
Amazon EC2 Reserved Instances and Other AWS Service Reservation Models

AWS Whitepaper



Amazon EC2 Reserved Instances and Other AWS Service Reservation Models: AWS Whitepaper

Copyright © 2018 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Amazon EC2 Reserved Instances and Other Reservation Models	1
Abstract	1
Introduction	1
Amazon EC2 Reserved Instances	2
Reserved Instance Payment Options	2
Standard vs. Convertible Offering Classes	2
Maximizing Utilization with Size Flexibility in Regional Reserved Instances	3
Reservation Models for Other AWS Services	5
Amazon RDS Reserved DB Instances	5
Amazon ElastiCache Reserved Nodes	5
Amazon Redshift Reserved Nodes	5
Amazon DynamoDB Reservations	6
Reserved Instance Billing	7
Usage Billing	7
Consolidated Billing	7
Reserved Instances: Capacity Reservations	7
Blended Rates	8
Regional vs. Zonal Reserved Instances	8
How Discounts Are Applied	8
Maximizing the Value of Reservations	10
Measure Success	10
Maximize Discounts by Standardizing Instance Type	10
Reservation Management Techniques	11
AWS Cost Explorer	11
AWS Cost and Usage Report	11
Reserved Instances on Your Cost and Usage Report	12
AWS Trusted Advisor	13
Conclusion	14
Resources	15
Document Details	16
Contributors	16
.....	16
AWS Glossary	17

Amazon EC2 Reserved Instances and Other Reservation Models

Publication date: **March 2018** ([Document Details](#) (p. 16))

Abstract

This is the third paper in a series of whitepapers designed to support your cloud journey. This paper seeks to empower you to maximize value from your investments, improve forecasting accuracy and cost predictability, create a culture of ownership and cost transparency, and continuously measure your optimization status.

This paper discusses Amazon EC2 Reserved Instances and reservation models for other AWS services.

Introduction

Although the cloud is well suited for variable workloads and rapid deployment, many cloud-based workloads display a more predictable pattern. For these stable applications, your organization can achieve significant cost savings by using [Amazon Elastic Compute Cloud \(Amazon EC2\) Reserved Instances](#). Amazon EC2 Reserved Instances enable you to commit to usage parameters at the time of purchase to achieve a lower hourly rate. Reservation models are also available for [Amazon Relational Database Service \(Amazon RDS\)](#), [Amazon ElastiCache](#), [Amazon Redshift](#), and [Amazon DynamoDB](#). This whitepaper discusses Amazon EC2 Reserved Instances and the reservation models for these other AWS services.

Amazon EC2 Reserved Instances

When you purchase Reserved Instances, you make a one-year or three-year commitment and receive a billing discount of up to 75 percent. For the right workloads, Reserved Instances can save you a lot of money.

A Reserved Instance is not a physical instance. It is a billing discount applied to the use of On-Demand Instances in your account. These On-Demand Instances must match certain attributes to benefit from the billing discount. You pay for the entire term of a Reserved Instance, regardless of actual usage, so your cost savings are closely tied to use. Therefore, it pays to plan and monitor your usage to make the most of your investment.

When you purchase a Reserved Instance in a specific Availability Zone, it provides a capacity reservation. This improves the likelihood that the compute capacity you need is available in a specific Availability Zone when you need it. A Reserved Instance purchased for an AWS Region does not provide capacity reservation.

Reserved Instance Payment Options

You can purchase Reserved Instances through the AWS Management Console. The following payment options are available for most Reserved Instances:

- **No Upfront** – No upfront payment is required. You are billed a discounted hourly rate for every hour within the term, regardless of whether the Reserved Instance is being used. No Upfront Reserved Instances are based on a contractual obligation to pay monthly for the entire term of the reservation. A successful billing history is required before you can purchase No Upfront Reserved Instances.
- **Partial Upfront** – A portion of the cost must be paid up front and the remaining hours in the term are billed at a discounted hourly rate, regardless of whether you're using the Reserved Instance.
- **All Upfront** – Full payment is made at the start of the term, with no other costs or additional hourly charges incurred for the remainder of the term, regardless of hours used.

