 consecutive period(s) of 1 Minute
Name of alarm: awsec2-Escalando-servidor-casadocodigo-CPU-U

CPU Utilization Percent

Escalonando-servidor-casadocodigo

Cancel Create Alarm

Para finalizar, devemos especificar que ação tomar quando esse limite for ultrapassado, vamos querer justamente fazer o escalonamento para "subir" uma instância, chegando ao valor máximo que estipulamos de 2 instâncias e clique no botão **Create** para finalizar a criação da política de escalonamento.

Name: Aumentando-servidores
Execute policy when: awsec2-Escalando-servidor-casadocodigo-CPU-Utilization
Create new alarm
breaches the alarm threshold: CPUUtilization >= 0.5 for 60 seconds
for the metric dimensions AutoScalingGroupName = Escalonando-servidor-casadocodigo

Take the action: Add 1 instances when 0.5 <= CPUUtilization < +infinity
Add step ⓘ

Agora só falta nós fazermos o teste, para isso vamos utilizar o Apache Bench para simular várias requisições de usuários, se estiver no Windows abra o Prompt e mude para o diretório onde teremos a pasta bin do XAMPP que por padrão estará em **C:\xampp\apache\bin** e faça o teste simulando por exemplo 500 requisições que já deverá ser suficiente para ultrapassar o limiar estipulado de 0.5% e deveremos ter assim a criação da segunda instância (Esse processo poderá demorar alguns minutos)

```
ab.exe -n 500 [IP Público instância EC2 us-east1-b]:8080/casadocodigo
```

Caso esteja no Linux ou no Mac abra o terminal e coloque (Esse processo poderá demorar alguns minutos):

```
ab -n 500 [IP Público instância EC2 us-east1-b]:8080/casadocodigo
```

Feito isso, volte para o painel de console com as instâncias do EC2, faça uma atualização e veja se o segundo servidor foi criado. Tente acessar a aplicação da Casa do Código nesse segundo servidor. Qual é o resultado? Você consegue acessar a aplicação da Casa do Código nesse servidor que fazia parte do grupo de auto-escalonamento?