MULTIPLE DATA CENTER HA NETWORK CONNECTIVITY *"How do I implement HA connectivity between my data centers and my VPC?"*

Overview

Amazon Web Services (AWS) offers customers the ability to achieve highly available network connections between Amazon Virtual Private Cloud (Amazon VPC) and their on-premises infrastructure. This capability extends customer access to AWS resources in a reliable, scalable, and cost-effective way. While many customers are able to quickly determine the type of connections they want to establish (either VPN or AWS Direct Connect[1]), they often struggle with understanding how to make these connections highly available and how to best leverage redundant connections, especially when these connections support remote networks that are geographically dispersed. This document provides AWS customers with best practices, recommendations, and common configurations to help build highly available, efficient connections between AWS and multiple remote data centers. For information about creating HA network connections from a single location, please see the Single Data Center HA Network Connectivity Solution Brief.

The following sections are applicable to customers who have already determined the appropriate remote connectivity option for their use case,[1] and assume basic knowledge of Amazon VPC, AWS Direct Connect, connecting and routing between remote networks, and dynamic routing protocols.

General Best Practices

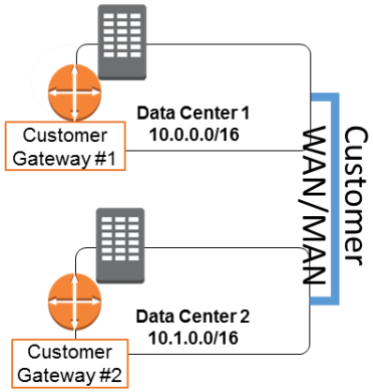
Highly available, fault-tolerant network connections are key to a well-architected system. When designing remote connections, consider using redundant hardware and telecommunications providers on both sides of the network connection. Additionally, it is a best practice to use dynamically routed, active/active connections for automatic load balancing and failover across redundant network connections. Keeping these topology best practices in mind, consider the following best practices when connecting to AWS:

- Leverage multiple dynamically routed, rather than statically routed, connections to AWS. This will allow remote connections to fail over automatically between redundant connections. Dynamic routing also enables remote connections to automatically leverage available preferred routes, if applicable, to the on-premises network.
- When selecting AWS Direct Connect network service providers, consider a dual-vendor approach, if financially feasible, to ensure private network diversity.
- Leverage more specific Border Gateway Protocol (BGP) route advertisements or BGP AS-path prepending to ensure AWS uses the most efficient routes to send traffic to your remote data centers.
- Provision sufficient network capacity to ensure that the failure of one network connection does not overwhelm and degrade redundant connections.

Application on the AWS Platform

The following sections provide high-availability options for redundant VPN connections, redundant AWS Direct Connect connections, and an AWS Direct Connect connection with a backup VPN connection. An organization's business-availability and application requirements will help determine the most appropriate configuration for each use case. For example, the most highly available option is to implement multiple AWS Direct Connect connections with circuits from different telecommunication providers, with an additional IPsec VPN backup connection.

While every customer network is unique, for the purposes of this document we will use the diagram to the right to represent a configuration of multiple customer data centers. The following configurations assume an internal, back-end network (depicted in the diagram as the Customer WAN/MAN) with internal routing protocols that are already configured to make the appropriate path selection to AWS.[2]



[1] Please see the *AWS Network Connectivity Options* whitepaper for more information about the available network connectivity options to AWS.

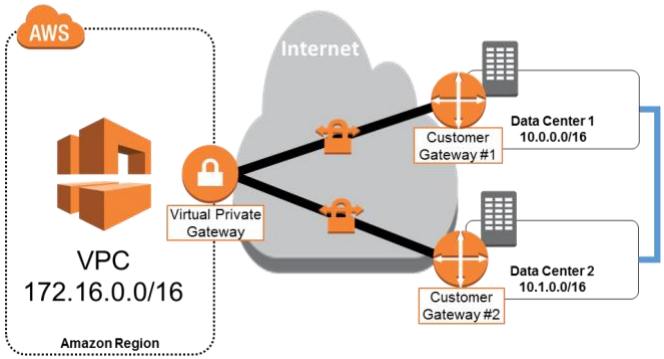
[2] It is important to maintain a common database of routes originating from AWS VPC to ensure a common routing topology and to improve route convergence in the event a route to AWS fails. A common approach is to configure iBGP between customer data centers to maintain this database, however any internal routing protocol that can accept or translate BGP metrics from AWS should be sufficient.

Redundant Active/Active VPN Connections

Many AWS customers choose to implement VPN connections because they are a quick, easy, and cost-effective way to set up remote connectivity to a VPC. To enable redundancy, each AWS Virtual Private Gateway (VGW) has two VPN endpoints[3] with capabilities for static and dynamic routing. Although statically routed VPN connections from a single customer gateway are sufficient for establishing remote connectivity to a VPC, this is not a highly available configuration. The best practice for making VPN connections highly available is to use redundant customer gateways and dynamic routing for automatic failover between AWS and customer VPN endpoints. For simplicity, the diagram in the next section depicts each VPN connection, consisting of two IPsec tunnels to both VGW endpoints, as a single line.