Reserved Instances with a higher upfront payment provide greater discounts. You can also find Reserved Instances offered by third-party sellers at lower prices and shorter terms on the [Reserved Instance Marketplace](#).

Volume discounts begin to apply as you purchase more Reserved Instances, and let you save even more.

For more information, see [Amazon EC2 Reserved Instance Pricing](#).

Standard vs. Convertible Offering Classes

When you purchase a Reserved Instance, you can choose between a Standard or Convertible offering class.

Standard Reserved Instance	Convertible Reserved Instance
One-year to three-year term	Three-year term

Standard Reserved Instance	Convertible Reserved Instance
Enables you to <i>modify</i> Availability Zone, scope, network platform, and instance size (within the same instance type) of your Reserved Instance. For more information, see Modifying Reserved Instances .	Enables you to <i>exchange</i> one or more Convertible Reserved Instances for another Convertible Reserved Instance with new attributes. These attributes include instance family, instance type, platform, scope, and tenancy, if the exchange results in the creation of a Reserved Instance of equal or greater value. For more information, see Exchanging Convertible Reserved Instances .
Can be sold in the Reserved Instance Marketplace.	Cannot be sold in the Reserved Instance Marketplace.

Standard Reserved Instances typically provide the highest discount levels. One-year Standard Reserved Instances provide a similar discount to three-year Convertible Reserved Instances.

Convertible Reserved Instances are useful when:

- Purchasing Reserved Instances in the payer account instead of a subaccount. You can more easily modify Convertible Reserved Instances to meet changing needs across your organization.
- Workloads are likely to change. In this case, a Convertible Reserved Instance enables you to adapt as needs evolve while still obtaining discounts and capacity reservations.
- You want to hedge against possible future price drops.
- You don't want to ask teams to do capacity planning or forecasting.
- You expect compute usage to remain at the committed amount over the commitment period.

Maximizing Utilization with Size Flexibility in Regional Reserved Instances

For additional flexibility, all Regional Linux Reserved Instances with shared tenancy apply to all sizes of instances within an instance family and an AWS Region, even if you are using them across multiple accounts via [Consolidated Billing](#). The only attributes that must be matched are the instance type (e.g., m4), tenancy (must be default), and platform (must be Linux). All new and existing Reserved Instances are sized according to a normalization factor based on instance size, as follows.

Instance Size	Normalization Factor
nano	0.25
micro	0.5
small	1
medium	2
large	4
xlarge	8
2xlarge	16
4xlarge	32

Amazon EC2 Reserved Instances and Other AWS
Service Reservation Models AWS Whitepaper
Maximizing Utilization with Size
Flexibility in Regional Reserved Instances

Instance Size	Normalization Factor
8xlarge	64
10xlarge	80
16xlarge	128
32xlarge	256

For example, if you have a Reserved Instance for a c4.8xlarge, it applies to any usage of a Linux c4 instance with shared tenancy in the AWS Region, such as:

- One c4.8xlarge instance
- Two c4.4xlarge instances
- Four c4.2xlarge instances
- Sixteen c4.large instances

It also includes combinations of c4 instances, such as one c4.4xlarge and eight c4.large instances. If you own a Reserved Instance that is smaller than the instance you are running, you are charged the prorated, On-Demand price for the excess.

Reservation Models for Other AWS Services

In addition to Amazon EC2, reservation models are available for Amazon RDS, Amazon ElastiCache, Amazon Redshift, and Amazon DynamoDB.

