



Ambiente de computação em nuvem – AWS

As principais soluções de computação, armazenamento e rede na Amazon Web Services.

Prof. Gustavo Ribeiro

Propósito

Desenvolver conhecimento e capacidade de provisionamento e gestão de recursos computacionais, de armazenamento e de rede na nuvem da Amazon Web Services, seguindo boas práticas.

Objetivos

- Reconhecer as principais soluções de computação na Amazon Web Services e as máquinas virtuais EC2.
- Identificar as principais opções de armazenamento da Amazon Web Services: EBS, S3 e EFS.
- Descrever o uso de VPC e recursos de rede como sub-redes, gateways da internet, rotas e IPs elásticos.
- Analisar aplicações web estáticas e dinâmicas na AWS.

Introdução

O modelo de computação em nuvem surge como resposta para um grande problema das companhias, a gestão complexa e custosa de infraestruturas de datacenters e seus equipamentos de computação, armazenamento e rede. Ao invés de gerir tudo isso por conta própria, as empresas consomem sob demanda recursos de TI e pagam somente por seu uso, deixando para empresas como a Amazon Web Services (AWS) a responsabilidade de custear toda a estrutura física e gerir boa parte da complexidade operacional desses ativos.

Já em 2006 a AWS foi pioneira em computação em nuvem, sendo hoje a plataforma de computação em nuvem mais abrangente e a mais utilizada no mundo. Essa escala massiva permite à empresa beneficiar vários clientes no uso de seu agregado da nuvem que, por sua vez, podem usufruir de economia em escala, com preços mais competitivos e pagando conforme o uso.

É importante compreender que a infraestrutura global da AWS é dividida em regiões, que representam localizações geográficas onde ficam hospedados seus datacenters. Dentro de cada região existem subdivisões conhecidas como zonas de disponibilidade (AZs), que consistem em um ou mais datacenters com energia, rede e conectividade redundante. A escolha de qual região será utilizada é do usuário que provisiona os recursos de TI, que deve avaliar critérios como latência, preço, disponibilidade e possíveis regulações de conformidade nesse ambiente. Para garantir a resiliência de ambientes na nuvem AWS, é sempre recomendado o uso de pelo menos duas AZ. Dessa forma, havendo falha em uma AZ, o ambiente continuará em pleno funcionamento em outra AZ da região.

Essas e outras decisões fazem parte do modelo de responsabilidade compartilhada, onde a AWS e você, como operador, dividem o esforço e a gestão sobre manter ambientes seguros e disponíveis para o usuário final. Neste conteúdo, apresentaremos essas responsabilidades e como é possível seguir um caminho de boas práticas no uso de computação em nuvem com a Amazon Web Services.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Opções de computação na AWS

Confira métodos para conhecer e identificar o melhor uso das principais soluções de computação na Amazon Web Services, aprender quais são as principais características de uma máquina virtual no Amazon EC2, e como provisionar, usando de boas práticas, um servidor na AWS.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Para profissionais que irão gerir recursos computacionais na Amazon Web Services, a primeira decisão a ser tomada é sobre qual tipo de computação será provisionada. Essa é uma decisão muito importante, pois afeta toda a arquitetura da aplicação/serviço a ser hospedado e, por isso, você precisa conhecer qual é o produto ideal para cada caso. Basicamente, existem 3 opções de computação na AWS:

Máquinas virtuais (VMs)

Geralmente, são a opção de computação mais fácil de se entender na AWS para quem tem conhecimento prévio de infraestrutura de TI, pois trata-se da virtualização de um servidor físico, que possui disco, placa de rede, e permite instalar e personalizar o ambiente de forma similar. A Amazon oferece opções de máquinas virtuais com sistema operacional e até algumas opções de softwares pré-instalados. Na AWS, as máquinas virtuais são chamadas de Amazon Elastic Compute Cloud (EC2).

Containers

Com a escalada de aplicações na nuvem, soluções que oferecem maior velocidade de provisionamento e consistência de funcionamento independente do ambiente (no on-premise ou na cloud, em desenvolvimento ou em produção), tornaram-se cada vez mais populares e esses são alguns benefícios que estimulam o uso de containers na computação em nuvem. Um container é um padrão de empacotamento de código e dependências projetado para ser executado de forma confiável em qualquer plataforma. Na AWS é possível executar containers no Amazon Elastic Container Service (Amazon ECS) ou no Amazon Elastic Kubernetes Service (Amazon EKS).

Computação sem servidor (serverless computing)

Uma das maiores vantagens da computação em nuvem é a abstração de hardware da camada de infraestrutura. Na computação sem servidor, a abstração sobe mais um nível, no qual não somente a camada física é abstraída, mas também a de sistema operacional e stack. Com esse nível de abstração, o foco passa a ser no código das suas aplicações, sem precisar gastar tempo mantendo e atualizando infraestrutura, servidores ou sistema operacional. Nesse modelo, você pagará apenas pelo tempo que sua aplicação executar. Na AWS, o principal serviço de computação sem servidor é o AWS Lambda.

Em linhas gerais, para aplicações que necessitam de armazenamento local e que possuem forte dependência do sistema operacional, têm características de monolito e não escalam horizontalmente, o mais recomendado é utilizar EC2. Para equipes que dominam docker ou kubernetes, que já utilizam uma arquitetura de microsserviços e armazenamento de rede ou de objetos, é recomendado o ECS/EKS.

O AWS Lambda pode ser uma ótima opção se o time técnico tem um perfil desenvolvedor e não quer gerir detalhes de infraestrutura ou de rede, as aplicações processam tarefas rapidamente (em menos de 15 minutos), não existindo a necessidade de armazenamento local no servidor e o que permite escalar horizontalmente.

Serviço	EC2	ECS	Lambda
Tipo de Computação	Instância; Infraestrutura como serviço (IaaS)	Container; Container como serviço (CaaS)	Função; Função como serviço (FaaS)
Caso de uso	De uso geral; controle completo sobre o servidor	Executar containers docker; Tarefas/ execuções de +15 minutos	Pequenas aplicações que executam tarefas em menos de 15 minutos
Disponibilidade	SLA: 99.99%	SLA: 99.99%	SLA: 99.95%
Escalabilidade	Uso de políticas de auto scaling groups para aumento e diminuição de instâncias	Escalabilidade nativa baseado em métricas do cluster	Escalabilidade automática
Tempo limite de execução	Sem limite	Sem limite	300 segundos (15 minutos)
Preço	Varia pelo tipo, tempo de execução e opção de compra.	ECS no EC2: mesmos custos do EC2; ECS no Fargate: quantidade vCPU e memória usada, tempo de execução e opção de compra.	Números de requisições e tempo de execução.

Opções de computação na AWS e suas características.
Gustavo Ribeiro

Máquinas virtuais – Amazon EC2

O **Amazon Elastic Compute Cloud** é um serviço que provê capacidade computacional segura e redimensionável na nuvem, em formato de máquinas virtuais, conhecidas como instâncias do Amazon EC2. Em poucos minutos você pode ligar um novo servidor, usufruir imediatamente da capacidade computacional dele e desligar quando quiser, encerrando o custo dessa infraestrutura. Comparado ao modelo on-premise, no qual você é o responsável pela infraestrutura de TI de uma empresa, caso você queira provisionar uma nova aplicação na internet como, por exemplo, um sistema de sites/blog Wordpress, você precisará seguir alguns passos:

1. Encontrar um fornecedor de equipamentos de TI e pagar adiantado a aquisição do equipamento.
2. Esperar o processo logístico de entrega dos equipamentos para a sua empresa.
3. Montar e ligar todos os equipamentos dentro da sua infraestrutura de datacenter.
4. Fazer instalação de sistema operacional com todas as configurações mínimas para iniciar a implantação da sua aplicação.
5. Baixar e configurar o Wordpress e todas suas dependências.

Se o mesmo desafio fosse proposto para ser realizado utilizando uma instância EC2, você seguiria os seguintes passos:

1. Iniciar o processo de lançamento de uma nova instância, optando por uma configuração básica na qual você pode escolher o sistema operacional, plataforma e até a possibilidade de já escolher um aplicativo (como o Wordpress).
2. Selecionar o tipo de hardware virtual que deseja, escolhendo o conjunto de vCPU e memória.
3. Selecionar detalhes como segurança, armazenamento local e rede.
4. Fazer uma revisão e iniciar a instância, podendo em seguida conectá-la remotamente.
5. A partir desse momento você já pode ter um Wordpress pré-instalado com suas dependências, ou fazer uma instalação da mesma forma que um servidor físico.

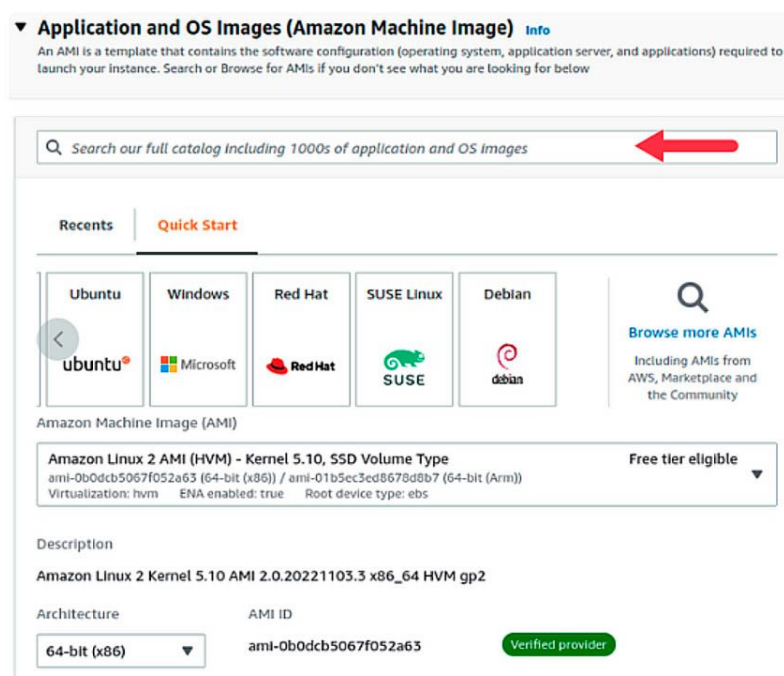
Um processo de semanas foi resolvido em poucos minutos com o uso da AWS, de forma self-service e com a vantagem de, ao final, caso você não precise mais do recurso, ter a possibilidade de desligar e encerrar os custos dessa operação. Essa mudança de paradigma na TI foi o que tornou o Amazon EC2 um serviço tão amplamente utilizado e estimulou várias empresas a migrarem do on-premise para a nuvem. Veremos agora, em detalhes, cada passo para fazer esse processo, seguindo boas práticas.

Amazon Machine Image (imagens de aplicações e sistema operacional)

No mundo tradicional de infraestrutura de TI, após a instalação física do servidor, o primeiro passo seria a instalação do sistema operacional a ser utilizado, utilizando discos, drivers ou pela rede. No Amazon EC2, a responsabilidade da instalação do sistema operacional não é do usuário, pois a AWS fornece imagens prontas, conhecidas como Amazon Machine Images (AMI).

Nesse caso, sua única responsabilidade é apenas de escolher qual sistema operacional é mais adequado ao seu cenário.

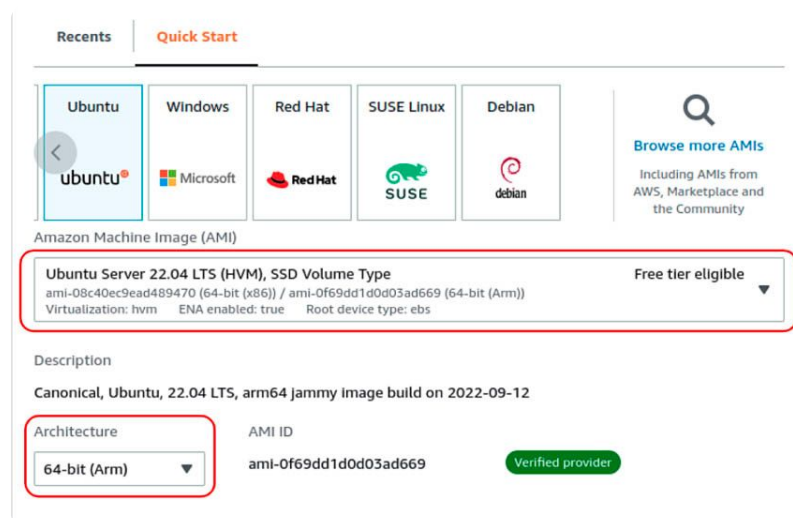
Além do sistema operacional, algumas imagens já podem vir com plataforma, aplicativos pré-instalados e suas dependências. A AWS fornece um catálogo de milhares de imagens prontas para uso fornecidas pela própria Amazon e por empresas parceiras, por meio de um marketplace ou da comunidade, que oferece imagens customizadas disponibilizadas por outros usuários da nuvem.



Campo de busca de imagens da AWS, marketplace e comunidade, no painel de lançamento de uma nova instância.

