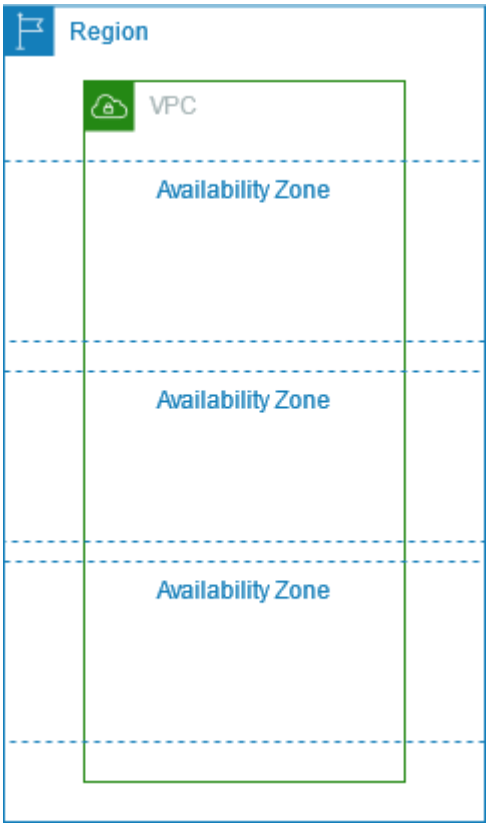
VPC Diagram

VPC is created and deployed on a region-wise basis, it covers all availability zones in the region.



VPC configuration options

Availability Zones

You can use multiple AZs to operate production applications and databases that are more highly available, fault-tolerant, and scalable than would be possible from a single data center.

CIDR blocks

You must specify IP address ranges for your VPC and subnets. For more information, see IP addressing for your VPCs and subnets.

DNS options

If you need public IPv4 DNS hostnames for the EC2 instances launched into your subnets, you must enable both of the DNS options. For more information, see DNS attributes for your VPC.

- **Enable DNS hostnames:** EC2 instances launched in the VPC receive public DNS hostnames that correspond to their public IPv4 addresses.
- **Enable DNS resolution:** DNS resolution for private DNS hostnames is provided for the VPC by the Amazon DNS server, called the Route 53 Resolver.

Internet gateway

Connects your VPC to the internet. The instances in a public subnet can access the internet because the subnet route table contains a route that sends traffic bound for the internet to the internet gateway. If a server doesn't need to be directly reachable from the internet, you should not deploy it into a public subnet. For more information, see Internet gateways.

Name

The names that you specify for the VPC and the other VPC resources are used to create Name tags. If you use the name tag auto-generation feature in the console, the tag values have the format *name-resource*.

NAT gateways

Enables instances in a private subnet to send outbound traffic to the internet, but prevents resources on the internet from connecting to the instances. In production, we recommend that you deploy a NAT gateway in each active AZ. For more information, see NAT gateways.

Route tables

Contains a set of rules, called routes, that determine where network traffic from your subnet or gateway is directed. For more information, see Route tables.

Subnets

A range of IP addresses in your VPC. You can launch AWS resources, such as EC2 instances, into your subnets. **Each subnet resides entirely within one Availability Zone.** By launching instances in at least two Availability Zones, you can protect your applications from the failure of a single Availability Zone.

A public subnet has a direct route to an internet gateway. Resources in a public subnet can access the public internet. A private subnet does not have a direct route to an internet gateway. **Resources in a private subnet require another component, such as a NAT device, to access the public internet.**

For more information, see Subnets.

Tenancy

This option defines if EC2 instances that you launch into the VPC will run on hardware that's shared with other AWS accounts or on hardware that's dedicated for your use only.

If you choose the tenancy of the VPC to be `Default`, EC2 instances launched into this VPC will use the tenancy attribute specified when you launch the instance -- For more information, see Launch an instance using defined parameters in the *Amazon EC2 User Guide*.

If you choose the tenancy of the VPC to be `Dedicated`, the instances will always run as Dedicated Instances on hardware that's dedicated for your use.

If you're using AWS Outposts, your Outpost requires private connectivity; you must use `Default` tenancy.