Topics

- [Amazon RDS Reserved DB Instances \(p. 5\)](#)
- [Amazon ElastiCache Reserved Nodes \(p. 5\)](#)
- [Amazon Redshift Reserved Nodes \(p. 5\)](#)
- [Amazon DynamoDB Reservations \(p. 6\)](#)

Amazon RDS Reserved DB Instances

Similar to Amazon EC2 Reserved Instances, there are three payment options for Amazon RDS reserved DB instances: No Upfront, Partial Upfront, and All Upfront. All reserved DB instance types are available for Aurora, MySQL, MariaDB, PostgreSQL, Oracle, and SQL Server database engines.

Size-flexible reserved DB instances are available for Amazon Aurora, MariaDB, MySQL, PostgreSQL, and the “Bring Your Own License” (BYOL) edition of the Oracle database engine.

For more information about Amazon RDS reserved DB instances, see the following:

- [Amazon RDS Reserved Instances](#)
- [Working with Reserved DB Instances](#)
- [Amazon DynamoDB Pricing](#)

Amazon ElastiCache Reserved Nodes

Amazon ElastiCache reserved nodes give you the option to make a low, one-time payment for each cache node you want to reserve. In turn, you receive a significant discount on the hourly charge for that node. Amazon ElastiCache provides three reserved cache node types (Light Utilization, Medium Utilization, and Heavy Utilization) that enable you to balance the amount you pay up front with your effective hourly price.

For more information, see [Amazon ElastiCache Reserved Nodes](#).

Amazon Redshift Reserved Nodes

In AWS, the charges that you accrue for using Amazon Redshift are based on compute nodes. Each compute node is billed at an hourly rate. The hourly rate varies depending on factors such as AWS Region, node type, and whether the node receives on-demand node pricing or reserved node pricing.

If you intend to keep an Amazon Redshift cluster running continuously for a prolonged period, you should consider purchasing reserved-node offerings. These offerings provide significant savings over on-

demand pricing. However, they require you to reserve compute nodes and commit to paying for those nodes for either a one-year or a three-year duration.

For more information about Amazon Redshift reserved node pricing, see [Reserved Instance Pricing](#) and [Purchasing Amazon Redshift Reserved Nodes](#).

Amazon DynamoDB Reservations

If you can predict your need for Amazon DynamoDB read-and-write throughput, reserved capacity offers significant savings over the normal price of DynamoDB provisioned throughput capacity. You pay a one-time upfront fee and commit to paying for a minimum usage level at specific hourly rates for the duration of the reserved capacity term. Any throughput you provision in excess of your reserved capacity is billed at standard rates for provisioned throughput.

Reserved Instance Billing

All Reserved Instances provide you with a discount compared to On-Demand pricing. With Reserved Instances, you pay for the entire term regardless of actual use. You can choose to pay for your Reserved Instance up front, partially up front, or monthly, depending on the [payment option](#) specified for the Reserved Instance.

When Reserved Instances expire, you are charged On-Demand rates. You can set up a billing alert to warn you when your bill exceeds a threshold you define. For more information, see [Monitoring Charges with Alerts and Notifications](#).

Usage Billing

Except for DynamoDB reservations, which are billed based on throughput, reservations are billed for every clock-hour during the term you select, regardless of whether an instance is running or not. A clock-hour is defined as the standard 24-hour clock that runs from midnight to midnight and is divided into 24 hours (for example, 1:00:00 to 1:59:59 is one clock-hour).

You can find out about the charges and fees to your account by viewing the [AWS Billing and Cost Management](#) console. You can also examine your utilization and coverage, and receive reservation purchase recommendations, via [AWS Cost Explorer](#). You can dive deeper into your reservations and Reserved Instance discount allocation via the [AWS Cost and Usage Report](#).

For more information on Reserved Instance usage billing, see [Usage Billing](#).

Consolidated Billing

The pricing benefits of Reserved Instances are shared when the purchasing account is billed under a consolidated billing payer account. The instance usage across all member accounts is aggregated in the payer account every month. This is useful for companies that have different functional teams or groups; then, the normal Reserved Instance logic is applied to calculate the bill.

