

Book of Network Services - Flow Virtual Networking

[PDF generated December 27 2023. For all recent updates please see the Nutanix Bible releases notes located at https://nutanixbible.com/release_notes.html. Disclaimer: Downloaded PDFs may not always contain the latest information.]

Flow Virtual Networking allows you to create completely isolated virtual networks that are separated from the physical network. You can do multi-tenant network provisioning with overlapping IP addresses, self-service network provisioning, and IP address preservation. Best of all, these isolated virtual networks provide security by default.

Supported Configurations

Core Use Cases:

- Multi-tenant networking
- Network isolation for security
- Overlapping IP addresses
- Self-service network creation
- VM IP mobility
- Hybrid Cloud connectivity

Management interfaces(s):

- Prism Central (PC)

Supported Environment(s):

- On-Premises:
 - AHV

Upgrades:

- Major Feature Upgrades Depend on
 - Prism Central
 - AHV
- Minor Feature Upgrades Included in LCM
 - Advanced Network Controller
 - Network Gateway (VPN, Subnet Extension, BGP)

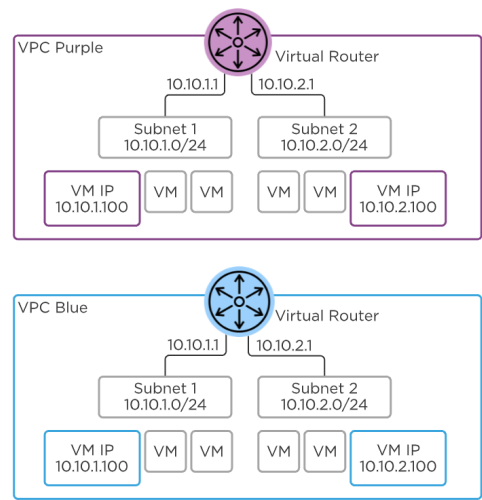
Implementation Constructs

Flow Virtual Networking introduces a number of new constructs to provide a complete networking solution.

- VPCs
- Overlay Subnets
- Routes
- Policies
- External Networks
 - NAT
 - Routed (NoNAT)
 - Multiple External Networks
- Network Gateway
 - Layer 3 VPN
 - Layer 2 Extended Subnets with VXLAN
 - BGP

VPCs (Virtual Private Clouds)

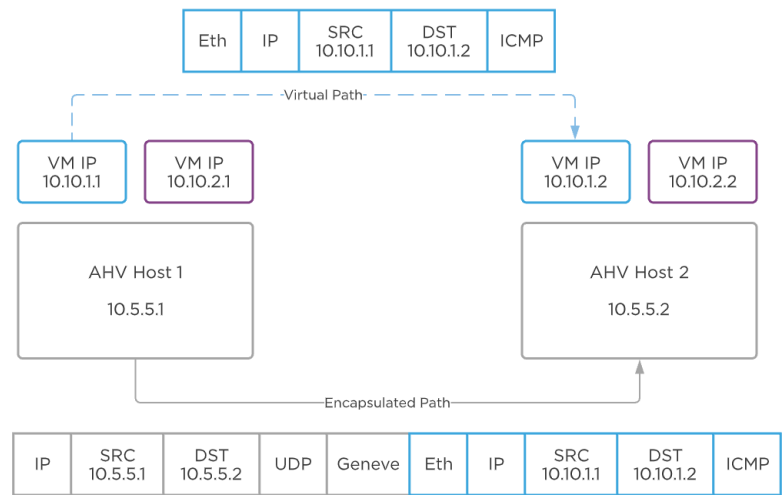
The VPC or Virtual Private Cloud is the basic unit of Flow Virtual Networking. Each VPC is an isolated network namespace with a virtual router instance to connect all of the subnets inside the VPC. This is what allows the IP addresses inside of one VPC to overlap with any other VPC, or even with the physical network. A VPC can expand to include any cluster managed by the same Prism Central, but generally a VPC should exist only within a single AHV cluster, or within clusters in the same availability zone. More on that when we get to External Networks.



Flow Virtual Networking - VPC

Overlay Subnets

Each VPC can have one or more subnets and they're all connected to the same VPC virtual router. Behind the scenes, a VPC leverages Geneve encapsulation to tunnel traffic between AHV hosts as needed. This means the subnets inside the VPC don't need to be created or even present on the top-of-rack switches for VMs on different hosts to communicate. When two VMs in a VPC on two different hosts send traffic to each other, packets are encapsulated in Geneve on the first host and sent to the other host where they're decapsulated and sent to the destination VM.