What is DHCP?

IP addresses are assigned dynamically by DHCP servers using the Dynamic Host Configuration Protocol (DHCP).

Applications running on EC2 instances can communicate with Amazon DHCP servers as needed to retrieve their IP address lease or other network configuration information (such as the IP address of an Amazon DNS server or the IP address of the router in your VPC).

You can specify the network configurations that are provided by Amazon DHCP servers by using DHCP option sets.

If you have a VPC configuration that requires your applications to make direct requests to the Amazon IPv6 DHCP server, note the following:

- **An EC2 instance in a dual-stack subnet** can only retrieve its IPv6 address from the IPv6 DHCP server. *It cannot retrieve any additional network configurations from the IPv6 DHCP server, such as DNS server names or domain names.*
- **An EC2 instance in an IPv6-only subnet** can retrieve its IPv6 address from the IPv6 DHCP server *and can retrieve additional networking configuration information, such as DNS server names and domain names.*
- For an EC2 instance in an IPv6-only subnet, the IPv4 DHCP Server will return 169.254.169.253 as the name server if "AmazonProvidedDNS" is explicitly mentioned in the DHCP option set.

  If "AmazonProvidedDNS" is missing from the option set, the IPv4 DHCP Server won't return an address, whether other IPv4 name servers are mentioned in the option set or not.

The Amazon DHCP servers can also provide an entire IPv4 or IPv6 prefix to a network interface in your VPC using prefix delegation (see Assigning prefixes to Amazon EC2 network interfaces in the *Amazon EC2 User Guide*).

IPv4 prefix delegation is not provided in DHCP responses.

IPv4 prefixes assigned to the interface can be retrieved using IMDS (see Instance metadata categories in the *Amazon EC2 User Guide*).

DHCP option set concepts

A *DHCP option set* is a group of network settings used by resources in your VPC, such as EC2 instances, to communicate over your virtual network.

Each Region has a default DHCP option set. Each VPC uses the default DHCP option set for its Region unless you either create and associate a custom DHCP option set with the VPC or configure the VPC with no DHCP option set.

If your VPC has no DHCP option set configured:

- For EC2 instances built on the Nitro System, AWS configures `169.254.169.253` as the default domain name server.

- For EC2 instances built on Xen, no domain name servers are configured and, because instances in the VPC have no access to a DNS server, they can't access the internet.

You can associate a DHCP option set with multiple VPCs, but each VPC can have only one associated DHCP option set.

If you delete a VPC, the DHCP option set that is associated with the VPC is disassociated from the VPC.

Default DHCP option set

The default DHCP option set contains the following settings:

- **Domain name servers**: The DNS servers that your network interfaces use for domain name resolution. For a default DHCP option set, this is always AmazonProvidedDNS. For more information, see Amazon DNS server.
- **Domain name:** The domain name that a client should use when resolving hostnames using the Domain Name System (DNS). For more information about the domain names used for EC2 instances, see Amazon EC2 instance hostnames.
- **IPv6 Preferred Lease Time:** How frequently a running instance with an IPv6 address assigned to it goes through DHCPv6 lease renewal. The default lease time is 140 seconds. Lease renewal typically occurs when half of the lease time has elapsed.

When you use a default DHCP options set, the following settings are not used, but there are defaults for EC2 instances:

- **NTP servers:** By default, EC2 instances use the Amazon Time Sync Service to retrieve the time.
- **NetBIOS name servers:** For EC2 instances running Windows, the NetBIOS computer name is a friendly name assigned to the instance to identify it on the network. The NetBIOS name server maintains a list of mappings between NetBIOS computer names and network addresses for networks that use NetBIOS as their naming service.

