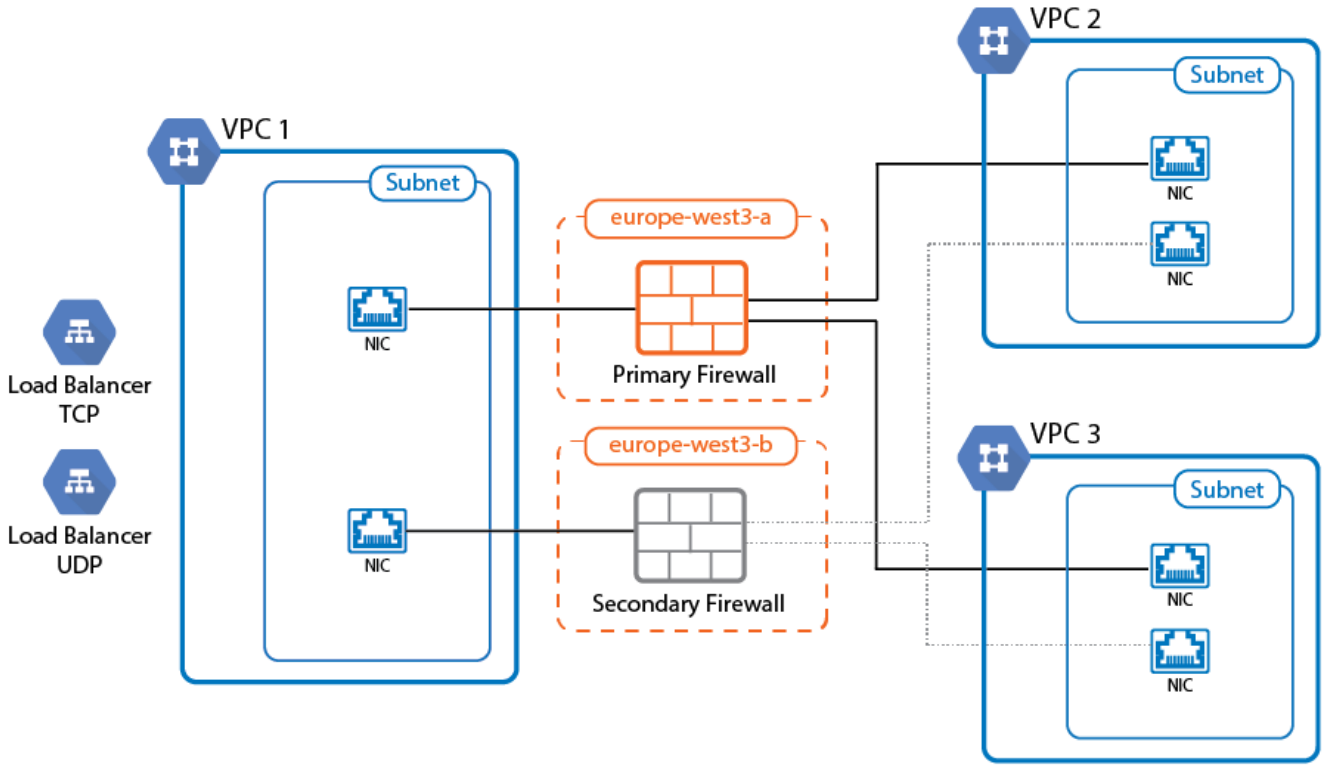
How to Configure a High Availability Cluster in Google Cloud

https://campus.barracuda.com/doc/73719432/

Running your CloudGen Firewall in a High Availability cluster in the Google Cloud ensures that even in the event of a datacenter failure in the cloud the other firewall can take over and your applications will remain reachable. All VPC networks must be in the same region; however, the two firewall instances are deployed into two different zones inside this region. The firewall instances are configured with one network interface per VPC network. Routing table in the VPC networks are configured to use the firewall as the target for traffic to the Internet and to other VPC networks. This allows the firewall to act both as the default gateway for Internet-bound traffic and as a segmentation firewall to VPC-to-VPC traffic. The number of network interfaces is determined by the number of CPU cores of the selected instance types. For example: for three VPCs, you need an instance with 3 CPU cores or more.

To rewrite the routes using the firewall as the target, a script must be placed in the /opt/phion/hooks/ha/ directory of each firewall. The script is executed every time the virtual server fails over and rewrites the routes to use the active firewall as the target.

To use the High Availability cluster with a single public IP address, add a TCP and/or UDP Google Network Load balancer. To use the load balancer, there must be a service on port 80 or 433 running on or behind the firewall because the Google legacy health check only allows HTTP and HTTP health checks. Use the SSL VPN service or the Cloud landing page. Alternatively, it is also possible to probe a web service behind the firewall, but an outage of the web service would result in the firewall to be considered unhealthy.



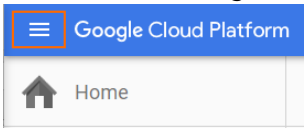
Before You Begin

- Download the Google Cloud Image from the Barracuda Networks Download Portal: https://dlportal.barracudanetworks.com.
- Create a custom service account and role for the High Availability cluster. For more information, see How to Create a Custom Role and Service Account for the CloudGen Firewall in the Google Cloud.
- Download the Google Cloud Takeover script needed for Step 18: gcp-ha-takeover.sh

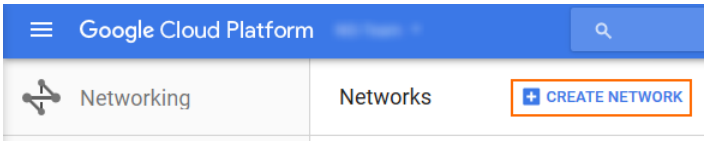
Step 1. Create the Hub VPC Network

Create the virtual private network where the two firewall instances will be running. Create a subnet for the firewall instances.

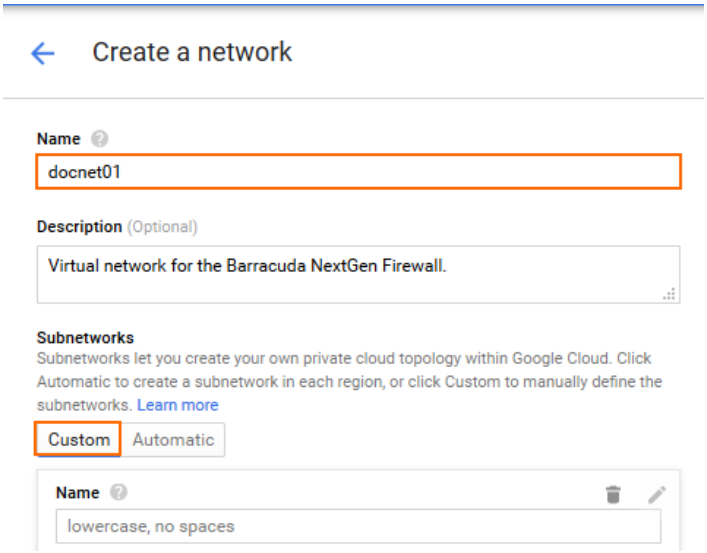
1. Log into the Google Cloud Platform. https://console.cloud.google.com/
2. Click the hamburger menu in the upper-left corner.



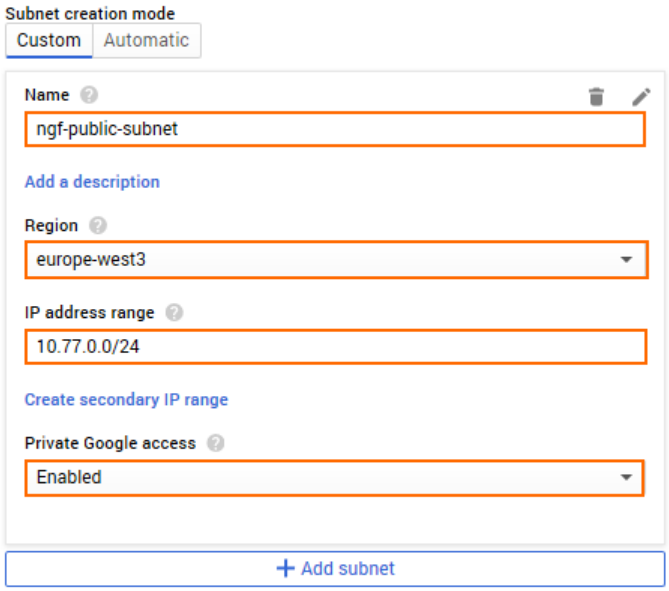
3. In the **Compute** section, click **Networking**.
4. In the main area, click **Create Network**.



