Building a Scalable and Secure Multi-VPC AWS Network Infrastructure

AWS Whitepaper

Building a Scalable and Secure Multi-VPC AWS Network Infrastructure: AWS Whitepaper

Copyright © 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

Building a Scalable and Secure Multi-VPC AWS Network Infrastructure

Publication date: *June 10, 2020 (last update (p. 41): February 2022)*

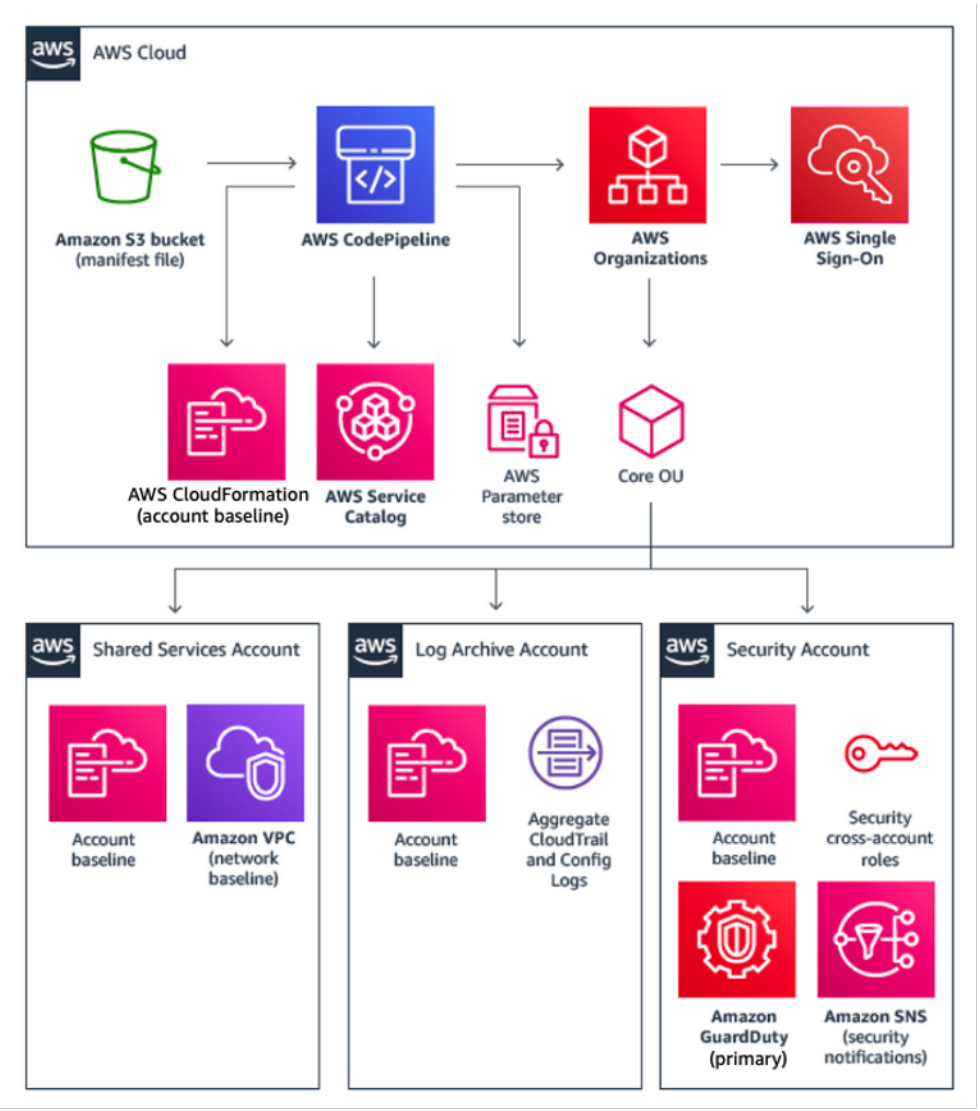
Abstract

Amazon Web Services (AWS) customers often rely on hundreds of accounts and virtual private clouds (VPCs) to segment their workloads and expand their footprint.This level of scale often creates challenges around resource sharing, inter-VPC connectivity, and on-premises facilities to VPC connectivity.

This whitepaper describes bestpractices for creating scalable and secure network architectures in a large network using AWS services such as Amazon Virtual Private Cloud (Amazon VPC), AWS Transit Gateway, AWS PrivateLink, AWS Direct Connect, Gateway Load Balancer, AWS Network Firewall, and Amazon Route 53. It demonstrates solutions for managing growing infrastructure—ensuring scalability, high availability, and security while keeping overhead costs low.

Introduction

AWS customers begin by building resources in a single AWS account that represents a management boundary which segments permissions, costs, and services. However, as the customer's organization grows, greater segmentation of services becomes necessary to monitor costs, control access, and provide easier environmental management. A multi-account solution solves these issues by providing specific accounts for IT services and users within an organization. AWS provides several tools to manage and configure this infrastructure, including AWS Landing Zone and AWS Control Tower.



Landing Zone account structure

AWS Landing Zone and AWS Control Tower automate the setup and integration of multiple AWS services to provide a baseline, highly controlled, multi-account environment with identity and access management (IAM), governance, data security, network design, and logging.

The AWS Landing Zone solution in the preceding figure includes four accounts:

- The **AWS Organizations** account (used to manage configuration and access to AWS Landing Zone managed accounts)
- The **Shared Services** account (used for creating infrastructure shared services such as directory services)
- The **Log Archive** account (centralized logging into Amazon Simple Storage Service (Amazon S3) buckets)
- The **Security** account (to be used by a company's security and compliance team to audit or perform emergency security operations in case of an incident in the spoke accounts).

> **Note**
> In this whitepaper, "Landing Zone" is a broad term for the scalable, secure, and performant multi-account/multi-VPC setup where you deploy your workloads. This setup can be built using any tool.

Most customers begin with a few VPCs to deploy their infrastructure. The number of VPCs a customer owns is usually related to their number of accounts, users, and staged environments (production, development, test, and so on). As cloud usage grows, the number of users, business units, applications, and Regions that a customer interacts with also grow, leading to the creation of new VPCs.

As the number of VPCs grows, cross-VPC management becomes essential for the operation of the customer's cloud network. This whitepaper covers best practices for three specific areas in cross-VPC and hybrid connectivity:

- **Network connectivity** – Interconnecting VPCs and on-premises networks at scale.
- **Network security** – Building centralized egress points for accessing the internet and endpoints such as network address translation (NAT) gateway, VPC endpoints , AWS PrivateLink , AWS Network Firewall , and Gateway Load Balancer.
- **DNS management** – Resolving DNS within the Landing Zone and hybrid DNS.

VPC to VPC connectivity

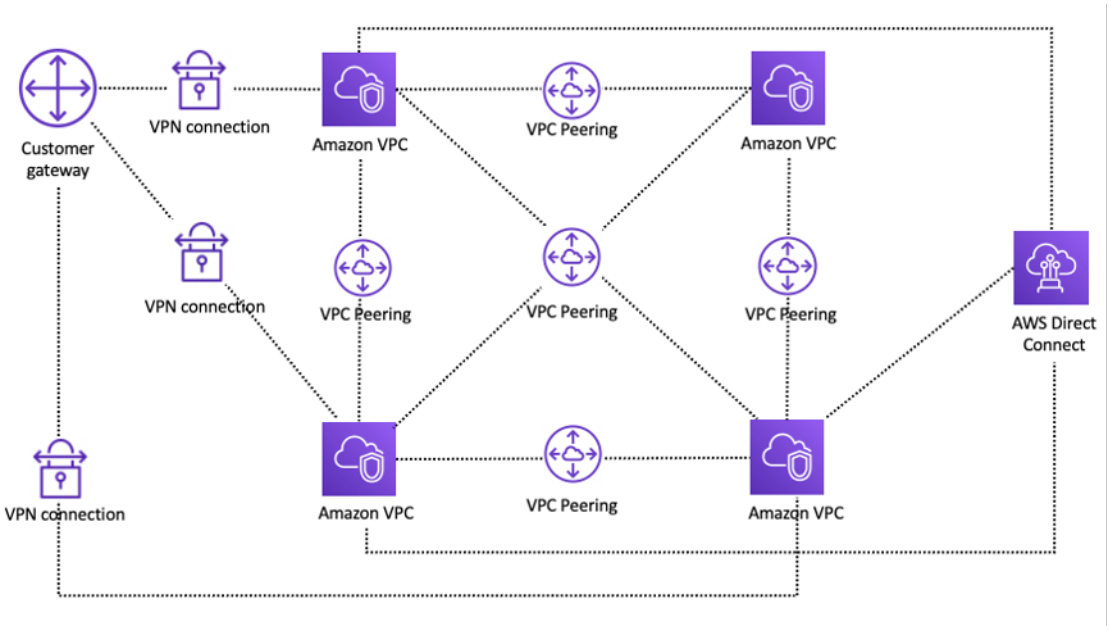
Customers can use two different VPC flow patterns to set up multi-VPC environments: *many-to-many*, or *hub-and-spoke*. In the many-to-many approach, the traffic between each VPC is managed individually between each VPC. In the hub-and-spoke model, all inter-VPC traffic flows through a central resource, which routes traffic based on established rules.

VPC peering

The simplest way to connect two VPCs is to use VPC Peering. In this setup, a connection enables full bidirectional connectivity between the VPCs. This peering connection is used to route traffic between the VPCs. VPCs across accounts and AWS Regions can also be peered together. VPC peering incurs costs only for traffic traveling over the connection (there is no hourly infrastructure fee).

VPC peering is point-to-point connectivity, and it does not support transitive routing. For example, if you have a VPC peering connection between VPC A and VPC B and between VPC A and VPC C, an instance in VPC B cannot transit through VPC A to reach VPC C. To route packets between VPC B and VPC C, you are required to create a direct VPC peering connection.

At scale, when you have tens to hundreds of VPCs, interconnecting them with peering results in a mesh of hundreds to thousands of peering connections, which are difficult to manage and scale. For example, if you have 100 VPCs and you want to setup a full mesh peering between them, it will take 4,950 peering connections [n(n-1)/2] where n=total number of VPCs. There is a maximum limit of up to 125 active peering connections per VPC.



Network setup using VPC peering

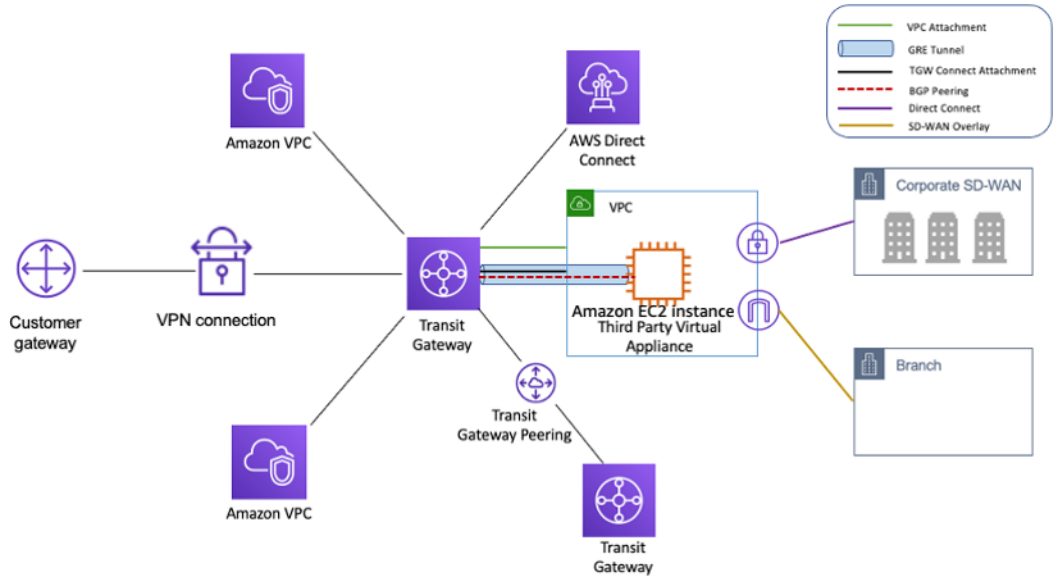
If you are using VPC peering, on-premises connectivity (VPN and/or Direct Connect) must be made to each VPC. Resources in a VPC cannot reach on-premises using the hybrid connectivity of a peered VPC, as shown in the preceding figure.

VPC peering is best used when resources in one VPC must communicate with resources in another VPC, the environment of both VPCs is controlled and secured, and the number of VPCs to be connected is less than 10 (to allow for the individual management of each connection). VPC peering offers the lowest overall cost when compared to other options for inter-VPC connectivity.

AWS Transit Gateway

AWS Transit Gateway provides a hub and spoke design for connecting VPCs and on-premises networks as a fully managed service without requiring you to provision virtual appliances like the Cisco CSRs. No VPN overlay is required, and AWS manages high availability and scalability.

Transit Gateway enables customers to connect thousands of VPCs. You can attach all your hybrid connectivity (VPN and Direct Connect connections) to a single Transit Gateway instance, consolidating and controlling your organization's entire AWS routing configuration in one place (refer to the figure under Transit VPC Solution). Transit Gateway controls how traffic is routed among all the connected spoke networks using route tables. This hub-and-spoke model simplifies management and reduces operational costs because VPCs only connect to the Transit Gateway instance to gain access to the connected networks.



Hub-and-spoke design with AWS Transit Gateway

Transit Gateway is a Regional resource and can connect thousands of VPCs within the same AWS Region. You can create multiple Transit Gateway instances per Region, and you can connect to a maximum of three Transit Gateway instances over a single Direct Connect connection for hybrid connectivity. Typically, you can use just one Transit Gateway instance connecting all your VPC instances in a given Region, and use Transit Gateway routing tables to isolate them wherever needed. There is a valid case for creating multiple Transit Gateway instances to limit misconfiguration blast radius, segregate control plane operations and administrative ease-of-use.

With Transit Gateway peering, customers can peer their Transit Gateway instances within same or multiple Regions and route traffic between them. It uses the same underlying infrastructure as VPC peering, and is therefore encrypted. For more information, refer to Building a global network using AWS Transit Gateway Inter-Region peering and AWS Transit Gateway now supports Intra-Region Peering.

Place your organization's Transit Gateway instance in its Network Services account. This enables centralized management by network engineers who manage the Network services account. Use AWS

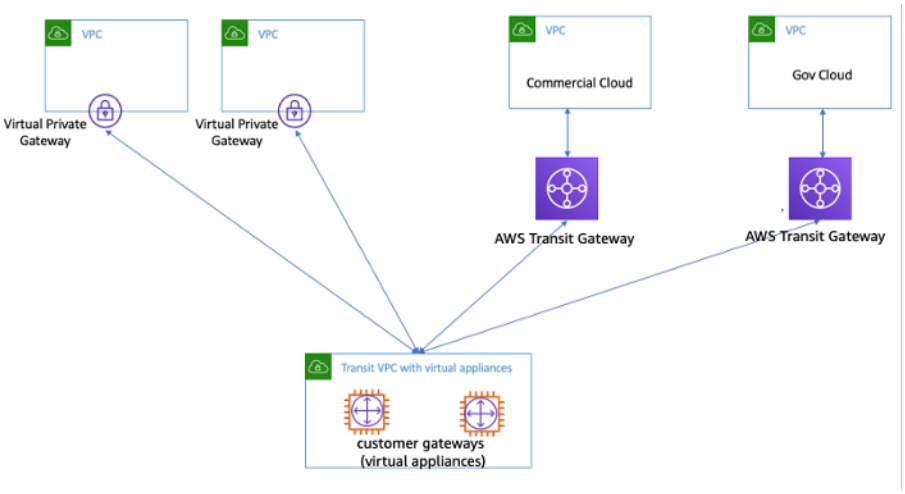
Resource Access Manager (RAM) to share a Transit Gateway instance for connecting VPCs across multiple accounts in your AWS Organization within the same Region. AWS RAM enables you to easily and securely share AWS resources with any AWS account, or within your AWS Organization. For more information, refer to the Automating AWS Transit Gateway attachments to a transit gateway in a central account blog post.

