EC2 FAQ

EC2

A Spot Instance is an unused EC2 instance that is available for less than the On-Demand price. Your Spot Instance runs whenever capacity is available and the maximum price per hour for your request exceeds the Spot price. Any instance present with unused capacity will be allocated.

You can specify that Amazon EC2 should do one of the following when it interrupts a Spot Instance:

Stop the Spot Instance

Hibernate the Spot Instance

Terminate the Spot Instance

The default is to terminate Spot Instances when they are interrupted.

**Reboot the Spot Instance** - This is an invalid option.

Most resources can have up to 50 tags.

User data is limited to 16KB & User data and metadata are not encrypted.

Instance metadata is available at http://169.254.169.254/latest/meta-data/ (the trailing “/” is required).

The Instance Metadata Query tool allows you to query the instance metadata without having to type out the full URI or category names.

Instance user data is available at: <http://169.254.169.254/latest/user-data>

You can use the RequestSpotFleet API operation to launch thousands of Spot Instances and diversify resources automatically.

To reduce the impact of interruptions and optimize Spot Instances, diversify and run your application across multiple capacity pools

To further reduce the impact of interruptions, you can also set up Spot Instances and Spot Fleets to respond to an interruption notice by stopping or hibernating rather than terminating instances when capacity is no longer available

Instance type modifications are supported for Linux only.

Amazon EC2 uses Amazon EBS and Amazon S3 to provide reliable, scalable storage of your AMIs so that we can boot them when you ask us to do so

With the Amazon CloudWatch metrics integration, you can view EC2 usage against limits in the Service Quotas console.

Service Quotas also enables customers to use CloudWatch for configuring alarms to warn customers of approaching limits.

In addition, you can continue to track and inspect your instance usage in Trusted Advisor and Limit Monitor

EC2 status checks are performed every minute and each returns a pass or a fail status. If all checks pass, the overall status of the instance is OK. >>> If one or more checks fail, the overall status is impaired.

Each dedicated host can only run one EC2 instance size and type.

. Status checks are built into Amazon EC2, so they cannot be disabled or deleted.You can, however create or delete alarms that are triggered based on the result of the status checks.

You can create Amazon CloudWatch alarms that monitor Amazon EC2 instances and automatically perform an action if the status check fails.

Actions can include:

Recover the instance (only supported on specific instance types and can be used only with StatusCheckFailed\_System).

Stop the instance (only applicable to EBS-backed volumes).

Terminate the instance (cannot terminate if termination protection is enabled).

