

Amazon Web Services

- **Amazon Web Services** (abbreviated **AWS**) is a **collection of remote computing services** (also called web services) that together make up a cloud computing platform, offered over the Internet by Amazon.com. The most central and well-known of these services are Amazon EC2 and Amazon S3.



Developer(s)

[Amazon.com](https://amazon.com)

Initial release

July 2002

Website

aws.amazon.com

History

- **Launched in July 2002, Amazon** Web Services provide online services for other web sites or client-side applications. Most of these services are not exposed directly to end users, but instead offer functionality that other developers can use in their applications. Amazon Web Services' offerings are accessed over HTTP, using REST and SOAP protocols. All services are billed based on usage, but how usage is measured for billing varies from service to service
- In 2004 an engineer at Amazon presented a paper proposing how the company could make a profit on the infrastructure required to run the Amazon.com store. As a result, Amazon EC2 was built by a team in Cape Town, South Africa.
- In June 2007, Amazon claimed that more than **330,000 developers** had signed up to use Amazon Web Services.
- On April 20, 2011, some parts of Amazon Web Services suffered a major outage
- On June 29th, 2012, several websites that rely on Amazon Web Services were taken offline due to an electrical storm in the Northern Virginia area where Amazon's datacenters are located

List of AWS products

- **Amazon CloudFront**, a content delivery network (CDN) for distributing objects to so-called "edge locations" near the requester.
- **Amazon CloudWatch**, provides monitoring for AWS cloud resources and applications, starting with EC2.
- **Amazon DevPay**, currently in limited beta version, is a billing and account management system for applications that developers have built atop Amazon Web Services.
- **Amazon Elastic Beanstalk** provides quick deployment and management of applications in the cloud.
- **Amazon Elastic Block Store (EBS)** provides persistent block-level storage volumes for EC2.
- **Amazon Elastic Compute Cloud (EC2)** provides scalable virtual private servers using Xen.
- **Amazon Elastic MapReduce** allows businesses, researchers, data analysts, and developers to easily and cheaply process vast amounts of data. It uses a hosted Hadoop framework running on the web-scale infrastructure of EC2 and Amazon S3.
- **Amazon ElastiCache** provides in-memory caching for web applications.

List of AWS products

- **Amazon Flexible Payments Service (FPS)** provides an interface for micropayments.
- **Amazon Fulfillment Web Service** provides a programmatic web service for sellers to ship items to and from Amazon using Fulfillment by Amazon.
- **Amazon Historical Pricing** provides access to Amazon's historical sales data from its affiliates. (It appears that this service has been discontinued.)
- **Amazon Mechanical Turk (Mturk)** manages small units of work distributed among many humans.
- **Amazon Product Advertising API** formerly known as Amazon Associates Web Service (A2S) and Amazon E-Commerce Service (ECS), provides access to Amazon's product data and electronic commerce functionality.

List of AWS products

- **Amazon Relational Database Service (RDS)** provides a scalable database server with MySQL and Oracle support.
- **Amazon Route 53** provides a highly available and scalable Domain Name System (DNS) web service.
- **Amazon CloudSearch**, currently in limited beta version, provides basic full text search and indexing of textual content.
- **Amazon Simple Email Service (SES)** provides bulk and transactional email sending.
- **Amazon Simple Storage Service (S3)** provides Web Service based storage.
- **Amazon Simple Queue Service (SQS)** provides a hosted message queue for web applications.
- **Amazon Simple Notification Service (SNS)** provides a hosted multiprotocol "push" messaging for web applications.
- **Amazon Simple Workflow (SWF)** is a workflow service for building scalable, resilient applications.
- **Amazon SimpleDB**, allows developers to run queries on structured data. It operates in concert with EC2 and S3 to provide "the core functionality of a database."

List of AWS products

- **Amazon Virtual Private Cloud (VPC)** creates a logically isolated set of Amazon EC2 instances which can be connected to an existing network using a VPN connection.
- **AWS CloudFormation**, create a collection of related AWS resources and provision them in an orderly and predictable fashion.
- **AWS Import/Export**, accelerates moving large amounts of data into and out of AWS using portable storage devices for transport.
- **AWS Management Console (AWS Console)**, A web-based point and click interface to manage and monitor the Amazon infrastructure suite including EC2, EBS, Amazon Elastic MapReduce, and Amazon CloudFront.
- **AWS Storage Gateway**, an iSCSI block storage appliance with cloud-based backup.

