

Compute

EC2

O Amazon Elastic Compute Cloud (Amazon EC2) é um serviço web com o qual você pode executar “instâncias” de servidor virtual na nuvem.

As instâncias do Amazon EC2 podem executar os sistemas operacionais Windows, Linux ou MacOS.

A interface de serviço web simples do EC2 permite obter e configurar a capacidade com o mínimo de atrito.

O EC2 foi projetado para tornar a computação em nuvem em escala da Web mais fácil para os desenvolvedores.

O Amazon EC2 muda a economia da computação permitindo que você pague apenas pela capacidade que você realmente usa.

O Amazon EC2 fornece aos desenvolvedores as ferramentas para criar aplicativos resilientes a falhas e isolá-los de cenários de falhas comuns.

Os benefícios do EC2 incluem:

- Computação Elastic Web-Scale – você pode aumentar ou diminuir a capacidade em minutos, não em horas, e comissionar de uma a milhares de instâncias simultaneamente.
- Totalmente controlado – Você tem controle total, incluindo acesso root a cada instância e pode parar e iniciar instâncias sem perder dados e usando APIs de serviço da web.
- Serviços flexíveis de hospedagem em nuvem – você pode escolher entre vários tipos de instâncias, sistemas operacionais e pacotes de software, bem como instâncias com configurações variadas de memória, CPU e armazenamento.
- Integrado – O EC2 é integrado à maioria dos serviços da AWS, como S3, RDS e VPC, para fornecer uma solução completa e segura.
- Confiável – O EC2 oferece um ambiente altamente confiável em que instâncias de substituição podem ser comissionadas de forma rápida e previsível com SLAs de 99,99% para cada região.
- Seguro – O EC2 funciona em conjunto com o VPC para fornecer um local seguro com um intervalo de endereços IP especificado e oferece grupos de segurança, ACLs de rede e recursos de VPN IPSec.
- Barato – a Amazon repassa os benefícios financeiros da escala cobrando taxas muito baixas e com base na capacidade consumida.

Uma Amazon Machine Image (AMI) é um tipo especial de dispositivo virtual usado para criar uma máquina virtual no Amazon Elastic Compute Cloud (“EC2”).

Uma AMI inclui o seguinte:

- Um ou mais snapshots do EBS ou, para AMIs com armazenamento de instância, um modelo para o volume raiz da instância (por exemplo, um sistema operacional, um servidor de aplicativos e aplicativos).
- Permissões de execução que controlam quais contas da AWS podem usar a AMI para iniciar instâncias.
- Um mapeamento de dispositivo de bloco que especifica os volumes a serem anexados à instância quando ela for iniciada

AMIs vêm em três categorias principais:

- Community AMIs – grátis para usar, geralmente você apenas seleciona o sistema operacional que deseja.
- AMIs do AWS Marketplace – pagas para usar, geralmente vêm empacotadas com software licenciado adicional.
- Minhas AMIs – AMIs que você mesmo cria.

Metadados e dados do usuário:

- Dados do usuário são dados fornecidos pelo usuário na inicialização da instância na forma de um script.
- Os metadados da instância são dados sobre sua instância que você pode usar para configurar ou gerenciar a instância em execução.
- Os dados do usuário são limitados a 16 KB.
- Os dados e metadados do usuário não são criptografados.
- Os metadados da instância estão disponíveis em <http://169.254.169.254/latest/meta-data> (<http://169.254.169.254/latest/meta-data>) .

A ferramenta Instance Metadata Query permite consultar os metadados da instância sem precisar digitar o URI completo ou os nomes de categoria.

Preços

On-demand:

- Bom para usuários que desejam o baixo custo e a flexibilidade do EC2 sem nenhum pagamento antecipado ou compromisso de longo prazo.
- Aplicativos com cargas de trabalho de curto prazo, pontiagudas ou imprevisíveis que não podem ser interrompidas.
- Aplicativos sendo desenvolvidos ou testados no EC2 pela primeira vez.