Transit Gateway also allows you to establish connectivity between SD-WAN infrastructure and AWS using Transit Gateway Connect. Use Transit GatewayConnect Border Gateway Protocol (BGP) for dynamic routing and Generic Routing Encapsulation (GRE) tunnel protocol for high performance, delivering up to 20 Gbps total bandwidth per Connect attachment (up to four Transit Gateway Connect peers per Connect attachment). By using Transit Gateway Connect, you can integrate both on-premises SD-WAN infrastructure or SD-WAN appliances running in the cloud through VPC attachment, or AWS Direct Connect attachment as the underlying transport layer. Refer to Simplify SD-WAN connectivity with AWS Transit Gateway Connect for reference architectures and detailed configuration.

Transit VPC solution

Transit VPCs can solve some of the shortcomings of VPC peering by introducing a hub and spoke design for inter-VPC connectivity. In a transit VPC network, one central VPC (the hub VPC) connects with every other VPC (spoke VPC) through a VPN connection typically leveraging BGP over IPsec. The central VPC contains Amazon Elastic Compute Cloud (Amazon EC2) instances running software appliances that route incoming traffic to their destinations using the VPN overlay. Transit VPC peering has the following advantages:

- Transitive routing is enabled using the overlay VPN network — allowing for a simpler hub and spoke design.
- When using third-party vendor software on the EC2 instance in the hub transit VPC, vendor functionality around advanced security (layer 7 firewall/Intrusion Prevention System (IPS)/Intrusion Detection System (IDS)) can be used. If customers are using the same software on-premises, they benefit from a unified operational/monitoring experience.
- The Transit VPC architecture enable connectivity that may be desired in some use-cases. For example, you can connect an AWS GovCloud instance and Commercial Region VPC or a Transit Gateway instance to a Transit VPC and enable inter-VPC connectivity between the two Regions. Evaluate your security and compliance requirements when considering this option. For additional security, you may deploy a centralized inspection model using design patterns described later in this whitepaper.



Transit VPC with virtual appliances

Transit VPC comes with its own challenges, such as higher costs for running third-party vendor virtual appliances on EC2 based on the instance size/family, limited throughput per VPN connection (up to 1.25 Gbps per VPN tunnel), and additional configuration, management and resiliency overhead (customers are responsible for managing the HA and redundancy of EC2 instances running the third-party vendor virtual appliances).

VPC peering vs. Transit VPC vs. Transit Gateway

Table 1 — Comparison between VPC peering, Transit VPC, and Transit Gateway

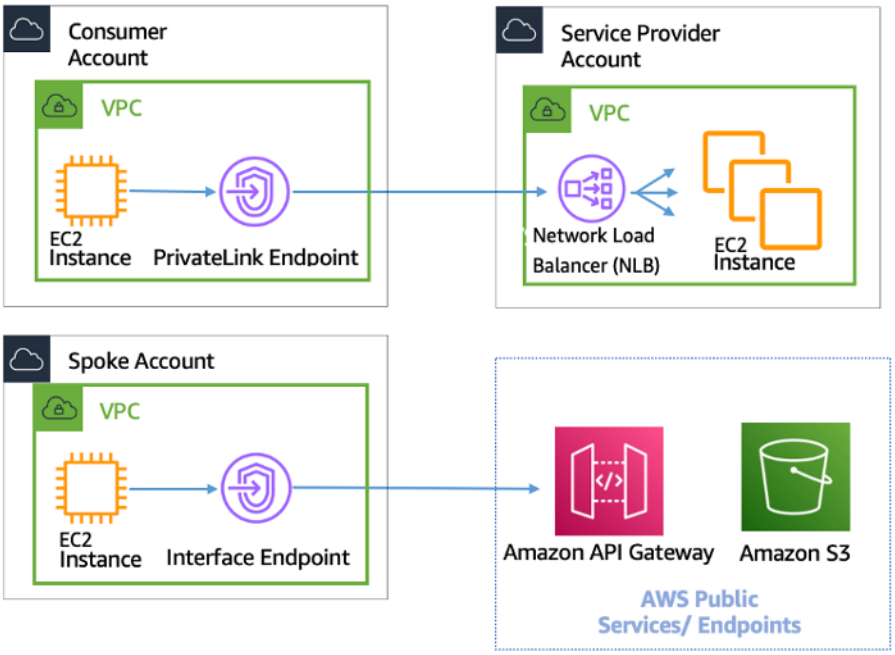
Criteria	VPC peering	Transit VPC	Transit Gateway
Architecture	Full mesh	VPN-based hub-and-spoke	Attachments-based hub-and-spoke. Can be peered with other TGWs.
Complexity	Increases with VPC count	Customer needs to maintain EC2 instance/HA	AWS managed service; increases with Transit Gateway count
Scale	125 active Peers/VPC	Depends on virtual router/EC2	5000 attachments per Region
Segmentation	Security groups	Customer managed	Transit Gateway route tables
Latency	Lowest	Extra, due to VPN encryption overhead	Additional Transit Gateway hop
Bandwidth limit	No limit	Subject to EC2 instance bandwidth limits based on size/family	Up to 50 Gbps (burst)/attachment
Visibility	VPC Flow Logs	VPC Flow Logs and CloudWatch Metrics	Transit Gateway Network Manager, VPC Flow Logs, CloudWatch Metrics
Security group cross-referencing	Supported	Not supported	Not supported
Cost	Data transfer	EC2 hourly cost, VPN tunnels cost and data transfer	Hourly per attachment, data processing, and data transfer

AWS PrivateLink

AWS PrivateLink provides private connectivity between VPCs, AWS services, and your on-premises networks without exposing your traffic to the public internet. AWS PrivateLink makes it easy to connect services across different accounts and VPCs to significantly simplify your network architecture. This allows customers who may want to privately expose a service/application residing in one VPC (service provider) to other VPCs (consumer) within an AWS Region in a way that only consumer VPCs initiate connections to the service provider VPC. An example of this is the ability for your private applications to access service provider APIs.

To use AWS PrivateLink, create a Network Load Balancer for your application in your VPC, and create a VPC endpoint service configuration pointing to that load balancer. A service consumer then creates an interface endpoint to your service. This creates an elastic network interface (ENI) in your subnet with a private IP address that serves as an entry point for traffic destined to the service. The consumer and service are not required to be in the same VPC. If the VPC is different, the consumer and service provider VPCs can have overlapping IP address ranges. In addition to creating the interface VPC endpoint to access services in other VPCs, you can create interface VPC endpoints to privately access supported AWS services through AWS PrivateLink, as shown in the following figure.

With Application Load Balancer (ALB) as target of NLB, you can now combine ALB advanced routing capabilities with AWS PrivateLink. Refer to Application Load Balancer-type Target Group for Network Load Balancer for reference architectures and detailed configuration.



AWS PrivateLink for connectivity to other VPCs and AWS Services

The choice between Transit Gateway, VPC peering, and AWS PrivateLink is dependent on connectivity.

- **AWS PrivateLink** — Use AWS PrivateLink when you have a client/server set up where you want to allow one or more consumer VPCs unidirectional access to a specific service or set of instances in the service provider VPC. Only the clients in the consumer VPC can initiate a connection to the service in the service provider VPC. This is also a good option when client and servers in the two VPCs have overlapping IP addresses as AWS PrivateLink uses ENIs within the client VPC in a manner that ensures that are no IP conflicts with the service provider. You can access AWS PrivateLink endpoints over VPC Peering, VPN, and AWS Direct Connect.
- **VPC peering and Transit Gateway** — Use VPC peering and Transit Gateway when you want to enable layer-3 IP connectivity between VPCs.

  Your architecture will contain a mix of these technologies in order to fulfill different use cases. All of these services can be combined and operated with each other. For example, AWS PrivateLink handling API style client-server connectivity, VPC peering for handling direct connectivity requirements where placement groups may still be desired within the Region or inter-Region connectivity is needed, and Transit Gateway to simplify connectivity of VPCs at scale as well as edge consolidation for hybrid connectivity.

VPC sharing

Sharing VPCs is useful when network isolation between teams does not need to be strictly managed by the VPC owner, but the account level users and permissions must be. With Shared VPC, multiple AWS accounts create their application resources (such as Amazon EC2 instances) in shared, centrally managed Amazon VPCs. In this model, the account that owns the VPC (owner) shares one or more subnets with other accounts (participants). After a subnet is shared, the participants can view, create, modify, and delete their application resources in the subnets shared with them. Participants cannot view, modify, or delete resources that belong to other participants or the VPC owner. Security between resources in shared VPCs is managed using security groups, network access control lists (NACLs), or through a firewall between the subnets.

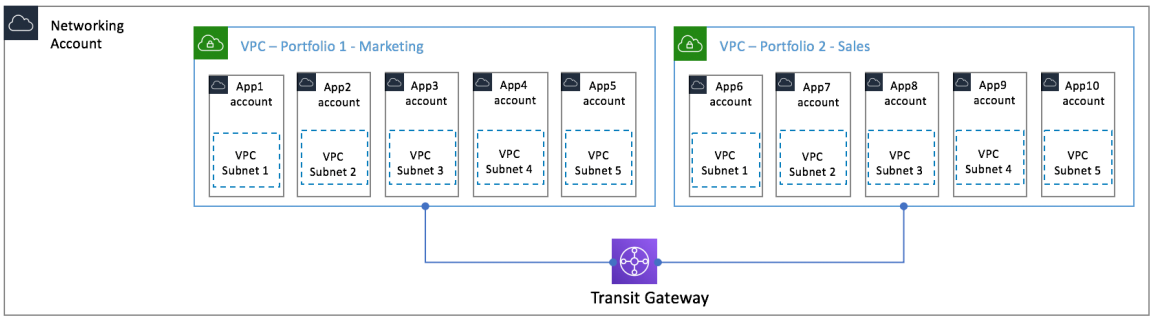
VPC sharing benefits:

- Simplified design — no complexity around inter-VPC connectivity
- Fewer managed VPCs
- Segregation of duties between network teams and application owners
- Better IPv4 address utilization
- Lower costs — no data transfer charges between instances belonging to different accounts within the same Availability Zone

> **Note**
> When you share a subnet with multiple accounts, your participants should have some level of cooperation since they're sharing IP space and network resources. If necessary, you can choose to share a different subnet for each participant account. One subnet per participant enables network ACL to provide network isolation in addition to security groups.

Most customer architectures will contain multiple VPCs, many of which will be shared with two or more accounts. Transit Gateway and VPC peering can be used to connect the shared VPCs. For example, suppose you have 10 applications. Each application requires its own AWS account. The apps can be categorized into two application portfolios (apps within the same portfolio have similar networking requirements, App 1–5 in 'Marketing' and App 6–10 in 'Sales').

You can have one VPC per application portfolio (two VPCs total), and the VPC is shared with the different application owner accounts within that portfolio. App owners deploy apps into their respective shared VPC (in this case, in the different subnets for network route segmentation and isolation using NACLs). The two shared VPCs are connected via the Transit Gateway. With this setup, you could go from having to connect 10 VPCs to just two, as seen in the following figure.



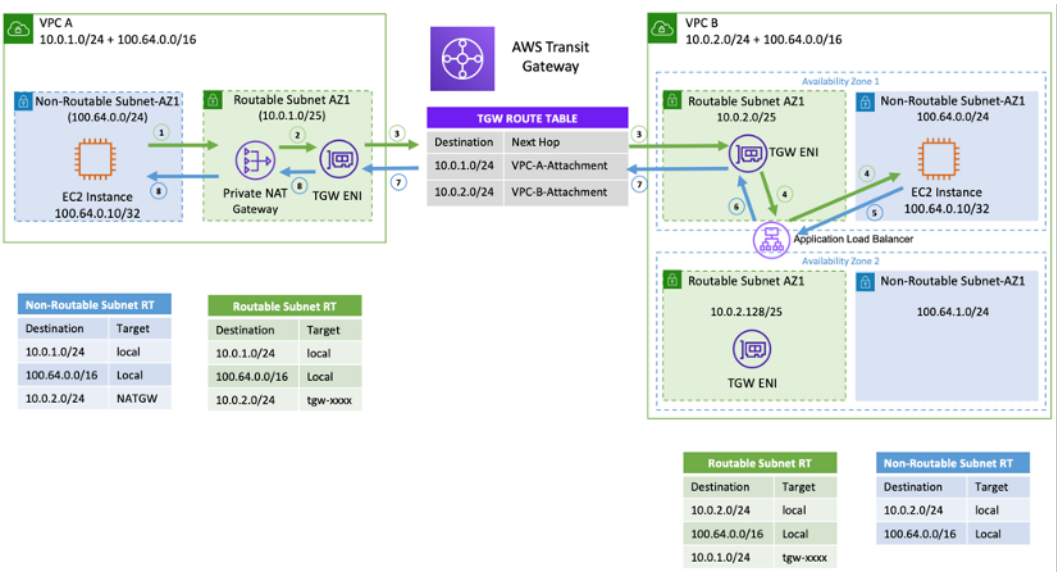
Example setup – shared VPC

> **Note**
> VPC sharing participants cannot create all AWS resources in a shared subnet. For more information, refer to the Limitations section in the VPC Sharing documentation.

For more information about the key considerations and best practices for VPC sharing, refer to the VPC sharing: key considerations and best practicesblog post.