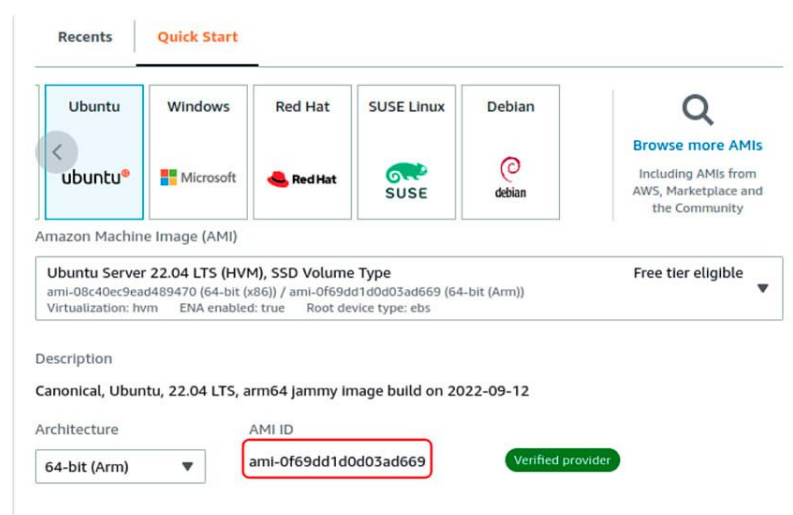
Como uma boa prática para iniciantes, devem ser utilizadas apenas AMI recomendadas pela AWS e com garantia de integridade, que estão localizadas na seção "Quick Start". Existe uma diversidade de opções de

sistemas operacionais e suas variações, que podem ser selecionadas a depender da versão e, em alguns casos, da arquitetura disponível.



AMI do sistema Ubuntu tem disponível a versão 22.04 LTS e a arquitetura ARM.

Note que as AMIs possuem um identificador único (AMI ID), com prefixo “ami-” seguido de um hash com números e letras, que representa todo o conjunto de características, como o sistema, a versão, a arquitetura e a região. Isso significa que cada região AWS terá suas próprias AMIs para as características selecionadas, não podendo uma mesma AMI ser utilizada em regiões diferentes, pois são únicas.



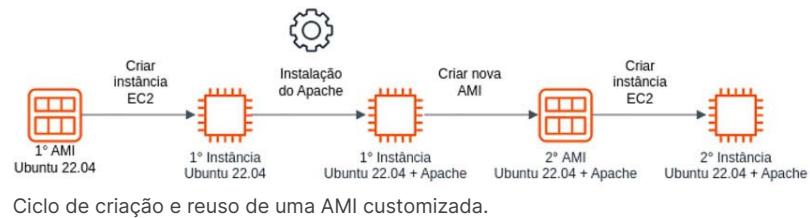
AMI ID para Ubuntu 22.04 LTS, arquitetura ARM, na região N. Virginia.

AMIs podem ser criadas pelo usuário a partir de uma instância em execução e isso permite sua customização e reutilização para novos provisionamentos de sistemas, com configurações próprias para o seu cenário. Isso é de grande utilidade para cenários de autoescalabilidade, nos quais a instância precisa ser ligada automaticamente e já deve estar pronta para funcionamento.

Podemos imaginar a AMI como um código DNA, que possui suas próprias características, mas pode ser modificado e melhorado para seu uso específico.

Como exemplo, podemos usar a AMI do Ubuntu 22.04 LTS, recomendada pela AWS, mas que só possui instalado o sistema operacional base. Para seu cotidiano operacional, você deseja lançar várias máquinas virtuais com o servidor web Apache para hospedagem de sites. Ao invés de sempre instalar o Apache para cada instância recém-criada, você criará uma AMI a partir de uma instância com o Apache instalado, gerando uma nova AMI, com um novo “DNA”. Essa será sua AMI de partida, com Ubuntu 22.04 LTS e o Apache pré-instalado.

No futuro, se você quiser adicionar novas características como, por exemplo, permitir o uso de sites na plataforma PHP, pode repetir o processo modificando o “DNA” da sua AMI de referência.



Na criação de uma nova AMI, a partir de uma instância, será importante você definir um nome para a imagem, para facilitar a identificação futura. Você também terá a opção de criar uma AMI com a instância em execução, sem precisar desligar, marcando a opção “no reboot”.



Atenção

Eventuais dados que sejam modificados durante a criação da imagem podem ficar inconsistentes. Para evitar isso, a AMI deve ser feita com a instância desligada.

A seguir, podemos ver a criação de uma AMI.

Create image [Info](#)

An image (also referred to as an AMI) defines the programs and settings that are applied when you launch an EC2 instance. You can create an image from the configuration of an existing instance.

Instance ID: [i-0ab...](#)

Image name:

Image description - optional:

No reboot: ☐ Enable

Volume type	Device	Snapshot	Size	Volume type	IOPS	Throughput	Delete on termination	Encrypted
EBS	/dev/...	Create new snapshot fr...	250	EBS General Purpose S...	3000		<input type="checkbox"/> Enable	<input checked="" type="checkbox"/> Enable
EBS	/dev/...	Create new snapshot fr...	10	EBS General Purpose S...	3000	125	<input type="checkbox"/> Enable	<input type="checkbox"/> Enable

[Add volume](#)

During the image creation process, Amazon EC2 creates a snapshot of each of the above volumes.

Exemplo de tela de criação de AMI de uma instância em execução.

Para cada AMI criada em sua conta vai existir ao menos um snapshot associado, podendo existir mais de um se a instância que foi usada de base para criação possuir mais de um disco. Esses snapshots não podem ser apagados até que a AMI vinculada seja desregistrada antes. As AMIs em si não possuem custo, mas sim os snapshots vinculados que você armazenará em sua conta, que serão cobrados por seu tamanho.

Tipos de instâncias EC2

Confira como identificar qual tipo é mais indicado para cada uso.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Instâncias EC2 são uma combinação de processadores virtuais (vCPU), memória, rede e, em alguns casos, armazenamento e processadores gráficos (GPUs). Quando você cria uma instância EC2, precisa escolher de quantos desses componentes você vai precisar. A Amazon define tipos de instâncias EC2 que são as combinações desses componentes otimizadas para diferentes casos de uso. Os tipos de instância seguem uma nomenclatura que consiste em um prefixo, que identifica a família e sua geração, seguido pelo tamanho. Em alguns casos, características adicionais como o processador utilizado, especificidades de armazenamento, rede ou virtualização podem estar indicados no prefixo. Vamos tomar como exemplo dois tipos: m5.large e t4g.small.

m5

Quinta geração da família M.

t4g

Quarta geração da família T, que passou a adotar a letra “g” para identificar que esse tipo utiliza processadores Graviton 2, baseados em ARM.

Large e small

Tamanhos que definem qual a capacidade daquela instância.

Vamos conhecer as principais famílias e qual é o caso de uso indicado para cada uma delas:

De uso geral

É aquele que provê um equilíbrio entre memória, processamento e rede, e pode ser utilizado para uma ampla gama de workloads. Ex.: t4g, t3, t3a, m6i, m6g, m6a, m5, m5a, m5n, m4.

Otimizado para computação

É ideal para workloads de uso intensivo de CPU, beneficiando-se de processadores de alta performance. Ex.: c6g, c6i, c6a, c5, c5a, c4.

Otimizado para memória

É ideal para workloads que precisam processar grandes conjuntos de dados em memória. Ex.: r6a, r6g, r6i, r5, r5a, r4.

Com cada vez mais tipos utilizando processadores Graviton (baseados na arquitetura ARM), a AWS passou a identificar as instâncias que utilizam processadores Intel e AMD com as letras “i” e “a”, respectivamente, após o número que identifica a geração (c6i, m6i, r6i, c5a, c6a e m5a, por exemplo). Porém, alguns tipos utilizam processadores Intel e não possuem a letra “i”, como também podem existir tipos com outras letras que indicam uma característica que não está diretamente relacionada ao modelo do processador e/ou processadores que não foram citados entre os tipos mais populares.

Par de chaves

Durante o processo de lançamento de uma instância EC2, é solicitado que você selecione um key pair (par de chaves) ou que crie um, caso ainda não tenha feito. Esse par de chaves consiste em uma chave pública e uma

chave privada. O Amazon EC2 armazena a chave pública em sua instância e você deve armazenar de forma segura a chave privada, pois qualquer um de posse dela pode conectar-se na sua instância. Para instâncias do Windows, a chave privada é necessária para descriptografar a senha do administrador e em instâncias Linux, para conectar remotamente, usando o SSH.

Create key pair

X

Key pairs allow you to connect to your instance securely.

Enter the name of the key pair below. When prompted, store the private key in a secure and accessible location on your computer. **You will need it later to connect to your instance.** [Learn more](#)

Key pair name

Enter key pair name

The name can include upto 255 ASCII characters. It can't include leading or trailing spaces.

Key pair type

☒ RSA
RSA encrypted private and public key pair

☐ ED25519
ED25519 encrypted private and public key pair (Not supported for Windows instances)

Private key file format

☒ .pem
For use with OpenSSH

☐ .ppk
For use with PuTTY

Cancel

Create key pair

Tela de criação de uma key pair onde é solicitado um nome, tipo e formato.

Rede e firewall de EC2

Durante o processo de criação de uma instância no Amazon EC2 são solicitadas, ou até mesmo pré-selecionadas, algumas informações relacionadas à rede (VPC) e ao firewall (security group) da máquina que você está prestes a criar.

O Amazon VPC permite que você execute recursos da AWS em uma rede virtual dedicada à sua conta, conhecida como **nuvem privada virtual (VPC)**. Ao iniciar uma instância, você pode selecionar uma sub-rede da VPC (ou deixar que a AWS escolha por você). A instância é configurada com uma interface de rede primária, que é uma placa de rede virtual lógica. A instância recebe um endereço IP privado primário do endereço IPv4 da sub-rede, que é atribuído à interface de rede primária.

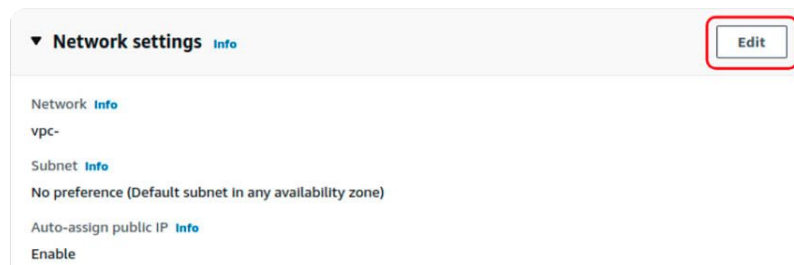
Você pode controlar também se a instância receberá um endereço IP público do pool de endereços IP públicos da Amazon. O endereço IP público de uma instância é associado à sua instância somente até que ela seja desligada ou encerrada. Se você precisar de um endereço IP público persistente, poderá alocar um endereço IP elástico para sua conta da AWS e associá-lo a uma instância ou interface de rede.



Saiba mais

Um endereço IP elástico permanece associado à sua conta da AWS até que você o libere, e você pode movê-lo de uma instância para outra conforme necessário.

O botão “Edit” habilitará a escolha de VPC, subnets e se deseja associar IP público automaticamente, como podemos ver a seguir.



▼ Network settings [Info](#)

Network [Info](#)

vpc-

Subnet [Info](#)

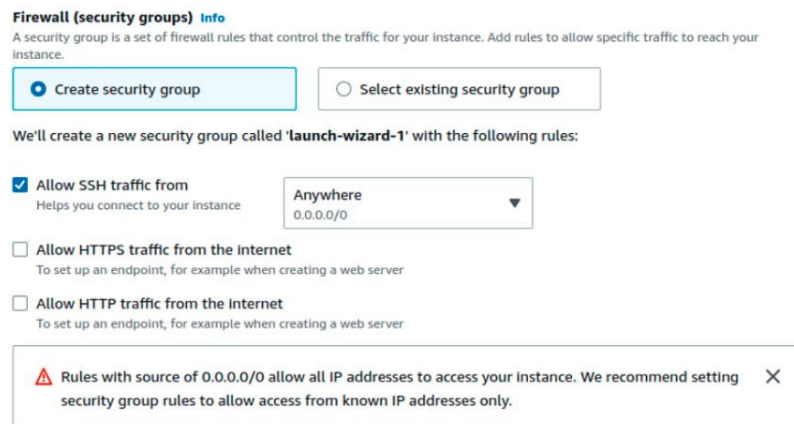
No preference (Default subnet in any availability zone)

Auto-assign public IP [Info](#)

Enable

Indicação do botão Edit para escolha do VPC.

Já o **security group** (SG) atua como um firewall virtual para suas máquinas EC2, controlando o tráfego de entrada e saída. Você pode especificar um ou mais security groups durante o processo de criação de uma instância e, caso você não especifique um grupo de segurança, o Amazon EC2 usará o grupo de segurança padrão. Você pode adicionar regras a cada grupo de segurança que permite o tráfego de saída ou entrada para suas instâncias associadas e pode modificar essas regras a qualquer momento. As regras novas ou modificadas são aplicadas automaticamente a todas as instâncias associadas ao grupo de segurança.



Firewall (security groups) [Info](#)

A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.