Reserved:

- Aplicações com estado estacionário ou uso previsível.
- Aplicativos que exigem capacidade reservada.
- Os usuários podem fazer pagamentos adiantados para reduzir ainda mais seus custos totais de computação.
- As instâncias reservadas padrão (RIs) oferecem até 75% de desconto no preço sob demanda.
- RIs conversíveis oferecem até 54% de desconto no preço sob demanda – oferece a capacidade de alterar os atributos da RI desde que a troca resulte na criação de RIs de valor igual ou superior.
- As RIs agendadas estão disponíveis para serem lançadas dentro da janela de tempo que você reservar. Essa opção permite que você combine sua reserva de capacidade com uma programação recorrente previsível que requer apenas uma fração de um dia, uma semana ou um mês.

Spot:

- Aplicativos que possuem horários de início e término flexíveis.
- Aplicativos que só são viáveis a preços de computação muito baixos.
- Usuários com necessidade urgente de uma grande quantidade de capacidade de computação adicional.
- Se a Amazon encerrar suas instâncias, você não paga, se encerrar, você paga por hora.

Dedicated hosts:

- Servidores físicos dedicados apenas para seu uso.
- Você então tem controle sobre quais instâncias são implantadas nesse host.
- Disponível como Sob Demanda ou com Reserva de Host Dedicado.
- Útil se você tiver licenças de software vinculadas ao servidor que usam métricas como por núcleo, por soquete ou por VM.
- Cada host dedicado pode executar apenas um tamanho e tipo de instância do EC2.
- Bom para conformidade regulatória ou requisitos de licenciamento.
- Desempenho previsível.
- Isolamento completo.
- Opção mais cara.
- O faturamento é por host.

Dedicated instances:

- Instâncias virtualizadas em hardware apenas para você.
- Também usa servidores EC2 fisicamente dedicados.
- Não fornece visibilidade e controles adicionais de hosts dedicados (por exemplo, como as instâncias são colocadas em um servidor).
- O faturamento é por instância.
- Pode compartilhar hardware com outras instâncias não dedicadas na mesma conta.
- Disponível como instâncias sob demanda, reservadas e instâncias spot.
- Custo adicional de US\$ 2 por hora por região.

Savings Plans:

- Savings Plans é um modelo de definição de preço flexível que oferece economia de até 72% no uso de computação da AWS.
- Esse modelo de definição de preço oferece preços mais baixos no uso de instâncias do Amazon EC2, independentemente da família de instâncias, tamanho, SO, localização ou região da AWS.
- Também se aplica ao uso do AWS Fargate e do AWS Lambda.

Tipos de instância

O Amazon EC2 oferece uma ampla seleção de tipos de instância otimizados para atender a diferentes casos de uso.

Os tipos de instância compreendem várias combinações de CPU, memória, armazenamento e capacidade de rede e oferecem a flexibilidade de escolher a combinação apropriada de recursos para seus aplicativos.

Cada tipo de instância inclui um ou mais tamanhos de instância, permitindo que você dimensione seus recursos de acordo com os requisitos de sua carga de trabalho de destino.

A tabela abaixo fornece uma visão geral dos diferentes tipos de instância do EC2:

| Category | Families | Purpose/Design |
|------------------------------|--|---|
| General Purpose | A1, T3, T3a, T2, M5, M5a, M4 | General purpose instances provide a balance of compute, memory and networking resources, and can be used for a variety of diverse workloads |
| Compute Optimized | C5, C5n, C4 | Compute Optimized instances are ideal for compute bound applications that benefit from high performance processors |
| Memory Optimized | R5, R5a, R4, X1e, X1, High Memory, z1d | Memory optimized instances are designed to deliver fast performance for workloads that process large data sets in memory |
| Accelerated Computing | P3, P2, G4, G3, F1 | Accelerated computing instances use hardware accelerators, or co-processors, to perform functions, such as floating-point number calculations, graphics processing, or data pattern matching |
| Storage Optimized | I3, I3en, D2, H1 | This instance family provides Non-Volatile Memory Express (NVMe) SSD-backed instance storage optimized for low latency, very high random I/O performance, high sequential read throughput and provide high IOPS at a low cost |

Amazon Elastic Container Service (ECS)

O Amazon Elastic Container Service (ECS) é outro produto na categoria AWS Compute. Ele fornece um serviço de gerenciamento de contêineres altamente escalável e de alto desempenho que oferece suporte a contêineres do Docker e permite que você execute aplicativos facilmente em um cluster gerenciado de instâncias do Amazon EC2.