5. Enter the **Name** and, in the **Subnetworks** section, click **Custom**.

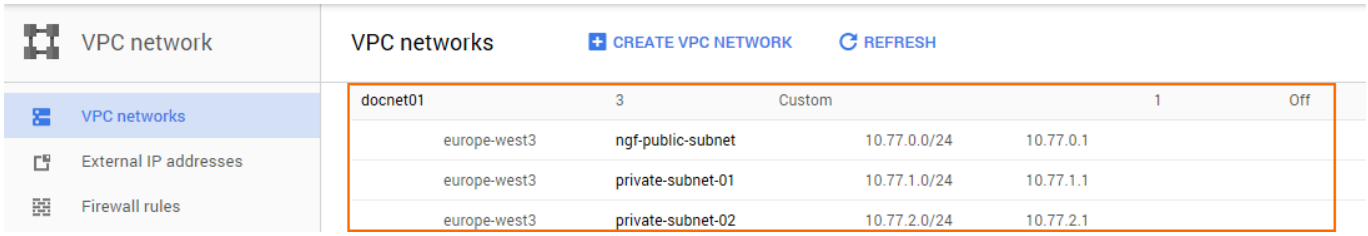


6. Create the public subnet:
   - **Name** – Enter **ngf-public-subnet**
   - **Region** – Select your region. All virtual networks must be in the same region.
   - **IP address range** – Enter the network in CIDR format. Do not use a network that overlaps with your on-premises network.
   - **Private Google access** – Select **Enabled**.



7. (optional) For each additional subnet in this virtual network, click **Add subnet**.
8. Click **Create**.

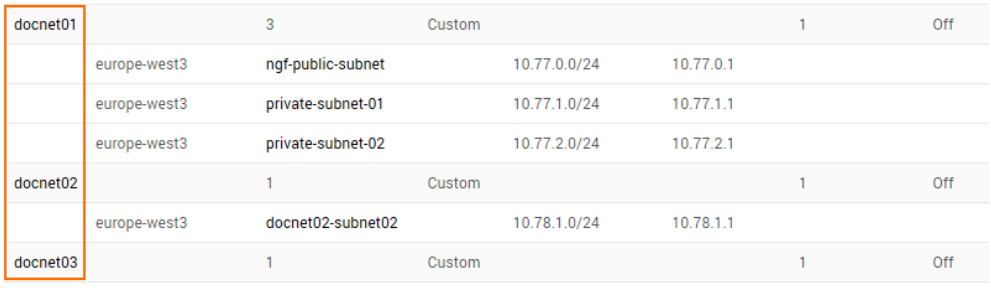
The VPC network for the firewall instances are now listed in the **VPC Networks** list.



Step 2. Create Additional VPC Networks

Create additional virtual networks with subnets in the same region. The number of virtual networks may not exceed the number of CPU cores on the firewall instance. Verify that the networks of the VPC networks do not overlap.

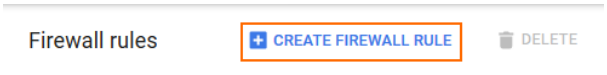
The VPC networks are now listed in the **VPC Networks** list.



Step 3. Create Google Firewall Rules

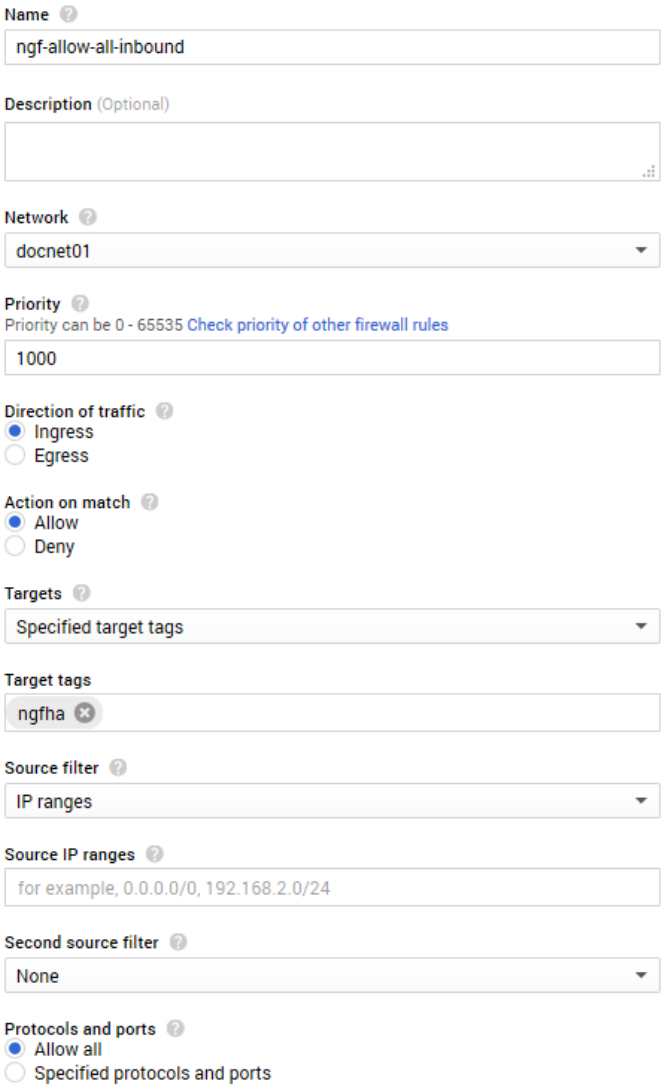
Google firewall rules must be configured for traffic to reach the firewall instances.

1. Go to https://console.cloud.google.com.
2. Click the hamburger menu in the upper-left corner.
3. In the **Compute** section, click **Networking**.
4. In the left menu, click **Firewall rules**.
5. In the main area, click **Create firewall rule**.



6. Create a firewall rule to allow incoming traffic from the Internet to your firewall instances:
   - **Name** – Enter the firewall rule name.

- **Network** – Select the network created in Step 1.
- **Priority** – Set a priority lower than 1000.
- **Direction of traffic** – Select **Ingress**.
- **Action on match** – Select **Allow**.
- **Targets** – Select **Specified target tags**.
- **Target tags** – Enter the tag `ngfha` that will be assigned to the firewall instances.
- **Source filter** – Select **IP ranges**.
- **Source IP ranges** – Enter `0.0.0.0/0`.
- **Protocols and ports** – Enter a semicolon-delimited, lower-case list of protocols and ports, or select **Allow all**.



7. Click **Create**.
8. In each VPC network, create a firewall rule to allow traffic from selected subnets to the firewall:
   - **Name** – Enter the firewall rule name.
   - **Network** – Select one of the VPC network created in Step 2. Select the VPC network created in Step 1 to allow traffic from private subnets in the hub VPC network to the firewall.
   - **Priority** – Set a priority lower than 1000.

- **Action on match** – Select **Allow**.
- **Targets** – Select **Specified target tags**.
- **Target tags** – Enter the tag `ngfha` that will be assigned to the firewall instances.
- **Source filter** – Select **Subnetworks**.
- **Subnetworks** – Select the subnets and click **OK**.
- **Protocols and ports** – Enter a semicolon-delimited, lower-case list of protocols and ports, or select **Allow all**.

9. Click **Create**.

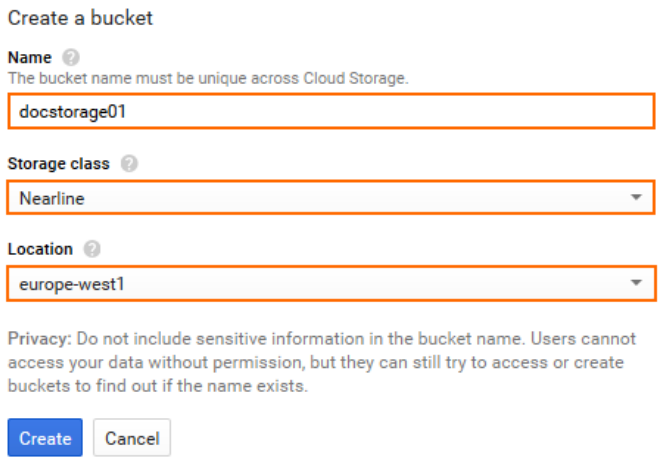
Traffic is now allowed to and from the firewall instances from the Internet and the additional VPC networks, as well as the private networks in the hub VPC network.

