Konzepte und Planung der vSphere IaaS-Steuerungsebene

Update 3
VMware vSphere 8.0
VMware vCenter 8.0
VMware ESXi 8.0

vmware
by Broadcom

Die aktuellste technische Dokumentation finden Sie auf der VMware by Broadcom-Website unter:

https://docs.vmware.com/de/

VMware by Broadcom
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

Inhalt

Konzepte und Planung der vSphere IaaS-Steuerungsebene

Das Handbuch *Konzepte und Planung der vSphere IaaS-Steuerungsebene* enthält Informationen zu den wichtigsten Konzepten und der Architektur von vSphere IaaS control plane (zuvor vSphere with Tanzu) sowie zu den Anforderungen, die Ihre vSphere-Umgebung erfüllen muss, damit Sie vSphere IaaS control plane auf vSphere-Clustern aktivieren und Arbeitslasten in Tanzu Kubernetes Grid-Clustern, vSphere-Pods und mithilfe des VM-Diensts erstellten VMs ausführen können.

Zielgruppe

Diese Informationen richten sich an vSphere-Administratoren und DevOps-Ingenieure, die sich mit den Anforderungen für die Aktivierung von vSphere IaaS control plane auf vSphere sowie mit den wichtigsten Konzepten und der Architektur der Plattform vertraut machen möchten.

Aktualisierte Informationen

Konzepte und Planung der vSphere IaaS-Steuerungsebene wird mit jeder Produktversion oder bei Bedarf aktualisiert.

Diese Tabelle enthält den Updateverlauf für *Konzepte und Planung der vSphere IaaS-Steuerungsebene*.

Revision	Beschreibung
18. APR 2024	Die Lizenzierungsinformationen für die neuen Lösungslizenzen wurden aktualisiert. Weitere Informationen finden Sie unter Lizenzierung für vSphere IaaS control plane.
29. FEB 2024	Es wurden Inhalte für Clouds hinzugefügt. Weitere Informationen finden Sie unter Komponenten für NSX Advanced Load Balancer.
07. FEB 2024	Ein Link in Supervisor-Netzwerk wurde aktualisiert.
13. DEZ 2023	Die NSX-Anforderungen wurden mit einem Hinweis zur Vorbereitung aller ESXi-Hosts aktualisiert, die am vSphere-Cluster als NSX-Transportknoten beteiligt sind. Weitere Informationen finden Sie unter Voraussetzungen für zonalen Supervisor mit NSX und Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX.
29. SEP 2023	Die Anforderungen des Lastausgleichsdiensts für die Bereitstellung von HAProxy wurden aktualisiert. Siehe Voraussetzungen für die zonale Supervisor-Bereitstellung mit HAProxy-Lastausgleichsdienst und Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst
21. SEP 2023	Es wurden Inhalte für Supervisor-Netzwerk mit NSX und NSX Advanced Load Balancer hinzugefügt. Weitere Informationen finden Sie unter vSphere IaaS control plane-Benutzerrollen und -Workflows und Supervisor-Netzwerk.
3. AUG 2023	Nebenversionen.
30. JUN 2023	■ Es wurde eine Aussage über die Verteilung von vSphere-Zonen über physische Sites hinweg hinzugefügt: Supervisor-Bereitstellungen mit Zonen und Clustern
9. JUN 2023	■ Die folgende Anweisung wurde vSphere IaaS control planeSicherheit hinzugefügt: Dasselbe Verschlüsselungsmodell gilt für die Daten in der Datenbank (etcd), die auf der Steuerungsebene für jeden Tanzu Kubernetes Grid-Cluster installiert sind. ■ Es wurde eine Aussage hinzugefügt, dass Storage vMotion von dauerhaften Volumes nicht unterstützt wird. Weitere Informationen finden Sie unter Wie ist Supervisor in vSphere-Speicher integriert? und Speicher für Tanzu Kubernetes Grid-Cluster. ■ Es wurde eine Empfehlung zum Trennen der Verwaltungs- und Arbeitslastdomänen als Best Practice in Kapitel 5 Voraussetzungen für die zonale Supervisor-Bereitstellung und Kapitel 6 Voraussetzungen für die Cluster-Supervisor-Bereitstellung hinzugefügt.
2. JUN 2023	■ Ein Link zum NSX Reference Design Guide wurde hinzugefügt. Weitere Informationen finden Sie unter Voraussetzungen für zonalen Supervisor mit NSX und Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX. ■ Kleinere Updates.

Revision	Beschreibung
30. MAI 2023	■ Es wurde eine Anforderung zur Unterstützung der GENEVE-Kapselung für NSX-Bereitstellungen hinzugefügt. Weitere Informationen finden Sie unter Anforderungen an den Zonen-Supervisor mit NSX and Anforderung an die Cluster-Supervisor-Bereitstellung mit NSX. ■ Kleinere Updates.
12. MAI 2023	Es wurde ein Hinweis hinzugefügt, dass Sie, wenn Sie für Ihre vSphere IaaS control plane-Umgebung ein Upgrade von einer vSphere-Version vor 8.0 durchgeführt haben und vSphere-Zones verwenden möchten, eine neue Supervisor mit drei Zonen erstellen müssen. Weitere Informationen finden Sie unter Kapitel 5 Voraussetzungen für die zonale Supervisor-Bereitstellung.
9. MAI 2023	■ Ein separates Thema über Was ist ein vSphere-Namespace? wurde hinzugefügt. ■ Der Inhalt zu vSphere IaaS control planeldentitäts- und Zugriffsverwaltung wurde aktualisiert. ■ Es wurde ein Hinweis hinzugefügt, dass vSphere-Pods nur mit dem NSX-Netzwerk-Stack unterstützt werden. Weitere Informationen finden Sie unter Was ist ein vSphere Pod?.
01. MAI 2023	Nebenversionen.
18. APR 2023	Allgemeine Updates für die Version vSphere 8 Update 1.

vSphere IaaS control plane-Konzepte

1

Mit der vSphere IaaS control plane können Sie vSphere-Cluster in eine Plattform für die Ausführung von Kubernetes-Arbeitslasten in dedizierten Ressourcenpools in vSphere umwandeln. Wenn vSphere IaaS control plane auf vSphere-Clustern aktiviert ist, wird direkt innerhalb der Hypervisor-Schicht eine Kubernetes-Steuerungsebene erstellt. Sie können dann Kubernetes-Container ausführen, indem Sie vSphere-Pods bereitstellen. Alternativ können Sie über VMware Tanzu™ Kubernetes Grid™ auch Upstream-Kubernetes-Cluster erstellen und Ihre Anwendungen in diesen Clustern ausführen.

Lesen Sie als Nächstes die folgenden Themen:

- Was ist vSphere IaaS control plane?
- Was ist ein vSphere-Namespace?
- vSphere IaaS control plane-Benutzerrollen und -Workflows
- Wie ändert vSphere IaaS control plane die vSphere-Umgebung?
- Lizenzierung für vSphere IaaS control plane
- vSphere IaaS control planeldentitäts- und Zugriffsverwaltung
- vSphere IaaS control planeSicherheit

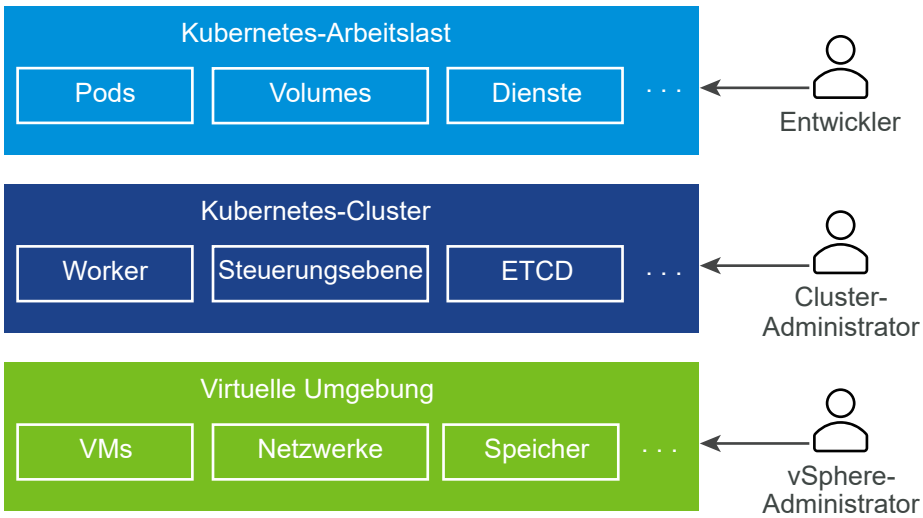
Was ist vSphere IaaS control plane?

Sie können vSphere IaaS control plane verwenden, um vSphere in eine Plattform zur nativen Ausführung von Kubernetes-Arbeitslasten auf der Hypervisor-Schicht umzuwandeln. Bei Aktivierung auf vSphere-Clustern bietet vSphere IaaS control plane die Möglichkeit, Kubernetes-Arbeitslasten direkt auf ESXi-Hosts auszuführen und Upstream-Kubernetes-Cluster in dedizierten Namespaces zu erstellen, die als vSphere-Namespace bezeichnet werden.

Herausforderungen des heutigen Anwendungsstapels

Die heutigen verteilten Systeme bestehen aus mehreren Mikrodiensten, in denen in der Regel eine Vielzahl von Kubernetes-Pods und VMs ausgeführt werden. Ein typischer Stack, der nicht auf vSphere IaaS control plane basiert, besteht aus einer zugrunde liegenden virtuellen Umgebung mit einer Kubernetes-Infrastruktur, die innerhalb von VMs bereitgestellt wird, bzw. Kubernetes-Pods, die auch auf diesen VMs ausgeführt werden. Jeder Teil des Stacks wird von einer von drei separaten Rollen ausgeführt: Anwendungsentwickler, Kubernetes-Cluster-Administratoren und vSphere-Administratoren.

Abbildung 1-1. Heutiger Anwendungsstapel



Die verschiedenen Rollen verfügen über keinen Einblick oder Kontrolle über die Umgebungen der jeweils anderen.

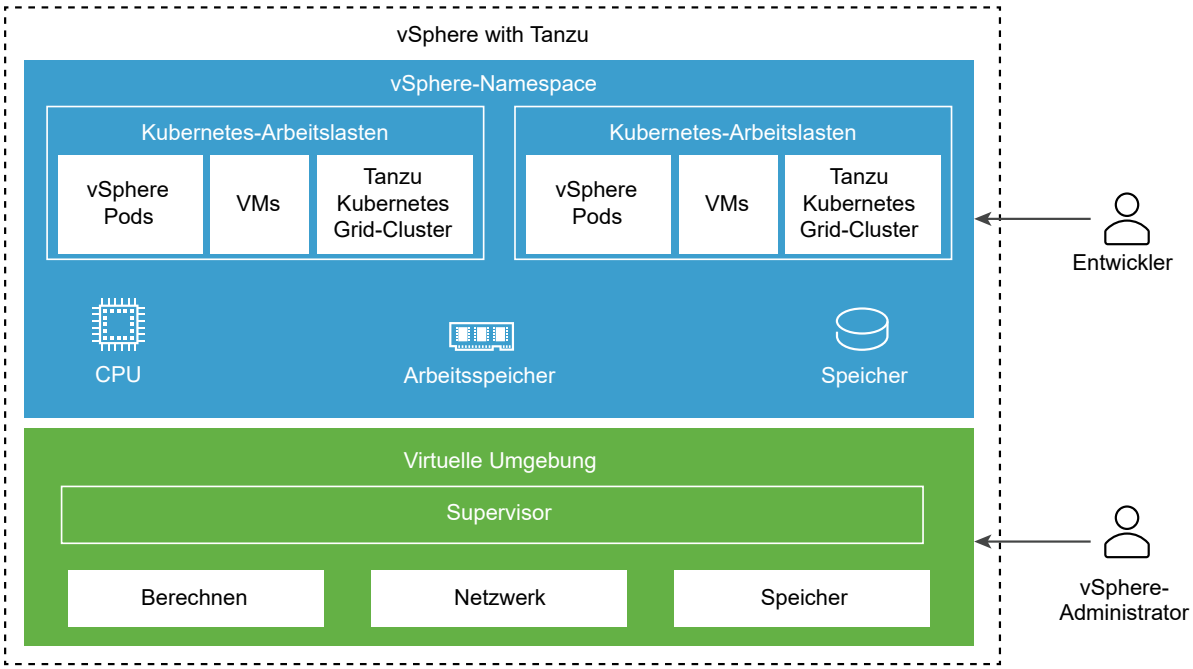
- Als Anwendungsentwickler können Sie Kubernetes-Pods ausführen und auf Kubernetes basierte Anwendungen bereitstellen und verwalten. Sie haben keinen Überblick über den gesamten Stack, auf dem Hunderte von Anwendungen ausgeführt werden.
- Als DevOps-Ingenieur oder Cluster-Administrator haben Sie nur die Kontrolle über die Kubernetes-Infrastruktur, ohne die Tools zum Verwalten oder Überwachen der virtuellen Umgebung und zur Behebung von ressourcenbezogenen und anderen Problemen.
- Als vSphere-Administrator haben Sie die volle Kontrolle über die zugrunde liegende virtuelle Umgebung, aber keinen Einblick in die Kubernetes-Infrastruktur, die Platzierung der verschiedenen Kubernetes-Objekte in der virtuellen Umgebung und deren Ressourcenverbrauch.

Die Ausführung von Vorgängen im gesamten Stack kann schwierig sein, da hierzu Kommunikation zwischen allen drei Rollen erforderlich ist. Auch die fehlende Integration zwischen den verschiedenen Schichten des Stacks kann Herausforderungen mit sich bringen. So hat der Kubernetes Scheduler beispielsweise keinen Einblick in die vCenter Server-Bestandsliste und kann Pods nicht intelligent platzieren.

Wie hilft vSphere IaaS control plane?

vSphere IaaS control plane erstellt eine Kubernetes-Steuerungsebene direkt auf der Hypervisor-Schicht. Als vSphere-Administrator aktivieren Sie vorhandene vSphere-Cluster für die vSphere IaaS control plane und erstellen so eine Kubernetes-Schicht innerhalb der ESXi-Hosts, die Teil der Cluster sind. Die für die vSphere IaaS control plane aktivierten vSphere-Cluster werden als Supervisoren bezeichnet.

Abbildung 1-2. vSphere IaaS control plane



Wenn eine Kubernetes-Steuerungsebene auf der Hypervisor-Schicht vorhanden ist, werden die folgenden Funktionen in vSphere ermöglicht:

- Als vSphere-Administrator können Sie Namespaces für Supervisor, die als vSphere-Namespaces bezeichnet werden, erstellen und mit Arbeitsspeicher und Speicher sowie CPU im angegebenen Umfang konfigurieren. Sie stellen diese vSphere-Namespaces den DevOps-Ingenieuren zur Verfügung.
- Als DevOps-Ingenieur können Sie Kubernetes-Arbeitslasten auf derselben Plattform mit gemeinsam genutzten Ressourcenpools innerhalb eines vSphere-Namespace ausführen. Sie können mehrere mit Tanzu Kubernetes Griderstellte Upstream-Kubernetes-Cluster bereitstellen und verwalten. Kubernetes-Container können Sie auch direkt auf dem Supervisor innerhalb eines speziellen VM-Typs mit der Bezeichnung vSphere Pod bereitstellen. Sie können auch reguläre VMs bereitstellen.
- Als vSphere-Administrator können Sie vSphere-Pods, VMs und Tanzu Kubernetes Grid-Cluster mit dem vSphere Client verwalten und überwachen.

- Als vSphere-Administrator haben Sie vollständigen Einblick in vSphere-Pods, VMs und Tanzu Kubernetes Grid-Cluster, die in unterschiedlichen Namespaces ausgeführt werden, deren Platzierung in der Umgebung und deren Ressourcenverbrauch.

Die Verwendung von Kubernetes auf der Hypervisor-Schicht erleichtert auch die Zusammenarbeit zwischen vSphere-Administratoren und DevOps-Teams, da beide Rollen mit denselben Objekten arbeiten.

Was ist eine Arbeitslast?

In vSphere IaaS control plane sind Arbeitslasten Anwendungen, die auf eine der folgenden Arten bereitgestellt werden:

- Anwendungen, die aus Containern bestehen, die innerhalb von vSphere-Pods ausgeführt werden.
- Über den VM-Dienst bereitgestellte Arbeitslasten.
- Tanzu Kubernetes Grid-Cluster, die mithilfe von Tanzu Kubernetes Grid bereitgestellt werden.
- Anwendungen, die innerhalb der Tanzu Kubernetes Grid-Cluster ausgeführt werden.

Was sind vSphere-Zonen?

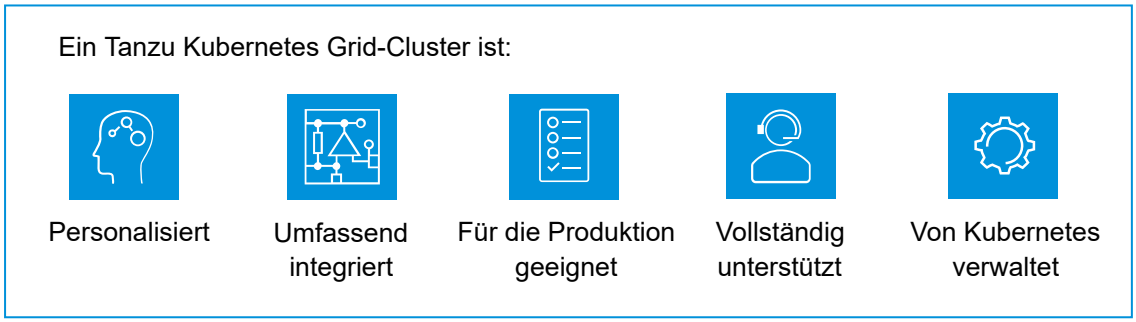
vSphere-Zonen bieten Hochverfügbarkeit gegen Ausfälle auf Clusterebene bei Arbeitslasten, die auf vSphere IaaS control plane bereitgestellt werden. Als vSphere-Administrator erstellen Sie vSphere-Zonen im vSphere Client und ordnen dann den Zonen vSphere-Cluster zu. Sie verwenden die Zonen zum Bereitstellen von Supervisoren in Ihrer vSphere IaaS control plane-Umgebung.

Für Hochverfügbarkeit auf Clusterebene können Sie einen Supervisor in drei vSphere-Zonen bereitstellen. Alternativ haben Sie die Möglichkeit, einen Supervisor auf einem einzelnen vSphere-Cluster bereitzustellen, der automatisch eine vSphere-Zone erstellt und diese dem Cluster zuweist, oder Sie können einen Cluster verwenden, der bereits einer Zone zugeordnet ist. Weitere Informationen finden Sie unter Supervisor – Architektur und Supervisor-Bereitstellungen mit Zonen und Clustern.

Was ist ein Tanzu Kubernetes Grid-Cluster?

Ein Tanzu Kubernetes Grid-Cluster ist eine vollständige Distribution von Kubernetes, die von VMware als Paket erstellt, signiert und unterstützt wird. Sie können Upstream-Tanzu Kubernetes Grid-Cluster mithilfe von Tanzu Kubernetes Grid auf Supervisorenn bereitstellen und betreiben.

Ein Tanzu Kubernetes Grid-Cluster, der von Tanzu Kubernetes Grid bereitgestellt wird, weist die folgenden Merkmale auf:



- Personalisierte Installation von Kubernetes. Tanzu Kubernetes Grid bietet durchdachte Standardeinstellungen, die für vSphere zur Bereitstellung von Tanzu Kubernetes Grid-Clustern optimiert sind. Mithilfe von Tanzu Kubernetes Grid können Sie die Zeit und den Aufwand reduzieren, die Sie normalerweise für die Bereitstellung und Ausführung eines unternehmensgerechten Kubernetes-Clusters investieren müssen.
- Integriert in die vSphere-Infrastruktur. Ein Tanzu Kubernetes Grid-Cluster ist in den vSphere SDDC-Stack integriert, einschließlich Speicher, Netzwerk und Authentifizierung. Darüber hinaus wird ein Tanzu Kubernetes Grid-Cluster in einem Supervisor erstellt, der vSphere-Clustern zugeordnet ist. Aufgrund der engen Integration bietet der Tanzu Kubernetes Grid-Cluster bei der Ausführung eine einheitliche Produkterfahrung.
- Für die Produktion geeignet. Tanzu Kubernetes Grid stellt für die Produktion geeignete Tanzu Kubernetes Grid-Cluster bereit. Sie können Produktionsarbeitslasten ohne zusätzliche Konfiguration ausführen. Darüber hinaus können Sie Verfügbarkeit sicherstellen, parallele Upgrades der Kubernetes-Software zulassen und verschiedene Versionen von Kubernetes in separaten Clustern ausführen.
- Hochverfügbarkeit für Kubernetes-Arbeitslasten. Tanzu Kubernetes Grid-Cluster, die in einem Supervisor mit drei vSphere-Zonen bereitgestellt werden, sind auf vSphere-Clusterebene vor Ausfällen geschützt. Die Arbeitslast- und Steuerungsebenenknoten von Tanzu Kubernetes Grid-Clustern sind über alle drei vSphere-Zonen verteilt, sodass die darin ausgeführten Kubernetes-Arbeitslasten hochverfügbar sind. Tanzu Kubernetes Grid-Cluster, die in einem Supervisor mit einer Zone ausgeführt werden, sind durch vSphere HA auf ESXi-Hostebene vor Ausfällen geschützt.
- Vollständig unterstützt von VMware. Tanzu Kubernetes Grid-Cluster verwenden das Open Source- und Linux-basierte Betriebssystem von VMware, werden auf einer vSphere-Infrastruktur bereitgestellt und werden auf ESXi-Hosts ausgeführt. Wenn Probleme mit einer beliebigen Schicht des Stacks – von Hypervisor bis zum Kubernetes-Cluster – auftreten, ist VMware der einzige Anbieter, den Sie kontaktieren müssen.
- Von Kubernetes verwaltet. Tanzu Kubernetes Grid-Cluster werden zusätzlich zum Supervisor erstellt, der selbst ein Kubernetes-Cluster ist. Ein Tanzu Kubernetes Grid-Cluster wird mithilfe einer benutzerdefinierten Ressource im vSphere-Namespace definiert. Sie stellen Tanzu Kubernetes Grid-Cluster mithilfe von bekannten kubectl-Befehlen und der Tanzu-

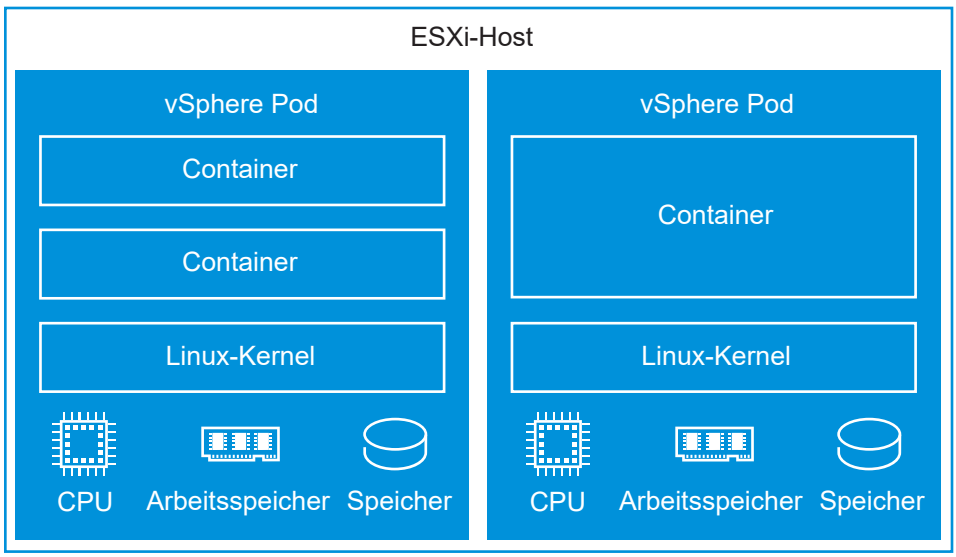
CLI als Self-Service-Komponenten bereit. In der gesamten Toolkette besteht Konsistenz, und zwar unabhängig davon, ob Sie einen Cluster oder Arbeitslasten bereitstellen. Sie verwenden dieselben Befehle, die schon bekannte YAML-Auszeichnungssprache und gängige Workflows.

Weitere Informationen finden Sie unter Kapitel 3 Tanzu Kubernetes Grid-Architektur und -Komponenten und *Verwenden des TKG-Dienstes mit der vSphere IaaS-Steuerungsebene*.

Was ist ein vSphere Pod?

vSphere IaaS control plane führt ein Konstrukt mit dem Namen vSphere Pod ein, das einem Kubernetes-Pod entspricht. Eine vSphere Pod ist eine VM mit einem kleinen Footprint, die einen oder mehrere Linux-Container ausführt. Jede vSphere Pod wird genau für die Arbeitslast angepasst, die sie unterstützt, und weist explizite Ressourcenreservierungen für diese Arbeitslast auf. Sie weist die genaue Menge an Speicherplatz, Arbeitsspeicher und CPU-Ressourcen zu, die für die Ausführung der Arbeitslast erforderlich ist. vSphere-Pods werden nur mit Supervisoren unterstützt, die mit NSX als Netzwerk-Stack konfiguriert sind.

Abbildung 1-3. vSphere-Pods



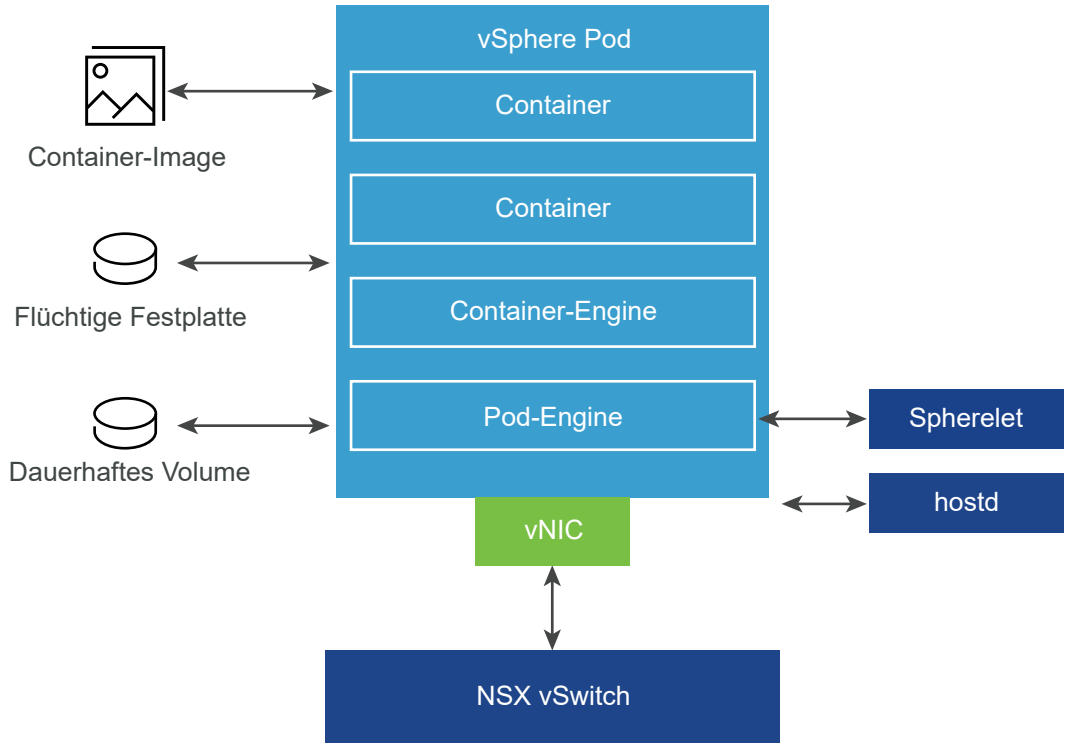
vSphere-Pods sind Objekte in vCenter Server. Sie ermöglichen die folgenden Funktionen für Arbeitslasten:

- Starke Isolierung. Ein vSphere Pod wird in gleicher Weise wie eine virtuelle Maschine isoliert. Jeder vSphere Pod hat seinen eigenen Linux-Kernel, der auf dem im Photon OS verwendeten Kernel basiert. Statt dass viele Container einen Kernel teilen, wie in einer Bare-Metal-Konfiguration, hat in einem vSphere Pod jeder Container einen eindeutigen Linux-Kernel.
- Ressourcenverwaltung. vSphere DRS übernimmt die Platzierung von vSphere-Pods im Supervisor.

- Hochleistung. vSphere-Pods erhalten die gleiche Ressourcenisolierung wie VMs und eliminieren Störungen durch benachbarte Elemente, während die schnelle Startzeit und der geringe Overhead von Containern beibehalten werden.
- Diagnose. Als vSphere-Administrator können Sie alle Überwachungs- und Selbstprüfungs-Tools verwenden, die in vSphere für Arbeitslasten verfügbar sind.

vSphere-Pods sind OCI-kompatibel (Open Container Initiative) und können Container von jedem Betriebssystem aus ausführen, solange diese Container ebenfalls OCI-kompatibel sind.

Abbildung 1-4. vSphere Pod Netzwerk und Speicher



vSphere-Pods verwenden drei Speichertypen, je nachdem, welche Objekte gespeichert sind: flüchtige VMDKs, dauerhafte Volume-VMDKs und Container-Image-VMDKs. Als vSphere-Administrator konfigurieren Sie Speicherrichtlinien für die Platzierung von Container-Image-Cache und flüchtigen VMDKs auf der Supervisor-Ebene. Auf vSphere-Namespace-Ebene konfigurieren Sie Speicherrichtlinien für die Platzierung persistenter Volumes. Weitere Informationen zu den Speicheranforderungen und Konzepten für vSphere IaaS control plane finden Sie unter Persistenter Speicher für Arbeitslasten.

Für Netzwerke verwenden vSphere-Pods und die VMs der Tanzu Kubernetes Grid-Cluster die von NSX bereitgestellte Topologie. Weitere Informationen finden Sie unter Supervisor-Netzwerk.

Spherelet ist ein zusätzlicher Prozess, der auf jedem Host erstellt wird. Es handelt sich um ein Kubelet, das nativ auf ESXi portiert wird und dem ESXi-Host ermöglicht, Teil des Kubernetes-Clusters zu werden.

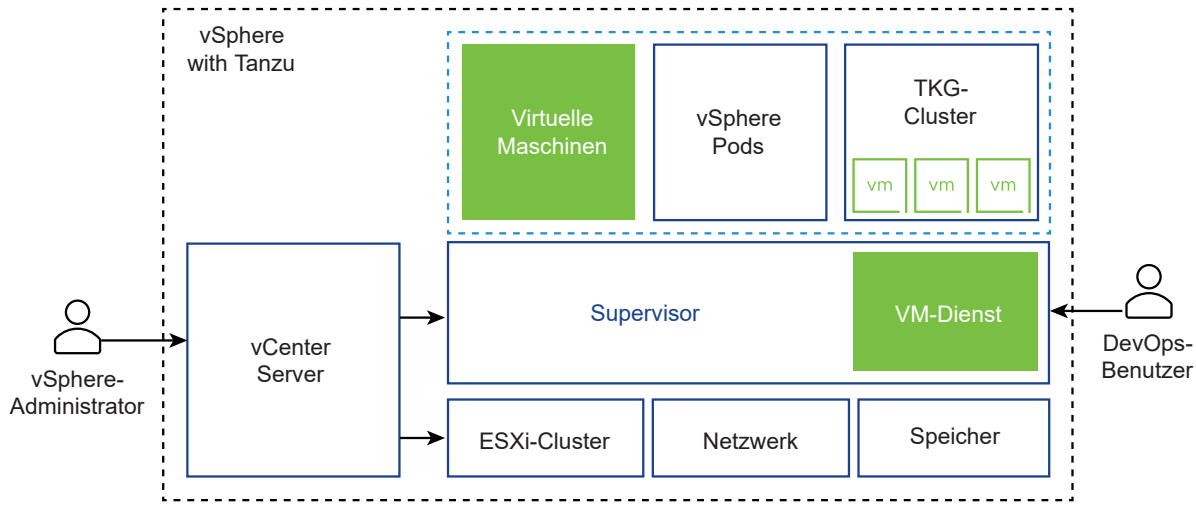
Informationen zur Verwendung von vSphere-Pods auf Supervisoren finden Sie unter Bereitstellen von Arbeitslasten in vSphere Pods in der Dokumentation *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

Verwenden von virtuellen Maschinen in vSphere IaaS control plane

vSphere IaaS control plane bietet eine VM-Service-Funktionalität, die es DevOps-Ingenieuren ermöglicht, neben Containern auch VMs in einer gemeinsamen, freigegebenen Kubernetes-Umgebung bereitzustellen und auszuführen. Sowohl Container als auch VMs nutzen dieselben vSphere-Namespace-Ressourcen und können über eine einzelne vSphere IaaS control plane-Schnittstelle verwaltet werden.

Der VM-Dienst adressiert die Bedürfnisse von DevOps-Teams, die Kubernetes nutzen, aber bestehende VM-basierte Arbeitslasten haben, die nicht einfach containerisiert werden können. Es hilft Anwendern auch, den Aufwand für die Verwaltung einer Nicht-Kubernetes-Plattform neben einer Container-Plattform zu reduzieren. Wenn Container und VMs auf einer Kubernetes-Plattform ausgeführt werden, können DevOps-Teams ihren Arbeitslastbedarf auf nur einer Plattform konsolidieren.

Hinweis Zusätzlich zu eigenständigen VMs verwaltet der VM-Dienst die VMs, die die Tanzu Kubernetes Grid-Cluster bilden. Informationen über Cluster finden Sie in der Dokumentation zum Thema *Verwenden des TKG-Dienstes mit der vSphere IaaS-Steuerungsebene*.



Jede VM, die über den VM-Dienst bereitgestellt wird, funktioniert als vollständige Maschine, auf der alle Komponenten, einschließlich eines eigenen Betriebssystems, in der vSphere IaaS control plane-Infrastruktur ausgeführt werden. Die VM hat Zugriff auf Netzwerke und Speicher, die ein Supervisor bereitstellt, und wird mithilfe des standardmäßigen Kubernetes-Befehls `kubectl` verwaltet. Die VM wird als vollständig isoliertes System ausgeführt, das immun gegen Störungen durch andere VMs oder Arbeitslasten in der Kubernetes-Umgebung ist.

Wann werden virtuelle Maschinen auf einer Kubernetes-Plattform verwendet?

Im Allgemeinen hängt die Entscheidung zum Ausführen von Arbeitslasten in einem Container oder in einer VM von Ihren geschäftlichen Anforderungen und Zielen ab. Zu den Gründen für den Einsatz von VMs zählen unter anderem die folgenden:

- Ihre Anwendungen können nicht containerisiert werden.
- Anwendungen sind für einen benutzerdefinierten Kernel oder ein benutzerdefiniertes Betriebssystem konzipiert.
- Anwendungen sind besser für die Ausführung in einer VM geeignet.
- Sie wünschen sich ein konsistentes Kubernetes-Erlebnis und möchten Overhead vermeiden. Anstatt separate Infrastruktursätze für Ihre Nicht-Kubernetes- und Container-Plattformen zu betreiben, können Sie diese Stacks konsolidieren und mit einem vertrauten `kubectl`-Befehl verwalten.

Informationen zum Bereitstellen und Verwalten eigenständiger virtueller Maschinen auf einem Supervisor finden Sie in der Dokumentation *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene* im Thema über das Bereitstellen und Verwalten von virtuellen Maschinen.

Bereitstellung von Supervisor-Dienste in vSphere IaaS control plane

Supervisor-Dienste sind vSphere-zertifizierte Kubernetes-Operatoren, die Infrastructure-as-a-Service-Komponenten und eng integrierte Dienste von unabhängigen Softwareanbietern für Entwickler bereitstellen. Sie können Supervisor-Dienste in der vSphere IaaS control plane-Umgebung installieren und verwalten, um sie für die Verwendung mit Kubernetes-Arbeitslasten verfügbar zu machen. Wenn Supervisor-Dienste in Supervisoren installiert werden, können DevOps-Ingenieure über die Dienst-APIs Instanzen in Supervisoren in ihren Benutzer-Namespaces erstellen. Diese Instanzen können dann in vSphere-Pods und Tanzu Kubernetes Grid-Clustern verwendet werden.

Weitere Informationen zu den unterstützten Supervisor-Dienste und zur Vorgehensweise beim Herunterladen ihrer YAML-Dienstdateien finden Sie unter http://vmware.com/go/supervisor-service.