Flow Virtual Networking - Geneve Encapsulation

When you select a NIC for a VM, you can place that NIC into an overlay subnet, or a traditional VLAN backed subnet. When you choose an overlay subnet, that is also choosing the VPC.

Pro tip

Each VM can be placed inside only a single VPC. You can't connect a VM to both a VPC and a VLAN at the same time, or to two different VPCs at the same time.

Routes

Every VPC contains a single virtual router and there are a few types of routes:

- External Networks
- Direct Connections
- Remote Connections

External networks should be the default destination of the 0.0.0.0/0 network prefix for the whole VPC. You can choose an alternate network prefix route for each external network in use. For completely isolated VPCs, you may choose not to set a default route.

Directly connected routes are created for each subnet inside the VPC. Flow Virtual Networking assigns the first IP address of each subnet as the default gateway for that subnet. The default gateway and network prefix are determined by the subnet configuration and cannot be altered directly. Traffic between two VMs on the same host and same VPC, but in two different subnets, will be routed locally in that host.

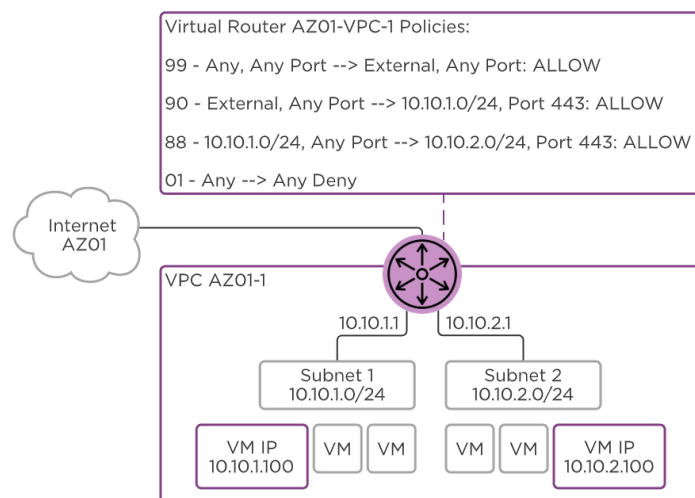
Remote connections such as VPN connections and External Networks can be set as the next hop destination for a network prefix.

Pro tip

Each network prefix inside a VPC routing table must be unique. Don't program two different next hop destinations with the same destination prefix.

Policies

The virtual router acts as a control point for traffic inside a VPC. You can apply simple stateless policies here, and any traffic that flows through the router will be evaluated by the policies. Traffic from one VM to another inside the same subnet won't go through a policy.



Flow Virtual Networking - Policies

Inside a VPC, policies are evaluated in priority order from highest (1,000) to lowest (10).

You can match traffic based on the following values:

- Source or Destination IPv4 Address
 - Any
 - External to the VPC (Any traffic entering the VPC)
 - Custom IP Prefix
- Protocol
 - Any
 - Protocol Number
 - TCP

- UDP
- ICMP

- Source or Destination Port Number

Once traffic is matched a policy can take the following actions:

- Permit
- Deny
- Reroute
 - Redirect traffic to another /32 IPv4 address in another subnet

The reroute policy is incredibly helpful to take action like routing all inbound traffic through a load balancer or firewall VM running inside another subnet in the VPC. This has the added value of requiring only a single Network Function VM (NFVM) for all traffic inside the VPC, rather than a traditional service chain that requires an NFVM per AHV host.

Pro tip

Stateless policies require separate rules defined in both the forward and reverse direction if a Permit rule is overriding a Drop rule. Otherwise, return traffic would be denied by the Drop rule. Use similar priorities to group these matching forward and reverse entries.

External Networks

An External Network is the primary way traffic enters and exits a VPC. External Networks are created in Prism Central and exist on only a single Prism Element cluster. This network defines the VLAN, the default gateway, the IP address pool, and the NAT type for all the VPCs using it. One External Network can be used by many VPCs.

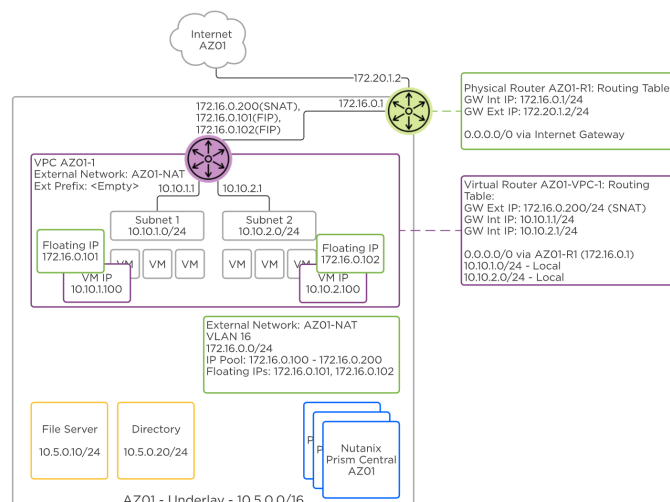
There are two types of External Networks:

- NAT
- Routed (NoNAT)

Starting in PC.2022.1 and AOS 6.1, you can select a maximum of two External Networks for a VPC. A VPC can have at most one NAT External Network and at most one Routed (NoNAT) External Network. You cannot configure two NAT, or two Routed External Networks in the same VPC.