Configuration Details

The configuration in this example consists of four fully meshed, dynamically routed IPsec tunnels between both VGW endpoints and two customer gateways. AWS provides configuration templates for a number of supported VPN devices to assist in establishing these IPsec tunnels and configuring BGP for dynamic routing. In addition to the AWS-provided VPN and BGP configuration details, customers must configure VPCs to efficiently route traffic to their data center networks. In this example, the VGW will prefer to send 10.0.0.0/16 traffic to Data Center 1 through Customer Gateway 1, and only reroute this traffic through Data Center 2 if the connection to Data Center 1 is down. Likewise, 10.1.0.0/16 traffic will prefer the VPN connection originating from Data Center 2.



AWS recommends using one of the following approaches for communicating these route preferences[4]:

- More specific routes: With this approach, both Customer Gateway 1 and Customer Gateway 2 advertise a summary route of 10.0.0.0/15. In addition, Customer Gateway 1 advertises 10.0.0.0/16 and Customer Gateway 2 advertises 10.1.0.0/16. AWS will use the more specific routes to send traffic to the appropriate data center, and will fail back to the other data center following the summarized route if the more specific route becomes temporarily unavailable.
- AS-path prepending: With this approach, both Customer Gateway 1 and Customer Gateway 2 advertise 10.0.0.0/16 and 10.1.0.0/16. However, Customer Gateway 1 uses AS-path prepending when advertising the 10.1.0.0/16 network to make this route less preferred. Likewise, Customer Gateway 2 uses AS-path prepending when advertising the 10.0.0.0/16 network to make this route less preferred. AWS will use the preferred routes to send traffic to the appropriate data center, and will fail back to the other data center following the less preferred routes when necessary.

If your organization already leverages AS-path prepending for influencing route preferences, then the latter approach will likely align more closely with your existing routing policies. Otherwise, the approach using more specific routes is a great place to start.

Considerations

This configuration relies on the Internet to carry traffic between on-premises networks and VPC. Although AWS leverages multiple Internet Service Providers (ISPs), and even if the customer leverages multiple ISPs, an Internet service disruption can still affect the availability of VPN network connectivity due to the interdependence of ISPs and Internet routing. The only way to control the exact network path of your traffic is to provision private network connectivity with AWS Direct Connect (see the next option).

[3] http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_VPN.html#VPNTunnels

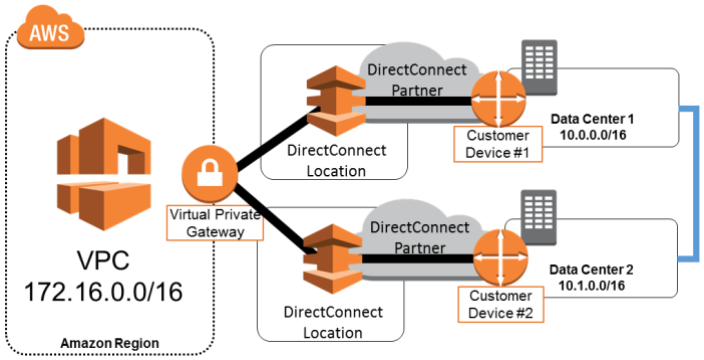
[4] For a full explanation of VPC routing rule algorithm, please see http://docs.aws.amazon.com/AmazonVPC/latest/NetworkAdminGuide/Introduction.html#MultipleVPNConnections

Redundant Active/Active AWS Direct Connect Connections

Many AWS customers establish private connectivity between AWS and their datacenter, office, or colocation environment with AWS Direct Connect to reduce network costs, increase bandwidth throughput, or provide a more consistent network experience than Internet-based connections. Because each dedicated, physical connection is in one AWS Direct Connect location, multiple dynamically routed AWS Direct Connect connections are necessary to achieve high availability, as depicted in the diagram below. Architectures with the highest levels of availability will leverage different AWS Direct Connect partner networks to ensure network-provider redundancy. For simplicity, the diagram in the next section depicts each AWS Direct Connect connection, consisting of both a physical connection that contains logical connections, as one continual line.[5]

Configuration Details

The configuration in this example consists of AWS Direct Connect connections to separate AWS Direct Connect routers in two locations from two independently configured customer devices. AWS provides example router configurations to assist in establishing AWS Direct Connect connections and configuring BGP for dynamic routing. In addition to the AWS-provided configuration details, customers must configure VPCs to efficiently route traffic to their data center networks. In this example, the VGW will prefer to send 10.0.0.0/16 traffic to Data Center 1, and only reroute this traffic to Data Center 2 if connectivity to Customer Device 1 is down. Likewise, 10.1.0.0/16 traffic will prefer the AWS Direct Connect connection from Data Center 2.



