



Whitepaper da AWS

Instâncias reservadas do Amazon EC2 e outros modelos de reserva da AWS



Instâncias reservadas do Amazon EC2 e outros modelos de reserva da AWS: Whitepaper da AWS

Copyright © Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

As marcas comerciais e o visual comercial da Amazon não podem ser usados em conexão com nenhum produto ou serviço que não seja da Amazon, nem de qualquer maneira que possa causar confusão entre os clientes ou que deprecie ou desacredite a Amazon. Todas as outras marcas comerciais que não pertencem à Amazon pertencem a seus respectivos proprietários, que podem ou não ser afiliados, conectados ou patrocinados pela Amazon.

Table of Contents

Resumo	1
Resumo	1
Introdução	2
Instâncias reservadas do Amazon EC2	3
Opções de pagamento das instâncias reservadas	3
Classes de oferta padrão versus conversíveis	4
Instâncias reservadas regionais e zonais	5
Diferenças entre instâncias reservadas regionais e zonais	5
Limites para a flexibilidade de tamanho da instância	6
Maximizar a utilização com flexibilidade de tamanho em instâncias reservadas regionais	7
Fator de normalização para instâncias do EC2 dedicadas	9
Fator de normalização para instâncias bare metal	9
Savings Plans	12
Modelos de reserva para outros serviços da AWS	14
Instâncias de banco de dados reservadas do Amazon RDS	14
Nós reservados do Amazon ElastiCache	14
Instâncias reservadas do Amazon OpenSearch Service	15
Nós reservados do Amazon Redshift	15
Reservas do Amazon DynamoDB	16
Cobrança de instâncias reservadas	17
Faturamento do uso	17
Faturamento consolidado	19
Instâncias reservadas: reservas de capacidade	19
Combinação de tarifas	20
Como os descontos são aplicados	20
Maximizar o valor das reservas	21
Medir o sucesso	21
Maximizar os descontos padronizando o tipo de instância	21
Técnicas de gerenciamento de reservas	22
Marketplace de instâncias reservadas	23
AWS Cost Explorer	23
AWS Cost and Usage Report	24
Instâncias reservadas em seu relatório de custo e uso	24
AWS Trusted Advisor	25

Conclusão	27
Colaboradores	28
Revisões do documento	29
Avisos	30

Instâncias reservadas do Amazon EC2 e outros modelos de reserva da AWS

Data de publicação: 29 de março de 2021 ([Revisões do documento](#))

Resumo

Este documento faz parte de uma série de whitepapers da AWS elaborados para apoiar sua jornada para a nuvem e discute as instâncias reservadas do Amazon EC2 e os modelos de reserva para outros serviços da AWS. Seu objetivo é capacitá-lo a maximizar o valor de seus investimentos, melhorar a precisão das previsões e a previsibilidade de custos, criar uma cultura de propriedade e transparência de custos e medir continuamente seu status de otimização.

Introdução

A nuvem é adequada para workloads variáveis e implantação rápida, mas muitas workloads baseadas em nuvem seguem um padrão mais previsível. Para essas aplicações, sua organização pode obter economias de custo significativas usando [instâncias reservadas do Amazon Elastic Compute Cloud \(Amazon EC2\)](#). As instâncias reservadas do Amazon EC2 permitem que sua organização se comprometa com os parâmetros de uso no momento da compra para obter uma taxa por hora mais baixa. Os modelos de reserva também estão disponíveis para o [Amazon Relational Database Service \(Amazon RDS\)](#), o [Amazon ElastiCache](#), [Amazon OpenSearch Service \(OpenSearch Service\)](#), o [Amazon Redshift](#) e o [Amazon DynamoDB](#). Este whitepaper aborda as instâncias reservadas do Amazon EC2 e os modelos de reserva para esses outros serviços da AWS.

Instâncias reservadas do Amazon EC2

Ao comprar instâncias reservadas, você assume um compromisso de um ou três anos e recebe um desconto de faturamento de até 72% em troca. Quando usadas para as workloads apropriadas, as instâncias reservadas podem gerar uma grande economia.

Uma instância reservada não é uma instância exclusiva de sua organização. Trata-se de um desconto de faturamento aplicado ao uso de instâncias sob demanda em sua conta. Essas instâncias sob demanda devem corresponder a determinados atributos das instâncias reservadas que você adquiriu para se beneficiar do desconto de faturamento. Você paga por todo o período de vigência de uma instância reservada, independentemente do uso real, para que sua economia de custos esteja intimamente ligada ao uso. Portanto, é importante planejar e monitorar o uso para aproveitar ao máximo seu investimento.

Ao adquirir uma instância reservada em uma zona de disponibilidade específica, você recebe uma reserva de capacidade. Isso aumenta a probabilidade de que a capacidade computacional de que você precisa esteja disponível em uma zona de disponibilidade específica quando você precisar dela. Uma instância reservada comprada para uma região da AWS não fornece reserva de capacidade.

Opções de pagamento das instâncias reservadas

Você pode comprar instâncias reservadas por meio do AWS Management Console. As opções de pagamento a seguir estão disponíveis para a maioria das instâncias reservadas:

- **Sem pagamento adiantado:** não é necessário pagamento adiantado. É cobrado de você a tarifa por hora com desconto para cada hora do período de vigência, independentemente de a instância reservada estar sendo usada ou não. Instâncias reservadas sem pagamento adiantado têm como base uma obrigação contratual de pagar mensalmente pelo período de vigência inteiro da reserva. É necessário ter um histórico de faturamento de sucesso para poder comprar instâncias reservadas sem pagamento adiantado.
- **Pagamento adiantado parcial:** parte do custo deve ser paga adiantada, e as horas restantes do período de vigência são faturadas a uma taxa por hora com desconto, independentemente de a instância reservada estar ou não sendo usada.
- **Adiantamento integral:** o pagamento integral é feito no início do período de vigência, sem outros custos ou cobranças por hora incorridos pelo restante do período, independentemente das horas usadas.