Private NAT Gateway

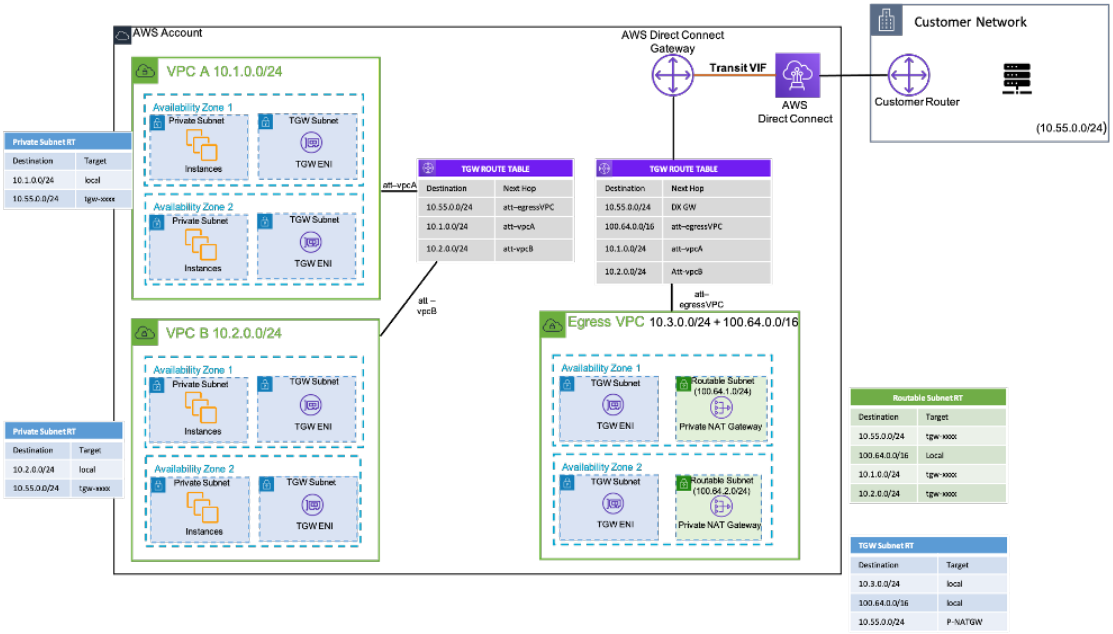
Teams often work independently and they may create a new VPC for a project, which may have overlapping classless inter-domain routing (CIDR) blocks. For integration, they might want to enable communication between networks with overlapping CIDRs, which is not achievable through features like VPC peering and Transit Gateway. A private NAT gateway can help with this use case. Private NAT gateway uses its private IP address to perform network address translation. You can route traffic from your private NAT gateway to other VPCs or on-premises network using Transit Gateway or virtual private gateway.



Example setup – Private NAT gateway

The preceding figure shows two non-routable (overlapping CIDRs, 100.64.0.0/16) subnets in VPC A and B. To establish connection between them, you can add a secondary non-overlapping / routable CIDRs (routable subnets, 10.0.1.0/24 and 10.0.2.0/24) to VPC A & B respectively. Routable CIDR should be allocated by network management team responsible for IP allocation. A private NAT gateway is added to the routable subnet in VPC A. Private NAT gateway performs source network address translation on requests from instance in non-routable subnet of VPC A to instance behind Application Load Balancer (ALB) in non-routable subnet of VPC. Traffic is then routed via Transit Gateway.

The private NAT gateway can also be used when your on-premise network restricts access to approved IPs. The on-premises networks of few customers are required by compliance to communicate only with private networks (no IGW) only through a limited contiguous block of approved IPs owned by the customer. Instead of allocating each instance a separate IP from the block, you can now run large workloads on AWS VPCs behind each allow listed IP using private NAT gateway. For details, refer to the How to solve Private IP exhaustion with Private NAT Solution blog post.



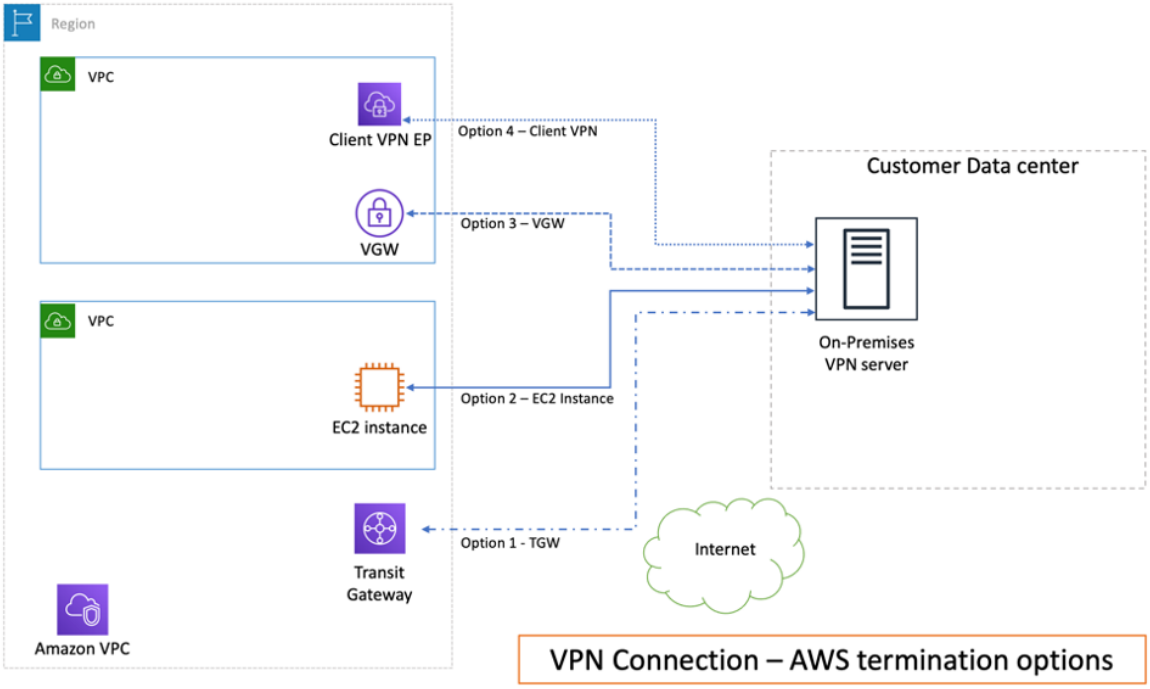
Example setup – How to use Private NAT gateway to provide approved IPs for on-premises network

Hybrid connectivity

This section focuses on securely connecting your cloud resources with your on-premises data centers. There are two approaches for enabling hybrid connectivity:

1. **One-to-one connectivity** — In this setup, a VPN connection and/or Direct Connect private VIF is created for every VPC. This is accomplished by using the virtual private gateway (VGW). This option is great for small numbers of VPCs, but as a customer scales their VPCs, managing hybrid connectivity per VPC can become difficult.
2. **Edge consolidation** — In this setup, customers consolidate hybrid IT connectivity for multiple VPCs at a single endpoint. All the VPCs share these hybrid connections. This is accomplished by using AWS Transit Gateway and the Direct Connect Gateway.

VPN



AWS VPN options

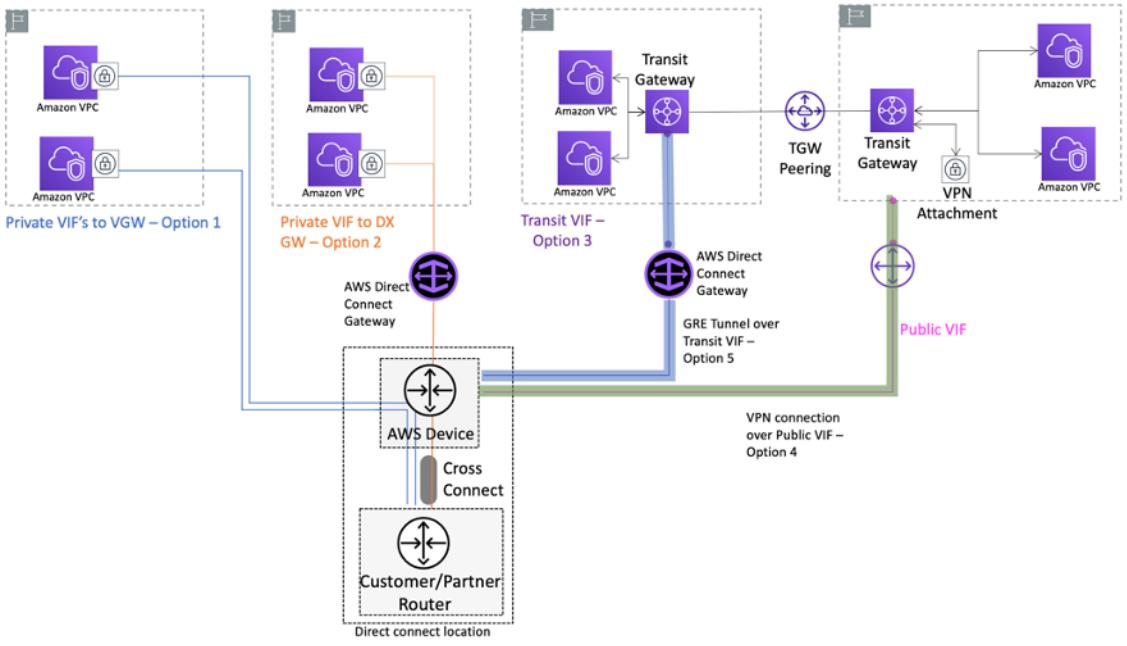
There are four ways to set up VPN to AWS:

1. **Consolidate VPN connectivity on Transit Gateway** — This option leverages the Transit Gateway VPN attachment on Transit Gateway. Transit Gateway supports IPsec termination for site-to-site VPN. Customers can create VPN tunnels to the Transit Gateway, and can access the VPCs attached to it.Transit Gateway supports both Static and BGP-based Dynamic VPN connections. Transit Gateway also supports Equal-Cost Multi-Path (ECMP) on VPN attachments. Each VPN connection has a maximum of 1.25-Gbps throughput per tunnel. Enabling ECMP allows you to aggregate throughput across VPN connections, allowing to scale beyond the default maximum limit of 1.25 Gbps. In this option, you pay for Transit Gateway pricing as well as AWS VPN pricing. AWS recommendeds using this option for VPN connectivity. For more information, refer to the Scaling VPN throughput using AWS Transit Gateway blog post.

2. **Terminate VPN on EC2 instance** — This option is leveraged by customers in edge cases when they want a particular vendor software feature set (such as Cisco DMVPN or Generic Routing Encapsulation (GRE)), or they want operational consistency across various VPN deployments. You can use the transit VPC design for edge consolidation, but it is important to remember that all the key considerations from the *VPC to VPC connectivity* (p. 4) section for transit VPC are applicable to hybrid VPN connectivity. You are responsible for managing high availability, and you pay for EC2 instance as well as any vendor software licensing and support costs.
3. **Terminate VPN on a virtual private gateway (VGW)** — This AWS Site-to-Site VPN service option enables a one-to-one connectivity design where you create one VPN connection (consisting of a pair of redundant VPN tunnels) per VPC. This is great way to get started with VPN connectivity into AWS, but as you scale the number of VPCs, managing a growing number of VPN connections can become challenging. Therefore, edge consolidation design leveraging Transit Gateway will eventually be a better option. VPN throughput to a VGW is limited to 1.25 Gbps per tunnel and ECMP load balancing is not supported. From a pricing perspective, you only pay for AWS VPN pricing, there is no charge for running a VGW. For more information, refer to AWS VPN Pricing and AWS VPN on virtual private gateway.
4. **Terminate VPN connection on client VPN endpoint** — AWS Client VPN is a managed client-based VPN service that enables you to securely access your AWS resources and resources in your on-premises network. With Client VPN, you can access your resources from any location using an OpenVPN or AWS provided VPN client. By setting up a Client VPN endpoint, clients and users can connect to establish a Transport Layer Security (TLS) VPN connection. For more information, refer to the AWS Client VPN documentation.

AWS Direct Connect

While VPN over internet is a great option to get started, internet connectivity may not be reliable for production traffic. Because of this unreliability, many customers choose AWS Direct Connect. AWS Direct Connect is a networking service that provides an alternative to using the internet to connect to AWS. Using AWS Direct Connect, data that would have previously been transported over the internet is delivered through a private network connection between your facilities and AWS. In many circumstances, private network connections can reduce costs, increase bandwidth, and provide a more consistent network experience than internet-based connections. There are five ways to use AWS Direct Connect to connect to VPCs:



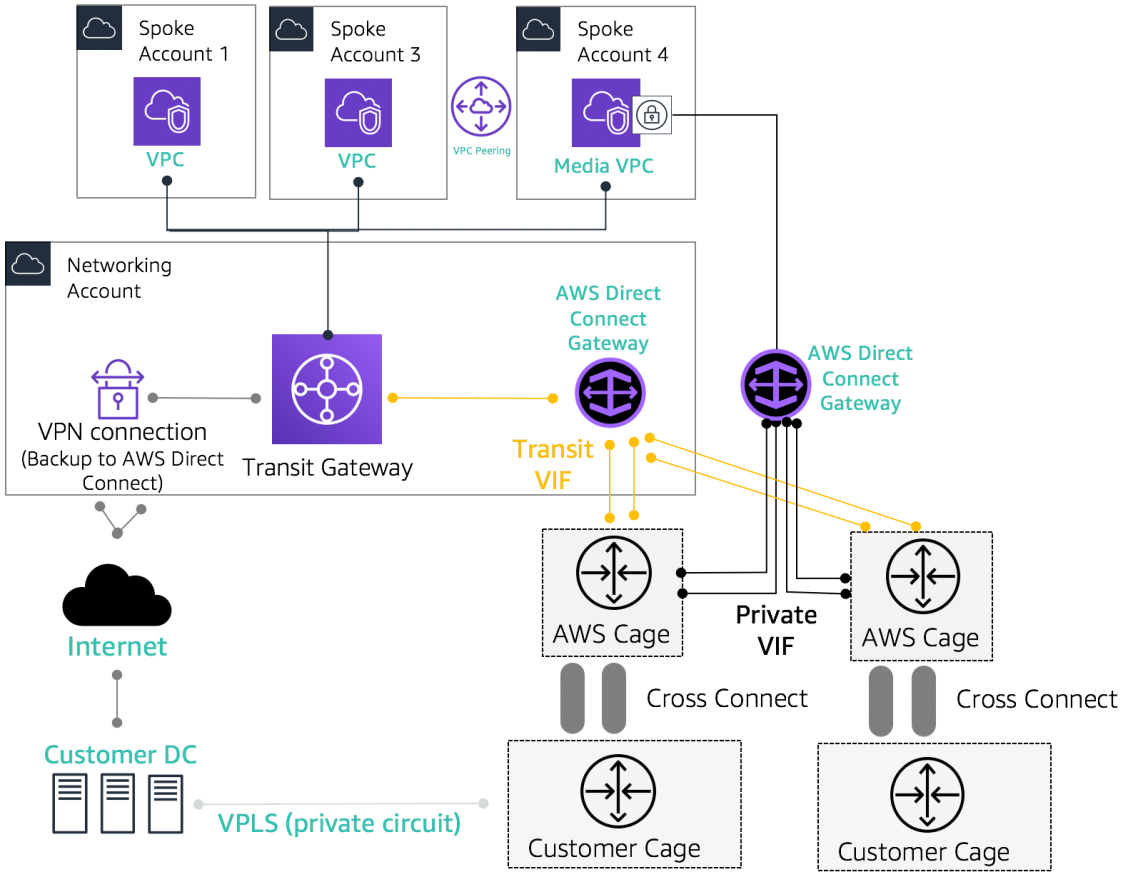
Ways to connect your on-premises data centers using AWS Direct Connect