Step 4. Create a Storage Bucket and Upload the Image

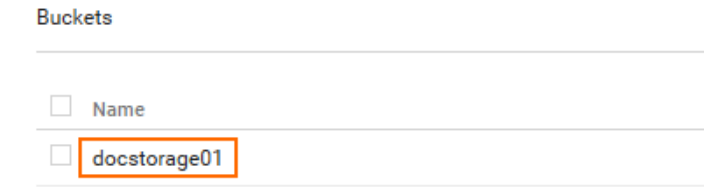
Upload the image to Google Cloud. If the upload through the browser does not work, you can instead use Google Cloud SDK to upload the image.

1. Go to https://console.cloud.google.com.
2. Click the hamburger menu in the upper left corner.
3. In the **Storage** section, click **Storage**.
4. In the main area, click **Create bucket**.

5. Create a storage bucket:
   - **Name** – Enter a unique name.
   - **Storage class** – Select a storage class depending on your preferences.
   - **Location** – Select the location matching the region you are deploying in.

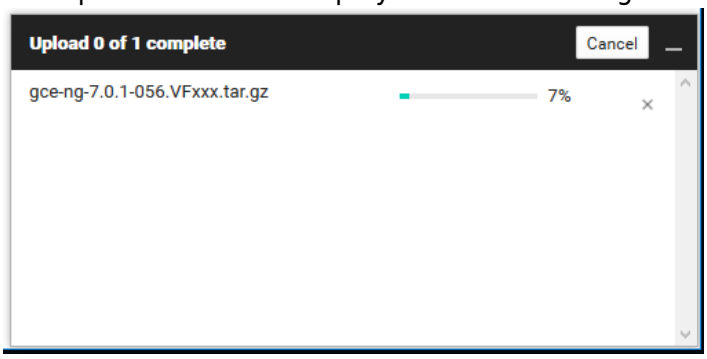


6. Click **Create**.
7. Click the storage bucket you just created.

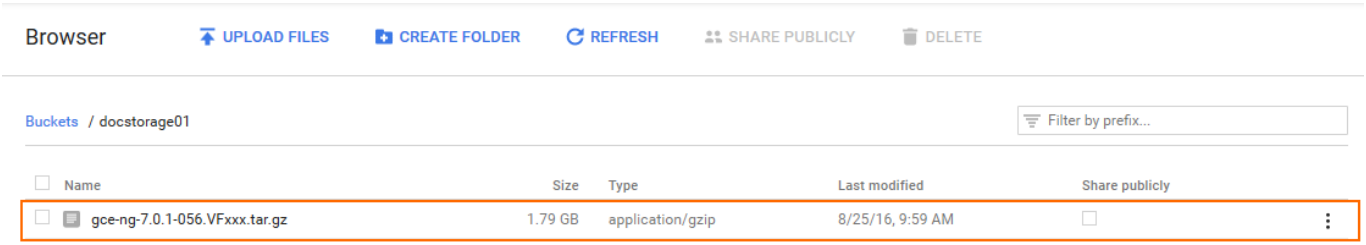


8. Click **Upload Files** and select the firewall image you previously downloaded from the Barracuda Download Portal.

9. The upload window is displayed in the lower-right corner.



The image is now listed in the file list of the storage bucket.

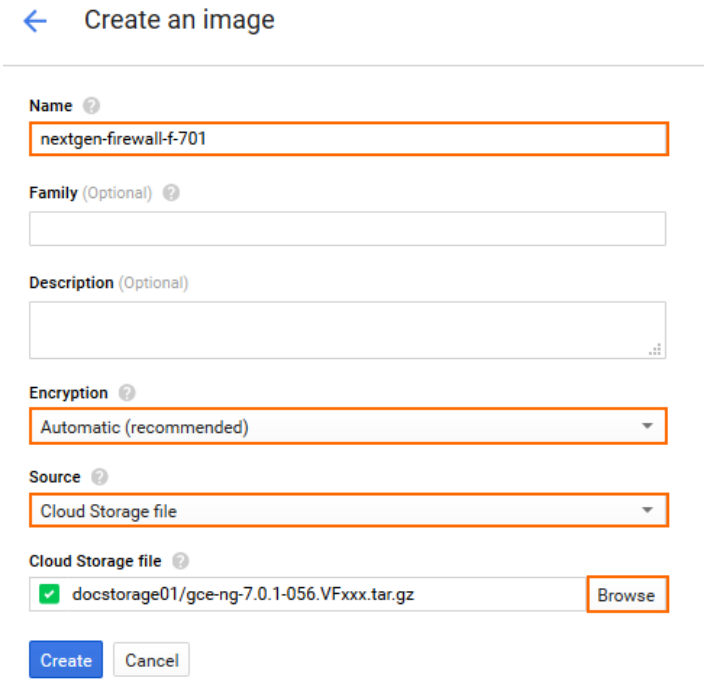


Step 5. Create a Compute Engine Image from the Uploaded Disk Image

To be able to deploy a firewall from the disk image uploaded in Step 3, you must create a Google Compute Engine image. The firewall is created with one dhcp interface. DHCP reservation can be done manually (static) or automatically by Google during deployment. Once assigned, the internal IP address does not change.

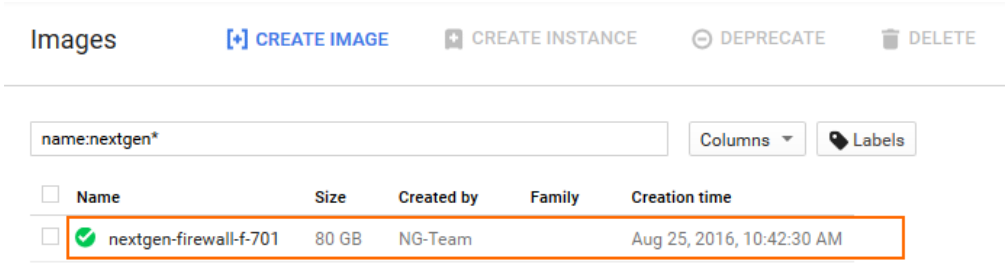
1. Go to https://console.cloud.google.com.
2. Click the hamburger menu in the upper-left corner.
3. In the **Compute** section, click **Compute Engine**.
4. In the left menu, click **Images**.
5. In the main area, click **Create Images**.

6. Create an image using the disk image uploaded in Step 3.
   - **Name** – Enter a name for the firewall image.
   - **Encryption** – Select **Automatic (recommended)**.
   - **Source** – Select **Cloud Storage file**.
   - **Cloud Storage File** – Click **Browse** and select the disk image in the storage bucket created in Step 3.



7. Click **Create**.

The firewall image is now listed in the **Images** list.



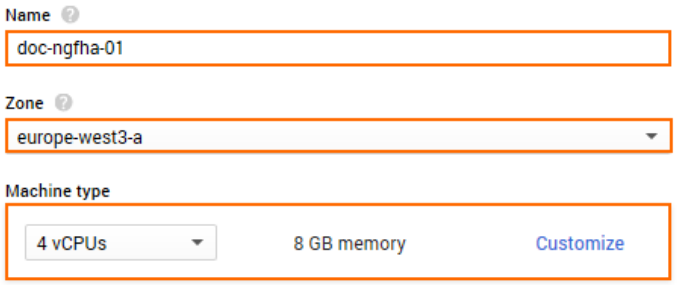
Step 6. Create the Primary Firewall Instance

Launch the primary firewall instance into the public subnet of the hub VPC network. Add one network interface per additional VPC network. The number of CPU cores must be at least equal to the required

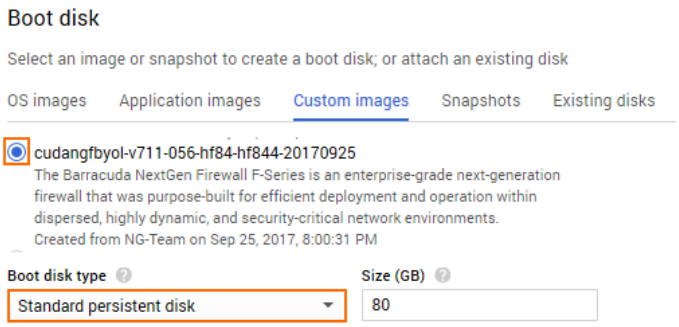
number of network interfaces.