☒ Create security group ☐ Select existing security group

We'll create a new security group called 'launch-wizard-1' with the following rules:

☒ Allow SSH traffic from
Helps you connect to your instance 0.0.0.0/0

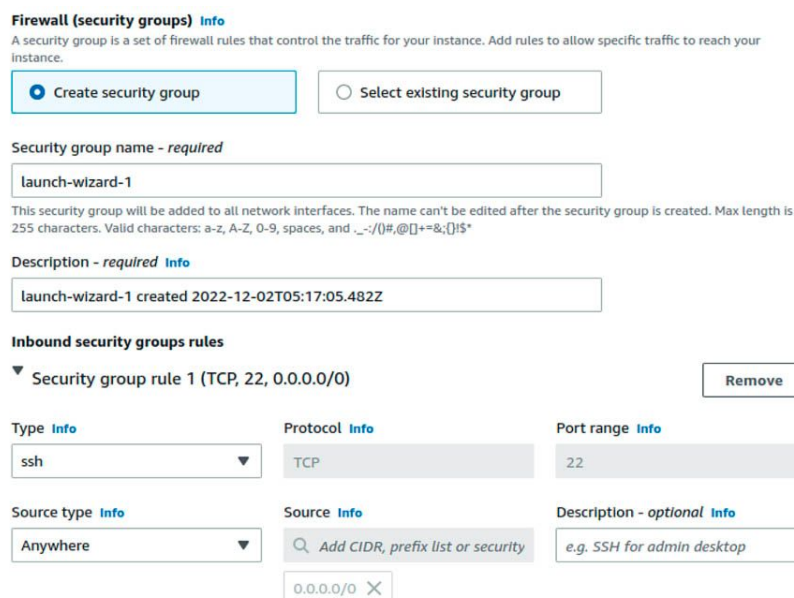
☐ Allow HTTPS traffic from the Internet
To set up an endpoint, for example when creating a web server

☐ Allow HTTP traffic from the Internet
To set up an endpoint, for example when creating a web server

⚠ Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only. ✕

Modo simplificado de criação de security group durante o passo a passo de lançamento de EC2.

Ao clicar no “Edit”, o painel oferece mais opções de personalização de um novo security group.



Firewall (security groups) [Info](#)

A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.

☒ Create security group ☐ Select existing security group

Security group name - *required*

launch-wizard-1

This security group will be added to all network interfaces. The name can't be edited after the security group is created. Max length is 255 characters. Valid characters: a-z, A-Z, 0-9, spaces, and _-./()#,@[]+=&:~!\$*

Description - *required* [Info](#)

launch-wizard-1 created 2022-12-02T05:17:05.482Z

Inbound security groups rules

▼ Security group rule 1 (TCP, 22, 0.0.0.0/0) Remove

Type Info	Protocol Info	Port range Info
ssh	TCP	22
Source type Info	Source Info	Description - <i>optional</i> Info
Anywhere	<input type="text" value="0.0.0.0/0"/>	e.g. SSH for admin desktop

Painel de personalização do security group.

Durante o processo de criação de uma nova instância, você poderá selecionar um SG previamente criado ou criar um novo, de forma simplificada ou personalizada. Ao criar um novo SG, é provável que já exista uma regra liberando a porta de acesso remoto (SSH para linux, RDP para Windows), sendo recomendado que você restrinja o acesso para que esse canal não fique aberto a qualquer um.

Disco do EC2

O Amazon EC2 oferece opções de armazenamento de dados para suas instâncias que são flexíveis, econômicas e fáceis de usar. Cada opção tem uma combinação única de desempenho e durabilidade.

Essas opções de armazenamento podem ser usadas independentemente ou em combinação para atender às suas necessidades.

Em geral, durante a criação de uma instância EC2, um disco raiz EBS (Elastic Block Store) de tamanho mínimo e tipo sugerido para a AMI selecionada já é adicionado, podendo ser alterado e até adicionados novos tipos de armazenamento.

▼ Configure storage [Info](#)

Advanced

1x GiB Root volume (Not encrypted)

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage

Add new volume

0 x File systems [Edit](#)

Indicação de tamanho (8GB) e tipo (gp2) de disco.

Como podemos ver na imagem, um disco de 8 GB do tipo gp2(SSD) é o padrão para instâncias Ubuntu 22.04.



Atenção

Deve-se respeitar a configuração mínima de tamanho e tipo que vêm pré-preenchidos, podendo ser alterados para tamanhos maiores ou tipos melhores. Do contrário, a instância pode ter problemas de performance ou até não dar boot no sistema operacional.

Scripts de boot nas instâncias EC2

Confira a técnica para aprender a executar o scripts durante o provisionamento de uma instância.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Detalhes avançados

Ao final do processo de criação de uma instância EC2, antes de lançar, ainda é possível selecionar algumas opções, onde podemos destacar:

Termination protection

Se ativado, a instância não pode ser encerrada até que a proteção contra encerramento seja desativada.

Detailed Cloudwatch

Diminui de 5 para 1 minuto o intervalo de métricas da instância, permitindo acompanhar com mais detalhes as variações de comportamento de recursos da máquina como vCPU e rede.

User Data

Campo que permite serem adicionados comandos ou scripts de comandos para serem executados pela instância logo após o boot.

Verificando o aprendizado

Questão 1

São tipos de instâncias de propósito geral:

A

m6i, m6g, r6g.

B

t3a, r5a, m5a.

C

c6g, c5a, c4.

D

t4g, m6i, m4.

E

r5, r6g, r4.



A alternativa D está correta.

Máquina do tipo "t" e "m" são de uso geral, pois são instâncias ideais para a maioria dos aplicativos que não exigem recursos especiais de hardware, como aceleração de GPU, e possuem um bom equilíbrio de quantidade de CPU e memória.

Questão 2

Qual das alternativas a seguir é a definição correta de Amazon Machine Image (AMI)?

A

É um tipo de imagem de máquina virtual que pode ser usada para instalar e executar sistemas operacionais e aplicativos em máquinas virtuais na nuvem da Amazon Web Services (AWS).

B

É um tipo de imagem de máquina virtual que pode ser usada para instalar e executar sistemas operacionais e aplicativos em máquinas virtuais em qualquer nuvem pública ou privada.

C

É um sistema operacional baseado em Linux que é usado exclusivamente para executar aplicativos na nuvem da AWS.

D

É um tipo de imagem de máquina virtual que pode ser usada para instalar e executar sistemas operacionais e aplicativos em máquinas virtuais localmente em um computador pessoal.

E

É um tipo de imagem de máquina virtual que pode ser usada para instalar e executar sistemas operacionais e aplicativos em máquinas virtuais em qualquer nuvem pública ou privada, exceto a AWS.



A alternativa A está correta.

Amazon Machine Image (AMI) é um tipo de imagem de máquina virtual que pode ser usada para instalar e executar sistemas operacionais e aplicativos em máquinas virtuais na nuvem da Amazon Web Services (AWS). AMIs são usadas para criar instâncias EC2, que são máquinas virtuais na nuvem da AWS, de maneira que contêm tudo o que é necessário para iniciar uma instância EC2, incluindo o sistema operacional, aplicativos e configurações.

Block Storage – Amazon EBS

O Amazon Elastic Block Storage é um serviço que fornece volumes de armazenamento em blocos, e que pode ser usado com instâncias EC2. Se você desligar ou apagar uma instância do Amazon EC2, todos os dados no volume do EBS anexo permanecerão disponíveis, permitindo reanexar a uma instância.

Para criar um volume do EBS, basta definir a configuração (como tamanho e tipo do volume) e a zona de disponibilidade da região onde será provisionado.

Depois de criar um volume do EBS, ele pode ser anexado a uma instância do Amazon EC2, similar à forma como você anexa um HD externo ao seu computador. Os volumes EBS agem de forma muito parecida a um HD externo. Grande parte dos volumes do Amazon EBS só pode ser conectada a uma instância por vez. A maioria dos volumes do EBS tem uma relação um para um com instâncias do EC2, portanto, eles não podem ser compartilhados ou anexados a várias instâncias ao mesmo tempo (recentemente, a AWS anunciou o recurso multi-attach do Amazon EBS, que permite que volumes sejam anexados a várias instâncias do EC2 ao mesmo tempo. Esse recurso não está disponível para todos os tipos de instância e todas as instâncias devem estar na mesma zona de disponibilidade).



Dica

Você pode desanexar um volume do EBS de uma instância do EC2 e anexá-lo a outra instância do EC2 que esteja na mesma zona de disponibilidade para acessar os dados nela contidos.

A unidade externa é separada do computador. Isso significa que, se ocorrer um acidente e o computador cair, você ainda terá seus dados na unidade externa. Isso também vale para volumes EBS em relação a uma instância Amazon EC2.

Soluções de armazenamento

Neste vídeo, você conhecerá as principais soluções de armazenamento na Amazon Web Services, características que as diferenciam e casos de uso de cada uma.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Snapshots de EBS

Como os volumes do EBS são para dados que precisam perdurar, é importante fazer backup dos dados. Você pode fazer backups complementares de volumes do EBS criando snapshots do Amazon EBS.

Um snapshot do EBS é um backup incremental. Isso significa que o primeiro backup de um volume copia todos os dados. Nos backups subsequentes, somente os blocos de dados que foram alterados desde o snapshot mais recente serão salvos.



Exemplo

Se você tiver 10 GB de dados em um volume e apenas 2 GB de dados tiverem sido modificados desde o último snapshot, apenas os 2 GB que foram alterados serão gravados no Amazon Simple Storage Service (Amazon S3).

Quando você faz um snapshot de qualquer um dos seus volumes EBS, os backups são armazenados de forma redundante em várias zonas de disponibilidade usando o Amazon S3. Esse aspecto de armazenar o backup no Amazon S3 é tratado pela AWS e, portanto, você não precisará interagir com o Amazon S3 para trabalhar com seus snapshots do EBS. Você os gerencia no console do Amazon EBS, que faz parte do console do Amazon EC2.

Os snapshots do EBS podem ser usados para criar vários volumes, estejam eles na mesma zona de disponibilidade ou em outra.

Quando você cria um novo volume a partir de um snapshot, ele é uma cópia exata do volume original no momento em que o snapshot foi obtido.

Casos de uso do EBS

O Amazon EBS é útil quando você precisa recuperar dados rapidamente e manter os dados por um longo prazo. Os volumes são comumente usados nos seguintes cenários:

Sistemas operacionais

Volumes de inicialização/raiz para armazenar um sistema operacional. O dispositivo raiz de uma instância executada a partir de uma imagem de máquina da Amazon (AMI) geralmente é um volume do Amazon EBS. Eles são comumente referidos como AMIs com suporte de EBS.

Bancos de dados

Uma camada no Amazon EC2 de armazenamento para bancos de dados em execução que dependem de leituras e gravações transacionais.

Aplicativos corporativos

O Amazon EBS fornece armazenamento de blocos confiável para executar aplicativos essenciais aos negócios.

Aplicativos com taxa de transferência intensiva

Aplicativos que executam leituras e gravações longas e contínuas.

Tipos de EBS

Neste vídeo, você aprenderá a escolher o tipo ideal de EBS para cada uso.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Os volumes do Amazon EBS são organizados em duas categorias principais: unidades de estado sólido (SSDs) e unidades de disco rígido (HDDs). Os SSDs oferecem forte desempenho para entrada/saída aleatória (E/S), enquanto os HDDs oferecem forte desempenho para E/S sequencial. A AWS oferece dois tipos de cada um. A tabela a seguir pode ajudar você a decidir qual volume de EBS é a melhor opção para sua carga de trabalho.

Tipos de volume EBS	Descrição	Caso de uso	Tamanho	IOPS Máximo	Throughput Máximo
EBS Provisioned IOPS SSD (io1)	SSD de desempenho mais alto, projetado para carga de trabalho transacional sensível à latência	NoSQL com uso intensivo de I/O e bancos de dados relacionais	4 GB–16 TB	64,000	1,000 MB/s
EBS General Purpose SSD (gp2/gp3)	SSD de uso geral que equilibra preço e desempenho para uma ampla variedade de cargas de trabalho transacionais	Volumes de boot, aplicativos interativos de baixa latência, desenvolvimento e teste	1 GB–16 TB	16,000	250 MB/s
Throughput Optimized HDD (st1)	HDD de baixo custo projetado para cargas de trabalho com alto rendimento e acesso frequente	Big data, data warehouses, processamento de logs	500 GB–16 TB	500	500 MB/s
Cold HDD (sc1)	HDD de menor custo projetado para cargas de trabalho acessadas com menos frequência	Dados mais frios que exigem menos varreduras por dia	500 GB–16 TB	250	250 MB/s

Volumes EBS e suas características.
Gustavo Ribeiro

Object storage - Amazon S3

Ao contrário do Amazon Elastic Block Store (Amazon EBS), o Amazon Simple Storage Service (Amazon S3) é uma solução de armazenamento independente, que não está vinculada à computação e permite que você recupere seus dados de qualquer lugar na web. Se você usou um serviço de armazenamento on-line para fazer backup dos dados de sua máquina local, provavelmente usou um serviço semelhante ao Amazon S3.

Nesse serviço, você armazena seus objetos em contêineres chamados de buckets (baldes).

Não é possível fazer upload de nenhum objeto, nem mesmo uma única foto, para o Amazon S3 sem criar um bucket primeiro.

Ao criar um bucket você especifica, no mínimo, dois detalhes: o nome desse bucket e a região da AWS na qual deseja que ele resida.



Para escolher uma região, você normalmente selecionará uma que tenha usado para outros recursos, como sua computação. Quando você escolhe uma região para seu bucket, todos os objetos que você coloca dentro dele serão armazenados de forma redundante em vários dispositivos, em várias zonas de disponibilidade. Esse nível de redundância foi projetado para fornecer aos clientes do Amazon S3, 99,999999999% de durabilidade e 99,99% de disponibilidade para objetos em um determinado ano.