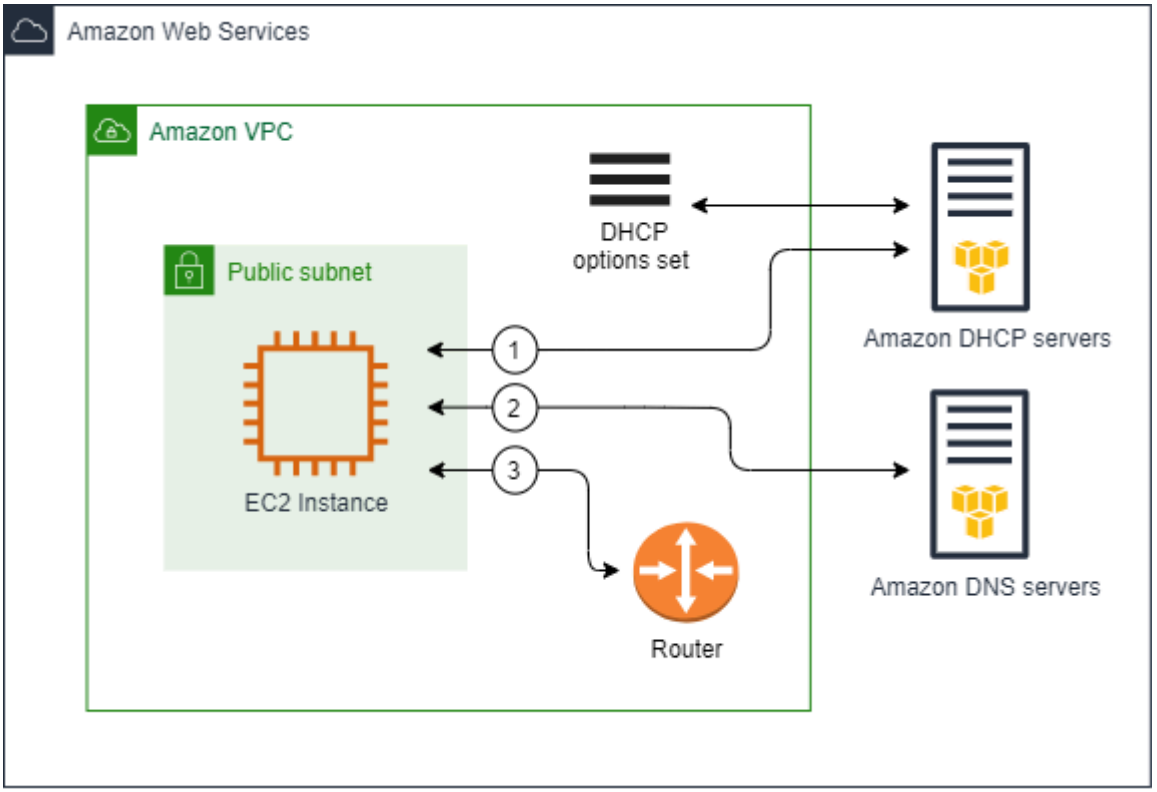
- **NetBIOS node type:** For EC2 instances running Windows, this is the method that the instances use to resolve NetBIOS names to IP addresses.

When you use the default option set, the Amazon DHCP server uses the network settings in the default option set. When you launch instances in your VPC, they do the following, as shown in the diagram:

(1) interact with the DHCP server

(2) interact with the Amazon DNS server

(3) connect to other devices in the network through the router for your VPC. The instances can interact with the Amazon DHCP server at any time to get their IP address lease and additional network settings.



Custom DHCP option set

You can create a custom DHCP option set with the following settings, and then associate it with a VPC:

