TKG Managed Service Networking

VMware App Modernization

Table of contents

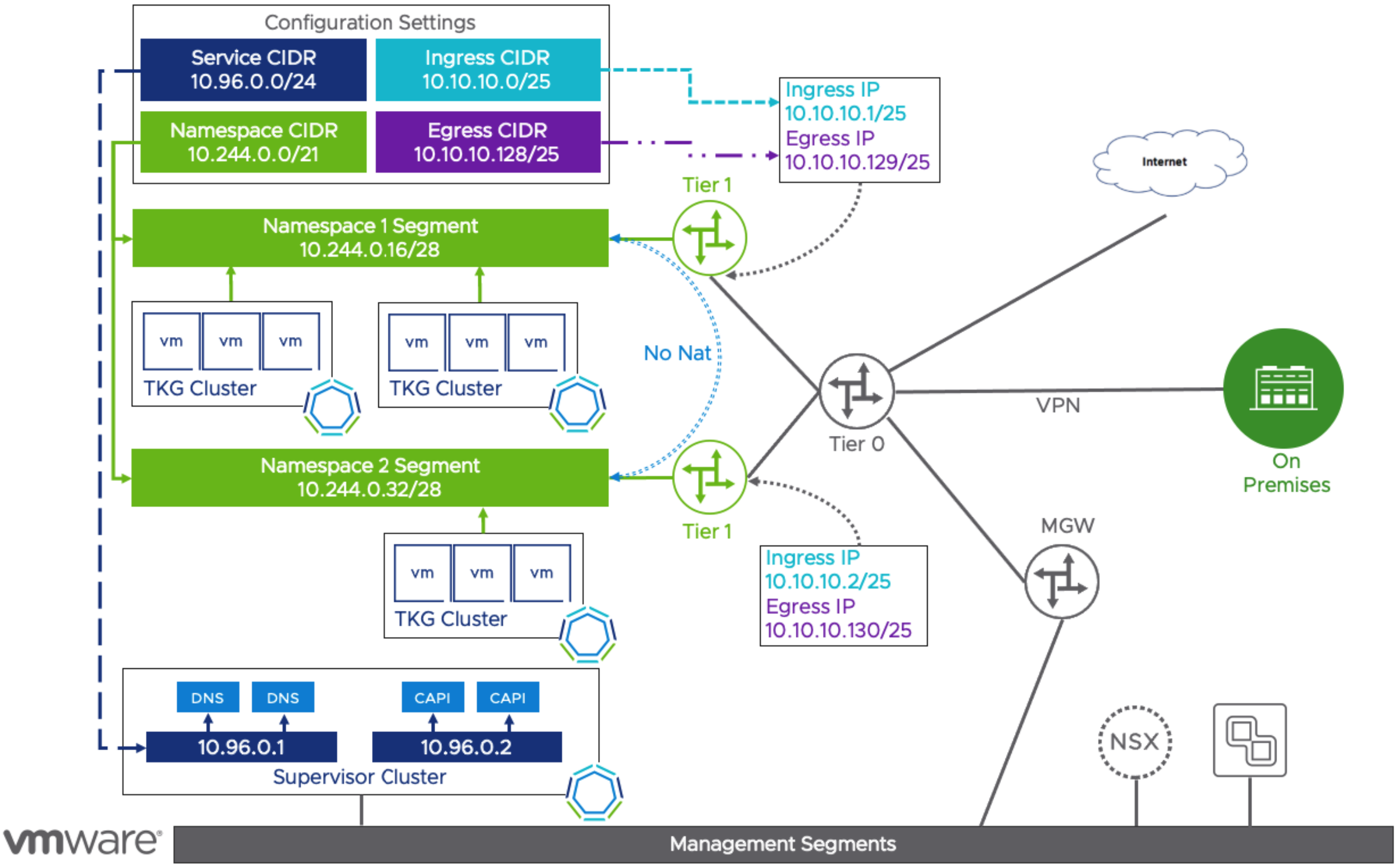
TKG Managed Service Networking

Tanzu Kubernetes Grid Service Networking in VMware Cloud on AWS

The Tanzu Kubernetes Grid Service in VMware Cloud on AWS is a managed service, and as such abstracts a lot of the complexity that you might normally deal with if you build something yourself. But there are some things you might need to know about how the service works. In this post, we'll discuss the networking that makes the Tanzu Kubernetes Grid Service work, and how you can take advantage of the solution when you're building your own applications to be run on Tanzu Kubernetes Clusters.