Ciclo de vida de objetos

Neste vídeo, você verá como criar regras de ciclo de vida no Amazon S3.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Versionamento no S3

Conforme vimos, o Amazon S3 identifica objetos em parte usando o nome do objeto. Por exemplo, ao carregar uma foto, você pode nomear o objeto `photo.gif` e armazená-lo em uma pasta chamada `PhotosFiles`. Se você não usar o controle de versão do Amazon S3, toda vez que fizer upload de um objeto chamado `photo.gif` para a pasta `PhotosFiles`, ele substituirá o arquivo original.

Isso pode ser um problema por vários motivos como, por exemplo:

1

Exemplo 1

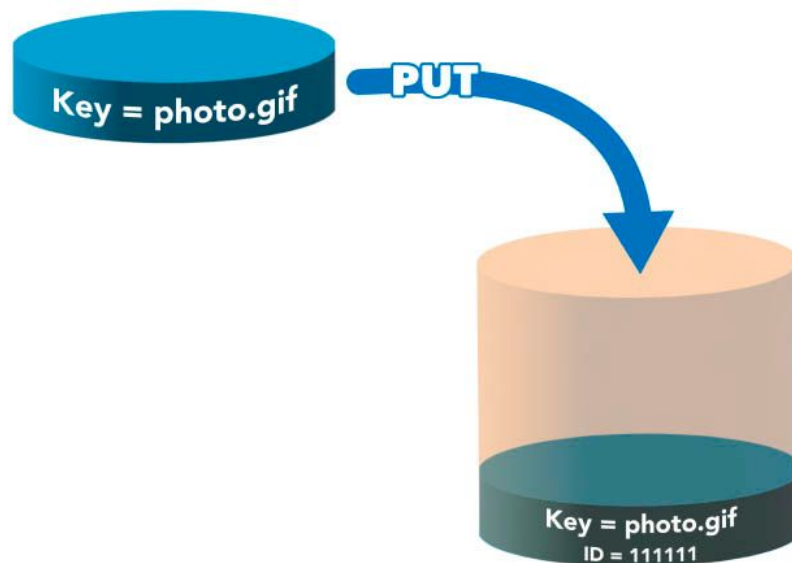
O nome do arquivo `photo.gif` é um nome comum para um objeto de foto. Você ou outra pessoa que tenha acesso ao bucket pode não ter a intenção de substituí-lo, mas uma vez substituído, o arquivo original não pode ser acessado.

2

Exemplo 2

Você pode querer preservar versões diferentes de `photo.gif`. Sem controle de versão, se você quiser criar uma nova versão de `photo.gif`, precisará carregar o objeto e escolher um nome diferente para ele. Ter vários objetos, todos com pequenas diferenças nas variações de nomenclatura, pode causar confusão e desordem nos buckets do S3.

Para neutralizar esses problemas, você pode usar o controle de versão do S3. O controle de versão mantém várias versões de um único objeto no mesmo bucket. Isso preserva versões antigas de um objeto sem usar nomes diferentes, o que ajuda na recuperação de arquivos de exclusões acidentais, substituições acidentais ou falhas de aplicativos.



Bucket PhotosFiles, com versionamento habilitado.

Se você habilitar o controle de versão para um bucket, o Amazon S3 gerará automaticamente um ID de versão exclusivo para o objeto.



Exemplo

Em um bucket você pode ter dois objetos com a mesma chave, mas diferentes IDs de versão, como photo.gif (versão 111111) e photo.gif (versão 111122).

Os buckets com versionamento, habilitados para controle de versão, permitem que você recupere objetos de exclusão ou substituição acidental.

A exclusão de um objeto não o remove permanentemente.

Em vez disso, o Amazon S3 coloca um marcador no objeto que mostra que você tentou excluí-lo. Se quiser restaurar o objeto, você pode remover o marcador e ele restabelece o objeto. Se você substituir um objeto, isso resultará em uma nova versão do objeto no bucket. Mas você ainda terá acesso às versões anteriores do objeto.

Classes de armazenamento no S3

Quando você carrega um objeto no Amazon S3 e não especifica a classe de armazenamento, você o carrega na classe de armazenamento padrão. Tudo que aprendemos até aqui é sobre a classe de armazenamento padrão do Amazon S3 mesmo sem você saber!

As classes de armazenamento do Amazon S3 permitem que você altere seu nível de armazenamento quando suas características de dados mudarem. Por exemplo, se você estiver acessando suas fotos antigas com pouca frequência, talvez queira alterar a classe de armazenamento das fotos para economizar custos.

Saiba mais um pouco sobre cada uma das principais classes de armazenamento no S3:

Standard

O S3 Standard oferece um armazenamento de objetos com altos níveis de resiliência, disponibilidade e performance para dados acessados com frequência. Como fornece baixa latência e alto throughput, o S3 Standard é adequado para uma grande variedade de casos de uso, como aplicativos na nuvem, sites dinâmicos, distribuição de conteúdo, aplicativos móveis e de jogos e dados analíticos de big data.

Standard-IA

O S3 Standard-IA é indicado para dados acessados com menos frequência, mas que exigem acesso rápido, quando esses dados são necessários. A categoria S3 Standard-IA oferece os altos níveis de resiliência e throughput e uma baixa latência da categoria S3 Standard, com taxas reduzidas por GB de armazenamento e por GB de recuperação. A combinação de baixo custo e alta performance torna a classe S3 Standard-IA ideal para armazenamento de longa duração, backups e data stores para arquivos de recuperação de desastres.

Glacier Instant Retrieval

A Amazon S3 Glacier Instant Retrieval é uma classe de armazenamento de arquivos que oferece o armazenamento de custo mais baixo para dados de longa duração, que raramente são acessados e exigem recuperação em milissegundos. Com a S3 Glacier Instant Retrieval, você pode economizar até 68% nos custos de armazenamento em comparação com o uso da classe de armazenamento S3 Standard-IA, quando seus dados são acessados uma vez por trimestre. A S3 Glacier Instant Retrieval oferece o acesso mais rápido ao armazenamento de arquivo, com a mesma taxa de transferência e acesso em milissegundos que as classes de armazenamento S3 Standard e S3 Standard – IA. O S3 Glacier Instant Retrieval é ideal para arquivar dados que precisam de acesso imediato, como imagens médicas, recursos de mídia de notícias ou arquivos de conteúdo gerado pelo usuário.

Glacier Flexible Retrieval

O S3 Glacier Flexible Retrieval oferece armazenamento de baixo custo, com custo até 10% menor (do que o S3 Glacier Instant Retrieval), para dados de arquivo que são acessados 1 a 2 vezes por ano e recuperados de forma assíncrona. Para dados de arquivo que não requerem acesso imediato, mas que precisam da flexibilidade para recuperar grandes conjuntos de dados sem nenhum custo, como casos de uso de backup ou recuperação de desastres, a S3 Glacier Flexible Retrieval é a classe de armazenamento ideal. É uma solução ideal para backup, recuperação de desastres, necessidades de armazenamento externo de dados e para quando alguns dados ocasionalmente precisam ser recuperados em minutos e você não quer se preocupar com custos.

Glacier Deep Archive

A S3 Glacier Deep Archive é a classe de armazenamento mais acessível do Amazon S3 e oferece suporte à retenção e preservação digital de longo prazo para dados que podem ser acessados uma ou duas vezes por ano. Essa classe é projetada para clientes que mantêm conjuntos de dados por 7 a 10 anos ou mais para cumprir requisitos de conformidade regulatória, especialmente em setores altamente regulados como serviços financeiros, saúde e setores públicos. O S3 Glacier Deep Archive também pode ser usado para casos de uso de backup e recuperação de desastres, além de ser uma alternativa mais barata e fácil de gerenciar em comparação aos sistemas de fita magnética como bibliotecas on-premises ou serviços externos, e podem ser restaurados em até 12 horas.

Casos de uso

O Amazon S3 é um serviço de armazenamento amplamente utilizado, com muito mais casos de uso do que caberia em uma tela. A lista a seguir resume algumas das formas mais comuns de usar o Amazon S3:

Backup e armazenamento

O Amazon S3 é um lugar natural para fazer backup de arquivos porque é altamente redundante. Conforme mencionado anteriormente, a AWS armazena seus snapshots do EBS no S3 para aproveitar sua alta disponibilidade.

Hospedagem de mídia

Como é possível armazenar objetos ilimitados e cada objeto individual pode ter até 5 TBs, o Amazon S3 é um local ideal para hospedar uploads de vídeos, fotos e músicas.

Data lakes

O Amazon S3 é uma base ideal para um data lake devido à sua escalabilidade virtualmente ilimitada. Você pode aumentar o armazenamento de gigabytes para petabytes de conteúdo, pagando apenas pelo que usar.

Sites estáticos

É possível configurar seu bucket S3 para hospedar um site estático de HTML, CSS e scripts do lado do cliente.

File storage – Amazon EFS

No armazenamento de arquivos, vários clientes (como usuários, aplicativos, servidores e assim por diante) podem acessar dados armazenados em pastas de arquivos compartilhadas. Nessa abordagem, um servidor de armazenamento organiza os arquivos por meio do uso de armazenamento em bloco, com um sistema local de arquivos. Os clientes acessam os dados por meio de caminhos de arquivo.

Comparado ao armazenamento em blocos e ao armazenamento de objetos, o armazenamento de arquivos é ideal para casos de uso em que muitos serviços e recursos precisam acessar os mesmos dados ao mesmo tempo.



O Amazon Elastic File System (Amazon EFS) é um sistema de arquivos escalável, usado com os serviços de nuvem AWS e recursos locais. À medida que você adiciona e remove arquivos, o Amazon EFS expande e retrai automaticamente, de forma que pode dimensionar sob demanda para petabytes sem interromper os aplicativos.

Verificando o aprendizado

Questão 1

O gp3 é o tipo de EBS mais recomendado para qual dos seguintes casos de uso?

A

Volumes de boot, aplicativos interativos de baixa latência, desenvolvimento e teste.

B

NoSQL com uso intensivo de I/O e bancos de dados relacionais.

C

Big data, data warehouses, processamento de logs.

D

Dados mais frios, que exigem menos varreduras por dia.

E

Backups e armazenamento de longo prazo.



A alternativa A está correta.

Volumes gp3 possuem boa performance de I/O e baixa latência o que o torna ideal para uso para volume de boot, aplicativos de baixa latência e pode ser utilizado até bancos de dados, seja relacional ou não relacional, porém os tipos io1 ou io2 possuem melhor performance para o caso de uso de bancos de dados relacionais e não relacionais.

Questão 2

Qual é o SLA de durabilidade de objetos no Amazon S3?

A

99,9%.

B

99,99%.

C

9,99%.

D

0,99%.

E

99,999999999%.



A alternativa E está correta.

O SLA de durabilidade se refere à capacidade de um sistema de armazenar dados de maneira segura e confiável ao longo do tempo. Isso inclui a capacidade de um sistema de proteger os dados contra perda ou corrupção e garantir que eles estejam sempre disponíveis para acesso. No S3 esse SLA é composto por "11 noves", ou seja, 99,999999999% em um ano. Já o SLA de disponibilidade se refere à capacidade de um sistema de estar disponível para uso pelos usuários. Isso inclui a capacidade de um sistema de estar sempre disponível para acesso e uso, mesmo em caso de falhas ou problemas. No Amazon S3, o SLA de disponibilidade em um ano é de 99,99%.

VPC – Virtual Private Cloud

Neste vídeo, você conhecerá a VPC (Amazon Virtual Private Cloud), a distribuição de endereçamento e recursos de conectividade com a internet.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Imagine os milhões de clientes que usam os serviços AWS e os milhões de recursos criados por eles, como as instâncias do Amazon EC2. Sem limites de acesso para todos esses recursos, o tráfego de rede fluiria sem restrições entre eles, permitindo que todos os recursos de todos os clientes tenham acesso uns aos outros.

Um serviço de rede que pode ser usado para definir limites para seus recursos AWS é o Amazon Virtual Private Cloud (Amazon VPC).

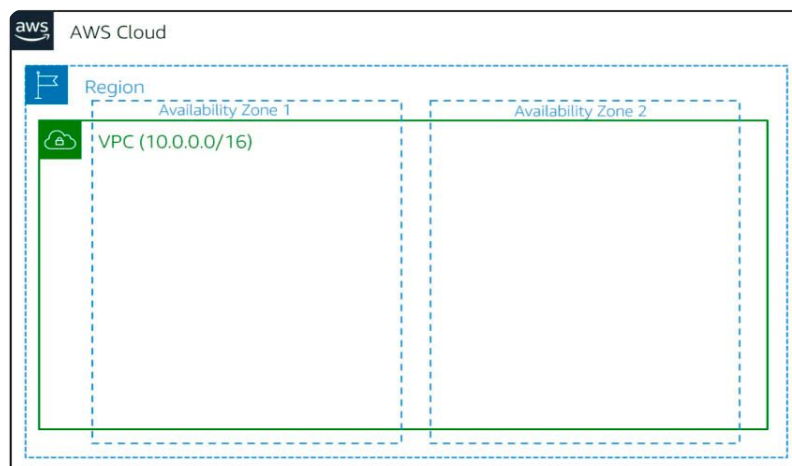
Esse serviço permite que você provisione uma seção isolada da nuvem AWS. Nessa seção, você pode executar os recursos em uma rede virtual que definir. Em uma Virtual Private Cloud (VPC), você pode organizar seus recursos em sub-redes. Uma sub-rede é uma seção de uma VPC que pode conter recursos como instâncias do Amazon EC2.