Informationen zur Verwendung der Supervisor-Dienste finden Sie in der Dokumentation zur Verwaltung von Supervisor-Diensten*Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

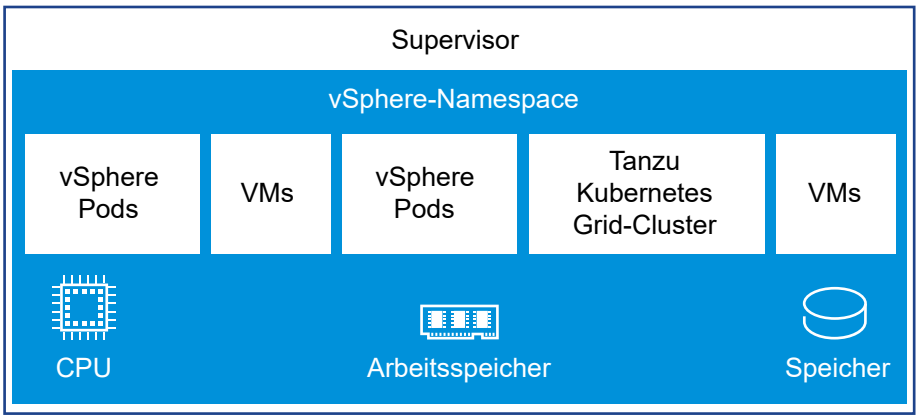
Was ist ein vSphere-Namespace?

Ein vSphere-Namespace legt die Ressourcengrenzen fest, in denen vSphere-Pods, VMs und Tanzu Kubernetes Grid-Cluster ausgeführt werden können. Als vSphere-Administrator erstellen und konfigurieren Sie vSphere-Namespaces über den vSphere-Client.

Bei der anfänglichen Erstellung verfügt ein vSphere-Namespace über unbegrenzte Ressourcen innerhalb des Supervisor. Als vSphere-Administrator können Sie Grenzwerte für CPU, Arbeitsspeicher und Speicher sowie die Anzahl der Kubernetes-Objekte festlegen, die innerhalb des vSphere-Namespace ausgeführt werden können. Speichereinschränkungen werden in Kubernetes als Speicherkontingente dargestellt. In vSphere wird für jeden vSphere-Namespace auf dem Supervisor ein Ressourcenpool erstellt.

In einem auf vSphere-Zonen aktivierten Supervisor wird ein Namespace-Ressourcenpool auf jedem vSphere-Cluster erstellt, der einer Zone zugeordnet ist. Die vSphere-Namespace verteilt sich auf alle drei vSphere Cluster, die Teil des vSphere-Zonen sind. Die für einen vSphere-Namespace auf einem Supervisor mit drei Zonen verwendeten Ressourcen werden zu gleichen Teilen aus allen drei zugrunde liegenden vSphere-Clustern bezogen. Wenn Sie beispielsweise 300 MHz an CPU zuweisen, kommen von jedem vSphere-Cluster 100 MHz.

Abbildung 1-5. vSphere-Namespace



Um dem DevOps-Ingenieur Zugriff auf Namespaces zu gewähren, weisen Sie als vSphere-Administrator den verfügbaren Benutzern oder Benutzergruppen Berechtigungen innerhalb einer Identitätsquelle zu, die mit vCenter Single Sign-On verknüpft ist. Alternativ können Sie Anmeldedaten von einem OIDC-Anbieter verwenden, der beim Supervisor registriert ist. Weitere Informationen finden Sie unter vSphere IaaS control planeIdentitäts- und Zugriffsverwaltung.

Nachdem ein Namespace mit Ressourcen- und Objektgrenzwerten sowie mit Berechtigungen und Speicherrichtlinien erstellt und konfiguriert wurde, können Sie als DevOps-Ingenieur auf den Namespace zugreifen, um Arbeitslasten wie Tanzu Kubernetes Grid-Cluster, vSphere-Pods und VMs, die über den VM-Dienst erstellt wurden, auszuführen.

Unterschiede zwischen einem vSphere-Namespace und einem Kubernetes-Namespace

Obwohl ein vSphere-Namespace grundlegend dieselbe Funktion wie ein Kubernetes-Namespace erfüllt, ist ein vSphere-Namespace spezifisch für die vSphere IaaS control plane. Sie sollten einen vSphere-Namespace nicht mit einem Kubernetes-Namespace verwechseln.

Ein vSphere-Namespace wird als Erweiterung eines vSphere Ressourcenpools implementiert. Seine Funktion besteht darin, Ressourcen für Arbeitslasten bereitzustellen, die im Supervisor ausgeführt werden. Ein vSphere-Namespace verfügt über eine direkte Zuordnung zu einem Kubernetes-Namespace, über den Objekt- und Speicherkontingente für Arbeitslasten erzwungen werden.

Ein weiterer Unterschied zu einem regulären Kubernetes-Namespace besteht darin, dass der vSphere-Administrator den Benutzerzugriff auf vSphere-Namespaces verwaltet, wie oben erwähnt. Der vSphere-Administrator kann auch VM-Klassen und Inhaltsbibliotheken zuordnen, die VM-Vorlagen enthalten, die DevOps-Ingenieure zum Self-Service von VMs verwenden können. Weitere Informationen finden Sie im Thema über das Bereitstellen und Verwalten von virtuellen Maschinen in *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

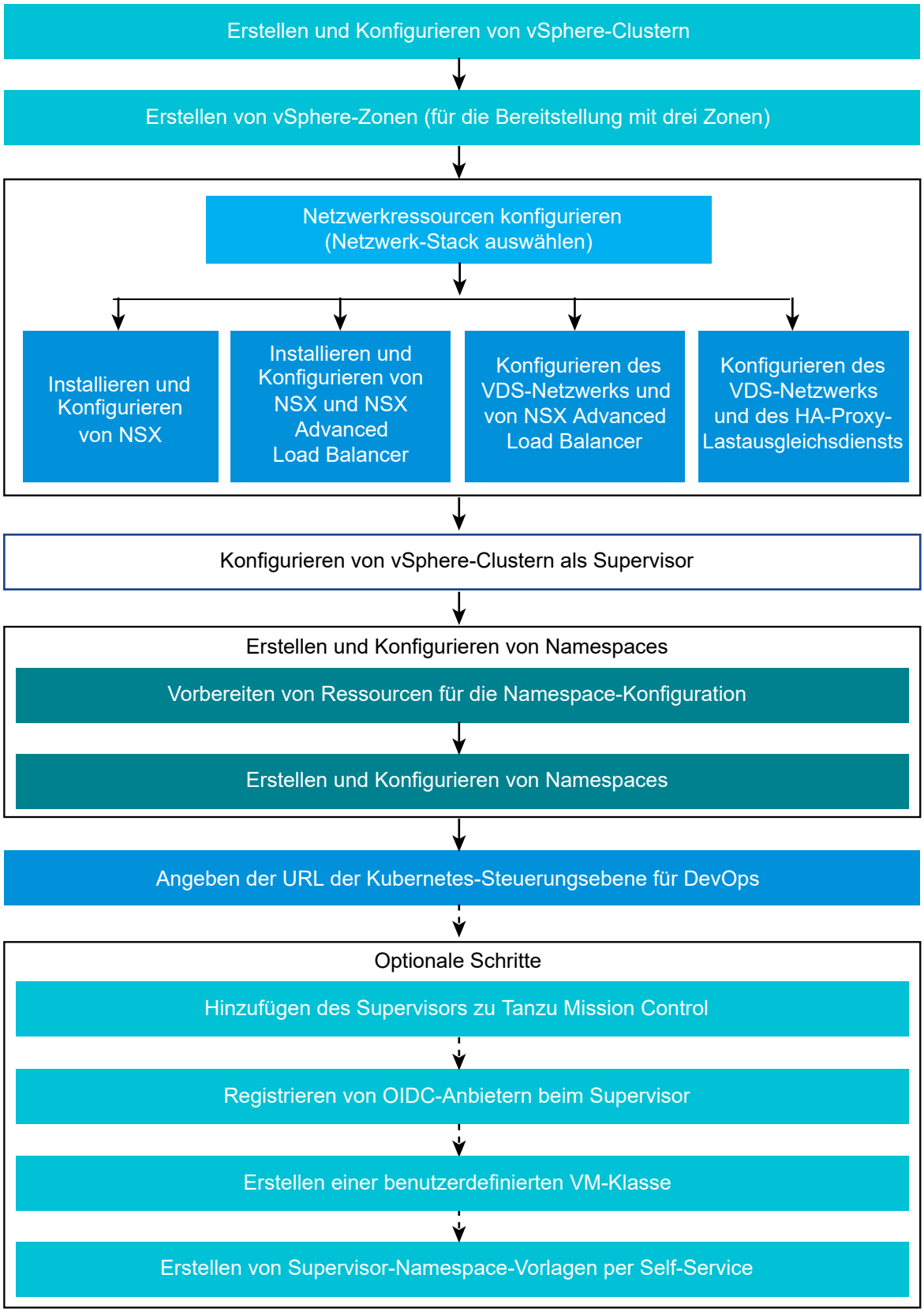
vSphere IaaS control plane-Benutzerrollen und -Workflows

vSphere IaaS control plane bietet zwei Rollen, den vSphere-Administrator und den DevOps-Ingenieur. Der DevOps-Ingenieur beinhaltet die Rollen DevOps, Anwendungsentwickler und Kubernetes-Administrator. Beide Rollen interagieren mit der Plattform über verschiedene Schnittstellen und können Benutzer oder Benutzergruppen aufweisen, die für sie in vCenter Server mit zugehörigen Berechtigungen definiert sind. Die Workflows für die Rollen „vSphere-Administrator“ und „DevOps-Ingenieur“ sind verschieden und werden durch den spezifischen Fachbereich bestimmt, die diese Rollen erfordern.

Benutzerrollen und Workflows

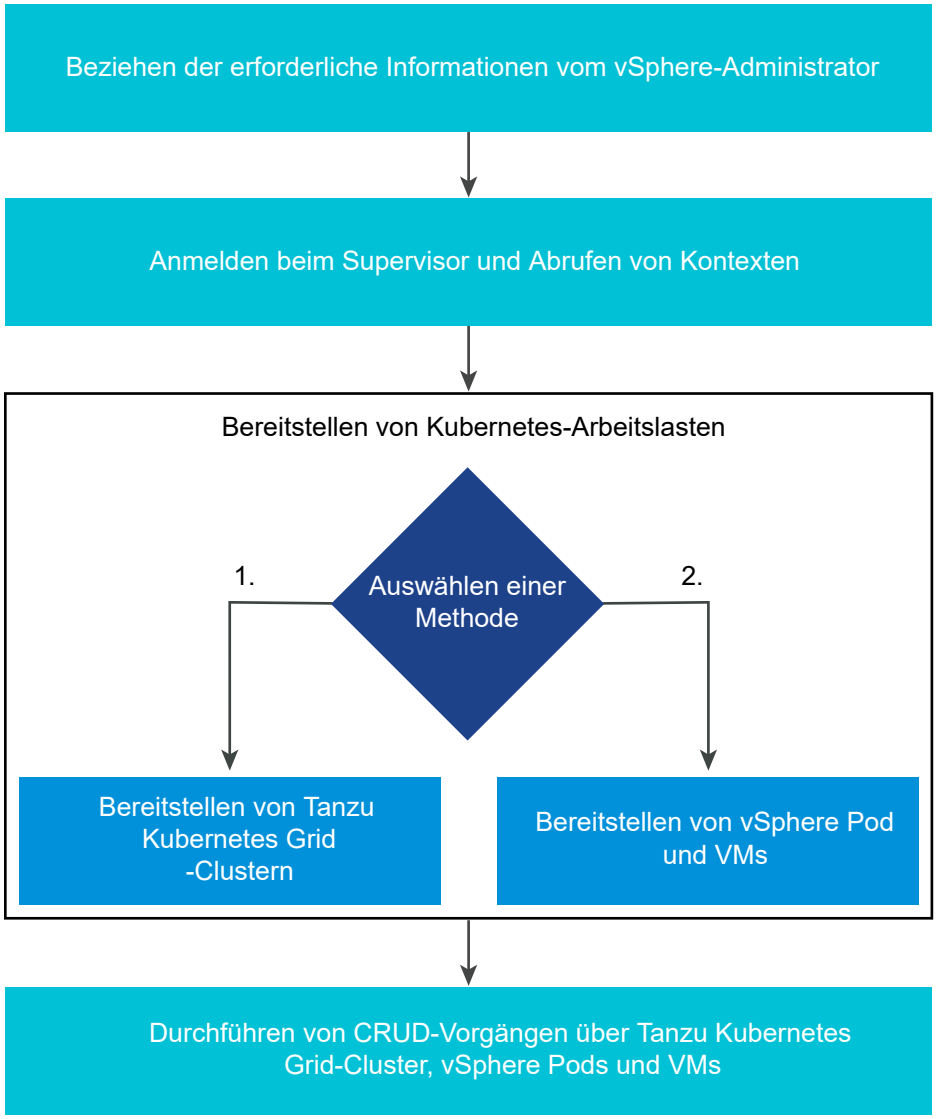
Als vSphere-Administrator ist die primäre Schnittstelle, über die Sie mit der vSphere IaaS control plane interagieren, der vSphere Client. Auf hoher Ebene zählt zu Ihren Zuständigkeiten die Konfiguration eines Supervisor und von Namespaces, in denen DevOps-Ingenieure Kubernetes-Arbeitslasten bereitstellen können. Sie sollten über exzellentes Wissen über vSphere, NSX Advanced Load Balancer oder den HAProxy-Lastausgleichsdienst, NSX (bei Auswahl dieses Netzwerk-Stacks) und grundlegende Kenntnisse von Kubernetes verfügen.

Abbildung 1-6. vSphere-Administrator-Workflow auf hoher Ebene



Als ein DevOps-Ingenieur können Sie als Kubernetes-Entwickler und Anwendungsbesitzer oder als Kubernetes-Administrator fungieren. Eine Kombination beider Funktionen ist auch möglich. Als DevOps-Ingenieur nutzen Sie kubectl-Befehle zum Bereitstellen von vSphere-Pods und VMs in vorhandenen Namespaces; zum Bereitstellen und Verwalten von Tanzu Kubernetes Grid-Clustern verwenden Sie kubectl und Tanzu-CLI. In der Regel müssen Sie als DevOps-Ingenieur kein Experte für vSphere, NSX, vDS oder NSX Advanced Load Balancer und HAProxy sein. Grundkenntnisse zu diesen Technologien und zur Plattform sind aber erforderlich, um effizienter mit vSphere-Administratoren interagieren zu können.

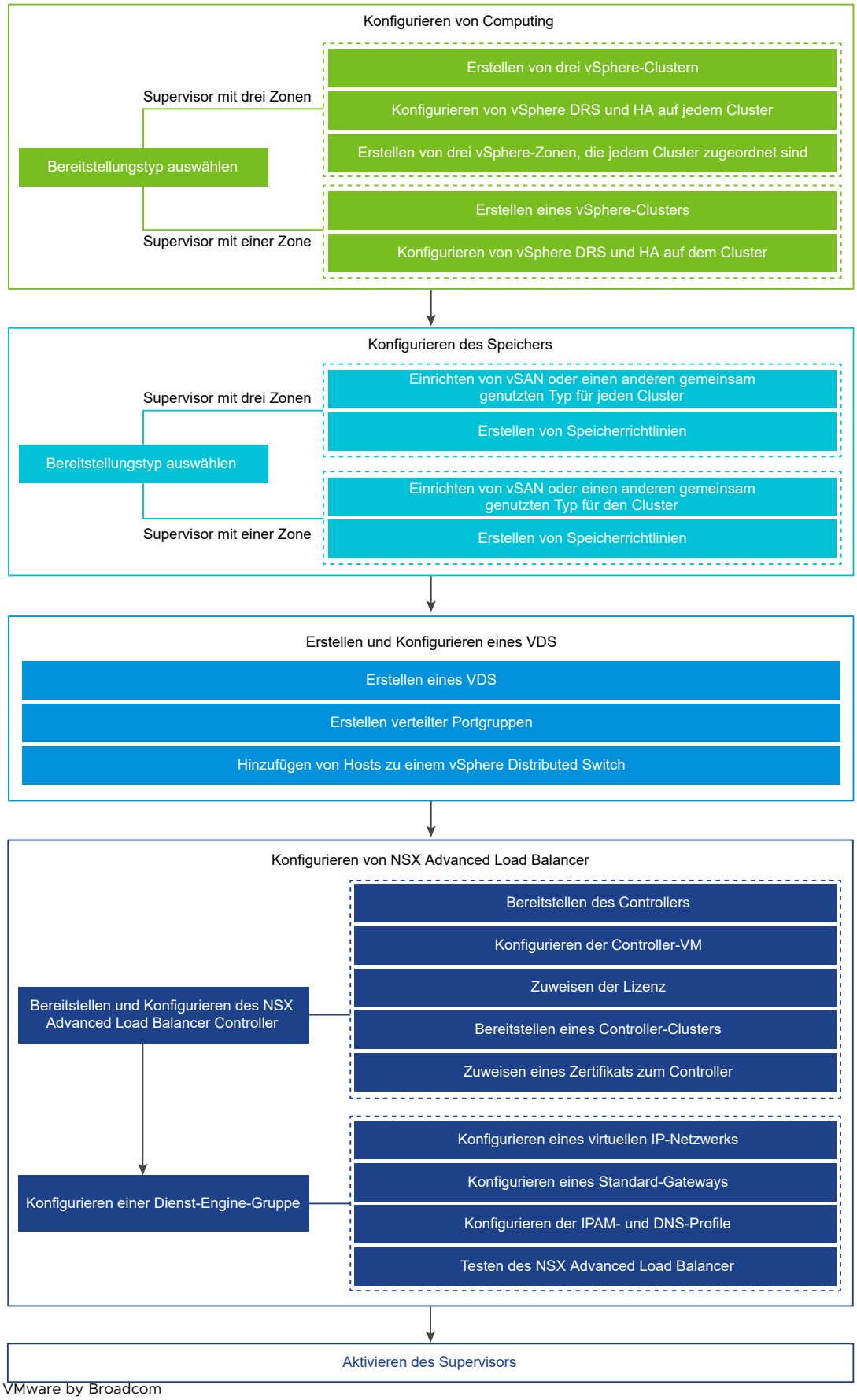
Abbildung 1-7. DevOps-Ingenieur-Workflow auf hoher Ebene



Workflow für Supervisor mit VDS-Netzwerk und NSX Advanced Load Balancer

Als vSphere-Administrator können Sie vSphere-Cluster als Supervisor mit dem vSphere-Netzwerk-Stack über eine VDS und NSX Advanced Load Balancer konfigurieren. Sie können einen Supervisor mit einer Zone und Zuordnung zu einem vSphere-Cluster konfigurieren oder einen Supervisor mit drei Zonen und Zuordnung zu drei vSphere-Clustern. Weitere Informationen zu den Systemanforderungen finden Sie im Abschnitt zu Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX Advanced Load Balancer und VDS-Netzwerk sowie im Abschnitt zu Voraussetzungen für die Zonen-Supervisor-Bereitstellung mit NSX Advanced Load Balancer und VDS-Netzwerk. Informationen zum Aktivieren eines Supervisor mit VDS-Netzwerk finden Sie unter Installieren und Konfigurieren in*Installieren und Konfigurieren der vSphere IaaS-Steuerungsebene*.

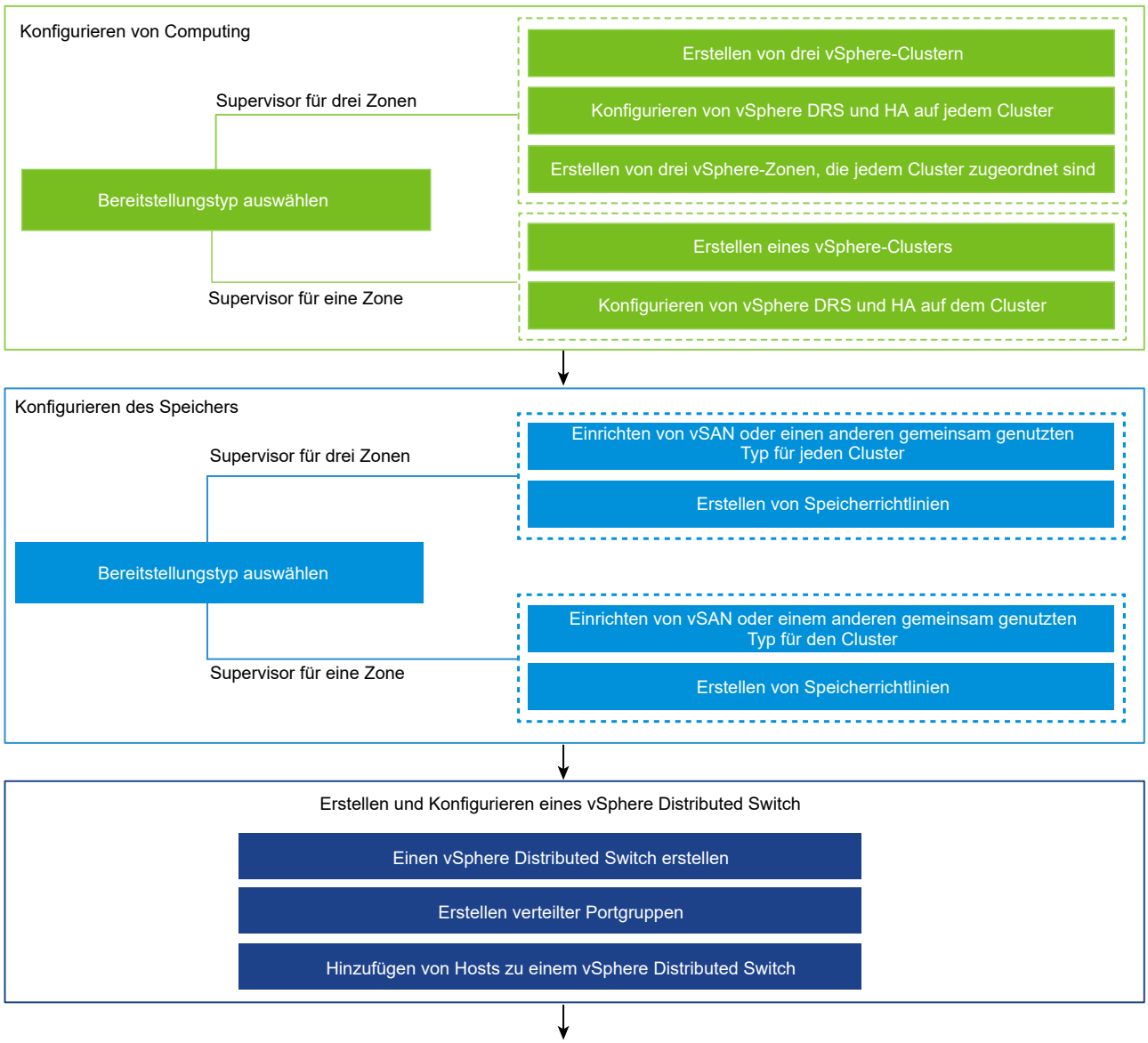
Abbildung 1-8. Workflow zum Aktivieren eines Supervisor mit VDS-Netzwerk und NSX Advanced Load Balancer



Workflow für Supervisor mit NSX-Netzwerk und NSX Advanced Load Balancer Controller

Sie können eine Supervisor mit einer Zone oder mit drei Zonen mit demNSX-Netzwerk-Stack und dem NSX Advanced Load Balancer Controller konfigurieren. Weitere Informationen zu den Voraussetzungen finden Sie unter Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX und NSX Advanced Load Balancer und Voraussetzungen für zonalen Supervisor mit NSX und NSX Advanced Load Balancer. Informationen zum Installationsvorgang finden Sie unter Installieren und Konfigurieren von NSX und NSX Advanced Load Balancer.

Abbildung 1-9. Workflow zur Aktivierung eines Supervisors mit NSX-Netzwerk und NSX Advanced Load Balancer Controller

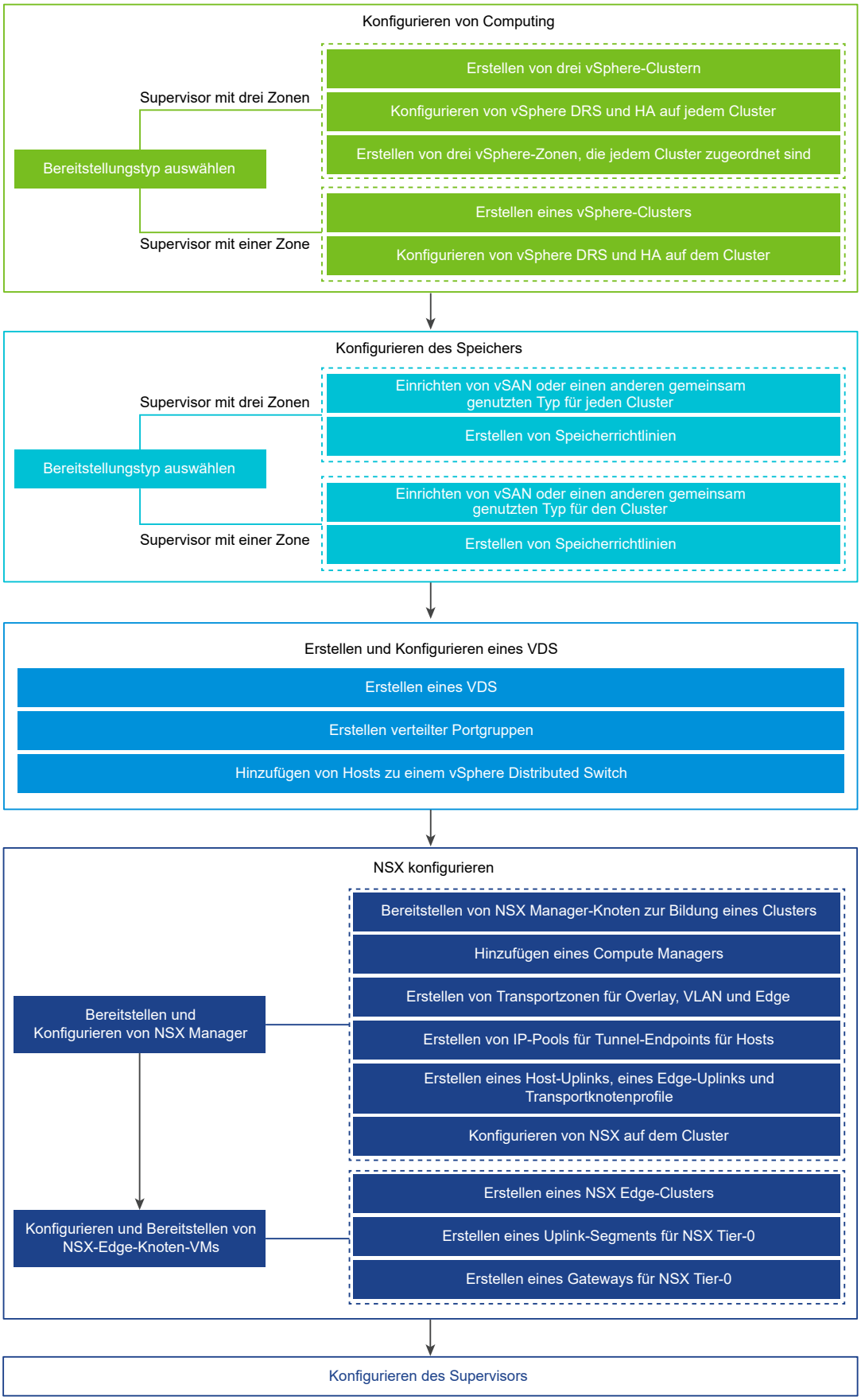



NSX konfigurieren
Bereitstellen und Konfigurieren von NSX Manager
Bereitstellen von NSX Manager-Knoten zur Bildung eines Clusters
Hinzufügen eines Compute Managers
Erstellen von Transportzonen für Overlay, VLAN und Edge
Erstellen von IP-Pools für Tunnel-Endpoints für Hosts
Erstellen eines Host-Uplinks, Edge-Uplinks, Transportknotens und Edge-Clusterprofilen
Konfigurieren von NSX auf dem Cluster
Konfigurieren und Bereitstellen von NSX Edge-Knoten-VMs
Erstellen eines NSX Edge-Clusters
Erstellen eines Gateways für NSX Tier-0
Erstellen eines Tier-0-Uplink-Segments
Erstellen eines Gateways für NSX Tier-1
Bereitstellen und Konfigurieren des NSX Advanced Load Balancer Controller
Bereitstellen des Controllers
Bereitstellen eines Controller-Clusters
Einschalten des Controllers
Konfigurieren der Controller-VM und Zuweisen einer Lizenz
Konfigurieren einer Dienst-Engine-Gruppe
Registrieren des Controllers
Zuweisen eines Zertifikats zum Controller
Testen des NSX Advanced Load Balancer
Konfigurieren des Supervisors

Workflow für Supervisor mit NSX-Netzwerk

Sie können auch einen Supervisor mit einer Zone oder mit drei Zonen unter Verwendung von NSX als Netzwerk-Stack konfigurieren. Weitere Informationen zu den Systemanforderungen finden Sie im Abschnitt zu Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX sowie im Abschnitt zu Voraussetzungen für zonalen Supervisor mit NSX. Installationsanweisungen finden Sie unter Installieren und Konfigurieren in *Installieren und Konfigurieren der vSphere IaaS-Steuerungsebene*.

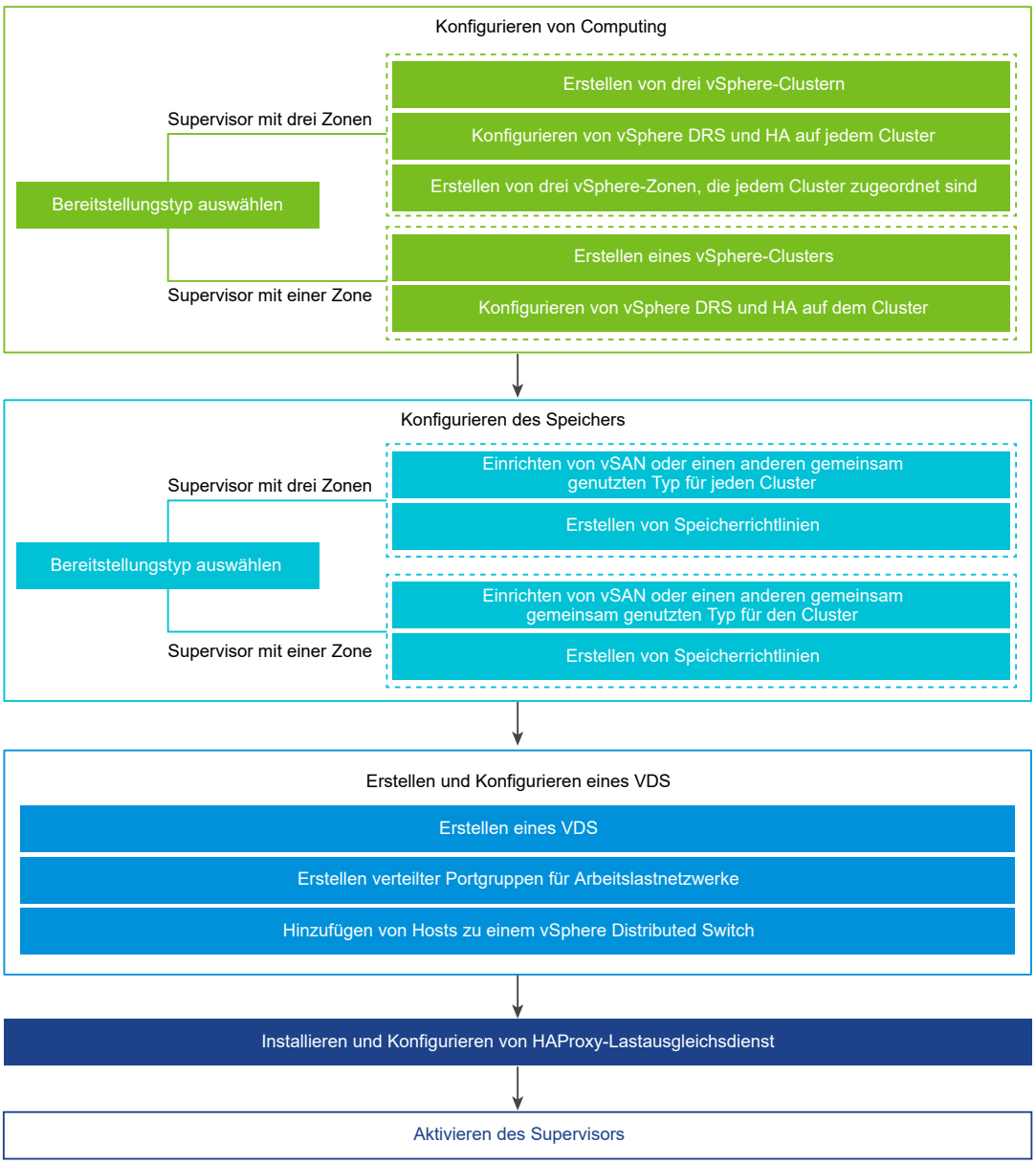
Abbildung 1-10. Workflow zum Aktivieren eines Supervisors mit NSX-Netzwerk



Workflow für Supervisor mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst

Als vSphere-Administrator können Sie einen Supervisor in einer oder drei vSphere-Zonen mit Zuordnung zu vSphere-Clustern aktivieren, indem Sie den VDS-Netzwerk-Stack und den HAProxy-Lastausgleichsdienst verwenden. Weitere Informationen zu den Systemanforderungen finden Sie im Abschnitt zu Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst sowie im Abschnitt zu Voraussetzungen für die zonale Supervisor-Bereitstellung mit HAProxy-Lastausgleichsdienst. Installationsanweisungen finden Sie unter Installieren und Konfigurieren in *Installieren und Konfigurieren der vSphere IaaS-Steuerungsebene*.

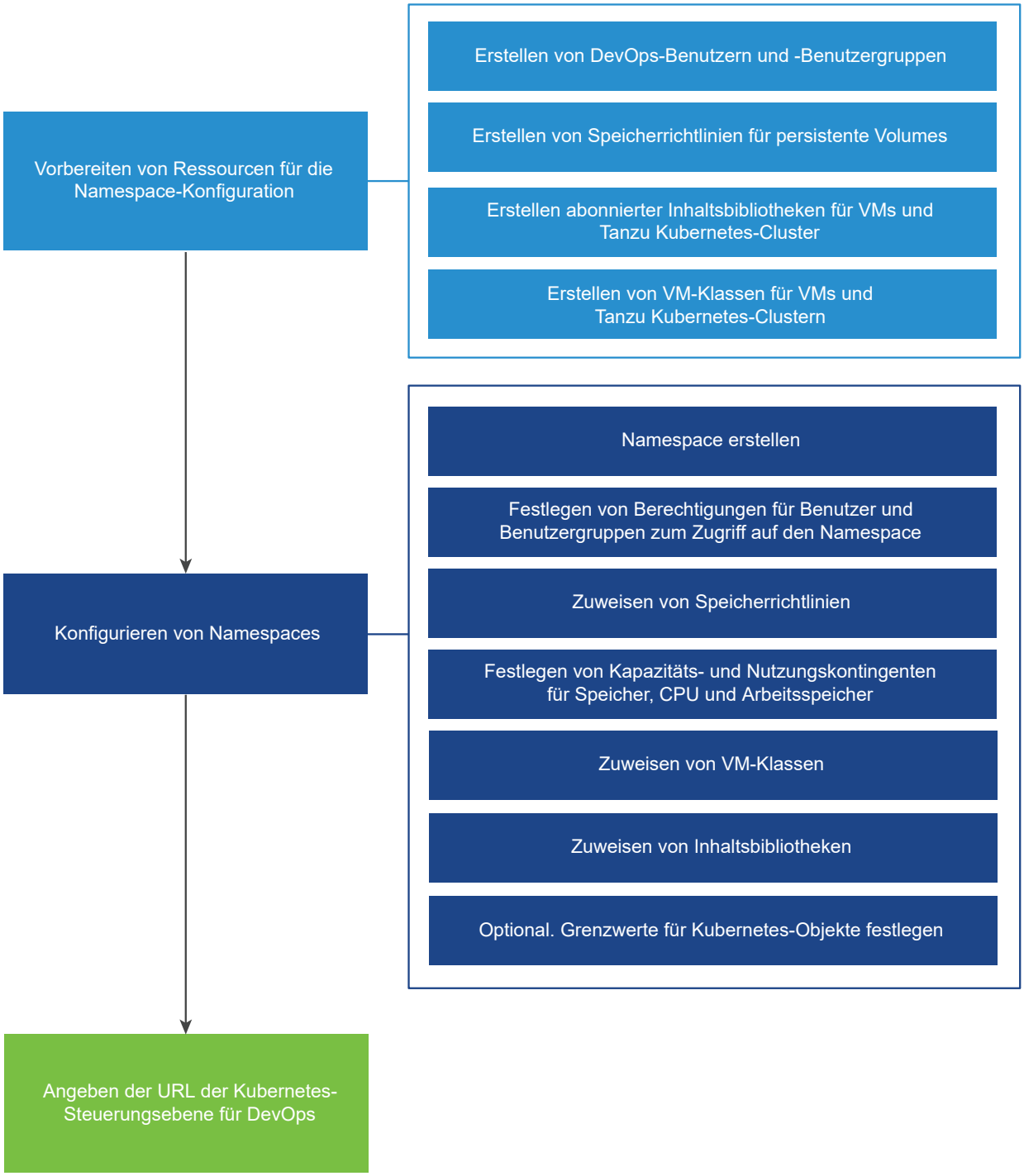
Abbildung 1-11. Workflow zum Aktivieren eines Supervisors mit VDS-Netzwerk und HAProxy



Workflow zum Erstellen und Konfigurieren von Namespaces

Nachdem Sie einen Supervisor als vSphere-Administrator aktiviert haben, erstellen und konfigurieren Sie vSphere-Namespaces im Supervisor. Sie müssen die spezifischen Ressourcenanforderungen von DevOps-Ingenieuren über die Anwendungen und Arbeitslasten, die sie ausführen möchten, erfassen und die Namespaces entsprechend konfigurieren. Weitere Informationen finden Sie unter Konfigurieren und Verwalten von vSphere-Namespaces.