An Overview

If you've activated the Tanzu Kubernetes Grid Service through the launchpad, you've probably seen the diagram below. We'll walk through each of the CIDR blocks that are needed to enable the TKG Service in the sections below, but as a review, lets cover the networking to the right of the Tier-0 router in the diagram before we proceed.

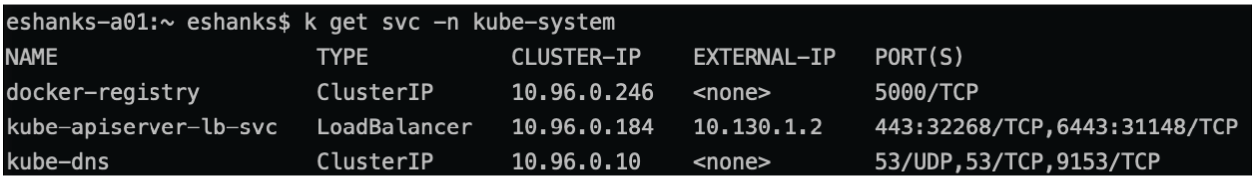

Configuration Settings
Service CIDR 10.96.0.0/24
Ingress CIDR 10.10.10.0/25
Namespace CIDR 10.244.0.0/21
Egress CIDR 10.10.10.128/25
Ingress IP 10.10.10.1/25
Egress IP 10.10.10.129/25
Internet
Tier 1
Namespace 1 Segment 10.244.0.16/28
vm
TKG Cluster
No Nat
Tier 0
VPN
On Premises
Namespace 2 Segment 10.244.0.32/28
Tier 1
Ingress IP 10.10.10.2/25
Egress IP 10.10.10.130/25
MGW
DNS
CAPI
10.96.0.1
10.96.0.2
Supervisor Cluster
NSX
vmware
Management Segments

Now we won't go into immense detail here as there are plenty of other articles that describe all the networking options for getting access to your SDDC, but at a high level lets recognize that the Tier-0 router is directly in the path for traffic from on-premises (through VPN, direct connect, transit gateway, connected AWS VPC, or the Internet). Before Tanzu services was introduced, there were only two other routers that you needed to be aware of. The MGW which is seen in the overview diagram, is the management gateway where some of the managed components are connected, such as vCenter and NSX. The other router which is not pictured is a Customer Gateway which is where your VMC segments are attached. Typically, your VMs would then attach to those networks.

Service CIDR

The Service CIDR is one of the four CIDR pools that needs to be entered to setup the TKG Service. Its purpose is to be a pool of addresses used within the Tanzu Supervisor Cluster for Kubernetes services. For example, there are several pods running in the Tanzu Supervisor Cluster such as the ClusterAPI components, CoreDNS, and etcd. These components are accessed through their Kubernetes service endpoint. As new services in the Supervisor Cluster are needed, the ClusterIP address used for the service comes from this Service CIDR.