1. **Create a private virtual interface (VIF) to a VGW attached to a VPC** — You can create 50 VIFs per Direct Connect connection, allowing you to connect to a maximum of 50 VPCs (one VIF provides connectivity to one VPC). There is one BGP peering per VPC. Connectivity in this setup is restricted to the AWS Region that the Direct Connect location is homed to. The one-to-one mapping of VIF to VPC (and lack of global access) makes this the least preferred way to access VPCs in the Landing Zone.
2. **Create a private VIF to a Direct Connect gateway associated with multiple VGWs (each VGW is attached to a VPC) —** A Direct Connect gateway is a globally available resource. You can create the Direct Connect gateway in any Region and access it from all other Regions (except China). A Direct Connect Gateway can connect to up to 10 VPCs (via VGWs) globally in any AWS account over a single private VIF. This is a great option if a Landing Zone consists of a small number of VPCs (ten or fewer VPCs) and/or you need global access. There is one BGP peering session per Direct Connect Gateway per Direct Connect connection. Direct Connect gateway is only for north/south traffic flow and does not permit VPC-to-VPC connectivity. Refer to Virtual private gateway associations in the AWS Direct Connect documentation for more details. With this option, the connectivity is not restricted to the AWS Region where the Direct Connect location is homed to.
3. **Create a transit VIF to a Direct Connect gateway associated with Transit Gateway –** You can associate a Transit Gateway instance to a Direct Connect gateway via a Transit VIF, which is only available over dedicated or hosted connections with speeds of one Gbps or greater. This option allows you to connect your on-premises data center to up to three Transit Gateway instances per Direct Connect Gateway (which can connect to thousands of VPCs) across different AWS Regions and AWS accounts over single transit VIF and BGP peering. This is the simplest setup among the four options for connecting multiple VPCs at scale, but you should be mindful of the Transit Gateway quotas. One key limit to note is that you can advertise only 20 prefixes from a Transit Gateway to on-premises router over the transit VIF. With the previous two options, you pay for Direct Connect pricing. For this option, you also pay for the Transit Gateway attachment and data processing charges. For more information, refer to the Transit Gateway Associations on Direct Connect documentation.
4. **Create a VPN connection to Transit Gateway over Direct Connect public VIF –** A public VIF allows you to access all AWS public services and endpoints using the public IP addresses. When you create a VPN attachment on a Transit Gateway, you get two public IP addresses for VPN endpoints at the AWS side. These public IPs are reachable over the public VIF. You can create as many VPN connections to as many Transit Gateway instances as you want over Public VIF. When you create a BGP peering over the public VIF, AWS advertises the entire AWS public IP range to your router. To ensure that you only permit certain traffic (for example, allowing traffic only to the VPN termination endpoints) you are advised to use a firewall on-premises facilities. This option can be used to encrypt your Direct Connect at the network layer.
5. **Create GRE tunnels to Transit Gateway over a transit VIF** – The Transit Gateway Connect attachment type supports GRE. With Transit Gateway Connect, SD-WAN infrastructure can be natively connected to AWS without having to set up IPsec VPNs between SD-WAN network virtual appliances and Transit Gateway. The GRE tunnels can be established over a transit VIF, having Transit Gateway Connect as the attachment type, providing higher bandwidth performance compared to a VPN connection. For more information, refer to the Simplify SD-WAN connectivity with AWS Transit Gateway Connect blog post.

While the “transit VIF to Direct Connect gateway” option might appear to be the best option because it lets you consolidate all your on-premises connectivity for a given AWS Region at a single point, (Transit Gateway) using a single BGP session per Direct Connect connection. Given some of the limits and considerations around this option, AWS expects customers to use both the “Create a private VIF” option and the “Transit VIF to Direct Connect gateway” option for their Landing Zone connectivity requirements. The following figure illustrates a sample setup where Transit VIF is used as a default method for connecting to VPCs, and a private VIF is used for an edge use case where huge amount of data must be transferred from an on-premises Data Center to the media VPC. Private VIF is used to avoid Transit Gateway data processing charges. As a best practice, you should have at least two connections at two different Direct Connect locations for maximum redundancy—a total of four connections. You create

one VIF per connection for a total of four private VIFs and four transit VIFs. You can also create a VPN as a backup connectivity to AWS Direct Connect connections.

With the "Create GRE tunnels to Transit Gateway over a transit VIF" option, you get the capability to natively connect your SD-WAN infrastructure with AWS. It eliminates the need to setup IPsec VPNs between SD-WAN network virtual appliances and Transit Gateway.



Sample reference architecture for hybrid connectivity

Use the Network Services account for creating Direct Connect resources enabling demarcation of network administrative boundaries. Direct Connect connection, Direct Connect gateway and Transit Gateway can all reside in a Network Services account. To share the AWS Direct Connect connectivity with your Landing Zone, simply share the Transit Gateway through AWS RAM with other accounts.

MACsec security on Direct Connect connections

Customers can use MAC Security Standard (MACsec) encryption (IEEE 802.1AE) with their Direct Connect connections for 10 Gbps and 100 Gbps dedicated connections at select locations. With this capability, customers can secure their data on the layer 2 level, and Direct Connect delivers point-to-point encryption as well. To enable the Direct Connect MACsec feature, ensure that the MACsec pre-requisites are met. Because MACsec protects links on a hop-by-hop basis, you must have a dedicated connection that is transparent to layer 2 traffic and therefore MACsec compatible. Your device must have a direct layer 2 adjacency with our Direct Connect device. Your last-mile provider can help you verify that your connection will work with MACsec. For more information refer to Adding MACsec security to AWS Direct Connect connections.

AWS Direct Connect resiliency recommendations

With AWS Direct Connect, customers can achieve highly resilient connectivity into their Amazon VPCs and AWS resources from their on-premises networks. It is best practice that customers connect from multiple data centers, to eliminate any single point physical location failures. It is also recommended that, depending on the type of workloads, customers utilize more than one Direct Connect circuit for redundancy.

AWS also offers the AWS Direct Connect Resiliency Toolkit, which provides customers with a connection wizard with multiple redundancy models; to help them determine which model works best for their service level agreement (SLA) requirements and design their hybrid connectivity using Direct Connect connections accordingly. For more information, refer to AWS Direct Connect Resiliency Recommendations.

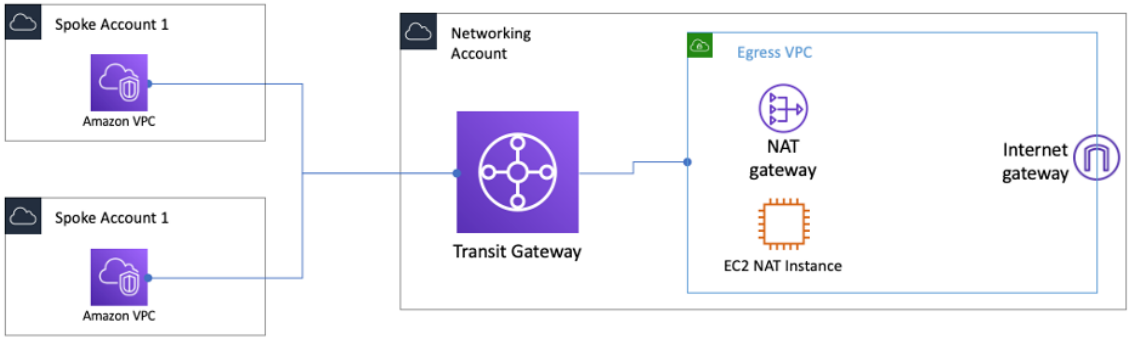
Centralized egress to internet

As you deploy applications in your Landing Zone, many apps will require outbound only internet access (for example, downloading libraries, patches, or OS updates). You can achieve this (preferable) by using a network address translation (NAT) gateway, or alternatively, an Amazon EC2 NAT instance as the next hop for all egress internet access. Internal applications reside in private subnets, while NAT gateway/ EC2 NAT instances reside in a public subnet. AWS recommends that you use NAT gateways because they provide better availability and bandwidth and require less effort on your part to administer. For more information, refer to Compare NAT gateways and NAT instances.

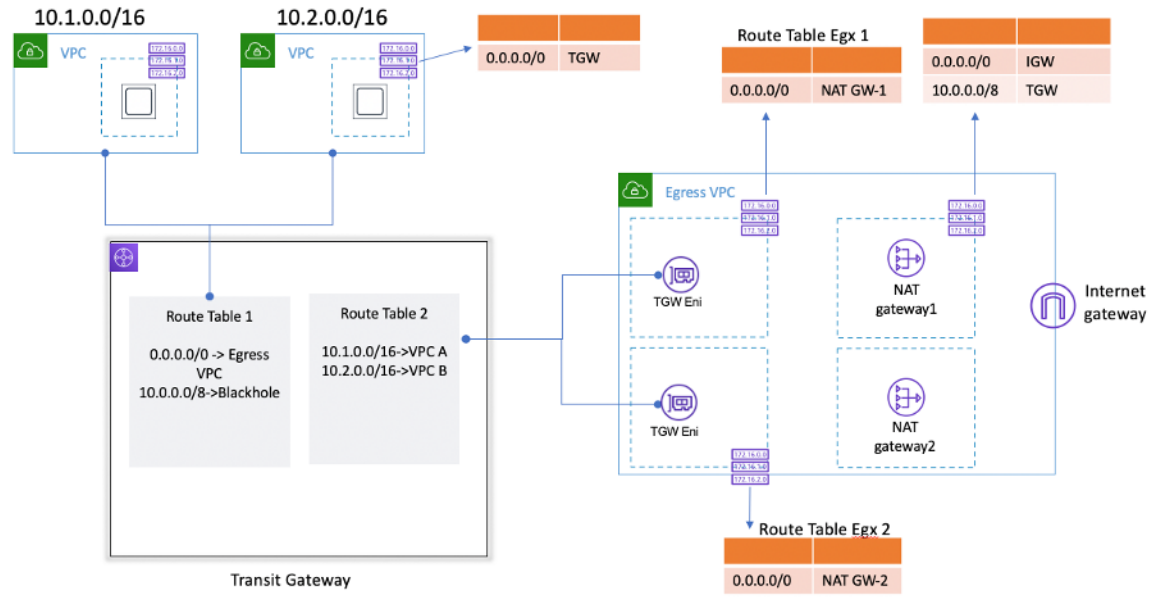
Using the NAT gateway for centralized egress

NAT gateway is a managed network address translation service. Deploying a NAT gateway in every spoke VPC can become cost prohibitive because you pay an hourly charge for every NAT gateway you deploy (refer to Amazon VPC pricing), so centralizing it could be a viable option. To centralize, you create a separate egress VPC in the network services account and route all egress traffic from the spoke VPCs via a NAT gateway sitting in this VPC using Transit Gateway, as shown in the following figure.

Note
When you centralize NAT gateway using Transit Gateway, you pay an extra Transit Gateway data processing charge — compared to the decentralized approach of running a NAT gateway in every VPC. In some edge cases when you send huge amounts of data through NAT gateway from a VPC, keeping the NAT local in the VPC to avoid the Transit Gateway data processing charge might be a more cost-effective option.



Centralized NAT gateway using Transit Gateway (overview)



Centralized NAT gateway using Transit Gateway (route table design)

In this setup, spoke VPC attachments are associated with Route Table 1 (RT1) and are propagated to Route Table 2 (RT2). There is a Blackhole route to disallow the two VPCs from communicating with each other. If you want to allow inter-VPC communication, you can remove the '10.0.0.0/8 -> Blackhole' route entry from RT1. This allows them to communicate via the NAT gateway. You can also propagate the spoke VPC attachments to RT1 (or alternatively, you can use one route table and associate/propagate everything to that), enabling direct traffic flow between the VPCs using Transit Gateway.

You add a static route in RT1 pointing all traffic to egress VPC. Because of this static route, Transit Gateway sends all internet traffic through its ENIs in the egress VPC. Once in the egress VPC, traffic follows the routes defined in the subnet route table where these Transit Gateway ENIs are present. You add a route in subnet route tables pointing all traffic towards the respective NAT gateway in the same Availability Zone to minimize cross-Availability Zone (AZ) traffic. The NAT gateway subnet route table has internet gateway (IGW) as the next hop. For return traffic to flow back, you must add a static route table entry in the NAT gateway subnet route table pointing all spoke VPC bound traffic to Transit Gateway as the next hop.

High availability

For high availability, you should use two NAT gateways (one in each Availability Zone). Within an Availability Zone, the NAT gateway has an availability SLA of 99.9%. Redundancy against component failure within an AZ is handled by AWS under the SLA agreement. Traffic is dropped during the 0.1% time when the NAT gateway may be unavailable in an Availability Zone. If one Availability Zone entirely fails, the Transit Gateway endpoint along with NAT gateway in that Availability Zone will fail, and all traffic will flow via the Transit Gateway and NAT gateway endpoints in the other Availability Zone.

Security

You rely on security groups on the source instances, blackhole routes in the Transit Gateway route tables, and the network ACL of the subnet in which the NAT gateway is located.

Scalability

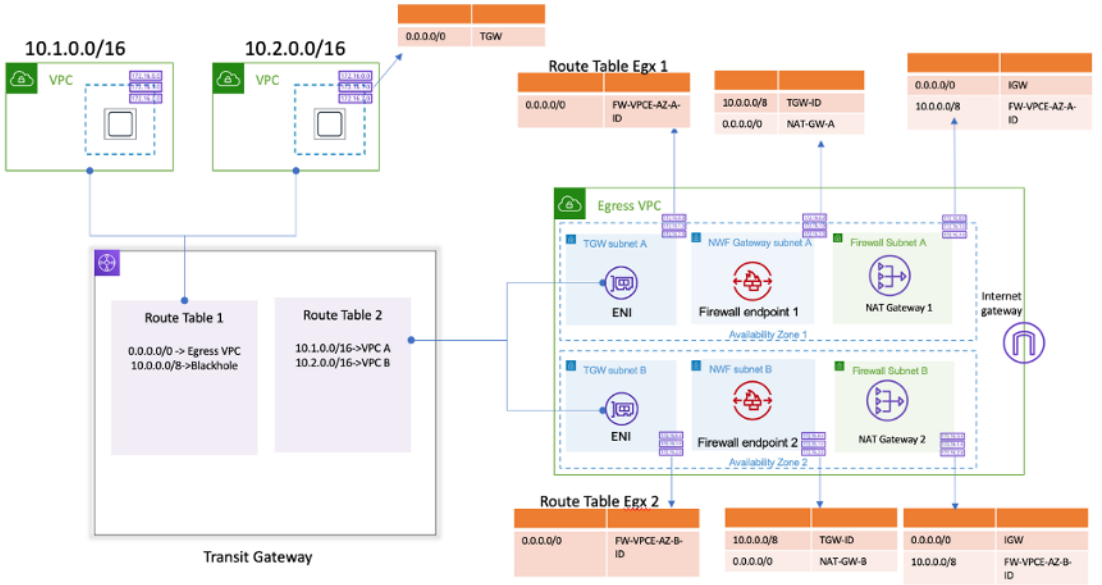
A NAT gateway can support up to 55,000 simultaneous connections to each unique destination. From a throughput standpoint, NAT gateway can scale from five Gbps to 45 Gbps. Transit Gateway generally does not act as a load balancer and would not distribute your traffic evenly across NAT gateway in the multiple Availability Zones. The traffic across the Transit Gateway will stay within an Availability Zone, if possible. If the Amazon EC2 instance initiating traffic is in Availability Zone 1, traffic will flow out of the Transit Gateway elastic network interface in the same Availability Zone 1 in the egress VPC and will flow to the next hop based on that subnet route table that elastic network interface resides in. However, to distribute the traffic evenly across both the Availability Zones, you can enable Appliance Mode on the VPC attachment connected to Egress VPC. For a complete list of rules, refer to NAT gateways in the Amazon Virtual Private Cloud documentation.