NAT

A NAT (Network Address Translation) External Network hides the IP addresses of VMs in the VPC behind either a Floating IP or the VPC SNAT (Source NAT) address. Each VPC has an SNAT IP address selected randomly from the External Network IP pool, and traffic exiting the VPC is rewritten with this source address.

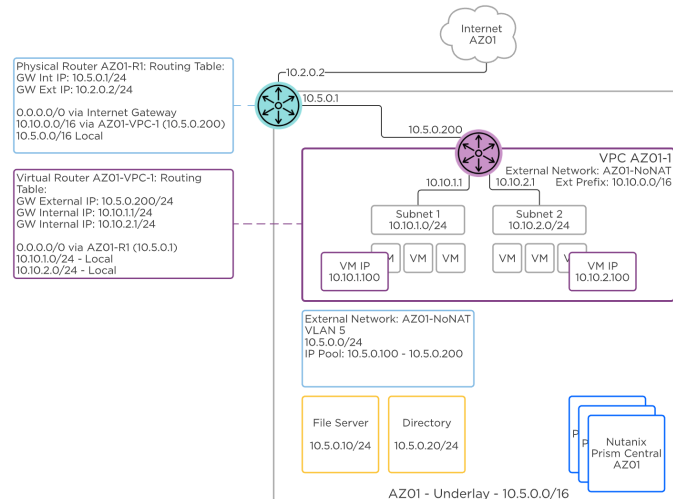


Floating IP addresses are also selected from the External Network IP pool and are assigned to VMs in a VPC to allow ingress traffic. When a Floating IP is assigned to a VM, the egress traffic is rewritten with the Floating IP instead of the VPC SNAT IP. This is useful for advertising public services outside the VPC without revealing the private IP address of the VM.

Routed

Routed, or NoNAT External Networks allow the IP address space of the physical network to be shared inside the VPC through routing. Instead of a VPC SNAT IP address, the VPC router IP is selected randomly from the External Network pool. Share this VPC router IP with the physical network team so they can set this virtual router IP as the next hop for all of the subnets provisioned inside the VPC.

For example, a VPC with an External Network of 10.5.0.0/24 may be assigned a virtual router IP of 10.5.0.200. If the subnets inside the VPC are created in the 10.10.0.0/16 network, the physical network team will create a route to 10.10.0.0/16 via 10.5.0.200. The 10.10.0.0/16 network becomes the externally routable prefix for the VPC.



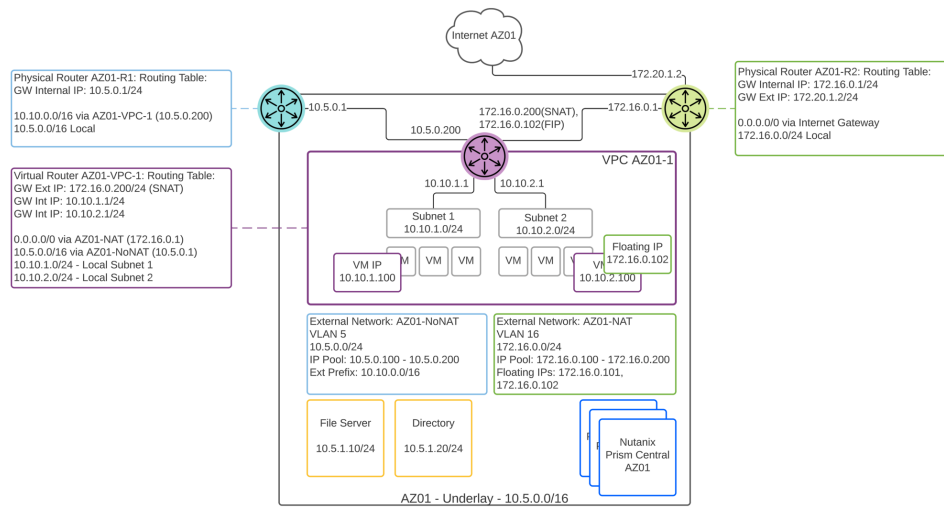
Flow Virtual Networking - Routed External Network

Multiple External Networks

Having two external networks is useful in the following scenarios.