Ao criar uma VPC, você deve escolher três fatores principais:

- O nome do VPC.
- A região onde o VPC vai ser provisionado, já que cada VPC abrange várias zonas de disponibilidade na região selecionada.
- O intervalo de IP para a VPC na notação CIDR. Isso determina o tamanho da sua rede. Cada VPC pode ter até quatro intervalos de IP /16.

Usando essas informações, a AWS provisionará uma rede e endereços IP para essa rede.

VPCs ficam dentro da região escolhida, mas podem se estender por múltiplas zonas de disponibilidade.



Zonas de disponibilidade das VPCs.

Sub-redes de VPC

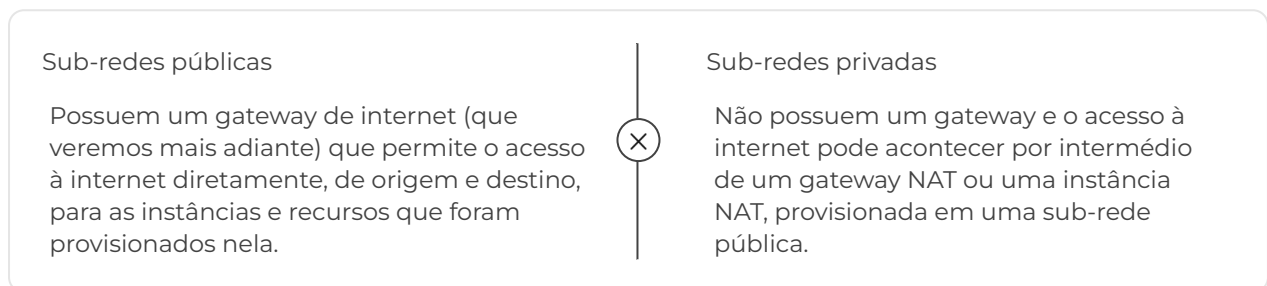
Depois de criar sua VPC, você deve criar sub-redes dentro da rede. Pense nelas como redes menores dentro de sua rede base – ou redes locais virtuais (VLANs) em uma rede local tradicional. Em uma rede local, o caso de uso típico para sub-redes é isolar ou otimizar o tráfego de rede. Na AWS, essas sub-redes são usadas para fornecer alta disponibilidade e opções de conectividade para seus recursos.

Ao criar uma sub-rede, você deve especificar:

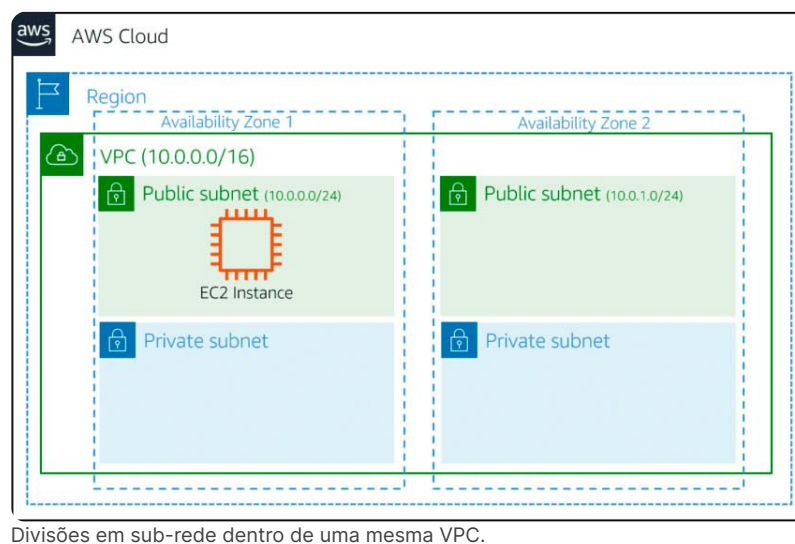
- A VPC na qual você deseja que sua sub-rede resida.
- A zona de disponibilidade na qual você deseja que sua sub-rede resida.
- O bloco CIDR para sua sub-rede, que deve ser um subconjunto do bloco VPC CIDR.

Ao iniciar uma instância do EC2, você a inicia dentro de uma sub-rede, que estará localizada dentro da zona de disponibilidade que você escolher.

As sub-redes são basicamente de dois tipos: públicas e privadas.

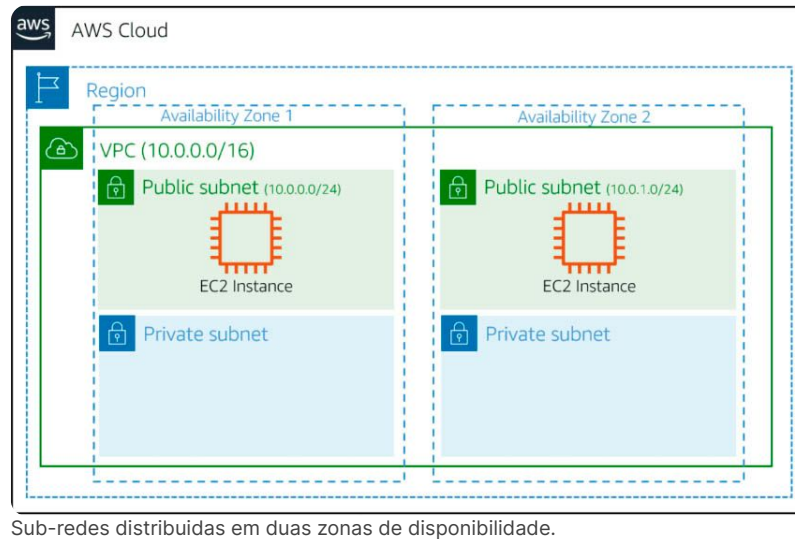


A seguir, podemos ver as divisões em sub-rede dentro de uma mesma VPC e em diferentes zonas de disponibilidade.



Alta disponibilidade em VPC

Ao criar suas sub-redes, tenha em mente a alta disponibilidade. Para manter a redundância e a tolerância a falhas crie, ao menos, duas sub-redes configuradas em duas zonas de disponibilidade. Lembre-se de que “tudo falha o tempo todo”. Com a rede de exemplo, se uma das AZs falhar, você ainda terá seus recursos disponíveis em outra AZ como backup/redundância. Recursos como instâncias EC2 não usufruem automaticamente de alta disponibilidade com múltiplas AZs pois, ao lançar um novo servidor virtual, este estará localizado em apenas uma AZ determinada. Aplicações hospedadas em instâncias EC2 e que possuem capacidade de escala horizontal (múltiplos servidores, com a mesma aplicação, funcionando como um cluster) podem usufruir de alta disponibilidade se for provisionada mais de uma máquina virtual em AZs diferentes e essas máquinas utilizarem, como exemplo, um balanceador de carga para distribuição do tráfego.



Endereço IP elástico

Neste vídeo, você verá como se beneficiar dos IPs elásticos.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Um endereço IP elástico (Elastic IP) é um endereço IPv4 público e estático projetado para computação em nuvem dinâmica. Você pode associar um endereço IP elástico a qualquer instância ou interface de rede em qualquer VPC em sua conta. Com um endereço IP elástico, você pode mascarar a falha de uma instância remapeando rapidamente o endereço para outra instância em sua VPC. A seguir, algumas considerações sobre IPs elásticos:

- Um endereço IP elástico pode ser associado a uma única instância ou interface de rede por vez.
- Você pode mover um endereço IP elástico de uma instância ou interface de rede para outra.
- Se você associar um endereço IP elástico à interface de rede primária de sua instância, seu endereço IPv4 público atual (se houver) será liberado para o pool de endereços IP públicos. Se você desassociar o endereço IP elástico, a interface de rede primária receberá automaticamente um novo endereço IPv4 público em alguns minutos. Isso não se aplica se você tiver anexado uma segunda interface de rede à sua instância.
- Existe uma cobrança de 0,005 dólares por hora quando eles não estão associados a uma instância em execução ou quando estão associados a uma instância interrompida ou a uma interface de rede não conectada. Enquanto sua instância estiver em execução, você não será cobrado por um endereço IP elástico associado à instância, mas por quaisquer endereços IP elásticos adicionais (a partir do segundo na mesma instância) associados à instância.
- Endereços IP elásticos para IPv6 não são suportados.

IPs reservados

Para que a AWS configure sua VPC adequadamente, ela reserva cinco endereços IP em cada sub-rede. Esses endereços IP são usados para roteamento, Domain Name System (DNS) e gerenciamento de rede.



Exemplo

Considere uma VPC com o intervalo de IP 10.0.0.0/22. A VPC inclui um total de 1.024 endereços IP. Isso é dividido em quatro sub-redes de tamanho igual, cada uma com um intervalo de IP /24 com 256 endereços IP. De cada um desses intervalos de IP, existem apenas 251 endereços IP que podem ser usados, já que a AWS reserva cinco. A primeira subnet pode ser 10.0.0.0/24.

Vejamos como são reservados pela AWS os cinco endereços IP:

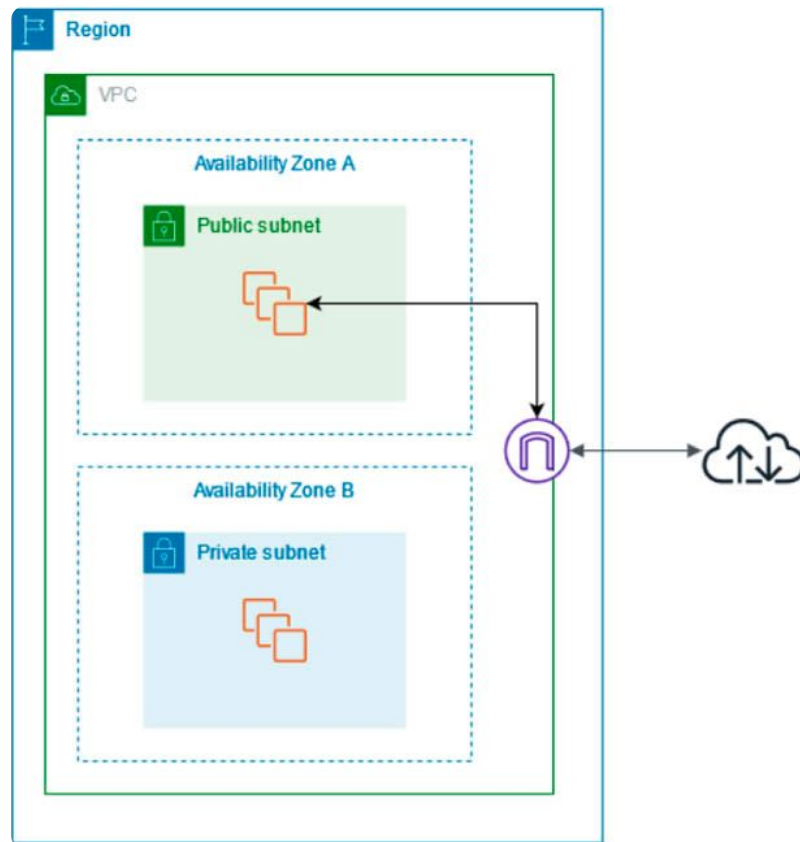
Endereço IP	Propósito
10.0.0.0	Endereço de rede
10.0.0.1	Roteador da VPC
10.0.0.2	Servidor DNS
10.0.0.3	De uso futuro
10.0.0.255	Endereço de broadcast

Especificação de IPs e seus propósitos no AWS.
Gustavo Ribeiro

Os cinco endereços IP reservados podem afetar a forma como você projeta sua rede. Um ponto de partida comum para quem é novo na nuvem é criar uma VPC com um intervalo de IP de /16 e criar sub-redes com um intervalo de IP de /24. Isso fornece uma grande quantidade de endereços IP para trabalhar nos níveis de VPC e sub-rede.

Gateway de internet (internet gateway)

Para habilitar a conectividade para sua VPC, você deve criar um gateway de internet. Pense nele como algo semelhante a um modem. Da mesma forma que um modem conecta seu computador à internet, o gateway conecta sua VPC. Ao contrário do seu modem em casa, que às vezes fica inativo ou offline, o gateway de internet é altamente disponível e escalável, abrangendo todas as AZs disponíveis. Depois de criar um gateway de internet, você o anexa à sua VPC.



Gateway anexado à sua VPC.

Gateway NAT

Um gateway NAT é um serviço de Network Address Translation (NAT) que pode ser usado para que as instâncias em uma sub-rede privada possam se conectar a serviços fora de sua VPC, mas os serviços externos não podem iniciar uma conexão com essas instâncias.

Ao criar um gateway NAT, você especifica um dos seguintes tipos de conectividade:

Pública (padrão)

As instâncias em sub-redes privadas podem se conectar à internet por meio de um gateway NAT público, mas não podem receber conexões de entrada não solicitadas. Você cria um gateway NAT público em uma sub-rede pública e deve associar a ele, na criação, um endereço IP elástico. Você roteia o tráfego do gateway NAT para o gateway da internet para a VPC. Como alternativa, pode ser usado um gateway NAT público para se conectar a outras VPCs ou à sua rede local. Nesse caso, você roteia o tráfego do gateway NAT por meio de um gateway de trânsito ou de um gateway privado virtual.