For more information, refer to the Creating a single internet exit point from multiple VPCs Using AWS Transit Gateway blog post.

Using the NAT gateway with AWS Network Firewall for centralized egress

If you want to inspect and filter your outbound traffic, you can incorporate AWS Network Firewall with NAT gateway in your centralized egress architecture. AWS Network Firewall is a managed service that makes it easy to deploy essential network protections for all of your VPCs. It provides control and visibility to Layer 3-7 network traffic for your entire VPC. You can do URL/domain name, IP address, and content-based outbound traffic filtering to stop possible data loss, help meet compliance requirements, and block known malware communications. AWS Network Firewall supports thousands of rules that can filter out network traffic destined for known bad IP addresses or bad domain names. You can also use Suricata IPS rules as part of the AWS Network Firewall service by importing open-source rulesets or authoring your own Intrusion Prevention System (IPS) rules using Suricata rule syntax. AWS Network Firewall also allows you to import compatible rules sourced from AWS partners.

In the centralized egress architecture with inspection, NAT gateway is a default route table target in the private subnet route table for the Egress VPC. Traffic between NAT gateway and the internet is inspected using AWS Network Firewall as shown in the following diagram. Additionally, Transit Gateway with AWS Network Firewall maintains symmetric routing to the same zonal firewall endpoint.



Centralized egress with AWS Network Firewall and NAT gateway (route table design)

For centralized deployment model with Transit Gateway, AWS recommends deploying AWS Network Firewall endpoints in two Availability Zones by allocating two subnets mapped to separate Availability Zones, as shown in the preceding diagram. As a best practice, do not use AWS Network Firewall subnet to deploy any other services because AWS Network Firewall is not able to inspect traffic from sources or destinations within a firewall subnet.

Similar to the previous setup, spoke VPC attachments are associated with Route Table 1 (RT1) and are propagated to Route Table 2 (RT2). A Blackhole route is explicitly added to disallow the two VPCs from communicating with each other.

Continue to use a default route in RT1 pointing all traffic to egress VPC. Transit Gateway will forward all traffic flows to one of the two availability zones in the egress VPC. Once traffic reaches one of Transit Gateway ENIs in the egress VPC, you hit a default route which will forward traffic to one of the AWS Network Firewall endpoints in their respective availability zone. AWS Network Firewall will then inspect traffic based on the rules you set before forwarding traffic to the NAT gateway using a default route.

This case doesn't require Transit Gateway appliance mode, because you aren't sending traffic between attachments.

> **Note**
> AWS Network Firewall does not perform network address translation for you, this function would be handled by the NAT gateway after traffic inspection through the AWS Network Firewall. Ingress routing is not required in this case as return traffic will be forwarded to the NATGW IPs by default.

Because you are using a Transit Gateway, here we can place the firewall prior to NAT gateway. In this model, the firewall can see the source IP behind the Transit Gateway.

If you were doing this in a single VPC, we can use the VPC routing enhancements that allow you to inspect traffic between subnets in the same VPC. For details, refer to the Deployment models for AWS Network Firewall with VPC routing enhancements blog post.

Scalability

AWS Network Firewall can automatically scale firewall capacity up or down based on traffic load to maintain steady, predictable performance to minimize costs. AWS Network Firewall is designed to support tens of thousands of firewall rules and can scale up to 45 Gbps throughput per Availability Zone.

Key considerations

- Each firewall endpoint can handle about 45 Gbps of traffic, if you require higher burst or sustained throughput, contact AWS support.
- If you choose to create a NAT gateway in your AWS account along with Network Firewall, standard NAT gateway processing and per-hour usage charges are waived on a one-to-one basis with the processing per GB and usage hours charged for your firewall.
- You can also consider distributed firewall endpoints through AWS Firewall Manager without a Transit Gateway.
- Test firewall rules before moving them to production, similar to a network access control list, as order matters.
- Advanced Suricata rules are required for deeper inspection. Currently network firewall does not support encrypted traffic inspection. If you have this requirement, consider using GWLB with third party inspection devices.
- The `HOME_NET` rule group variable defined the source IP range eligible for processing in the Stateful engine. Using a centralized approached, you must add all additional VPC CIDRs attached to the Transit Gateway to make them eligible for processing. Refer to the Network Firewall documentation for more details on the `HOME_NET` rule group variable.
- Consider deploying Transit Gateway and egress VPC in a separate Network Services account to segregate access based on delegation of duties; for example, only network administrators can access the Network Services account.

Using the NAT gateway and Gateway Load Balancer with EC2 instances for centralized egress

Using a software-based virtual appliance (on Amazon EC2) from AWS Marketplace and AWS Partner Network as an exit point is similar to the NAT gateway setup. This option can be used if you want to use the advanced layer 7 firewall/Intrusion Prevention/Detection System (IPS/IDS) and deep packet inspection capabilities of the various vendor offerings.

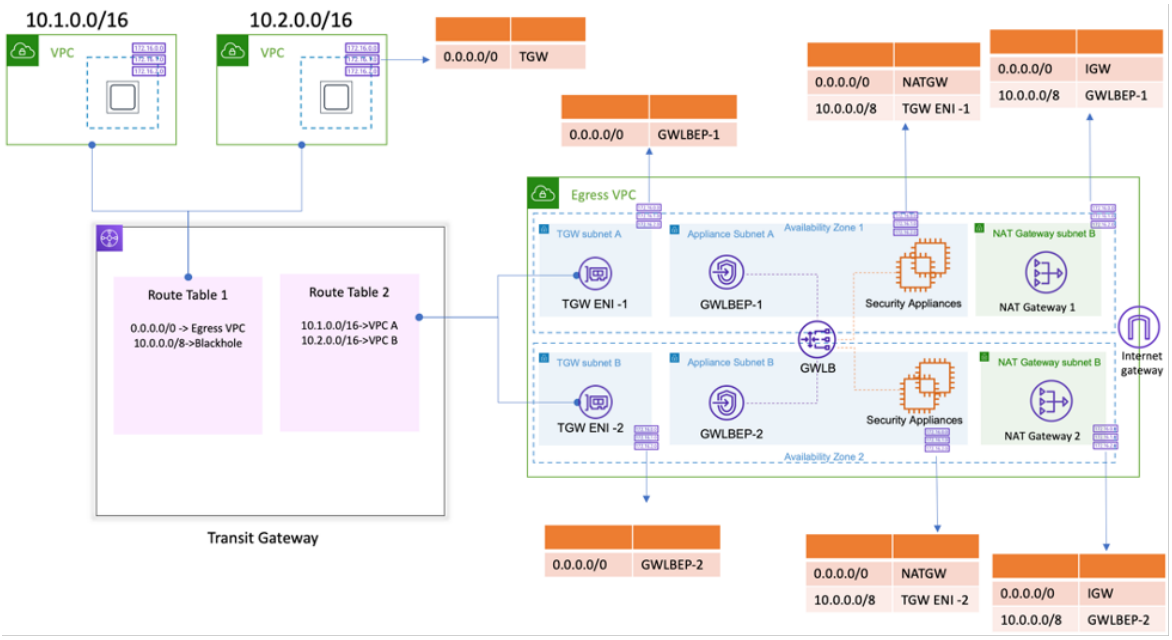
In the following figure, in addition to NAT gateway, you deploy virtual appliances using EC2 instances behind a Gateway Load Balancer (GWLB). In this setup, the GWLB, Gateway Load Balancer Endpoint (GWLBE), virtual appliances and NAT gateways are deployed in a centralized VPC which is connected to Transit Gateway using VPC attachment. The spoke VPCs are also connected to the Transit Gateway using a VPC Attachment. Because GWLBEs are a routable target, you can route traffic moving to and from Transit Gateway to the fleet of virtual appliances that are configured as targets behind a GWLB. GWLB acts as a bump-in-the-wire and transparently passes all Layer 3 traffic through third-party virtual appliances, and thus invisible to the source and destination of the traffic. Therefore, this architecture allows you to centrally inspect all of your egress traffic traversing through Transit Gateway.

For more information on how the traffic flows from the applications in the VPCs to the internet and back through this setup, refer to Centralized inspection architecture with AWS Gateway Load Balancer and AWS Transit Gateway.

You can enable appliance mode on Transit Gateway to maintain flow symmetry through virtual appliances. This means the bidirectional traffic is routed through the same appliance and the Availability

Zone for the life of the flow. This setting is particular important for stateful firewalls performing deep packet inspection. Enabling appliance mode removes the need for complex workarounds, such as source network address translation (SNAT), to force traffic to return to the correct appliance to maintain symmetry. Refer to Best practices for deploying Gateway Load Balancer for details.

It is also possible to deploy GWLB endpoints in a distributed manner without Transit Gateway to enable egress inspection. Learn more about this architectural pattern in the Introducing AWS Gateway Load Balancer: Supported architecture patterns blog post.



Centralized egress with Gateway Load Balancer and EC2 instance (route table design)

High availability

AWS recommends deploying Gateway Load Balancers and virtual appliances in multiple Availability Zones for higher availability.

Gateway Load Balancer can perform health checks to detect virtual appliance failures. In case of an unhealthy appliance, GWLB reroutes the new flows to healthy appliances. Existing flows always go to the same target regardless of health status of target. This allows for connection draining and accommodate health check failures due to CPU spikes on appliances. For more details, refer to section 4: Understand appliance and Availability Zone failure scenarios in the blog post Best practices for deploying Gateway Load Balancer. Gateway Load Balancer can use auto scaling groups as targets. This benefit takes out the heavy lifting of managing availability and scalability of the appliance fleets.

Scalability

Gateway Load Balancer operates at Layer 3 of the Open Systems Interconnection (OSI) model and is capable of handling several thousands of connections per second.

Advantages

Gateway Load Balancer and Gateway Load Balancer endpoints are powered by AWS PrivateLink, which allows for the exchange of traffic across VPC boundaries securely without the need to traverse the public internet.

Gateway Load Balancer is a managed service that removes undifferentiated heavy lifting of managing, deploying, scaling virtual security appliances so that you can focus on the things that matter. Gateway Load Balancer can expose the stack of firewalls as an endpoint service for customers to subscribe to using the AWS Marketplace. This is called Firewall as a Service (FWaaS).

Introduces a simplified deployment, and removes the need to rely on BGP and ECMP to distribute traffic across multiple Amazon EC2 instances.

Key considerations

- The appliances need to support Geneve encapsulation protocol to integrate with GWLB.
- Some third-party appliances can support SNAT and overlay routing (two-arm mode) therefore eliminating the need to create NAT gateways for saving costs. However, consult with an AWS partner of your choice before using this mode as this is dependent on vendor support and implementation.
- GWLB has a fixed idle timeout of 350 seconds for TCP flows and 120 seconds for non-TCP flows. This can result in connection timeouts on clients. You can tune your timeouts on client, server, firewall and OS level to avoid this. Refer to section 1: Tune TCP keep-alive or timeout values to support long-lived TCP flows in the Best practices for deploying Gateway Load Balancer blog post for more information.
- GWLBE are powered by AWS PrivateLink, so Transit Gateway and AWS PrivateLink charges will be applicable. You can learn more in the AWS PrivateLink pricing page.
- Consider deploying Transit Gateway and egress VPC in a separate Network Services account to segregate access based on delegation of duties, such as only network administrators can access Network Services Account.

Centralized network security for VPC-to-VPC and on-premises to VPC traffic

There may be scenarios where a customer wants to implement a layer 3-7 firewall/IPS/IDS within their Landing Zone to inspect traffic flows between VPCs or between an on- premises data center and a VPC. This can be achieved through different ways, depending on use case and requirements. For example, you could incorporate the Gateway Load Balancer, Network Firewall, Transit VPC, or use centralized architectures with Transit Gateways. These scenarios are discussed in the following section.

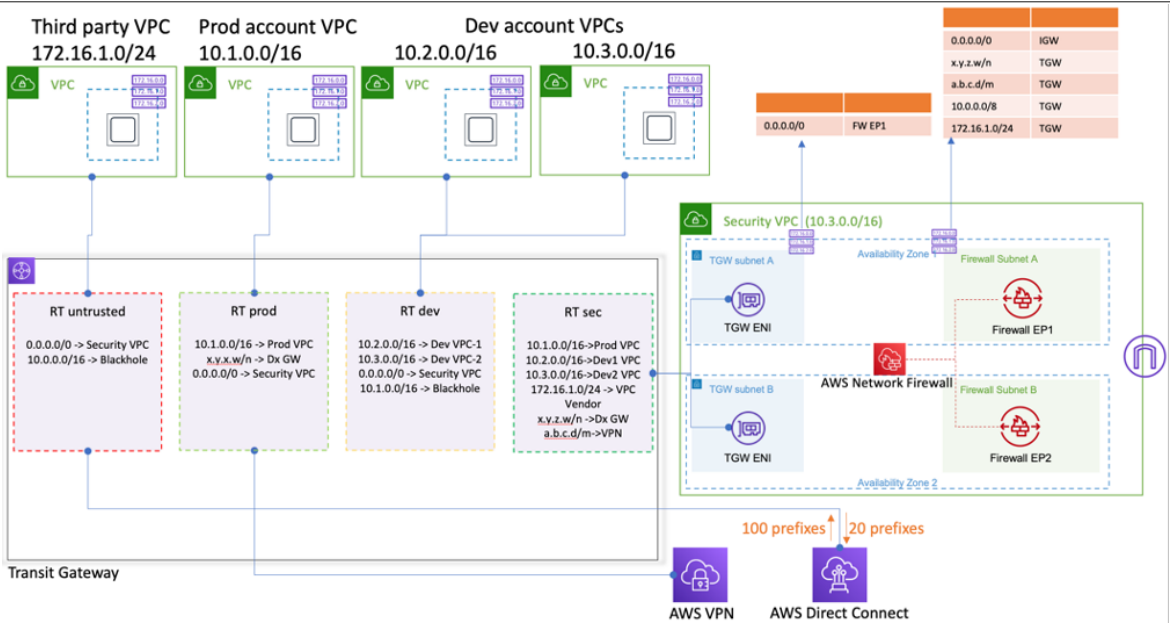
Considerations using a centralized network security inspection model

You should be selective in what traffic passes via your AWS Network Firewall or Gateway Load Balancer. One way to proceed is to define security zones and inspect traffic between untrusted zones. An untrusted zone can be a remote site managed by a third party, a vendor VPC you don't control/trust, or a sandbox/dev VPC, which has more relaxed security rules compared to rest of your environment. There are four zones in this example:

- **Untrusted Zone** — This is for any traffic coming from the 'VPN to remote untrusted site' or the third-party vendor VPC.
- **Production (Prod) Zone** — This contains traffic from the production VPC and on-premises customer DC.
- **Development (Dev) Zone** —This contains traffic from the two development VPCs.
- **Security (Sec) Zone** — Contains our firewall components Network Firewallor Gateway Load Balancer.