AWS Organizations is an account management service that lets you consolidate multiple AWS accounts into an *organization* that you create and centrally manage. AWS Organizations includes consolidated billing and account management capabilities that enable you to better meet the budgetary, security, and compliance needs of your business. For more information, see [What Is AWS Organizations?](#)

For more information on consolidated bills and how they are calculated, see [Understanding Consolidated Bills](#).

Reserved Instances: Capacity Reservations

AWS also offers discounted hourly rates in exchange for an upfront fee and term contract. Services such as [Amazon EC2](#) and [Amazon RDS](#) use this approach to sell reserved capacity for hourly use of Reserved Instances. For more information, see [Reserved Instances](#) in the *Amazon EC2 User Guide for Linux Instances* and [Working with Reserved DB Instances](#) in the *Amazon Relational Database Service User Guide*.

When you reserve capacity with Reserved Instances, your hourly usage is calculated at a discounted rate for instances of the same usage type in the same Availability Zone. When you launch additional instances of the same instance type in the same Availability Zone and exceed the number of instances in your

reservation, AWS averages the rates of the Reserved Instances and the On-Demand Instances to give you a blended rate.

Blended Rates

A line item for the blended rate of that instance is displayed on the bill of any member account that is running an instance that matches the specifications of a reservation in the organization.

The master account of an organization can turn off Reserved Instance sharing for member accounts in that organization via the AWS Billing Preferences. This means that Reserved Instances are not shared between that member account and other member accounts. Each estimated bill is computed using the most recent set of preferences. For information on how to configure sharing, see [Turning Off Reserved Instance Sharing](#).

Regional vs. Zonal Reserved Instances

Standard and Convertible Reserved Instances can be purchased to apply to instances in a specific Availability Zone (Zonal Reserved Instances) or to instances in an AWS Region (Regional Reserved Instances).

The following table summarizes the differences between Reserved Instances purchased for an Availability Zone and for a Region.

Zonal Reserved Instance	Regional Reserved Instance
Provides a capacity reservation	Does not provide a capacity reservation
Applies only to the specified Availability Zone for the duration of the contract	Applies to usage in any Availability Zone in the Region
Applies only to instances of a specific size within a given instance family	For Reserved Instances on the Linux or Unix platform with default or shared tenancy; applies to instances of any size within a given instance family (referred to as <i>instance size flexibility</i>)

How Discounts Are Applied

The application of Amazon EC2 Reserved Instances is based on instance attributes, including the following:

- **Instance type** – Instance types comprise varying combinations of CPU, memory, storage, and networking capacity (e.g., m4.xlarge). This gives you the flexibility to choose the appropriate mix of resources for your applications, such as compute optimized, storage optimized, and so on. Each instance type includes one or more instance sizes, enabling you to scale your resources to the requirements of your target workload.
- **Platform** – You can purchase Reserved Instances for Amazon EC2 instances running Linux or UNIX, SUSE Linux, Red Hat Enterprise Linux, Windows Server, and Microsoft SQL Server platforms.
- **Tenancy** – Reserved Instances can be default tenancy or dedicated tenancy.
- **Regional or Zonal** – See [the section called “Maximizing Utilization with Size Flexibility in Regional Reserved Instances”](#) (p. 3).

If you purchase a Reserved Instance and you already have a running instance that matches the attributes of the Reserved Instance, the billing benefit is immediately applied. You don't have to restart your

instances. If you do not have an eligible running instance, launch an instance and ensure that you match the same criteria that you specified for your Reserved Instance. For more information, see [Using Your Reserved Instances](#).

Maximizing the Value of Reservations

This section discusses how you can maximize the value of your reservations.