O Amazon ECS elimina a necessidade de instalar, operar e dimensionar sua própria infraestrutura de gerenciamento de cluster.

Usando chamadas de API, você pode iniciar e interromper aplicativos habilitados para contêiner, consultar o estado completo dos clusters e acessar muitos recursos familiares, como grupos de segurança, Elastic Load Balancing, volumes EBS e funções do IAM.

O Amazon ECS pode ser usado para agendar a colocação de contêineres em clusters com base nas necessidades de recursos e nos requisitos de disponibilidade.

Um tipo de inicialização do Amazon ECS determina o tipo de infraestrutura na qual suas tarefas e serviços estão hospedados.

Existem dois tipos de lançamento e a tabela abaixo descreve algumas das diferenças entre os dois tipos de lançamento:

| Amazon EC2 | Amazon Fargate |
|--|---|
| You explicitly provision EC2 instances | The control plane asks for resources and Fargate automatically provisions |
| You're responsible for upgrading, patching, care of EC2 pool | Fargate provisions compute as needed |
| You must handle cluster optimization | Fargate handles cluster optimization |
| More granular control over infrastructure | Limited control, as infrastructure is automated |

O Elastic Container Registry (ECR) é um serviço de registro gerenciado do AWS Docker para armazenar, gerenciar e implantar imagens do Docker.

Não há cobrança adicional para o Amazon ECS. Você paga pelos recursos da AWS (por exemplo, instâncias do EC2 ou volumes do EBS) que você cria para armazenar e executar seu aplicativo.

O Amazon ECR é integrado ao Amazon EC2 Container Service (ECS).

Com o Amazon ECR, não há taxas ou compromissos iniciais. Você paga apenas pela quantidade de dados que armazena em seus repositórios e pelos dados transferidos para a Internet.

AWS Lambda

O AWS Lambda é uma tecnologia de computação sem servidor que permite executar código sem provisionar ou gerenciar servidores.

O AWS Lambda executa código somente quando necessário e dimensiona automaticamente.

Você paga apenas pelo tempo de computação que consome (você não paga nada quando seu código não está em execução).

Acionado por eventos

Benefícios do AWS Lambda:

- Nenhum servidor para gerenciar.
- Escalonamento contínuo.
- Medição de subsegundos.
- Integra-se com quase todos os outros serviços da AWS.

Casos de uso primários para AWS Lambda:

- Processamento de dados.
- Processamento de arquivos em tempo real.
- Processamento de fluxo em tempo real.
- Crie back-ends sem servidor para solicitações de API da Web, dispositivos móveis, IOT e de terceiros.

Instâncias do Amazon LightSail

O Amazon Lightsail é um dos serviços mais recentes do pacote de produtos AWS Compute. O Amazon Lightsail é ótimo para usuários que não têm profundo conhecimento técnico da AWS, pois facilita muito o provisionamento de serviços de computação.

O Amazon Lightsail oferece aos desenvolvedores capacidade e recursos de computação, armazenamento e rede para implantar e gerenciar sites, aplicativos da web e bancos de dados na nuvem.

O Amazon Lightsail inclui tudo o que você precisa para iniciar seu projeto rapidamente: uma máquina virtual, armazenamento baseado em SSD, transferência de dados, gerenciamento de DNS e um IP estático.

O Amazon Lightsail fornece servidores privados virtuais (instâncias) pré-configurados que incluem tudo o que é necessário para implantar e aplicar ou criar um banco de dados.