- Connectivity to internal and external resources through separate routes.
 - Internal or Corporate through Routed, or NoNAT
 - External or Internet services through NAT and Floating IPs
- Providing Floating IPs for network gateway VMs inside the VPC. Network gateways inside a VPC require Floating IPs.
 - VPN Network Gateway VM
 - VXLAN VTEP Network Gateway VM
 - BGP Network Gateway VM inside the VPC

The following VPC is connected to a NAT network for Internet access, as shown in the VPC routing table **Virtual Router AZ01-VPC-1: Routing Table**. All Internet directed traffic (0.0.0.0/0) exits the VPC through AZ01-NAT external network. This VPC is also connected to a Routed or NoNAT external network and all internal traffic destined to 10.5.0.0/16 exits through the AZ01-NoNAT external network.



Flow Virtual Networking - Multiple External Networks

Network Gateways

A network gateway acts as a connector between subnets. These subnets can be of many different types and in different locations.

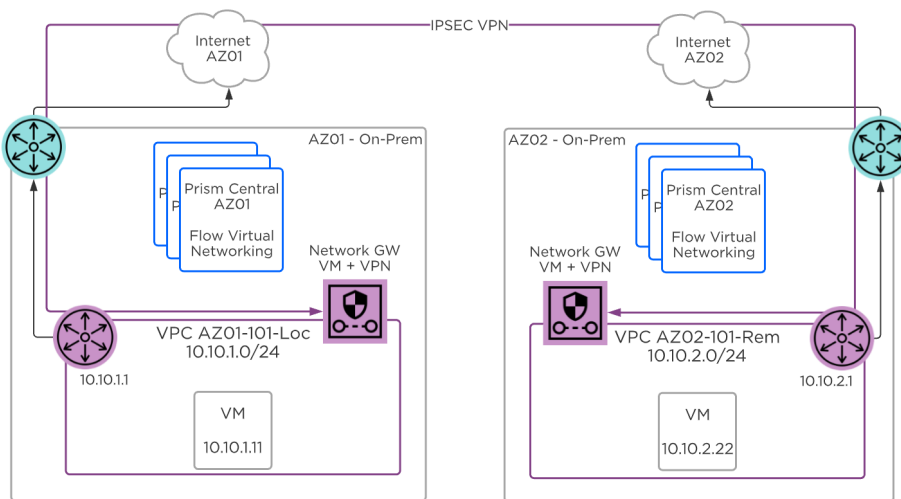
- Subnet Types
 - ESXi VLAN
 - AHV VLAN
 - VPC Overlay (Network gateways in a VPC require Floating IPs)
 - Physical Network VLAN
- Subnet Locations
 - On-premises VLANs
 - On-premises VPCs
 - Cloud VPCs

The network gateway has several methods of connecting subnets.

- Layer 3 VPN
 - Network Gateway to Network Gateway
 - Network Gateway to Physical Firewall or VPN
- Layer 2 VXLAN VTEP
 - Network Gateway to Network Gateway
 - Network Gateway to Physical Router or Switch VTEP
- Layer 2 VXLAN VTEP over VPN
 - Network Gateway to Network Gateway
- BGP
 - Dynamic VPC subnet advertisement to a remote BGP neighbor

Layer 3 VPN

In the Layer 3 VPN connection type, two subnets with two separate network prefixes are connected. For example, local subnet 10.10.1.0/24 can be connected to remote subnet 10.10.2.0/24. The following diagram assumes a NAT external network and the VPN VM is assigned a Floating IP from this NAT network.



Flow Virtual Networking - Layer 3 VPN

When using two network gateways inside a VPC, each gateway VM is assigned an external Floating IP address from the NAT external network DHCP pool, and they must be able to communicate over these addresses. When the network gateway is inside a VLAN, the assigned IP address must be reachable from the remote site.

You can also connect the network gateway VM to a remote physical firewall or VPN appliance or VM. The local network gateway must still be able to communicate with the remote physical or virtual appliance.

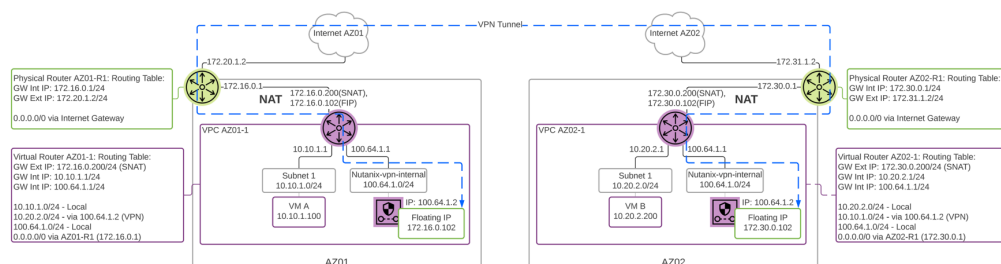
Traffic from each subnet to the remote subnet is directed over the created VPN connection using IP routing inside the VPC. Routes can be either static or shared between network gateways using BGP. Traffic inside this VPN tunnel is encrypted with IPSEC.