Topics

- [Measure Success](#) (p. 10)
- [Maximize Discounts by Standardizing Instance Type](#) (p. 10)
- [Reservation Management Techniques](#) (p. 11)
- [AWS Cost Explorer](#) (p. 11)
- [AWS Cost and Usage Report](#) (p. 11)
- [AWS Trusted Advisor](#) (p. 13)

Measure Success

Making the most of reservations means measuring your reservation coverage (portion of instances enjoying reservation discount benefits) and reservation utilization (degree to which purchased Reserved Instances are used). Establish a standardized review cadence in which you answer the following questions:

- Do we need to modify any of our existing reservations to increase utilization?
- Are any currently utilized reservations expiring?
- Do we need to purchase any reservations to increase our coverage?

A standardized review cadence ensures that issues are surfaced in a timely manner.

Maximize Discounts by Standardizing Instance Type

By standardizing the instance types that your organization uses, you can ensure that deployments match the characteristics of your reservations to maximize your discounts. Standardization maximizes utilization and minimizes the level of effort associated with management of reservations. Two services that can help you standardize your instances are:

- [AWS Config](#) – Enables you to assess, audit, and evaluate the configurations of your AWS resources. AWS Config continuously monitors and records your AWS resource configurations and lets you automate the evaluation of recorded configurations against desired configurations.
- [AWS Service Catalog](#) – Lets you create and manage catalogs of IT services that are approved for use on AWS. These IT services can include everything from virtual machine images, servers, software, and databases to complete multi-tier application architecture.

Reservation Management Techniques

You can manage reservations either by using a central IT operations or management team or by using a specific team or business unit. The following table summarizes the different reservation management techniques.

Central Reservation Management	Team/Business Unit Reservation Management
Maximizes reservation coverage by covering aggregate usage across a business	Increases likelihood of high reservation utilization (e.g., using already-purchased reservations), because a single team should understand its capacity commitment of RIs
Simplifies overall reservation management especially when combining central management and Convertible Reserved Instances	Reduces interfacing or planning between the business unit and the central team
Reduces the requirement for an individual team to understand reservations	Streamlines decisions about purchases, purchase process, and reservation account location

AWS Cost Explorer

[AWS Cost Explorer](#) lets you visualize, understand, and manage your AWS costs and usage over time. You can analyze your cost and usage data at a high level (e.g., total costs and usage across all accounts in your organization) or for highly specific requests (e.g., m2.2xlarge costs within account Y that are tagged *project: secretProject*).

You can dive deeper into your reservations using the Reserved Instance utilization and coverage reports, which are available out of the box. Using these reports, you can set custom Reserved Instance utilization and coverage targets, and visualize progress toward your goals. From there, you can refine the underlying data using the available filtering dimensions (e.g., account, instance type, scope, and more).

AWS Cost Explorer provides the following prebuilt reports:

- [EC2 RI Utilization %](#) offers relevant data to identify and act on opportunities to increase your Reserved Instance usage efficiency. It's calculated by dividing Reserved Instance used hours by total Reserved Instance purchased hours.
- [EC2 RI Coverage %](#) shows how much of your overall instance usage is covered by Reserved Instances. This lets you make informed decisions about when to purchase or modify a Reserved Instance to ensure maximum coverage. It's calculated by dividing Reserved Instance used hours by total EC2 On-Demand and Reserved Instance hours.

Furthermore, AWS Cost Explorer provides Reserved Instance purchase recommendations for Zonal and Size-Flexible Reserved Instances to help master accounts achieve greater cost efficiencies.

AWS Cost and Usage Report

The [AWS Cost and Usage Report](#) contains the most comprehensive set of data about your AWS costs and usage, including additional information regarding AWS services, pricing, and reservations. By using the AWS Cost and Usage report, you can gain a wealth of reservation-related insights about the Amazon

Resource Name (ARN) for a reservation, the number of reservations, the number of units per reservation, and more. It can help you do the following:

- **Calculate savings** – Each hourly line item of usage contains the discounted rate that was charged, as well as the public On-Demand rate for that usage type at that time. You can quantify your savings by calculating the difference between the public On-Demand rates and the rates you were charged.
- **Track the allocation of Reserved Instance discounts** – Each line item of usage that receives a discount contains information about where the discount came from. This makes it easier to trace which instances are benefitting from specific reservations.