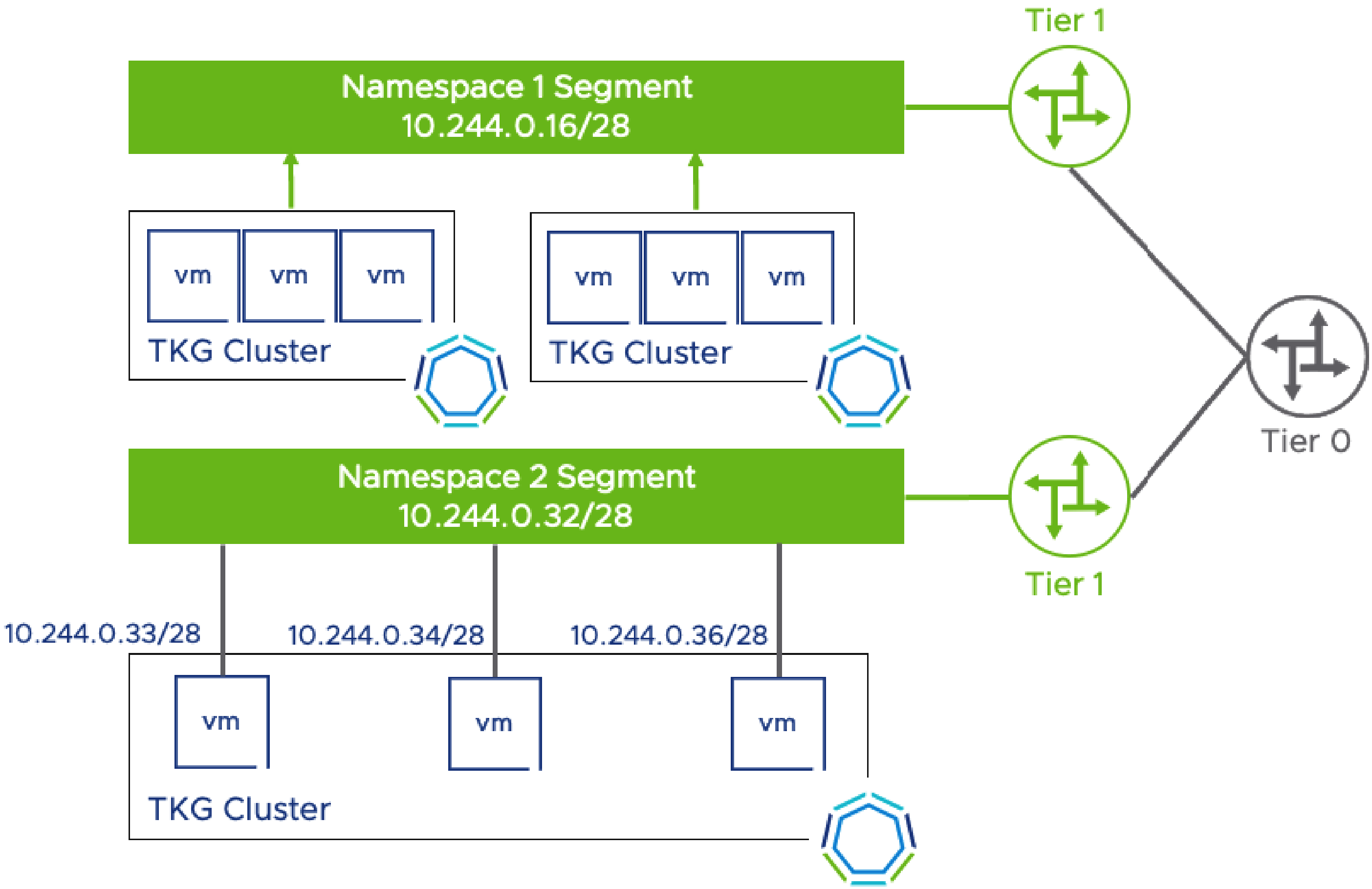
If you were to use the *cloudadmin@vmc.local* administrative role to authenticate to the Supervisor *Cluster*, you could list the services in each namespace to see these ClusterIPs. In the screenshot below, we're listing the services in the kube-system namespace of a Supervisor Cluster. You can see that the cluster-ip addresses are from a 10.96.0.0/24 CIDR which is the default configuration used to setup the TKG Service on VMware Cloud on AWS.



```
eshanks-a01:~ eshanks$ k get svc -n kube-system
NAME                     TYPE           CLUSTER-IP     EXTERNAL-IP    PORT(S)
docker-registry          ClusterIP      10.96.0.246    <none>         5000/TCP
kube-apiserver-lb-svc    LoadBalancer   10.96.0.184    10.130.1.2     443:32268/TCP,6443:31148/TCP
kube-dns                 ClusterIP      10.96.0.10     <none>         53/UDP,53/TCP,9153/TCP
```