This is a setup has four security zones, but you might have more. You can use multiple route tables and blackhole routes to achieve security isolation and optimal traffic flow. Choosing the right zones is dependent on your overall Landing Zone design strategy (account structure, VPC design). You can have zones to enable isolation between Business Units (BUs), applications, environments, and so on.

If you want to inspect and filter your VPC-to-VPC, inter-zone traffic, and VPC-on-premises traffic, you can incorporate AWS Network Firewall with Transit Gateway in your centralized architecture. By having the hub-and-spoke model of the AWS Transit Gateway; centralized deployment model can be achieved. The AWS Network Firewall is deployed in a separate security VPC. A separate security VPC provides a simplified and central approach to manage inspection. Such a VPC architecture gives AWS Network Firewall source and destination IP visibility. Both source and destination IPs are preserved. This security VPC consists of two subnets in each Availability Zone; where one subnet is a dedicated to AWS Transit Gateway attachment and the other subnet is dedicated to the firewall endpoint.



VPC-to-VPC and on-premises to VPC traffic inspection using Transit Gateway and AWS Network Firewall (route table design)

In the centralized architecture with inspection, the Transit Gateway subnets require a separate VPC route table to ensure the traffic is forwarded to firewall endpoint within the same Availability Zone. For the return traffic, a single VPC route table containing a default route towards the Transit Gateway is configured. Traffic is returned to AWS Transit Gateway in the same Availability Zone after it has been inspected by AWS Network Firewall. This is possible due to the appliance mode feature of the Transit Gateway. The appliance mode feature of the Transit Gateway also helps the AWS Network Firewall to have stateful traffic inspection capability inside the security VPC.

With the appliance mode enabled on a transit gateway, it selects a single network interface using flow hash algorithm for the entire life of the connection. The transit gateway uses the same network interface for the return traffic. This ensures that bidirectional traffic is routed symmetrically—it's routed through the same Availability Zone in the VPC attachment for the life of the flow. For more information on appliance mode, refer to Stateful appliances and appliance mode in the Amazon VPC documentation.

For different deployment options of security VPC with AWS Network Firewall and Transit Gateway, refer to the Deployment models for AWS Network Firewall blog post.

Using Amazon Gateway Load Balancer with Transit Gateway for centralized network security

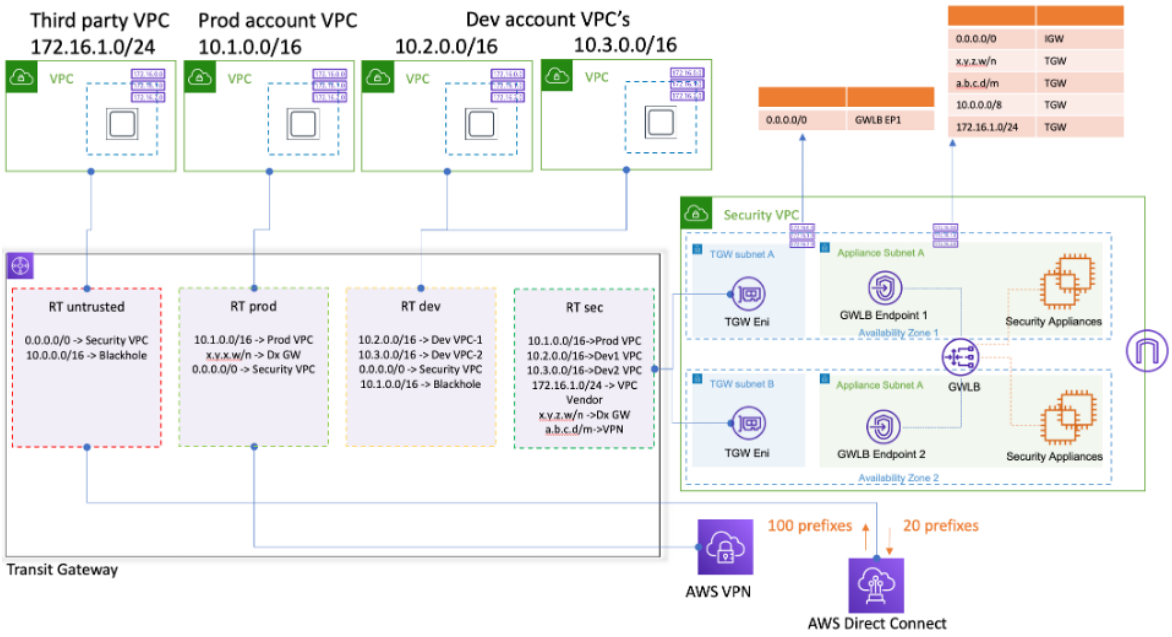
Often times, customers want to incorporate virtual appliances to handle the traffic filtering and to provide security inspection capabilities. In such use cases, they can integrate Gateway Load Balancer, virtual appliances, and Transit Gateway to deploy a centralized architecture for inspecting VPC-to-VPC and VPC-to-on-premises traffic.

Gateway Load Balancer, along with the virtual appliances, is deployed in a separate security VPC. The virtual appliances that will inspect the traffic are configured as targets behind the Gateway Load Balancer. Because Gateway Load Balancer endpoints are a routable target, customers can route traffic moving to and from Transit Gateway to the fleet of virtual appliances. To ensure flow symmetry, appliance mode is enabled on the Transit Gateway.

Each spoke VPC has a route table that is associated with the Transit Gateway, which has the default route to the Security VPC attachment as the next-hop.

The centralized Security VPC consists of appliance subnets in each Availabilityi Zone; which have the Gateway Load Balancer endpoints and the virtual appliances. It also has subnets for Transit Gateway attachments in each Availability Zone, as shown in the following figure.

For more information on centralized security inspection with Gateway Load Balancer and Transit Gateway, refer to the Centralized inspection architecture with AWS Gateway Load Balancer and AWS Transit Gateway blog post.



VPC-to-VPC and on-premises-to-VPC traffic inspection using Transit Gateway and AWS Gateway Load Balancer (route table design)

Key considerations for AWS Network Firewall and AWS Gateway Load Balancer

- Appliance mode should be enabled on the Transit Gateway.
- You can deploy the same model for inspection of traffic to other AWS Regions using AWS Transit Gateway Inter-Region peering.
- By default, each Gateway Load Balancer deployed in an Availability Zone distributes traffic across the registered targets within the same Availability Zone only. This is called Availability Zone affinity. If you enable cross-zone load balancing, Gateway Load Balancer distributes traffic across all registered and healthy targets in all enabled Availability Zones. If all targets across all Availability Zones are unhealthy, Gateway Load Balancer fails open. Refer to section 4: Understand appliance and AZ failure scenarios in the Best practices for deploying Gateway Load Balancer blog post for more details.
- For multi-Region deployment, AWS recommends that you set up separate inspection VPCs in the respective local Regions to avoid inter-Region dependencies and reduce associated data transfer costs. You should inspect traffic in the local Region instead of centralizing inspection to another Region.
- Cost of running an additional EC2-based high availability (HA) pair in multi-Region deployments can add up. For more information, refer to the Best practices for deploying Gateway Load Balancer blog post.

AWS Network Firewall vs Gateway Load Balancer

Table 2 — AWS Network Firewall vs Gateway Load Balancer

Criteria	AWS Network Firewall	Gateway Load Balancer
Use case	Stateful, managed, network firewall with intrusion detection and prevention service capability compatible with Suricata.	Managed service which makes it easy to deploy, scale and manage third-party virtual appliances
Complexity	AWS managed service. AWS handles the scalability and availability of the service.	AWS managed service. AWS will handle the scalability and availability of the service
Scale	AWS Network Firewall endpoints are powered by AWS PrivateLink. They support burst speeds up to 45 Gbps per endpoint.	Gateway Load Balancer endpoints support maximum bandwidth of up to 40 Gbps per endpoint
Cost	AWS Network Firewall endpoint cost + Data processing charges	Gateway Load Balancer + Gateway Load Balancer endpoints + virtual appliances + data processing charges

Centralized inbound inspection

Internet-facing applications, by their nature, have a larger attack surface and are exposed to categories of threats most other types of applications don't have to face. Having the necessary protection from attacks on these types of applications, and minimizing the impact surface area, are a core part of any security strategy.

As you deploy applications in your Landing Zone, many apps will be accessed by the users over the public internet (for example, through a Content Delivery Network (CDN), or through a public facing web-application) via a public facing load balancer, API gateway or directly through an internet gateway. You can secure your workloads and applications in this case by using AWS Web Application Firewall (AWS WAF) for Inbound Application Inspection, or alternatively IDS/IPS Inbound Inspection using AWS Gateway Load Balancer. or AWS Network Firewall.

As you continue to deploy applications in your Landing Zone, you might have a requirement to inspect inbound internet traffic. You can achieve this multiple ways, using either distributed, centralized, or combined inspection architectures using Gateway Load Balancer running your third-party firewall appliances, or alternatively AWS Network Firewall for advance IDS/IPS capabilities through the use of open source Suricata rules. This section covers both Gateway Load Balancer and AWS Network Firewall in a centralized deployment, using AWS Transit Gateway acting as a central hub for routing traffic.

AWS Web Application Firewall (AWS WAF) and AWS Firewall Manager for inspecting inbound traffic from the internet

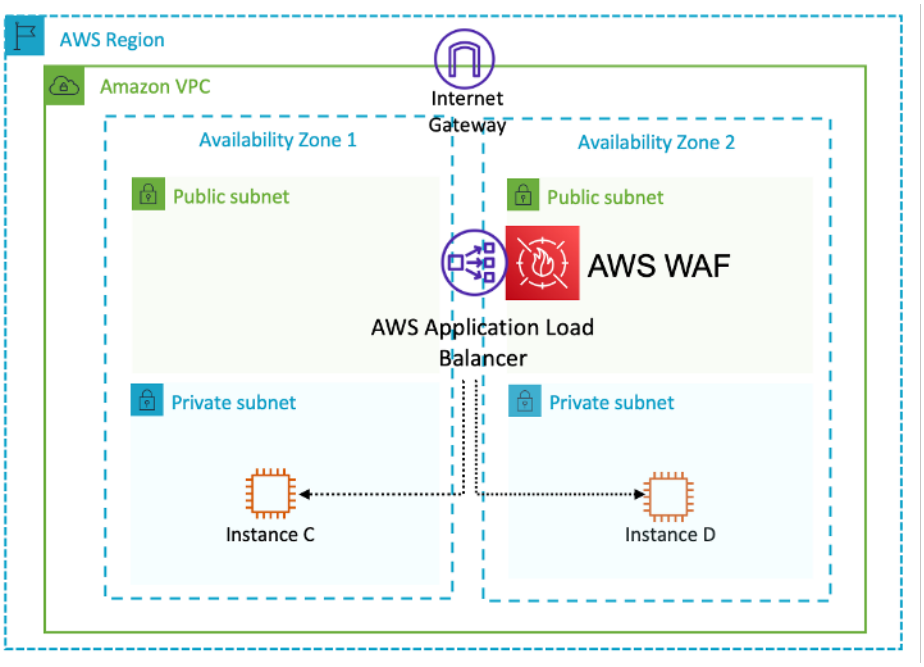
AWS WAF is a web application firewall that helps protect your web applications or APIs against common web exploits and bots that may affect availability, compromise security, or consume excessive resources. AWS WAF gives you control over how traffic reaches your applications by enabling you to create security rules that control bot traffic and block common attack patterns, such as SQL injection or cross-site scripting (XSS). You can also customize rules that filter out specific traffic patterns.

You can deploy AWS WAF on Amazon CloudFront as part of your CDN solution, the Application Load Balancer that fronts your web servers or origin servers running on Amazon EC2, Amazon API Gateway for your REST APIs, or AWS AppSync for your GraphQL APIs.

Once you deploy AWS WAF, you can then create your own traffic filter rules using the visual rule builder, code in JSON, or deploy managed rules maintained by AWS. These rules would filter out unwanted traffic by evaluating the traffic against the specified patterns. You can protect against exploits and vulnerabilities such as SQLi/XSS attacks. You can further use Amazon CloudWatch for monitoring incoming traffic metrics and logging.

For centralized management across all your accounts and applications in AWS Organizations, you can use AWS Firewall Manager. AWS Firewall Manager is a security management service which allows you to centrally configure and manage firewall rules. As your new applications are created, AWS Firewall Manager makes it easy to bring new applications and resources into compliance by enforcing a common set of security rules.

Using AWS Firewall Manager, you can easily roll out AWS WAF rules for your Application Load Balancers, API Gateway instances, and Amazon CloudFront distributions. AWS Firewall Manager integrates with AWS Managed Rules for AWS WAF, which gives you an easy way to deploy pre-configured, curated AWS WAF rules on your applications. For more information on centrally managing AWS WAF with AWS Firewall Manager, refer to Centrally manage AWS WAF (API v2) and AWS Managed Rules at scale with AWS Firewall Manager.



Centralized inbound traffic inspection using AWS WAF

In the preceding architecture, applications are running on Amazon EC2 instances in multiple availability zones in the private subnets. There is a public-facing Application Load Balancer (ALB) deployed in front of the Amazon EC2 instances, load balancing the requests between different targets. The AWS WAF is associated with the ALB.

Advantages

- With AWS WAF Bot Control, you get visibility and control over common and pervasive bot traffic to your applications.
- With Managed Rules for AWS WAF, you can quickly get started and protect your web application or APIs against common threats. You can select from many rule types, such as those that address issues such as the Open Web Application Security Project (OWASP) Top 10 security risks, threats specific to Content Management Systems (CMS) like WordPress or Joomla, or even emerging Common Vulnerabilities and Exposures (CVE). Managed rules are automatically updated as new issues emerge, so that you can spend more time building applications.
- AWS WAF is a managed service and no appliance is needed for inspection in this architecture.
- AWS WAF provides near-real-time logs through Amazon Kinesis Data Firehose.
- AWS WAF gives near real-time visibility into your web traffic, which you can use to create new rules or alerts in Amazon CloudWatch.