AWS recommends using one of the following approaches for communicating these route preferences[6]:

- More specific routes: With this approach, both Customer Device 1 and Customer Device 2 advertise a summary route of 10.0.0.0/15. In addition, Customer Device 1 advertises 10.0.0.0/16 and Customer Device 2 advertises 10.1.0.0/16. AWS will use the more specific routes to send traffic to the appropriate data center, and will fail back to the other data center following the summarized route if the more specific route becomes temporarily unavailable.
- AS-path prepending: With this approach, both Customer Device 1 and Customer Device 2 advertise 10.0.0.0/16 and 10.1.0.0/16. However, Customer Device 1 uses AS-path prepending when advertising the 10.1.0.0/16 network to make this route less preferred. Likewise, Customer Device 2 uses AS-path prepending when advertising the 10.0.0.0/16 network to make this route less preferred. AWS will use the preferred routes to send traffic to the appropriate data center, and will fail back to the other data center following the less preferred routes when necessary.

If your organization already leverages AS-path prepending for influencing route preferences, then the latter approach will likely align more closely with your existing routing policies. Otherwise, the approach using more specific routes is a great place to start.

Considerations

AWS Direct Connect allows you to create resilient connections to AWS because you have full control over the network path and network providers between your remote networks and AWS. Choose network providers and AWS Direct Connect locations that align with your organization's risk tolerance, financial expectations, and data-center-connectivity policies. For example, if your current data centers leverage multiple network providers to reduce the risk associated with an individual network-provider outage, then you should consider using different network providers for each AWS Direct Connect connection. Likewise, leveraging different AWS Direct Connect locations (e.g. CoreSite and Equinix in US East) will reduce the risk that a facility failure will interrupt your network connectivity with AWS.

[5] For more information on AWS Direct Connect physical and logical connections, please see http://docs.aws.amazon.com/directconnect/latest/UserGuide/Welcome.html

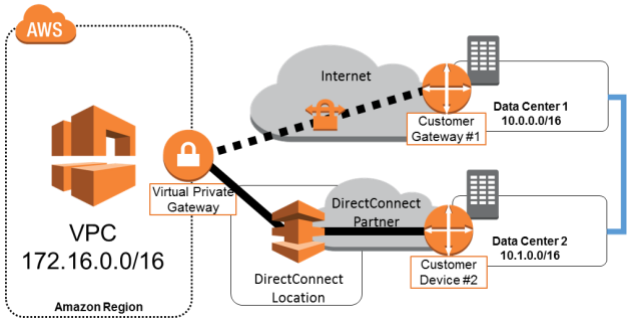
[6] For a full explanation of VPC routing rule algorithm, please see http://docs.aws.amazon.com/AmazonVPC/latest/NetworkAdminGuide/Introduction.html#MultipleVPNConnections

AWS Direct Connect with Backup VPN Connection

Some AWS customers would like the benefits of one or more AWS Direct Connect connections for their primary connectivity to AWS, coupled with a lower-cost backup connection. To achieve this objective, they can establish AWS Direct Connect connections with a VPN backup, as depicted in the diagram below.

Configuration Details

The configuration in this example consists of two dynamically routed connections, one using AWS Direct Connect and the other using a VPN connection from two different customer devices. AWS provides example router configurations to assist in establishing both AWS Direct Connect and VPN connections with BGP for dynamic routing. By default, AWS will always prefer to send traffic over an AWS Direct Connect connection, so no additional configuration is required to define primary and backup connections. In this example, both Customer Gateway 1 and Customer Device 2 advertise a summary route of 10.0.0.0/15 and AWS will send all traffic to Customer Device 2 as long as this network path is available.



Considerations

Although AWS prefers AWS Direct Connect to VPN connections by default, note that you still have the ability to influence AWS routing decisions by advertising more specific routes. If, for example, you want to leverage your backup VPN connection for a subset of traffic (e.g., developer traffic versus production traffic), you can advertise specific routes from Customer Gateway 1.

This approach allows you to choose the primary network path and network provider for your AWS traffic, with the option of using a different provider for a backup VPN connection. Choose network providers and AWS Direct Connect locations that align with your organization's risk tolerance, financial expectations, and data-center connectivity policies. For example, if you are concerned about the risk associated with an individual network-provider outage, consider different network providers for AWS Direct Connect and Internet connectivity. Additionally, make sure to monitor AWS Direct Connect utilization to ensure that a VPN connection will be a sufficient backup to support your application's latency and bandwidth requirements.

Resources

Amazon VPC Network Connectivity Options	https://aws.amazon.com/whitepapers/amazon-vpc-connectivity-options/ AWS whitepaper describing common network connectivity options, including the integration of customer networks with Amazon VPC.
AWS Direct Connect Documentation	https://aws.amazon.com/documentation/direct-connect/
Amazon VPC Documentation	https://aws.amazon.com/documentation/vpc/
Single Data Center HA Network Connectivity	https://d0.awsstatic.com/aws-answers/AWS_Single_Data_Center_HA_Network_Connectivity.pdf AWS Solution Brief describing options for creating highly available connections from a single data center to AWS.