As instâncias reservadas com um pagamento adiantado mais alto oferecem descontos maiores. Você também pode encontrar instâncias reservadas oferecidas por vendedores terceiros a preços menores e períodos de vigência mais curtos no [Marketplace de instâncias reservadas](#).

À medida que você compra mais instâncias reservadas, vão sendo aplicados descontos por volume que permitem uma economia ainda maior.

Para obter mais informações, consulte [Preço de instâncias reservadas do Amazon EC2](#).

Classes de oferta padrão versus conversíveis

Ao adquirir uma instância reservada, você pode escolher entre uma classe de oferta padrão ou conversível.

Tabela 1: Comparação entre instâncias reservadas padrão e conversíveis

Instância padrão reservada	Instância reservada conversível
Período de vigência de um a três anos	Período de vigência de um a três anos
Permite modificar a zona de disponibilidade, o escopo, o tipo de rede e o tamanho da instância (no mesmo tipo de instância) da sua instância reservada. Para obter mais informações, consulte Modificar instâncias reservadas .	Permite que você troque uma ou mais instâncias reservadas conversíveis por outra do mesmo tipo com uma configuração diferente, incluindo família de instâncias, sistema operacional e localização. Não há limites para quantas vezes você pode realizar uma troca, desde que a instância reservada conversível de destino tenha valor igual ou superior ao da instância que você está trocando. Para obter mais informações, consulte Trocar instâncias reservadas conversíveis .
Pode ser vendida no Marketplace de instâncias reservadas.	Não pode ser vendida no Marketplace de instâncias reservadas.

As instâncias padrão reservadas geralmente oferecem os maiores descontos. As instâncias padrão reservadas de um ano oferecem um desconto semelhante ao das instâncias reservadas conversíveis de três anos.

Se você quiser comprar reservas de capacidade, consulte [Reservas de capacidade sob demanda](#).

Instâncias reservadas conversíveis são úteis nos seguintes casos:

- Compra de instâncias reservadas na conta pagante em vez de em uma subconta. Você pode modificar mais facilmente as instâncias reservadas conversíveis para atender às necessidades de mudança em toda a organização.
- É provável que as workloads mudem. Nesse caso, uma instância reservada conversível permite que você se adapte à medida que as necessidades mudam e, ao mesmo tempo, obtenha descontos e reservas de capacidade.
- Você deseja se proteger contra possíveis quedas de preços futuras.
- Você não pode ou não quer pedir às equipes que façam planejamento ou previsão de capacidade.
- Você espera que o uso de computação permaneça no valor comprometido durante o período de compromisso.

Instâncias reservadas regionais e zonais

Ao comprar uma instância reservada, é preciso determinar o escopo dela. O escopo pode ser regional ou zonal.

- Regional: ao adquirir uma instância reservada para uma região, isso é chamado de instância reservada regional.
- Zonal: ao adquirir uma instância reservada para uma zona de disponibilidade específica, ela é chamada de instância reservada zonal.

Diferenças entre instâncias reservadas regionais e zonais

A tabela a seguir destaca algumas diferenças importantes entre as instâncias reservadas regionais e as instâncias reservadas zonais:

Tabela 2: Comparação entre instâncias reservadas regionais e zonais

	Instâncias reservadas regionais	Instâncias reservadas zonais
Flexibilidade da zona de disponibilidade	O desconto da instância reservada se aplica ao uso da instância em qualquer zona de disponibilidade na região especificada.	Sem flexibilidade da zona de disponibilidade: o desconto da instância reservada se aplica ao uso da instância somente na zona de disponibilidade especificada.
Reserva de capacidade	Sem reserva de capacidade: uma instância reservada regional não fornece uma reserva de capacidade.	Uma instância reservada zonal fornece uma reserva de capacidade na zona de disponibilidade especificada.
Flexibilidade de tamanho da instância	O desconto da instância reservada se aplica ao uso da instância na família de instâncias, independentemente do tamanho. Compatível somente com instâncias reservadas Linux/Unix da Amazon com locação padrão. Para obter mais informações, consulte Flexibilidade do tamanho da instância determinada pelo fator de normalização .	Sem flexibilidade de tamanho da instância: o desconto da instância reservada se aplica ao uso da instância somente para o tamanho e o tipo de instância especificados.

Limites para a flexibilidade de tamanho da instância

A flexibilidade de tamanho da instância não se aplica às seguintes instâncias reservadas:

- Instâncias reservadas que são compradas para uma zona de disponibilidade específica (instâncias reservadas zonais).
- Instâncias reservadas com locação dedicada.

- Instâncias reservadas para Windows Server, Windows Server com SQL Standard, Windows Server com SQL Server Enterprise, Windows Server com SQL Server Web, RHEL e SUSE Linux Enterprise Server.
- Instâncias reservadas para instâncias G4.

Maximizar a utilização com flexibilidade de tamanho em instâncias reservadas regionais

Para obter maior flexibilidade, todas as instâncias reservadas regionais Linux com locação compartilhada se aplicam a todos os tamanhos de instâncias em uma família de instâncias e uma região da AWS, mesmo que você as esteja usando em várias contas por meio do [Faturamento consolidado](#). Os únicos atributos que devem ser correspondidos são o tipo de instância (por exemplo, m4), a locação (deve ser padrão) e a plataforma (deve ser Linux). Todas as instâncias reservadas novas e existentes são dimensionadas de acordo com um fator de normalização com base no tamanho da instância, da forma a seguir.