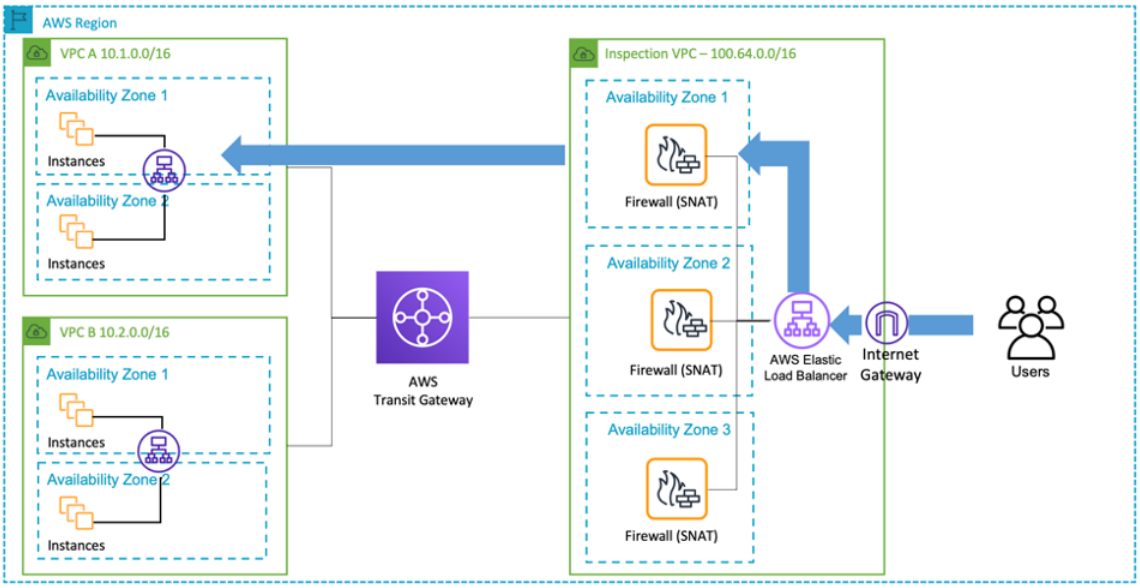
Key considerations

- This architecture is best suited for HTTP header inspection and distributed inspections, because AWS WAF is integrated on a per-ALB, CloudFront distribution and API Gateway.
- AWS WAF does not log the request body.
- Traffic going to a second set of ALB (if present) may not get inspected by the same AWS WAF instance; because a new request would be made to the second set of ALB.

Centralized inbound inspection with third-party appliances

In this architectural design pattern, you deploy third-party firewall appliances on Amazon EC2 across multiple availability zones behind an Elastic Load Balancer (ELB) such as an Application/Network Load Balancer in a separate Inspection VPC.

The Inspection VPC along with other Spoke VPCs are connected together through a Transit Gateway as VPC attachments. The applications in Spoke VPCs are frontend by an internal ELB which can be either ALB or NLB depending on the application type. The clients over the internet connect to the DNS of the external ELB in the inspection VPC which routes the traffic to one of the Firewall appliances. The Firewall inspects the traffic and then routes the traffic to the Spoke VPC through Transit Gateway using the DNS of the internal ELB as shown in the following figure. For more information regarding inbound security inspection with third-party appliances, refer to the How to integrate third-party firewall appliances into an AWS environment blog post.



Centralized ingress traffic inspection using third-party appliances and ELB

Advantages

- This architecture can support any application type for inspection and advanced inspection capabilities offered through third-party firewall appliances.
- This pattern supports DNS based routing from firewall appliances to spoke VPCs, which allows the applications in Spoke VPCs to scale independently behind an ELB.
- You can use Auto Scaling with the ELB to scale the firewall appliances in the Inspection VPC.

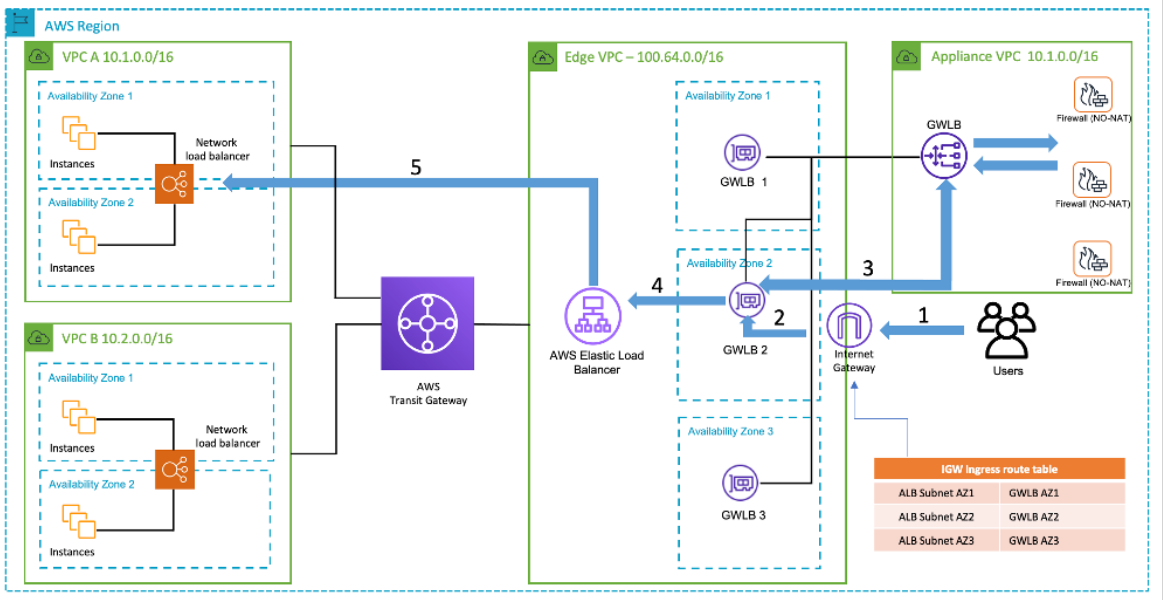
Key considerations

- You need to deploy multiple firewall appliances across Availability Zones for high availability.
- The firewall needs to be configured with and perform Source NAT in order to maintain flow symmetry, which means the client IP address won't be visible to the application.

- Consider deploying Transit Gateway and Inspection VPC in the Network Services account.
- Additional third-party vendor firewall licensing/support cost. Amazon EC2 charges are dependent on instance type.

Inspecting inbound traffic from the internet using firewall appliances with Gateway Load Balancer

Customers use third-party next-generation firewalls (NGFW) and intrusion prevention systems (IPS) as part of their defense in depth strategy. Traditionally these often are dedicated hardware or software/ virtual appliances. You can useGateway Load Balancer to scale these virtual appliances horizontally to inspect traffic from and to your VPC, as shown in the following figure.



Centralized ingress traffic inspection using firewall appliances with Gateway Load Balancer

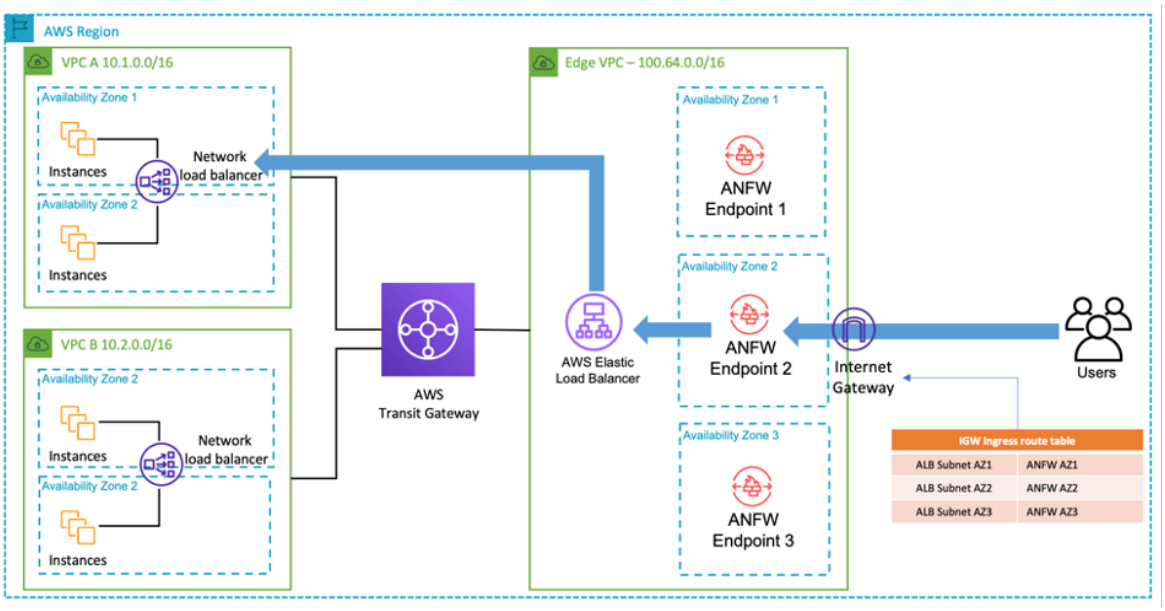
In the preceding architecture, Gateway Load Balancer endpoints are deployed into each Availability Zone in a separate edge VPC. The next-generation firewalls, intrusion prevention systems etc. are deployed behind the Gateway Load Balancer in the centralized appliance VPC. This appliance VPC can be in the same AWS account as the spoke VPCs or different AWS account. Virtual appliances can be configured to use Auto Scaling groups and are registered automatically with the Gateway Load Balancer, allowing auto scaling of the security layer.

These virtual appliances can be managed by accessing their management interfaces through an Internet Gateway (IGW) or using a bastion host setup in the appliance VPC.

Using the VPC ingress routing feature, the edge route table is updated to route inbound traffic from internet to firewall appliances behind Gateway Load Balancer. Inspected traffic is routed via Gateway Load Balancer endpoints to target VPC instance. Refer to the Introducing AWS Gateway Load Balancer: Supported architecture patterns blog post for details on various ways to use Gateway Load Balancer.

Using the AWS Network Firewall for centralized ingress

In this architecture, ingress traffic is inspected by AWS Network Firewall before reaching the rest of the VPCs. In this setup, traffic is split among all firewall endpoints deployed in the Edge VPC. You deploy a public subnet between the firewall endpoint and the Transit Gateway subnet. You can use an ALB or NLB, which contain IP targets in your spoke VPCs while handling Auto Scaling for targets behind them.



Ingress traffic inspection using AWS Network Firewall

Key considerations for AWS Network Firewall in a centralized ingress architecture

- Elastic Load Balancing in Edge VPC can only have IP addresses as target types, not a hostname. In the preceding figure, the targets are the private IPs of the Network Load Balancer in spoke VPCs. Using IP targets behind the ELB in the edge VPC results in the loss of Auto Scaling.
- Consider using AWS Firewall Manager as a single pane of glass for your firewall endpoints.
- This deployment model uses traffic inspection right as it enters the edge VPC, so it has the potential to reduce the overall cost of your inspection architecture.

DNS

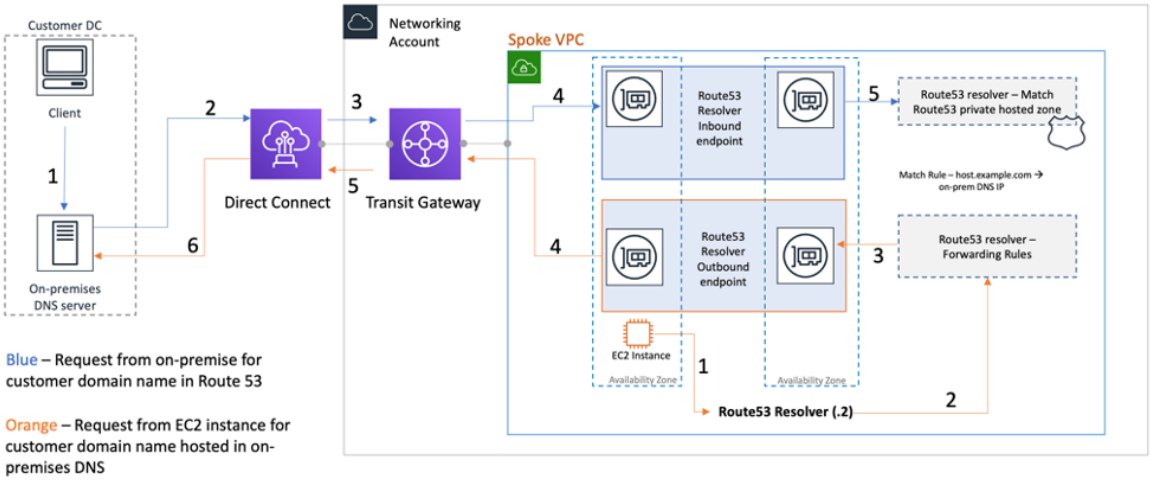
When you launch an instance into a nondefault VPC, AWS provides the instance with a private DNS hostname (and potentially a public DNS hostname) depending on the DNS attributes you specify for the VPC and if your instance has a public IPv4 address. When the 'enableDnsSupport' attribute is set to true, you get a DNS resolution within the VPC from Route 53 Resolver (+2 IP offset to the VPC CIDR). By default, Route 53 Resolver answers DNS queries for VPC domain names such as domain names for EC2 instances or Elastic Load Balancing load balancers. With VPC peering, hosts in one VPC can resolve public DNS hostnames to private IP addresses for instances in peered VPCs, provided the option to do so is enabled. The same is applicable for VPCs connected via AWS Transit Gateway. For more information, refer to Enabling DNS Resolution Support for a VPC Peering Connection.

If you want to map your instances to a custom domain name, you can use Amazon Route 53 to create a custom DNS-to-IP-mapping record. An Amazon Route 53 hosted zone is a container that holds information about how you want Amazon Route 53 to respond to DNS queries for a domain and its subdomains. Public Hosted Zones contain DNS information that is resolvable over the public internet while Private Hosted Zones are a specific implementation that only presents information to VPCs that have been attached to the specific private hosted zone. In a Landing Zone setup where you have multiple VPCs or accounts, you can associate a single private hosted zone with multiple VPCs across AWS accounts and across Regions (doable with SDK/CLI/API only).. The end hosts in the VPCs use their respective Route 53 Resolver IP (+2 offset the VPC CIDR) as the name server for DNS queries. The Route 53 Resolver in VPC accepts DNS queries only from resources within a VPC.

Hybrid DNS

DNS is a critical component of any infrastructure, hybrid or otherwise, as it provides the hostname-to-IP-address resolution that applications rely on. Customers implementing hybrid environments usually have a DNS resolution system already in place, and they want a DNS solution that works in tandem with their current system. Native Route 53 resolver (+2 offset of the base VPC CIDR) is not reachable from on-premises networks using VPN or AWS Direct Connect. Therefore, when you integrate DNS for the VPCs in an AWS Region with DNS for your network, you need a Route 53 Resolver inbound endpoint (for DNS queries that you are forwarding to your VPCs) and a Route 53 Resolver outbound endpoint (for queries that you are forwarding from your VPCs to your network).

As shown in the following figure, you can configure outbound Resolver endpoints to forward queries it receives from Amazon EC2 instances in your VPCs to DNS servers on your network. To forward selected queries, from a VPC to an on-premises network, create Route 53 Resolver rules that specify the domain names for the DNS queries that you want to forward (such as example.com), and the IP addresses of the DNS resolvers on your network where you want to forward the queries. For inbound queries from on-premises networks to Route 53 hosted zones, DNS servers on your network can forward queries to inbound Resolver endpoints in a specified VPC.