1. Go to https://console.cloud.google.com.
2. Click the hamburger menu in the upper-left corner.
3. In the **Compute** section, click **Compute Engine**.
4. In the main area, click **Create instance**.

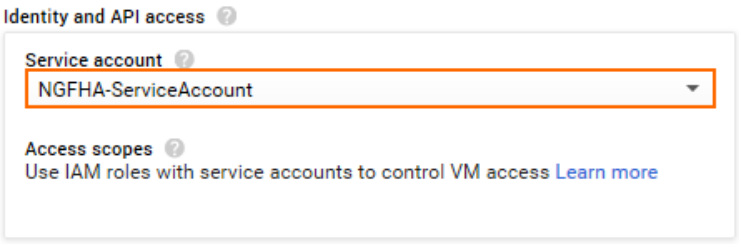
5. Enter a lowercase **Name** for the primary firewall instance.
6. Select the **Zone**. The zone must be in the same region as the public subnet in the network created in Step 1.
7. Select **Machine type**. Verify that the number of vCPUs matches the number of cores included in your CloudGen Firewall license and the number of network interfaces used by the instance.



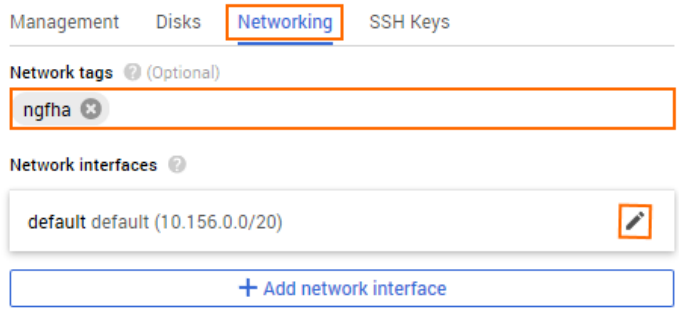
8. In the **Boot disk** section, click **Change**.
9. Click the **Custom Images** tab.
10. Select the image you created in Step 5.
11. Select the **Boot disk type**:
    - **Standard persistent disk**
    - **SSD persistent disk**



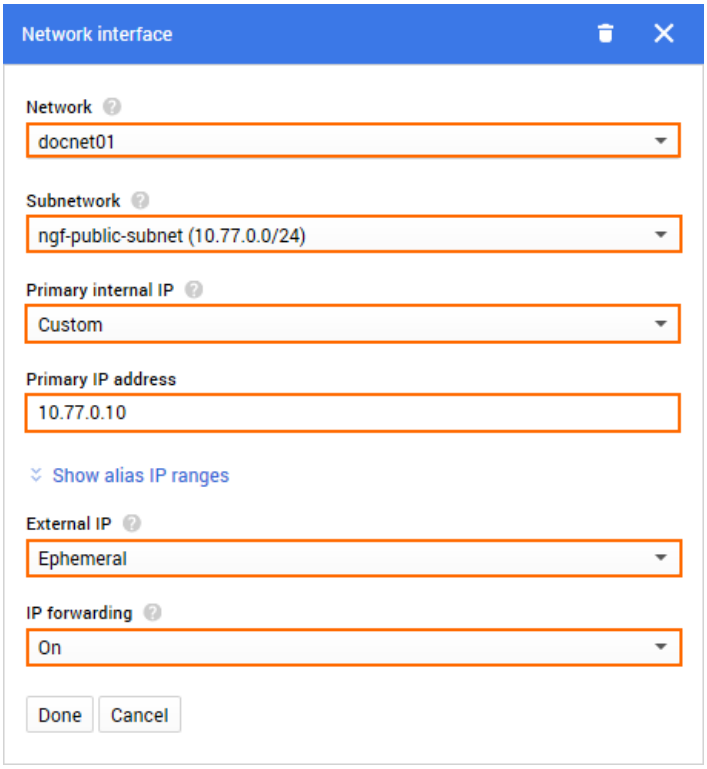
12. Click **Select**.
13. Select the dedicated **Service account** associated with the custom role created for the High Availability cluster. For more information, see How to Create a Custom Role and Service Account for the CloudGen Firewall in the Google Cloud.
14. In the **Access scopes** section, select **Allow full access to all Cloud APIs**.



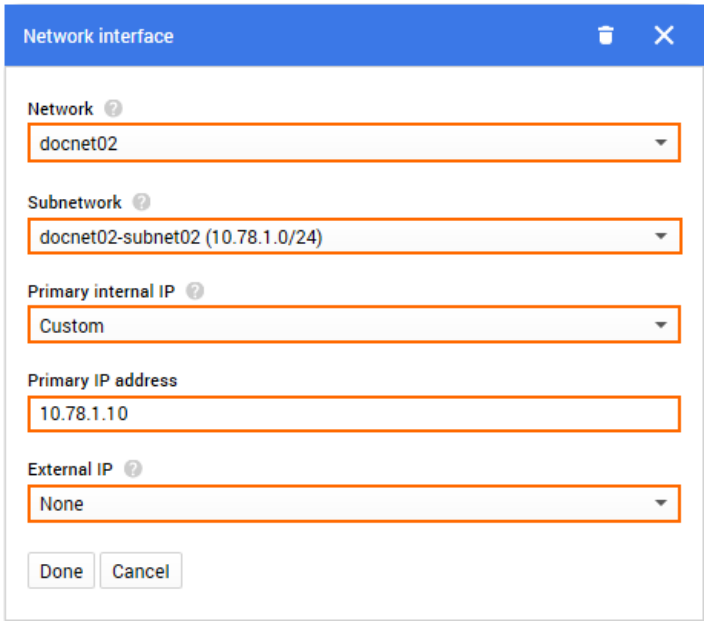
15. Below the **Firewall** section, click **Management, disk, networking, SSH keys**.
16. Click on the **Networking** tab.
17. Add `ngfha` to the **Network tags**.
18. In the **Network Interfaces** section, click the edit icon for the **default** network interface.



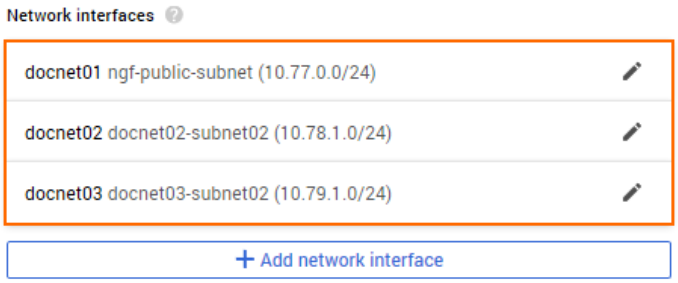
19. Configure the default network interface:
    - **Network** – Select the network created in Step 1.
    - **Subnetwork** – Select the public subnet created in Step 1.
    - **Internal IP** – Select **Custom**.
    - **Internal IP address** – Enter a free IP address in the subnet. The first IP address in the subnet is reserved for the gateway.
    - **External IP** – Select a reserved external IP address; otherwise, select **Ephemeral** to use a dynamic public IP address.
    - **IP forwarding** – Select **On**.



20. For each additional network interface, click **Add network interface**.
21. Configure the additional network interface:
    - **Network** – Select one of the additional VPC networks created in Step 2.
    - **Subnetwork** – Select a subnet in the VPC network that is in the same region as the firewall instance.
    - **Internal IP** – Select **Custom**.
    - **Internal IP address** – Enter a free IP address in the subnet. The first IP address in the subnet is reserved for the gateway.
    - **External IP** – Select **None**.



22. Click **Done**. All network interfaces are now listed in the **Network interfaces** section.



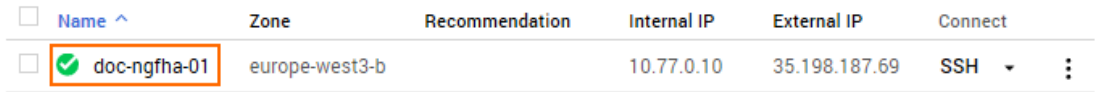
23. Click **Create**.

The primary firewall instance is now started.