Tabela 3: Tamanhos de instâncias reservadas regionais e fatores de normalização

Tamanho da instância	Fator de normalização
nano	0,25
micro	0,5
pequeno	1
médio	2
grande	4
xlarge	8
2xlarge	16
4xlarge	32
8xlarge	64

Tamanho da instância	Fator de normalização
9xlarge	72
10xlarge	80
12xlarge	96
16xlarge	128
24xlarge	192
32xlarge	256

Por exemplo, se você tiver uma instância reservada para uma c4.8xlarge, ela se aplicará a qualquer uso de uma instância Linux c4 com localização compartilhada na região da AWS, como:

- Uma instância c4.8xlarge
- Duas instâncias c4.4xlarge
- Quatro instâncias c4.2xlarge
- Dezesesseis instâncias c4.large

Também inclui combinações de instâncias, por exemplo, uma instância t2.medium tem um fator de normalização de 2. Se você tiver adquirido uma instância reservada Linux/Unix da Amazon na localização padrão t2.medium na região Leste dos EUA (Norte da Virgínia) e tiver duas instâncias t2.small em execução na sua conta nessa região, o benefício de faturamento será aplicado integralmente às duas instâncias.

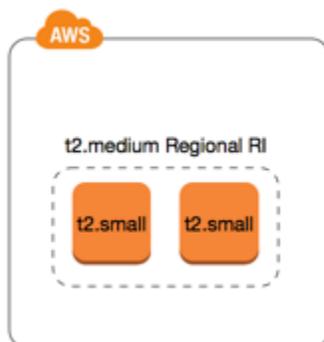


Figura 1: Duas instâncias t2.medium em execução em uma região

Ou, se você tiver uma instância t2.large em execução em sua conta na região Leste dos EUA (Norte da Virgínia), o benefício de faturamento será aplicado a 50% do uso da instância.

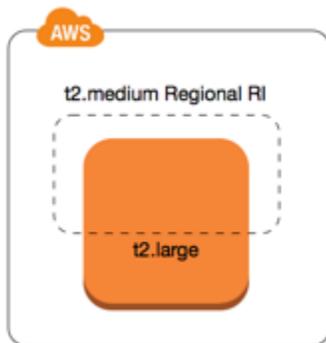


Figura 2: Uma instância t2.large em execução em uma região

O fator de normalização é aplicado também ao modificar instâncias padrão reservadas.

Fator de normalização para instâncias do EC2 dedicadas

Para RIs sem flexibilidade de tamanho, o fator de normalização é sempre 1. O fator de normalização não se aplica a instâncias do EC2 sem flexibilidade de tamanho. O único objetivo do fator de normalização é fornecer a capacidade de combinar várias instâncias do EC2 entre si em uma família para que você possa trocar um tipo por outro. Não oferecemos suporte a esse caso de uso para instâncias do EC2 sem flexibilidade de tamanho, portanto, o fator de normalização não é usado e, para manter nosso modelo de dados uniforme em diferentes casos de uso do EC2, atribuímos a ele um valor equivalente a 1.

Fator de normalização para instâncias bare metal

A flexibilidade de tamanho da instância também se aplica a instâncias bare metal na família de instâncias. Se você tem instâncias reservadas Linux/Unix regionais da Amazon com locação compartilhada em instâncias bare metal, é possível se beneficiar das economias de instâncias reservadas na mesma família de instâncias. O inverso também é verdadeiro: se você tem instâncias reservadas Linux/Unix regionais da Amazon com locação compartilhada em instâncias na mesma família que uma instância bare metal, é possível se beneficiar das economias de instâncias reservadas na instância bare metal.

Uma instância bare metal tem o mesmo tamanho que a maior instância da mesma família de instâncias. Por exemplo, uma i3.metal tem o mesmo tamanho que uma i3.16xlarge e, portanto, elas têm o mesmo fator de normalização.

Os tamanhos das instâncias .metal não têm um único fator de normalização. Eles variam de acordo com a família de instâncias específica. Para obter a lista mais atualizada, consulte [Tipos de instância do Amazon EC2](#).

Tabela 4: Tamanhos de instâncias bare metal e fatores de normalização

Tamanho da instância	Fator de normalização
a1.metal	32
c5.metal	192
c5d.metal	192
c5n.metal	144
c6g.metal	128
c6gd.metal	128
g4dn.metal	192
i3.metal	128
i3en.metal	192
m5.metal	192
m5d.metal	192
m5dn.metal	192
m5n.metal	192
m5zn.metal	96
m6g.metal	128
m6gd.metal	128
r5.metal	192

Tamanho da instância	Fator de normalização
r5b.metal	192
r5d.metal	192
r5dn.metal	192
r5n.metal	192
r6g.metal	128
r6gd.metal	128
x2gd.metal	128
z1d.metal	96

Por exemplo, uma instância i3.metal tem um fator de normalização de 128. Se você comprar uma instância reservada Linux/Unix da Amazon de locação padrão i3.metal na região Leste dos EUA (Norte da Virgínia), o benefício de faturamento poderá ser aplicado da seguinte forma:

- Se você tiver uma i3.16xlarge em execução na sua conta nessa região, o benefício de faturamento será aplicado integralmente à instância i3.16xlarge (fator de normalização i3.16xlarge = 128).
- Ou, se você tiver duas instâncias i3.8xlarge em execução em sua conta nessa região, o benefício de faturamento será aplicado integralmente às duas instâncias i3.8xlarge (fator de normalização da i3.8xlarge = 64).
- Ou, se você tiver quatro instâncias i3.4xlarge em execução em sua conta nessa região, o benefício de faturamento será aplicado integralmente a todas as quatro instâncias i3.4xlarge (fator de normalização da i3.4xlarge = 32).

O inverso também é verdadeiro. Por exemplo, se você comprar duas instâncias reservadas Linux/Unix da Amazon de locação padrão i3.8xlarge na região Leste dos EUA (Norte da Virgínia) e tiver uma instância i3.metal em execução nessa região, o benefício de faturamento será aplicado integralmente à instância i3.metal.