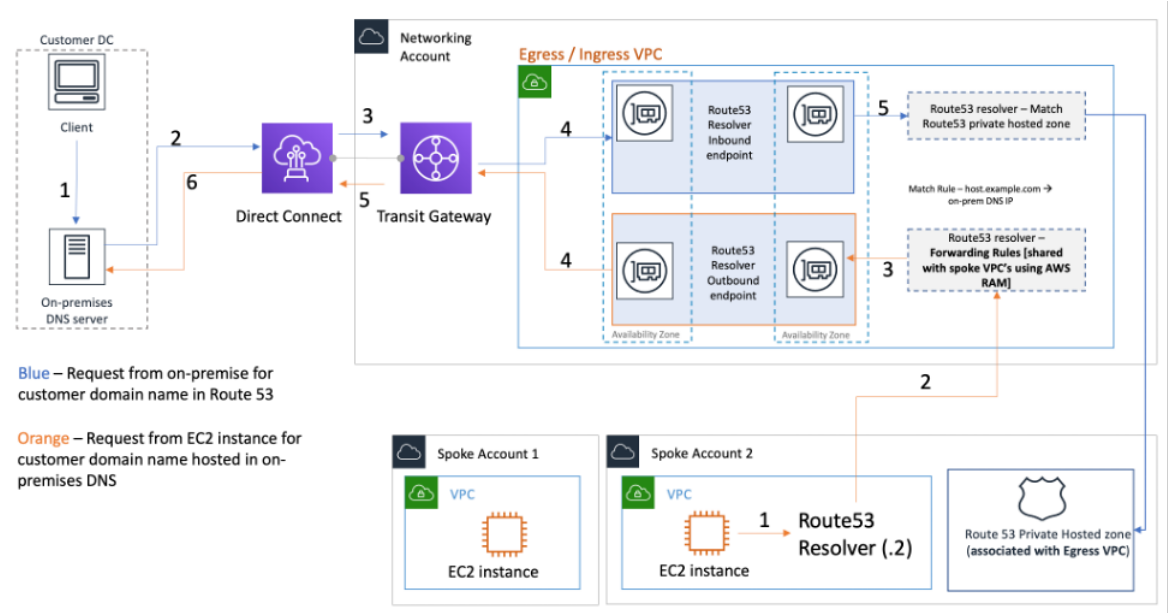


Hybrid DNS resolution using Route 53 Resolver

This allows your on-premises DNS resolvers to easily resolve domain names for AWS resources, such as Amazon EC2 instances or records in a Route 53 private hosted zone associated with that VPC. In addition, Route 53 Resolver endpoints can handle up to 10,000 queries per second per ENI, so can scale to much larger DNS query volume easily. Refer to Best practices for Resolver in the Amazon Route 53 documentation for more details.

It is not recommended that you create Route 53 Resolver endpoints in every VPC of the Landing Zone. Centralize them in a central egress VPC (in the Network services account). This approach allows for better manageability while keeping the costs low (you are charged an hourly fee for each inbound/outbound resolver endpoint you create). You share the centralized inbound and outbound endpoint with the rest of the Landing Zone.

- **Outbound resolution** — Use the Network Services account to write resolver rules (based on which DNS queries will be forwarded to on-premises DNS servers). Using Resource Access Manager (RAM), share these Route 53 Resolver rules with multiple accounts (and associate with VPCs in the accounts). EC2 instances in spoke VPCs can send DNS queries to Route 53 Resolver and Route 53 Resolver Service will forward these queries to on-premises DNS server via the outbound Route 53 Resolver endpoints in the egress VPC. You don't need to peer spoke VPCs to the egress VPC, or connect them via Transit Gateway. Do not use the IP of the outbound resolver endpoint as the primary DNS in the spoke VPCs. Spoke VPCs should use Route 53 Resolver (to offset of the VPC CIDR) in their VPC.



Centralizing Route 53 Resolver endpoints in ingress/egress VPC

- **Inbound DNS resolution** – Create Route 53 Resolver inbound endpoints in a centralized VPC and associate all the private hosted zones in your Landing Zone with this centralized VPC. For more information, refer to Associating More VPCs with a Private Hosted Zone. Multiple Private Hosted Zones (PHZ) associated with a VPC cannot overlap. As shown in the preceding figure, this association of PHZ with the centralized VPC will enable on-premises servers to resolve DNS for any entry in any private hosted zone (associated with central VPC) using the inbound endpoint in the centralized VPC. For further more information on hybrid DNS setups, refer to Centralized DNS management of hybrid cloud with Amazon Route 53 and AWS Transit Gateway and Hybrid Cloud DNS Options for Amazon VPC.

Route 53 DNS Firewall

Amazon Route 53 Resolver DNS Firewall helps filter and regulate outbound DNS traffic for your VPCs. A primary use of the DNS Firewall is to help prevent data exfiltration of your data by defining domain name allow-lists which allow resources in your VPC to make outbound DNS requests only for the sites your organization trusts. It also gives customers the ability to create "blocklists" for domains they don't want resources inside a VPC to communicate with via DNS. Amazon Route 53 Resolver DNS firewall has the following features:

Customers can create rules to define how DNS queries are answered. The actions that can be defined for the domain names include NODATA, OVERRIDE and NXDOMAIN.

Customers can create alerts for both allow-lists and deny-lists to monitor the rule activity. This can come in handy when customers want to test the rule before moving it to production.

For more information, refer to the How to Get Started with Amazon Route 53 Resolver DNS Firewall for Amazon VPC blog post.

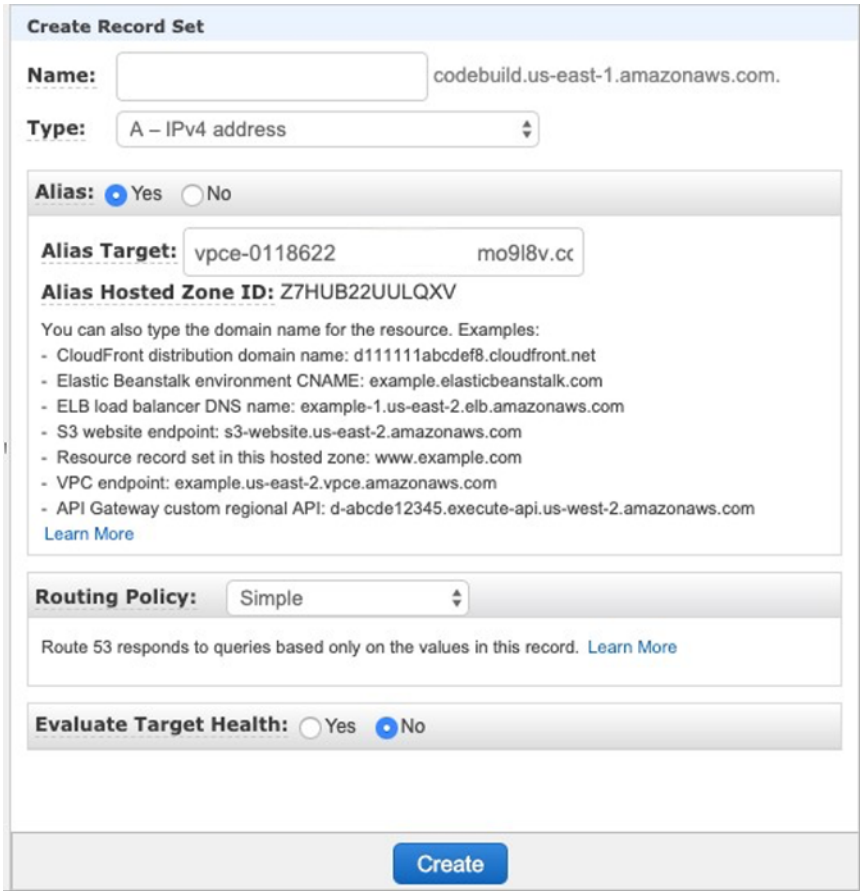
Centralized access to VPC private endpoints

A VPC endpoint allows you to privately connect your VPC to supported AWS services without requiring an internet gateway or a NAT device, VPN connection, or AWS Direct Connect connection. Therefore, your VPC is not exposed to the public internet. Instances in your VPC do not require public IP addresses to communicate with AWS service endpoints with this interface endpoint. Traffic between your VPC and other services does not leave the AWS network backbone. VPC endpoints are virtual devices. They are horizontally scaled, redundant, and highly available VPC components. Two types of endpoints can currently be provisioned: interface endpoints (powered by AWS PrivateLink) and gateway endpoints. Gateway endpoints can be utilized to access Amazon S3 and Amazon DynamoDB services privately. There is no additional charge for using gateway endpoints. Standard charges for data transfer and resource usage apply.

Interface VPC endpoints

An interface endpoint consists of one or more elastic network interfaces with a private IP address that serves as an entry point for traffic destined to a supported AWS service. When you provision an interface endpoint, a cost is incurred by users for every hour the endpoint is running along with data processing charges. By default, you create an interface endpoint in every VPC from which you want to access the AWS service. This can be cost prohibitive and challenging to manage in the Landing Zone setup where a customer wants to interact with a specific AWS service across multiple VPCs. To avoid this, you can host the interface endpoints in a centralized VPC. All the spoke VPCs will use these centralized endpoints via Transit Gateway.

When you create a VPC endpoint to an AWS service, you can enable private DNS. When enabled, the setting creates an AWS managed Route 53 private hosted zone (PHZ) which enables the resolution of public AWS service endpoint to the private IP of the interface endpoint. The managed PHZ only works within the VPC with the interface endpoint. In our setup, when we want spoke VPCs to be able to resolve VPC endpoint DNS hosted in a centralized VPC, the managed PHZ won't work. To overcome this, disable the option that automatically creates the private DNS when an interface endpoint is created. Next, manually create a Route 53 PHZ and add an Alias record with the full AWS service endpoint name pointing to the interface endpoint, as shown in the following figure.



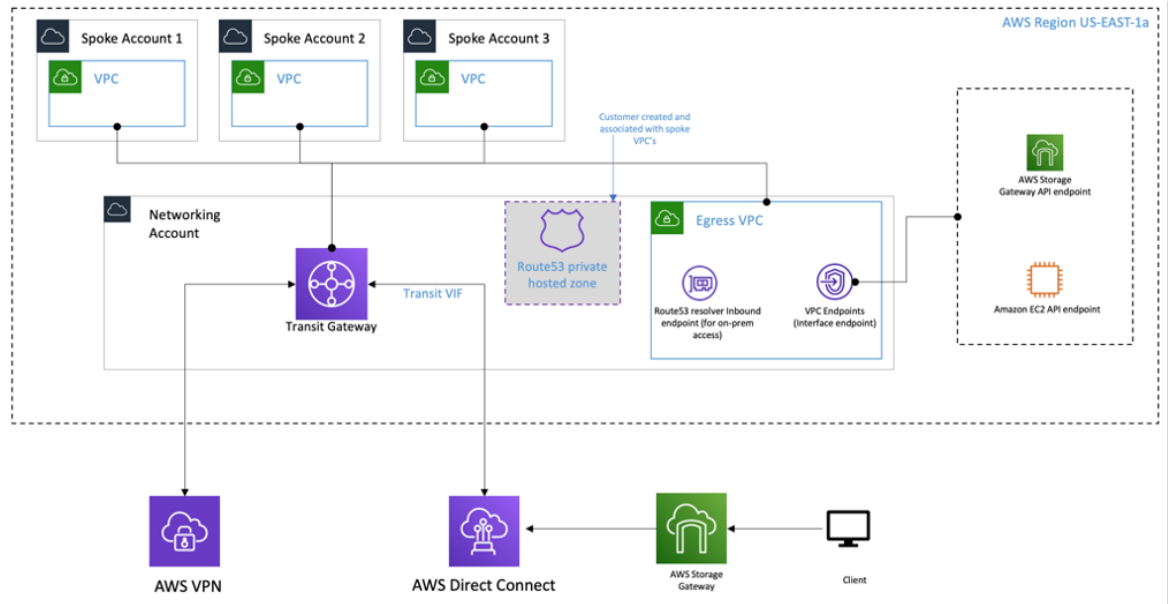
Manually created PHZ

You associate this private hosted zone with other VPCs within the Landing Zone. This configuration allows the spoke VPCs to resolve the full-service endpoint names to interface endpoints in the centralized VPC.

> **Note**
> To access the shared private hosted zone, the hosts in the spoke VPCs should use the Route 53 Resolver IP of their VPC. Interface endpoints are also accessible from on-premises networks over VPN and Direct Connect. Use conditional forwarding rules to send all DNS traffic for the full-service endpoint names to Route 53 Resolver inbound endpoints, which will resolve DNS requests according to the private hosted zone.

In the following figure, Transit Gateway enables traffic flow from the spoke VPCs to the centralized interface endpoints. Create VPC Endpoints and the private hosted zone for it in Network Services Account and share it with spoke VPCs in the spoke accounts. For more details on sharing endpoint information with other VPCs, refer to the Integrating AWS Transit Gateway with AWS PrivateLink and Amazon Route 53 Resolver blog post.

> **Note**
> A distributed VPC endpoint approach that is, an endpoint per VPC allows you to apply least privilege policies on VPC endpoints. In a centralized approach, you will apply and manage policies for all spoke VPC access on a single endpoint. With growing number of VPCs, the complexity of maintaining least privilege with a single policy document might grow. Single policy document also results in larger blast radius. You are also restricted on the size of the policy document (20,480 characters).



Centralizing interface VPC endpoints

Conclusion

As you scale your usage of AWS and deploy applications in the AWS Landing Zone, the number of VPCs and networking components increases. This whitepaper explained how you can manage this growing infrastructure ensuring scalability, high availability, and security while keeping costs low. Making the right design decisions when using services such as Transit Gateway, Shared VPC, AWS Direct Connect, VPC endpoints, Gateway Load Balancer, AWS Network Firewall, Amazon Route 53, and third-party software appliances becomes critical. It is important to understand the key considerations of each approach and work backwards from your requirements and analyze as to which option or combination of options fit you best.

Contributors

The following individuals contributed to this document:

- Sohaib Tahir, Solutions Architect, Amazon Web Services
- Shirin Bhambhani, Solutions Architect, Amazon Web Services
- Kunal Pansari, Solutions Architect, Amazon Web Services
- Eric Vasquez, Solutions Architect, Amazon Web Services

Document history

To be notified about updates to this whitepaper, subscribe to the RSS feed.

update-history-change	update-history-description	update-history-date
Minor update (p. 41)	Updated Transit Gateway section with Transit Gateway Connect, updated Transit VPC section; updated AWS Direct Connect section with MACsec and resiliency recommendations; updated AWS PrivateLink section.	February 22, 2022
Major update (p. 41)	Added VPC peering vs. Transit VPC vs. Transit Gateway comparison table; added centralized inbound inspection section; updated centralized network security for VPC-to-VPC and VPC-on-premises to VPC and centralized egress to internet with AWS Network Firewall and Gateway Load Balancer design patterns; added private NAT gateway and Amazon Route 53 DNS Firewall sections.	February 22, 2022
Minor update (p. 41)	Updated *Transit Gateway vs VPC peering* section	April 2, 2021
Whitepaper updated (p. 41)	Corrected text to match the options illustrated in Figure 7	June 10, 2020
Initial publication (p. 41)	Whitepaper published.	November 15, 2019

Notices

Customers are responsible for making their own independent assessment of the information in this document. This document: (a) is for informational purposes only, (b) represents current AWS product offerings and practices, which are subject to change without notice, and (c) does not create any commitments or assurances from AWS and its affiliates, suppliers, or licensors. AWS products or services are provided “as is” without warranties, representations, or conditions of any kind, whether express or implied. The responsibilities and liabilities of AWS to its customers are controlled by AWS agreements, and this document is not part of, nor does it modify, any agreement between AWS and its customers.

© 2022 Amazon Web Services, Inc. or its affiliates. All rights reserved.