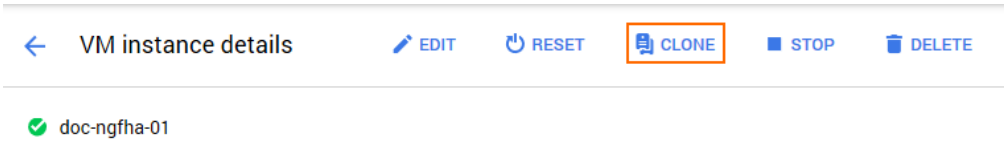
Step 7. Create the Secondary Firewall Instance

Deploy the secondary firewall of the High Availability cluster into the same subnet, but in a different zone. This ensures that one firewall of the cluster will always be running, even in case of a datacenter failure within the Google Cloud. To ease configuration clone the primary firewall and change the configuration to match the settings of the secondary firewall.

1. Go to https://console.cloud.google.com.
2. Click the hamburger menu in the upper-left corner.
3. In the **Compute** section, click **Compute Engine**.
4. Click on the primary firewall instance created in Step 4. The **VM instance details** page opens.



5. Click **CLONE**.



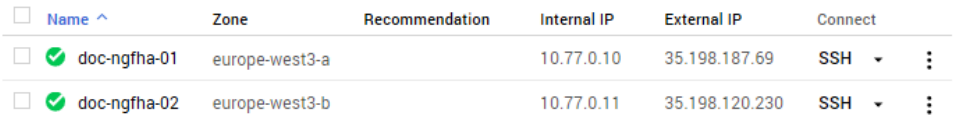
6. Enter the **Name** for the secondary firewall instance.
7. Select a **Zone**. Select different zones in the same region for the two firewalls in the High Availability cluster.

8. Below the **Firewall** section, click **Management, disk, networking, SSH keys**.
9. Click the **Networking** tab
10. Add `ngfha` to the **Network tags**.
11. In the **Network Interfaces** section, click the edit icon for the **default** network interface
12. Click the edit icon for the first network interface:

- **Network** – Select the network created in Step 1.
- **Subnetwork** – Select the public subnet created in Step 1.
- **Internal IP** – Select **Custom**.
- **Internal IP address** – Enter a free IP address in the subnet.
- **External IP** – Select a reserved external IP address; otherwise, select **Ephemeral** to use a dynamic public IP address.
- **IP forwarding** – Select **On**.

13. Click **Done**.
14. Click the edit icon for the other network interfaces, and assign free custom internal IP addresses in the subnets.
15. Click **Create**.

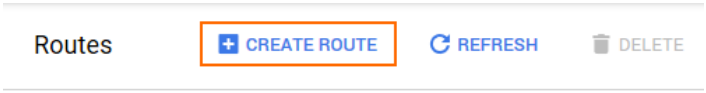
Both the primary and secondary firewalls of the High Availability cluster are now running.



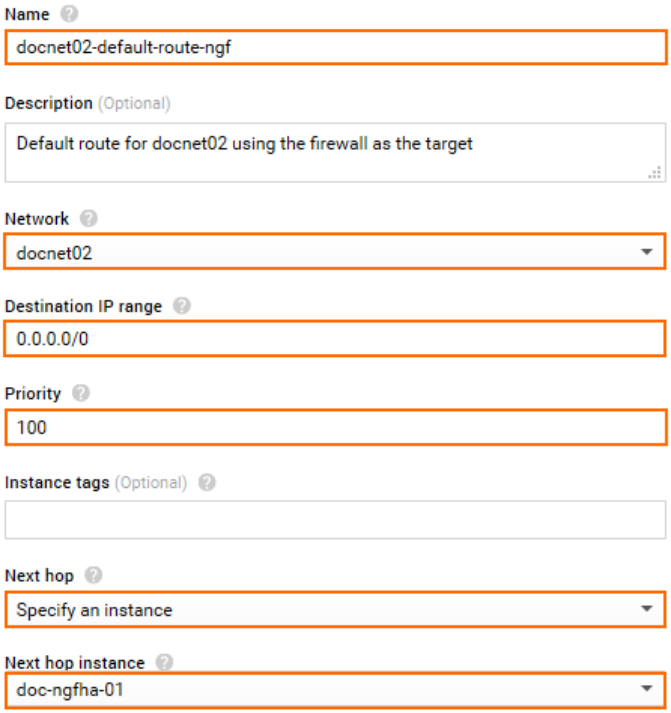
Step 8. Configure a Default Route for the VPC Networks to Use the Primary Firewall

For each VPC network, create a default route for the client instances to use the active firewall as the target.

1. Go to https://console.cloud.google.com.
2. Click the hamburger menu in the upper-left corner.
3. Click **VPC network**.
4. In the left menu, click **Routes**.
5. Click **Create Route**.



6. Configure the route:
   - **Name** – Enter a name for the route.
   - **Network** – Select the VPC network from the list.
   - **Destination IP range** – Enter `0.0.0.0/0`.
   - **Priority** – Enter `100`.
   - **Next hop** – Select **Specify an instance**.
   - **Next hop instance** – Select the active firewall.



7. Click **Create**.

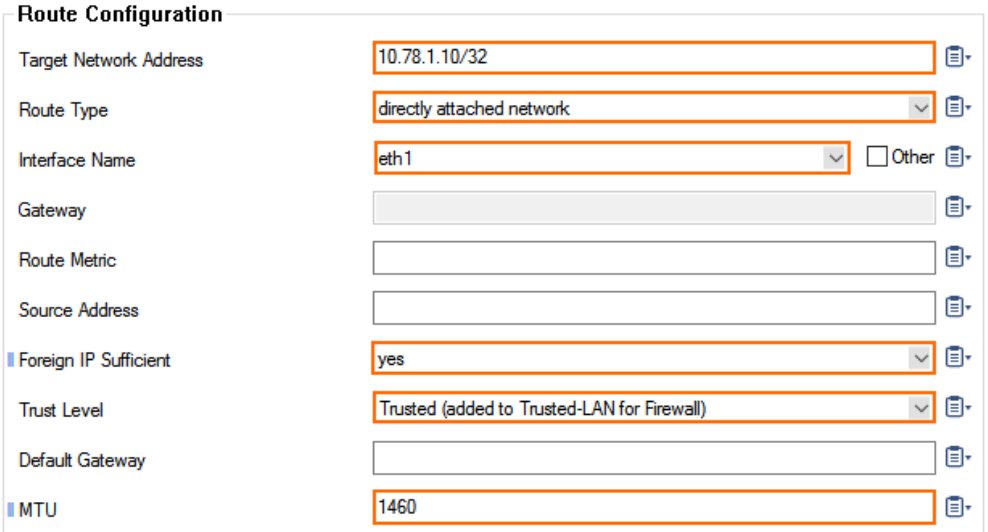
All traffic leaving the VPC is now being sent through the active firewall. If you have attached two additional VPC networks to the firewall, you should have at least two routes: one for each VPC network. If you also have private subnets in the hub VPC network, three routes must be created. The next hop is the IP address of the firewall's network interface in that VPC network subnet.

Step 9. Add an Additional Network Interface to the Primary Firewall Configuration

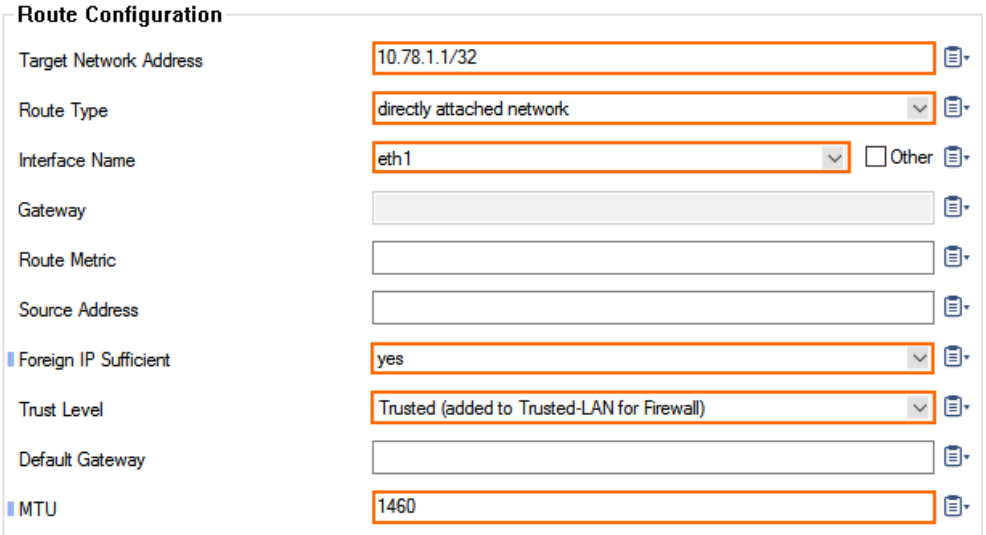
Add and configure the additional network interfaces on the primary firewall.