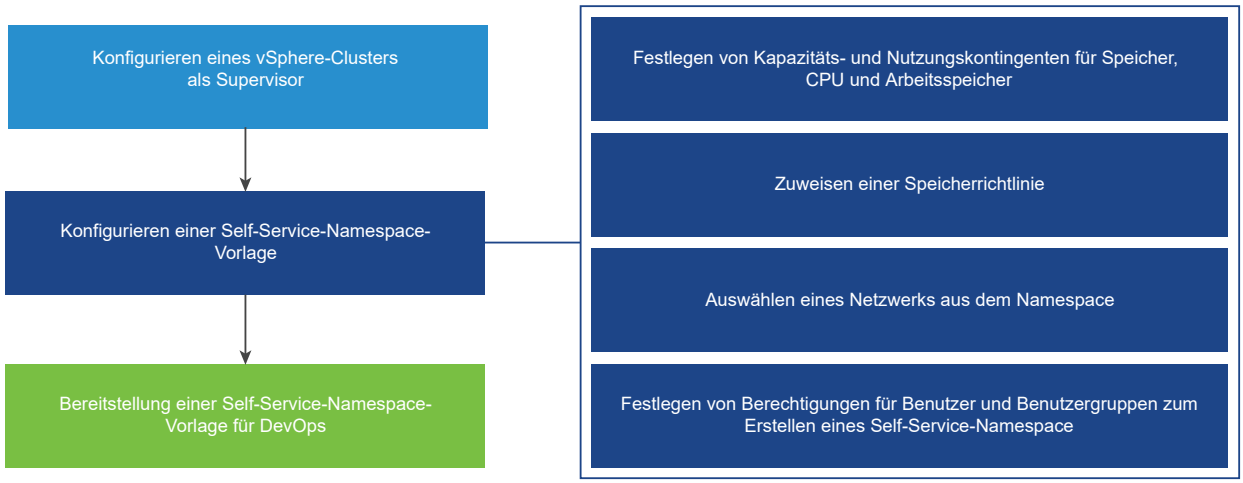
Abbildung 1-12. Workflow zum Konfigurieren von vSphere-Namespaces



Workflow zum Erstellen und Konfigurieren eines Self-Service-Namespace

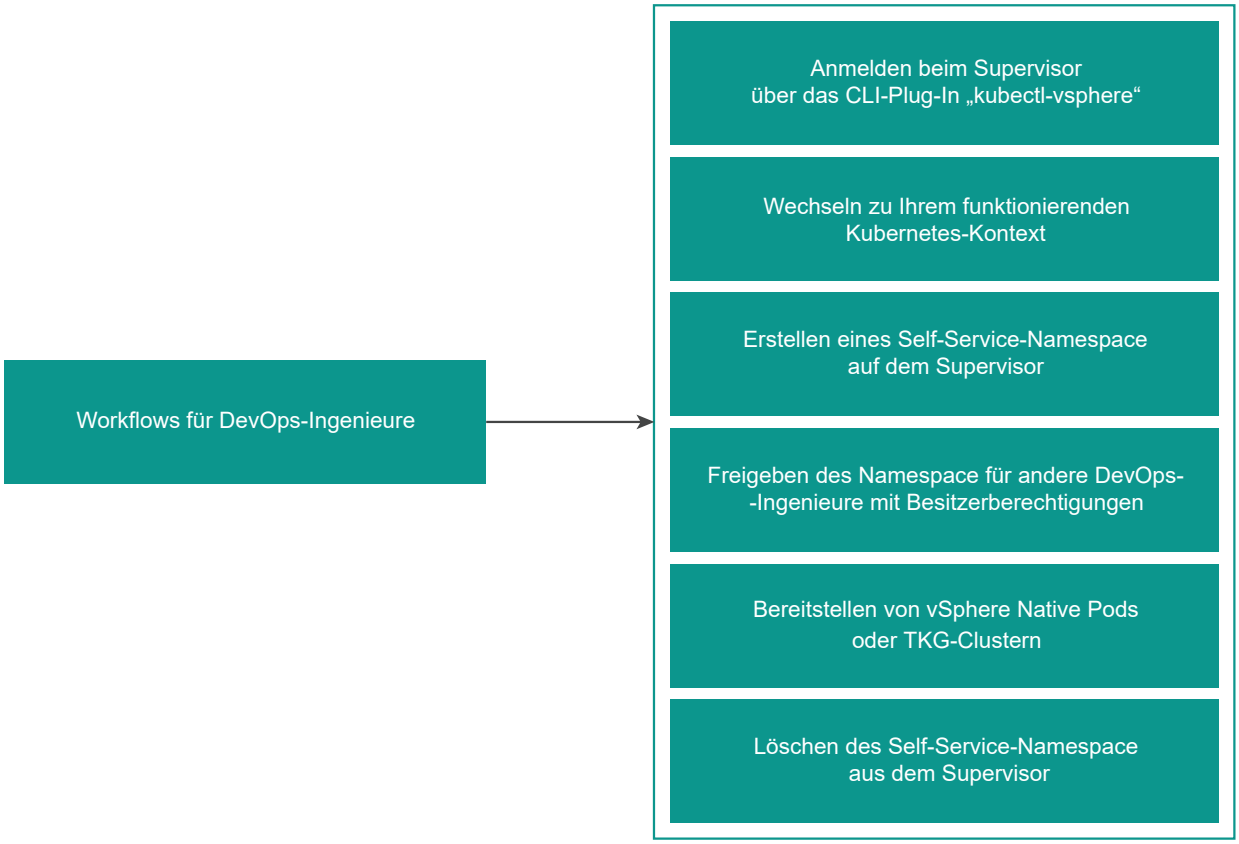
Als vSphere-Administrator können Sie einen vSphere-Namespace erstellen, CPU-, Arbeitsspeicher- und Speichergrenzwerte für den Namespace festlegen, Berechtigungen zuweisen und den Namespace-Dienst auf einem Cluster als Vorlage bereitstellen oder aktivieren. Weitere Informationen finden Sie unter Konfigurieren und Verwalten von vSphere-Namespaces.

Abbildung 1-13. Workflow zur Bereitstellung einer Self-Service-Namespace-Vorlage



Als DevOps-Ingenieur können Sie einen vSphere-Namespace in einer Self-Service-Form erstellen und Arbeitslasten darin bereitstellen. Sie können ihn für andere DevOps-Ingenieure freigeben oder löschen, wenn sie nicht mehr benötigt wird.

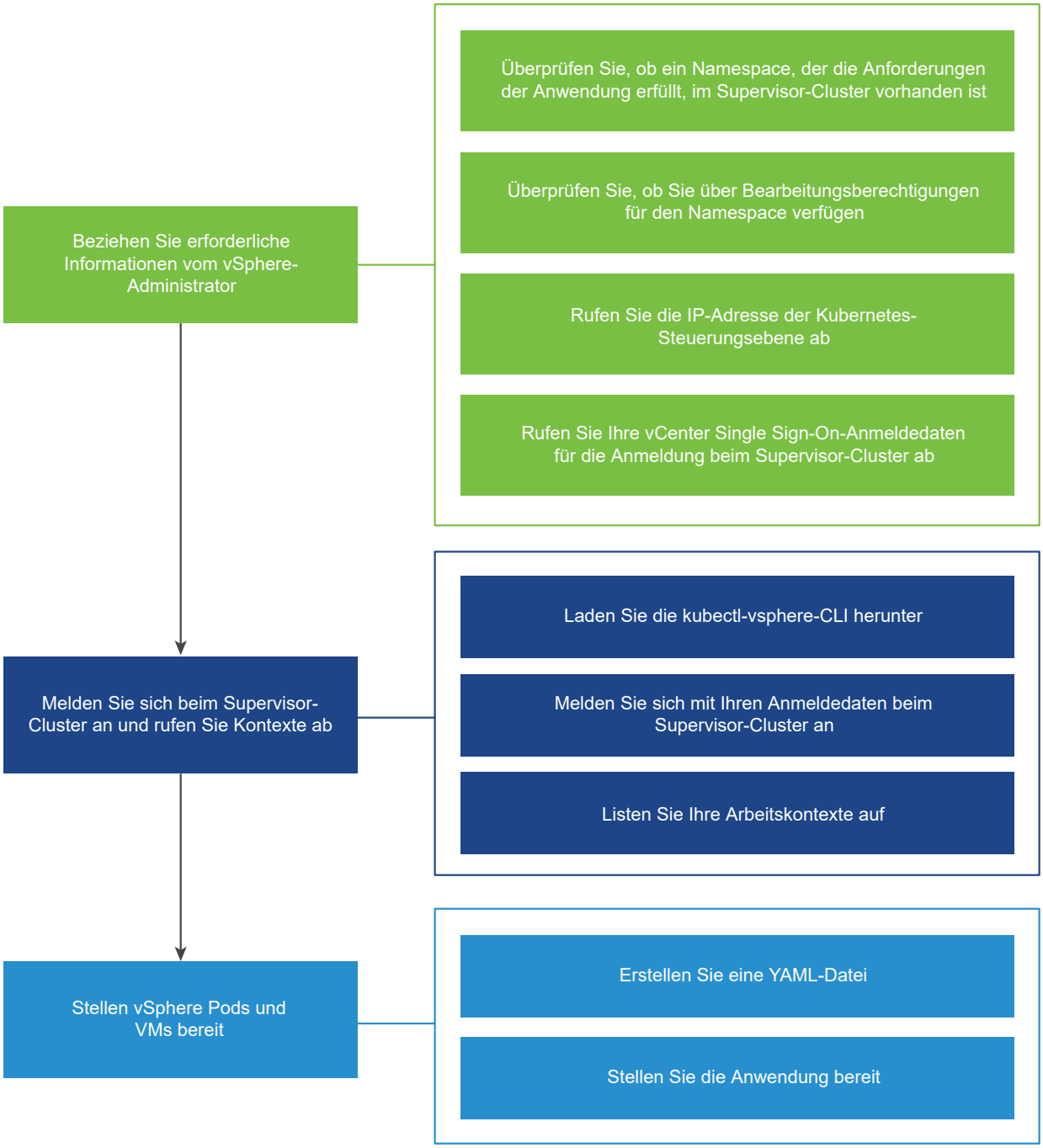
Abbildung 1-14. Workflow zur Erstellung von Self-Service-Namespaces



Workflow für vSphere Pod und VM-Bereitstellung

Als DevOps-Ingenieur können Sie vSphere-Pods und VMs innerhalb der Ressourcengrenzen eines auf einem Supervisor ausgeführten Namespace bereitstellen. Weitere Informationen finden Sie unter Bereitstellen von Arbeitslasten in vSphere Pods und Bereitstellen und Verwalten von virtuellen Maschinen in *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

Abbildung 1-15. Workflow für vSphere-Pods und VM-Bereitstellung



Tanzu Kubernetes Grid-Cluster – Bereitstellungs-Workflow

Als DevOps-Ingenieur erstellen und konfigurieren Sie Tanzu Kubernetes Grid-Cluster auf vSphere-Namespaces. Weitere Informationen finden Sie im Handbuch *Verwenden des TKG-Dienstes mit der vSphere IaaS-Steuerungsebene*.

Wie ändert vSphere IaaS control plane die vSphere-Umgebung?

Ein Supervisor fügt der vCenter Server-Bestandsliste Objekte hinzu, wie z. B. Namespaces, vSphere-Pods und Tanzu Kubernetes Grid-Cluster.

Unter jedem Supervisor können Sie Folgendes anzeigen:

- Namespaces, die im Cluster ausgeführte logische Anwendungen darstellen.
- Ressourcenpools für jeden Namespace im Supervisor. In einer Bereitstellung mit drei Zonen wird ein Ressourcenpool für jeden Namespace in jeder Clusterkomponente einer Zone erstellt.

In jedem Namespace können Sie Folgendes anzeigen:

- vSphere-Pods.
- Tanzu Kubernetes Grid-Cluster.
- Kubernetes-Steuerungsebenen-VMs und eigenständige VMs.
- Netzwerk- und Speicherressourcen.
- Benutzerberechtigungen für diesen Namespace.

Lizenzierung für vSphere IaaS control plane

Erfahren Sie, welche verschiedenen Lizenzen Sie dem Supervisor zuweisen können und wie die Lizenzkonformität, der Testzeitraum und der Lizenzablauf funktionieren.

Lizenzieren eines Supervisors

Nach der Aktivierung eines Supervisors auf vSphere-Clustern können Sie den gesamten Funktionssatz des Supervisors für einen 60-Tage-Testzeitraum nutzen. Sie müssen dem Supervisor vor Ablauf des 60-Tage-Testzeitraums eine gültige Lizenz zuweisen.

VCF- und VVF-Lösungslizenzen

Ab Version vSphere 8 Update 2b können Sie die Lösungslizenzen VMware vSphere Foundation (VVF) oder VMware Cloud Foundation (VCF) für vSphere IaaS control plane verwenden. Nach dem Upgrade von vCenter Server auf Version 8 Update 2b können Sie den Supervisoren in Ihrer vSphere-Umgebung die VVF- oder VCF-Lösungslizenz zuweisen.

Hinweis Einzelne Komponentenlizenzschlüssel werden weiterhin unterstützt. Sie werden zusammen mit der Lösungslizenz bereitgestellt. In Ihrer Umgebung können Sie die Lösungslizenz, die einzelnen Komponentenlizenzen oder eine Kombination aus beiden verwenden.

Lizenzen der Tanzu Edition

Wenn Sie vSphere 8 Update 2b ausführen und Ihr Supervisor bereits mit einer gültigen Tanzu Edition-Lizenz lizenziert ist, funktioniert die Lizenz weiterhin, bis sie abläuft. Nachdem die Tanzu-Lizenz abgelaufen ist, müssen Sie dem Supervisor eine VCF- oder VVF-Lösungslizenz oder eine gültige Tanzu-Lizenz zuweisen.

Lizenzablauf

Bei Ablauf einer Lösungslizenz oder einer Tanzu Edition-Lizenz können Sie den gesamten Funktionensatz von vSphere IaaS control plane weiterverwenden, bis Sie neue Lizenzen erwerben. Sie können die abgelaufene Lizenz jedoch nicht auf neuen Supervisoren zuweisen.

Ablauf des Testzeitraums

Wenn der Testzeitraum für einen Supervisor abläuft, können Sie als vSphere Administrator keine neuen vSphere-Namespaces erstellen oder die Kubernetes-Version des Supervisors aktualisieren. Als DevOps-Ingenieur können Sie weder neue Arbeitslasten bereitstellen noch Änderungen an der Konfiguration der vorhandenen Tanzu Kubernetes Grid-Cluster vornehmen. Zum Beispiel können Sie keine neuen Knoten hinzufügen.

Sie können weiterhin Arbeitslasten auf Tanzu Kubernetes Grid-Clustern bereitstellen. Alle vorhandenen Arbeitslasten werden weiterhin wie erwartet ausgeführt. Alle bereits bereitgestellten Kubernetes-Arbeitslasten werden weiterhin im normalen Betrieb fortgesetzt.

Lizenzkonformität

Eine Lösungslizenz oder ein Tanzu-Lizenzschlüssel hat eine Kapazität von bis zu 32 Kernen pro CPU, ähnlich wie die ESXi-Hostlizenzen. Wenn Sie einem Supervisor eine dieser Lizenzen zuweisen, richtet sich die verbrauchte Kapazität nach der Anzahl der CPUs auf den Hosts in den Clustern und nach der Anzahl der Kerne in jeder CPU. Sie können eine Lösungslizenz oder einen Tanzu Edition-Lizenzschlüssel mehreren Supervisoren gleichzeitig zuweisen, aber das Zuweisen mehrerer Lizenzschlüssel zu einem Supervisor ist nicht möglich.

Wenn Sie einen Supervisor (beispielsweise durch Hinzufügen neuer Hosts) erweitern und die Kapazität des Lizenzschlüssels, den Sie dem Supervisor zugewiesen haben, erschöpft ist, können Sie denselben Lizenzschlüssel weiterhin verwenden. Um jedoch EULA-konform zu bleiben, müssen Sie einen neuen Lizenzschlüssel mit ausreichender Kapazität erwerben, damit alle CPUs und Kerne im Supervisor abgedeckt werden.

Lizenzen für vSphere IaaS control plane

Je nach Netzwerk-Stack, mit dem Sie die vSphere IaaS control plane konfiguriert haben, variieren die bereitgestellten Lizenzen:

Einrichtung von Supervisor	Lizenzen für vSphere 8 Update 2b	Lizenzen vor vSphere 8 Update 2b
Supervisor mit VDS-Netzwerk und NSX Advanced Load Balancer	■ VCF-Lösungslizenz ■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition ■ NSX Advanced Load Balancer Essentials	■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition ■ NSX Advanced Load Balancer Essentials
Supervisor mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst	■ VVF-Lösungslizenz ■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition	■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition
Supervisor mit NSX	■ VVF-Lösungslizenz ■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition ■ NSX Advanced oder höher	■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition ■ NSX Advanced oder höher
Supervisor mit NSX und NSX Advanced Load Balancer	■ VCF-Lösungslizenz ■ NSX Advanced Load Balancer Enterprise ■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition ■ NSX Advanced oder höher ■ NSX Advanced Load Balancer Enterprise	■ Lizenz für vSphere Enterprise+ ■ Lizenz der Tanzu Edition ■ NSX Advanced oder höher ■ NSX Advanced Load Balancer Enterprise

vSphere IaaS control planeIdentitäts- und Zugriffsverwaltung

Als vSphere-Administrator benötigen Sie Rechte zum Konfigurieren eines Supervisors und zum Verwalten von vSphere-Namespaces. Sie definieren Berechtigungen für Namespaces, um festzulegen, welche DevOps-Ingenieure und Entwickler auf sie zugreifen können. Sie können den Supervisor auch mit einem externen OIDC-Anbieter (OpenID Connect) konfigurieren, um die Multifaktor-Authentifizierung zu aktivieren. Als DevOps-Ingenieur oder Entwickler authentifizieren Sie sich beim Supervisor, indem Sie entweder Ihre vCenter Single Sign-On-Anmeldedaten oder die Anmeldedaten eines OIDC-Anbieters verwenden, je nachdem, was Ihr vSphere-Administrator für Sie auf dem Supervisor konfiguriert hat. Sie können nur auf die vSphere-Namespaces zugreifen, für die Sie über Berechtigungen verfügen.

Unterstützte Identitätsanbieter

vSphere IaaS control plane unterstützt folgende Identitätsanbieter:

- vCenter Single Sign-On. Der Standardidentitätsanbieter, den Sie zur Authentifizierung bei der vSphere IaaS control plane-Umgebung verwenden, einschließlich Supervisoren und Tanzu

Kubernetes Grid-Clustern. vCenter Single Sign-On bietet Authentifizierung für die vSphere-Infrastruktur und kann in AD/LDAP-Systeme integriert werden. Weitere Informationen zu vCenter Single Sign-On finden Sie unter vSphere-Authentifizierung mit vCenter Single Sign-On.

- Externer Identitätsanbieter. Als vSphere-Administrator können Sie einen Supervisor mit einem externen Identitätsanbieter konfigurieren, der das OpenID Connect-Protokoll unterstützt. Nach der Konfiguration mit einem externen Identitätsanbieter fungiert der Supervisor als OAuth 2.0-Client und verwendet den Authentifizierungsdienst Pinniped, um mithilfe der Tanzu-CLI eine Verbindung zu Tanzu Kubernetes Grid-Clustern herzustellen. Die Tanzu-CLI unterstützt die Bereitstellung und Verwaltung des Lebenszyklus von Tanzu Kubernetes Grid-Clustern. Die Supervisor-Instanz kann einen externen Identitätsanbieter unterstützen.

Authentifizierung beim Supervisor

Die verschiedenen Rollen, die mit dem vSphere IaaS control plane interagieren, können sich mit den folgenden Methoden beim Supervisor authentifizieren:

- vSphere-Administrator. Als vSphere-Administrator verwenden Sie vCenter Single Sign-On, um sich über die vSphere Client bei vSphere zu authentifizieren. Sie können sich auch mit dem vSphere-Plug-In für kubectl beim Supervisor authentifizieren und Cluster über kubectl Tanzu Kubernetes Grid. Weitere Informationen finden Sie unter Herstellen einer Verbindung mit einem Supervisor als vCenter Single Sign-On-Benutzer.
- DevOps-Ingenieur oder -Entwickler. Als DevOps-Ingenieur oder -Entwickler verwenden Sie vCenter Single Sign-On, um sich über die vSphere-Plug-In für kubectl und kubectl beim Supervisor zu authentifizieren. Sie können auch eine Verbindung zum Supervisor herstellen, indem Sie Anmeldedaten von den externen Identitätsanbieter verwenden, der beim Supervisor konfiguriert ist. Weitere Informationen finden Sie unter Verbindung zu TKG-Clustern auf dem Supervisor über einen externen Identitätsanbieter.

Anmeldesitzungen mit dem Supervisor

Wenn Sie sich beim Supervisor als vCenter Single Sign-On-Benutzer anmelden, leitet ein Authentifizierungs-Proxy die Anforderung an vCenter Single Sign-On weiter. Das vSphere-Plug-In für kubectl richtet eine Sitzung mit vCenter Server ein und erhält ein Authentifizierungstoken von vCenter Single Sign-On. Es ruft auch eine Liste der vSphere-Namespaces ab, auf die Sie Zugriff haben, und füllt die Konfiguration mit diesen vSphere-Namespaces auf. Die Liste der vSphere-Namespaces wird bei der nächsten Anmeldung aktualisiert, wenn Änderungen an den Berechtigungen Ihres Benutzerkontos vorgenommen wurden.

Das Konto, mit dem Sie sich beim Supervisor anmelden, bietet Ihnen nur Zugriff auf die vSphere-Namespaces, die Ihnen zugewiesen sind. Um sich bei vCenter Server anzumelden, muss Ihr vSphere-Administrator die entsprechenden Berechtigungen für Ihr Konto bei einem oder mehr vSphere-Namespaces festlegen.

Hinweis Die Sitzung mit kubectl dauert 10 Stunden. Nach Ablauf der Sitzung müssen Sie sich erneut beim Supervisor authentifizieren. Bei der Abmeldung wird das Token aus der Konfigurationsdatei Ihres Benutzerkontos gelöscht. Es bleibt aber bis zum Ende der Sitzung gültig.

Authentifizierung bei Tanzu Kubernetes Grid-Clustern

Als DevOps-Ingenieur oder Entwickler stellen Sie auch eine Verbindung zu bereitgestellten Tanzu Kubernetes Grid-Clustern her, um sie zu betreiben und zu verwalten. Wenn Ihrem Benutzerkonto eine Bearbeitungs- oder Besitzerberechtigung für den vSphere-Namespace erteilt wird, in dem der Tanzu Kubernetes Grid-Cluster bereitgestellt wird, wird Ihr Konto der Rolle `cluster-admin` zugewiesen. Alternativ können Sie auch den Benutzertyp `kubernetes-admin` verwenden, um eine Verbindung zu Tanzu Kubernetes Grid aufzubauen. Außerdem haben Sie die Möglichkeit, Entwicklern Zugriff auf Tanzu Kubernetes Grid-Cluster zu gewähren, indem Sie einen Benutzer oder eine Gruppe an die standardmäßige oder benutzerdefinierte Pod-Sicherheitsrichtlinie binden. Weitere Informationen finden Sie unter Verbindung zu TKG-Clustern auf dem Supervisor mit vCenter SSO-Authentifizierung and Verbindung zu TKG-Clustern auf dem Supervisor mit einem externen Identitätsanbieter.

vSphere-NamespacesRollenberechtigungen

Als vSphere-Administrator erteilen Sie DevOps-Ingenieuren oder -Entwicklern Ansichts-, Bearbeitungs- oder Besitzerberechtigungen für vSphere-Namespaces. Ihre Benutzer oder Gruppen müssen in vCenter Single Sign-On oder in einem externen Identitätsanbieter verfügbar sein, der mit dem Supervisor konfiguriert ist. Ein Benutzer oder eine Gruppe können auf mehrere vSphere-Namespaces zugreifen. Jede vSphere-Namespace-Rolle ermöglicht die folgenden Aktionen:

Rolle	Beschreibung
Kann ansehen	Schreibgeschützter Zugriff für den Benutzer oder die Gruppe. Der Benutzer oder die Gruppe kann sich bei der Supervisor-Steuerungsebene anmelden und die Arbeitslasten, die in der vSphere-Namespace ausgeführt werden, wie z. B. vSphere-Pods- und Tanzu Kubernetes Grid-Cluster und VMs, auflisten.
Kann bearbeiten	Der Benutzer oder die Gruppe kann vSphere-Pods, Tanzu Kubernetes Grid Cluster und VMs erstellen, lesen, aktualisieren und löschen. Benutzer, die Teil der Administratorengruppe sind, verfügen über Bearbeitungsberechtigungen für alle Namespaces im Supervisor.
Besitzer	Benutzer oder Gruppen mit Besitzerberechtigungen können: ■ Stellen Sie Arbeitslasten im vSphere-Namespace bereit und verwalten Sie sie. ■ vSphere-Namespace für andere Benutzer oder Gruppen freigeben. ■ Erstellen und löschen Sie mithilfe von kubectl zusätzliche vSphere-Namespaces. Wenn Benutzer mit Besitzerberechtigungen den Namensraum freigeben, können sie anderen Benutzern oder Gruppen Ansichts-, Bearbeitungs- oder Besitzerberechtigungen zuweisen. **Hinweis** Die Besitzerrolle wird für Benutzer unterstützt, die in vCenter Single Sign-On verfügbar sind. Sie können die Rolle „Besitzer“ nicht mit einem Benutzer / einer Gruppe eines externen Identitätsanbieters verwenden.

Informationen zum Erstellen einer vSphere-Namespaces finden Sie unter Erstellen und Konfigurieren eines vSphere-Namespace.

Nachdem Sie als vSphere-Administrator einen vSphere-Namespace mit Rollenberechtigungen, Ressourcenkontingenten und Speicher konfiguriert haben, stellen Sie die URL der Supervisor-Steuerungsebene den DevOps-Ingenieuren oder Entwicklern zur Verfügung, die sich damit bei der Steuerungsebene anmelden können. Nach der Anmeldung können DevOps-Ingenieure und Entwickler auf die vSphere-Namespaces zugreifen, für die sie über Berechtigungen verfügen, in den zu einem vCenter Server-System gehörenden Supervisorenn, die mit demselben Identitätsanbieter konfiguriert wurden. Wenn sich vCenter Server-Systeme im erweiterten verknüpften Modus befinden, können DevOps-Ingenieure auf alle vSphere-Namespaces, für die sie über Berechtigungen verfügen, in allen in der Gruppe „Verknüpfter Modus“ verfügbaren Supervisoren zugreifen. Die IP-Adresse der Supervisor-Steuerungsebene ist eine virtuelle IP-Adresse, die von NSX oder bei einem VDS-Netzwerk von einem Lastenausgleichsdienst generiert wird und als Zugriffspunkt auf die Supervisor-Steuerungsebene dient.

vSphere-Administratorberechtigungen

Als vSphere-Administrator kann Ihr Benutzerkonto in der Regel über die folgenden Berechtigungen verfügen:

Objekt	Berechtigungen
Benutzer vCenter Single Sign-On	Administratorengruppe
Benutzer vSphere-Namespaces	Mitgliedern der Gruppe „Administratoren“ werden Bearbeitungsberechtigungen auf allen vSphere-Namespaces gewährt.

Abhängig von der Schnittstelle, die Sie für die Interaktion mit vSphere IaaS control plane verwenden, können Sie verschiedene Vorgänge mit Ihnen erteilten Berechtigungen durchführen:

Schnittstelle	Vorgänge
vSphere Client	Wenn Sie als Administrator bei der vSphere Client angemeldet sind, haben Sie folgende Möglichkeiten: ■ Aktivieren und Konfigurieren von Supervisoren ■ Erstellen und konfigurieren Sie vSphere-Namespaces mit Ressourcenzuteilungen und Rollenberechtigungen für DevOps-Ingenieure oder Entwickler. Rollenberechtigungen auf vSphere-Namespaces sind für Benutzer oder Gruppen erforderlich, die sich über kubectl bei der Supervisor-Steuerungsebene anmelden möchten, um Arbeitslasten verwalten zu können. ■ Bereitstellung und Verwaltung von Supervisor-Dienst auf Supervisoren
kubectl	Wenn Sie bei der Supervisor-Steuerungsebene mit einem vCenter Single Sign-On Administratorkonto angemeldet sind, haben Sie folgende Möglichkeiten: ■ Ressourcen in allen vSphere-Namespaces anzeigen, einschließlich system vSphere-Namespaces (kube-system und alle vmware-system-* Namespaces) ■ Bearbeiten Sie Ressourcen in allen Nicht-System-vSphere-Namespaces, bei denen es sich um Namespaces handelt, die über die vSphere Client- oder vCenter Server-APIs erstellt wurden. Wenn Sie jedoch bei der Supervisor-Steuerungsebene mit einem Konto, das Teil der Administratorgruppe ist, angemeldet sind, dürfen Sie keine Ressourcen auf Clusterebene bearbeiten, vSphere-Namespaces mithilfe von kubectl erstellen oder Rollenbindungen erstellen. Sie müssen das vSphere Client als primäre Schnittstelle verwenden, um Ressourcenkontingente festzulegen, vSphere-Namespaces zu erstellen und zu konfigurieren und Benutzerberechtigungen einzurichten.

DevOps-Ingenieure und Entwicklerberechtigungen

Als DevOps-Ingenieur oder -Entwickler benötigt Ihr Benutzerkonto in der Regel die folgenden Berechtigungen:

Objekt	Berechtigungen
vSphere-Namespaces	Bearbeiten oder Besitzer
Benutzer vCenter Single Sign-On	Nicht oder nur lesen

Als DevOps-Ingenieur oder -Entwickler verwenden Sie kubectl als primäre Schnittstelle für die Interaktion mit vSphere IaaS control plane. Sie müssen sich über die vSphere-Plug-In für kubectl bei der Supervisor-Steuerungsebene anmelden können, um Arbeitslasten in der ihnen zugewiesenen vSphere-Namespaces anzuzeigen, auszuführen und zu verwalten. Aus diesem Grund benötigt Ihr Benutzerkonto Bearbeitungs- oder Besitzerberechtigungen für eine oder mehrere vSphere-Namespaces.

In der Regel müssen Sie keine administrativen Vorgänge auf Supervisoren über das vSphere Client durchführen. In bestimmten Fällen möchten Sie sich jedoch beim vSphere Client anmelden können, um die Ressourcen und Arbeitslasten im vSphere-Namespaces anzuzeigen, die Ihrem Konto zugewiesen sind. Zu diesem Zweck benötigen Sie möglicherweise Leseberechtigungen in vSphere.

vSphere-Namespaces-Rechte

vSphere-Namespaces-Berechtigungen steuern, wie Sie mit vSphere IaaS control plane interagieren. Sie können eine Berechtigung auf verschiedenen Ebenen der Hierarchie festlegen. Wenn Sie beispielsweise ein Recht auf der Ordnerebene festgelegt haben, können Sie das Recht an ein oder mehrere Objekte innerhalb des Ordners weitergeben. Die in der Required On-Spalte aufgelisteten Objekte müssen, auf direkte oder geerbte Art und Weise, über das Recht verfügen.

Rechtename im vSphere Client	Beschreibung	Erforderlich bei	Rechtename in der API
Lässt Vorgänge zur Außerbetriebnahme von Festplatten zu	Ermöglicht die Außerbetriebnahme von Datenspeichern.	Datenspeicher	Namespaces.ManageDisks
Dateien der Arbeitslastkomponente sichern	Ermöglicht das Sichern der Inhalte des etcd-Clusters (wird nur in VMware Cloud on AWS verwendet).	Cluster	Namespaces.Backup
Verfügbare Namespaces auflisten	Ermöglicht das Auflisten der einsehbaren vSphere-Namespaces.	Cluster	Namespaces.ListAccess
Clusterweite Konfiguration ändern	Ermöglicht das Ändern der Supervisor-Konfiguration sowie das Erstellen und Löschen von vSphere-Namespaces.	Cluster	Namespaces.ManageCapabilities

Rechtename im vSphere Client	Beschreibung	Erforderlich bei	Rechtename in der API
Clusterweite Namespace-Self-Service-Konfiguration ändern	Ermöglicht das Ändern der vSphere-Namespace-Self-Service-Konfiguration.	Cluster (zum Aktivieren und Deaktivieren) Vorlagen (zum Ändern der Konfiguration) vCenter Server (zum Erstellen einer Vorlage)	Namespaces.SelfServiceManage
Namespace-Konfiguration ändern	Ermöglicht das Ändern der vSphere-Namespace-Konfigurationsoptionen wie Ressourcenzuteilung, Benutzerberechtigungen und Inhaltsbibliothekszuordnungen.	Cluster	Namespaces.Manage
Clusterfunktionen umschalten	Ermöglicht die Änderung des Status von Cluster-Supervisor-Funktionen (wird intern nur für VMware Cloud on AWS verwendet).	Cluster	–
Upgrade von Clustern auf neuere Versionen durchführen	Ermöglicht die Initiierung des Supervisor-Upgrades.	Cluster	Namespaces.Upgrade

Supervisor-Dienste-Rechte

Supervisor-Dienste-Rechte bestimmen, wer Supervisor-Dienste in der vSphere IaaS control plane-Umgebung erstellen und verwalten kann.

Tabelle 1-1. Supervisor-Dienste-Rechte

Rechtename im vSphere Client	Beschreibung	Erforderlich bei	Rechtename in der API
Supervisor-Dienste verwalten	Ermöglicht das Erstellen, Aktualisieren oder Löschen eines Supervisor-Dienst-Diensts. Darüber hinaus können Sie einen Supervisor-Dienst auf einem Supervisor installieren und eine Supervisor-Dienst-Version erstellen oder löschen.	Cluster	SupervisorServices.Manage

Rechte für VM-Klassen

Rechte für VM-Klassen steuern, wer VM-Klassen in einem vSphere-Namespace hinzufügen und entfernen kann.

Tabelle 1-2. Rechte für VM-Klassen

Rechtename im vSphere Client	Beschreibung	Erforderlich bei	Rechtename in der API
Klassen virtueller Maschinen verwalten	Ermöglicht die Verwaltung von VM-Klassen auf vSphere-Namespaces in einem Supervisor.	Cluster	VirtualMachineClasses.Manage

Speicheransichtsberechtigungen

In den Rechten für Speicheransichten können Sie Speicherrichtlinien in vCenter Server anzeigen, sodass Sie sie vSphere-Namespaces zuweisen können.