Namespace CIDR

A vSphere Namespace is how platform operators are assigned vSphere resources for their Kubernetes clusters. As Cloud Admins create a new namespace and assign resources, new networking components are deployed to segment the namespace from the rest of the environment. Each time a new namespace is created, a new Tier-1 router is created and connected to the Tier-0 router that's already deployed in VMware Cloud. Then a network segment is created and attached to this new Tier-1 router. The Virtual machines deployed to build Tanzu Kubernetes Clusters are connected to these namespace network segments. To successfully use Tanzu Kubernetes Grid, a namespace CIDR should be at least a /23 or better.



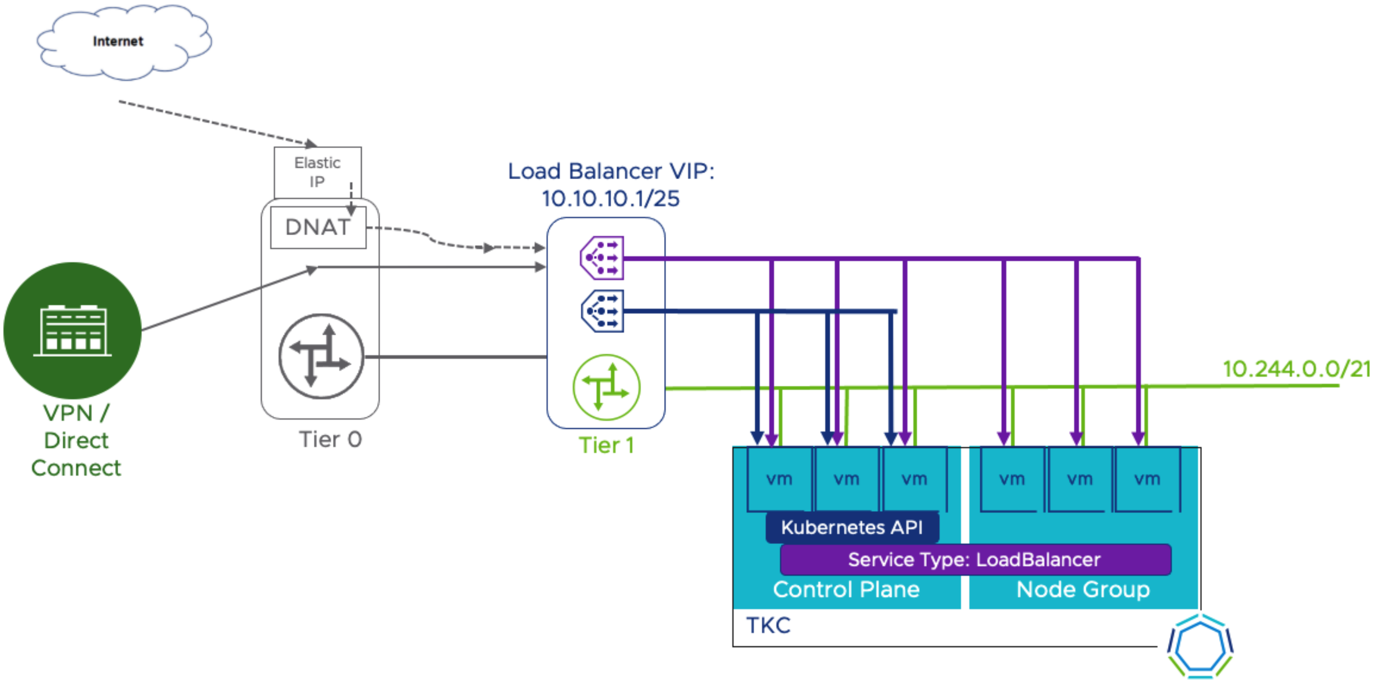
The networking addresses used to setup the new segments are derived from the Namespace Network CIDR entered during the TKG Activation steps.