1. Log into the primary firewall:
   - **IP Address** – The public IP address listed in the **External IP** column on the **VM Instances** page.
   - **User** – Enter `root`
   - **Password** – The name of the instance.
2. Go to **CONFIGURATION > Configuration Tree > Box > Network**.
3. Click **Lock**.
4. In the left menu, select **Interfaces**.
5. Double-click the entry in **Network Interface Cards**. The **Network Interface Configuration** window opens.
6. Change the **Number of Interfaces** to the number of interfaces attached to the firewall.
7. Click **Send Changes**.
8. In the left menu, select **Routing**.

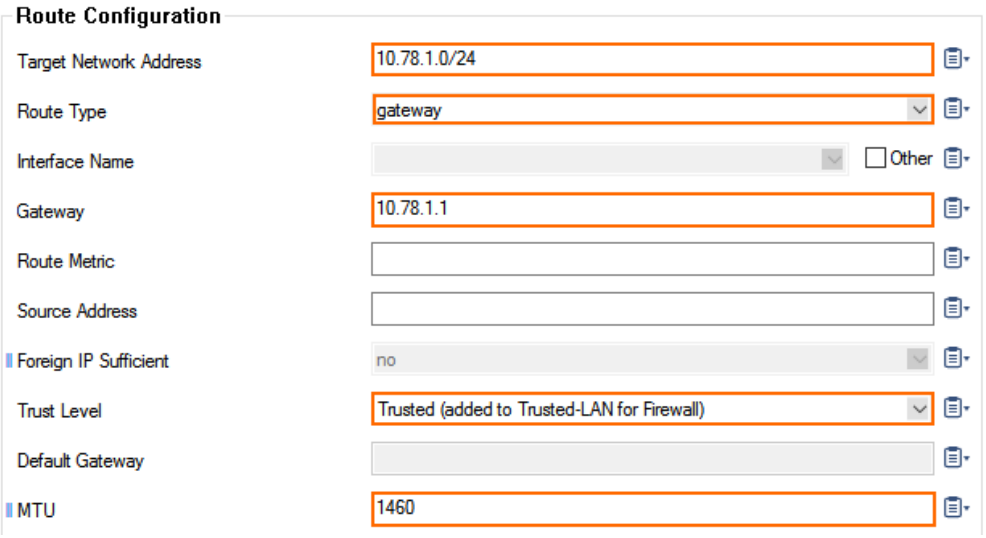
9. In the left menu, expand the **Configuration Mode** section and click **Switch to Advanced**.
10. Create a new **directly attached route** for private IP address of the network interface:
    - **Target Network Address** – Enter the private IP address of the network interface with a /32 subnet mask E.g., `10.78.1.10/32`
    - **Route Type** – Select **directly attached network**.
    - **Interface Name** – Select the network interface. E.g., eth1
    - **Foreign IP Sufficient** – Select **yes**.
    - **Trust Level** – Select **Trusted**.
    - **MTU** – Enter 1460.



11. Create a new **directly attached route** for the default subnet gateway assigned by Google. The default gateway is always the first IP address in the subnet:
    - **Target Network Address** – Enter the first IP address in the subnet with /32 subnet mask. E.g., 10.78.1.1/32
    - **Route Type** – Select **directly attached network**.
    - **Interface Name** – Select the network interface. E.g., eth1
    - **Foreign IP Sufficient** – Select **yes**.
    - **Trust Level** – Select **Trusted**.
    - **MTU** – Enter 1460.



12. Create a new **gateway route** for the subnet using the default subnet gateway:
    - **Target Network Address** – Enter the subnet in CIDR format. E.g., `10.78.1.0/24`
    - **Route Type** – Select **gateway**.
    - **Gateway** – Enter the first IP address in the subnet. E.g., `10.78.1.1`
    - **Trust Level** – Select **Trusted**.
    - **MTU** – Enter `1460`.

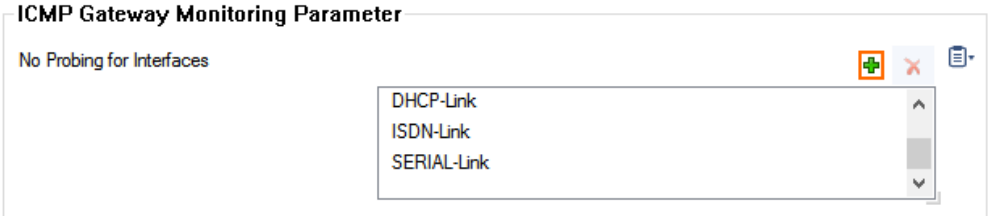


13. Click **Send Changes** and **Activate**.

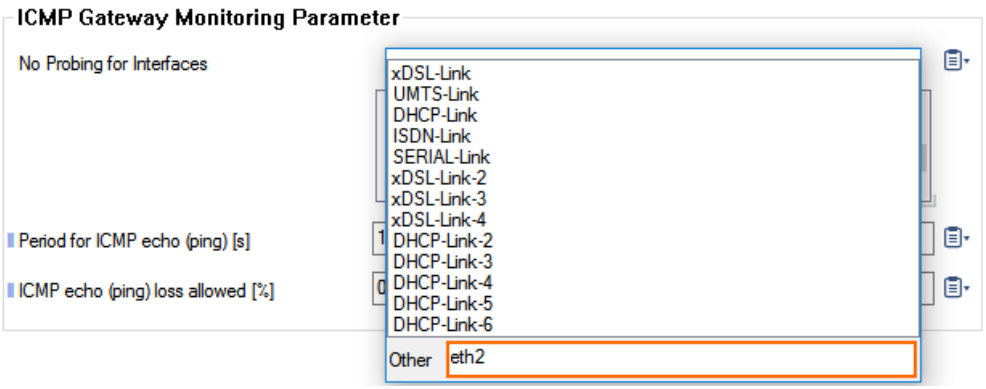
Step 10. Disable ICMP Gateway Monitoring for Additional Network Interfaces

Disable ICMP gateway monitoring for all additional network interfaces.

1. Log into the primary firewall.
2. Go to **CONFIGURATION > Configuration Tree > Box > Infrastructure Services > Control**.
3. Click **Lock**.
4. For each additional network interface click **+** to add an entry in the **No Probing for Interfaces** table,



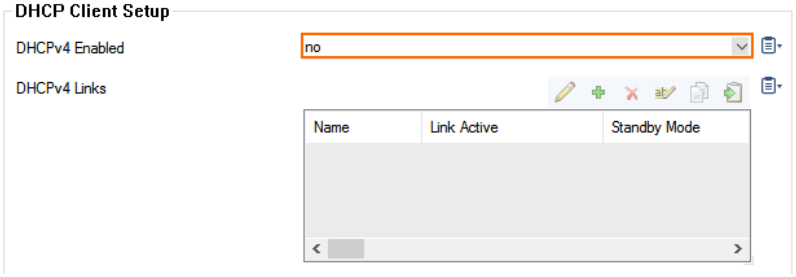
5. If the interface is not in the list enter it in the **Other** field.