These reports update up to three times per day.

Reserved Instances on Your Cost and Usage Report

The **Fee** line item is added to your bill when you purchase an All Upfront or Partial Upfront Reserved Instance, as shown.

lineltem/ LinelitemType	lineltem/ Product Code	lineltem/ UsageStartDate	lineltem/ Description	lineltem/ 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/f8c204c1-dd48-43f1-adb8-f38aa61e0dea

The **RI Fee** line item describes the recurring monthly charges that are associated with Partial Upfront and No Upfront Reserved Instances. The **RI Fee** is calculated by multiplying your discounted hourly rate by the number of hours in the month, as shown.

lineltem/ LinelitemType	lineltem/ Product Code	lineltem/ UsageStartDate	lineltem/ UsageType	lineltem/ Description	lineltem/ Normalization Factor	lineltem/ Unblended Cost	reservation/ AvailabilityZone	reservation/ ReservationARN	reservation/ TotalReservedUnits	reservation/ TotalReservedNormalizedUnits
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/f8c204c1-dd48-43f1-adb8-f38aa61e0dea	744	2976

The **Discounted Usage** line item describes the instance usage that received a matching Reserved Instance discount benefit. It's added to your bill when you have usage that matches one of your Reserved Instances, as shown.

lineltem/ LinelitemType	lineltem/ Product Code	lineltem/ UsageStartDate	lineltem/ UsageType	lineltem/ Description	lineltem/ ResourceID	lineltem/ AvailabilityZone
Discounted Usage	Amazon EC2	2016-01-01T00:00:00Z	BoxUsage: m4.large	Linux/Unix (AmazonVPC), m4.large reserved instance applied	i-1bd250bc	us-east-1b

AWS Trusted Advisor

[AWS Trusted Advisor](#) is an online resource to help you reduce cost, increase performance, and improve security by optimizing your AWS environment. AWS Trusted Advisor provides real-time guidance to help you provision your resources following AWS best practices. To help you maximize utilization of Reserved Instances, AWS Trusted Advisor checks your Amazon EC2 computing-consumption history and calculates an optimal number of Partial Upfront Reserved Instances. Recommendations are based on the previous calendar month's hour-by-hour usage aggregated across all consolidated billing accounts. Note that Trusted Advisor does not provide size-flexible Reserved Instance recommendations.

For more information about how the recommendation is calculated, see *Reserved Instance Optimization Check Questions* in the [Trusted Advisor FAQs](#).

Conclusion

Effectively planned and managed, reservations can help you achieve significant discounts for AWS workloads that run on a predictable schedule. It's important to analyze your current AWS usage to select the right reservation attributes from the start and to devise a longer-term strategy for monitoring and managing your Reserved Instances. Using tools such as the AWS Cost and Usage report, and the Reserved Instance Utilization and Coverage reports in AWS Cost Explorer, you can examine your overall usage and discover opportunities for greater cost efficiencies.

Resources

- [AWS Architecture Center](#)
- [AWS Whitepapers](#)
- [AWS Architecture Monthly](#)
- [AWS Architecture Blog](#)
- [This Is My Architecture videos](#)
- [AWS Answers](#)
- [AWS Documentation](#)

Document Details

Contributors

The following individuals and organizations contributed to this document:

- Amilcar Alfaro, Sr. Product Marketing Manager, AWS
- Erin Carlson, Marketing Manager, AWS
- Keith Jarrett, WW BD Lead - Cost Optimization, AWS Business Development

Date	Description
March 2018	First publication

AWS Glossary

For the latest AWS terminology, see the [AWS Glossary](#) in the *AWS General Reference*.