Ingress CIDR

These Kubernetes clusters aren't very useful if we don't have a solution in place to access them. The Tanzu Kubernetes Grid Service uses NSX to provide access to both the Kubernetes API as well as applications through a Kubernetes service of type LoadBalancer.

The namespace Tier-1 router will create a Load Balancer and a virtual IP (VIP) address to be used to access the Tanzu Kubernetes

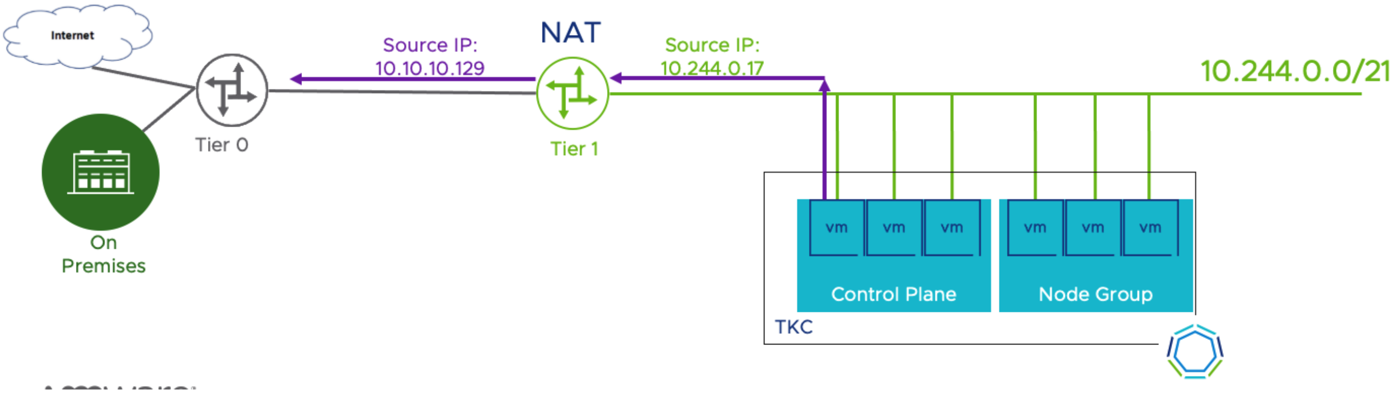
Clusters (TKCs). The VIP address used by the load balancer comes from the ingress CIDR configured during the TKG activation steps. When an admin creates a new Tanzu Kubernetes Cluster a Load Balancer is created in the Namespace Tier-1 router, an address is pulled from the pool of Ingress addresses, and then assigned as the VIP so that you can access the API server through kubectl commands. When a new service of type LoadBalancer is deployed, another VIP is created where the load balancer targets are the VMs within the Tanzu Kubernetes Cluster, so that you may publish apps. This process is explained in more detail in the NSX-T Container Plugin blog post.