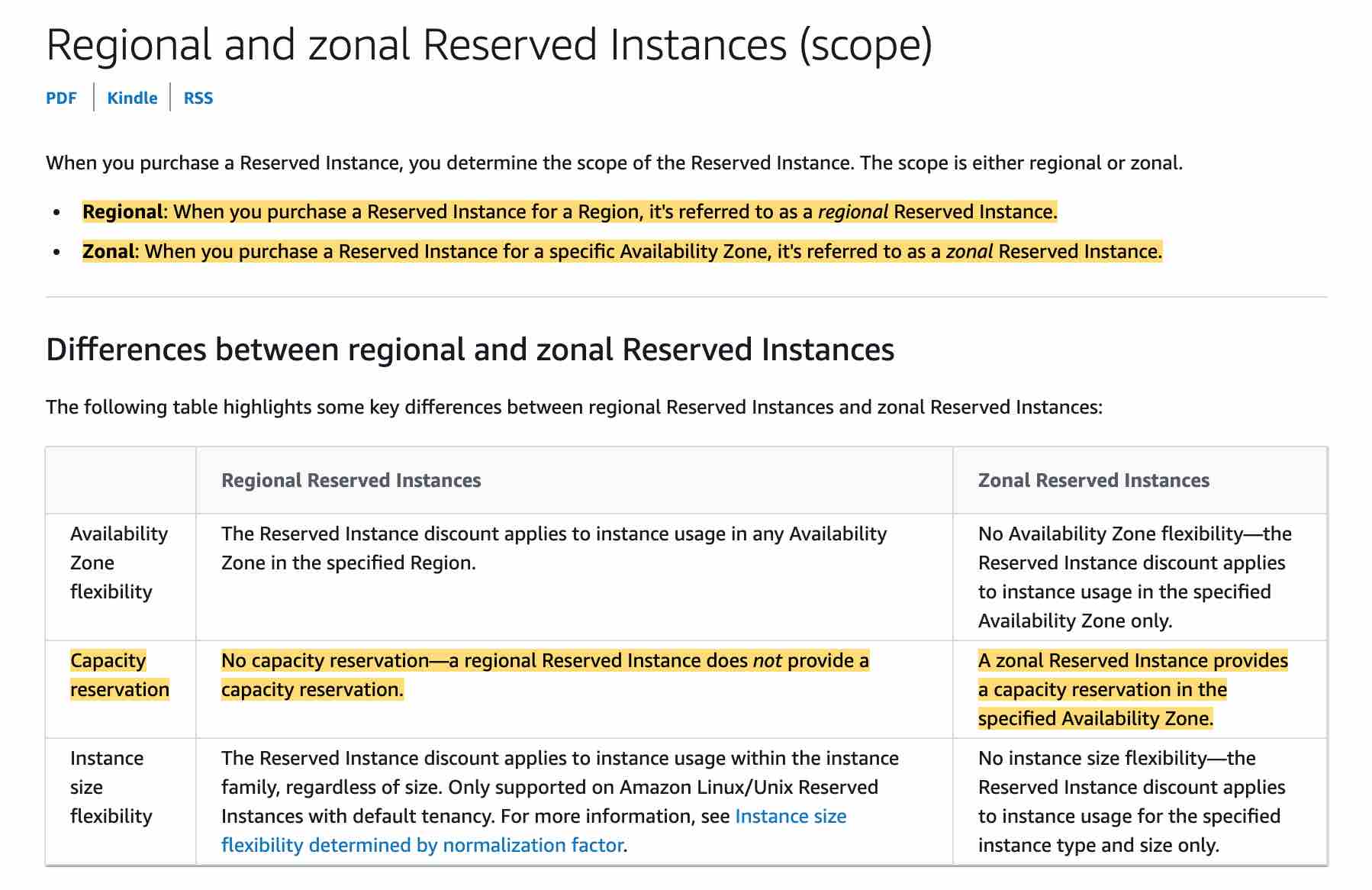
Reboot the instance

check on the status of your instances using the DescribeInstances API call

RunInstances API StopInstances API.

When you purchase a Reserved Instance for a specific Availability Zone, it's referred to as a Zonal Reserved Instance. Zonal Reserved Instances provide capacity reservations as well as discounts.

**Zonal Reserved Instances** - A zonal Reserved Instance provides a capacity reservation in the specified Availability Zone. Capacity Reservations enable you to reserve capacity for your Amazon EC2 instances in a specific Availability Zone for any duration. This gives you the ability to create and manage Capacity Reservations independently from the billing discounts offered by Savings Plans or regional Reserved Instances.



A Classic Load Balancer with HTTP or HTTPS listeners might route more traffic to higher-capacity instance types. This distribution aims to prevent lower-capacity instance types from having too many outstanding requests. It’s a best practice to use similar instance types and configurations to reduce the likelihood of capacity gaps and traffic imbalances.

Q. When should I use Compute Optimized instances?

Compute Optimized instances are designed for applications that benefit from high compute power. These applications include compute-intensive applications like high-performance web servers, high-performance computing (HPC), scientific modelling, distributed analytics and machine learning inference.

Q. What is the lifecycle of a Dedicated Host?

Once a Dedicated Host is allocated within your account, it will be standing by for your use. You can then launch an instance with a tenancy of "host" using the RunInstances API, and can also stop/start/terminate the instance through the API. You can use the AWS Management Console to manage the Dedicated Host and the instance. The Dedicated Host will be allocated to your account for the period of 3-year reservation. After the 3-year reservation expires, you can continue using the host or release it anytime.

Q. Can I launch, stop/start, and terminate High Memory instances using AWS CLI/SDK?

You can launch, stop/start, and terminate instances on your EC2 Dedicated Hosts using AWS CLI/SDK.

Q. When should I use Memory-optimized instances?

Memory-optimized instances offer large memory size for memory intensive applications including in-memory applications, in-memory databases, in-memory analytics solutions, High Performance Computing (HPC), scientific computing, and other memory-intensive applications.

Q. What are the storage options available with High Memory instances?

High Memory instances support Amazon EBS volumes for storage. High Memory instances are EBS-optimized by default, and offer up to 38Gbps of storage bandwidth to both encrypted and unencrypted EBS volumes.