Tabelle 1-3. Speicheransichtsberechtigungen

Rechtename im vSphere Client	Beschreibung	Erforderlich bei	Rechtename in der API
Dienst konfigurieren	Erlaubt berechtigten Benutzern die Verwendung aller Speicherüberwachungsdienst-APIs. Verwenden Sie **Speicheransichten.Anzeigen** für Rechte bezüglich schreibgeschützter Speicherüberwachungsdienst-APIs.	Root-vCenter Server	StorageViews.ConfigureService
Anzeigen	Erlaubt berechtigten Benutzern die Verwendung schreibgeschützter Speicherüberwachungsdienst-APIs.	Root-vCenter Server	StorageViews.View

vSphere IaaS control planeSicherheit

vSphere IaaS control plane nutzt vSphere-Sicherheitsfunktionen und stellt Tanzu Kubernetes Grid-Cluster bereit, die standardmäßig sicher sind.

vSphere IaaS control plane ist ein Add-On-Modul für vSphere, das die in vCenter Server und ESXi integrierten Sicherheitsfunktionen nutzen kann. Weitere Informationen finden Sie in der Dokumentation zu vSphere-Sicherheit .

Der Supervisor verschlüsselt alle geheimen Schlüssel, die in der Datenbank (etcd) gespeichert werden. Die geheimen Schlüssel werden über eine lokale Verschlüsselungsschlüsseldatei verschlüsselt, die beim Start von vCenter Server bereitgestellt wird. Der Entschlüsselungsschlüssel wird im Arbeitsspeicher (tempfs) auf den Supervisor-Knoten und auf der Festplatte in verschlüsselter Form innerhalb der vCenter Server-Datenbank gespeichert. Der Schlüssel steht den Root-Benutzern jedes Systems als Klartext zur Verfügung. Die in der Datenbank befindlichen geheimen Schlüssel aller Arbeitslastcluster werden in Klartext gespeichert. Alle etcd-Verbindungen werden mit Zertifikaten authentifiziert, die bei der Installation generiert und während Upgrades rotiert werden. Eine manuelle Rotation oder Aktualisierung der Zertifikate ist derzeit nicht möglich. Dasselbe Verschlüsselungsmodell gilt für die Daten in der Datenbank (etcd), die auf der Steuerungsebene für jeden Tanzu Kubernetes Grid-Cluster installiert sind.

In Supervisor können Sie vertrauliche vSphere-Pods auf kompatiblen Systemen ausführen. Sie können vertrauliche vSphere-Pods erstellen, indem Sie SEV-ES (Secure Encrypted Virtualization-Encrypted State) als Sicherheitserweiterung hinzufügen. Weitere Informationen finden Sie unter Bereitstellen eines vertraulichen vSphere-Pods in *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

Ein Tanzu Kubernetes Grid-Cluster ist standardmäßig sicher. Für alle Tanzu Kubernetes Grid-Cluster ist eine eingeschränkte PodSecurityPolicy-Ressource (PSP) verfügbar. Wenn Entwickler berechtigte Pods oder Root-Container ausführen müssen, muss ein Clusteradministrator mindestens ein RoleBinding-Objekt erstellen, das dem Benutzer Zugriff auf die berechtigten Standard-PSP gewährt. Weitere Informationen finden Sie unter *Verwenden des TKG-Dienstes mit der vSphere IaaS-Steuerungsebene*.

Ein Tanzu Kubernetes Grid-Cluster verfügt nicht über Anmeldedaten für die Infrastruktur. Die Anmeldedaten, die in einem Tanzu Kubernetes Grid-Cluster gespeichert werden, reichen nur für den Zugriff auf den vSphere-Namespace aus, in dem der Tanzu Kubernetes Grid-Cluster mandantenfähig ist. Infolgedessen gibt es für Clusteroperatoren oder Benutzer keinen Eskalationsweg für Rechte.

Das für den Zugriff auf einen Tanzu Kubernetes Grid-Cluster verwendete Authentifizierungstoken wird so skaliert, dass das Token nicht für den Zugriff auf den Supervisor oder andere Tanzu Kubernetes Grid-Cluster genutzt werden kann. Dadurch wird verhindert, dass Clusteroperatoren oder Personen, die möglicherweise versuchen, einen Cluster zu kompromittieren, ihren Zugriff auf Root-Ebene nutzen, um das Token eines vSphere-Administrators zu erfassen, wenn sie sich bei einem Tanzu Kubernetes Grid-Cluster anmelden.

Supervisor-Architektur und -Komponenten

2

Die mit vSphere IaaS control plane aktivierten Cluster werden als Supervisoren bezeichnet. Sie können wählen zwischen einer Bereitstellung mit drei Zonen, bei der Sie einen Supervisor auf drei vSphere-Clustern aktivieren, und einer 1:1-Zuordnung zwischen einem vSphere-Cluster und einem Supervisor. Supervisoren bilden die Basis von vSphere IaaS control plane. Sie bieten die erforderlichen Komponenten und Ressourcen für das Ausführen von Arbeitslasten, die vSphere-Pods, VMs und Tanzu Kubernetes Grid-Cluster enthalten.

Lesen Sie als Nächstes die folgenden Themen:

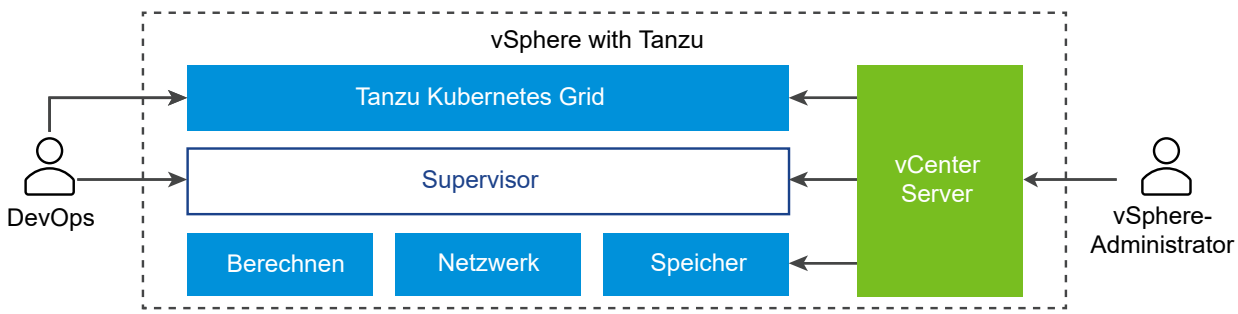
- Supervisor – Architektur
- Supervisor-Netzwerk
- Supervisor-Speicher

Supervisor – Architektur

Wenn Sie vSphere IaaS control plane auf vSphere-Clustern aktivieren und diese zu Supervisoren werden, wird eine Kubernetes-Steuerungsebene innerhalb der Hypervisor-Schicht erstellt. Diese Schicht enthält bestimmte Objekte, die das Ausführen von Kubernetes-Arbeitslasten innerhalb von ESXi ermöglichen.

Abbildung 2-1. Allgemeine Supervisor-Architektur

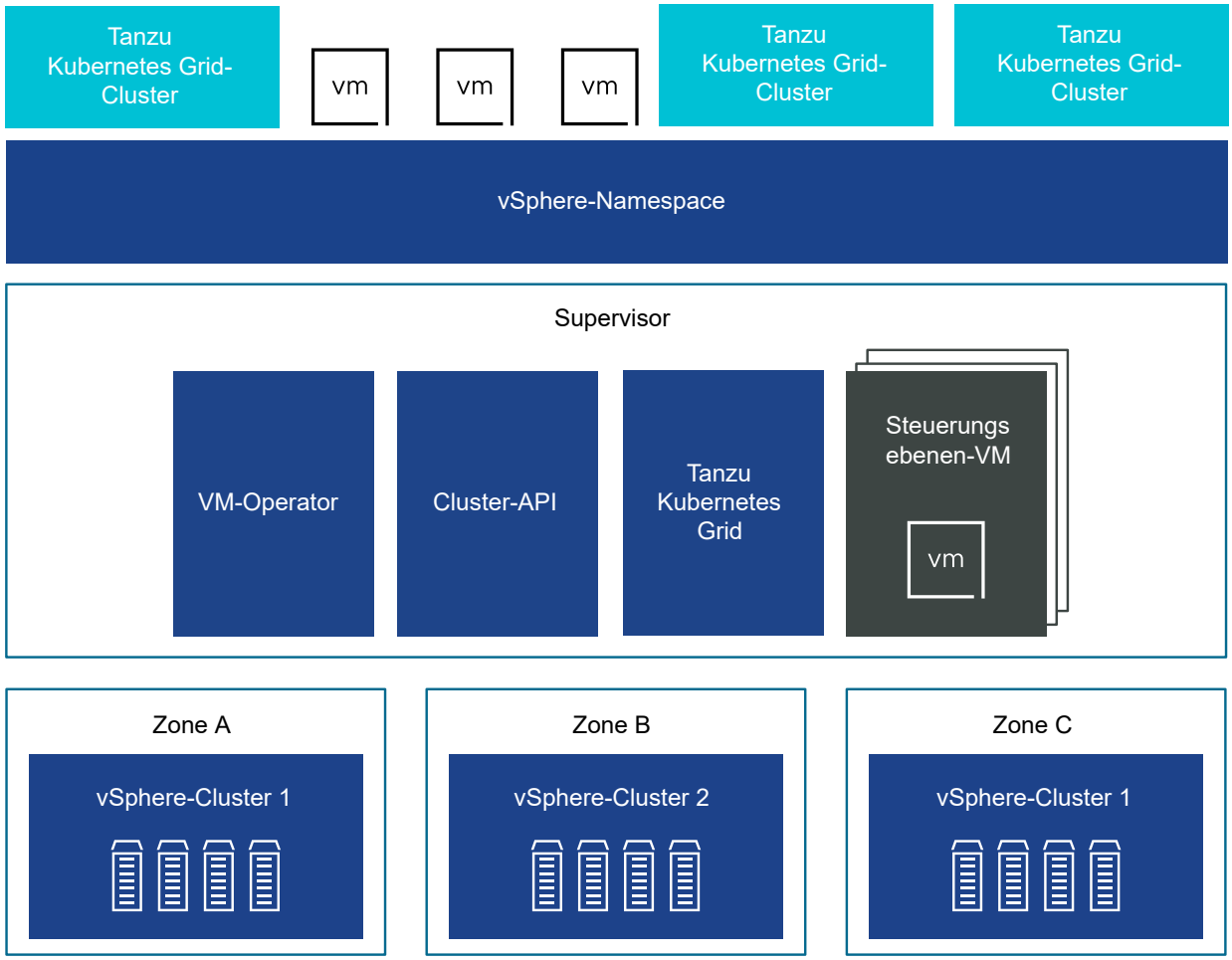
Das Diagramm zeigt die allgemeine vSphere IaaS control plane-Architektur von an – mit dem Tanzu Kubernetes Grid oben, dem Supervisor in der Mitte, dann ESXi, Netzwerk und Speicher im unteren Bereich. Alle Komponenten werden vom vCenter Server verwaltet.



Ein Supervisor wird oberhalb einer SDDC-Schicht ausgeführt, die aus ESXi für Computing, NSX oder VDS-Netzwerke und vSAN oder einer anderen Lösung für freigegebenen Speicher besteht. Freigegebener Speicher wird für persistente Volumes bei vSphere-Pods, für im Supervisor ausgeführte VMs und für Pods in einem Tanzu Kubernetes Grid-Cluster verwendet. Nachdem ein Supervisor erstellt wurde, können Sie als vSphere-Administrator vSphere-Namespaces innerhalb des Supervisor erstellen. Als DevOps-Ingenieur können Sie Arbeitslasten ausführen, die aus Containern bestehen, die innerhalb von vSphere-Pods ausgeführt werden, Vms durch den VM-Dienst bereitstellen und Tanzu Kubernetes Grid-Cluster erstellen.

Sie können einen Supervisor in drei vSphere-Zonen bereitstellen, um Hochverfügbarkeit auf Clusterebene bereitzustellen, die Ihre Kubernetes-Arbeitslasten vor Ausfällen auf Clusterebene schützt. Eine vSphere-Zone wird einem vSphere-Cluster zugeordnet, den Sie als unabhängige Ausfalldomäne einrichten können. Bei einer Bereitstellung mit drei Zonen werden alle drei vSphere-Cluster zu einem Supervisor. Sie haben auch die Möglichkeit, einen Supervisor auf einem vSphere-Cluster bereitzustellen, der automatisch eine vSphere-Zone erstellt und dem Cluster zuweist, sofern Sie nicht einen vSphere-Cluster verwenden, der bereits einer Zone zugeordnet ist. Bei einer Bereitstellung mit einem einzelnen Cluster verfügt der Supervisor nur über Hochverfügbarkeit auf Hostebene, bereitgestellt von vSphere HA.

Abbildung 2-2. Supervisor-Architektur mit drei Zonen

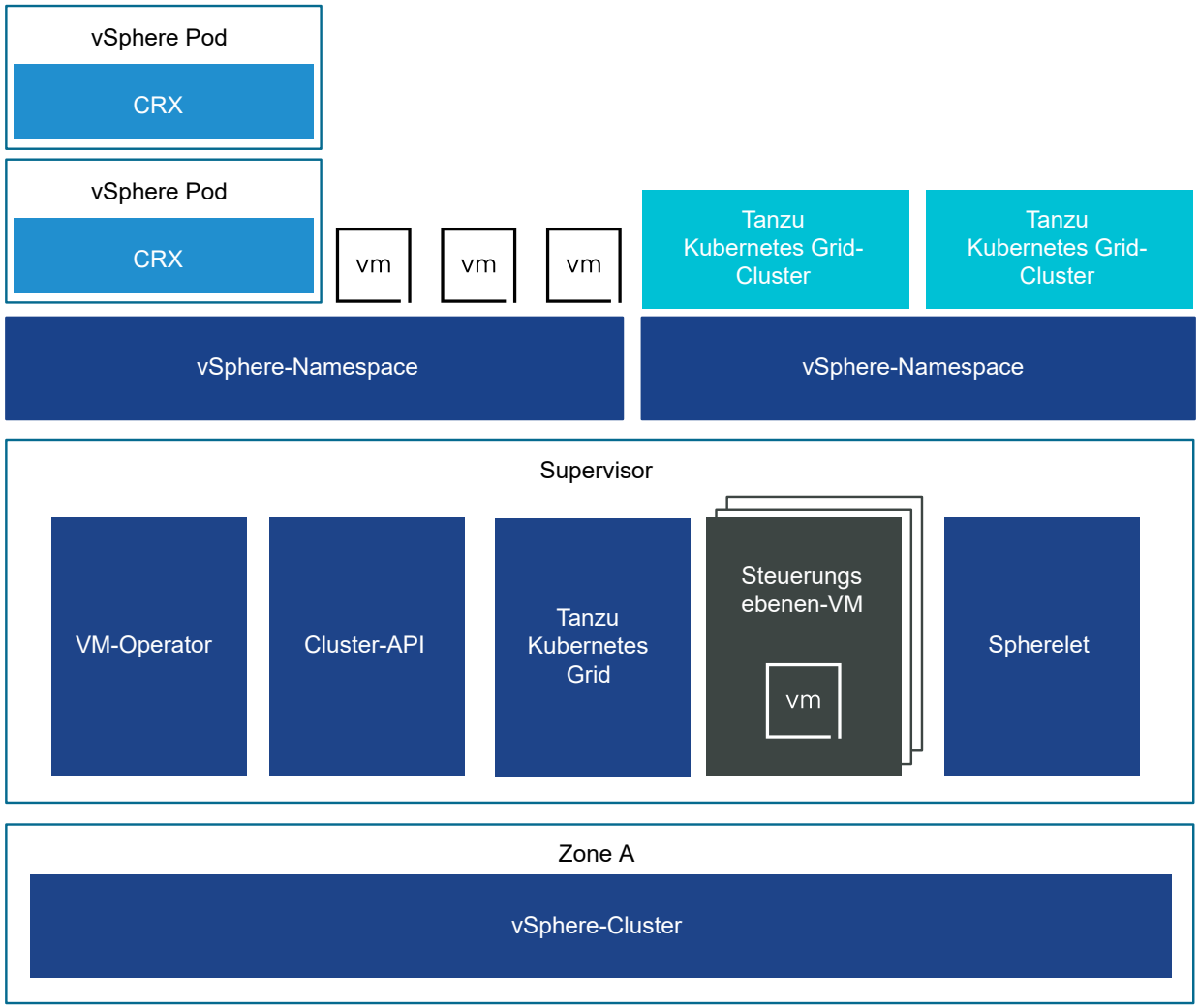


Bei einem Supervisor mit drei Zonen können Sie Kubernetes-Arbeitslasten auf Tanzu Kubernetes Grid-Clustern und -VMs ausführen, die mit dem VM-Dienst erstellt wurden. Ein Supervisor mit drei Zonen hat die folgenden Komponenten:

- Supervisor-Steuerungsebenen-VM. Auf dem Supervisor werden insgesamt drei Supervisor-Steuerungsebenen-VMs erstellt. Bei einer Bereitstellung mit drei Zonen befindet sich in jeder Zone eine Steuerungsebenen-VM. Die drei Supervisor-Steuerungsebenen-VMs verfügen über Lastausgleich, da jede eine eigene IP-Adresse hat. Darüber hinaus wird einer der VMs eine dynamische IP-Adresse zugewiesen, und eine 5. IP-Adresse wird für Patchzwecke reserviert. vSphere DRS bestimmt die exakte Platzierung der Steuerungsebenen-VMs auf den ESXi-Hosts des Supervisors und migriert sie bei Bedarf.
- Tanzu Kubernetes Grid und Cluster-API. Dies sind Module, die im Supervisor ausgeführt werden und die Bereitstellung und Verwaltung von Tanzu Kubernetes Grid-Clustern ermöglichen.
- VM-Dienst. Dieses Modul ist zuständig für die Bereitstellung und Ausführung von eigenständigen VMs sowie von VMs, die die Tanzu Kubernetes Grid-Cluster bilden.

In einem Supervisor mit drei Zonen wird ein Namespace-Ressourcenpool auf jedem vSphere-Cluster erstellt, der einer Zone zugeordnet ist. Der Namespace verteilt sich auf alle drei vSphere-Cluster in jeder Zone. Die für einen Namespace auf einem Supervisor mit drei Zonen verwendeten Ressourcen werden zu gleichen Teilen aus allen drei zugrunde liegenden vSphere-Clustern bezogen. Wenn Sie beispielsweise 300 MHz an CPU zuweisen, kommen von jedem vSphere-Cluster 100 MHz.

Abbildung 2-3. Supervisor-Architektur mit einem einzelnen Cluster



Ein Supervisor, der auf einem einzelnen vSphere-Cluster bereitgestellt wird, verfügt auch über drei Steuerungsebenen-VMs, die sich auf den ESXi-Hosts des Clusters befinden. Auf einem Supervisor mit einem einzelnen Cluster können Sie vSphere-Pods zusätzlich zu Tanzu Kubernetes Grid-Clustern und VMs ausführen. vSphere DRS ist mit dem Kubernetes Scheduler in den Supervisor-Steuerungsebenen-VMs vernetzt, sodass DRS die Platzierung von vSphere-Pods bestimmt. Wenn Sie als DevOps-Ingenieur einen vSphere Pod planen, wird die Anforderung über den regulären Kubernetes-Workflow an DRS übertragen, wodurch die endgültige Platzierungsentscheidung getroffen wird.

Durch die vSphere Pod-Unterstützung verfügt ein Supervisor mit einem einzelnen Cluster über die folgenden zusätzlichen Komponenten:

- Spherelet. Ein zusätzlicher Prozess namens Spherelet wird auf jedem Host erstellt. Es handelt sich um ein Kubelet, das nativ auf ESXi portiert wird und dem ESXi-Host ermöglicht, Teil des Kubernetes-Clusters zu werden.

- CRX-Komponente (Container Runtime Executive). Hinsichtlich Hostd und vCenter Server ist CRX mit einer VM vergleichbar. CRX umfasst einen paravirtualisierten Linux-Kernel, der mit dem Hypervisor zusammenarbeitet. CRX verwendet die gleichen Hardwarevirtualisierungstechniken wie VMs und ist von einer VM-Begrenzung umgeben. Es wird eine Direktstarttechnik verwendet, mit der der Linux-Gast von CRX den Hauptinitialisierungsprozess initiieren kann, ohne die Kernel-Initialisierung durchlaufen zu müssen. Auf diese Weise können vSphere-Pods fast so schnell wie Container starten.

Supervisor-Netzwerk

In einer vSphere IaaS control plane-Umgebung kann ein Supervisor entweder den vSphere-Netzwerk-Stack oder NSX verwenden, um Konnektivität für Supervisor-Steuerungsebenen-VMs, Dienste und Arbeitslasten bereitzustellen.

Wenn ein Supervisor mit dem vSphere-Netzwerk-Stack konfiguriert ist, werden alle Hosts aus dem Supervisor mit einem vDS verbunden, der Arbeitslasten und Supervisor-Steuerungsebenen-VMs Konnektivität bereitstellt. Ein Supervisor, der den vSphere-Netzwerk-Stack verwendet, benötigt einen Lastausgleichsdienst im vCenter Server-Verwaltungsnetzwerk, um DevOps-Benutzern und externen Diensten Konnektivität bereitzustellen.

Ein Supervisor, der mit NSX konfiguriert ist, verwendet die softwarebasierten Netzwerke der Lösung sowie einen NSX Edge-Lastausgleichsdienst oder den NSX Advanced Load Balancer, der externen Diensten und DevOps-Benutzern Konnektivität bereitstellt. Sie können den NSX Advanced Load Balancer auf NSX konfigurieren, wenn Ihre Umgebung die folgenden Bedingungen erfüllt:

- NSX-Version ist 4.1.1 oder höher.
- Die NSX Advanced Load Balancer Version ist 22.1.4 oder höher mit der Enterprise-Lizenz.
- Der NSX Advanced Load Balancer Controller, den Sie konfigurieren möchten, ist auf NSX registriert.
- Ein NSX-Lastausgleichsdienst ist auf dem Supervisor noch nicht konfiguriert.

Supervisor-Netzwerk mit VDS

In einem Supervisor, der von VDS als Netzwerk-Stack gestützt wird, müssen alle Hosts aus den vSphere-Clustern, die den Supervisor unterstützen, mit demselben VDS verbunden sein. Der Supervisor verwendet verteilte Portgruppen als Arbeitslastnetzwerke für Kubernetes-Arbeitslasten und Datenverkehr auf Steuerungsebene. Sie weisen Arbeitslastnetzwerke Namespaces im Supervisor zu.

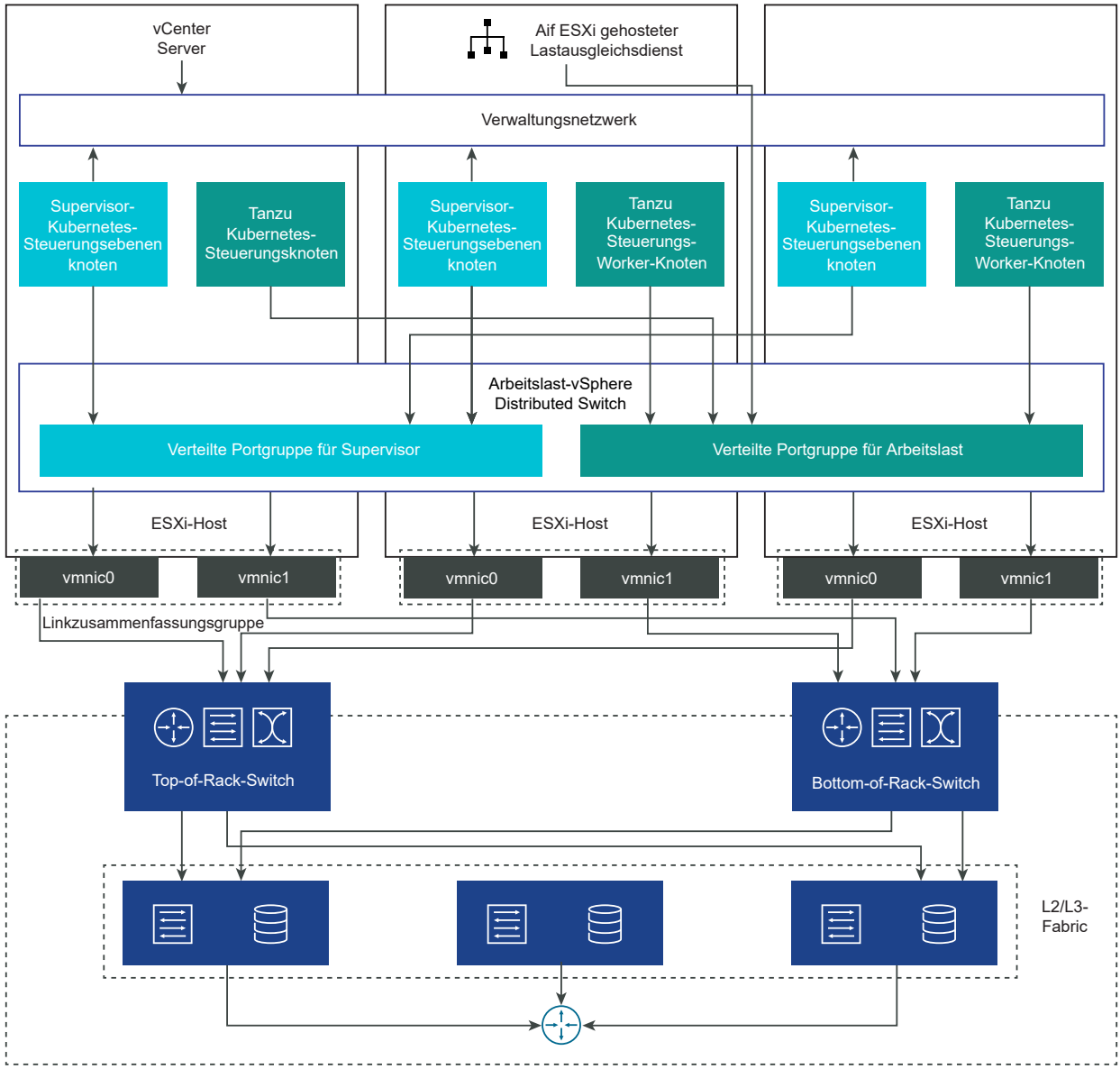
Abhängig von der für den Supervisor implementierten Topologie können Sie eine oder mehrere verteilte Portgruppen als Arbeitslastnetzwerke verwenden. Das Netzwerk, das den Supervisor-Steuerungsebenen-VMs Konnektivität bereitstellt, wird als primäres Arbeitslastnetzwerk bezeichnet. Sie können dieses Netzwerk allen Namespaces auf dem Supervisor zuweisen. Sie können aber auch verschiedene Netzwerke für jeden Namespace verwenden. Die Tanzu Kubernetes Grid-Cluster stellen eine Verbindung zu dem Arbeitslastnetzwerk her, das dem Namespace zugewiesen wird, in dem sich der Cluster befindet.

Ein von einem VDS gestützter Supervisor verwendet einen Lastausgleichsdienst, um DevOps-Benutzern und externen Diensten Konnektivität bereitzustellen. Sie können die -NSX Advanced Load Balancer oder den HAProxy-Lastausgleichsdienst verwenden.

Weitere Informationen finden Sie unter Installieren und Konfigurieren von NSX Advanced Load Balancer und Installieren und Konfigurieren von HAProxy-Lastausgleichsdienst.

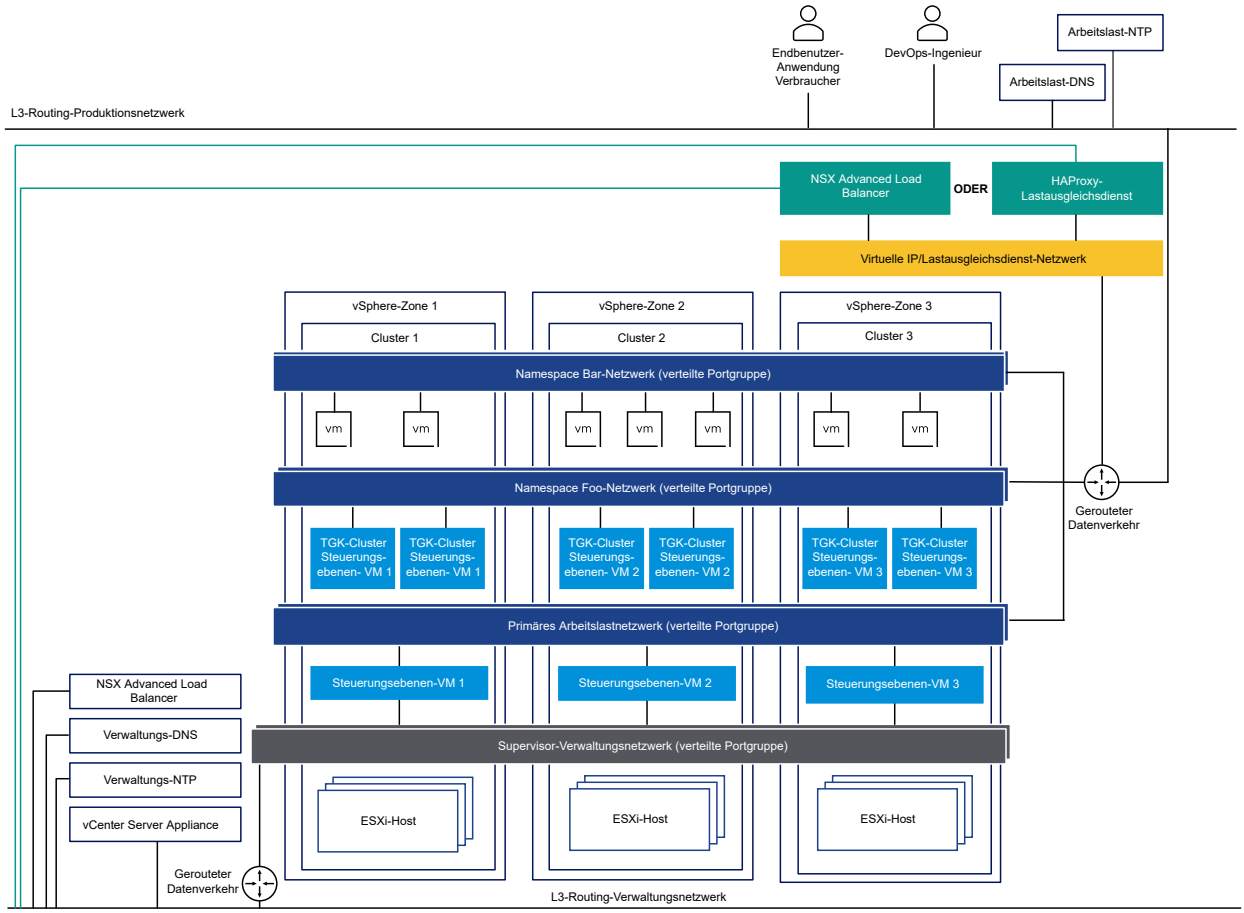
In einer Supervisor-Einrichtung mit einem Cluster wird Supervisor nur von einem vSphere-Cluster gestützt. Alle Hosts aus dem Cluster müssen mit einem VDS verbunden sein.

Abbildung 2-4. Supervisor-Netzwerk mit einem Cluster und VDS



In einem Supervisor für drei Zonen stellen Sie Supervisor auf drei vSphere-Zonen bereit, die jeweils einem vSphere-Cluster zugeordnet sind. Alle Hosts aus diesen vSphere-Clustern müssen mit demselben VDS verbunden sein. Alle physischen Server müssen mit einem L2-Gerät verbunden sein. Arbeitslastnetzwerke, die Sie für den Namespace konfigurieren, erstrecken sich über alle drei vSphere-Zonen.

Abbildung 2-5. Supervisor-Netzwerk für drei Zonen und VDS



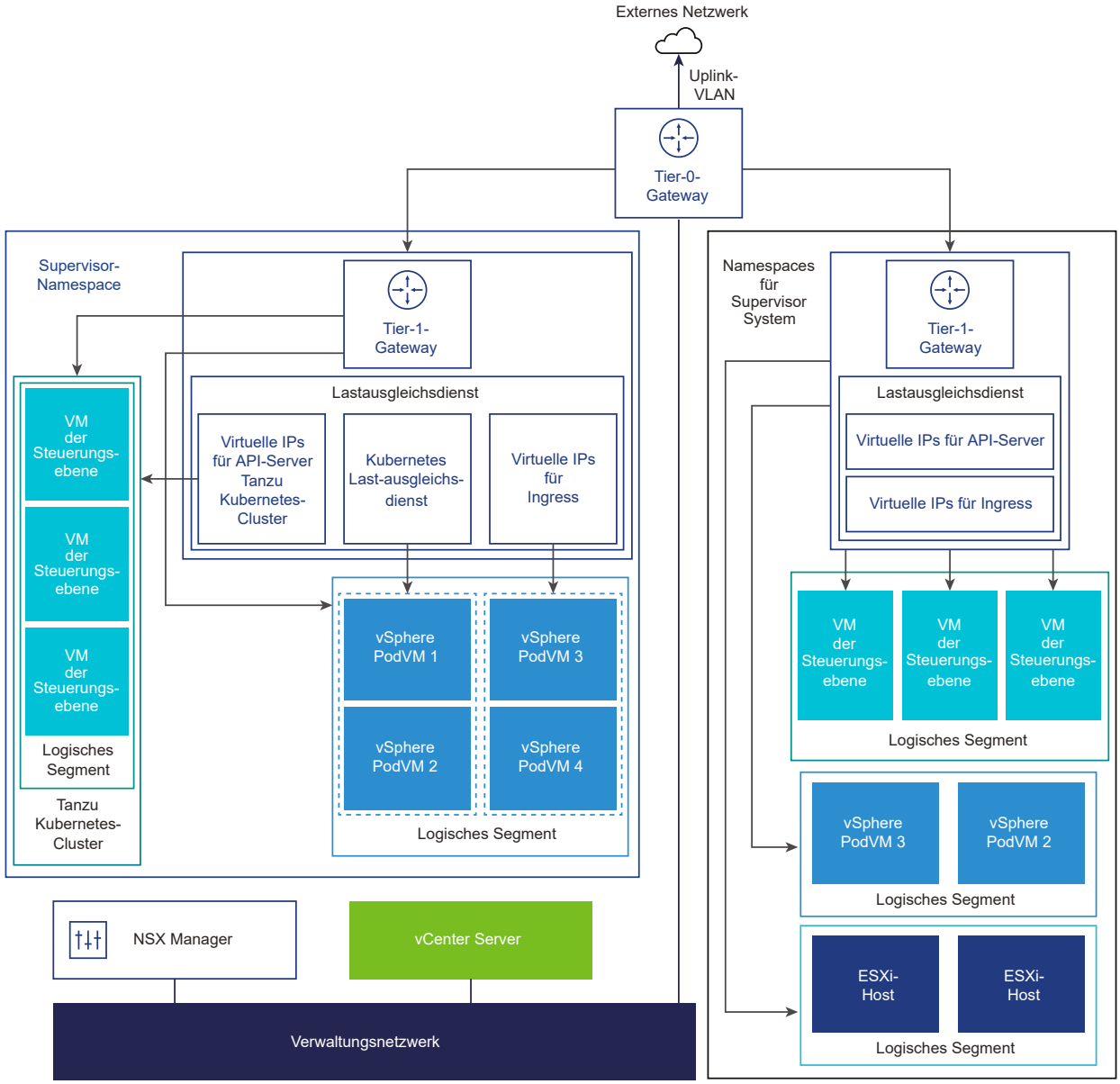
Supervisor-Netzwerk mit NSX

NSX bietet Netzwerkkonnektivität zu den Objekten innerhalb des Supervisors und externen Netzwerken. Die Konnektivität mit den ESXi-Hosts, die den Cluster umfassen, wird von den standardmäßigen vSphere-Netzwerken verarbeitet.

Sie können das Supervisor-Netzwerk auch manuell konfigurieren, indem Sie eine vorhandene NSX-Bereitstellung verwenden oder eine neue NSX-Instanz bereitstellen.

Weitere Informationen finden Sie unter Installieren und Konfigurieren von NSX for vSphere IaaS control plane.

Abbildung 2-6. Supervisor-Netzwerk mit NSX



- NSX Container Plugin (NCP) bietet Integration zwischen NSX und Kubernetes. Die Hauptkomponente von NCP wird in einem Container ausgeführt und kommuniziert mit NSX Manager und mit der Kubernetes-Steuerungsebene. NCP überwacht Änderungen an Containern und anderen Ressourcen und verwaltet Netzwerkressourcen wie logische Ports, Segmente, Router und Sicherheitsgruppen für die Container durch Aufrufen der NSX API.