Savings Plans

Os [Savings Plans](#) são outro 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 de família de instâncias, tamanho, sistema operacional, localização ou região da AWS, e também é aplicado ao uso do AWS Fargate e do AWS Lambda.

Os Savings Plans oferecem economias significativas em relação às instâncias sob demanda, assim como as instâncias reservadas do EC2, em troca do compromisso de usar uma quantidade específica de potência computacional (medido em USD/hora) por um período de um ou três anos. Você pode se cadastrar nos Savings Plans por um período de vigência de um ou três anos e gerenciar seus planos com facilidade, aproveitando as recomendações, os relatórios de performance e os alertas de orçamento no AWS Cost Explorer.

A AWS oferece dois tipos de Savings Plans:

- Os Compute Savings Plans fornecem a maior flexibilidade e ajudam a reduzir seus custos em até 66% (como RIs conversíveis). Esses planos se aplicam automaticamente ao uso da instância do EC2, independentemente da família de instâncias, tamanho, AZ, região, sistema operacional ou localização da instância, e também se aplicam ao uso do Fargate e do Lambda. Por exemplo, com os Compute Savings Plans, você pode mudar das instâncias C4 para M5, deslocar uma workload da região Europa (Irlanda) para a região Europa (Londres) ou mover uma workload do Amazon EC2 para o Fargate ou o Lambda a qualquer momento e continuar pagando automaticamente o preço dos Savings Plans.
- Os EC2 Instance Savings Plans fornecem os preços mais baixos, oferecendo economia de até 72% (como as RIs padrão) em troca do comprometimento com o uso de famílias de instâncias individuais em uma região (por exemplo, usar a M5 no Norte da Virginia). Isso reduz automaticamente o custo na família de instâncias selecionadas nessa região, qualquer que seja a AZ, o tamanho, o sistema operacional ou a localização. Os EC2 Instance Savings Plans oferecem a flexibilidade de alterar o uso entre instâncias de uma família nessa região. Por exemplo, você pode passar da c5.xlarge executando o Windows para a c5.2xlarge executando o Linux e se beneficiar automaticamente dos preços dos Savings Plans.

Os Savings Plans não fornecem uma reserva de capacidade. No entanto, você pode reservar capacidade com reservas de capacidade sob demanda e pagar preços mais baixos com os Savings Plans.

Você pode continuar comprando RIs para manter a compatibilidade com seus processos de gerenciamento de custos existentes, e suas RIs funcionarão em conjunto com os Savings Plans para reduzir sua fatura geral. No entanto, à medida que suas RIs expiram, recomendamos que você se cadastre nos Savings Plans, pois eles oferecem as mesmas economias que as RIs, mas com flexibilidade adicional.

Modelos de reserva para outros serviços da AWS

Além do Amazon EC2, os modelos de reserva estão disponíveis para o Amazon RDS, o Amazon ElastiCache, o OpenSearch Service, o Amazon Redshift e o Amazon DynamoDB.

Tópicos

- [Instâncias de banco de dados reservadas do Amazon RDS](#)
- [Nós reservados do Amazon ElastiCache](#)
- [Instâncias reservadas do Amazon OpenSearch Service](#)
- [Nós reservados do Amazon Redshift](#)
- [Reservas do Amazon DynamoDB](#)

Instâncias de banco de dados reservadas do Amazon RDS

Semelhante às instâncias reservadas do Amazon EC2, há três opções de pagamento para instâncias de banco de dados reservadas do Amazon RDS: sem pagamento adiantado, com pagamento adiantado parcial e com adiantamento integral. Todos os tipos de instâncias de banco de dados reservadas estão disponíveis para os mecanismos de banco de dados Aurora, MySQL, MariaDB, PostgreSQL, Oracle e SQL Server.

Instâncias de banco de dados reservadas com tamanho flexível estão disponíveis para Amazon Aurora, MariaDB, MySQL, PostgreSQL e a edição Bring Your Own License “Traga Sua Própria Licença” do mecanismo de banco de dados Oracle.

Para obter mais informações sobre instâncias de banco de dados reservadas do Amazon RDS, consulte o seguinte:

- [Instâncias reservadas do Amazon RDS](#)
- [Trabalhar com instâncias de bancos de dados reservadas](#)
- [Definição de preço do Amazon DynamoDB](#)

Nós reservados do Amazon ElastiCache

Os nós reservados do Amazon ElastiCache oferecem a opção de fazer um pagamento único e baixo para cada nó de cache que você deseja reservar. Por sua vez, você recebe um desconto significativo

na cobrança por hora desse nó. O Amazon ElastiCache oferece três tipos de nó de cache reservados (Utilização leve, média e pesada) que permitem equilibrar o valor inicial pago antecipadamente e o preço efetivo por hora. Com base na workload da aplicação e do período planejado de execução, os nós reservados do Amazon ElastiCache podem oferecer uma economia significativa em relação à execução de nós sob demanda. Os nós de cache reservados estão disponíveis para Redis e Memcached.

Para obter mais informações, consulte [Nós reservados do Amazon ElastiCache](#).

Instâncias reservadas do Amazon OpenSearch Service

As instâncias reservadas (RIs) do Amazon OpenSearch Service (OpenSearch Service) oferecem descontos significativos em comparação com instâncias sob demanda padrão. As instâncias propriamente ditas são idênticas; as RIs são apenas um desconto de faturamento aplicado a instâncias sob demanda na conta. Para aplicações de longa duração com uso previsível, as RIs podem oferecer uma economia considerável ao longo do tempo. As RIs do OpenSearch Service exigem prazos de um ou três anos e têm três opções de pagamento que afetam a taxa de desconto.

Para obter mais informações, consulte [Instâncias reservadas do Amazon OpenSearch Service](#).

