Sign up for an AWS account

If you do not have an AWS account, complete the following steps to create one.

To sign up for an AWS account

- 1. Open https://portal.aws.amazon.com/billing/signup.
- 2. Follow the online instructions.

Part of the sign-up procedure involves receiving a phone call and entering a verification code on the phone keypad.

When you sign up for an AWS account, an AWS account root user is created. The root user has access to all AWS services and resources in the account. As a security best practice, assign administrative access to an administrative user, and use only the root user to perform tasks that require root user access.

AWS sends you a confirmation email after the sign-up process is complete. At any time, you can view your current account activity and manage your account by going to https://aws.amazon.com/ and choosing **My Account**.

Verify permissions

Before you can use Amazon VPC, you must have the required permissions. For more information, see Identity and access management for Amazon VPC and Amazon VPC policy examples.

Determine your IP address ranges

The resources in your VPC communicate with each other and with resources over the internet using IP addresses. When you create VPCs and subnets, you can select their IP address ranges. When you deploy resources in a subnet, such as EC2 instances, they receive IP addresses from the IP address range of the subnet. For more information, see IP addressing for your VPCs and subnets.

As you choose a size for your VPC, consider how many IP addresses you'll need across your AWS accounts and VPCs. Ensure that the IP address ranges for your VPCs don't overlap with the IP address ranges for your own network. If you need connectivity between multiple VPCs, you must ensure that they have no overlapping IP addresses.

IP Address Manager (IPAM) makes it easier to plan, track, and monitor the IP addresses for your application. For more information, see the IP Address Manager Guide.

Select your Availability Zones

An AWS Region is a physical location where we cluster data centers, known as Availability Zones. Each Availability Zone has independent power, cooling, and physical security, with redundant power, networking, and connectivity. The Availability Zones in a Region are physically separated by a meaningful distance, and interconnected through high-bandwidth, low-latency networking. You can design your application to run in multiple Availability Zones to achieve even greater fault tolerance.

Production environment

For a production environment, we recommend that you select at least two Availability Zones and deploy your AWS resources evenly in each active Availability Zone.

Development or test environment

For a development or test environment, you might choose to save money by deploying your resources in only one Availability Zone.

Plan your internet connectivity

Plan to divide each VPC into subnets based on your connectivity requirements. For example:

- If you have web servers that will receive traffic from clients on the internet, create a subnet for these servers in each Availability Zone.
- If you also have servers that will receive traffic only from other servers in the VPC, create a separate subnet for these servers in each Availability Zone.
- If you have servers that will receive traffic only through a VPN connection to your network, create a separate subnet for these servers in each Availability Zone.

If your application will receive traffic from the internet, the VPC must have an internet gateway. Attaching an internet gateway to a VPC does not automatically make your instances accessible from the internet. In addition to attaching the internet gateway, you must update the subnet route table with a route to the internet

gateway. You must also ensure that the instances have public IP addresses and an associated security group that allows traffic from the internet over specific ports and protocols required by your application.

Alternatively, register your instances with an internet-facing load balancer. The load balancer receives traffic from the clients and distributes it across the registered instances in one or more Availability Zones. For more information, see Elastic Load Balancing. To allow instances in a private subnet to access the internet (for example, to download updates) without allowing unsolicited inbound connections from the internet, add a public NAT gateway in each active Availability Zone and update the route table to send internet traffic to the NAT gateway. For more information, see Access the internet from a private subnet.

Create your VPC

After you've determined the number of VPCs and subnets that you need, what CIDR blocks to assign to your VPCs and subnets, and how to connect your VPC to the internet, you are ready to create your VPC. If you create your VPC using the AWS Management Console and include public subnets in your configuration, we create a route table for the subnet and add the routes required for direct access to the internet. For more information, see Create a VPC.

Deploy your application

After you've created your VPC, you can deploy your application.

Production environment

For a production environment, you can use one of the following services to deploy servers in multiple Availability Zones, to configure scaling so that you maintain the minimum number of servers required by your application, and to register your servers with a load balancer to distribute traffic evenly across your servers.

- Amazon EC2 Auto Scaling
- EC2 Fleet
- Amazon Elastic Container Service (Amazon ECS)

Development or test environment

For a development or test environment, you might choose to launch a single EC2 instance. For more information, see Get started with Amazon EC2 in the Amazon EC2 User Guide.