Amazon Elastic Compute Cloud

Amazon Elastic Compute Cloud



Amazon Elastic Compute Cloud

<u>Original author(s)</u>	Amazon.com, Inc.
<u>Developer(s)</u>	Amazon.com
Initial release	August 25, 2006; 5 years ago
Development status	Active
<u>Operating system</u>	• Microsoft Windows • Mac OS X • Linux
<u>Available in</u>	English
<u>Type</u>	Virtual Private Server
<u>License</u>	Proprietary software
Website	aws.amazon.com/ec2/

Amazon Elastic Compute Cloud

- **Amazon Elastic Compute Cloud (EC2)** is a **central part of Amazon.com's cloud computing platform**, Amazon Web Services (AWS):
- EC2 allows users to rent virtual computers on which to run their own computer applications. EC2 allows scalable deployment of applications by providing a Web service through which a user can boot an Amazon Machine Image to create a virtual machine, which Amazon calls an "instance", containing any software desired.
- A user can create, launch, and terminate server instances as needed, paying by the hour for active servers, hence the term "elastic". EC2 provides users with control over the geographical location of instances that allows for latency optimization and high levels of redundancy
- In November 2010, Amazon made the switch of its own retail website to EC2 and AWS

History

- Amazon announced a limited public beta of EC2 on August 25, 2006.[Access to EC2 was granted on a first come first served basis. Amazon added two new instance types (Large and Extra-Large) on October 16, 2007. On May 29, 2008, two more types were added, High-CPU Medium and High-CPU Extra Large. There are currently twelve types of instances available.
- Amazon added three new features on March 27, 2008. These features included static IP addresses, Availability Zones, and User Selectable Kernels. Amazon added Elastic Block Store (EBS) on August 20, 2008. This provides persistent storage, a basic feature that had been lacking since the service was introduced.
- Amazon EC2 is in full production since it dropped the beta label on October 23, 2008. On the same day, Amazon announced the following features: a service level agreement for EC2, Microsoft Windows in beta form on EC2, Microsoft SQL Server in beta form on EC2, plans for an AWS management console, and plans for load balancing, autoscaling, and cloud monitoring services. These features were subsequently added on May 18, 2009.
- Amazon EC2 was mostly developed by a team in Cape Town, South Africa. The team was led by Chris Pinkham. Pinkham provided the initial architecture guidance for EC2 and then built the team and led the development of the project. Other members of the early team included Chris Brown, Quinton Hoole, Roland Paterson-Jones and Willem Van Biljon.

Elastic compute units

- The **Elastic Compute Unit (ECU)** was introduced by Amazon EC2 as an abstraction of computer resources. Amazon's Definition of ECU notes "We use several benchmarks and tests to manage the consistency and predictability of the performance of an EC2 Compute Unit."
- One EC2 Compute Unit provides the **equivalent CPU capacity** of a
- 1.0-1.2 GHz 2007 Opteron
- or
- 2007 Xeon processor.

- This is also the equivalent to an early-2006
- 1.7 GHz Xeon processor referenced in our original documentation"

Elastic compute units

■ Instance types

- **EC2 uses Xen virtualization.** Each virtual machine, called an "instance", functions as a virtual private server. Amazon.com sizes instances based on "Elastic Compute Units".

- The performance of multiple identical virtual machines may vary

■ Cost

- As of June 2011, Amazon charges (\$0.02 to \$0.03)/hour (\$14.40 to \$21.60)/month) for the smallest "**Micro Instance**" (t1.micro) virtual machine running Linux or Windows, respectively. High CPU and memory instances cost as much as \$2.48/hour. "**Reserved**" instances can go as low as \$6.43/month for a three-year prepaid plan. The data transfer charge ranges from free to \$0.12 per gigabyte, depending on the direction and monthly volume (inbound data transfer is free on all AWS services)

■ Free Tier

- As of December 2010 Amazon offers a bundle of **free resource credits** to new account holders. The credits are designed to run a "**micro**" sized server, storage (EBS), and **bandwidth for one year**. Charges are applied on demand so the credit need not be used in the first month

■ Reserved Instances

- Reserved instances enable EC2 service users to reserve an instance **for one or three years**. There is a fee associated with reserving an instance. The corresponding per hour rate charged by Amazon to operate the instance is much less than the rate charged for on-demand instances

Features

■ Operating systems

- When it launched in August 2006, the EC2 service offered
 - ◆ Linux
 - ◆ Sun Microsystems' OpenSolaris and
 - ◆ Solaris Express Community Edition.
- In October 2008, EC2 added
 - ◆ the Windows Server 2003
 - ◆ and
 - ◆ Windows Server 2008 operating systems to the list of available operating systems.
- As of December 2010 it has also been reported to run FreeBSD; in March 2011, NetBSD AMIs became available.