Nós reservados do Amazon Redshift

Na AWS, as cobranças acumuladas pela utilização do Amazon Redshift são baseadas em nós de computação. Cada nó de computação é faturado por uma taxa horária. A taxa horária varia de acordo com fatores como a região da AWS, o tipo de nó e se é possível ou não aplicar ao nó as definições de preços sob demanda ou de preços de nós reservados.

Se você pretende manter seu cluster do Amazon Redshift em execução continuamente durante um período prolongado, deve considerar a compra de ofertas de nós reservados. Essas ofertas oferecem economias significativas em relação aos preços sob demanda. No entanto, os nós reservados exigem que você reserve nós de computação e se comprometa a pagar por esses nós por um período de um a três anos.

Para obter mais informações sobre o preço dos nós reservados do Amazon Redshift, consulte [Preço de instâncias reservadas](#) e [Comprar nós reservados do Amazon Redshift](#).

Reservas do Amazon DynamoDB

Se você puder prever a necessidade de taxa de transferência de leitura e gravação do Amazon DynamoDB, a capacidade reservada oferecerá economia considerável sobre o preço normal da taxa de transferência provisionada do DynamoDB. Você paga uma taxa única adiantada e se compromete a pagar por um nível mínimo de utilização, de acordo com taxas horárias específicas, durante o período de vigência da capacidade reservada. Toda a taxa de transferência provisionada além da capacidade reservada será cobrada usando as taxas padrão de taxa de transferência provisionada.

O modo de capacidade provisionada poderá ser a melhor opção se você:

- Tiver tráfego de aplicações previsível.
- Executar aplicações com tráfego consistente e que aumenta gradualmente.
- Poderá prever os requisitos de capacidade para controlar os custos.

Para obter informações, consulte [Definição de preço de capacidade provisionada](#).

Cobrança de instâncias reservadas

As instâncias reservadas fornecem um desconto significativo em comparação com o preço de instâncias sob demanda. Com instâncias reservadas, você paga por todo o período de vigência, e não pelo uso real. É possível optar por pagamento adiantado, parcial adiantado ou por mês pela instância reservada, dependendo da [opção de pagamento](#) especificada para a instância reservada.

Quando as instâncias reservadas expiram, são cobradas taxas de instâncias sob demanda. Você pode colocar uma instância reservada na fila para compra com até três anos de antecedência. Isso pode ajudar a garantir que você tenha cobertura ininterrupta. Para obter mais informações, consulte [Colocar sua compra na fila](#).

Você pode configurar um alerta de faturamento para avisá-lo quando sua conta ultrapassar um limite definido. Para obter mais informações, consulte [Monitorar as cobranças com alertas e notificações](#).

Faturamento do uso

Exceto reservas do DynamoDB, que são cobradas com base na taxa de transferência, as reservas são cobradas para cada hora fechada durante o período de vigência selecionado, independentemente de uma instância estar em execução ou não. Uma hora fechada é definida como um relógio padrão de 24 horas, que vai da meia-noite à meia-noite, e é dividida em 24 horas (por exemplo, 1:00:00 a 1:59:59 é uma hora fechada).

Um benefício do faturamento de instância reservada pode ser aplicado a uma instância em execução com base em uma taxa por segundo. O faturamento por segundo está disponível para instâncias que usam uma distribuição de código aberto do Linux, como o Amazon Linux e o Ubuntu. O faturamento por hora é usado para distribuições comerciais do Linux, como o Red Hat Enterprise Linux e o SUSE Linux Enterprise Server.

Um dos benefícios de faturamento de instância reservada pode ser aplicado a no máximo 3.600 segundos (uma hora) de uso de instância por hora fechada. Você pode executar várias instâncias simultaneamente, mas só pode receber o benefício do desconto de instância reservada por um total de 3.600 segundos por hora fechada. O uso da instância que exceder 3.600 segundos em uma hora fechada será cobrado de acordo com a taxa de instâncias sob demanda.

Por exemplo, se você adquirir uma instância reservada m4.xlarge e executar quatro instâncias m4.xlarge simultaneamente por uma hora, uma instância será cobrada por uma hora de uso de

instância reservada enquanto as outras três instâncias serão cobradas por três horas de uso de instância sob demanda.

Contudo, se você adquirir uma instância reservada m4.xlarge e executar quatro instâncias m4.xlarge por 15 minutos (900 segundos) cada na mesma hora, o tempo total de execução das instâncias será uma hora, o que resultará em uma hora de uso de instância reservada e 0 hora de uso sob demanda.

	1:00	1:15	1:30	1:45
Instance 1				
Instance 2				
Instance 3				
Instance 4				

Figura 3: Executar quatro instâncias por 15 minutos cada na mesma hora

Se várias instâncias qualificadas estiverem sendo executadas simultaneamente, o benefício de faturamento da instância reservada será aplicado a todas as instâncias ao mesmo tempo até um máximo de 3.600 segundos em uma hora fechada. Depois disso, as taxas de instância sob demanda serão aplicadas.

	1:00	1:15	1:30	1:45
Instance 1				
Instance 2				
Instance 3				
Instance 4				

Uses Reserved Instance Rate for first 3600 seconds of use Uses On-Demand Rate

Figura 4: Executar quatro instâncias simultaneamente ao longo da hora

Você pode saber mais sobre as cobranças e taxas da sua conta ao visualizar o console [Gerenciamento de Faturamento e Custos da AWS](#). Você também pode examinar sua utilização e cobertura e receber recomendações de compra de reserva, por meio do [AWS Cost Explorer](#). É possível se aprofundar nas reservas e na alocação de descontos de instâncias reservadas com o [AWS Cost and Usage Report](#).

Para obter mais informações sobre faturamento de uso de instâncias reservadas, consulte [Faturamento de uso](#).