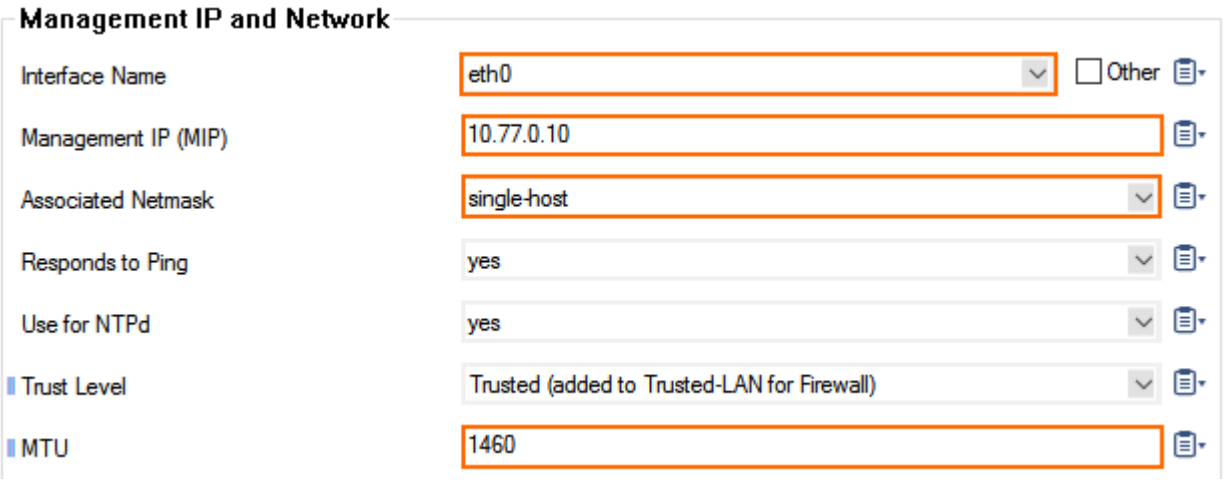
6. Click **Send Changes** and **Activate**.

Step 11. Change the Primary Firewall Configuration to Use the Static Network Interface

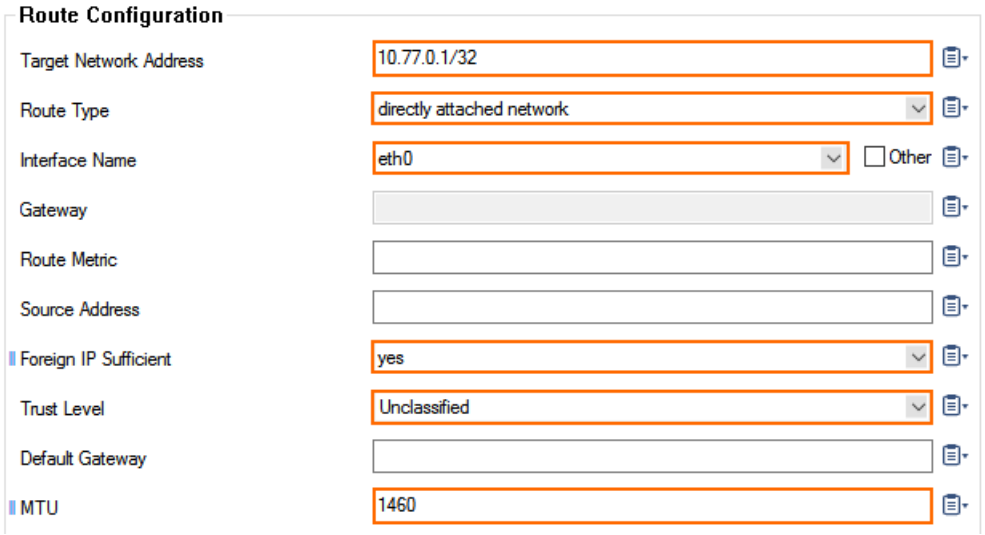
1. Log into the primary firewall.
2. Go to **CONFIGURATION > Configuration Tree > Box > Network**.
3. In the left menu, expand the **Configuration Mode** section and click **Switch to Advanced**.
4. Click **Lock**.
5. in the left menu, click **xDSL/DHCP/ISDN**.
6. Delete the **DHCP01** entry in the **DHCP Links** list.
7. Select **No** from the **DHCPv4 Enabled** drop-down list.



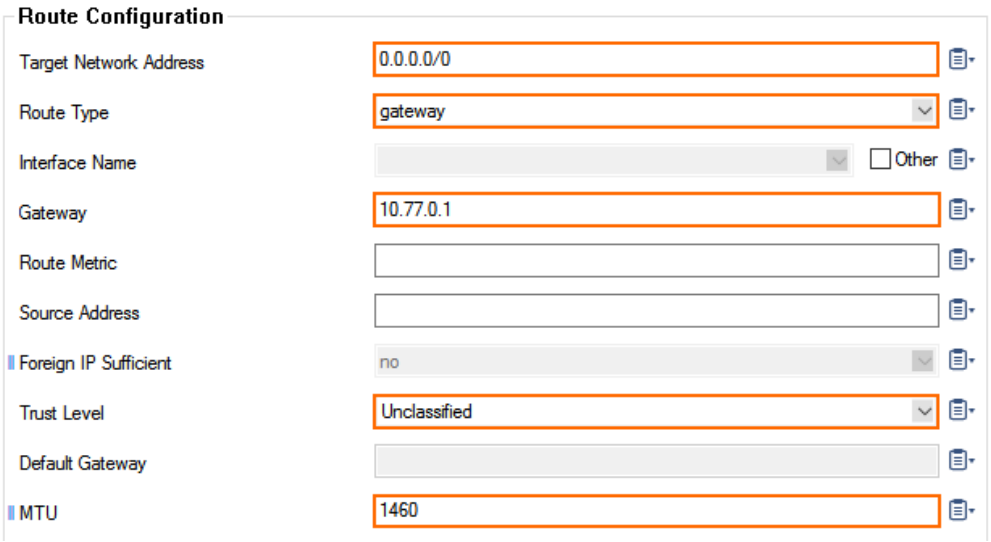
8. In the left menu, click **IP Configuration**.
9. In the **Management IP and Network** section, reconfigure the management IP:
   - **Interface Name** – Select **Other** and enter `eth0`.
   - **Management IP** – Enter the private IP address of the primary firewall. Go to **CONTROL> Network**. The private IP address is assigned to the dhcp interface.
   - **Associated Netmask** – Select **single-host**.
   - **MTU** – Enter 1460.



10. In the left menu, click **Routing**.
11. Create a new **directly attached route** for the default subnet gateway assigned by Google. The default gateway is always the first IP address in the subnet:
    - **Target Network Address** – Enter the first IP address in the subnet with /32 subnet mask. E.g., 10.77.0.1/32
    - **Route Type** – Select **directly attached network**.
    - **Interface Name** – Select **Other** and enter `eth0`.
    - **Foreign IP Sufficient** – Select **yes**.
    - **Trust Level** – Select **Unclassified**.
    - **MTU** – Enter 1460.



12. Click **OK**.
13. Click **+** in the **Routes** table and add the default route:
    - **Target Network Address** – Enter `0.0.0.0/0.`
    - **Route Type** – Select **gateway**.
    - **Gateway** – Enter the first IP address in the subnet. E.g., `10.77.0.1`
    - **Trust Level** – Select **Unclassified**.
    - **MTU** – Enter 1460.



14. Click **OK**.
15. Click **Send Changes** and **Activate**.

Open the **CONTROL > Network** page. Your interface and IP address are now static.

Step 12. Activate the Network Changes

1. Go to **CONTROL > Box**.
2. In the left menu, expand the **Network** section and click **Activate new network configuration**.
3. Select **Failsafe**.

Step 13. Configure the DNS Server

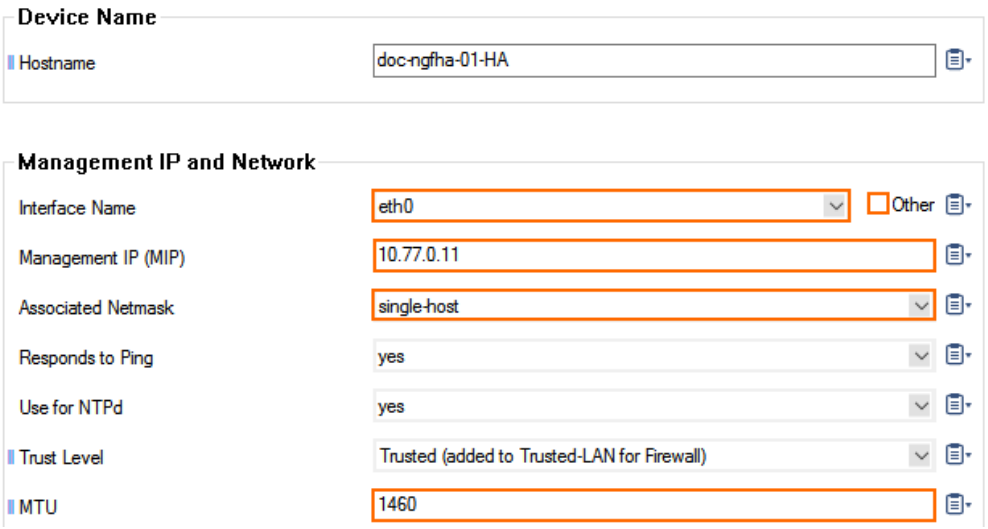
Add the first IP address of the subnet as the DNS server (e.g., 10.77.0.1). Do not use external DNS servers because, otherwise, it is not possible to resolve the internal Google metadata service used by the HA failover script.