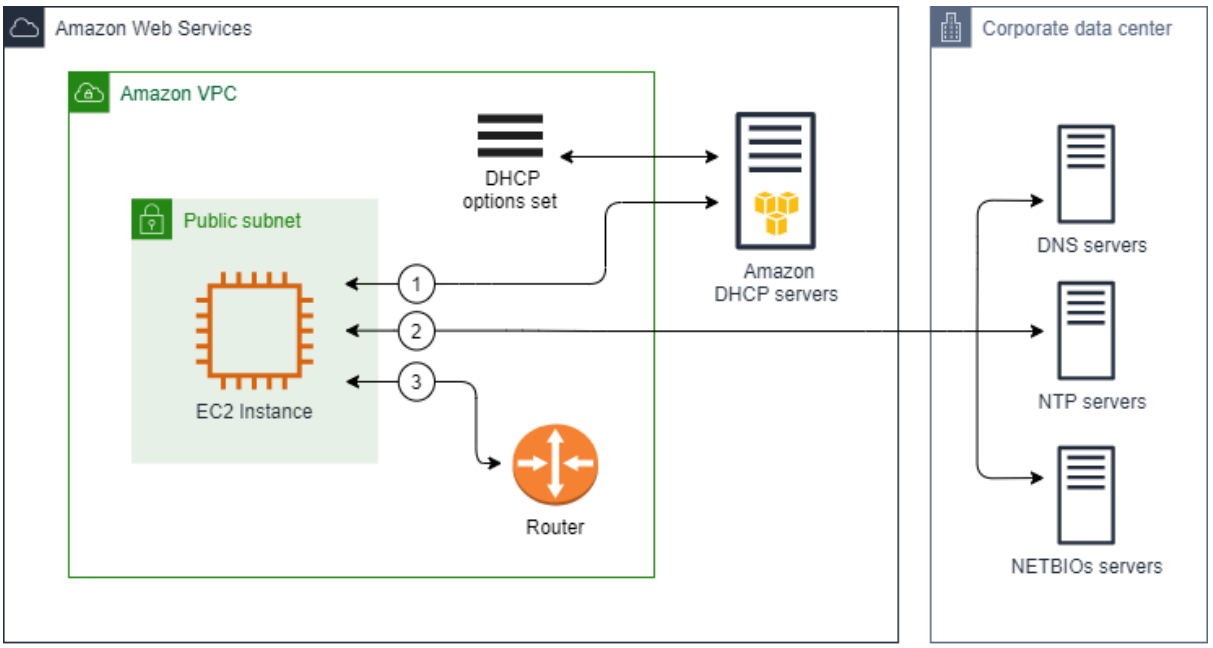
- **Domain name servers:** The DNS servers that your network interfaces use for domain name resolution.
- **Domain name:** The domain name that a client uses when resolving hostnames using the Domain Name System (DNS).
- **NTP servers:** The NTP servers that provide the time to the instances.
- **NetBIOS name servers:** For EC2 instances running Windows, the NetBIOS computer name is a friendly name assigned to the instance to identify it on the network. A NetBIOS name server maintains a list of mappings between NetBIOS computer names and network addresses for networks that use NetBIOS as their naming service.
- **NetBIOS node type:** For EC2 instances running Windows, the method that the instances use to resolve NetBIOS names to IP addresses.
- **IPv6 Preferred Lease Time (optional):** A value (in seconds, minutes, hours, or years) for how frequently a running instance with an IPv6 address assigned to it goes through DHCPv6 lease renewal. Acceptable values are between 140 and 4294967295 seconds (approximately 138 years). If no value is entered, the default lease time is 140 seconds. If you use long-term addressing for EC2 instances, you can increase the lease time and avoid frequent lease renewal requests. Lease renewal typically occurs when half of the lease time has elapsed.

When you use a custom option set, instances launched into your VPC do the following, as shown in the diagram:

(1) Use the network settings in the custom DHCP option set

(2) interact with the DNS, NTP, and NetBIOS servers specified in the custom DHCP option set

(3) Connect to other devices in the network through the router for your VPC.



DNS attributes for your VPC

Amazon provides a DNS server (the Amazon Route 53 Resolver) for your VPC. To use your own DNS server instead, create a new set of DHCP options for your VPC. For more information, see DHCP option sets in Amazon VPC.

Amazon DNS server, also known as the Route 53 Resolver. This DNS resolver service is natively integrated into each Availability Zone within your AWS Region, providing a reliable and scalable solution for domain name resolution within your Virtual Private Cloud (VPC).

Amazon DNS server

The Route 53 Resolver (also called "Amazon DNS server" or "AmazonProvidedDNS") is a DNS Resolver service that is built into each Availability Zone in an AWS Region.

The Route 53 Resolver is located at 169.254.169.253 (IPv4), fd00:ec2::253 (IPv6), and at the primary private IPV4 CIDR range provisioned to your VPC plus two.