Faturamento consolidado

O AWS Organizations é um serviço de gerenciamento de contas que permite consolidar várias contas da AWS em uma organização que você cria e gerencia centralmente. O AWS Organizations inclui as funcionalidades de gerenciamento de contas e faturamento consolidado que permitem melhor atender as necessidades de orçamento, segurança e conformidade de sua empresa. Para obter mais informações, consulte [O que é o AWS Organizations?](#)

Para obter mais informações sobre as faturas consolidadas e como elas são calculadas, consulte [Noções básicas das faturas consolidadas](#).

Os benefícios de preços das instâncias reservadas são compartilhados quando a conta que faz a compra é cobrada sob uma conta pagante de faturamento consolidado. O uso da instância em todas as contas-membro é agregada na conta pagante todos os meses. Isso é útil para empresas em que há equipes ou grupos funcionais diferentes; assim, a lógica usual da instância reservada é aplicada para calcular a conta.

Instâncias reservadas: reservas de capacidade

A AWS também oferece taxas por hora com desconto em troca de uma taxa inicial e período de vigência do contrato. Serviços como o Amazon EC2 e o Amazon RDS usam essa abordagem para vender capacidade reservada para uso por hora de instâncias reservadas. Para obter mais informações, consulte [Instâncias reservadas no Guia do usuário do Amazon EC2 para instâncias Linux](#) e [Trabalhar com instâncias de banco de dados reservadas no Guia do usuário do Amazon Relational Database Service](#).

Quando você reserva capacidade com instâncias reservadas, o uso por hora é calculado a uma tarifa com desconto para as instâncias do mesmo tipo de uso na mesma zona de disponibilidade (AZ). Quando você inicia instâncias adicionais do mesmo tipo de na mesma zona de disponibilidade e excede o número de instâncias na reserva, a AWS calcula as médias das tarifas de instâncias reservadas e de instâncias sob demanda para oferecer uma combinação de tarifas.

Combinação de tarifas

Um item de linha para a combinação de tarifas dessa instância é exibido na fatura de qualquer conta-membro que esteja executando uma instância correspondente às especificações de uma reserva na organização.

A conta pagante de uma organização pode desativar o compartilhamento de instâncias reservadas para contas-membro nessa organização por meio das Preferências de faturamento da AWS. Isso significa que as instâncias reservadas não são compartilhadas entre essa conta-membro e outras contas-membro. Cada fatura estimada é calculada usando o último conjunto de preferências.

Para obter informações sobre como configurar o compartilhamento, consulte [Desativar o compartilhamento de instâncias reservadas](#).

Como os descontos são aplicados

A aplicação das instâncias reservadas do Amazon EC2 é baseada em atributos de instância, incluindo o seguinte:

- Tipo de instância: os tipos de instância abrangem combinações variáveis de capacidade de CPU, memória, armazenamento e rede (por exemplo, m4.xlarge). Isso oferece a flexibilidade de escolher a combinação adequada de recursos para suas aplicações, como otimização para computação, otimização para armazenamento etc. Cada tipo de instância inclui um ou mais tamanhos, permitindo escalar seus recursos de acordo com os requisitos da workload de destino.
- Plataforma: você pode comprar instâncias reservadas para as instâncias do Amazon EC2 que executam as plataformas Linux, Unix, SUSE Linux, Red Hat Enterprise Linux, Windows Server e Microsoft SQL Server.
- Locação: as instâncias reservadas podem ser locação padrão ou locação dedicada.
- Regional ou zonal: consulte [Instâncias reservadas regionais e zonais](#).

Se você tiver adquirido uma instância reservada e já tiver uma instância em execução que corresponda aos atributos da instância reservada, o benefício de faturamento será imediatamente aplicado. Você não precisa reiniciar suas instâncias. Se você não tiver uma instância em execução adequada, inicie uma instância e faça a correspondência dos mesmos critérios especificados para sua instância reservada. Para obter mais informações, consulte [Usar suas instâncias reservadas](#).

Maximizar o valor das reservas

Esta seção explica como maximizar o valor de suas reservas.

Tópicos

- [Medir o sucesso](#)
- [Maximizar os descontos padronizando o tipo de instância](#)
- [Técnicas de gerenciamento de reservas](#)
- [Marketplace de instâncias reservadas](#)
- [AWS Cost Explorer](#)
- [AWS Cost and Usage Report](#)
- [AWS Trusted Advisor](#)

Medir o sucesso

Aproveitar ao máximo as reservas significa medir a cobertura da sua reserva (parte das instâncias que desfrutam de benefícios de desconto de reserva) e a utilização da reserva (grau em que as Instâncias reservadas compradas são usadas). Estabeleça uma frequência de revisão padronizada na qual você se concentre nas seguintes questões:

- Você precisa modificar alguma de nossas reservas existentes para aumentar a utilização?
- As reservas utilizadas atualmente estão expirando?
- Você precisa comprar alguma reserva para aumentar sua cobertura?

Uma cadência de revisão padronizada garante que os problemas sejam apresentados e abordados em tempo hábil. À medida que suas RIs expiram, recomendamos que você se cadastre nos Savings Plans, pois eles oferecem as mesmas economias que as RIs, mas com flexibilidade adicional.

