

Criando grupo autoescalamento

Transcrição

[00:00] Temos agora a nossa aplicação funcionando inteiramente na Amazon. Mudamos acesso do banco de dados, não colocando mais uma interface pública, garantindo assim, uma maior segurança dessa nossa aplicação. Então vamos ver o cenário que temos até o momento?

[00:16] Temos aqui esse nosso servidor, na Amazon, com aplicação da Casa do Código, e eu aqui em São Paulo, quando coloco o endereço "IP" público nesse servidor, e a porta do Tomcat "8080/casadocodigo", identificamos a nossa aplicação funcionando na Amazon. O que acontece se tiver vários usuários, e todos eles acessarem esse único servidor que temos com aplicação da Casa do Código?

[00:43] A tendência é que esse servidor, ele acaba tendo uma sobrecarga e com isso ele pode vir a ficar indisponível. E a nossa aplicação da casa do código, não estaria mais disponível para ser acessada pelos usuários. E não queremos isso, queremos que sempre a nossa aplicação da Casa do Código esteja disponível para os usuários acessarem.

[01:03] Então o que poderíamos pensar inicialmente? Poderíamos pensar que, para evitar esse cenário, poderíamos subir inicialmente, 2, 3 servidores, com a aplicação da Casa do Código, com objetivo de distribuir esses vários acessos entre esses dois, três servidores que já poderia configurar inicialmente.

[01:22] Mas tudo isso, sempre que subimos um servidor na Amazon, isso gera um custo para nós, temos aquele consumo gratuito de 750 horas por mês, e quanto mais máquina vamos subindo, mais rápido esses recursos vão consumindo.

[01:36] Então o ideal é o que? Que configure essas outras máquinas, esses outros servidores, conforme vamos tendo essa demanda de acesso, conforme ela for aumentando. Então qual vai ser a nossa ideia aqui? A nossa ideia vai colocar esse servidor, que está com a aplicação da Casa do Código, e vamos configurá-lo para um grupo de "auto escalamento", ou em inglês, chamada de "auto-scaling".

[02:01] Então qual é a ideia com esse grupo de "auto-escalamento"? É configurar algumas métricas, algumas políticas, que conforme essas políticas, esses valores forem ultrapassados, do consumo dos recursos do nosso, o que vamos fazer?

[02:16] Se passar desse limiar, vamos pedir para Amazon criar um outro servidor com a aplicação da Casa do Código. Nesse grupo de "auto-escalamento", vamos configurar essas métricas, para que caso esse nosso servidor comece ficar sobrecarregado, saberemos que está tendo uma grande demanda de acessos. Vamos pedir para a Amazon criar um outro servidor com a aplicação da Casa do Código para poder fazer melhor essa distribuição de acessos dos usuários.

[02:44] Então vamos fazer essa etapa agora, de realizar essa configuração desse grupo de "auto-escalamento". Vamos voltar para o "painel de console" da Amazon, e vamos voltar para o nosso serviço do "EC2". E o que queremos? Vamos aqui nessa única instância rodando na Amazon, nosso único servidor com a aplicação da Casa do Código, e queremos fazer com que essa Instância seja vinculada a esse grupo de "auto-escalamento".

[03:12] Vamos na aba "action", e vamos nessa opção "instance settings", que seria as configurações desse servidor, dessa instância. Então vamos aqui "instances settings", e queremos atachar essa máquina, essa instância no grupo de "auto-escalamento".

[03:26] Então vamos em "attach to auto scaling group", clicamos, e não temos ainda, na Amazon, nenhum grupo de "auto-escalamento" criado. O que temos que fazer? Temos que clicar nessa opção para criar um novo grupo de "auto-

escalonamento".

[03:40] Vamos colocar um nome para esse grupo? Vou colocar "Escalonando-servidor-casadocodigo". E vamos atachar essa instância, esse servidor, nesse grupo do "auto-escalonamento", para poder definir as métricas depois.

[06:58] Então temos a informação que esse nosso grupo de "auto-escalonamento", ele já foi criado. Vamos aqui e podemos fechar. E se for nesse menu lateral esquerdo, temos a aba do "auto-scaling", do "auto-escalonamento", e nessa segunda opção, temos os grupos de "auto-escalonamento" existentes.

[04:19] Então esperamos que tenha agora esse grupo que acabamos de criar "Escalonando-servidor-casadocodigo". Vamos só confirmar se ele foi criado? Vou clicar aqui, e temos o grupo de auto-escalonamento que criamos, chamado de "Escalonando-servidor-casadocodigo".

[04:41] Esses outros servidores que vão fazer parte desse grupo de "auto-escalonamento", que vamos vai subir conforme a demanda de acessos aumentar. Eles vão ter que ter uma imagem para ser inicializada pelos servidores. Por isso, esse grupo de auto-escalonamento tem que ter a ciência de qual imagem que será configurada para esses demais servidores que irão fazer parte desse grupo de escalonamento.

[05:09] Então se a gente voltar na aba do menu lateral esquerdo, dentro dessa opção "auto scaling", temos o que? Temos as configurações que vão ser utilizadas para subir esses outros servidores. Então se for aqui em "launch configurations", e vamos ter que os demais servidores que for subir, que farão parte desse grupo de "auto-escalonamento" vão utilizar essa imagem para subir esses outros servidores.

[05:39] Que imagem será? Vamos pesquisar? Vamos copiar o número de identificação dessa imagem, e aqui também no menu lateral esquerdo, dentro da opção da aba "imagens" clica em "AMIs", e vamos fazer um filtro para descobrir que imagem é essa. Então vamos fechar o estava marcado antes, e vamos colar esse número de identificação que temos e vou clicar em "enter" para pesquisar.

[06:08] Temos esse número de identificação que foi vinculado no nosso grupo de "auto-escalonamento", é a imagem base do Ubuntu. Estamos usando a imagem base do Ubuntu, mas realizamos algumas customizações. Instalamos o Tomcat, levamos a nossa aplicação da Casa do Código, e não queremos ter que ficar fazendo isso manualmente para todos os outros servidores que vão fazer parte desse grupo de "auto-escalonamento".

[06:36] O ideal é que esses outros servidores que vão subir nesse grupo já tivessem toda a configuração pronta que vamos precisar. Então, o que vamos ter que fazer? Vamos ter que voltar para essa aba "instance", que tem no nosso servidor, e eu vou pedir para a Amazon criar essa imagem customizada desse servidor que montamos, com um Tomcat também com a nossa aplicação da Casa do Código.

[07:02] Para isso o que vamos fazer? Vamos vir nesse único servidor, rodando na Amazon, clicamos na aba "action", e quero criar uma imagem customizada que temos desse servidor. Vamos em "image", "criate image", eu quero criar essa imagem customizada que já vai o Tomcat e a aplicação da Casa do Código. Qual é o nome da imagem? Vamos colocar como sendo "imagem-casadocodigo", também descrição, vou colocar o mesmo nome "imagem-casadocodigo", ele vai colocar que essa imagem será salva nesse volume, podemos manter assim mesmo, e colocamos aqui para criar essa imagem.

[07:43] Então agora com isso temos a nossa imagem customizada com o Ubuntu, o Tomcat, mais a Casa do Código. Agora temos que voltar para o nosso grupo do "auto-escalonamento", e falar que ele deve usar essa imagem customizada que acabamos de criar, que tem o Ubuntu e o Tomcat e a Casa do Código. Vamos fazer isso na sequência.