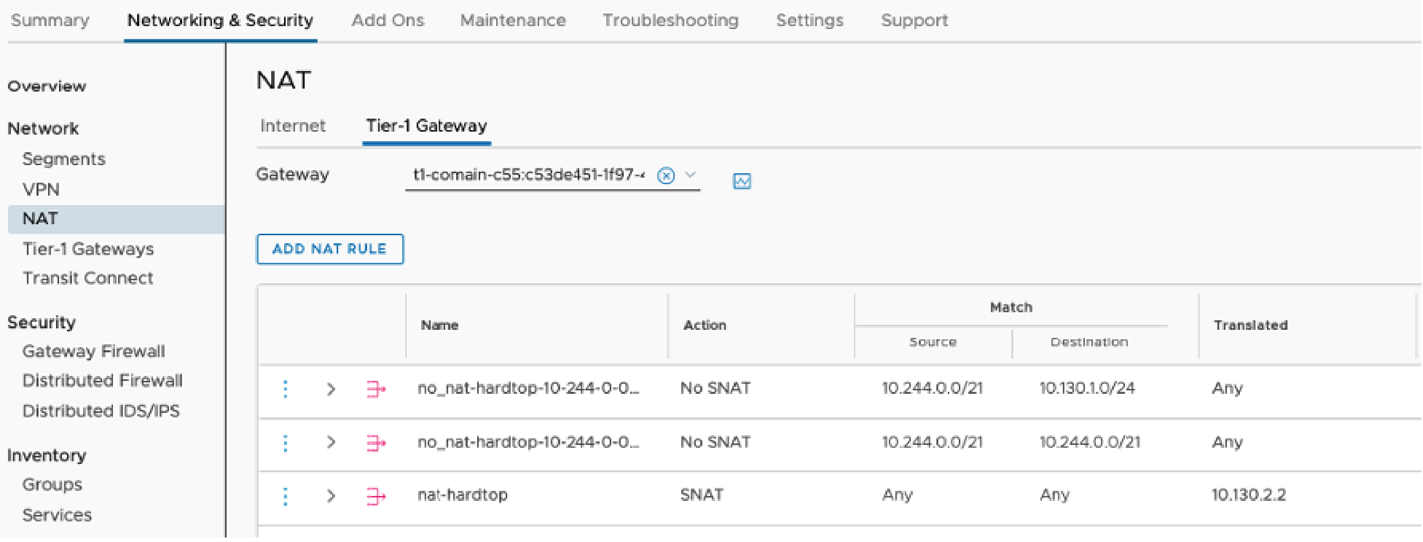
Traffic entering the Tier-0 router through a VPN Tunnel, Direct Connect or Transit Gateway, should be able to access these Load Balancer VIPs provided firewall rules were created to allow access. However, if you are attempting to access the Ingress VIP over the internet connection, you must first assign a public IP address for the Tier-0 router, and a Destination NAT (DNAT) rule to create the appropriate translation to reach the ingress address. You can find more information here.

Egress CIDR

We also need a way to control how traffic leaving the cluster is handled as well. To provide access to resources outside the namespace, a source network address translation (SNAT) is created on the namespace Tier-1 router. As traffic from any of the TKC nodes leaves the cluster, the source IP Address is translated into an egress IP address before leaving the Tier-1 router.



You can see this translation rule for the cluster through the Cloud Services Console in the Tier-1 Gateway NAT settings.



Gateway t1-comain-c55:c53de451-1f97-‹

Name	Action	Match		Translated
		Source	Destination	
no_nat-hardtop-10-244-0-0...	No SNAT	10.244.0.0/21	10.130.1.0/24	Any
no_nat-hardtop-10-244-0-0...	No SNAT	10.244.0.0/21	10.244.0.0/21	Any
nat-hardtop	SNAT	Any	Any	10.130.2.2

Summary and Additional Resources

Just because managing Kubernetes clusters has become easy to do in VMware Cloud on AWS, doesn't mean that there isn't some networking complexity to make it work. Luckily, most of this complexity is taken care of for us, but it's important to have a view into what's happening behind the scenes so that we can make good engineering decisions. Hopefully, this look into how routing, load balancing, and NAT are used with the Tanzu Kubernetes Grid Service will assist you in your Kubernetes deployments on VMware Cloud.

Additional Resources

Using Tanzu Kubernetes Grid Service with VMware Cloud on AWS

VMware Cloud on AWS: A Technical Overview

VMware Cloud on AWS: SDDC Network Architecture

An Introduction to VMware Transit Gateway for VMware Cloud on AWS

NSX-T Container Plugin for the Managed Tanzu Kubernetes Grid Service

Changelog

The following updates were made to this guide.

Date	**Description of Changes**
2021-11-3	Initial publication

About the Author and Contributors

Eric Shanks has spent two decades working with VMware and cloud technologies focusing on hybrid cloud and automation. Eric has obtained some of the industry's highest distinctions, including two VMware Certified Design Expert (VCDX #195) certifications and many others across various solutions, including Kubernetes, Microsoft, Cisco, and Amazon Web Services.

Eric's acted as a community contributor through work as a Chicago VMUG Users Group leader, blogger at theITHollow.com, and Tech Field Day delegate.

- Eric Shanks, Sr. Technical Marketing Architect, Cloud Infrastructure Business Group - VMware