Pro tip

Use VPN with IPSEC encryption when traffic between subnets will go over a public link such as the Internet.

Layer 3 VPN Detailed View for NAT

The following diagram shows a more detailed view of the assigned interfaces and addresses for a VPN network gateway VM connecting between two VPCs with NAT external networks.

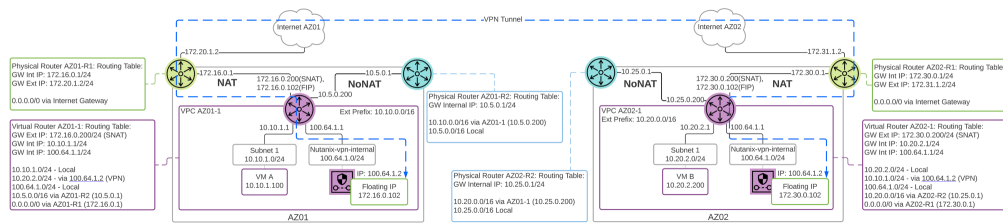


Flow Virtual Networking - Layer 3 VPN NAT Detail

Note that a new internal subnet in the 100.64.1.0/24 overlay network is automatically created when the VPN gateway VM is deployed. This VM reserves a Floating IP from the NAT external network pool. Pay special attention to the routing table within the VPC. Traffic to the remote subnet is routed through the VPN connection.

Layer 3 VPN Detailed View for Routed or NoNAT

The following diagram shows the detailed interface view when a routed external network is attached to the VPC. Multiple external networks are mandatory in this case to provide a NAT external network for the VPN VM Floating IP.

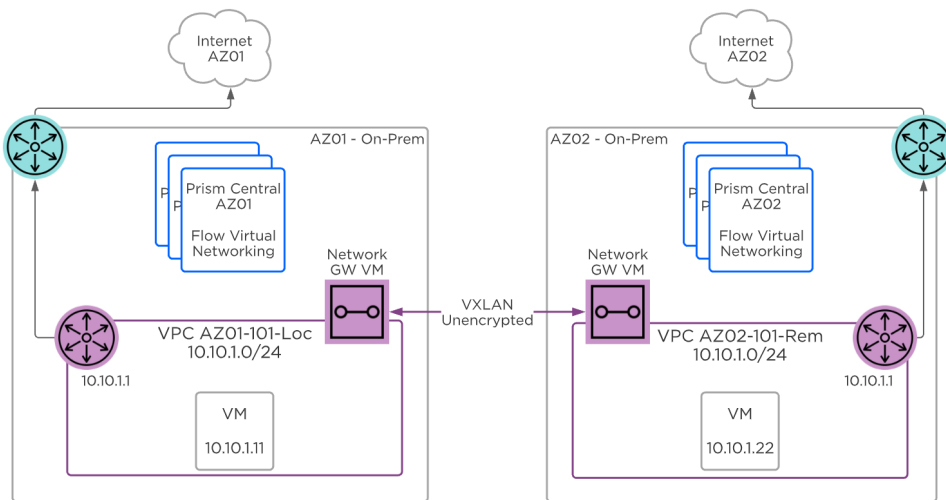


Flow Virtual Networking - Layer 3 VPN NoNAT Detail

In this case the AZ01-1 and AZ01-2 VPC routing tables contain default routes to the NAT network, and more specific routes to the private internal NoNAT or Routed network. The VPN tunnel traffic traverses the NAT external network to reach the remote site.

Layer 2 VXLAN VTEP Subnet Extension

In the Layer 2 VXLAN VTEP case, called subnet extension, two subnets that share the same network prefixes are connected. For example, local overlay subnet 10.10.1.0/24 is connected to a remote overlay subnet that also uses 10.10.1.0/24.