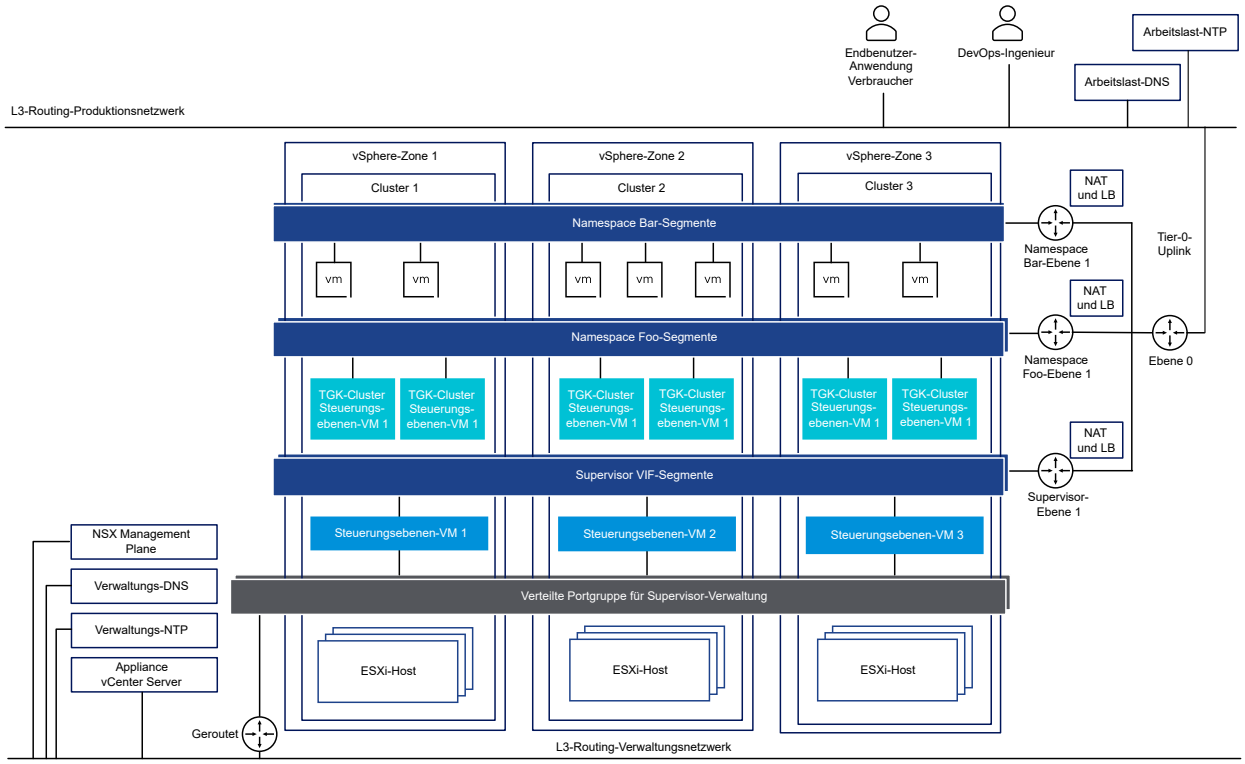
  Das NCP erstellt standardmäßig ein freigegebenes Tier-1-Gateway für System-Namespaces und ein Tier-1-Gateway und einen Lastausgleichsdienst für jeden Namespace. Das Tier-1-Gateway ist mit dem Tier-0-Gateway und einem Standardsegment verbunden.

System-Namespaces sind Namespaces, die von den Kernkomponenten verwendet werden, die für das Funktionieren des Supervisors und der Tanzu Kubernetes Grid-Cluster von wesentlicher Bedeutung sind. Die freigegebenen Netzwerkressourcen, die das Tier-1-Gateway, den Lastausgleichsdienst und die SNAT-IP enthalten, sind in einem System-Namespace gruppiert.

- NSX Edge bietet Konnektivität von externen Netzwerken zu Supervisor-Objekten. Der NSX Edge-Cluster verfügt über einen Lastausgleichsdienst, der Redundanz für die Kubernetes-API-Server auf den Supervisor-Steuerungsebenen-VMs sowie für alle Anwendungen bietet, die veröffentlicht werden müssen und auf die ein Zugriff von außerhalb des Supervisors möglich sein muss.
- Ein Tier-0-Gateway ist mit dem NSX Edge-Cluster verknüpft, um das Routing zum externen Netzwerk bereitzustellen. Die Uplink-Schnittstelle verwendet entweder das dynamische Routing-Protokoll, BGP oder statisches Routing.
- Jeder vSphere-Namespace verfügt über ein separates Netzwerk und eine Reihe von Netzwerkressourcen, die von Anwendungen innerhalb des Namespace gemeinsam genutzt werden, wie z. B. das Tier-1-Gateway, den Lastausgleichsdienst und die SNAT-IP-Adresse.
- Arbeitslasten, die in vSphere-Pods, regulären VMs oder Tanzu Kubernetes Grid-Clustern ausgeführt werden, die sich im selben Namespace befinden, nutzen dieselbe SNAT-IP für die Nord-Süd-Konnektivität.
- Arbeitslasten, die in vSphere-Pods oder Tanzu Kubernetes Grid-Clustern ausgeführt werden, haben dieselbe Isolationsregel, die von der Standard-Firewall implementiert wird.
- Eine eigene SNAT-IP ist nicht für jeden Kubernetes-Namespace erforderlich. Die Ost-West-Konnektivität zwischen Namespaces ist kein SNAT.
- Die Segmente für die einzelnen Namespaces befinden sich auf dem im Standardmodus funktionierenden VDS, der dem NSX Edge-Cluster zugeordnet ist. Das Segment stellt dem Supervisor ein Overlay-Netzwerk zur Verfügung.
- Supervisoren verfügen über getrennte Segmente innerhalb des gemeinsam genutzten Tier-1-Gateways. Für jeden Tanzu Kubernetes Grid-Cluster werden Segmente innerhalb des Tier-1-Gateways des Namespace definiert.
- Die Spherelet-Prozesse auf den einzelnen ESXi-Hosts kommunizieren mit vCenter Server über eine Schnittstelle im Verwaltungsnetzwerk.

In einer Supervisor-Konfiguration für drei Zonen mit NSX als Netzwerk-Stack müssen alle Hosts aus allen drei vSphere-Clustern, die den Zonen zugeordnet sind, mit demselben VDS verbunden sein und an derselben NSX-Overlay-Transportzone teilnehmen. Alle Hosts müssen mit demselben physischen L2-Gerät verbunden sein.

Abbildung 2-7. Supervisor-Netzwerk für drei Zonen mit NSX



Supervisor-Netzwerk mit NSX und NSX Advanced Load Balancer

NSX bietet Netzwerkkonnektivität zu den Objekten innerhalb des Supervisors und externen Netzwerken. Ein mit NSX konfigurierter Supervisor kann den NSX Edge oder den NSX Advanced Load Balancer verwenden.

Zu den Komponenten des NSX Advanced Load Balancer gehören der NSX Advanced Load Balancer Controller-Cluster, Dienst-Engines (Datenebenen-) VMs und der Avi Kubernetes Operator (AKO).

Der NSX Advanced Load Balancer Controller interagiert mit dem vCenter Server, um den Lastausgleich für die Tanzu Kubernetes Grid-Cluster zu automatisieren. Er ist für die Bereitstellung von Dienst-Engines, die Koordination von Ressourcen anhand von Dienst-Engines sowie die Zusammenfassung von Dienst-Engine-Metriken und -Protokollen zuständig. Der Controller bietet eine Web-Schnittstelle, Befehlszeilenschnittstelle und API für den Benutzerbetrieb und die programmgesteuerte Integration. Nachdem Sie die Controller-VM bereitgestellt und konfiguriert haben, können Sie einen Controller-Cluster bereitstellen, um den Steuerungsebenen-Cluster für HA einzurichten.

Die Dienst-Engine ist die virtuelle Maschine der Datenebene. Eine Dienst-Engine führt einen oder mehrere virtuelle Dienste aus. Eine Dienst-Engine wird vom NSX Advanced Load Balancer Controller verwaltet. Der Controller stellt Dienst-Engines für das Hosten virtueller Dienste zur Verfügung.

Die Dienst-Engines verfügen über zwei Arten von Netzwerkschnittstellen:

- Die erste Netzwerkschnittstelle, `vnic0` der VM, wird mit dem Verwaltungsnetzwerk verbunden, wo sie eine Verbindung zum NSX Advanced Load Balancer Controller herstellen kann.
- Die restlichen Schnittstellen, `vnic1 – 8`, verbinden sich mit dem Datennetzwerk, in dem virtuelle Dienste ausgeführt werden.

Die Dienst-Engine-Schnittstellen stellen automatisch eine Verbindung mit den richtigen vDS-Portgruppen her. Jede Dienst-Engine kann bis zu 1.000 virtuelle Dienste unterstützen.

Ein virtueller Dienst stellt Ebene-4- und Ebene-7-Lastausgleichsdienste für Tanzu Kubernetes Grid-Clusterarbeitslasten zur Verfügung. Ein virtueller Dienst ist mit einer virtuellen IP und mehreren Ports konfiguriert. Wenn ein virtueller Dienst bereitgestellt wird, wählt der Controller automatisch einen ESX-Server aus, startet eine Dienst-Engine und verbindet sie mit den richtigen Netzwerken (Portgruppen).

Die erste Dienst-Engine wird erst erstellt, nachdem der erste virtuelle Dienst konfiguriert wurde. Alle nachfolgenden virtuellen Dienste, die konfiguriert werden, verwenden die vorhandene Dienst-Engine.

Jeder virtuelle Server macht einen Load Balancer der Ebene 4 mit einer eindeutigen IP-Adresse des Typs Load Balancer für einen Tanzu Kubernetes Grid verfügbar. Die IP-Adresse, die jedem virtuellen Server zugewiesen ist, wird aus dem IP-Adressblock ausgewählt, der dem Controller bei der Konfiguration zugewiesen wurde.

Der Avi-Kubernetes-Operator (AKO) überwacht Kubernetes-Ressourcen und kommuniziert mit dem NSX Advanced Load Balancer Controller, um die entsprechenden Lastausgleichsressourcen anzufordern. Der Avi-Kubernetes-Operator wird im Rahmen des Aktivierungsprozesses auf den Supervisoren installiert.

Weitere Informationen finden Sie unter Installieren und Konfigurieren von NSX und NSX Advanced Load Balancer.

Abbildung 2-8. Supervisor-Netzwerk mit NSX und NSX Advanced Load Balancer Controller

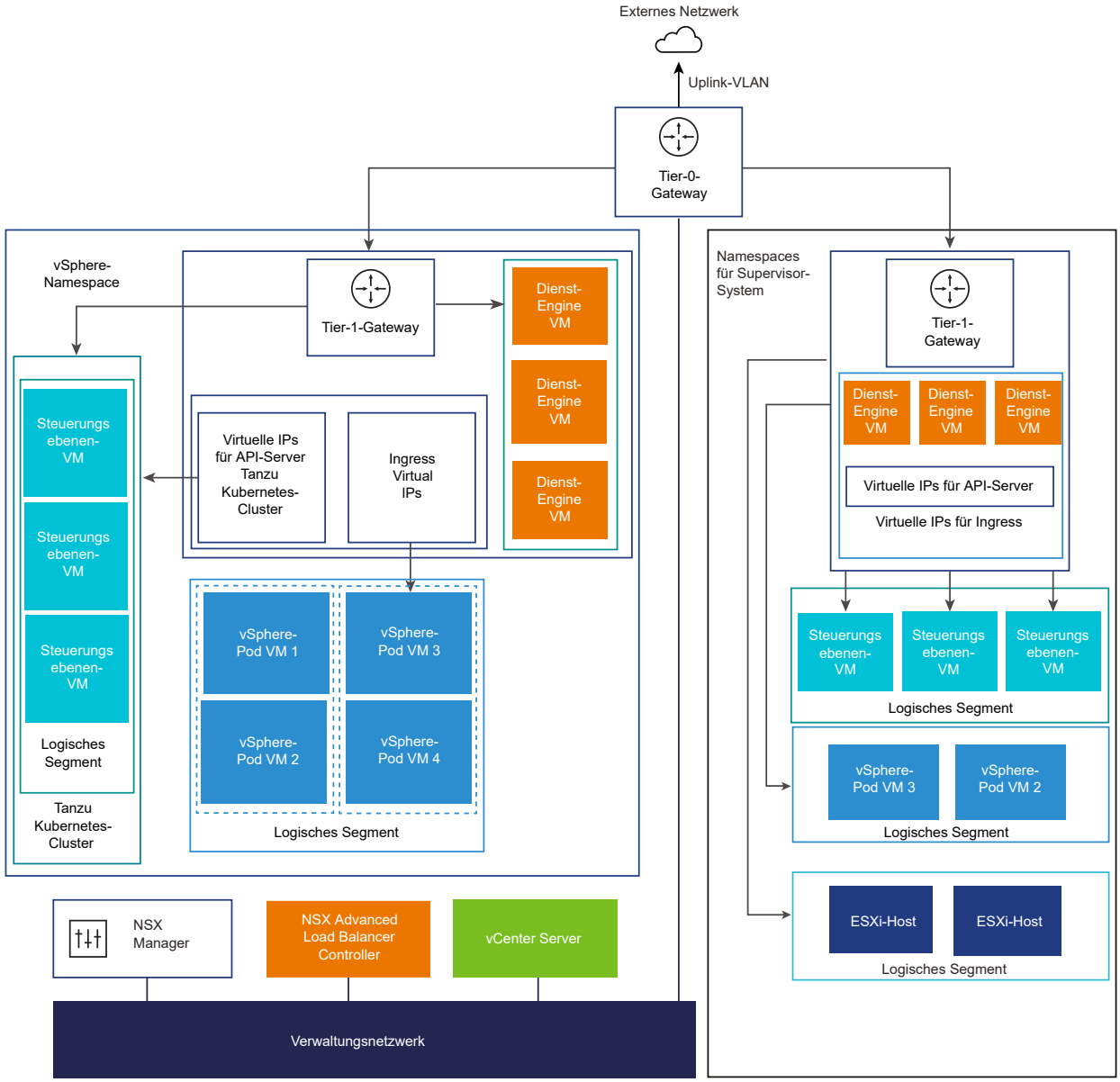
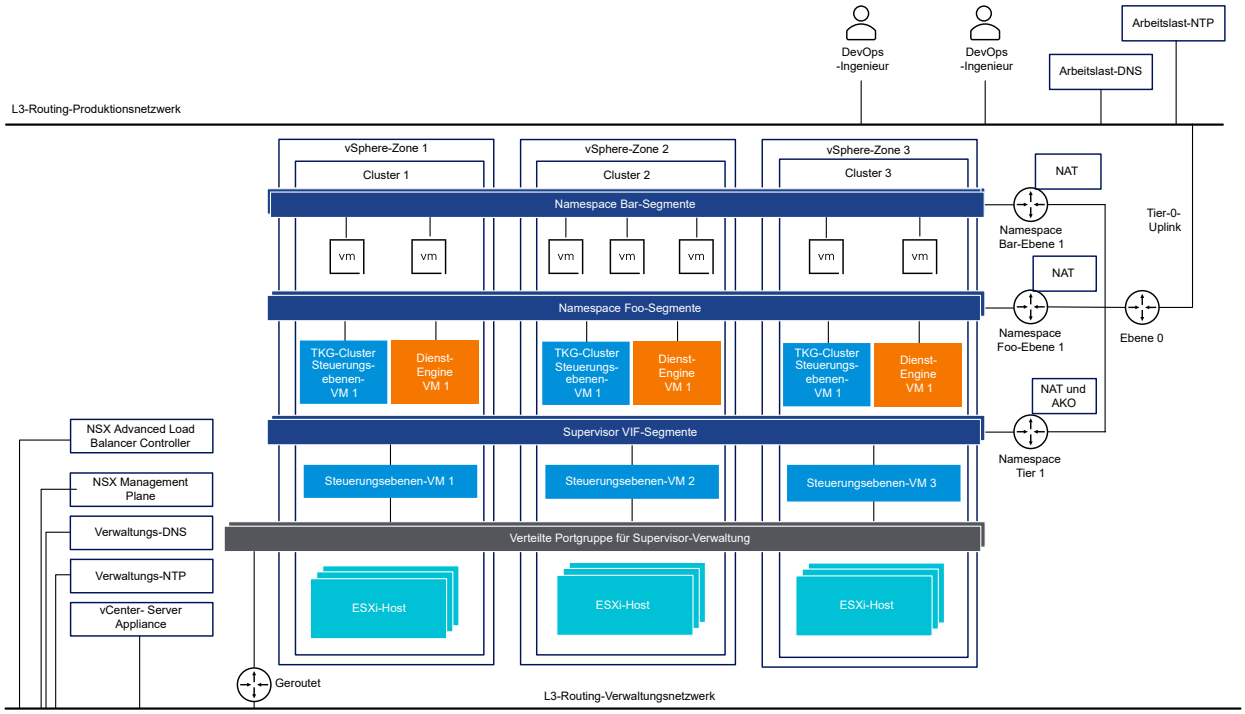


Abbildung 2-9. Supervisor-Netzwerk für drei Zonen mit NSX und NSX Advanced Load Balancer Controller



Wichtig Wenn Sie den NSX Advanced Load Balancer Controller in einer NSX-Bereitstellung konfigurieren, sollten Sie Folgendes berücksichtigen:

- Sie können den NSX Advanced Load Balancer Controller nicht in einer vCenter Server-Bereitstellung im erweiterten verknüpften Modus bereitstellen. Sie können den NSX Advanced Load Balancer Controller nur in einer Bereitstellung mit einem einzelnen vCenter Server bereitstellen. Wenn mehrere vCenter Server verknüpft sind, kann nur einer davon beim Konfigurieren des NSX Advanced Load Balancer Controller verwendet werden.
- Sie können den NSX Advanced Load Balancer Controller nicht in einer mehrschichtigen Tier-0-Topologie konfigurieren. Wenn die NSX-Umgebung mit einer Tier-0-Topologie mit mehreren Ebenen eingerichtet ist, können Sie beim Konfigurieren des NSX Advanced Load Balancer Controller nur ein Tier-0-Gateway verwenden.

Netzwerkkonfigurationsmethoden mit NSX

Supervisor verwendet eine Opinionated-Netzwerkkonfiguration. Zum Konfigurieren des Supervisor-Netzwerks mit NSX gibt es zwei Methoden, die zur Bereitstellung desselben Netzwerkmodells für einen Supervisor für eine Zone führen:

- Die einfachste Möglichkeit zur Konfiguration des Supervisor-Netzwerks besteht darin, VMware Cloud Foundation SDDC Manager zu verwenden. Weitere Informationen finden Sie in der Dokumentation zu VMware Cloud Foundation SDDC Manager. Weitere Informationen finden Sie im Administratorhandbuch für VMware Cloud Foundation.

- Sie können das Supervisor-Netzwerk auch manuell konfigurieren, indem Sie eine vorhandene NSX-Bereitstellung verwenden oder eine neue NSX-Instanz bereitstellen. Weitere Informationen finden Sie unter Installieren und Konfigurieren von NSX for vSphere IaaS control plane.

Supervisor-Speicher

Supervisor-Komponenten, -Anwendungen und -Arbeitslasten müssen Daten speichern und abrufen. Einige Anwendungen und Objekte verwenden möglicherweise schnellen flüchtigen Speicher, während für andere persistenter Speicher erforderlich ist.

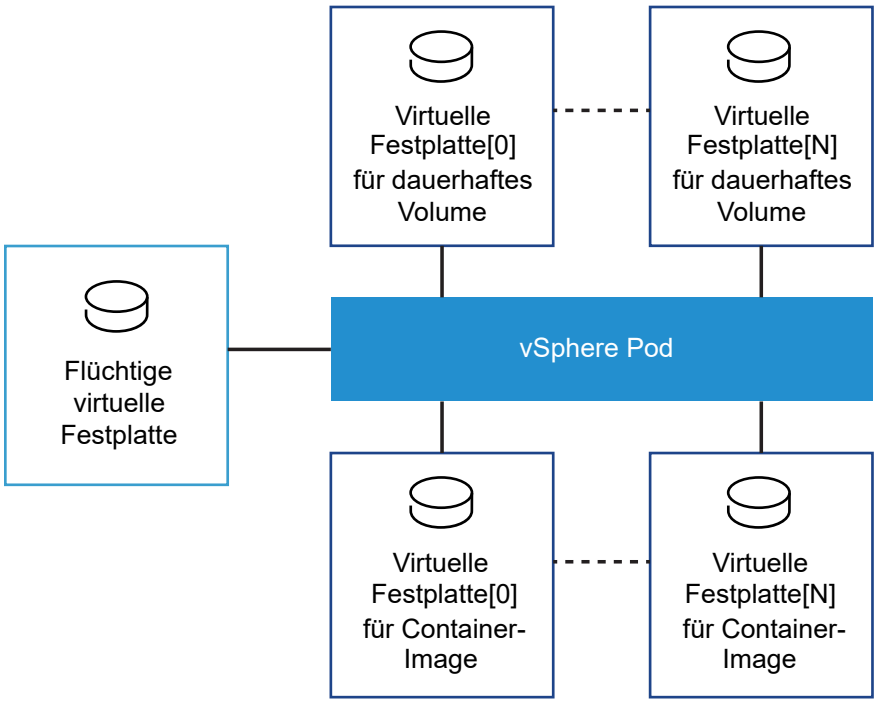
Grundlegendes zu Speicherrichtlinien

Supervisor verwendet Speicherrichtlinien zur Integration in verfügbaren Speicher in Ihrer vSphere-Umgebung. Die Richtlinien stellen Datenspeicher dar und steuern die Speicherplatzierung von Komponenten und Objekten wie Steuerungsebenen-VMs, flüchtigen vSphere Pod-Festplatten und Container-Images. Möglicherweise benötigen Sie auch Richtlinien für die Speicherplatzierung dauerhafter Volumes und VM-Inhaltsbibliotheken. Falls Sie Tanzu Kubernetes Grid-Cluster verwenden, bestimmen die Speicherrichtlinien auch, wie die Tanzu Kubernetes Grid-Clusterknoten bereitgestellt werden.

Speicherrichtlinien unterstützen alle gemeinsam genutzten Datenspeicher in Ihrer Umgebung, wie etwa VMFS, NFS und vSAN, einschließlich vSAN ESA oder vVols.

Je nach vSphere-Speicherumgebung und den Anforderungen von DevOps können Sie mehrere Speicherrichtlinien für verschiedene Speicherklassen erstellen. Wenn Sie einen Supervisor aktivieren und Namespaces einrichten, können Sie verschiedene Speicherrichtlinien zuweisen, die von verschiedenen Objekten, Komponenten und Arbeitslasten verwendet werden sollen.

Wenn beispielsweise ein vSphere Pod drei Arten von virtuellen Festplatten mountet und Ihre vSphere-Speicherumgebung über drei Klassen von Datenspeichern (Bronze, Silber und Gold) verfügt, können Sie Speicherrichtlinien für alle Datenspeicher erstellen. Sie können den Bronze-Datenspeicher dann für flüchtige virtuelle Festplatten und virtuelle Festplatten mit Container-Images und die Silber- und Gold-Datenspeicher für virtuelle Festplatten mit dauerhaften Volumes verwenden.



Informationen zum Erstellen von Speicherrichtlinien finden Sie in Dokumentation zum Erstellen von Speicherrichtlinien*Installieren und Konfigurieren der vSphere IaaS-Steuerungsebene*.

Allgemeine Informationen zu Speicherrichtlinien finden Sie in Kapitel Speicherrichtlinienbasierte Verwaltung in der *vSphere-Speicher*-Dokumentation.

Speicherrichtlinien für einen Supervisor

Auf Supervisor-Ebene konfigurieren Sie eine Speicherrichtlinie für Supervisor-Steuerungsebenen-VMs. Wenn Ihre Bereitstellung vSphere-Pods unterstützt, weisen Sie außerdem Speicherrichtlinien zu und geben die Datenspeicherorte für flüchtige Festplatten und Container-Images an. Informationen zum Einstellen des Speichers beim Aktivieren von Supervisor finden Sie in der Dokumentation *Installieren und Konfigurieren der vSphere IaaS-Steuerungsebene*. Informationen zum Ändern der Speichereinstellungen finden Sie unter Ändern der Speichereinstellungen im Supervisor.

Speicherrichtlinie für Steuerungsebene

Mit dieser Richtlinie wird dafür gesorgt, dass die VMs der Steuerungsebene auf den Datenspeichern platziert werden, die von den Richtlinien dargestellt werden.

Flüchtige virtuelle Festplatten

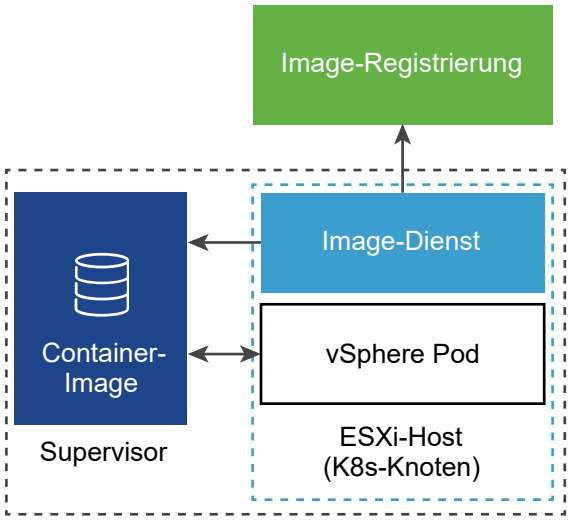
Eine vSphere Pod erfordert flüchtigen Speicher, um Kubernetes-Objekte wie Protokolle, `emptyDir`-Volumes und `ConfigMaps` während der Vorgänge zu speichern. Dieser flüchtige (oder transiente) Speicher steht so lange zur Verfügung, wie der Pod vorhanden ist. Flüchtige Daten werden über Container-Neustarts hinweg beibehalten, aber sobald der Pod das Ende seiner Lebensdauer erreicht, wird die flüchtige virtuelle Festplatte nicht mehr angezeigt.

Jeder Pod verfügt über eine flüchtige virtuelle Festplatte. Als vSphere-Administrator verwenden Sie beim Konfigurieren des Speichers für den Supervisor eine Speicherrichtlinie, um den Speicherort des Datenspeichers für alle flüchtigen virtuellen Festplatten festzulegen.

Virtuelle Festplatten des Container-Images

Container innerhalb der vSphere Pod verwenden Images, die die auszuführende Software enthalten. Der Pod mountet Images, die von ihren Containern als virtuelle Image-Festplatten verwendet werden. Wenn der Pod seinen Lebenszyklus abgeschlossen hat, werden die virtuellen Image-Festplatten vom Pod getrennt.

Image-Dienst, eine ESXi-Komponente, ist dafür verantwortlich, Container-Images aus der Image-Registrierung abzurufen und sie in virtuelle Festplatten zu transformieren, um sie innerhalb des Pods auszuführen.



ESXi kann Images zwischenspeichern, die für die im Pod ausgeführten Container heruntergeladen werden. Nachfolgende Pods, die dasselbe Image verwenden, rufen dieses aus dem lokalen Cache und nicht aus der externen Container-Registrierung ab.

Persistenter Speicher für Arbeitslasten

Bestimmte Kubernetes-Arbeitslasten, die DevOps in einem Namespace ausführen, benötigen persistenten Speicher, um Daten dauerhaft zu speichern.

Persistenter Speicher kann von vSphere-Pods, Tanzu Kubernetes Grid-Clustern, VMs und anderen Arbeitslasten genutzt werden, die Sie im Namespace ausführen. Um dem DevOps-Team persistenten Speicher zur Verfügung zu stellen, erstellt der vSphere-Administrator Speicherrichtlinien, die verschiedene Speicheranforderungen und Dienstklassen beschreiben. Anschließend weist der Administrator Speicherrichtlinien zu und konfiguriert Speichergrenzwerte auf Namespace-Ebene.

Machen Sie sich mit den wichtigen Kubernetes-Konzepten wie Speicherklassen, persistenten Volumes und der Anforderungen von persistenten Volumes vertraut, um zu verstehen, wie vSphere IaaS control plane mit persistentem Speicher arbeitet. Weitere Informationen dazu finden Sie in der Kubernetes-Dokumentation unter https://kubernetes.io/docs/home/.

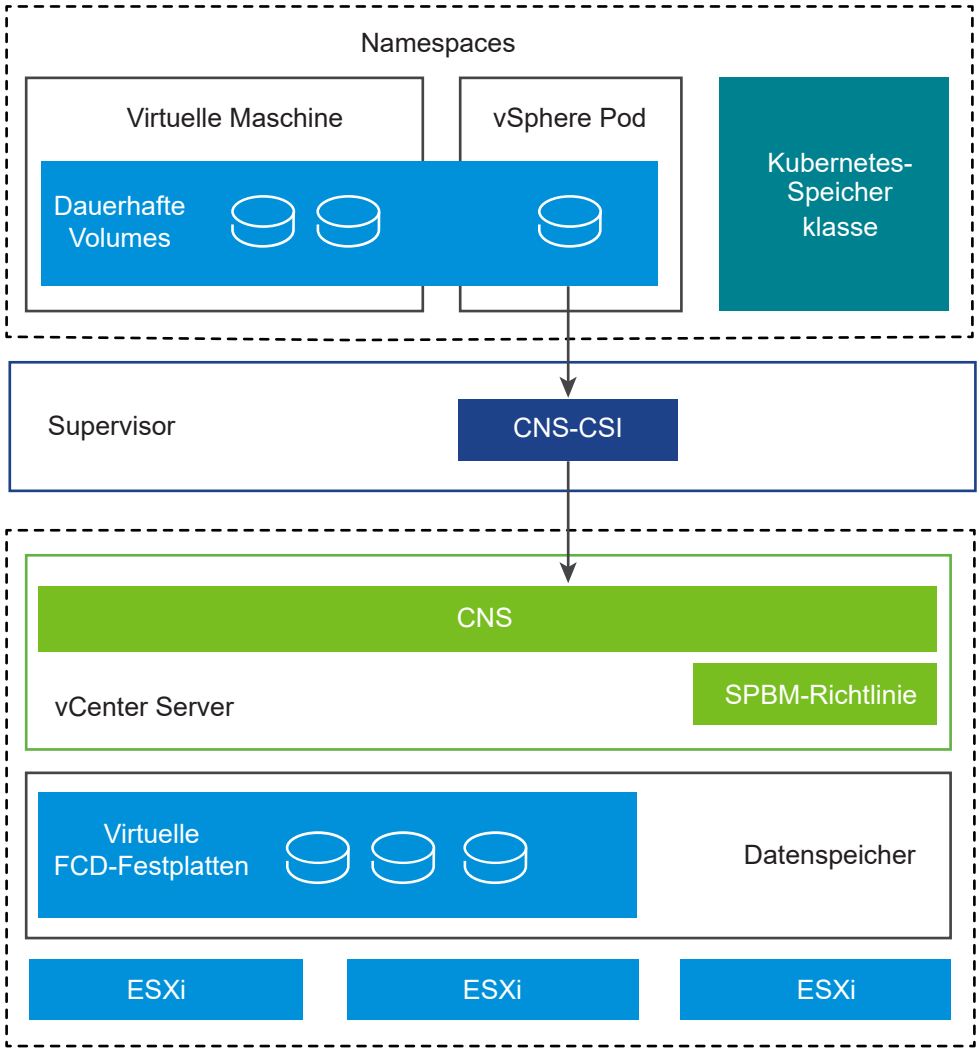
Informationen zum persistenten Speicher für Tanzu Kubernetes Grid-Cluster finden Sie unter Speicher für Tanzu Kubernetes Grid-Cluster.

Informationen zur Verwendung von persistentem Speicher finden Sie im Abschnitt zum Verwenden von persistentem Speicher bei Arbeitslasten in der Dokumentation *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

Plant Ihr DevOps-Team die Bereitstellung von Drittanbieterdiensten, die vSAN Direct für ihren Bedarf an persistentem Speicher verwenden? Weitere Informationen dazu finden Sie im Abschnitt zum Aktivieren von statusbehafteten Diensten in vSphere with Tanzu in der Dokumentation *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

Wie ist Supervisor in vSphere-Speicher integriert?

Supervisor verwendet mehrere Komponenten für die Integration in vSphere Storage.



Cloudnativer Speicher (CNS) in vCenter Server

Die CNS-Komponente befindet sich in vCenter Server. Es handelt sich um eine Erweiterung der vCenter Server-Verwaltung, die die Bereitstellungs- und Lebenszyklusvorgänge für dauerhafte Volumes implementiert.

Bei der Bereitstellung von dauerhaften Volumes interagiert die Komponente mit der First-Class Disks-Funktionalität (erstklassige Festplatten) von vSphere, um virtuelle Festplatten zur Unterstützung der Volumes zu erstellen. Darüber hinaus kommuniziert die CNS-Serverkomponente mit der speicherrichtlinienbasierten Verwaltung, um den Festplatten die erforderliche Befehlseben zu garantieren.

Die CNS-Komponente führt auch Abfragevorgänge aus, mit denen vSphere-Administratoren dauerhafte Volumes und deren unterstützende Speicherobjekte über vCenter Server verwalten und überwachen können.

Erstklassige Festplatte (First Class Disk, FCD)

Wird auch als verbesserte virtuelle Festplatte bezeichnet. Diese Festplatten befinden sich auf Datenspeichern und unterstützen dauerhafte ReadWriteOnce-Volumes.

Beachten Sie bei der Verwendung von FCDs Folgendes:

- FCDs unterstützen NFS 4.x-Protokolle nicht. Verwenden Sie stattdessen NFS 3.
- vCenter Server serialisiert keine Vorgänge auf demselben FCD. Dies führt dazu, dass Anwendungen nicht gleichzeitig Vorgänge auf derselben FCD ausführen können. Das gleichzeitige Klonen, Verlagern, Löschen, Abrufen und so weiter über verschiedene Threads führt zu unvorhersehbaren Ergebnissen. Um Probleme zu vermeiden, müssen Anwendungen Vorgänge auf derselben FCD in sequenzieller Reihenfolge ausführen.
- FCD ist kein verwaltetes Objekt und bietet keine Unterstützung für eine globale Sperre, die mehrere Schreibvorgänge für eine einzelne FCD schützt. Dies führt dazu, dass FCD nicht mehrere vCenter Server-Instanzen unterstützt, die dieselbe FCD verwalten. Wenn Sie mehrere vCenter Server-Instanzen mit FCDs verwenden müssen, sind folgenden Optionen verfügbar:
  - Mehrere vCenter Server-Instanzen können verschiedene Datenspeicher verwalten.
  - Mehrere vCenter Server-Instanzen funktionieren nicht auf derselben FCD.

Speicherrichtlinienbasierte Verwaltung

Die speicherrichtlinienbasierte Verwaltung ist ein vCenter Server-Dienst, der die Bereitstellung von dauerhaften Volumes und deren unterstützenden virtuellen Festplatten gemäß den in einer Speicherrichtlinie beschriebenen Speicheranforderungen unterstützt. Nach der Bereitstellung überwacht der Dienst die Konformität des Volumes mit den Merkmalen der Speicherrichtlinien. Weitere Informationen zur Verwaltung auf der Basis von Speicherrichtlinien finden Sie im Kapitel Speicherrichtlinienbasierte Verwaltung in der Dokumentation zu *vSphere-Speicher*.

vSphere CNS-CSI

Die vSphere CNS-CSI-Komponente entspricht der CSI-Spezifikation (Container Storage Interface), einem Branchenstandard, durch den eine Schnittstelle für die Containerorchestrierung zur Verfügung gestellt wird, die u. a. von Kubernetes zur Bereitstellung von dauerhaftem Speicher genutzt wird. Der CNS-CSI-Treiber wird im Supervisor ausgeführt und stellt die Verbindung zwischen vSphere-Speicher und der Kubernetes-Umgebung in einem Namespace her. Der vSphere CNS-CSI kommuniziert direkt mit der CNS-Komponente für alle Speicherbereitstellungsanforderungen, die aus dem Namespace stammen.