Privada

As instâncias em sub-redes privadas podem se conectar a outras VPCs ou à sua rede local por meio de um gateway NAT privado. Serve basicamente para isolar, mas dar conectividade a outras redes privadas, como em outras VPCs ou com redes on-premise via VPN ou conectividade direta. Você não pode associar um endereço IP elástico a um gateway NAT privado. É possível anexar um gateway da internet a uma VPC usando um gateway NAT privado. Porém, se rotear o tráfego do gateway NAT para o gateway da internet, este último vai descartar o tráfego.

O gateway NAT substitui o endereço IP de origem das instâncias por endereço IP próprio. Para um gateway NAT público, este é o seu endereço IP elástico. Para um gateway NAT privado, este é o endereço IP privado do

gateway NAT. Ao enviar tráfego de resposta para as instâncias, o dispositivo NAT converte os endereços de volta para o endereço IP original.

Os principais casos de uso do gateway NAT são:

1

Acesso à internet a partir de uma sub-rede privada

Você pode usar um gateway NAT público para permitir que instâncias em uma sub-rede privada enviem tráfego de saída para a internet, evitando sua conexão com as instâncias.

2

Acesso a uma rede usando endereços IP permitidos

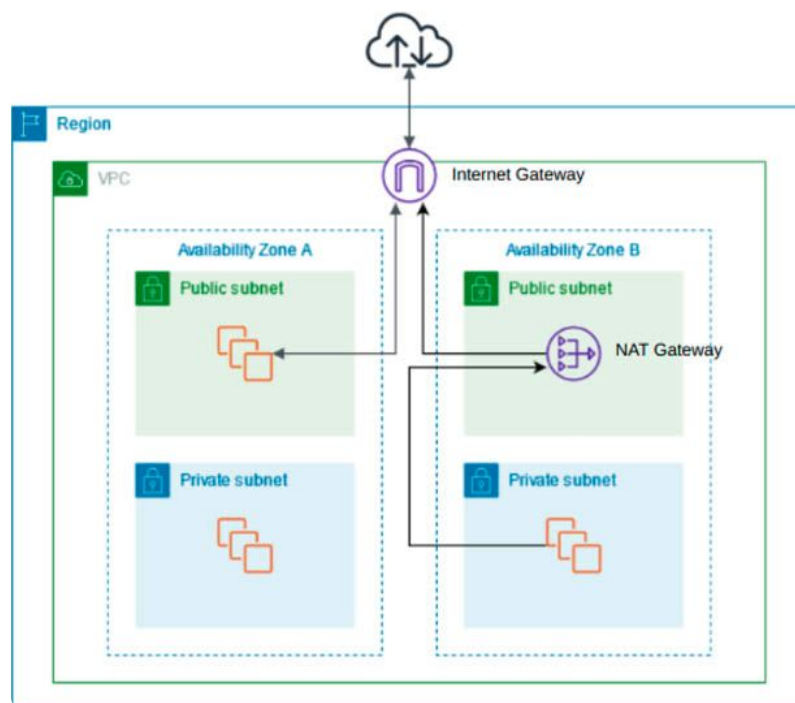
Se você precisa se comunicar com outra rede, pública ou privada, mas precisa restringir o acesso a determinados IPs, pode usar um gateway NAT, público ou privado, para representar um pool de IPs de instâncias, ao invés de criar regras de liberação para cada uma delas.

3

Ativar a comunicação entre redes sobrepostas

Você pode usar um gateway NAT privado para permitir a comunicação entre as redes, mesmo que tenham intervalos CIDR sobrepostos. Por exemplo, suponhamos que as instâncias na VPC A precisem acessar os serviços fornecidos pelas instâncias na VPC B. Elas utilizarão o mesmo bloco de IPs 10.0.0.0/24.

A seguir, veremos um diagrama com gateway NAT público em conjunto com o gateway de Internet para garantir conectividade.



Gateway NAT público em conjunto com o gateway de internet para garantir conectividade.

Tabelas de rota

Quando você cria uma VPC, a AWS cria tabelas de rota principal. As tabelas de rota contêm um conjunto de regras, chamadas rotas, que são usadas para determinar para onde é direcionado o tráfego de rede. A AWS supõe que, ao criar uma nova VPC com sub-redes, você deseja que o tráfego flua entre elas. Portanto, a configuração padrão da tabela de rota principal tem por objetivo permitir o tráfego entre todas as sub-redes da rede local. A seguir, apresentamos um exemplo de uma tabela de rota principal.

Vamos considerar que o destino e o alvo são duas partes principais dessa tabela de rotas.

- O destino (destination) é um intervalo de endereços IP para o qual você deseja que seu tráfego vá. No exemplo do envio de uma carta, você precisa de um destino para encaminhá-la ao local apropriado. O mesmo ocorre para o tráfego de roteamento. Nesse caso, o destino é o intervalo de IP da rede VPC.
- O alvo (target) é a conexão por meio da qual será enviado o tráfego. Nesse caso, o tráfego é roteado pela rede VPC local.

Destination	Target	Status	Propagated
0.0.0.0/0	igw-63082	Active	No
172.31.0.0/16	local	Active	No

Tabela de rota indicando destino (destination) e alvo (target).

VPC padrão

Neste vídeo, você conhecerá a VPC que a AWS já pré-provisiona na região.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Ao criar uma conta na AWS, você já encontra uma VPC padrão (default) em cada região (a própria AWS já deixou isso pré-provisionado). Uma VPC padrão vem com uma sub-rede pública em cada zona de disponibilidade, um gateway de internet e configurações para habilitar a resolução de DNS, permitindo que você possa iniciar imediatamente as instâncias do Amazon EC2.

Uma VPC padrão é adequada para começar rapidamente e iniciar instâncias públicas, como um blog ou site simples. Você pode modificar os componentes de sua VPC padrão conforme necessário para adequar às suas necessidades.

A VPC padrão possui as seguintes características:

- VPC com um bloco CIDR IPv4 de tamanho /16 (172.31.0.0/16), o que permite fornecer até 65.536 endereços IPv4 privados.
- Sub-redes de tamanho /20 em cada zona de disponibilidade. Isso fornece até 4.096 endereços por sub-rede, lembrando que alguns são de uso reservado.
- Um gateway de internet já conectado à VPC padrão.
- Uma tabela de rotas principal que aponta todo o tráfego (0.0.0.0/0) para o gateway da internet.
- Grupo de segurança padrão.
- Opções de DHCP padrão definidas para sua conta da AWS com sua VPC padrão.

Verificando o aprendizado

Questão 1

Em uma sub-rede 10.0.0.0/24, de uma VPC na AWS, qual endereço privado entre as alternativas a seguir poderia ser de uma instância EC2, ou seja, não seria reservado?

A

10.0.0.255.

B

10.0.0.254.

C

10.0.0.2.

D

10.0.0.3.

E

10.0.0.1.



A alternativa B está correta.

O 10.0.0.1 é endereço de gateway padrão da rede (roteador) reservado para esse uso. O endereço 10.0.0.2 é reservado para DNS e o 10.0.0.3 é reservado para uso especial futuro. O endereço 10.0.0.255 é o endereço de broadcast da rede. Dessa forma, apenas o 10.0.0.254 dentre as opções listadas está livre para uso em um EC2.

Questão 2

Quantas AZs (Availability Zones) são necessárias para ter alta disponibilidade em uma VPC (Virtual Private Cloud)?

A

Uma.

B

Duas.

C

Três.

D

Quatro.

E

Cinco ou mais.



A alternativa B está correta.

Para ter alta disponibilidade em uma VPC, é recomendável usar ao menos duas AZs. Isso permite que o sistema continue funcionando mesmo se uma delas falhar ou ficar indisponível por algum motivo. Por exemplo, se você tiver instâncias EC2 em duas AZs diferentes, elas poderão continuar funcionando mesmo se uma das AZs ficar indisponível devido a um problema de hardware ou a um desastre natural.

Aplicação web estática e web dinâmica

Confira um caso de uso e arquiteturas de referência para provisionamento de aplicações na AWS, com os recursos a serem utilizados.

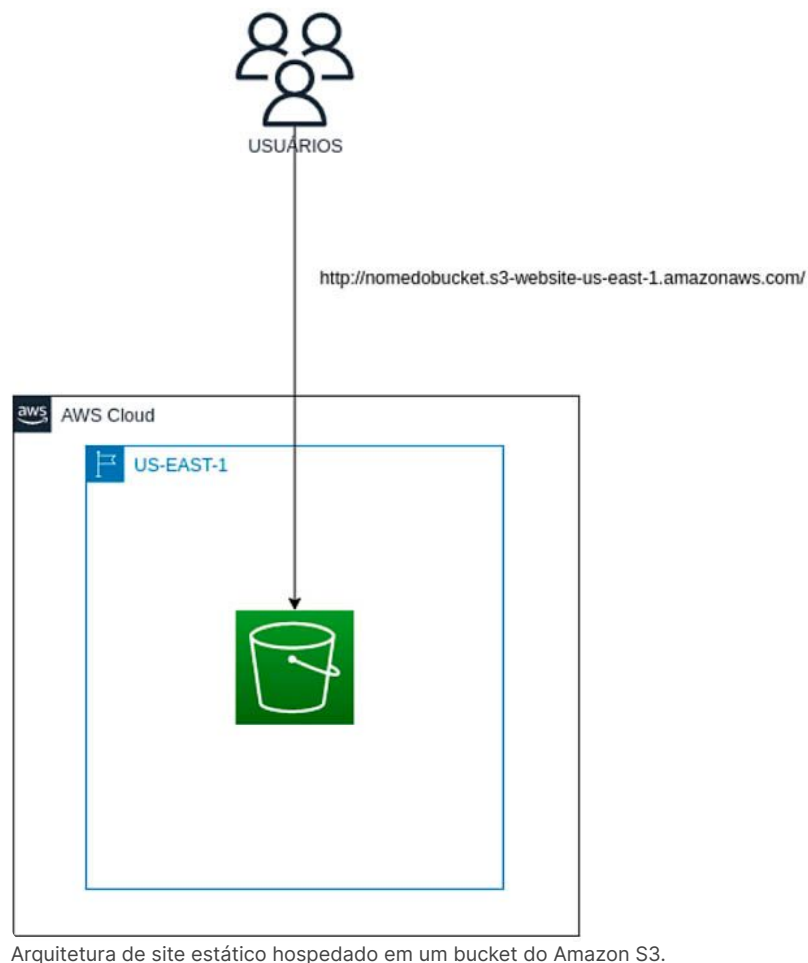


Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Com uma infinidade de propósitos, a Amazon Web Services permite que empresas e profissionais experimentem soluções antes mesmo de provisionarem suas cargas definitivas, a um baixo custo e sem compromissos de longo prazo. Vamos conhecer e experimentar alguns cenários que podem ser expandidos e customizados para se tornarem ambientes de produção, aplicáveis em nosso cotidiano.

A aplicação web é uma solução de software executada diretamente no navegador, não havendo instalação na máquina do usuário. Todo processamento da aplicação se concentra nos servidores de hospedagem e o usuário interage de forma remota. Na AWS existem várias arquiteturas para esse tipo de solução, que dependem das características e dos requisitos do ambiente da aplicação.



O Amazon S3 tem capacidade de atuar como serviço de hospedagem para sites estáticos, sendo uma opção muito interessante para esse tipo de aplicação quando comparado ao Amazon EC2, pois não possui custo de execução por hora, além do processo ser bastante simples.

Um dos primeiros passos é habilitar, nas propriedades do bucket, o Static website hosting, informando o nome de arquivo que será tratado como página padrão do site.

Static website hosting
Use this bucket to host a website or redirect requests. [Learn more](#)

Static website hosting

☐ Disable

☒ **Enable**

Hosting type

☒ **Host a static website**
Use the bucket endpoint as the web address. [Learn more](#)

☐ Redirect requests for an object
Redirect requests to another bucket or domain. [Learn more](#)

i For your customers to access content at the website endpoint, you must make all your content publicly readable. To do so, you can edit the S3 Block Public Access settings for the bucket. For more information, see [Using Amazon S3 Block Public Access](#)

Index document
Specify the home or default page of the website.

Error document - optional
This is returned when an error occurs.

Como habilitar e designar o arquivo que será hospedado.

Após salvar essa alteração, já estará disponível na seção Static website hosting o endereço que será usado para acesso público. Salve para uso posterior.

Static website hosting
Use this bucket to host a website or redirect requests. [Learn more](#)

Static website hosting

Enabled

Hosting type

Bucket hosting

Bucket website endpoint
When you configure your bucket as a static website, the website is available at the AWS Region-specific website endpoint of the bucket. [Learn more](#)

☒ **http://gustavoensiname.s3-website-us-east-1.amazonaws.com** [Learn more](#)

URL gerada para acesso público.

Será necessário alterar as permissões do bucket que, em geral, estará com bloqueios de acesso público. Porém, justamente por se tratar de um site de acesso público, liberaremos esse tipo de acesso. Na aba de permissões, todas as opções de bloqueio devem ficar desmarcadas no “Block public access”.