For example

```
if you have a VPC with an IPv4 CIDR of 10.0.0.0/16 and an IPv6
CIDR of 2001:db8::/32.
you can reach the Route 53 Resolver at 169.254.169.253 (IPv4),
fd00:ec2::253 (IPv6), or 10.0.0.2 (IPv4).
```

Resources within a VPC use a link local address for DNS queries. These queries are transported to the Route 53 Resolver privately and are not visible on the network.

In an IPv6-only subnet, the IPv4 link-local address (169.254.169.253) is still reachable as long as "AmazonProvidedDNS" is the name server in the DHCP option set.

When you launch an instance into a VPC, AWS provide the instance with a private DNS hostname. AWS also provide a public DNS hostname if the instance is configured with a public IPv4 address and the ***VPC DNS attributes are enabled***.

The Amazon DNS server in your VPC is used to resolve the DNS domain names that you specify in a private hosted zone in Route 53. For more information about private hosted zones, see Working with private hosted zones in the Amazon Route 53 Developer Guide.

Rules and considerations

When using the Amazon DNS server, the following rules and considerations apply.

- You cannot filter traffic to or from the Amazon DNS server using network ACLs or security groups.
- Services that use the Hadoop framework, such as Amazon EMR, require instances to resolve their own fully qualified domain names (FQDN). In such cases, DNS resolution can fail if the domain-name-servers option is set to a custom value. To ensure proper DNS resolution, consider adding a conditional forwarder on your DNS server to forward queries for the domain *region-name.compute.internal* to the Amazon DNS server. For more information, see Setting up a VPC to host clusters in the *Amazon EMR Management Guide*.
- The Amazon Route 53 Resolver only supports recursive DNS queries.

DNS hostnames

When you launch an instance, it always receives a private IPv4 address and a private DNS hostname that corresponds to its private IPv4 address. If your instance has a public IPv4 address, the **DNS attributes for its VPC** determines whether it receives a public DNS hostname that corresponds to the public IPv4 address. For more information, see DNS attributes in your VPC.

With the Amazon provided DNS server enabled, DNS hostnames are assigned and resolved as follows.

`Private IP DNS name (IPv4 only)`

You can use the Private IP DNS name (IPv4 only) hostname **for** communication between instances **in** the same VPC.

You can resolve the Private IP DNS name (IPv4 only) hostnames of other instances **in** other VPCs as long as the instances are **in** the same AWS Region **and** the hostname of the other instance is **in** the private address space range defined by RFC 1918: 10.0.0.0 - 10.255.255.255 (10/8 prefix), 172.16.0.0 - 172.31.255.255 (172.16/12 prefix), **and** 192.168.0.0 - 192.168.255.255 (192.168/16 prefix).

Private resource DNS name

The RBN-based DNS name that can resolve **to** the A **and** AAAA DNS records selected **for** this instance.

This DNS hostname is visible **in** the instance details **for** instances **in** dual-stack **and** IPv6-only subnets.

For more information about RBN, see EC2 instance hostname types.

Public IPv4 DNS

A public (external) IPv4 DNS hostname takes the form ***ec2-public-ipv4-address.compute-1.amazonaws.com*** **for** the **us-east-1** Region, **and** **ec2-public-ipv4-address.region.compute.amazonaws.com** **for** other Regions.

The Amazon DNS server resolves a public DNS hostname **to** the public IPv4 address of the instance outside the network of the instance, **and to** the private IPv4 address of the instance **from** within the network of the instance.

DNS attributes in your VPC

The following VPC attributes determine the DNS support provided for your VPC. If both attributes are enabled, an instance launched into the VPC receives a public DNS hostname if it is assigned a public IPv4 address or an Elastic IP address at creation. If you enable both attributes for a VPC that didn't previously have them both enabled, instances that were already launched into that VPC receive public DNS hostnames if they have a public IPv4 address or an Elastic IP address.

To check whether these attributes are enabled for your VPC, see View and update DNS attributes for your VPC.