Maximizar os descontos padronizando o tipo de instância

Ao padronizar os tipos de instância que sua organização usa, você pode garantir que as implantações correspondam às características de suas reservas para maximizar seus descontos. A padronização maximiza a utilização e minimiza o nível de esforço associado ao gerenciamento de reservas. Três serviços que podem ajudar você a padronizar suas instâncias são:

- [AWS Config](#): permite avaliar, auditar e examinar as configurações de seus recursos da AWS. O AWS Config monitora e registra continuamente suas configurações de recursos da AWS e permite automatizar a avaliação das configurações registradas em relação às configurações desejadas.
- [Service Catalog](#): permite criar e gerenciar catálogos de serviços de TI aprovados para uso na AWS. Esses serviços de TI podem incluir tudo, de imagens de máquinas virtuais (VM), servidores, software e bancos de dados a uma arquitetura completa de aplicações multicamada.
- [AWS Compute Optimizer](#): recomenda os recursos computacionais ideais da AWS para suas workloads a fim de reduzir custos e melhorar a performance usando algoritmos de machine learning para analisar métricas de utilização históricas. O Compute Optimizer analisa a configuração e a utilização de recursos da workload para identificar dezenas de características determinantes. Por exemplo, se uma workload faz uso intenso de CPU, se ela apresenta um padrão diário ou se acessa o armazenamento local com frequência. O serviço processa essas características e identifica a margem de recursos de hardware exigida pela workload. Ele infere como seria a performance da workload em várias plataformas de hardware (por exemplo, tipos de instância do Amazon EC2) e faz recomendações.

Técnicas de gerenciamento de reservas

Você pode gerenciar reservas usando uma equipe central de operações ou gerenciamento de TI ou usando uma equipe ou unidade de negócios específica. A tabela a seguir resume as diferentes técnicas de gerenciamento de reservas.

Tabela 5: Comparação entre diferentes técnicas de gerenciamento de reservas

Gerenciamento central de reservas	Gerenciamento de reservas de equipe/unidade de negócios
Maximiza a cobertura de reservas abrangendo o uso agregado em uma empresa	Aumenta a probabilidade de alta utilização de reservas (por exemplo, usando reservas já adquiridas), porque uma única equipe deve entender seu compromisso de capacidade das RIs
Simplifica o gerenciamento geral de reservas, especialmente ao combinar o gerenciamento central e as instâncias reservadas conversíveis	Reduz a interface ou o planejamento entre a unidade de negócios e a equipe central

Gerenciamento central de reservas	Gerenciamento de reservas de equipe/unidade de negócios
Reduz a necessidade de uma equipe individual entender as reservas	Simplifica as decisões sobre compras, processo de compra e localização da conta de reserva

Marketplace de instâncias reservadas

O [Marketplace de instâncias reservadas](#) é compatível com a venda de instâncias padrão reservadas não utilizadas de clientes da AWS e terceiros, que varia em termos de comprimento e opções de preço. Por exemplo, você pode querer vender instâncias reservadas depois de movê-las para uma nova região da AWS, mudar para um novo tipo de instância, concluir projetos antes da validade do período de vigência quando suas necessidades de negócio mudarem ou se você tiver capacidade desnecessária.

Se você quiser vender suas instâncias reservadas não utilizadas Marketplace de instâncias reservadas, deverá atender a determinados critérios de elegibilidade. Para obter mais informações, consulte [Marketplace de instâncias reservadas](#).

AWS Cost Explorer

O [AWS Cost Explorer](#) permite visualizar, entender e gerenciar os custos e o uso da AWS ao longo do tempo. É possível analisar seus dados de custo e uso em um alto nível (por exemplo, custos totais e uso em todas as contas da organização) ou para solicitações altamente específicas (por exemplo, custos da m2.2xlarge na conta Y marcados como projeto: secretProject).

Você pode se aprofundar em suas reservas usando os relatórios de cobertura e utilização de instâncias reservadas. Com esses relatórios, é possível definir metas personalizadas de utilização e cobertura de instâncias reservadas e visualizar o andamento em relação às suas metas. A partir daí, você pode refinar os dados subjacentes usando as dimensões de filtragem disponíveis (por exemplo, conta, tipo de instância, escopo etc.).

O AWS Cost Explorer fornece os seguintes relatórios predefinidos:

- [% de utilização do EC2](#) oferece dados relevantes para identificar e agir de acordo com as oportunidades de aumentar a eficiência de uso das instâncias reservadas. Ele é calculado

dividindo as horas de instâncias reservadas usadas pelo total de horas compradas de instâncias reservadas.

- [% de cobertura de RI do EC2](#) mostra quanto do uso geral da instância é coberto por instâncias reservadas. Isso permite que você tome decisões embasadas sobre quando comprar ou modificar uma instância reservada para garantir a cobertura máxima. Ele é calculado dividindo as horas de instâncias reservadas usadas pelo total de horas de instâncias reservadas e sob demanda do EC2.

Além disso, o AWS Cost Explorer faz recomendações de compra de instâncias reservadas zonais e de tamanho flexível para ajudar as contas pagantes a obter maior eficiência de custo. Para obter mais informações, consulte [AWS Cost Explorer](#).

AWS Cost and Usage Report

O [AWS Cost and Usage Report](#) contém o conjunto mais abrangente de dados sobre seus custos e uso da AWS, incluindo informações adicionais sobre serviços, preços e reservas da AWS. Com esse relatório, você pode obter uma variedade de insights relacionados a reservas sobre o Nome de recurso da Amazon (ARN) para uma reserva, o número de reservas, o número de unidades por reserva e muito mais. Ele pode ajudar a fazer o seguinte:

- Calcular economias: cada item de linha por hora de uso contém a taxa com desconto que foi cobrada, além da taxa de Instância sob demanda pública para esse tipo de uso naquele momento. Você pode quantificar suas economias calculando a diferença entre as taxas de instâncias sob demanda públicas e as taxas cobradas.
- Monitorar a alocação de descontos de instância reservada: cada item de linha de uso que recebe um desconto contém informações sobre a origem do desconto. Essas informações facilitam a monitoração de quais instâncias estão se beneficiando de reservas específicas.

Esses relatórios são atualizados até três vezes por dia.