For more information, see How to Configure DNS Settings.

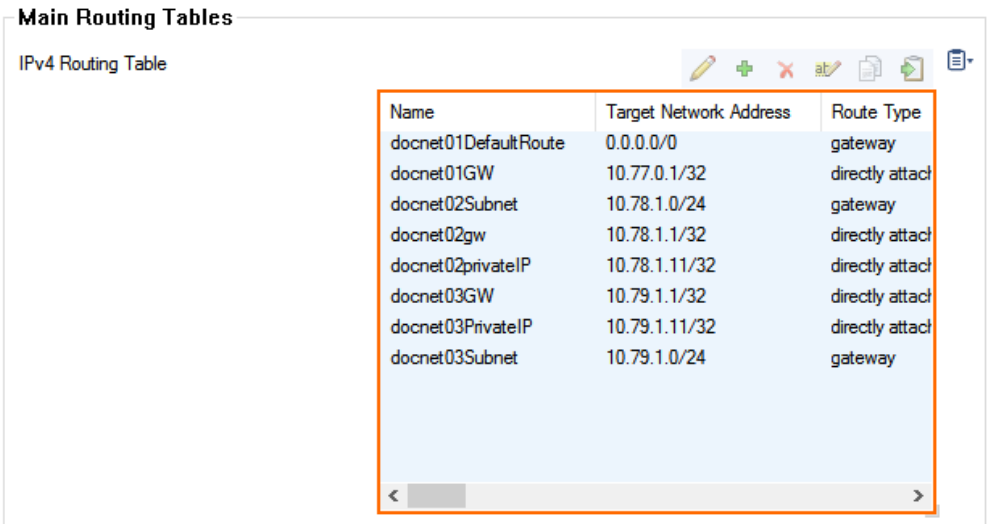
Step 14. Create the DHA Cluster Configuration

Create the DHA cluster configuration for the secondary firewall and configure the routing configuration.

1. Log into the primary firewall.
2. Go to **CONFIGURATION > Configuration Tree > Box**.
3. Right-click **Box** and select **Create DHA box**. At the bottom of the **Config Tree**, the **HA Box** configuration node is added.
4. Go to **CONFIGURATION > Configuration Tree > HA Box > HA Network**.
5. In the left menu, expand the **Configuration Mode** section and click **Switch to Advanced**.
6. In the **Management IP and Network** section, reconfigure the management IP:
   - **Interface Name** – Select **Other** and enter `eth0`.
   - **Management IP** – Enter the private IP address of the secondary firewall. On the secondary firewall, go to **CONTROL> Network**. The private IP address is assigned to the dhcp interface.
   - **Associated Netmask** – Select **single-host**.
   - **MTU** – Enter `1460`.



7. Edit the directly attached routes for the private IP addresses to match the secondary firewall custom internal IP address on that network interface.
8. Verify that the routing is configured analog to the primary firewall:
   - **For the hub VPC network** – One gateway route and on directly attached route.
   - **For each additional VPC network** – One gateway and two directly attached routes. The directly attached routes for the private IP addresses must be changed to match the custom internal IP addresses of the secondary firewall on that interface.

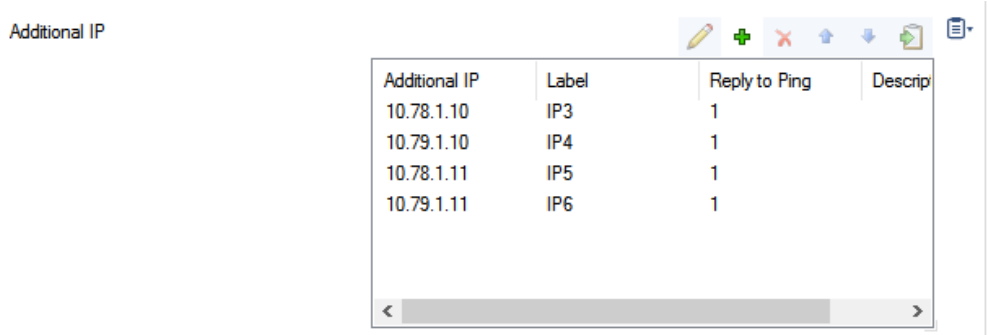


9. Click **Send Changes** and **Activate**.

Step 15. Add the Private IPs to the Virtual Server IPs and Add the Secondary Firewall to the Virtual Server

Add the custom private IP addresses of both firewalls to the additional network interfaces to the virtual server IP addresses.

1. Log into the primary firewall
2. Go to **CONFIGURATION > Configuration Tree > Box > Virtual Servers > S1 > Server Properties**.
3. Click **Lock**.
4. From the **Backup Box** drop-down list, select **Other-Box**.
5. In the **Additional IP** table, add the private IP addresses for both firewalls.



6. Click **Send Changes** and **Activate**.

Step 16. Join the High Availability Cluster

Step 16.1 Export PAR File

1. On the primary firewall, create the PAR file:
2. Go to **CONFIGURATION > Configuration Tree > Box**.
3. From the **Config Tree**, right-click **Box** and select **Create PAR file for HA box**.
4. Save the PAR file to your local hard disk drive.

Step 16.2 Import the PAR File on the Secondary Firewall

1. Log in to the secondary firewall:
   - **IP Address** – The public IP address listed in the **External IP** column on the **VM Instances** page.
   - **User** – Enter `root`
   - **Password** – The name of the instance.
2. Go to **CONFIGURATION > Configuration Tree > Box**.
3. From the **Config Tree**, right-click **Box** and select **Restore from PAR file**.
4. Click **OK**.
5. Select the boxha.par file created in Step 15.1 and click **OK**.
6. Click **Activate**.

Step 16.3 Activate the Network Configuration on the Secondary Firewall

1. Go to **CONTROL > Box**.
2. Click **Trust**.
3. In the left menu, expand **Network** and click **Activate new network configuration**.
4. Select **Failsafe** as the activation mode.

Step 17. Activate and License the High Availability Cluster

Activate and license the High Availability cluster. Activate the secondary firewall first. Then, activate the primary firewall.

For more information, see How to Activate and License a Standalone High Availability Cluster.

Step 18. Add the High Availability Failover Script to the Primary and Secondary Firewalls

To rewrite the default routes using the firewall as the default gateway to always use the active firewall in the High Availability cluster, copy the failover script to the /opt/phion/hooks/ha/ on both firewalls. The script is executed automatically every time the virtual server fails over.

Step 18.1. Enable SSH Root Access

1. Log into the primary firewall.
2. Go to **CONFIGURATION > Configuration Tree > Box > Advanced Configuration > SSH**.
3. Click **Lock**.
4. In the left menu, select **Basic Setup**.
5. From the **Permit Root Login** drop-down list, select **key-only** .
6. Click **Send Changes** and **Activate**.

Step 18.2. Create the Failover Script on both Firewalls

Create this script on both firewalls.

1. Download the gcp-ha-takeover.sh script.
2. Log into the primary firewall via SSH.
3. Go to the
4. Copy the **gcp-ha-takeover.sh** script to the **/opt/phion/hooks/ha/** directory on the firewall
5. Make the script executable:

   ```
   chmod +x /opt/phion/hooks/ha/gcp-ha-takeover.sh
   ```

6. Repeat on the secondary firewall.

Step 19. (optional) Add the Google Network Load Balancer

To use only one public-facing IP address, it is also possible to use the Google Network Load Balancer in front of the High Availability cluster. To use the load balancer, a service on port 80 or 443 must be reachable for the health check of the load balancer. TCP and UDP services require separate load balancers. To use a service on the firewall for probing create an **App Redirect** rule redirecting HTTP traffic to the fwauth daemon running on 127.0.0.1:451.

For more information, see https://cloud.google.com/compute/docs/load-balancing/network/

Figures