A infraestrutura subjacente e o sistema operacional são gerenciados pelo Amazon Lightsail.

Mais adequado para projetos que exigem algumas dezenas de instâncias ou menos.

Fornecer uma interface de gerenciamento simples.

Bom para blogs, sites, aplicativos da web, e-commerce etc.

Pode implantar balanceadores de carga e anexar armazenamento em bloco.

API pública.

Limitado a 20 instâncias do Amazon Lightsail, 5 IPs estáticos, 3 zonas DNS, armazenamento em bloco de 20 TB, 40 bancos de dados e 5 load balancers por conta.

Até 20 certificados por ano civil.

Podem se conectar uns aos outros e a outros recursos da AWS por meio de rede pública de Internet e rede privada (emparelhamento de VPC).

Os modelos de aplicativos incluem WordPress, WordPress Multisite, Drupal, Joomla!, Magento, Redmine, LAMP, Nginx (LEMP), MEAN, Node.js e muito mais.

Atualmente, o Amazon Lightsail oferece suporte a 6 distribuições Linux ou do tipo Unix: Amazon Linux, CentOS, Debian, FreeBSD, OpenSUSE e Ubuntu, bem como 2 versões do Windows Server: 2012 R2 e 2016.

Bancos de dados do Amazon LightSail

Os bancos de dados do Amazon Lightsail são instâncias dedicadas à execução de bancos de dados.

Um banco de dados do Amazon Lightsail pode conter vários bancos de dados criados pelo usuário e você pode acessá-lo usando as mesmas ferramentas e aplicativos que usa com um banco de dados autônomo.

Os bancos de dados gerenciados do Amazon Lightsail fornecem uma maneira fácil e de baixa manutenção de armazenar seus dados na nuvem.

O Amazon Lightsail gerencia uma série de atividades de manutenção e segurança para seu banco de dados e sua infraestrutura subjacente.

O Amazon Lightsail faz backup automático de seu banco de dados e permite a restauração pontual dos últimos 7 dias usando a ferramenta de restauração de banco de dados.

Os bancos de dados do Amazon Lightsail são compatíveis com as versões principais mais recentes do MySQL. Atualmente, essas versões são 5.6, 5.7 e 8.0 para MySQL.

Os bancos de dados do Amazon Lightsail estão disponíveis nos planos Standard e High Availability.

Os planos de alta disponibilidade adicionam redundância e durabilidade ao seu banco de dados, criando automaticamente um banco de dados em espera em uma zona de disponibilidade separada.

O Amazon Lightsail é muito acessível.

Os planos do Amazon Lightsail são cobrados por hora sob demanda, portanto, você paga apenas pelo que usa.

Para cada plano do Amazon Lightsail que você usa, cobramos o preço fixo por hora, até o custo mensal máximo do plano.

AWS Elastic Beanstalk

O AWS Elastic Beanstalk é a maneira mais rápida e simples de colocar aplicativos da web em funcionamento na AWS.

Os desenvolvedores simplesmente carregam o código do aplicativo e o serviço lida automaticamente com todos os detalhes, como provisionamento de recursos, balanceamento de carga, dimensionamento automático e monitoramento.

O Elastic Beanstalk é ideal se você tiver um aplicativo Web PHP, Java, Python, Ruby, Node.js, .NET, Go ou Docker.

O Elastic Beanstalk usa os principais serviços da AWS, como Amazon EC2, Amazon Elastic Container Service (Amazon ECS), Auto Scaling e Elastic Load Balancing para oferecer suporte fácil a aplicativos que precisam ser dimensionados para atender a milhões de usuários.

AWS Batch

O AWS Batch permite que desenvolvedores, cientistas e engenheiros executem com facilidade e eficiência centenas de milhares de trabalhos de computação em lote na AWS.

O AWS Batch provisiona dinamicamente a quantidade e o tipo ideais de recursos de computação (por exemplo, instâncias otimizadas para CPU ou memória) com base no volume e nos requisitos de recursos específicos dos trabalhos em lote enviados.