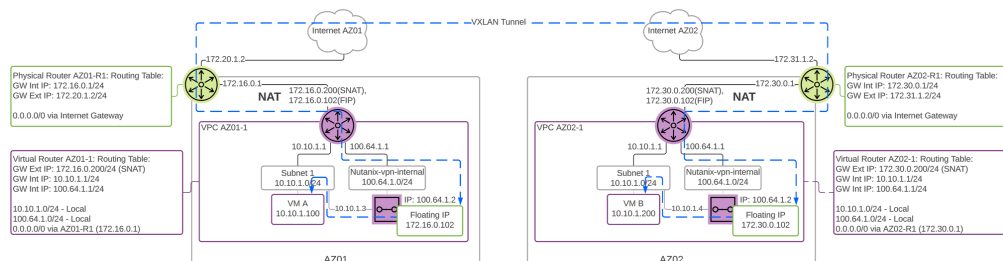


Flow Virtual Networking - Layer 2 VXLAN VTEP

When using two network gateway VMs to connect two VPCs, each VPN VM is assigned an external floating IP address, and the two VPN VMs must be able to communicate over these addresses. Therefore a VPC that uses layer 2 subnet extension must have a NAT external subnet with available floating IPs. If you are performing subnet extension for a VPC that uses routed, or NoNAT external subnets exclusively, you must add a NAT external subnet to the VPC.

Layer 2 Subnet Extension Detailed View for NAT

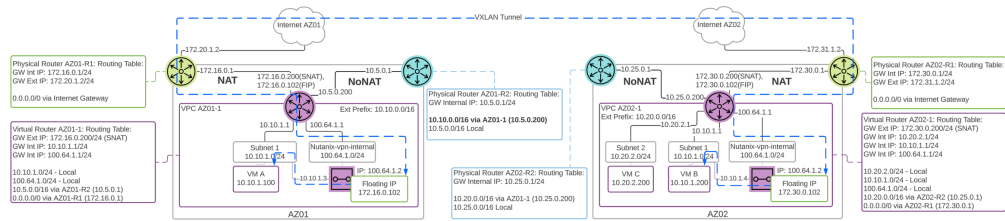
The following diagram shows the detailed interface view of a network gateway VM used for subnet extension between two VPCs with NAT external subnets. The network gateway VM for subnet extension has two interfaces, one interface inside the internal 100.64.1.0/24 overlay subnet, and another interface inside the overlay subnet being extended. The floating IP address of the network gateway VM is assigned to the 100.64.1.0/24 internal interface. You can think of subnet extension like plugging a virtual cable in between two networks. The network gateway VM acts as a layer 2 bridge with an interface in the target subnet.



Flow Virtual Networking - Subnet Extension NAT VPC Detail

Layer 2 Subnet Extension Detail View for Routed or NoNAT

The following detailed diagram shows both the Routed and NAT subnets needed for subnet extension in a Routed VPC. Careful planning is needed with routed networks and subnet extension to ensure traffic is directed to the right endpoints. In this example, traffic from overlay subnet 10.10.1.0/24 in AZ1 destined for 10.20.2.0/24 will route through the physical NoNAT network and not the subnet extension. Traffic from AZ2 10.10.1.0/24 destined to 10.20.2.0/24 will instead route through the Nutanix virtual router. Consider these routing paths when using subnet extension.



Flow Virtual Networking - Subnet Extension Routed VPC Detail

Layer 2 Subnet Extension Physical Device

The local network gateway can also connect to a remote physical VXLAN VTEP terminating device such as a physical switch. The physical device can be any standard VXLAN device from popular vendors such as, but not limited to, Cisco, Arista, and Juniper. Just enter the remote physical device IP address in the VXLAN VTEP connection.

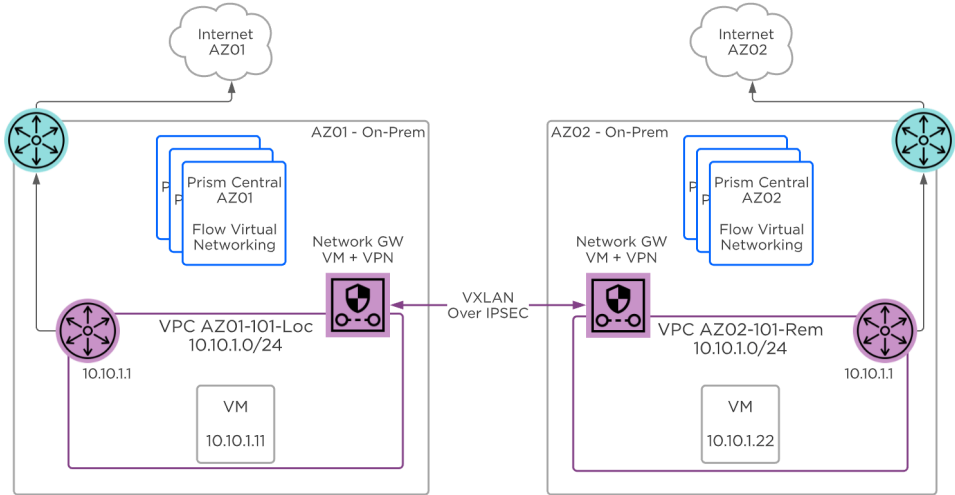
Traffic from the local subnet to the remote subnet is exchanged via layer 2 switching and encapsulated in unencrypted VXLAN. Each network gateway maintains a source MAC address table and can forward unicast or flooded packets to the remote subnet.