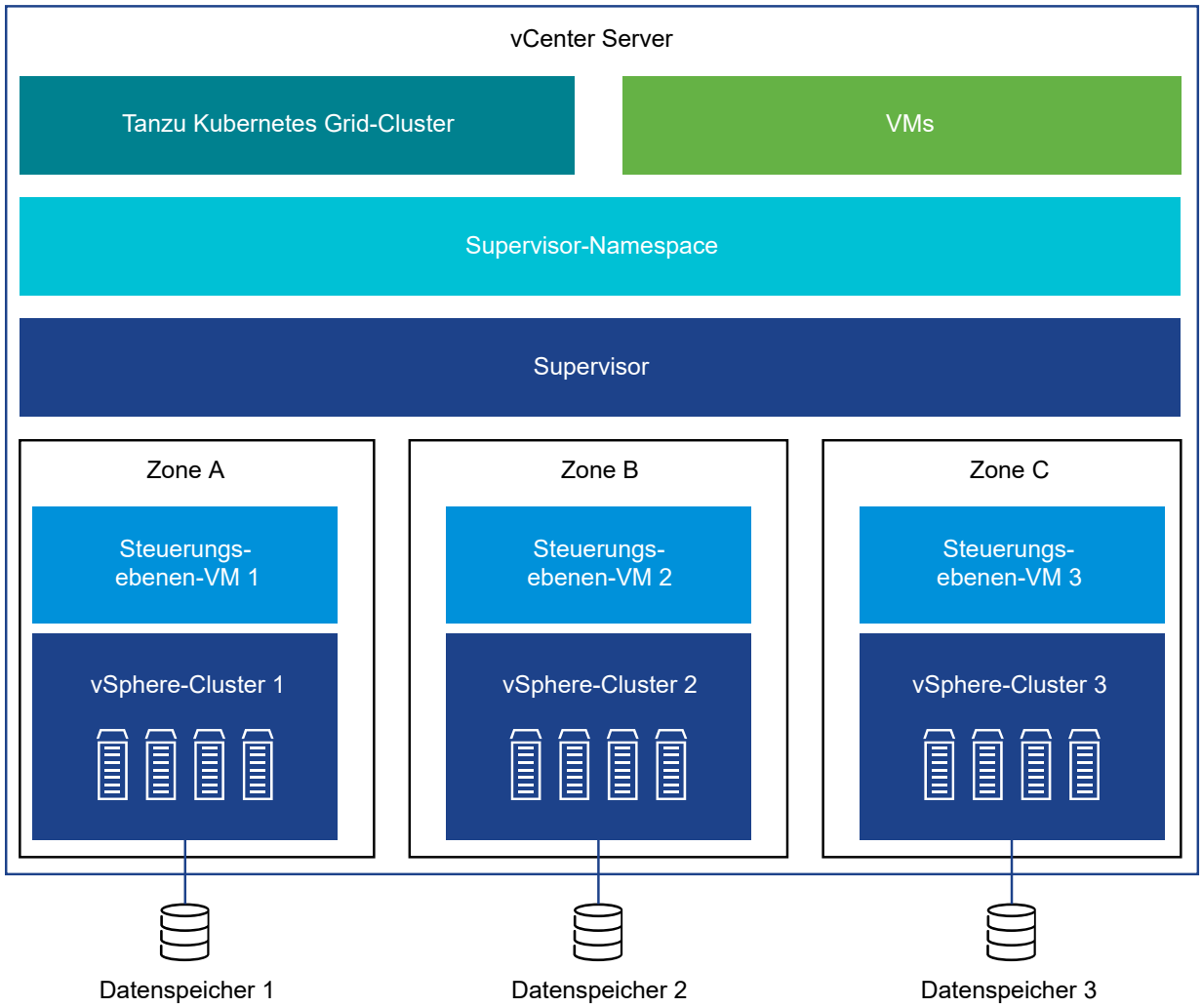
Von vSphere CNS-CSI unterstützte Funktionalität

Die im Supervisor ausgeführte vSphere CNS-CSI-Komponente unterstützt mehrere vSphere- und Kubernetes-Speicherfunktionen. Hierbei gelten allerdings bestimmte Einschränkungen.

Unterstützte Funktionen	vSphere CNS-CSI mit Supervisor
CNS-Unterstützung in vSphere Client	Ja
Verbesserte Objektintegrität in vSphere Client	Ja (nur vSAN)
Dynamisches dauerhaftes Volume für Blöcke (ReadWriteOnce-Zugriffsmodus)	Ja
Dynamisches dauerhaftes Volume für Dateien (ReadWriteMany-Zugriffsmodus)	Nein
vSphere-Datenspeicher	VMFS, NFS, vSAN (einschließlich vSAN ESA), vVols
Statisches dauerhaftes Volume	Ja
Verschlüsselung	Nein
Offline-Volume-Erweiterung	Ja
Online-Volume-Erweiterung	Ja
Volume-Topologie und -Zonen	Ja. Volumes können nur von Tanzu Kubernetes Grid-Clustern verbraucht werden.
Mehrere Steuerungsebenen-Instanzen in Kubernetes	Ja
WaitForFirstConsumer	Nein
VolumeHealth	Ja
Storage vMotion mit dauerhaften Volumes	Nein

Persistenter Speicher und Supervisor mit vSphere-Zonen

Ein Supervisor mit drei Zonen unterstützt den zonenspezifischen Speicher, bei dem ein Datenspeicher von allen Hosts in einer einzelnen Zone gemeinsam genutzt wird.



Beachten Sie Folgendes bei der Vorbereitung von Speicherressourcen für den Supervisor mit drei Zonen:

- Der Speicher in allen drei Zonen muss nicht vom gleichen Typ sein. Ein einheitlicher Speicher in allen drei Clustern bietet jedoch eine konsistente Leistung.
- Verwenden Sie für den Namespace auf dem Supervisor mit drei Zonen eine Speicherrichtlinie, die mit dem freigegebenen Speicher in jedem der Cluster kompatibel ist. Die Speicherrichtlinie muss topologiefähig sein.
- Entfernen Sie keine Topologieeinschränkungen aus der Speicherrichtlinie, nachdem Sie sie dem Namespace zugewiesen haben.
- Mounten Sie keine Zonendatenspeicher in anderen Zonen.
- Folgendes unterstützt ein Supervisor mit drei Zonen nicht:
  - Zonenübergreifende Volumes
  - vSAN File-Volumes (ReadWriteMany-Volumes)

- Bereitstellung statischer Volumes mithilfe der Volume-Registrierungs-API
- Arbeitslasten, die die vSAN Data Persistence-Plattform verwenden
- vSphere Pod
- vSAN Stretched Cluster
- VMs mit vGPU und Instanzspeicher

Weitere Informationen finden Sie unter Verwenden von persistentem Speicher auf einem Supervisor mit drei Zonen in der Dokumentation zu *Dienste und Arbeitslasten der vSphere IaaS-Steuerungsebene*.

Tanzu Kubernetes Grid-Architektur und -Komponenten

3

Informieren Sie sich über die Architektur von Tanzu Kubernetes Grid sowie über die Art der Integration in den Supervisor und dessen Komponenten. Hier erfahren Sie, wie Netzwerke und Speicher für Tanzu Kubernetes Grid-Cluster funktionieren, was Hochverfügbarkeit im Zusammenhang mit Tanzu Kubernetes Grid bedeutet und welche Supervisor-Bereitstellung dafür Unterstützung bietet.

Lesen Sie als Nächstes die folgenden Themen:

- Tanzu Kubernetes Grid – Architektur
- Tanzu Kubernetes Grid-Cluster-Netzwerk
- Speicher für Tanzu Kubernetes Grid-Cluster
- Hochverfügbarkeit für Tanzu Kubernetes Grid-Cluster
- Authentifizierung von Tanzu Kubernetes Grid

Tanzu Kubernetes Grid – Architektur

Tanzu Kubernetes Grid ermöglicht die Self-Service-Lebenszyklusverwaltung von Tanzu Kubernetes Grid-Clustern. Sie verwenden den Tanzu Kubernetes Grid zum Erstellen und Verwalten von Tanzu Kubernetes Grid-Clustern anhand eines deklarativen Ansatzes, der den Kubernetes-Anwendern und Entwicklern vertraut ist.

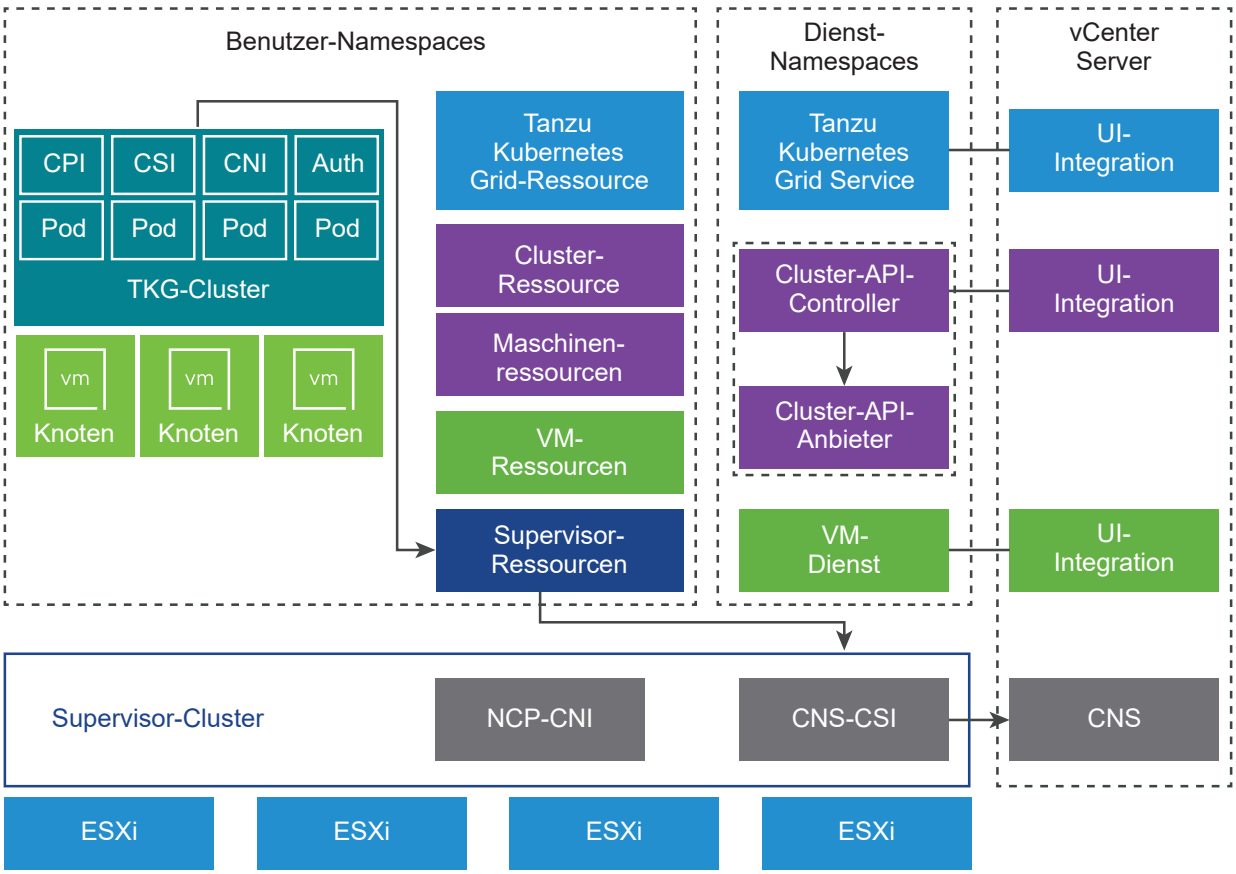
Komponenten für Tanzu Kubernetes Grid

Der Tanzu Kubernetes Grid verfügt über drei Ebenen von Controllern zum Verwalten des Lebenzyklus eines Tanzu Kubernetes Grid-Clusters.

- Der Tanzu Kubernetes Grid stellt Cluster bereit, die die Komponenten enthalten, die für die Integration der zugrunde liegenden vSphere-Namespace-Ressourcen erforderlich sind. Zu diesen Komponenten gehört ein Cloud-Anbieter-Plug-In, das in den Supervisor integriert wird. Darüber hinaus übergibt ein Tanzu Kubernetes Grid-Cluster Anforderungen bezüglich persistenter Volumes an den Supervisor, der in den cloudnativen Speicher von VMware (CNS) integriert ist. Weitere Informationen hierzu finden Sie unter Persistenter Speicher für Arbeitslasten.

- Die Cluster-API stellt deklarative APIs im Kubernetes-Stil für die Erstellung, Konfiguration und Verwaltung von Clustern bereit. Die Eingaben für die Cluster-API umfassen eine Ressource, die den Cluster beschreibt, eine Gruppe von Ressourcen, die die virtuellen Maschinen beschreiben, aus denen sich der Cluster zusammensetzt, sowie eine Reihe von Ressourcen, die Cluster-Add-Ons beschreiben.
- Der VM-Dienst stellt eine deklarative API im Kubernetes-Stil für die Verwaltung von VMs und zugeordneten vSphere-Ressourcen bereit. Der VM-Dienst führt das Konzept „Klassen virtueller Maschinen“ ein, die abstrakte wiederverwendbare Hardwarekonfigurationen darstellen. Die vom VM-Dienst bereitgestellten Funktionen werden verwendet, um den Lebenszyklus der Steuerungsebene und der Worker-Knoten-VMs zu verwalten, die einen Tanzu Kubernetes Grid-Cluster hosten.

Abbildung 3-1. Tanzu Kubernetes Grid-Architektur und -Komponenten



Tanzu Kubernetes Grid-Cluster Komponenten

Die Komponenten, die in einem Tanzu Kubernetes Grid-Cluster ausgeführt werden, erstrecken sich über vier Bereiche: Authentifizierung und Autorisierung, Speicherintegration, Pod-Netzwerk und Lastausgleich.

- Authentifizierungs-Webhook: ein Webhook, der als Pod innerhalb des Clusters ausgeführt wird, um Benutzerauthentifizierungstoken zu validieren.

- Container-Speicherschnittstellen-Plug-In: ein paravirtuelles CSI-Plug-In, das in CNS über den Supervisor integriert wird.
- Container Network Interface-Plug-In: ein CNI-Plug-In, das Pod-Netzwerke bereitstellt.
- Cloud-Anbieterimplementierung: Unterstützt das Erstellen von Kubernetes-Lastausgleichsdiensten.

Tanzu Kubernetes Grid-API

Sie verwenden die Tanzu Kubernetes Grid-API zum Bereitstellen und Verwalten von Tanzu Kubernetes Grid-Clustern. Es handelt sich dabei um eine deklarative API, die Sie mithilfe von kubectl und YAML aufrufen. Sie können die erweiterte ausführbare VMware-Datei „kubectl“ unter der IP-Adresse des Supervisor-API-Endpoints herunterladen.

Mit einer deklarativen API geben Sie anstelle von imperativen Befehlen an das System den gewünschten Zustand des Tanzu Kubernetes Grid-Clusters an: Anzahl der Knoten, verfügbarer Speicher, VM-Größen und Kubernetes-Softwareversion. Der Tanzu Kubernetes Grid übernimmt die Aufgabe der Bereitstellung eines Clusters, dessen Zustand mit dem gewünschten Zustand übereinstimmt.

Zum Aufrufen der Tanzu Kubernetes Grid-API rufen Sie kubectl mithilfe einer YAML-Datei auf, die ihrerseits die API aufruft. Nach der Erstellung des Clusters aktualisieren Sie die YAML-Datei, um den Cluster zu aktualisieren.

Tanzu Kubernetes Grid-Cluster-Netzwerk

Ein Tanzu Kubernetes Grid-Cluster, der vom Tanzu Kubernetes Grid bereitgestellt wird, unterstützt zwei CNI-Optionen: Antrea (Standard) und Calico. Bei beiden Optionen handelt es sich um Open-Source-Software, die ein Netzwerk für Cluster-Pods, Dienste und Ingress bereitstellt.

Tanzu Kubernetes Grid-Cluster, die vom Tanzu Kubernetes Grid bereitgestellt werden, unterstützen die folgenden Container Network Interface (CNI)-Optionen:

- Antrea
- Calico

Antrea fungiert als standardmäßige CNI für neue Tanzu Kubernetes Grid-Cluster. Wenn Sie Antrea verwenden, müssen Sie diese während der Bereitstellung des Clusters nicht als CNI angeben. Zum Verwenden von Calico als CNI stehen zwei Optionen zur Verfügung:

- Geben Sie die CNI direkt in der Cluster-YAML an. Weitere Informationen finden Sie unter v1alpha3-Beispiel: TKC mit benutzerdefiniertem Netzwerk.
- Ändern Sie die Standard-CNI. Weitere Informationen finden Sie unter v1beta1-Beispiel: Cluster mit Calico-CNI.

Hinweis Die Verwendung von Antrea als standardmäßige CNI erfordert eine Mindestversion der OVA-Datei für Tanzu Kubernetes Grid-Cluster. Weitere Informationen finden Sie unter Aktualisierung von TKG 2-Clustern auf Supervisor.

In der Tabelle werden Netzwerkfunktionen von Tanzu Kubernetes Grid-Clustern und deren Implementierung zusammengefasst.

Tabelle 3-1. Tanzu Kubernetes Grid-Cluster-Netzwerk

Endpoint	Anbieter	Beschreibung
Pod-Konnektivität	Antrea oder Calico	Containernetzwerkschnittstelle für Pods. Antrea verwendet Open vSwitch. Calico verwendet die Linux-Bridge mit BGP.
Diensttyp: ClusterIP	Antrea oder Calico	Standardmäßiger Kubernetes-Diensttyp, auf den nur innerhalb des Clusters zugegriffen werden kann.
Diensttyp: NodePort	Antrea oder Calico	Ermöglicht externen Zugriff über einen Port, der vom Kubernetes-Netzwerk-Proxy auf jedem Worker-Knoten geöffnet wird.
Diensttyp: LoadBalancer	NSX-T-Lastausgleichsdienst, NSX Advanced Load Balancer, HAProxy	Für NSX-T, ein virtueller Server pro Diensttypdefinition. Weitere Informationen zu NSX Advanced Load Balancer finden Sie in diesem Abschnitt dieser Dokumentation. **Hinweis** Einige Lastausgleichsfunktionen sind möglicherweise mit HAProxy nicht verfügbar, z. B. die Unterstützung für statische IPs.
Cluster-Ingress	Ingress-Controller eines Drittanbieters	Routing für eingehenden Pod-Datenverkehr. Sie können einen beliebigen Ingress-Controller eines Drittanbieters verwenden, wie z. B. Contour.
Netzwerkrichtlinie	Antrea oder Calico	Steuert den Datenverkehr, der von und zu ausgewählten Pods und Netzwerk-Endpoints zulässig ist. Antrea verwendet Open vSwitch. Calico verwendet Linux-IP-Tabellen.

Speicher für Tanzu Kubernetes Grid-Cluster

Für Tanzu Kubernetes Grid-Cluster ist – wie auch für einige andere Komponenten und Arbeitslasten, die in Supervisor-Namespaces ausgeführt werden – persistenter Speicher erforderlich.

Speicherrichtlinien für Tanzu Kubernetes Grid-Cluster

Um persistente Speicherressourcen für die Tanzu Kubernetes Grid-Cluster bereitzustellen, konfiguriert ein vSphere-Administrator Speicherrichtlinien, die unterschiedliche Speicheranforderungen beschreiben. Der Administrator fügt die Speicherrichtlinien dann dem Namespace hinzu, in dem der Tanzu Kubernetes Grid-Cluster bereitgestellt wird. Mit den für den Namespace sichtbaren Speicherrichtlinien wird festgelegt, auf welche Datenspeicher der Namespace zugreifen und welche Datenspeicher er für persistenten Speicher verwenden kann. Die Richtlinien bestimmen, wie die Clusterknoten und Arbeitslasten in der vSphere-Speicherumgebung platziert werden.

Basierend auf den für den Namespace zugewiesenen Speicherrichtlinien erstellt vSphere IaaS control plane passende Kubernetes-Speicherklassen, die automatisch im Namespace angezeigt werden. Sie werden auch an den Tanzu Kubernetes Grid-Cluster in diesem Namespace weitergegeben.

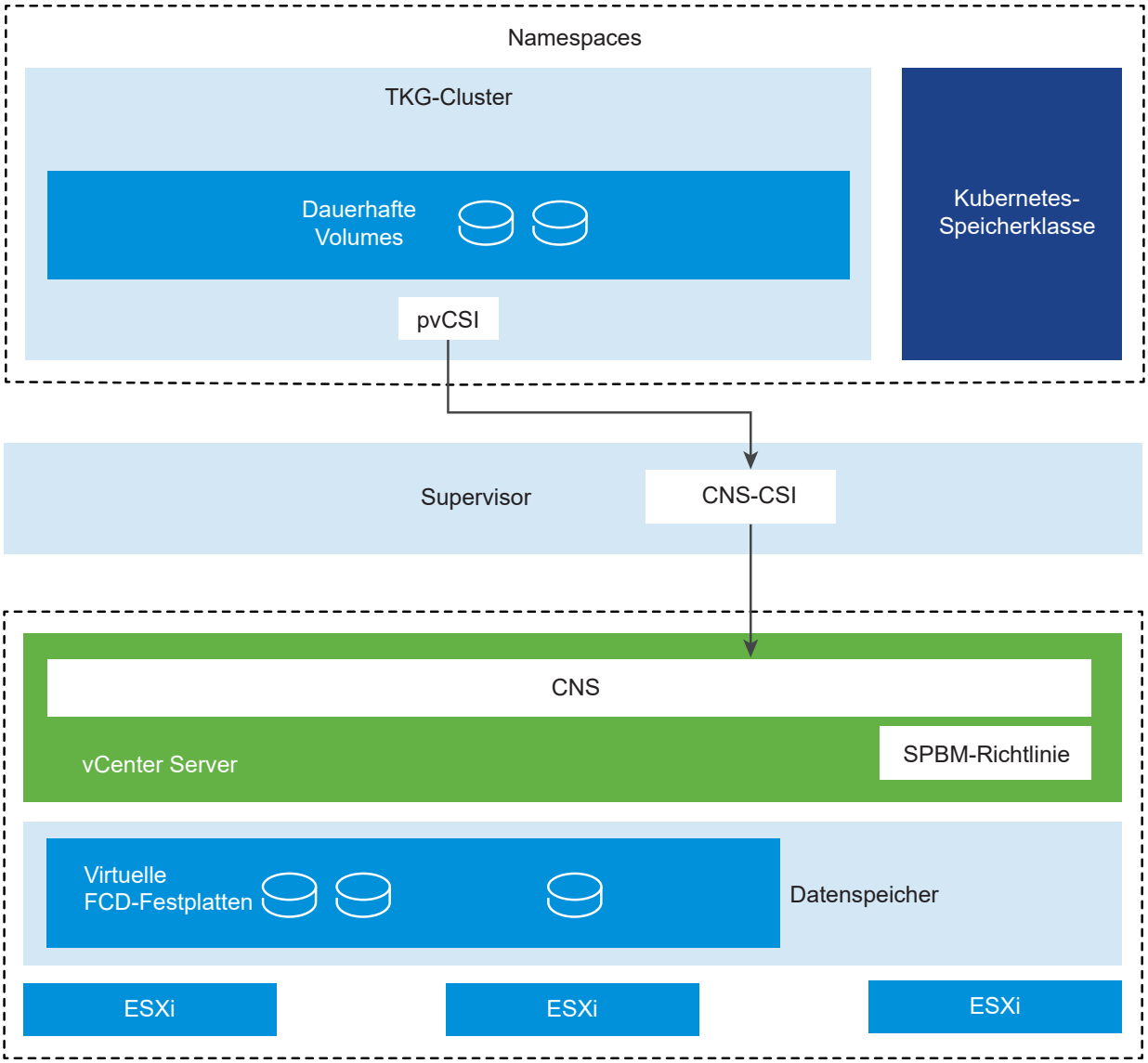
Im Tanzu Kubernetes Grid-Cluster werden die Speicherklassen in zwei Editionen angezeigt: eine mit dem `Immediate`- und eine mit dem `WaitForFirstConsumer`-Bindungsmodus. Die vom DevOps-Team ausgewählte Edition hängt von den jeweiligen Anforderungen ab.

Weitere Informationen zu Speicherklassen in Tanzu Kubernetes Grid-Clustern finden Sie unter Verwenden von Speicherklassen für persistente Volumes.

Integrieren von Tanzu Kubernetes Grid-Clustern in vSphere-Speicher

Für die Integration in den Supervisor und in vSphere-Speicher verwenden Tanzu Kubernetes Grid-Cluster paravirtuelle CSI (pvCSI).

pvCSI ist die für Tanzu Kubernetes Grid-Cluster geänderte Version des vSphere CNS-CSI-Treibers. Die pvCSI-Komponente befindet sich im Tanzu Kubernetes Grid-Cluster und ist für alle aus dem Tanzu Kubernetes Grid-Cluster stammenden, speicherbezogenen Anforderungen zuständig. Die Anforderungen werden an CNS-CSI übermittelt und von dort an CNS in vCenter Server weitergeleitet. Dies führt dazu, dass pvCSI nicht direkt mit der CNS-Komponente kommuniziert, sondern bei allen Speicherbereitstellungsvorgängen auf CNS-CSI angewiesen ist. Im Gegensatz zu CNS-CSI benötigt pvCSI keine Anmeldedaten für die Infrastruktur. pvCSI ist mit einem Dienstkonto im Namespace konfiguriert.

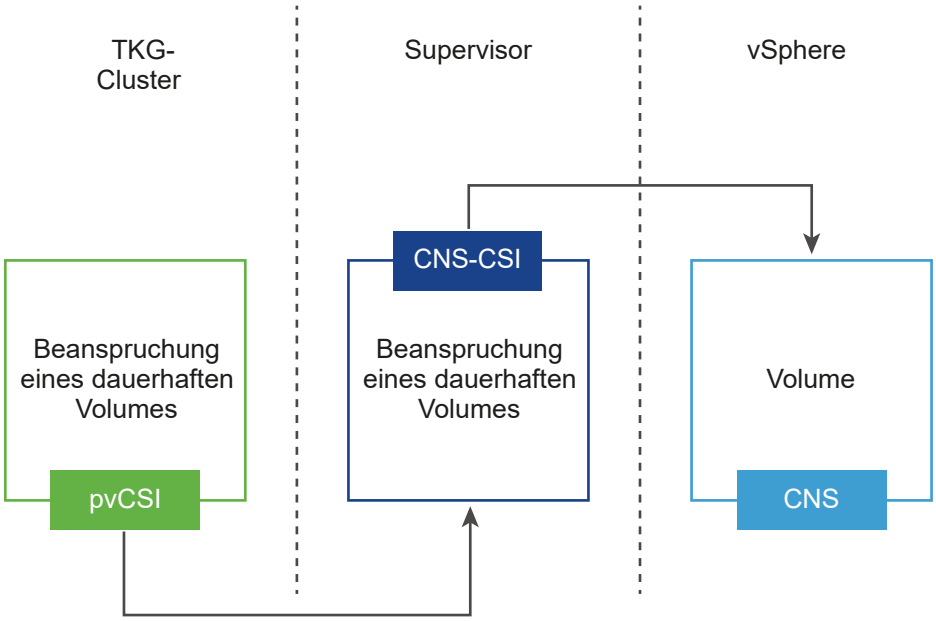


Weitere Informationen zu Supervisor-Komponenten für die Integration in vSphere-Speicher finden Sie unter Persistenter Speicher für Arbeitslasten.

Erstellen eines persistenten Volumes

Im Folgenden ist dargestellt, wie unterschiedliche Komponenten miteinander interagieren, wenn ein DevOps-Ingenieur einen speicherbezogenen Vorgang im Tanzu Kubernetes Grid-Cluster durchführt und z. B. eine Beanspruchung eines dauerhaften Volumes (Persistent Volume Claim, PVC) erstellt.

Der DevOps-Ingenieur erstellt ein PVC über die Befehlszeile im Tanzu Kubernetes Grid-Cluster. Durch diese Aktion wird ein übereinstimmendes PVC auf dem Supervisor generiert. Außerdem wird CNS-CSI ausgelöst. CNS-CSI ruft die CNS-API zum Erstellen von Volumes auf.



Nach erfolgreicher Erstellung eines Volumes wird der Vorgang über den Supervisor wieder an den Tanzu Kubernetes Grid-Cluster zurückgegeben. Infolge dieser Weitergabe können die Benutzer das persistente Volume und die Anforderung des persistenten Volumes im gebundenen Zustand im Supervisor sehen. Außerdem wird ihnen beides im gebundenen Zustand auch im Tanzu Kubernetes Grid-Cluster angezeigt.

Von pvCSI unterstützte Funktionen

Die im Tanzu Kubernetes Grid-Cluster ausgeführte pvCSI-Komponente unterstützt eine Reihe von vSphere- und Kubernetes-Speicherfunktionen.

Unterstützte Funktionen	pvCSI mit Tanzu Kubernetes Grid-Cluster
CNS-Unterstützung in vSphere Client	Ja
Verbesserte Objektintegrität in vSphere Client	Ja (nur vSAN)
Dynamisches dauerhaftes Volume für Blöcke (ReadWriteOnce-Zugriffsmodus)	Ja
Dynamisches dauerhaftes Volume für Dateien (ReadWriteMany-Zugriffsmodus)	Ja (mit vSAN-Dateidiensten)
vSphere-Datenspeicher	VMFS/NFS/vSAN/vVols
Statisches dauerhaftes Volume	Ja
Verschlüsselung	Nein
Offline-Volume-Erweiterung	Ja
Online-Volume-Erweiterung	Ja
Volume-Topologie und -Zonen	Ja
Mehrere Steuerungsebenen-Instanzen in Kubernetes	Ja
WaitForFirstConsumer	Ja

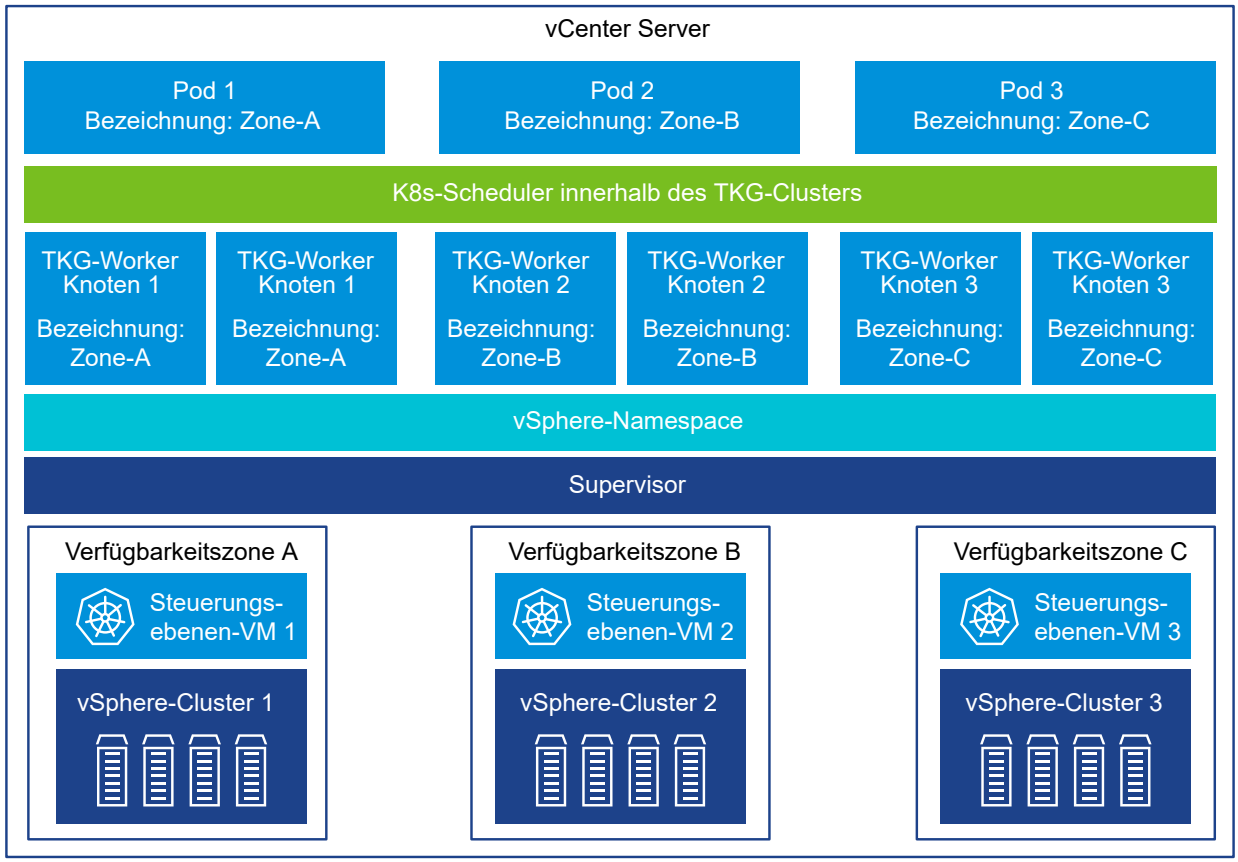
Unterstützte Funktionen	pvCSI mit Tanzu Kubernetes Grid-Cluster
VolumeHealth	Ja
Storage vMotion mit dauerhaften Volumes	Nein

Hochverfügbarkeit für Tanzu Kubernetes Grid-Cluster

Sie können Tanzu Kubernetes Grid-Clustern eine Hochverfügbarkeit bieten, wenn sie auf einem Supervisor in drei vSphere-Zonen bereitgestellt werden. Eine vSphere-Zone ist einem vSphere-Cluster zugeordnet, was bedeutet, dass bei einer Bereitstellung eines Supervisor in drei vSphere-Zonen die Ressourcen aller drei zugrunde liegenden vSphere-Cluster verwendet werden. Dadurch werden Ihre innerhalb Tanzu Kubernetes Grid-Clustern ausgeführten Kubernetes-Arbeitslasten vor Ausfällen auf vSphere-Clusterebene geschützt. In einer Bereitstellung mit einer Zone wird die Hochverfügbarkeit für Tanzu Kubernetes Grid-Cluster auf einer ESXi-Hostebene durch vSphere HA bereitgestellt.

In einem Supervisor mit drei Zonen werden Steuerungsebenenknoten von Tanzu Kubernetes Grid-Clustern automatisch über die vSphere-Zonen hinweg platziert. Sie können jedoch steuern, wie Worker-Knoten auf die Zonen verteilt werden. Sie können ein NodePool-Objekt für die Worker-Knoten von Tanzu Kubernetes Grid-Clustern definieren und jede vSphere-Zone einer FailureDomain innerhalb jedes NodePools zuordnen. Auf diese Weise sorgt die Cluster-API dafür, dass die Worker-Knoten entsprechend auf vSphere-Zonen verteilt werden. Wenn Sie die Angabe einer FailureDomain für einen oder alle NodePools überspringen, verteilt die Cluster-API die NodePools automatisch über Zonen.

Abbildung 3-2. High Availability für Tanzu Kubernetes Grid-Cluster in mehreren Zonen



Authentifizierung von Tanzu Kubernetes Grid

Informieren Sie sich über die verschiedenen Authentifizierungsmechanismen und deren Verwendung mit Tanzu Kubernetes Grid-Clustern.

Aufbauen einer Verbindung zum Supervisor

Als DevOps-Ingenieur stellen Sie eine Verbindung zum Supervisor her, um Tanzu Kubernetes Grid-Cluster bereitzustellen. Sie haben nur Zugriff auf die Namespaces, für die Ihnen der vSphere-Administrator Berechtigungen erteilt hat.

Für einen Verbindungsaufbau zum Supervisor auf der Kubernetes-Steuerungsebenen-IP oder zu bereitgestellten Tanzu Kubernetes Grid-Clustern haben Sie die Wahl zwischen zwei Verfahren:

- vCenter Single Sign-On und Kubernetes-CLI-Tools für vSphere. In diesem Fall wird ein Authentifizierungstoken erstellt, das alle 10 Stunden abläuft.
- Anmeldedaten von einem beim Supervisor registrierten OIDC-Anbieter und Tanzu-CLI. Die Sitzung mit und dem OIDC-Anbieter wird durch die Einstellungen im Anbieter selbst gesteuert.

Weitere Informationen finden Sie in der Dokumentation *Verwenden des TKG-Dienstes mit der vSphere IaaS-Steuerungsebene*.

Herstellen einer Verbindung mit Tanzu Kubernetes Grid-Clustern

Als DevOps-Ingenieur stellen Sie auch eine Verbindung zu bereitgestellten Tanzu Kubernetes Grid-Clustern her, um sie zu betreiben und zu verwalten. Wenn Ihrem Benutzerkonto eine Bearbeitungsberechtigung für den vSphere-Namespace erteilt wird, in dem der Tanzu Kubernetes Grid-Cluster bereitgestellt wird, wird Ihr Konto der Rolle `cluster-admin` zugewiesen. Alternativ können Sie auch den Benutzertyp `kubernetes-admin` verwenden, um eine Verbindung zu Tanzu Kubernetes Grid -Clustern aufzubauen. Außerdem haben Sie die Möglichkeit, Entwicklern Zugriff auf Tanzu Kubernetes Grid-Cluster zu gewähren, indem Sie einen Benutzer oder eine Gruppe an die standardmäßige oder benutzerdefinierte Pod-Sicherheitsrichtlinie binden. Weitere Informationen finden Sie in der Dokumentation *Verwenden des TKG-Dienstes mit der vSphere IaaS-Steuerungsebene*.