Amazon S3 > Buckets > gustavoensiname > Edit Block public access (bucket settings)

Edit Block public access (bucket settings) [Info](#)

Block public access (bucket settings)

Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, access point policies, or all. In order to ensure that public access to all your S3 buckets and objects is blocked, turn on Block all public access. These settings apply only to this bucket and its access points. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to your buckets or objects within, you can customize the individual settings below to suit your specific storage use cases. [Learn more](#)

☐ **Block all public access**

Turning this setting on is the same as turning on all four settings below. Each of the following settings are independent of one another.

- ☐ **Block public access to buckets and objects granted through new access control lists (ACLs)**
S3 will block public access permissions applied to newly added buckets or objects, and prevent the creation of new public access ACLs for existing buckets and objects. This setting doesn't change any existing permissions that allow public access to S3 resources using ACLs.
- ☐ **Block public access to buckets and objects granted through any access control lists (ACLs)**
S3 will ignore all ACLs that grant public access to buckets and objects.
- ☐ **Block public access to buckets and objects granted through new public bucket or access point policies**
S3 will block new bucket and access point policies that grant public access to buckets and objects. This setting doesn't change any existing policies that allow public access to S3 resources.
- ☐ **Block public and cross-account access to buckets and objects through any public bucket or access point policies**
S3 will ignore public and cross-account access for buckets or access points with policies that grant public access to buckets and objects.

[Cancel](#) [Save changes](#)

Aba de permissões de acesso público.

Ainda na parte de permissões do bucket, uma política precisa ser preenchida no “Bucket policy”, liberando o acesso de leitura dos objetos que estarão armazenados. A política apresentada a seguir pode ser usada para qualquer bucket, precisando apenas readequar o “nomedobucket” para o nome do bucket que você está utilizando.

```
python
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PublicReadGetObject",
      "Effect": "Allow",
      "Principal": "*",
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::nomedobucket/*"
      ]
    }
  ]
}
```

Bucket policy

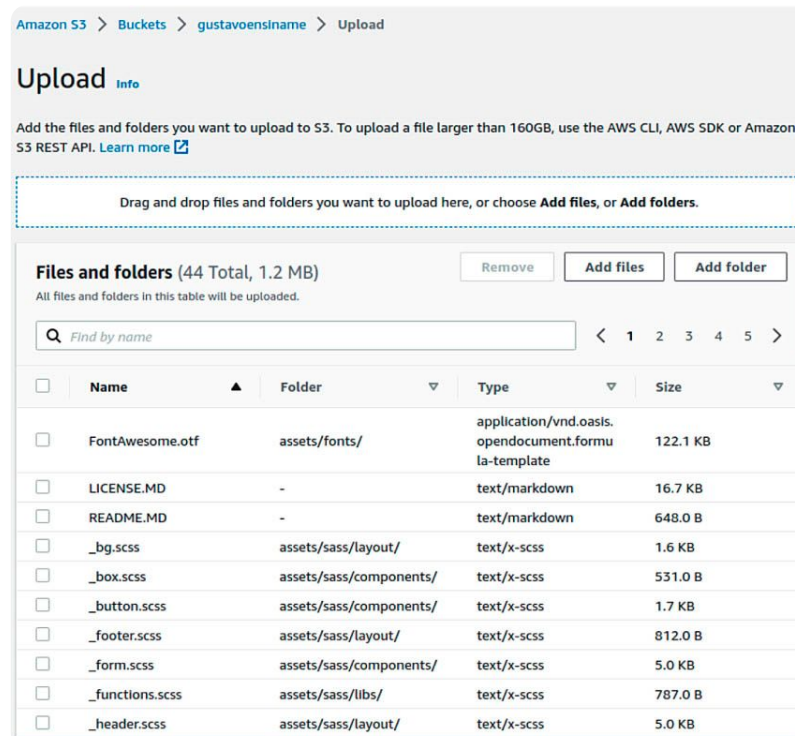
The bucket policy, written in JSON, provides access to the objects stored in the bucket. Bucket policies don't apply to objects owned by other accounts. [Learn more](#)