Pro tip

Use VXLAN VTEP only when traffic is traversing private or secured links, because this traffic is not encrypted.

Layer 2 VXLAN VTEP over VPN

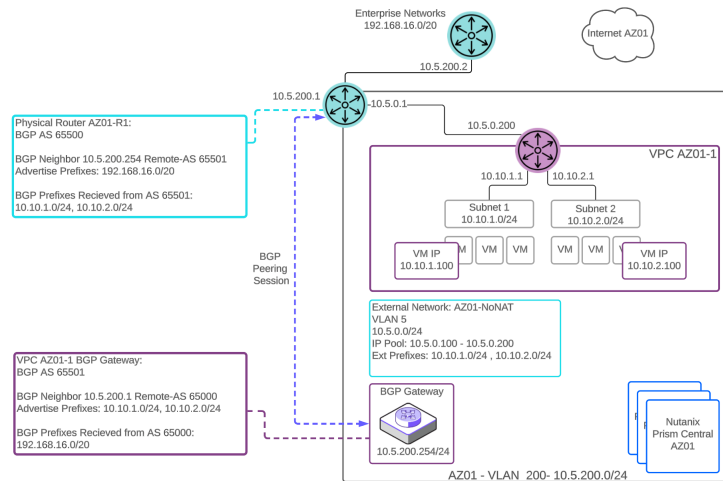
For extra security, the VXLAN connection can be tunneled through an existing VPN connection to add encryption. In this case, the network gateway VM provides the VXLAN and the VPN connections, so a network gateway VM is required in both the local and remote subnet.



Flow Virtual Networking - Layer 2 VXLAN VTEP over VPN

BGP Gateway

The BGP Gateway automates the exchange of routing information between a VPC and an external BGP peer. Use the BGP Gateway when a VPC has a Routed or NoNAT external network along with an overlay subnet whose IP address space has a requirement to communicate with external networks without using NAT, and this communication needs dynamic address advertisement. Unlike other network gateway options in Flow Virtual Networking, the BGP gateway is not in the data path. This gateway is part of the control plane.



BGP Gateway Deployed on a VLAN

BGP Gateway Deployment

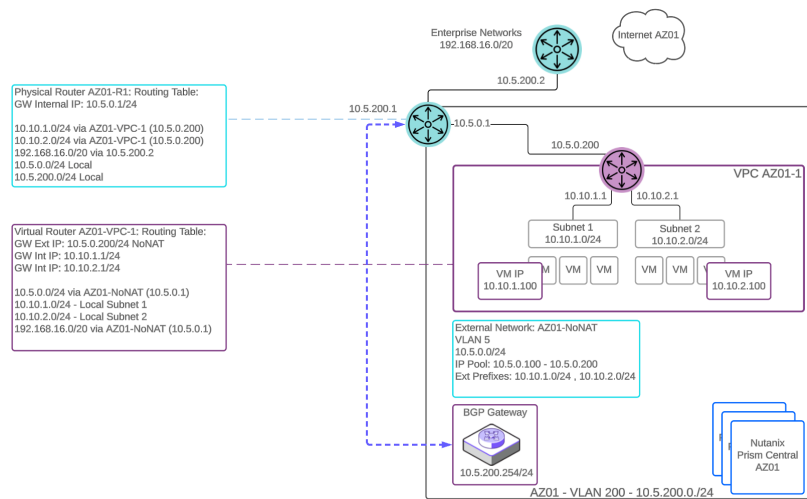
The BGP gateway can be deployed on an underlay VLAN-backed network or within the VPC in an overlay subnet.

The BGP gateway establishes peering sessions with upstream infrastructure routers. A peering session is used to advertise the reachability of the configured routable VPC subnets to the BGP neighbor. The gateway can advertise up to 20 routable prefixes. The next hop advertised from the VPC BGP gateway is the VPC router interface that is connected to the NoNAT, or Routed, external subnet. The gateway supports one session per BGP peer and up to 10 unique BGP peers.