Supervisor-Bereitstellungsoptionen

4

Informieren Sie sich über die Möglichkeiten zur Bereitstellung und Konfiguration eines Supervisors. Je nach Netzwerk-Stack oder der für den Supervisor implementierten Bereitstellungsoption unterscheiden sich die unterstützten Topologien und Arbeitslasttypen.

Lesen Sie als Nächstes die folgenden Themen:

- Supervisor-Bereitstellungen mit Zonen und Clustern
- Topologie für einen Supervisor mit VDS-Netzwerk und NSX Advanced Load Balancer
- Topologien für einen Supervisor mit einer Zone und NSX als Netzwerk-Stack
- Topologien für einen Supervisor mit einer Zone und NSX als Netzwerk-Stack und NSX Advanced Load Balancer
- Topologien für die Bereitstellung des HAProxy-Lastausgleichsdiensts

Supervisor-Bereitstellungen mit Zonen und Clustern

Informieren Sie sich über die Unterschiede zwischen der Bereitstellung des Supervisors auf drei vSphere-Clustern, die vSphere-Zonen zugeordnet sind, und der Bereitstellung des Supervisors auf einem einzelnen Cluster mit Zuordnung zu einer vSphere-Zone.

Hinweis Nachdem Sie einen Supervisor auf einem einzelnen vSphere Cluster bereitgestellt haben, was zum Erstellen einer vSphere-Zone führt, können Sie den Supervisor nicht auf eine Bereitstellung mit drei Zonen erweitern. Sie können einen Supervisor entweder in einer vSphere-Zone (Bereitstellung mit einem einzelnen Cluster) oder in drei vSphere-Zonen bereitstellen.

Supervisor-Bereitstellung mit drei Zonen für Hochverfügbarkeit auf Clusterebene

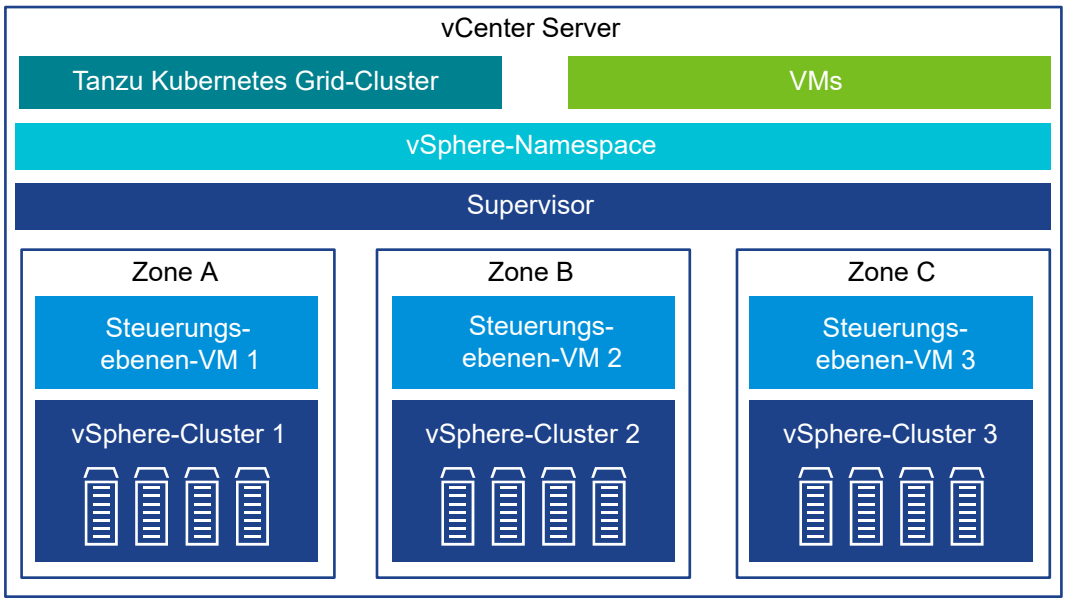
Sie können vSphere IaaS control plane auf drei vSphere-Clustern aktivieren, die drei vSphere-Zonen zugeordnet sind. Sie konfigurieren jeden vSphere-Cluster als unabhängige Ausfalldomäne und ordnen ihn einer vSphere-Zone zu. Bei einer Bereitstellung mit drei Zonen werden alle drei vSphere-Cluster zu einem Supervisor. Bei einer Bereitstellung mit drei Zonen haben Sie folgende Möglichkeiten:

- Bereitstellen von Hochverfügbarkeit auf Clusterebene für den Supervisor, da jeder vSphere-Cluster eine unabhängige Ausfalldomäne darstellt.

- Verteilen der Knoten Ihrer Tanzu Kubernetes Grid-Cluster auf alle drei vSphere-Zonen, sodass Hochverfügbarkeit für Ihre Kubernetes-Arbeitslasten auf vSphere-Clusterebene bereitgestellt wird.
- Skalieren des Supervisors durch Hinzufügen von Hosts zu jedem der drei vSphere-Cluster.

Sie können Arbeitslasten auf einem Supervisor mit drei Zonen ausführen, indem Sie Tanzu Kubernetes Grid-Cluster, vSphere-Pods und VMs verwenden.

Abbildung 4-1. Supervisor-Bereitstellung mit drei Zonen



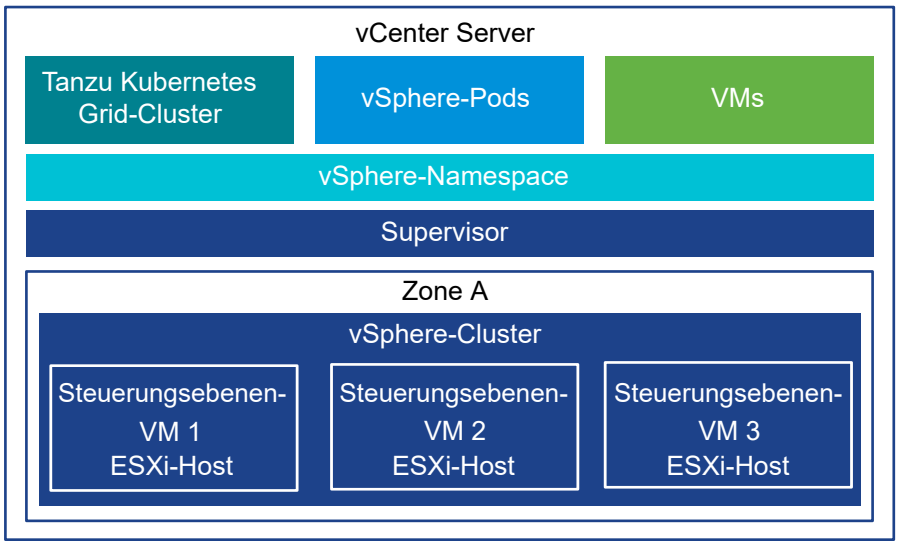
Platzierung von vSphere-Zonen über physische Sites hinweg

Sie können vSphere-Zonen auf verschiedene physische Sites verteilen, solange die Latenz zwischen den Sites 100 ms nicht überschreitet. Sie können beispielsweise die vSphere-Zonen auf zwei physische Sites verteilen: eine vSphere-Zone auf der ersten Site und zwei vSphere-Zonen auf der zweiten Site.

Supervisor-Bereitstellung mit einem einzelnen Cluster

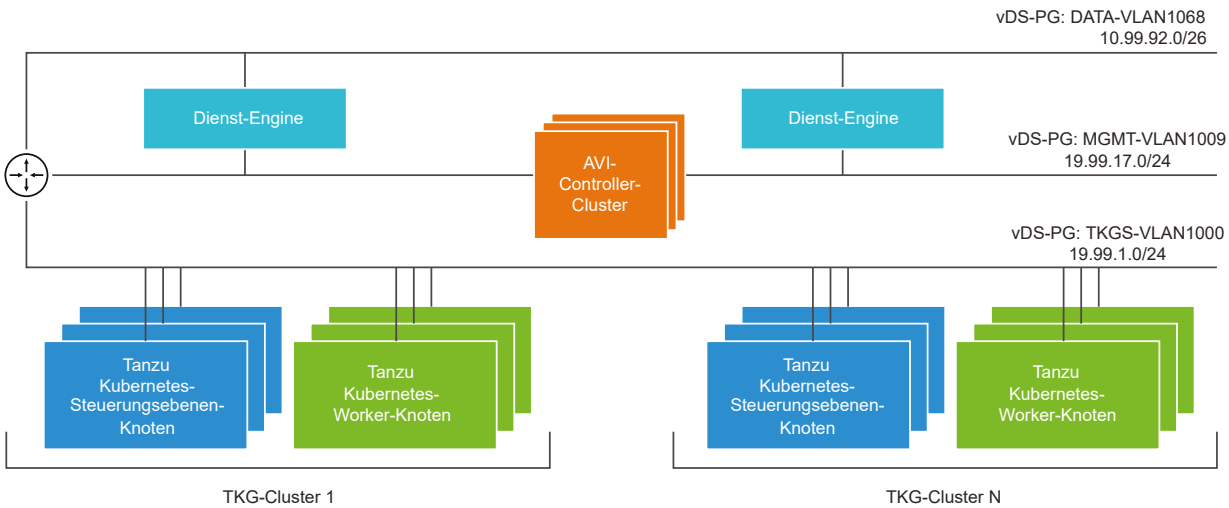
Sie können einen Supervisor auch auf einem einzelnen vSphere-Cluster aktivieren. In diesem Fall wird automatisch eine einzelne Zone für den Supervisor erstellt. Alternativ haben Sie die Möglichkeit, eine Zone zu verwenden, die Sie vorab erstellt haben. In einer Bereitstellung mit einem einzelnen Cluster verfügen Sie durch vSphere HA immer noch über Hochverfügbarkeit auf Clusterebene, und Sie können Ihre vSphere IaaS control plane-Konfiguration nur skalieren, indem Sie dem vSphere-Cluster, der dem Supervisor zugeordnet ist, Hosts hinzufügen. In einer Bereitstellung mit einem einzelnen Cluster können Sie Arbeitslasten über vSphere-Pods, Tanzu Kubernetes Grid-Cluster und VMs ausführen, die über den VM-Dienst bereitgestellt werden.

Abbildung 4-2. Supervisor-Bereitstellung mit einem einzelnen Cluster



Topologie für einen Supervisor mit VDS-Netzwerk und NSX Advanced Load Balancer

Der AVI-Controller wird immer im Verwaltungsnetzwerk bereitgestellt, in dem er mit dem vCenter Server, ESXi-Hosts und Supervisor-Knoten der Steuerungsebene verbunden werden kann. Die Dienst-Engines werden mit Schnittstellen zum Verwaltungsnetzwerk und zum Datennetzwerk bereitgestellt.



Im Verwaltungsnetzwerk, z. B. `MGMT-VLAN1009`, befindet sich der Controller, und die Verwaltungsschnittstelle der Dienst-Engines ist mit dem Netzwerk verbunden.

Die Dienst-Engine-Schnittstellen, die eine Verbindung zur VIP-Platzierung herstellen, befinden sich im Datennetzwerk, wie z. B. `DATA-VLAN1068`. Der Client-Datenverkehr erreicht die VIP, und die Dienst-Engines verteilen den Datenverkehr zu den Arbeitslastnetzwerk-IPs über dieses Netzwerk.

Im Arbeitslastnetzwerk, wie z. B. `TKGS-VLAN1000`, werden die Tanzu Kubernetes Grid-Cluster ausgeführt. Die Dienst-Engines benötigen keine Schnittstellen zum Arbeitslastnetzwerk.

Die Dienst-Engines werden im einarmigen Modus ausgeführt. Sie leiten den Lastausgleichsdatenverkehr über den Router an das Arbeitslastnetzwerk weiter. Die Dienst-Engines erhalten nicht die Standard-Gateway-IP von DHCP in den Datennetzwerken. Sie müssen statische Routen konfigurieren, damit die Dienst-Engines den Datenverkehr ordnungsgemäß an die Arbeitslastnetzwerke und die Client-IP weiterleiten können.

Diese Topologie bietet die Möglichkeit, dass sich die Dienst-Engine in einem einzelnen Netzwerk befindet. Die Erstellung der Dienst-Engine und die Netzwerkverbindungen werden vom AVI-Controller automatisiert.

Informationen zum Installieren und Konfigurieren des NSX Advanced Load Balancers finden Sie unter Installieren und Konfigurieren des NSX Advanced Load Balancers

Komponenten für NSX Advanced Load Balancer

Zu den Komponenten des NSX Advanced Load Balancer, auch bekannt als Avi-Lastausgleichsdienst, gehören der Controller-Cluster, VMs der Dienst-Engines (Datenebene) und der Avi-Kubernetes-Operator (AKO).

Informationen zum Installieren und Konfigurieren von NSX Advanced Load Balancer-Komponenten finden Sie unter Installieren und Konfigurieren von NSX Advanced Load Balancer.

Controller

Der NSX Advanced Load Balancer Controller, auch als Controller bezeichnet, interagiert mit dem vCenter Server, um den Lastausgleich für die Tanzu Kubernetes Grid-Cluster zu automatisieren. Er ist für die Bereitstellung von Dienst-Engines, die Koordination von Ressourcen anhand von Dienst-Engines sowie die Zusammenfassung von Dienst-Engine-Metriken und -Protokollen zuständig. Der Controller bietet eine Webschnittstelle, eine Befehlszeilenschnittstelle und eine API für Benutzervorgänge und die programmgesteuerte Integration.

Nachdem Sie die Controller-VM in vSphere bereitgestellt und konfiguriert haben, können Sie einen Controller-Cluster bereitstellen, um den Steuerungsebenen-Cluster für HA einzurichten.

Clouds sind Container für die Umgebung, in der der NSX Advanced Load Balancer installiert ist oder betrieben wird. Während der Erstkonfiguration des Controllers wird automatisch eine Cloud mit dem Namen **Default-cloud** erstellt. Sie können die **Default-cloud** als **VMware vCenter**-Cloud verwenden oder eine oder mehrere benutzerdefinierte Clouds vom Typ **VMware vCenter** erstellen.

Wenn Sie eine Cloud vom Typ **VMware vCenter** konfigurieren, wird diese einem eindeutigen vCenter und einem Datencenter innerhalb dieses vCenter zugeordnet. Alle Ressourcen, die diesem vCenter und Datencenter zur Verfügung stehen, sind für die Cloud verfügbar.

Damit der Lastausgleichsdienst mehrere vCenter-Server oder mehrere Datencenter bedienen kann, können Sie mehrere benutzerdefinierte Clouds vom Typ **VMware vCenter** erstellen, eine für jede Kombination aus vCenter und Datencenter. Dies senkt die Arbeitslast, da weniger Lastausgleichsdienst-Instanzen und somit weniger Kerne zur Unterstützung der Umgebung erforderlich sind. Weitere Informationen zu Clouds finden Sie in der Dokumentation zu NSX Advanced Load Balancer.

Dienst-Engine

Bei der NSX Advanced Load Balancer-Dienst-Engine, auch als Dienst-Engine bezeichnet, handelt es sich um die virtuelle Maschine der Datenebene. Eine Dienst-Engine führt einen oder mehrere virtuelle Dienste aus. Eine Dienst-Engine wird vom Controller verwaltet. Der Controller stellt Dienst-Engines für das Hosten virtueller Dienste zur Verfügung.

Die Dienst-Engines verfügen über zwei Arten von Netzwerkschnittstellen:

- Die erste Netzwerkschnittstelle, `vnic0` der VM, wird mit dem Verwaltungsnetzwerk verbunden, wo sie eine Verbindung zum NSX Advanced Load Balancer Controller herstellen kann.
- Die restlichen Schnittstellen, `vnic1 – 9`, verbinden sich mit dem Datennetzwerk, in dem virtuelle Dienste ausgeführt werden.

Die Dienst-Engine-Schnittstellen stellen automatisch eine Verbindung mit den richtigen vDS-Portgruppen her. Nicht verwendete Schnittstellen sind mit einer Verwaltungsnetzwerk-Portgruppe im getrennten Zustand verbunden. Jede Dienst-Engine kann bis zu 1.000 virtuelle Dienste unterstützen.

Ein virtueller Dienst stellt Ebene-4- und Ebene-7-Lastausgleichsdienste für Tanzu Kubernetes Grid-Clusterarbeitslasten zur Verfügung. Ein virtueller Dienst ist mit einer virtuellen IP und mehreren Ports konfiguriert. Wenn ein virtueller Dienst bereitgestellt wird, wählt der Controller automatisch einen ESX-Server aus, startet eine Dienst-Engine und verbindet sie mit den richtigen Netzwerken (Portgruppen).

Die erste Dienst-Engine wird erst erstellt, nachdem der erste virtuelle Dienst konfiguriert wurde. Alle nachfolgenden virtuellen Dienste, die konfiguriert werden, verwenden die vorhandene Dienst-Engine.

Jeder virtuelle Server macht einen Load Balancer der Ebene 4 mit einer eindeutigen IP-Adresse des Typs Load Balancer für einen Tanzu Kubernetes Grid verfügbar. Die IP-Adresse, die jedem virtuellen Server zugewiesen ist, wird aus dem IP-Adressblock ausgewählt, der dem Controller bei der Konfiguration zugewiesen wurde.

AVI umfasst natives IPAM und die Unterstützung externer IPAM-Anbieter. In vSphere wird AVI-natives IPAM genutzt.

Avi-Kubernetes-Operator

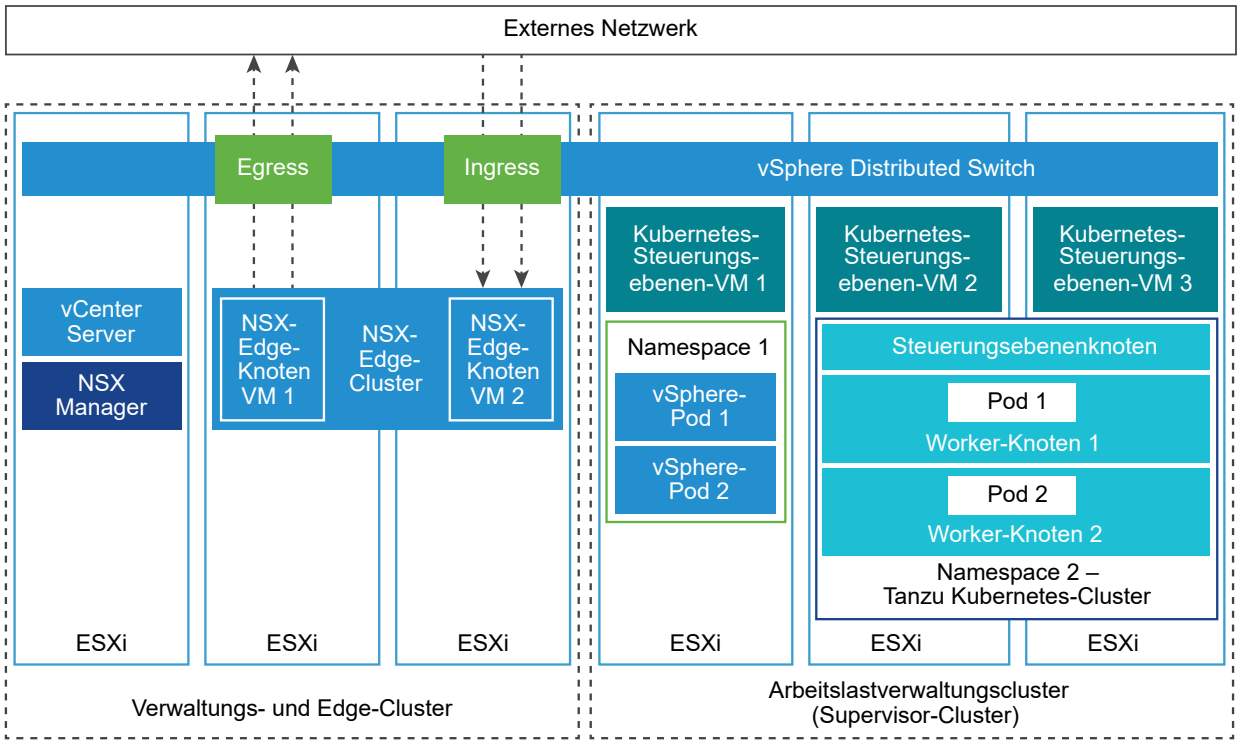
Der Avi-Kubernetes-Operator (AKO) überwacht Kubernetes-Ressourcen und kommuniziert mit dem Controller, um die entsprechenden Lastausgleichsressourcen anzufordern.

Der Avi-Kubernetes-Operator wird im Rahmen des Aktivierungsprozesses auf den Supervisoren installiert.

Topologien für einen Supervisor mit einer Zone und NSX als Netzwerk-Stack

Sie können vSphere IaaS control plane in zwei Clustern bereitstellen: einem Cluster für die Verwaltungs- und Edge-Funktionen und einem weiteren für die Arbeitslastverwaltung.

Abbildung 4-3. Verwaltungs- und Edge- und Arbeitslastverwaltungscluster



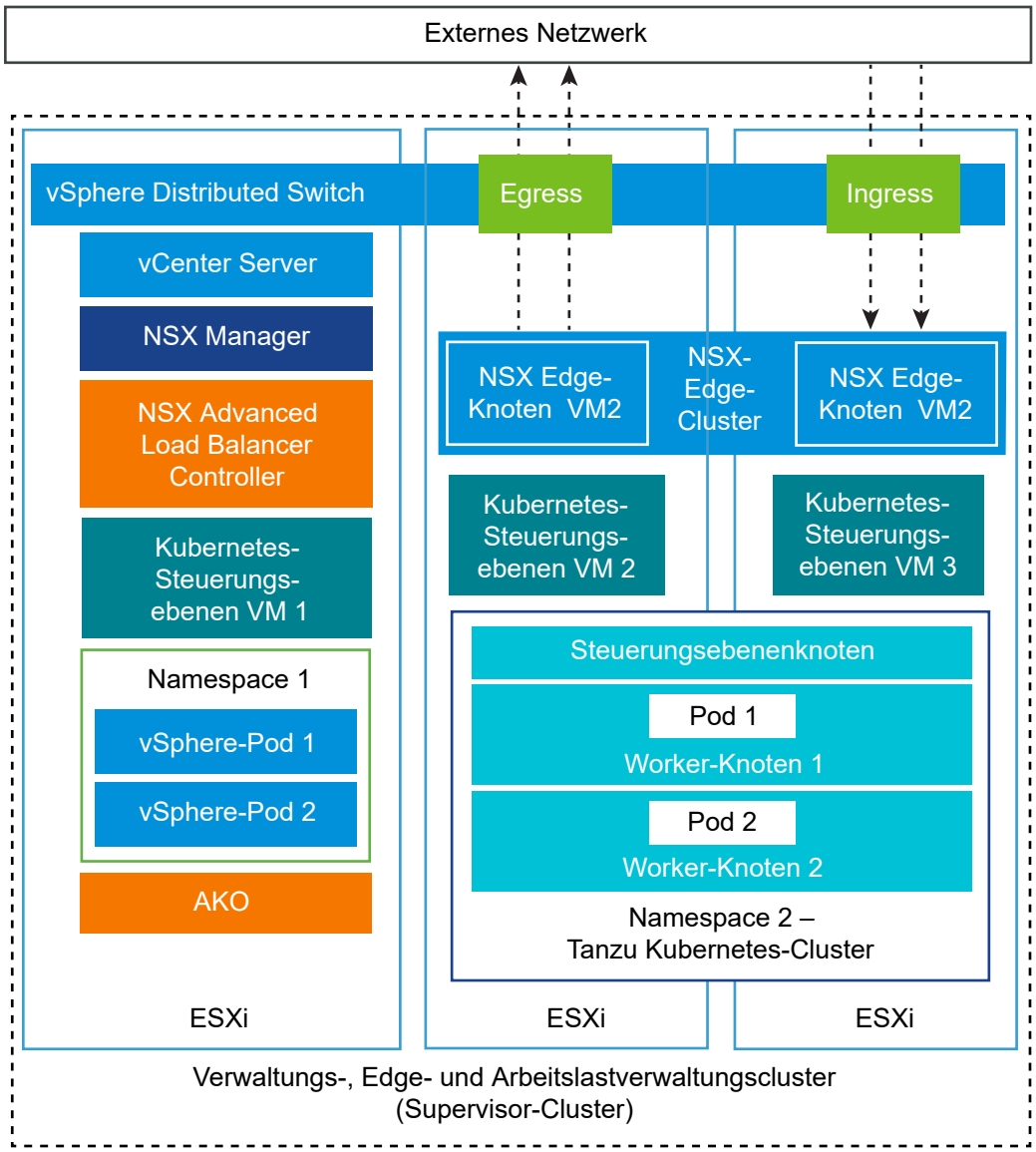
Topologien für einen Supervisor mit einer Zone und NSX als Netzwerk-Stack und NSX Advanced Load Balancer

Je nach den Anforderungen Ihrer Kubernetes-Arbeitslasten und der zugrunde liegenden Netzwerkinfrastruktur können Sie verschiedene Topologien auf Ihren Supervisor anwenden.

Topologie für einen Verwaltungs-, Edge- und Arbeitslastdomänen-Cluster

Sie können vSphere IaaS control plane mit kombinierten Verwaltungs-, Edge- und Arbeitslastverwaltungsfunktionen in einem einzelnen vSphere-Cluster bereitstellen.

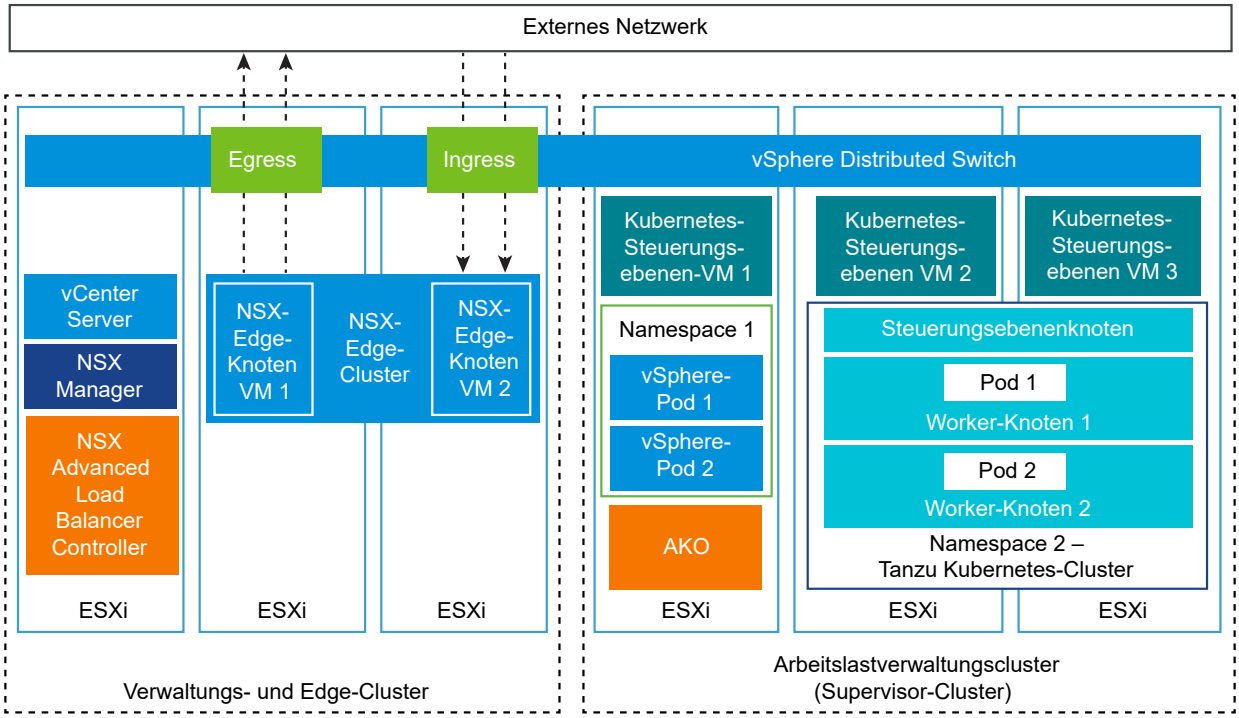
Abbildung 4-4. Verwaltungs-, Edge- und Arbeitslastverwaltungscluster



Topologie mit separatem Verwaltungs- und Edge-Cluster und Arbeitslastverwaltungscluster

Sie können vSphere IaaS control plane in zwei Clustern bereitstellen: einem Cluster für die Verwaltungs- und Edge-Funktionen und einem weiteren für die Arbeitslastverwaltung.

Abbildung 4-5. Verwaltungs- und Edge- und Arbeitslastverwaltungscluster



Topologien für die Bereitstellung des HAProxy-Lastausgleichsdiensts

Überprüfen Sie die möglichen Topologien, die Sie für den HAProxy-Lastausgleichsdienst bei einem mit VDS-Netzwerk konfigurierten Supervisor implementieren können. Bei Verwendung von vSphere IaaS control plane mit VDS-Netzwerk bietet HAProxy Lastausgleich für Entwickler, die auf die Tanzu Kubernetes Grid-Steuerungsebene zugreifen, sowie für Kubernetes-Dienste vom Typ „Lastausgleichsdienst“.

Arbeitslastnetzwerke auf dem Supervisor

Zum Konfigurieren eines Supervisors mit dem VDS-Netzwerk müssen Sie alle Hosts aus dem Cluster mit einem VDS verbinden. Je nach der von Ihnen für die Supervisor-Arbeitslastnetzwerke implementierten Topologie können Sie eine oder mehrere verteilte Portgruppen erstellen. Sie bestimmen die Portgruppen als Arbeitslastnetzwerke für vSphere-Namespaces.

Arbeitslastnetzwerke bieten eine Verbindung zu den Knoten von Tanzu Kubernetes Grid-Clustern und zu den VMs der Supervisor Control Plane. Das Arbeitslastnetzwerk, das Konnektivität zu den Kubernetes-Steuerungsebenen-VMs bereitstellt, wird als „primäres Arbeitslastnetzwerk“ bezeichnet. Jeder Supervisor muss über ein primäres Arbeitslastnetzwerk verfügen. Sie müssen eine der verteilten Portgruppen als primäres Arbeitslastnetzwerk für den Supervisor festlegen.

Hinweis Arbeitslastnetzwerke werden nur während der Supervisor-Aktivierung hinzugefügt und können später nicht mehr hinzugefügt werden.

Die Kubernetes-Steuerungsebenen-VMs im Supervisor verwenden drei IP-Adressen aus dem IP-Adressbereich, der dem primären Arbeitslastnetzwerk zugewiesen ist. Jeder Knoten eines Tanzu Kubernetes Grid-Clusters verfügt über eine eigene, aus dem Adressbereich des Arbeitslastnetzwerks zugewiesene IP-Adresse. Das Arbeitslastnetzwerk ist mit dem Namespace konfiguriert ist, in dem der Tanzu Kubernetes Grid-Cluster ausgeführt wird.

Zuteilung von IP-Bereichen

Wenn Sie die Planung für die Netzwerktopologie des Supervisors mit dem HAProxy-Lastausgleichsdienst vornehmen, sollten Sie die Verwendung von zwei IP-Bereichstypen einplanen:

- Ein Bereich für die Zuteilung virtueller IPs für HAProxy. Der IP-Bereich, den Sie für die virtuellen Server von HAProxy konfigurieren, ist von der Lastausgleichsdienst-Appliance reserviert. Wenn der Bereich für virtuelle IPs beispielsweise `192.168.1.0/24` lautet, sind sämtliche Hosts in diesem Bereich für keinen anderen Datenverkehr als den Datenverkehr über virtuelle IPs verfügbar.

  Hinweis Sie dürfen kein Gateway innerhalb des virtuellen HAProxy-IP-Bereichs konfigurieren, da alle Routen zu einem solchen Gateway fehlschlagen.

- Ein IP-Bereich für die Knoten des Supervisors und der Tanzu Kubernetes Grid-Cluster. Jeder Kubernetes-Steuerungsebenen-VM im Supervisor ist eine IP-Adresse zugewiesen, in der Summe sind dies drei IP-Adressen. Jedem Knoten eines Tanzu Kubernetes Grid-Clusters ist auch eine eigene IP-Adresse zugewiesen. Sie müssen jedem Arbeitslastnetzwerk auf dem Supervisor, das Sie für einen Namespace konfigurieren, einen eindeutigen IP-Bereich zuweisen.

Beispiel für eine Konfiguration mit einem /24-Netzwerk:

- Netzwerk: 192.168.120.0/24
- HAProxy-VIPs: 192.168.120.128/25
- 1 IP-Adresse für die HAProxy-Arbeitslastschnittstelle: 192.168.120.5

Abhängig von den IPs, die innerhalb der ersten 128 Adressen frei sind, können Sie IP-Bereiche für Arbeitslastnetzwerke auf dem Supervisor definieren, z. B.:

- 192.168.120.31-192.168.120.40 für das primäre Arbeitslastnetzwerk
- 192.168.120.51-192.168.120.60 für ein weiteres Arbeitslastnetzwerk

Hinweis Die Bereiche, die Sie für Arbeitslastnetzwerke definieren, dürfen sich nicht mit dem HAProxy-VIP-Bereich überschneiden.

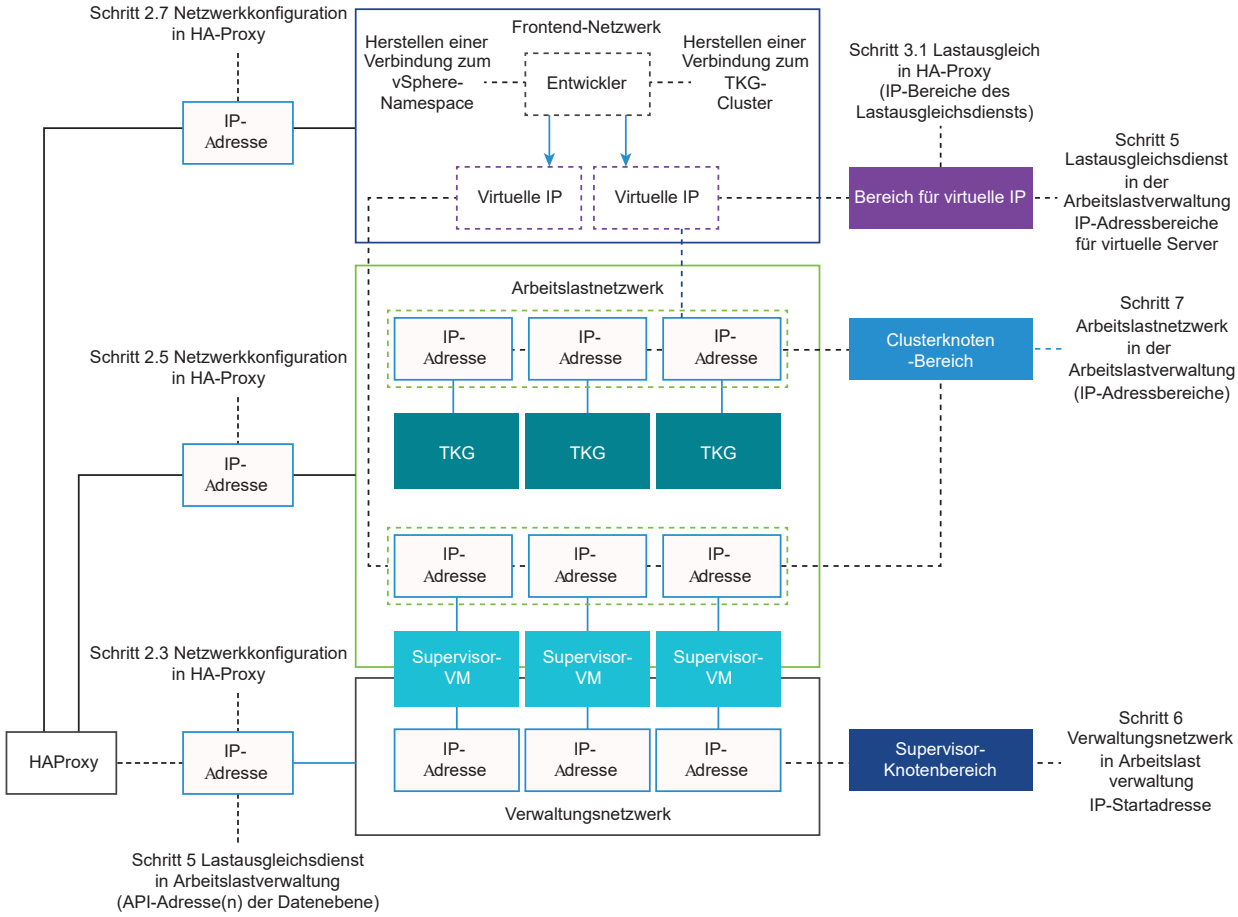
HAProxy-Netzwerktopologie

Es gibt zwei Optionen für die Netzwerkkonfiguration für die Bereitstellung von HAProxy: **Standard** und **Frontend**. Das Standardnetzwerk verfügt über zwei Netzwerkkarten: eine für das Verwaltungsnetzwerk und eine für das Arbeitslastnetzwerk. Das Frontend-Netzwerk verfügt über 3 NICs: Verwaltungsnetzwerk, Arbeitslastnetzwerk und das Frontend-Netzwerk für Clients. In der Tabelle sind die Merkmale jedes Netzwerks aufgelistet und beschrieben.