Attribute	Description
`enableDnsHostnames`	Determines whether the VPC supports assigning public DNS hostnames to instances with public IP addresses. The default for this attribute is `false` unless the VPC is a default VPC. Note the Rules and considerations for this attribute below.
`enableDnsSupport`	Determines whether the VPC supports DNS resolution through the Amazon provided DNS server. If this attribute is `true`, queries to the Amazon provided DNS server succeed. For more information, see Amazon DNS server.

	The default for this attribute is `true`. Note the Rules and considerations for this attribute below.

Rules and considerations

- If both attributes are set to true, the following occurs:
  - Instances with public IP addresses receive corresponding public DNS hostnames.
  - The Amazon Route 53 Resolver server can resolve Amazon-provided private DNS hostnames.

- If at least one of the attributes is set to false, the following occurs:
  - Instances with public IP addresses do not receive corresponding public DNS hostnames.
  - The Amazon Route 53 Resolver cannot resolve Amazon-provided private DNS hostnames.
  - Instances receive custom private DNS hostnames if there is a custom domain name in the DHCP options set.

- If you are not using the Amazon Route 53 Resolver server, your custom domain name servers must resolve the hostname as appropriate.
- If you use custom DNS domain names defined in a private hosted zone in Amazon Route 53, or use private DNS with interface VPC endpoints (AWS PrivateLink), you must set both the enableDnsHostnames and enableDnsSupport attributes to true.
- The Amazon Route 53 Resolver can resolve private DNS hostnames to private IPv4 addresses for all address spaces, including where the IPv4 address range of your VPC falls outside of the private IPv4 addresses ranges specified by RFC 1918. However, if you created your VPC before October 2016, the Amazon Route 53 Resolver does not resolve private DNS hostnames if your VPC's IPv4 address range falls outside of these ranges. To enable support for this, contact Support.
- If you use VPC peering, you must enable both attributes for both VPCs, and you must enable DNS resolution for the peering connection. For more information, see Enable DNS resolution for a VPC peering connection.

DNS quotas

There is a 1024 packet per second (PPS) limit to services that use link-local addresses. This limit includes the aggregate of Route 53 Resolver DNS queries, Instance Metadata Service (IMDS) requests, Amazon Time Service Network Time Protocol (NTP) requests, and Windows Licensing Service (for Microsoft Windows based instances) requests. This quota cannot be increased.

The number of DNS queries per second supported by Route 53 Resolver varies by the type of query, the size of the response, and the protocol in use. For more information and recommendations for a scalable DNS architecture, see the AWS Hybrid DNS with Active Directory Technical Guide.

If you reach the quota, the Route 53 Resolver rejects traffic. Some of the causes for reaching the quota might be a DNS throttling issue, or instance metadata queries that use the Route 53 Resolver network interface. For information about how to solve VPC DNS throttling issues, see How can I determine whether my DNS queries to the Amazon provided DNS server are failing due to VPC DNS throttling. For information about instance metadata retrieval, see Retrieve instance metadata in the *Amazon EC2 User Guide*.

Private hosted zones

To access the resources in your VPC using custom DNS domain names, such as `example.com`, instead of using private IPv4 addresses or AWS-provided private DNS hostnames, you can create a private hosted zone in Route 53.

A private hosted zone is a container that holds information about how you want to route traffic for a domain and its subdomains within one or more VPCs without exposing your resources to the internet.

You can then create Route 53 resource record sets, which determine how Route 53 responds to queries for your domain and subdomains.

For example,

If you want browser requests for example.com to be routed to a web server in your VPC, you'll create an A record in your private hosted zone and specify the IP address of that web server.

To access resources using custom DNS domain names, you must be connected to an instance within your VPC. From your instance, you can test that your resource in your private hosted zone is accessible from its custom DNS name by using the `ping` command; for example, `ping mywebserver.example.com`. (You must ensure that your instance's security group rules allow inbound ICMP traffic for the `ping` command to work.)

Private hosted zones do not support transitive relationships outside of the VPC; for example, you cannot access your resources using their custom private DNS names from the other side of a VPN connection.