Features

■ Persistent Storage

- An EC2 instance may be launched with a choice of two types of storage for its boot disk or "root device." The **first option** is a local "instance-store" disk as a root device (originally the only choice). The **second option** is to use an EBS volume as a root device.
- **Instance-store volumes** are temporary storage, which survive rebooting an EC2 instance, but when the instance is terminated (e.g., by an API call, or due to a failure), this store is lost.
- **EBS volumes** provide persistent storage independent of the lifetime of the EC2 instance, and act much like hard drives on a real server.
- More accurately, they appear as block devices to the operating system that are backed by Amazon's disk arrays. The OS is free to use the device however it wants. In the most common case, a file system is loaded and the volume acts as a hard drive. Another possible use is the creation of RAID arrays by combining two or more EBS volumes.
- RAID allows increases of speed and/or reliability of EBS. Users can set up and manage storage volumes of sizes from 1GB to 1TB. The volumes support snapshots, which can be taken from a GUI tool or the API. EBS volumes can be attached or detached from instances while they are running, and moved from one instance to another.

Features

- **Persistent Storage**
- Simple Storage Service (S3) is a storage system in which data is accessible to EC2 instances, or directly over the network to suitably authenticated callers (all communication is over HTTP).
- Amazon does not charge for the bandwidth for communications between EC2 instances and S3 storage "in the same region."
- Accessing S3 data stored in a different region (for example, data stored in Europe from a US East Coast EC2 instance) will be billed at Amazon's normal rates.
- S3-based storage is priced per gigabyte per month.
- Applications access S3 through an API.
- For example, Apache Hadoop supports a special s3: filesystem to support reading from and writing to S3 storage during a MapReduce job.
- There are also S3 filesystems for Linux, which mount a remote S3 filestore on an EC2 image, as if it were local storage. As S3 is not a full POSIX filesystem, things may not behave the same as on a local disk (e.g., no locking support).

Features

- **Elastic IP Addresses**
- Amazon's Elastic IP Address feature is similar to static IP address in traditional data centers, with one key difference.
- A user can programmatically map an Elastic IP Address to any virtual machine instance without a network administrator's help and without having to wait for DNS to propagate the new binding.
- In this sense an Elastic IP Address belongs to the account and not to a virtual machine instance.
- It exists until it is explicitly removed. It remains associated with the account, even while it is associated with no instance.
- Partial IPv6 is provided in the US East (Northern Virginia), EU (Ireland) and Asia Pacific (Tokyo & Singapore) regions.

Features

■ Amazon CloudWatch

- Amazon CloudWatch is a Web service that provides real-time monitoring to Amazon's EC2 customers on their resource utilization such as CPU, disk and network. Cloudwatch does not provide any memory, disk space, or load average metrics. An Amazon engineer has stated that this is due to the requirement to install software in the VM - something they wish to avoid. The data is aggregated and provided through AWS management console. It can also be accessed through command line tools and Web API's, if the customer desires to monitor their EC2 resources through their enterprise monitoring software.
 - The metrics collected by Amazon CloudWatch enables Auto Scaling feature to dynamically add or remove EC2 instances. The customers are charged by the number of monitoring instances.
 - Since May 2011, Amazon CloudWatch accepts custom metrics that can be submitted programmatically via Web Services API and then monitored the same way as all other internal metrics, including setting up the alarms for them.
- ## ■ Automated Scaling
- Amazon's Auto Scaling feature of EC2 allows it to automatically adapt computing capacity to site traffic.

Reliability

- To make EC2 more fault-tolerant, Amazon engineered *Availability Zones* that are designed to be insulated from failures in other availability zones.
- Availability zones do not share the same infrastructure.
- Applications running in more than one availability zone can achieve higher availability.
- EC2 provides users with control *over the geographical location* of instances that allows for latency optimization and high levels of redundancy.
- For example, to minimize downtime, a user can set up server instances in multiple zones that are insulated from each other for most causes of failure such that one backs up the other.
- High availability database services, like *Amazon Relational Database Service* run on top of EC2 instances.

Issues

- In early July 2008 the anti-spam organizations Outblaze and Spamhaus.org began blocking Amazon's EC2 address pool due to problems with the distribution of spam and malware.
- On December 1, 2010 Amazon pulled its service to WikiLeaks after coming under political pressure in the USA. The Internet group Anonymous attempted to attack EC2 in revenge; however, Amazon was not affected by the attack.
- Amazon's websites were temporarily offline on December 12, 2010, although it was initially unclear if this was due to attacks or a hardware failure. An Amazon official later stated that it was due to a hardware failure

Issues

- On Sunday August 6th 2011, Amazon **suffered a power outage in one of their Ireland availability zones**. Lightning was originally blamed for the outage; however, on August 11th, Irish energy supplier ESB Networks dismissed this as a cause, but at time of writing, could not confirm what the cause of the problem was. The power outage raised multiple questions regarding Amazon's EBS infrastructure, which caused several bugs in their software to be exposed. The bugs resulted in some customers' data being deleted when recovering EBS volumes in a mid-write operation during the crash.
- August 8th, 2011 saw another network connectivity outage of Amazon's Northern Virginia datacenter, knocking out the likes of Reddit, Quora, Netflix and FourSquare. **The outage lasted around 25 minutes.**