Q: How many instances can I run in Amazon EC2?

You are limited to running On-Demand Instances per your vCPU-based [On-Demand Instance limit](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-on-demand-instances.html#ec2-on-demand-instances-limits), purchasing 20 Reserved Instances, and requesting Spot Instances per your [dynamic Spot limit](http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-spot-limits.html) per region. New AWS accounts may start with limits that are lower than the limits described here.

If you need more instances, complete the [Amazon EC2 limit increase request form](https://aws.amazon.com/support/createCase?type=service_limit_increase&serviceLimitIncreaseType=ec2-instances) with your use case, and your limit increase will be considered. Limit increases are tied to the region they were requested for.

Q: What happens to my data when a system terminates?

The data stored on a local instance store will persist only as long as that instance is alive. However, data that is stored on an Amazon EBS volume will persist independently of the life of the instance. Therefore, we recommend that you use the local instance store for temporary data and, for data requiring a higher level of durability, we recommend using Amazon EBS volumes or backing up the data to Amazon S3. If you are using an Amazon EBS volume as a root partition, you will need to set the Delete On Terminate flag to "N" if you want your Amazon EBS volume to persist outside the life of the instance.

Q: Which volume type should I choose?

Amazon EBS includes two major categories of storage: SSD-backed storage for transactional workloads (performance depends primarily on IOPS) and HDD-backed storage for throughput workloads (performance depends primarily on throughput, measured in MB/s). SSD-backed volumes are designed for transactional, IOPS-intensive database workloads, boot volumes, and workloads that require high IOPS. SSD-backed volumes include Provisioned IOPS SSD (io1 and io2) and General Purpose SSD (gp2 and gp3). HDD-backed volumes are designed for throughput-intensive and big-data workloads, large I/O sizes, and sequential I/O patterns. HDD-backed volumes include Throughput Optimized HDD (st1) and Cold HDD (sc1).

Q: Do you support multiple instances accessing a single volume?

Yes, you can enable Multi-Attach on an EBS Provisioned IOPS io1 volume to allow a volume to be concurrently attached to up to sixteen Nitro-based EC2 instances within the same Availability Zone. For more information on Amazon EBS Multi-Attach, see the [EBS product page](https://aws.amazon.com/ebs/).

Q: Will I be able to access my EBS snapshots using the regular Amazon S3 APIs?

No, EBS snapshots are only available through the Amazon EC2 APIs.

Q: Do volumes need to be un-mounted in order to take a snapshot? Does the snapshot need to complete before the volume can be used again?

No, snapshots can be done in real time while the volume is attached and in use. However, snapshots only capture data that has been written to your Amazon EBS volume, which might exclude any data that has been locally cached by your application or OS. In order to ensure consistent snapshots on volumes attached to an instance, we recommend cleanly detaching the volume, issuing the snapshot command, and then reattaching the volume. For Amazon EBS volumes that serve as root devices, we recommend shutting down the machine to take a clean snapshot.

Q: What charges apply when using Amazon EBS shared snapshots?

If you share a snapshot, you won’t be charged when other users make a copy of your snapshot. If you make a copy of another user’s shared volume, you will be charged normal EBS rates.

Q: How can I discover Amazon EBS snapshots that have been shared with me?

You can find snapshots that have been shared with you by selecting “Private Snapshots” from the viewing dropdown in the Snapshots section of the AWS Management Console. This section will list both snapshots you own and snapshots that have been shared with you.

Q: How can I find what Amazon EBS snapshots are shared globally?

You can find snapshots that have been shared globally by selecting “Public Snapshots” from the viewing dropdown in the Snapshots section of the AWS Management Console.

Q: Do you offer encryption on Amazon EBS volumes and snapshots?

Yes. EBS offers seamless encryption of data volumes and snapshots. EBS encryption better enables you to meet security and encryption compliance requirements.

Q. How do I access a file system from an Amazon EC2 instance?

To access your file system, you mount the file system on an Amazon EC2 Linux-based instance using the standard Linux mount command and the file system’s DNS name. Once you’ve mounted, you can work with the files and directories in your file system just like you would with a local file system.

Amazon EFS uses the NFSv4.1 protocol. For a step-by-step example of how to access a file system from an Amazon EC2 instance,

. What Amazon EC2 instance types and AMIs work with Amazon EFS?