Important

```
If you use custom DNS domain names defined in a private hosted
zone in Amazon Route 53, you must set both the enableDnsHostnames
and enableDnsSupport attributes to true.
```

Network Address Usage for your VPC

Network Address Usage (NAU) is a metric applied to resources in your virtual network to help you plan for and monitor the size of your VPC. Each NAU unit contributes to a total that represents the size of your VPC.

It's important to understand the total number of units that make up the NAU of your VPC because the following VPC quotas limit the size of a VPC:

- Network Address Usage – The maximum number of NAU units that a single VPC can have. Each VPC can have up to 64,000 NAU units by default. You can request a quota increase up to 256,000.

- Peered Network Address Usage – The maximum number of NAU units for a VPC and all of its peered VPCs. If a VPC is peered with other VPCs in the same Region, the VPCs combined can have up to 128,000 NAU units by default. You can request a quota increase up to 512,000. VPCs that are peered across different Regions do not contribute to this limit.

You can use the NAU in the following ways:

- Before you create your virtual network, calculate the NAU units to help you decide if you should spread workloads across multiple VPCs.
- After you've created your VPC, use Amazon CloudWatch to monitor the NAU usage of the VPC so that it doesn't grow beyond the NAU quota limits. For more information, see CloudWatch metrics for your VPCs.

How NAU is calculated

Resource	NAU units
Each private or public IPv4 and each IPv6 address assigned to a network interface for an EC2 instance in the VPC	1
Additional network interfaces attached to an EC2 instance	1
Prefix assigned to a network interface	1
Network Load Balancer per AZ	6
Gateway Load Balancer per AZ	6
VPC endpoint per AZ	6
Transit gateway attachment	6
Lambda function	6
NAT gateway	6
EFS mount target	6
EFA interface (EFA with an ENA device) or an EFA-only interface	1

NAU examples

The following examples show how to calculate NAU.

Example 1 - Two VPCs connected using VPC peering

Peered VPCs in the same Region contribute to a combined NAU quota.

- VPC 1
    - 50 Network Load Balancers in 2 subnets in separate Availability Zones - 600 NAU units
    - 5,000 instances (each with an IPv4 address and IPv6 address) in one subnet and 5,000 instances (each with an IPv4 address and IPv6 address) in another subnet - 20,000 units
    - 100 Lambda functions - 600 NAU units
- VPC 2
    - 50 Network Load Balancers in 2 subnets in separate Availability Zones - 600 NAU units
    - 5,000 instances (each with an IPv4 address and IPv6 address) in one subnet and 5,000 instances (each with an IPv4 address and IPv6 address) in another subnet - 20,000 units
    - 100 Lambda functions - 600 NAU units
- Total peering NAU count: 42,400 units
- Default peering NAU quota: 128,000 units

Example 2 - Two VPCs connected using a transit gateway

VPCs that are connected using a transit gateway do not contribute to a combined NAU quota as they do for peered VPCs.

- VPC 1
    - 50 Network Load Balancers in 2 subnets in separate Availability Zones - 600 NAU units

    - 5,000 instances (each with an IPv4 address and IPv6 address) in one subnet and 5,000 instances (each with an IPv4 address and IPv6 address) in another subnet - 20,000 units
    - 100 Lambda functions - 600 NAU units
- VPC 2
    - 50 Network Load Balancers in 2 subnets in separate Availability Zones - 600 NAU units
    - 5,000 instances (each with an IPv4 address and IPv6 address) in one subnet and 5,000 instances (each with an IPv4 address and IPv6 address) in another subnet - 20,000 units
    - 100 Lambda functions - 600 NAU units
- Total NAU count per VPC: 21,200 units
- Default NAU quota per VPC: 64,000 units

Share your VPC subnets with other accounts

VPC subnet sharing allows multiple AWS accounts to create their application resources.

The account that owns the VPC (owner) shares one or more subnets with other accounts (participants) that belong to the same organization from AWS Organizations.

You can simplify network topologies by interconnecting shared Amazon VPC subnets using connectivity features, such as AWS PrivateLink, transit gateways, and VPC peering. For more information about the benefits of VPC subnet sharing, see VPC sharing: A new approach to multiple accounts and VPC management