Instâncias reservadas em seu relatório de custo e uso

O item de linha Fee (Taxa) é adicionado à sua fatura quando você compra uma Instância reservada com adiantamento integral ou adiantamento parcial, conforme mostrado.

lineitem/ LineItemType	lineitem/ Product Code	lineitem/ UsageStartDate	lineitem/ Description	lineitem/ Unblended Cost	reservation/ ReservationARN
Fee	Amazon EC2	2016-01-01T00:00:00Z	Sign up charge for subscription: 363836886, planId: 1026576	68	Am:aws:ec2:us-east-1:572481847476:reserved-instances/1-0048-4371-...

Figura 5: Item de linha Fee (Taxa) do AWS Cost and Usage Report

O item de linha RI Fee (Taxa de RI) descreve as cobranças mensais recorrentes associadas a instâncias reservadas com pagamento adiantado parcial e sem pagamento adiantado. A RI Fee (Taxa de RI) é calculada multiplicando sua taxa por hora com desconto pelo número de horas no mês, conforme mostrado.

lineitem/ LineItemType	lineitem/ Product Code	lineitem/ UsageStartDate	lineitem/ UsageType	lineitem/ Description	lineitem/ Normalization Factor	lineitem/ Unblended Cost	reservation/ AvailabilityZone	reservation/ ReservationARN	reservation/ TotalReservedUnits	reservation/ TotalReserved NormalizedUnits
RI fee	Amazon EC2	2016-01-01T00:00:00Z	HeavyUsage: m4.large	USD 0.0309 hourly fee per Linux/Unix (Amazon VPC) m4.large instance	4	23		Am:aws:ec2:us-east-1:572481847476:reserved-instances/1-0048-4371-...	744	2976

Figura 6: Item de linha RI Fee (Taxa de RI) do AWS Cost and Usage Report

O item de linha Discounted Usage (Uso com desconto) descreve o uso da instância que recebeu um benefício de desconto de Instância reservada correspondente. Ele é adicionado à sua fatura quando você tem um uso correspondente a uma de suas instâncias reservadas, conforme mostrado.

lineitem/ LineItemType	lineitem/ Product Code	lineitem/ UsageStartDate	lineitem/ UsageType	lineitem/ Description	lineitem/ ResourceID	lineitem/ AvailabilityZone	lineitem/ UsageAmount	lineitem/ Normalization Factor	lineitem/ Normalized UsageAmount	lineitem/ Unblended Rate	lineitem/ Unblended Cost	reservation/ ReservationARN
Discounted Usage	Amazon EC2	2016-01-01T00:00:00Z	BoxUsage: m4.large	Linux/Unix (Amazon VPC), m4.large reserved instance applied	i-70e200bc	us-east-1b	4	4	4	0	0	Am:aws:ec2:us-east-1:572481847476:reserved-instances/1-0048-4371-...

Figura 7: Item de linha Discounted Usage (Uso com desconto) do AWS Cost and Usage Report

AWS Trusted Advisor

O [AWS Trusted Advisor](#) é um recurso online para ajudar a reduzir custos, aumentar a performance e melhorar a segurança ao otimizar seu ambiente da AWS; o AWS Trusted Advisor disponibiliza orientações em tempo real para auxiliar você a provisionar seus recursos seguindo as práticas recomendadas da AWS. Para ajudar a maximizar a utilização de instâncias reservadas, o AWS Trusted Advisor verifica seu histórico de consumo de computação do Amazon EC2 e calcula um número ideal de instâncias reservadas com pagamento adiantado parcial. As recomendações são baseadas no uso horário do mês anterior no calendário agregado por todas as contas de

faturamento consolidadas. O Trusted Advisor não faz recomendações de instâncias reservadas com tamanho flexível.

Para obter mais informações sobre como a recomendação é calculada, consulte “Perguntas de verificação de otimização de instâncias reservadas” nas perguntas frequentes sobre o Trusted Advisor.

Conclusão

Planejadas e gerenciadas com eficiência, as reservas podem ajudar você a obter descontos significativos para workloads da AWS executadas em um cronograma previsível. É importante analisar o uso atual da AWS para selecionar os atributos de reserva corretos desde o início e criar uma estratégia de longo prazo para monitorar e gerenciar suas instâncias reservadas. Usando ferramentas como o AWS Compute Optimizer, o AWS Cost and Usage report e os relatórios de utilização e cobertura de instâncias reservadas no AWS Cost Explorer, você pode examinar o uso geral e descobrir oportunidades para maior eficiência de custos.

Colaboradores

Os colaboradores desse documento incluem:

- Pritam Pal, arquiteto de soluções especialista sênior, EC2 Spot, Amazon Web Services

Revisões do documento

Para ser notificado sobre atualizações deste whitepaper, inscreva-se no RSS feed.

update-history-change

[Tipos de instância bare metal e fatores de normalização atualizados. Link removido para instâncias agendadas.](#)

update-history-description

Atualização secundária.

update-history-date

29 de março de 2021

[Informações de faturamento e fatores de normalização de instâncias reservadas atualizadas. Seção Savings Plan adicionada.](#)

Whitepaper atualizado.

31 de agosto de 2020

[Publicação inicial](#)

Whitepaper publicado.

1 de março de 2018

Avisos

Os clientes são responsáveis por fazer sua própria avaliação independente das informações neste documento. Este documento é: (a) fornecido apenas para fins informativos, (b) representa as ofertas e práticas de produtos atuais da AWS, que estão sujeitas a alterações sem aviso prévio e (c) não cria nenhum compromisso ou garantia da AWS e suas afiliadas, fornecedores ou licenciadores. Os produtos ou serviços da AWS são fornecidos no “estado em que se encontram”, sem garantias, declarações ou condições de qualquer tipo, explícitas ou implícitas. As responsabilidades e obrigações da AWS com seus clientes são regidas por contratos da AWS, e este documento não modifica nem faz parte de nenhum contrato entre a AWS e seus clientes.

© 2021, Amazon Web Services, Inc. ou suas afiliadas. Todos os direitos reservados.