The remote BGP neighbor in turn advertises this reachability to the rest of the network. The remote BGP neighbor also advertises external routes to the VPC BGP gateway. The gateway can learn and install up to 250 routes into the VPC route table per BGP peering session. The gateway will advertise the configured externally routable IP addresses. These IP addresses are defined in CIDR format and support x.x.x.x/16 through x.x.x.x/28 networks. This provides the flexibility to advertise your entire subnet or smaller subsets of network addresses. The gateway will learn routes from configured external peers and install these routes into the routing table of the VPC.

BGP Gateway in a VLAN

When the BGP gateway is deployed in a VLAN-backed subnet, it should have IP reachability to access the BGP peers. The VPC is configured with a Routed or NoNAT external subnet and the BGP gateway advertises the externally routable IP addresses of the VPC with the next-hop of the VPC virtual router.

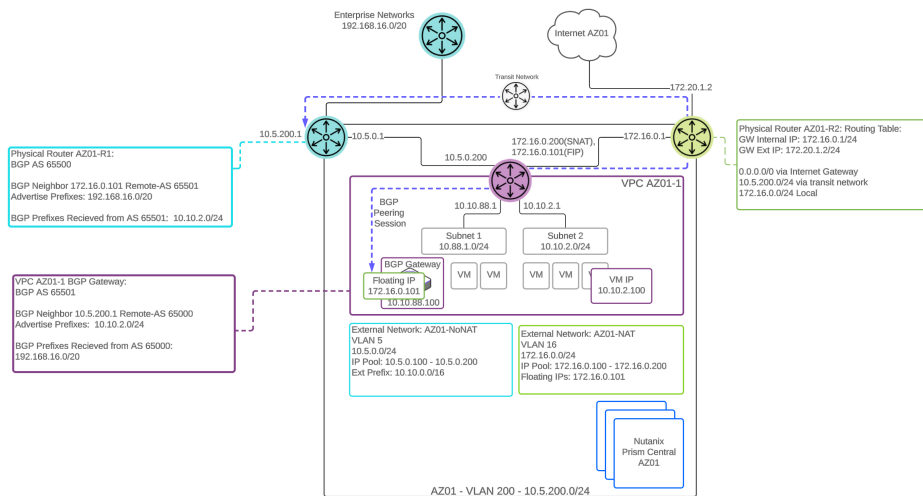


Routing Tables with BGP Established

Pro tip

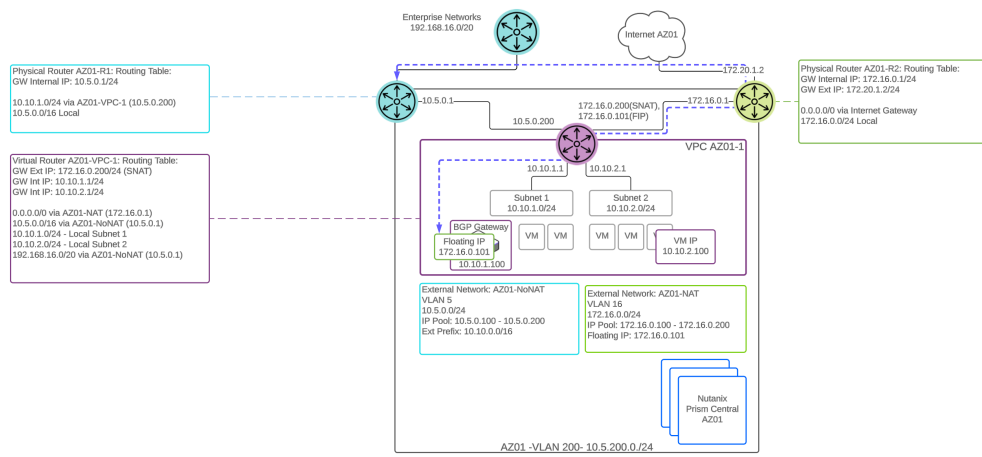
Deploying the BGP gateway in a VLAN-backed underlay network makes configuration easier as there are no requirements for a NAT external network and floating IP assignment.

BGP Gateway Deployed in Overlay Subnet



BGP Gateway Deployed in VPC Overlay Subnet

There is also the option to deploy the BGP gateway on an overlay subnet within the VPC. Placing the BGP gateway in an overlay subnet requires two types of external subnets, NoNat and NAT. The BGP gateway requires an assigned floating IP address from the NAT external subnet pool. The gateway uses this floating IP address to establish BGP sessions with remote neighbors. These peering sessions will be routed through the NAT external network path.



Routing Table with BGP Gateway Deployed in a VPC Overlay Subnet

The next hop advertised from the VPC BGP gateway is the VPC router interface that is connected to the NoNAT (routed) external subnet. All traffic to destinations learned from a BGP neighbor will exit the VPC via the NoNAT external subnet.