Amazon EFS is compatible with all Amazon EC2 instance types and is accessible from Linux-based AMIs. You can mix and match the instance types connected to a single file system.

Q. How do I load data into a file system?

You can load data into an Amazon EFS file system from your Amazon EC2 instances or from your on-premises datacenter servers.

Amazon EFS file systems can be mounted on an Amazon EC2 instance, so any data that is accessible to an Amazon EC2 instance can also be read and written to Amazon EFS. To load data that is not currently stored on the Amazon cloud, you can use the same methods you use to transfer files to Amazon EC2 today, such as Secure Copy (SCP).

Amazon EFS file systems can also be mounted on an on-premises server, so any data that is accessible to an on-premises server can be read and written to Amazon EFS using standard Linux tools. For more information about accessing a file system from an on-premises server,

Q. How do I access my file system from outside my VPC?

Amazon EC2 instances within your VPC can access your file system directly, and Amazon EC2 Classic instances outside your VPC can mount a file system via [ClassicLink](http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/vpc-classiclink.html). On-premises servers can mount your file systems via an [AWS Direct Connect](https://aws.amazon.com/directconnect/) connection to your VPC.

Q. How many Amazon EC2 instances can connect to a file system?

Amazon EFS supports one to thousands of Amazon EC2 instances connecting to a file system concurrently.

Q: How long does it take to remap an Elastic IP address?

The remap process currently takes several minutes from when you instruct us to remap the Elastic IP until it fully propagates through our system

Q: How do I prevent other people from viewing my systems?

You have complete control over the visibility of your systems. The Amazon EC2 security systems allow you to place your running instances into arbitrary groups of your choice. Using the web services interface, you can then specify which groups may communicate with which other groups, and also which IP subnets on the Internet may talk to which groups. This allows you to control access to your instances in our highly dynamic environment. Of course, you should also secure your instance as you would any other server.

Q: Can I get a history of all EC2 API calls made on my account for security analysis and operational troubleshooting purposes?

Yes. To receive a history of all EC2 API calls (including VPC and EBS) made on your account, you simply turn on CloudTrail in the [AWS Management Console](https://console.aws.amazon.com/cloudtrail/home).

Q: What is the minimum time interval granularity for the data that Amazon CloudWatch receives and aggregates?

Metrics are received and aggregated at 1 minute intervals.

Q: Will I lose the metrics data if I disable monitoring for an Amazon EC2 instance?

You can retrieve metrics data for any Amazon EC2 instance up to 2 weeks from the time you started to monitor it. After 2 weeks, metrics data for an Amazon EC2 instance will not be available if monitoring was disabled for that Amazon EC2 instance. If you want to archive metrics beyond 2 weeks you can do so by calling mon-get-stats command from the command line and storing the results in Amazon S3 or Amazon SimpleDB.

Q: Can I access the metrics data for a terminated Amazon EC2 instance or a deleted Elastic Load Balancer?

Yes. Amazon CloudWatch stores metrics for terminated Amazon EC2 instances or deleted Elastic Load Balancers for 2 weeks.

Q: Why does the graphing of the same time window look different when I view in 5 minute and 1 minute periods?

If you view the same time window in a 5 minute period versus a 1 minute period, you may see that data points are displayed in different places on the graph. For the period you specify in your graph, Amazon CloudWatch will find all the available data points and calculates a single, aggregate point to represent the entire period. In the case of a 5 minute period, the single data point is placed at the beginning of the 5 minute time window. In the case of a 1 minute period, the single data point is placed at the 1 minute mark. We recommend using a 1 minute period for troubleshooting and other activities that require the most precise graphing of time periods.

Q: Why should I hibernate an instance?

You can hibernate an instance to get your instance and applications up and running quickly, if they take long time to bootstrap (e.g. load memory caches). You can start instances, bring them to a desired state and hibernate them. These “pre-warmed” instances can then be resumed to reduce the time it takes for an instance to return to service. Hibernation retains memory state across Stop/Start cycl

Q: How much does it cost to hibernate an instance?

Hibernating instances are charged at standard EBS rates for storage. As with a stopped instance, you do not incur instance usage fees while an instance is hibernating.

Q: How can I hibernate an instance?