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PublicReadGetObject",
      "Effect": "Allow",
      "Principal": "*",
      "Action": "s3:GetObject",
      "Resource": "arn:aws:s3:::gustavoensiname/*"
    }
  ]
}
```

Print do campo de política do bucket.

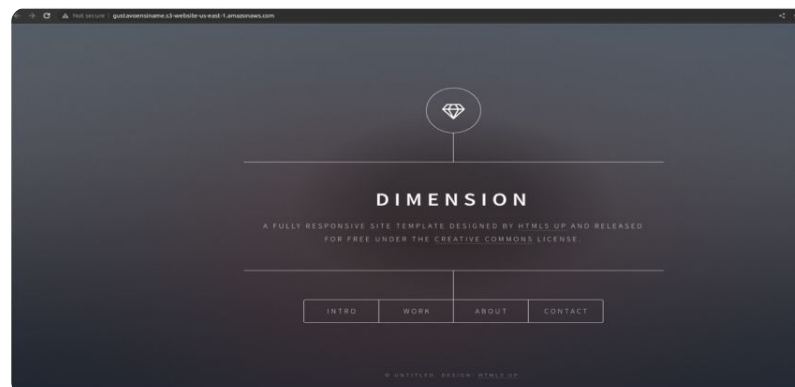
O bucket deve estar preparado para fazer hospedagem de sites estáticos a essa altura, caso tenha sido criado com todas as opções padrão, e você poderá fazer upload dos arquivos do site e utilizar o endereço gerado para fazer o acesso pelo seu navegador. Para fins didáticos, vamos utilizar um site estático de exemplo, da Cloud Academy, que pode ser baixado direto do GitHub, jogando todo conteúdo na raiz do bucket:

<https://github.com/cloudacademy/static-website-example>



Tela de upload de site estático.

Este será o site que você visualizará, indicando que a configuração da sua hospedagem está correta:



Página inicial do site estático usado como exemplo.

Nome do bucket para web hosting

Neste vídeo, você verá a importância do nome do bucket para uso de hospedagem web estática.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

Aplicação web dinâmica

Neste vídeo, você verá como encontrar e escolher AMIs de comunidade e marketplace.



Conteúdo interativo

Acesse a versão digital para assistir ao vídeo.

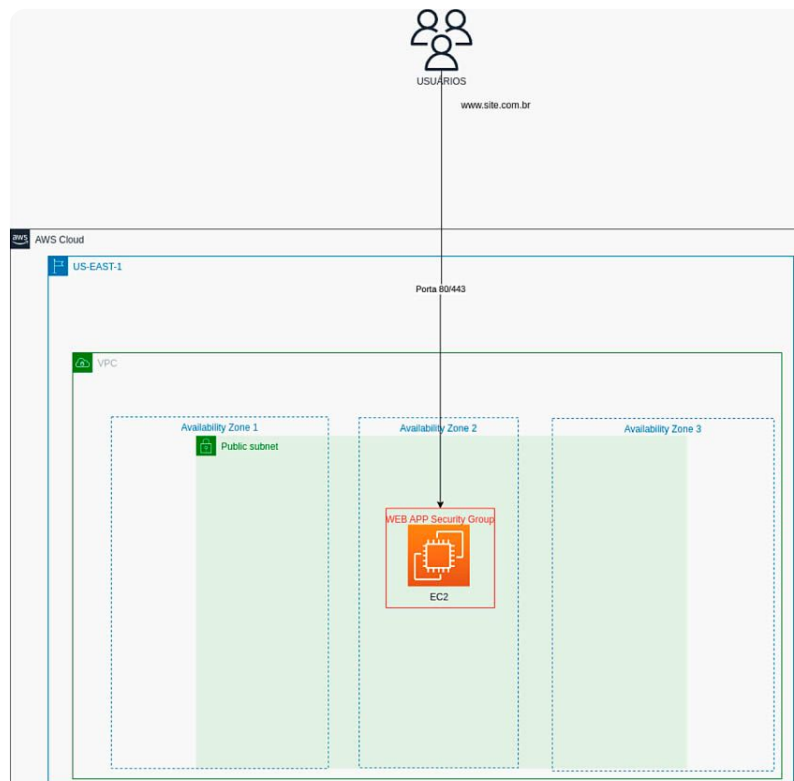
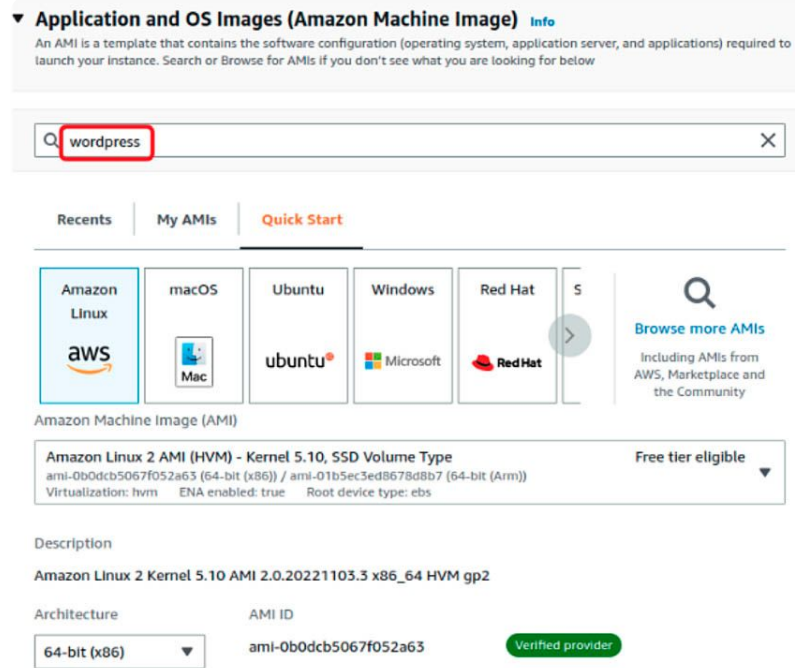


Diagrama de acesso ao EC2.

A maior parte dos softwares web existentes no mercado utilizam uma tecnologia dinâmica de conteúdo e, por isso, não conseguem ser atendidos pela solução anterior. Assim, precisamos de servidores de aplicação com capacidade de processar essas tecnologias. Com o Amazon EC2 é possível criar um ambiente aplicável a qualquer tecnologia disponível no mercado de linguagens e frameworks de programação e servir conteúdo dinâmico e estático.

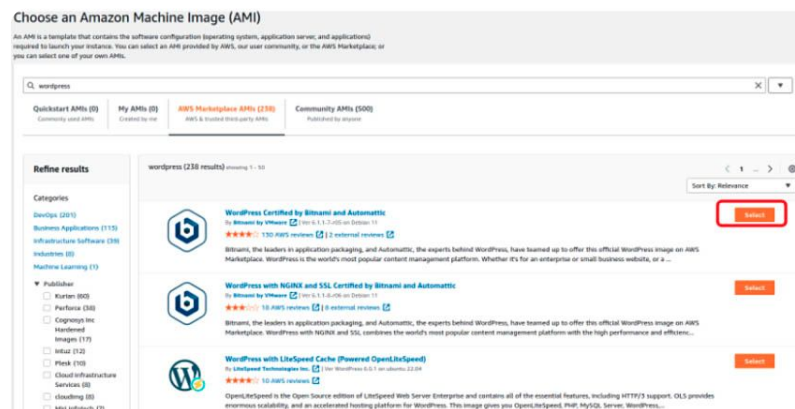
Utilizaremos como exemplo o Wordpress, um software web de blogging, mas que também é amplamente utilizado para sites de diversos propósitos. Como ponto de partida pode ser utilizada uma AMI pronta, com toda plataforma necessária para permitir que o Wordpress rode na instância, concentrando todo o trabalho de provisionamento na infraestrutura AWS.

Faça a busca pela palavra “wordpress” em “Application and OS Images” e uma lista de AMI de Wordpress aparecerá para a sua seleção.



Indicação de palavra-chave para busca de AMI de WordPress.

As AMIs são ordenadas por padrão e por relevância, de forma que as mais recomendadas aparecerão primeiro. Você pode escolher a imagem fornecida pela Bitnami by VMware, que costuma ser validada pela própria AWS.



Lista de AMIs com destaque para a da Bitnami.

Uma vez selecionada a AMI, podemos seguir o processo tradicional de provisionamento de uma nova EC2, escolhendo tamanho e um par de chaves. Para esse tipo de aplicação, é recomendada, em ambiente de testes, ao menos uma t3.small de tipo/tamanho.

▼ Application and OS Images (Amazon Machine Image) [Info](#)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below

AMI from catalog

Recents

My AMIs

Quick Start

Amazon Machine Image (AMI)

bitnami-wordpress-6.1.1-7-r05-linux-debian-11-x86_64-hvm-ebs-nami-7d426cb7-9522-4dd7-a56b-55dd8cc1c8d0
ami-082209a936ac9c95a

Verified provider

[Browse more AMIs](#)
Including AMIs from AWS, Marketplace and the Community

Catalog	Published	Architecture	Virtualization	Root device type	ENA Enabled
AWS	2022-12-	x86_64	hvm		Yes
Marketplace	02T22:05:29.0			ebs	
AMIs	00Z				

If you have an existing license entitlement to use this software, then you can launch this software without creating a new subscription. If you do not have an existing entitlement, then by launching this software, you will be subscribed to this software and agree that your use of this software is subject to the pricing terms and the seller's [End User License Agreement](#)

AMI da Bitnami com indicação de fornecedor verificado pela AWS, que confere garantia sobre a solução fornecida.

Na seção de configurações de rede e segurança da instância, encontraremos um security group personalizado e com configurações específicas para a AMI que selecionamos. Algumas AMIs facilitam o processo de provisionamento, adequando configurações necessárias para o pleno funcionamento do software que está embarcado. No caso do Wordpress, as portas 80 e 443 (http e https) são liberadas para o pleno acesso via web da aplicação.

▼ Network settings [Info](#)

Edit

Network [Info](#)

vpc-e6598d9b

Subnet [Info](#)

No preference (Default subnet in any availability zone)

Auto-assign public IP [Info](#)

Enable

Firewall (security groups) [Info](#)

A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.

Create security group

Select existing security group

We'll create a new security group called 'WordPress Certified by Bitnami and Automattic-6.1.1-7-r05 on Debian 11-AutogenByAWSMP--2' with the following rules:

☒ Allow SSH traffic from

Recommended rule from AMI

Anywhere
0.0.0.0/0

☒ Allow HTTPS traffic from the Internet

To set up an endpoint, for example when creating a web server

☒ Allow HTTP traffic from the Internet

To set up an endpoint, for example when creating a web server

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

Detalhes do security group na seção de configurações de rede.

A instância pode ser provisionada sem mais ajustes e, após inicializada, localizar as informações de acesso. Busque pelo endereço de DNS público e o utilize no seu navegador.

Instances (1/1) Info

Connect Instance state Actions Launch instances

Find Instance by attribute or tag (case-sensitive)

Instance ID = i-0bc4a6f988f4ad75f Clear filters

Name	Instance ID	Instance state
-	i-0bc4a6f988f4ad75f	Running

Instance: i-0bc4a6f988f4ad75f

Details Security Networking Storage Status checks Monitoring Tags

▼ Instance summary Info

Instance ID i-0bc4a6f988f4ad75f	Public IPv4 address 54.147.217.91 open address	Private IPv4 addresses 172.31.27.242
IPv6 address -	Instance state Running	Public IPv4 DNS ec2-54-147-217-91.compute-1.amazonaws.com open address

Indicação do DNS público nos detalhes da instância.

Você deve visualizar o conteúdo da imagem a seguir, indicando que carregamos um blog do Wordpress:

Mindblown: a blog about philosophy.

Hello world!

Welcome to WordPress. This is your first post. Edit or delete it, then start writing!
December 6, 2022

Got any book recommendations?

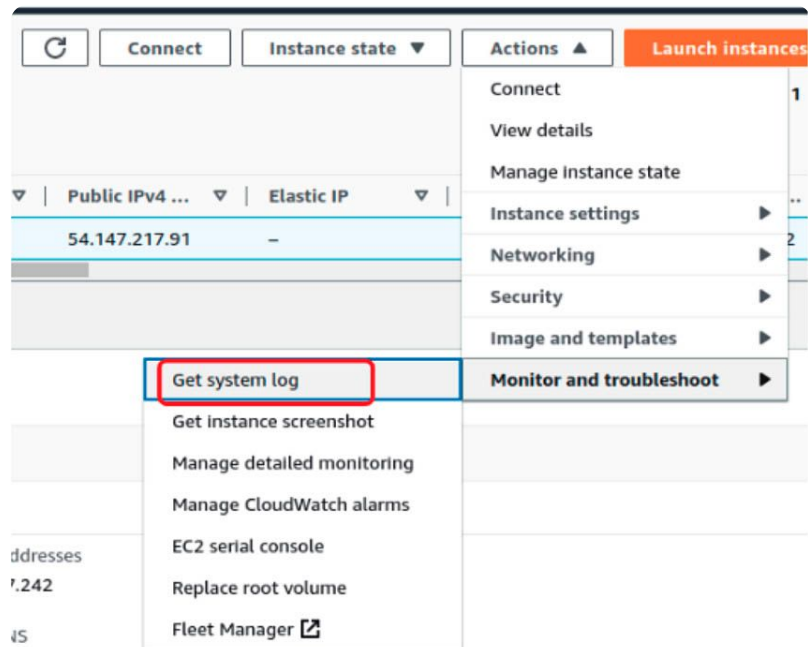
Get In Touch

User's blog

Proudly powered by WordPress

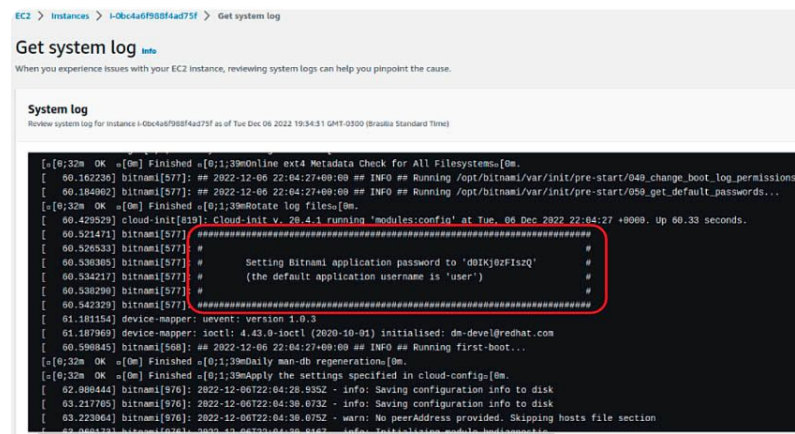
Tela que indica que o o blog foi instalado com sucesso.

O trabalho de provisionamento da nossa aplicação dinâmica está completo, mas podemos avançar um pouco mais e fazer um novo post na plataforma. Para isso, precisamos recuperar o usuário e senha da parte administrativa do Wordpress. A documentação da AMI (que pode ser encontrada no Marketplace da AWS ou uma busca na internet) indica que as credenciais de acesso podem ser encontradas no log de boot da instância. O acesso ao log pode ser feito selecionando a instância, clicando no botão "action", e será exibido o menu com a seguinte opção:



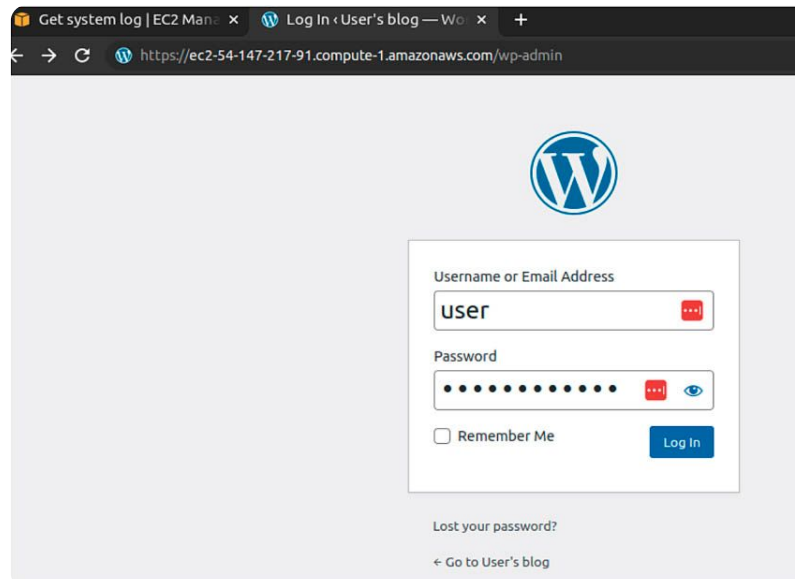
Indicação do caminho para obter o log de boot da instância.

Role a tela de log até encontrar o bloco onde estão as credenciais de acesso ao admin do Wordpress:



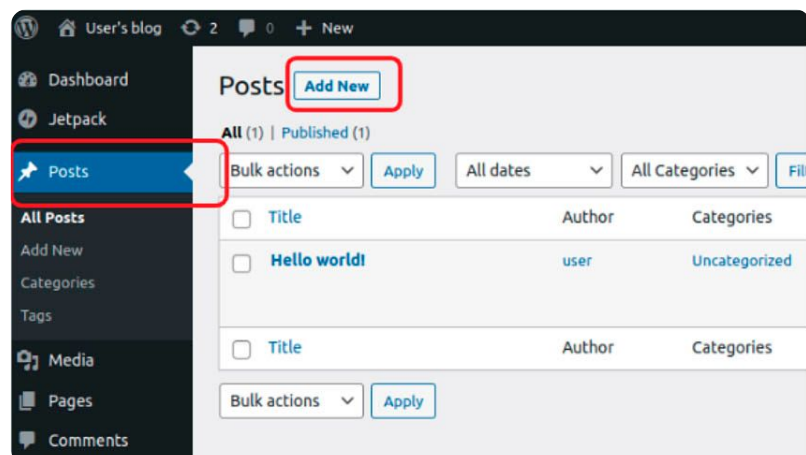
Bloco com as credenciais de acesso.

Com as credenciais em mãos, o endereço do admin do Wordpress é o path `/wp-admin` do endereço que utilizamos para acessar a plataforma:



Tela de login do Wordpress.

Após feito o login, caso não conheça a plataforma, faça um novo post, indo na seção “posts” e clicando no botão “add new”. Um editor aparecerá e você pode editar da forma que for mais conveniente, salvando em seguida no botão “publish”.



Interface administrativa do Wordpress permite fazer novas publicações na plataforma.

O conteúdo deverá ser refletido no endereço principal do site, indicando que houve modificação e consulta da plataforma para exibição dos dados salvos.



Verificando o aprendizado

Questão 1

Quais os tipos de site/página web o Amazon S3 possui capacidade para fazer hospedagem e executar como um servidor web?

A

Dinâmicos e estáticos.

B

Dinâmicos.

C

Estáticos.

D

PHP.

E

ASP.NET.



A alternativa C está correta.

Os buckets no Amazon S3 não possuem capacidade de rodar sites dinâmicos de nenhum tipo de linguagem e podem funcionar apenas como hospedagem de sites estáticos, em HTML.

Questão 2

Qual a vantagem em utilizarmos AMIs de fornecedores verificados pela AWS?

A

Não possuem custo de utilização.

B

Oferecem suporte gratuito.

C

Possuem garantia de origem confiável.

D

Sempre estão atualizadas com o sistema operacional mais recente.

E

Indicam que foram criadas pela própria AWS.



A alternativa C está correta.

A verificação de fornecedores de AMIs é uma garantia que a AWS dá a parceiros que se qualificam, conferindo confiabilidade de origem.

Considerações finais

A Amazon Web Services pode ser vista como uma grande caixa de blocos de encaixe, cabendo ao profissional escolher as peças ideais para cada projeto. É comum encontrarmos arquiteturas de referência, desenhadas pela própria AWS, indicando quais recursos utilizar para vários casos de uso, porém, é importante adaptá-los para os requisitos de cada projeto da vida real e suas eventuais restrições, adequando assim a expectativa de nível de serviço da solução.

Não existe apenas um caminho correto de entregar os benefícios da computação em nuvem. A variedade de serviços e ferramentas permite optar por caminhos distintos e que oferecerão resultados parecidos.

Nesse mercado, é importante que você se mantenha atualizado, pois novos serviços e produtos são lançados com muita frequência, entregando mais benefícios por um custo menor. Ao final do ano, acontece uma série de lançamentos que podem modificar a forma de construir infraestruturas, trazendo soluções que facilitam ou diminuem o custo do projeto. Você viu aqui como construir as bases, mas esse conhecimento é continuado e acompanhar o que a AWS lança de novidades vai te ajudar a fazer mais por menos.

Podcast

Neste podcast trataremos das bases da computação em nuvem na AWS, abordaremos computação, armazenamento, rede e os principais cenários de uso na vida real, direcionando sempre para as boas práticas e como escolher a solução correta para cada caso de uso.



Conteúdo interativo

Acesse a versão digital para ouvir o áudio.

Explore +

Como estudo complementar para avançar ainda mais na capacidade de execução de infraestruturas de nuvem, que é parte essencial de arquiteturas modernas de aplicação, pesquise os bancos de dados como serviço, em especial o **Amazon RDS (Relational Database Service)** e o **Amazon Aurora**, duas soluções que abstraem, facilitam e agilizam a gestão de bancos relacionais.

Outro tema essencial no contexto de infraestruturas na AWS são serviços de CI/CD (continuous integration and continuous delivery) que permitem automatizar e disponibilizar, de forma contínua e frequente, aplicativos na infraestrutura. Nesse contexto, pesquise o **AWS CodeDeploy** como solução, especialmente em rotinas onde há interação com times de desenvolvimento e existem tarefas de provisionamento de código/aplicações desenvolvidas por esse time.

Referências

AMAZON WEB SERVICES. **Overview of Amazon Web Services** – AWS Whitepapers. Consultado na internet em: 28 nov. 2022.

AMAZON WEB SERVICES. **Amazon Elastic Compute Cloud** – User Guide for Linux Instances. Consultado na internet em: 28 nov. 2022.

AMAZON WEB SERVICES. **Amazon Elastic Compute Cloud** – User Guide for Windows Instances. Consultado na internet em: 28 nov. de 2022.

AMAZON WEB SERVICES. **Amazon Simple Storage Service** – User Guide. Consultado na internet em: 28 nov. de 2022.

AMAZON WEB SERVICES. **Amazon Virtual Private Cloud** – User Guide. Consultado na internet em: 28 nov. de 2022.