Hibernation needs to be enabled when you launch the instance. Once enabled, you can use the StopInstances API with an additional ‘Hibernate’ parameter to trigger hibernation. You can also do this through the console by selecting your instance, then clicking Actions> Instance State > Stop - Hibernate

Q: How can I resume a hibernating instance?

You can resume by calling the StartInstances API as you would for a regular stopped instance. You can also do this through the console by selecting your instance, then clicking Actions > Instance State > Start

Q: Can I enable hibernation on an existing instance?

No, you cannot enable hibernation on an existing instance (running or stopped). This needs to be enabled during instance launch.

Q: How can I tell that an instance is hibernated?

You can tell that an instance is hibernated by looking at the state reason. It should be ‘Client.UserInitiatedHibernate’. This is visible on the console under “Instances - Details” view or in the DescribeInstances API response as “reason” field.

Q: What is the state of an instance when it is hibernating?

Hibernated instances are in ‘Stopped’ state.

Q: What if my EBS root volume is not large enough to store memory state (RAM) for hibernate?

To enable hibernation, space is allocated on the root volume to store the instance memory (RAM). Make sure that the root volume is large enough to store the RAM contents and accommodate your expected usage, e.g. OS, applications. If the EBS root volume does not enough space, hibernation will fail and the instance will get shutdown instead.

Q. How can I tell if an application needs more CPU resources than a Micro instance is providing?

The CloudWatch metric for CPU utilization will report 100% utilization if the instance bursts so much that it exceeds its available CPU resources during that CloudWatch monitored minute. CloudWatch reporting 100% CPU utilization is your signal that you should consider scaling – manually or via Auto Scaling – up to a larger instance type or scale out to multiple Micro instances.

Q. If an instance in a cluster placement group is stopped then started again, will it maintain its presence in the cluster placement group?

Yes. A stopped instance will be started as part of the cluster placement group it was in when it stopped. If capacity is not available for it to start within its cluster placement group, the start will fail

Q: Can I use stop or Hibernation interruption behaviors with Spot Fleet?

Yes, stop-start and hibernate-resume are supported with Spot Fleet with “maintain” fleet option enabled.

Q. Can I resume a hibernated instance?

No, you will not be able to resume a hibernated instance directly. Hibernate-resume cycle is controlled by Amazon EC2. If an instance is hibernated by Spot, it will be resumed by Amazon EC2 when the capacity becomes available

Q. When should I use Savings Plans, EC2 RIs, and Capacity Reservations?

Use Savings Plans or Regional RIs to reduce your bill while committing to a one- or three-year term. Savings Plans offer significant savings over On Demand, just like EC2 RIs, but automatically reduce customers’ bills on compute usage across any AWS region, even as usage changes. Use Capacity Reservations if you need the additional confidence in your ability to launch instances. Capacity Reservations can be created for any duration and can be managed independently of your Savings Plans or RIs. If you have Savings Plans or Regional RIs, they will automatically apply to matching Capacity Reservations. This gives you the flexibility to selectively add Capacity Reservations to a portion of your instance footprint and still reduce your bill for that usage.

Q: When should I purchase a zonal RI?

If you want to take advantage of the capacity reservation, then you should buy an RI in a specific Availability Zone.

Q: Do RIs provide a capacity reservation?

Yes, when a Standard or Convertible RI is scoped to a specific Availability Zone (AZ), instance capacity matching the exact RI configuration is reserved for your use (these are referred to as “zonal RIs”). Zonal RIs give you additional confidence in your ability to launch instances when you need them.

You can also choose to forego the capacity reservation and purchase Standard or Convertible RIs that are scoped to a region (referred to as “regional RIs”). Regional RIs automatically apply the discount to usage across Availability Zones and instance sizes in a region, making it easier for you to take advantage of the RI’s discounted rate.

Q: What are the differences between Standard RIs and Convertible RIs?

Standard RIs offer a significant discount on EC2 instance usage when you commit to a particular instance family. Convertible RIs offer you the option to change your instance configuration during the term, and still receive a discount on your EC2 usage

Q: When should I purchase a zonal RI?

If you want to take advantage of the capacity reservation, then you should buy an RI in a specific Availability Zone.

Q: When should I purchase a regional RI?

If you do not require the capacity reservation, then you should buy a regional RI. Regional RIs provide AZ and instance size flexibility, which offers broader applicability of the RI’s discounted rate.