Für Produktionsinstallationen empfiehlt es sich, den HAProxy-Lastausgleichsdienst mit der Konfiguration für das **Frontend-Netzwerk** zu implementieren. Wenn Sie den HAProxy-Lastausgleichsdienst mit der **Standardkonfiguration** bereitstellen, wird empfohlen, dass Sie dem Arbeitslastnetzwerk eine /24-IP-Adressblockgröße zuweisen. Für beide Konfigurationsoptionen wird DHCP nicht empfohlen.

Netzwerk	Merkmale
Verwaltung	Der Supervisor-Cluster nutzt das Verwaltungsnetzwerk, um eine Verbindung zum HAProxy-Lastausgleichsdienst herzustellen und diesen zu programmieren. ■ Der HAProxy-Datenebenen-API-Endpoint ist an die mit dem Verwaltungsnetzwerk verbundene Netzwerkschnittstelle gebunden. ■ Die der HAProxy-Steuerungsebenen-VM zugewiesene Verwaltungs-IP-Adresse muss eine statische IP im Verwaltungsnetzwerk sein, damit der Supervisor-Cluster zuverlässig eine Verbindung mit der Lastausgleichsdienst-API herstellen kann. ■ Das Standard-Gateway für die HAProxy-VM sollte sich in diesem Netzwerk befinden. ■ DNS-Abfragen sollten in diesem Netzwerk ausgeführt werden.
Arbeitslast	Die HAProxy-Steuerungsebenen-VM nutzt das Arbeitslastennetzwerk, um auf die Dienste im Supervisor-Cluster und auf den Tanzu Kubernetes-Clusterknoten zuzugreifen. ■ Die HAProxy-Steuerungsebenen-VM leitet Datenverkehr an die Supervisor und die Tanzu Kubernetes-Clusterknoten in diesem Netzwerk weiter. ■ Wenn die HAProxy-Steuerungsebenen-VM im Standardmodus (zwei Netzwerkkarten) bereitgestellt wird, muss das Arbeitslastnetzwerk die logischen Netzwerke bereitstellen, die für den Zugriff auf die Lastausgleichsdienste verwendet werden. ■ In der **Standardkonfiguration** kommen die virtuellen IPs des Lastausgleichsdiensts und die IPs der Kubernetes-Clusterknoten von diesem Netzwerk. Sie werden als getrennte, nicht überlappende Bereiche innerhalb des Netzwerks definiert. **Hinweis** Das Arbeitslastnetzwerk muss sich in einem anderen Subnetz als das Verwaltungsnetzwerk befinden. Weitere Informationen finden Sie in den Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst.
Frontend (optional)	Externe Clients (z. B. Benutzer oder Anwendungen), die auf Clusterarbeitslasten zugreifen, verwenden das Frontend-Netzwerk, um über virtuelle IP-Adressen auf Backend-Lastausgleichsdienste zuzugreifen. ■ Das Frontend-Netzwerk wird nur dann verwendet, wenn die HAProxy-Steuerungsebenen-VM mit drei Netzwerkkarten bereitgestellt wird. ■ Empfohlen für Produktionsinstallationen. ■ Das Frontend-Netzwerk ist der Ort, an dem Sie die virtuelle IP-Adresse (VIP) angeben. HAProxy sorgt für einen Ausgleich des Datenverkehrs und leitet diesen an das entsprechende Backend weiter.

Das folgende Diagramm veranschaulicht eine HAProxy-Bereitstellung mithilfe einer **Frontend-Netzwerk**-Topologie. Das Diagrammverzeichnis gibt an, wo während des Installations- und Konfigurationsvorgangs Konfigurationsfelder erwartet werden.



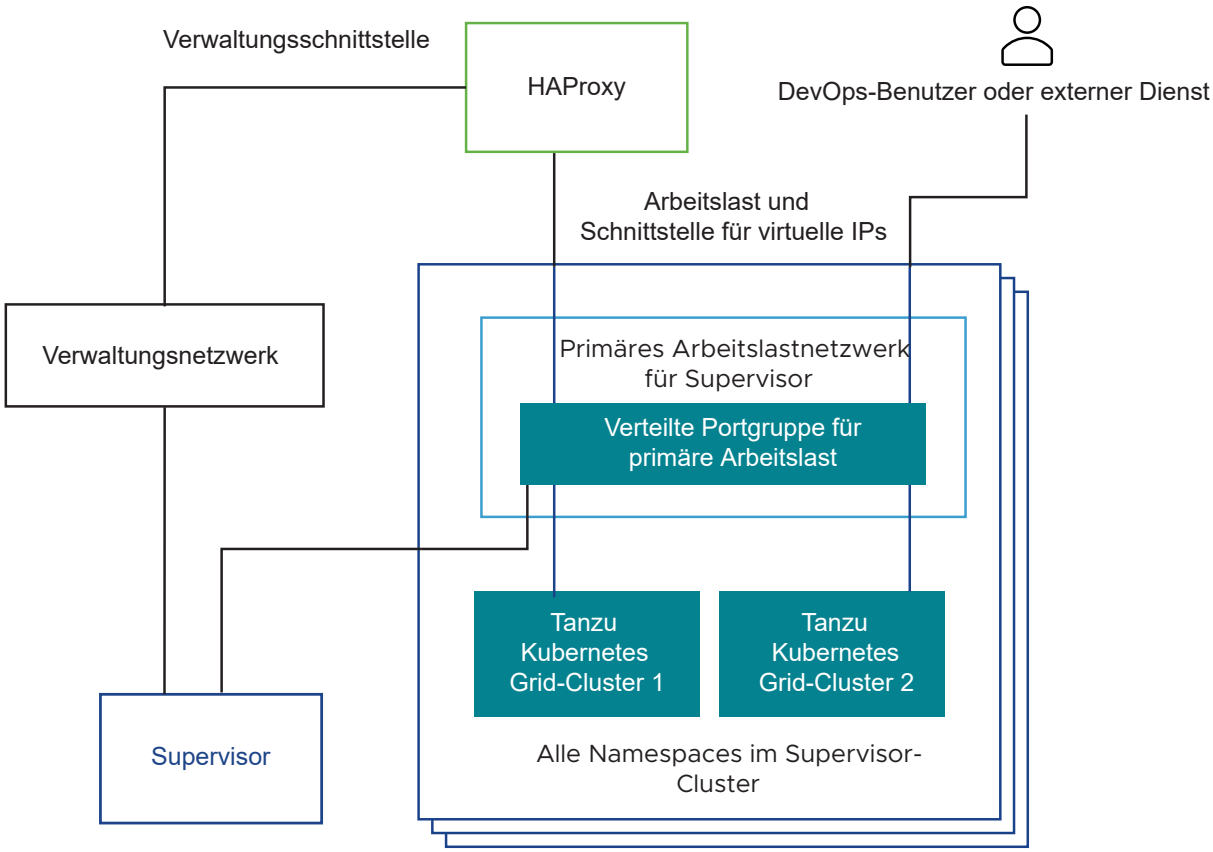
Supervisor-Topologie mit einem Arbeitslastnetzwerk und HAProxy mit zwei virtuellen Netzwerkkarten

In dieser Topologie konfigurieren Sie einen Supervisor mit einem Arbeitslastnetzwerk für die folgenden Komponenten:

- Kubernetes-Steuerungsebenen-VMs
- Die Knoten von Tanzu Kubernetes Grid-Clustern.
- Der virtuelle HAProxy-Bereich für Verbindungen externer Dienste und DevOps-Benutzer. In dieser Konfiguration wird HAProxy mit zwei virtuellen NICs (**Standard**-Konfiguration) bereitgestellt, wobei eine mit dem Verwaltungsnetzwerk und eine zweite mit dem primären Arbeitslastnetzwerk verbunden ist. Sie müssen die Zuteilung virtueller IPs in einem Subnetz planen, das vom primären Arbeitslastnetzwerk getrennt ist.

Sie legen eine Portgruppe als primäres Arbeitslastnetzwerk für den Supervisor fest und verwenden dann dieselbe Portgruppe als Arbeitslastnetzwerk für vSphere-Namespaces. Supervisor, Tanzu Kubernetes Grid-Cluster, HAProxy, DevOps-Benutzer und externe Dienste stellen alle eine Verbindung mit derselben verteilten Portgruppe her, die als primäres Arbeitslastnetzwerk festgelegt ist.

Abbildung 4-6. Von einem Netzwerk gestützter Supervisor



Der Datenverkehrspfad für die DevOps-Benutzer oder externe Anwendungen lautet wie folgt:

1 Der DevOps-Benutzer oder der externe Dienst sendet Datenverkehr an eine virtuelle IP-Adresse in einem Subnetz des Arbeitslastnetzwerks der verteilten Portgruppe.

2 HAProxy gleicht die Last des virtuellen IP-Datenverkehrs aus, der zur IP des Tanzu Kubernetes Grid-Clusterknotens oder zur IP der Steuerungsebenen-VM fließt. HAProxy beansprucht die virtuelle IP-Adresse, damit er die Last des auf dieser IP eingehenden Datenverkehrs ausgleichen kann.

3 Die Steuerungsebenen-VM oder der Tanzu Kubernetes Grid-Clusterknoten leitet den Datenverkehr an die Ziel-Pods weiter, die jeweils innerhalb des Supervisors oder des Tanzu Kubernetes Grid-Clusters ausgeführt werden.

Supervisor-Topologie mit einem isolierten Arbeitslastnetzwerk und einem HAProxy mit zwei virtuellen Netzwerkkarten

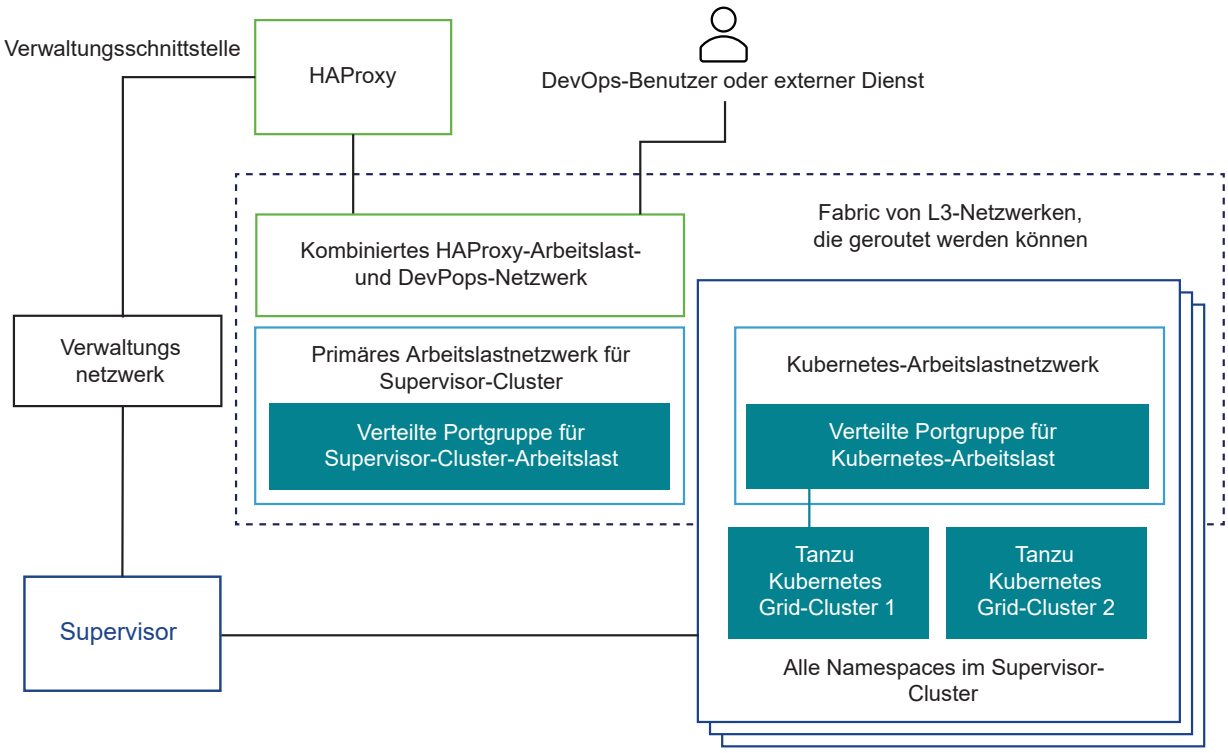
In dieser Topologie konfigurieren Sie Netzwerke für die folgenden Komponenten:

- Kubernetes-Steuerungsebenen-VMs. Ein primäres Arbeitslastnetzwerk für die Verarbeitung des Datenverkehrs für die Kubernetes-Steuerungsebenen-VMs.

- Tanzu Kubernetes Grid-Clusterknoten Ein Arbeitslastnetzwerk, das Sie allen Namespaces auf dem Supervisor zuweisen. Dieses Netzwerk verbindet die Tanzu Kubernetes Grid-Clusterknoten.
- Virtuelle HAProxy-IPs. In dieser Konfiguration wird die HAProxy-VM mit zwei virtuellen NICs bereitgestellt (**Standard**-Konfiguration). Sie können die HAProxy-VM entweder mit dem primären Arbeitslastnetzwerk oder mit dem Arbeitslastnetzwerk verbinden, das Sie für Namespaces verwenden. Sie können HAProxy auch mit einem VM-Netzwerk verbinden, das bereits in vSphere vorhanden ist und für das Routing in das primäre Arbeitslastnetzwerk und die Arbeitslastnetzwerke möglich ist.

Der Supervisor ist mit der verteilten Portgruppe verbunden, die das primäre Arbeitslastnetzwerk unterstützt, und Tanzu Kubernetes Grid-Cluster sind mit einer verteilten Portgruppe verbunden, die das Arbeitslastnetzwerk unterstützt. Die beiden Portgruppen müssen für Schicht-3-Routing geeignet sein. Sie können die Schicht 2-Isolierung über VLANs implementieren. Die Filterung des Schicht-3-Datenverkehrs ist über IP-Firewalls und -Gateways möglich.

Abbildung 4-7. Supervisor mit einem isolierten Arbeitslastnetzwerk



Der Datenverkehrspfad für DevOps-Benutzer oder den externen Dienst lautet wie folgt:

1 Der DevOps-Benutzer oder der externe Dienst sendet Datenverkehr an eine virtuelle IP. Der Datenverkehr wird an das Netzwerk weitergeleitet, mit dem HAProxy verbunden ist.

2 HAProxy gleicht die Last des virtuellen IP-Datenverkehrs aus, der zur IP des Tanzu Kubernetes Grid-Knotens oder zur Steuerungsebenen-VM fließt. HAProxy beansprucht die virtuelle IP-Adresse, damit er die Last des auf dieser IP eingehenden Datenverkehrs ausgleichen kann.

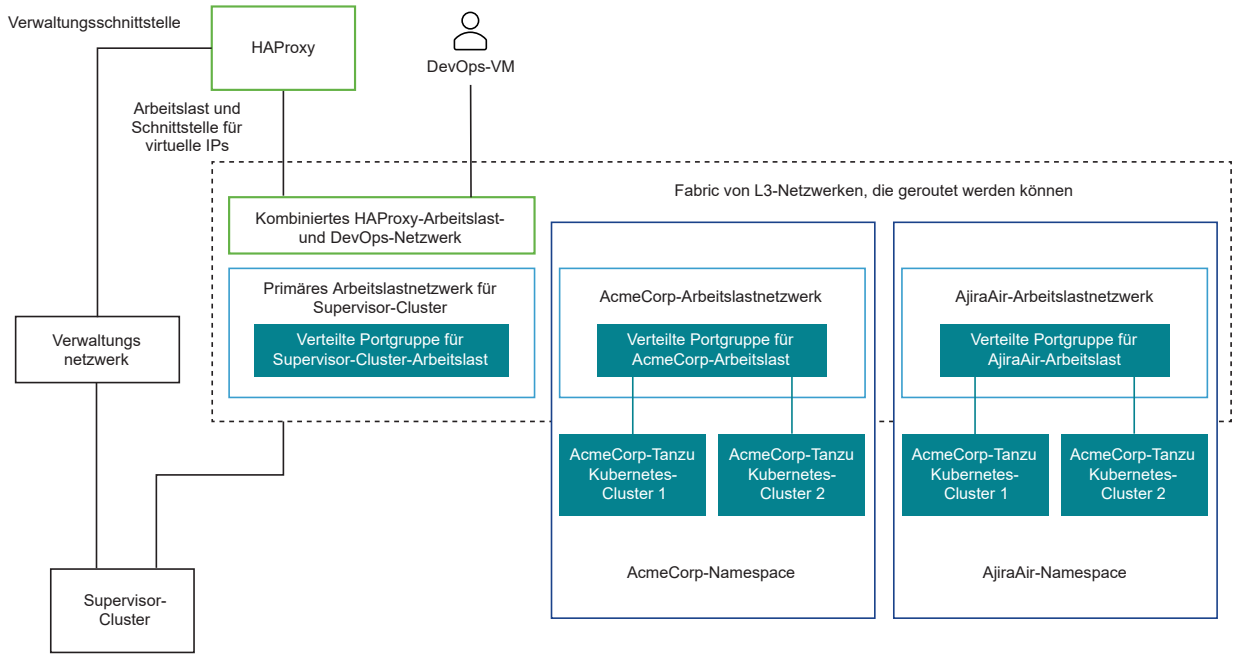
3 Die Steuerungsebenen-VM oder der Tanzu Kubernetes Grid-Clusterknoten leitet den Datenverkehr an die Ziel-Pods weiter, die innerhalb des Tanzu Kubernetes Grid-Clusters ausgeführt werden.

Supervisor-Topologie mit mehreren Arbeitslastnetzwerken und HAProxy mit zwei virtuellen Netzwerkkarten

In dieser Topologie können Sie eine Portgruppe so konfigurieren, dass sie als primäres Arbeitslastnetzwerk agiert, und Sie können eine dedizierte Portgruppe als Arbeitslastnetzwerk für jeden Namespace konfigurieren. HAProxy wird mit zwei virtuellen Netzwerkkarten bereitgestellt (**Standard**-Konfiguration), und Sie können eine Verbindung mit dem primären Arbeitslastnetzwerk oder den anderen Arbeitslastnetzwerken herstellen. Sie können auch ein vorhandenes VM-Netzwerk verwenden, für das Routing in das primäre Arbeitslastnetzwerk und die Arbeitslastnetzwerke möglich ist.

Der Datenverkehrspfad für die DevOps-Benutzer und externe Dienste in dieser Topologie ist identisch mit demjenigen der Topologie mit einem isolierten Arbeitslastnetzwerk.

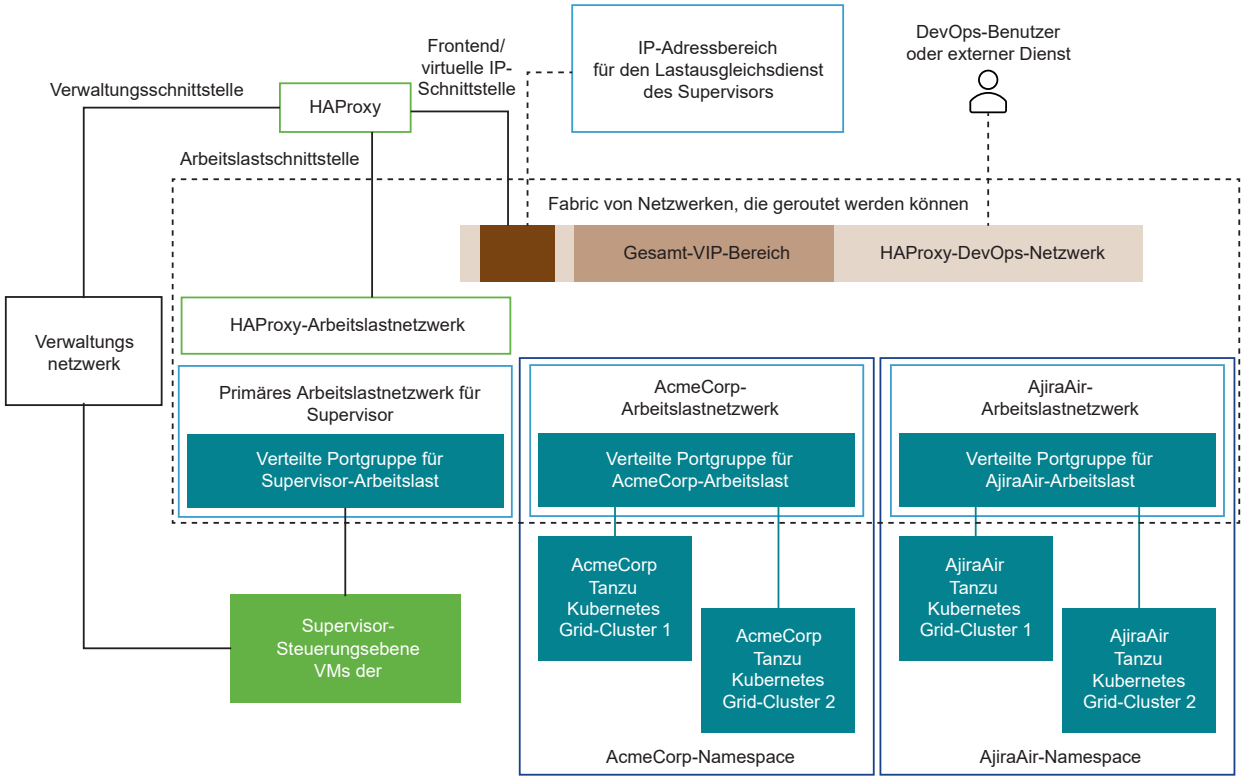
Abbildung 4-8. Von mehreren isolierten Arbeitslastnetzwerken gestützter Supervisor



Supervisor-Topologie mit mehreren Arbeitslastnetzwerken und HAProxy mit drei virtuellen Netzwerkkarten

In dieser Konfiguration stellen Sie die HAProxy-VM mit drei virtuellen NICs bereit und verbinden damit HAProxy mit einem Frontend-Netzwerk. DevOps Benutzer und externe Dienste können über virtuelle IPs im Frontend-Netzwerk auf HAProxy zugreifen. Die Bereitstellung von HAProxy mit drei virtuellen Netzwerkkarten wird für Produktionsumgebungen empfohlen.

Abbildung 4-9. HAProxy-Bereitstellung mit drei virtuellen NICs



Auswählen zwischen den möglichen Topologien

Bevor Sie zwischen den möglichen Topologien auswählen, analysieren Sie die Anforderungen Ihrer Umgebung:

1 Benötigen Sie eine Schicht-2-Isolierung zwischen dem Supervisor und Tanzu Kubernetes Grid-Clustern?

   a Nein: die einfachste Topologie mit einem Arbeitslastnetzwerk, das alle Komponenten bedient.

   b Ja: die isolierte Arbeitslastnetzwerk-Topologie mit getrennten primären und Arbeitslastnetzwerken.

2 Benötigen Sie eine weitere Schicht-2-Isolierung zwischen Ihren Tanzu Kubernetes Grid-Clustern?

   a Nein: isolierte Arbeitslastnetzwerk-Topologie mit getrennten primärem Arbeitslastnetzwerk und Arbeitslastnetzwerken.

   b Ja: mehrere Arbeitslastnetzwerk-Topologien mit einem getrennten Arbeitslastnetzwerk für jeden Namespace und einem dedizierten primären Arbeitslastnetzwerk.

3 Möchten Sie verhindern, dass Ihre DevOps-Benutzer und externe Dienste ein direktes Routing zu VMs der Steuerungsebene von Kubernetes und Tanzu Kubernetes Grid-Clusterknoten ausführen?

   a Nein: HAProxy-Konfiguration mit zwei NICs.

   b Ja: HAProxy-Konfiguration mit drei NICs Diese Konfiguration wird für Produktionsumgebungen empfohlen.

Überlegungen zur Verwendung des HAProxy-Lastausgleichsdiensts mit vSphere IaaS control plane

Beachten Sie beim Planen eines vSphere IaaS control plane mit dem HAProxy-Lastausgleichsdienst Folgendes.

- Ein Support-Vertrag mit HAProxy ist erforderlich, um technischen Support für den HAProxy-Lastausgleichsdienst zu erhalten. VMware GSS kann für die HAProxy-Appliance keinen Support bieten.
- Die HAProxy-Appliance ist ein Singleton ohne Möglichkeit für eine hochverfügbare Topologie. Für hochverfügbare Umgebungen empfiehlt VMware die Verwendung einer vollständigen Installation von NSX oder den NSX Advanced Load Balancer.
- Es ist nicht möglich, den für das Front-End verwendeten IP-Adressbereich zu einem späteren Zeitpunkt zu erweitern. Dies bedeutet, dass das Netzwerk für das gesamte zukünftige Wachstum dimensioniert werden sollte.

Voraussetzungen für die zonale Supervisor-Bereitstellung

5

Informieren Sie sich über die Voraussetzungen zum Aktivieren eines Supervisors auf vSphere-Zonen. Ein auf vSphere-Zonen aktivierter Supervisor bietet Hochverfügbarkeit für Ihre Kubernetes-Arbeitslasten auf vSphere-Clusterebene.

Hinweis Wenn Sie Ihre vSphere IaaS control plane-Umgebung von einer vSphere-Version vor 8.0 aktualisiert haben und vSphere Zonen für Ihre Bereitstellungen wie Tanzu Kubernetes Grid-Cluster verwenden möchten, müssen Sie einen neuen Supervisor mit drei Zonen erstellen. vSphere IaaS control plane unterstützt die Konvertierung des Supervisors von einem Supervisor mit einem einzelnen Cluster in einen mit drei Zonen nicht.

Lesen Sie als Nächstes die folgenden Themen:

- Voraussetzungen für die Zonen-Supervisor-Bereitstellung mit NSX Advanced Load Balancer und VDS-Netzwerk
- Voraussetzungen für zonalen Supervisor mit NSX
- Voraussetzungen für zonalen Supervisor mit NSX und NSX Advanced Load Balancer
- Voraussetzungen für die zonale Supervisor-Bereitstellung mit HAProxy-Lastausgleichsdienst

Voraussetzungen für die Zonen-Supervisor-Bereitstellung mit NSX Advanced Load Balancer und VDS-Netzwerk

Informieren Sie sich über die Voraussetzungen zum Aktivieren eines Supervisors mit VDS-Netzwerk und NSX Advanced Load Balancer auf drei vSphere-Clustern, die drei vSphere-Zonen zugeordnet sind. Um vSphere IaaS control plane mit dem NSX Advanced Load Balancer, auch bekannt als Avi-Load Balancer, zu konfigurieren, muss Ihre Umgebung bestimmte Anforderungen erfüllen. vSphere IaaS control plane unterstützt mehrere Topologien: ein einzelnes VDS-Netzwerk für die Avi-Dienst-Engine und die Load Balancer-Dienste, einen VDS für die Avi Management Plane und einen weiteren VDS für den NSX Advanced Load Balancer.

Arbeitslastnetzwerke

Zum Konfigurieren eines Supervisors mit dem VDS-Netzwerk-Stack müssen Sie alle Hosts aus dem Cluster mit einem VDS verbinden. Je nach der von Ihnen für den Supervisor implementierten Topologie erstellen Sie eine oder mehrere verteilte Portgruppen. Sie bestimmen die Portgruppen als Arbeitslastnetzwerke für vSphere-Namespaces.

Arbeitslastnetzwerke bieten eine Verbindung zu den Knoten von Tanzu Kubernetes Grid-Clustern, zu VMs, die über den VM-Dienst erstellt wurden, sowie zu den VMs der Supervisor-Steuerungsebene. Das Arbeitslastnetzwerk, das Konnektivität zu den Kubernetes-Steuerungsebenen-VMs bereitstellt, wird als „primäres Arbeitslastnetzwerk“ bezeichnet. Jeder Supervisor muss über ein primäres Arbeitslastnetzwerk verfügen. Sie müssen eine der verteilten Portgruppen als primäres Arbeitslastnetzwerk für den Supervisor festlegen.

Die Kubernetes-Steuerungsebenen-VMs im Supervisor verwenden drei IP-Adressen aus dem IP-Adressbereich, der dem primären Arbeitslastnetzwerk zugewiesen ist. Jeder Knoten eines Tanzu Kubernetes Grid-Clusters verfügt über eine eigene, aus dem Adressbereich des Arbeitslastnetzwerks zugewiesene IP-Adresse. Das Arbeitslastnetzwerk ist mit dem Namespace konfiguriert ist, in dem der Tanzu Kubernetes Grid-Cluster ausgeführt wird.

Netzwerkanforderungen

Der NSX Advanced Load Balancer erfordert zwei routingfähige Subnetze:

- Das Verwaltungsnetzwerk. Der AVI-Controller, auch als Controller bezeichnet, befindet sich im Verwaltungsnetzwerk. Das Verwaltungsnetzwerk stellt dem Controller Konnektivität zum vCenter Server, zu den ESXi-Hosts und den Supervisor-Knoten der Steuerungsebene bereit. In diesem Netzwerk befindet sich die Verwaltungsschnittstelle der AVI-Dienst-Engine. Für dieses Netzwerk sind ein VDS und eine verteilte Portgruppe erforderlich.
- Das Datennetzwerk. Die Datenschnittstelle der AVI-Dienst-Engines, auch als Dienst-Engines bezeichnet, stellt eine Verbindung zu diesem Netzwerk her. Die virtuellen IPs (VIPs) des Lastausgleichsdiensts werden von diesem Netzwerk zugewiesen. Für dieses Netzwerk sind ein VDS und verteilte Portgruppen erforderlich. Sie müssen den vDS und die Portgruppen konfigurieren, bevor Sie den Lastausgleichsdienst installieren.

Zuteilung der IP-Adressen

Der Controller und die Dienst-Engine sind mit dem Verwaltungsnetzwerk verbunden. Wenn Sie den virtuellen NSX Advanced Load Balancer installieren und konfigurieren, stellen Sie für jede Controller-VM eine statische, routingfähige IP-Adresse bereit.

Die Dienst-Engines können DHCP verwenden. Wenn DHCP nicht verfügbar ist, können Sie einen Pool von IP-Adressen für die Dienst-Engines konfigurieren.

Platzierung von vSphere-Zonen über physische Sites hinweg

Sie können vSphere-Zonen auf verschiedene physische Sites verteilen, solange die Latenz zwischen den Sites 100 ms nicht überschreitet. Sie können beispielsweise die vSphere-Zonen auf zwei physische Sites verteilen: eine vSphere-Zone auf der ersten Site und zwei vSphere-Zonen auf der zweiten Site.

Computing-Mindestanforderungen für Testzwecke

Wenn Sie die Funktionen des vSphere IaaS control plane testen möchten, können Sie die Plattform auf einem sehr minimalen Testbed bereitstellen. Sie sollten sich sicherstellen, dass ein solches Testbed nicht für die Ausführung von Arbeitslasten im Produktionsumfang geeignet ist und keine HA auf Clusterebene bereitstellt.

Tabelle 5-1. Computing-Mindestanforderungen für Testzwecke

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8.0	Klein	2	21 GB	290 GB
vSphere-Cluster	■ 3 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf jedem vSphere-Cluster. vSphere DRS muss sich im vollautomatischen oder teilweise automatisierten Modus befinden. ■ Unabhängige Speicher- und Netzwerkkonfiguration für jeden vSphere-Cluster.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8.0	Bei jedem vSphere-Cluster: ■ Ohne vSAN: 1 ESXi-Host mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 2 ESXi-Hosts pro Cluster mit mindestens 2 physischen NICs. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die dem Cluster beitreten, Kleinbuchstaben verwenden. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8 pro Host	64 GB pro Host	Nicht anwendbar

Tabelle 5-1. Computing-Mindestanforderungen für Testzwecke (Fortsetzung)

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB
NSX Advanced Load Balancer-Controller	Enterprise	4 (Klein) 8 (Mittel) 24 (Groß)	12 GB 24 GB 128 GB	128 GB 128 GB 128 GB

Computing-Mindestanforderungen für die Produktion

In der Tabelle sind die Computing-Mindestanforderungen bei Aktivierung eines Supervisors mit VDS-Netzwerk und NSX Advanced Load Balancer in drei vSphere-Zonen aufgeführt. Es wird empfohlen, die Verwaltungs- und die Arbeitslastdomäne zu trennen. Die Arbeitslastdomäne hostet den Supervisor, in dem Sie Arbeitslasten ausführen. Die Verwaltungsdomäne hostet alle Verwaltungskomponenten wie vCenter Server.

Tabelle 5-2. Mindestanforderungen für Computing

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8.0	Klein	2	21 GB	290 GB
vSphere-Cluster	■ 3 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf jedem vSphere-Cluster. vSphere DRS muss sich im vollautomatischen oder teilweise automatisierten Modus befinden. ■ Unabhängige Speicher- und Netzwerkkonfiguration für jeden vSphere-Cluster.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8.0	Bei jedem vSphere-Cluster: ■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 4 ESXi-Hosts pro Cluster mit mindestens 2 physischen Netzwerkkarten. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die dem Cluster beitreten, Kleinbuchstaben verwenden. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8 pro Host	64 GB pro Host	Nicht anwendbar

Tabelle 5-2. Mindestanforderungen für Computing (Fortsetzung)

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB
NSX Advanced Load Balancer-Controller	Enterprise Bei Produktionsumgebungen wird empfohlen, einen Cluster mit 3 Controller-VMs zu installieren. Für HA sind mindestens 2 Dienst-Engine-VMs erforderlich.	4 (Klein) 8 (Mittel) 24 (Groß)	12 GB 24 GB 128 GB	128 GB 128 GB 128 GB

Mindestanforderungen für das Netzwerk

In der Tabelle sind die Mindestnetzwerkanforderungen bei Aktivierung eines Supervisors mit VDS-Netzwerk und NSX Advanced Load Balancer aufgeführt.

Tabelle 5-3. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Layer-2-Gerät	1	Das Verwaltungsnetzwerk, das den Supervisor-Datenverkehr verarbeitet, muss sich für alle Cluster, die zum Supervisor gehören, auf demselben Layer-2-Gerät befinden. Das primäre Arbeitslastnetzwerk muss sich ebenfalls auf demselben Layer-2-Gerät befinden.
Physische Netzwerk-MTU	1500	Die MTU-Größe muss für jede verteilte Portgruppe mindestens 1500 betragen.

Tabelle 5-4. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
Latenz	100 ms	Die maximal empfohlene Latenz zwischen jedem Cluster, der Teil einer vSphere-Zone ist, die in einem Supervisor verbunden ist.
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)

Tabelle 5-5. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.
Verwaltungsnetzwerk-Subnetz	1	Der NSX Advanced Load Balancer-Controller, auch als Controller bezeichnet, befindet sich im Verwaltungsnetzwerk. Hier wird auch die Verwaltungsschnittstelle der Dienst-Engine verbunden. Der Controller muss mit den vCenter Server- und ESXi-Verwaltungs-IPs aus diesem Netzwerk verbunden sein. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.

Tabelle 5-6. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
vSphere Distributed Switch	1	Alle Hosts aus allen drei vSphere-Clustern müssen mit einem VDS verbunden sein.
Arbeitslastnetzwerke	1	Mindestens eine verteilte Portgruppe muss auf dem VDS erstellt werden, den Sie als primäres Arbeitslastnetzwerk konfigurieren. Je nach der gewählten Topologie können Sie dieselbe verteilte Portgruppe verwenden wie für das Arbeitslastnetzwerk der Namespaces oder mehrere Portgruppen erstellen und sie als Arbeitslastnetzwerke konfigurieren. Arbeitslastnetzwerke müssen die folgenden Anforderungen erfüllen: ■ Routing-Fähigkeit zwischen einem beliebigen Arbeitslastnetzwerk und dem vom NSX Advanced Load Balancer für die Zuteilung von virtuellen IPs verwendeten Netzwerk. ■ Keine Überschneidung von IP-Adressen über alle Arbeitslastnetzwerke innerhalb eines Supervisors hinweg.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Tabelle 5-7. Netzwerkanforderungen für den Lastausgleichsdienst

NTP- und DNS-Server	1	Die DNS-Server-IP ist erforderlich, damit der NSX Advanced Load Balancer-Controller die vCenter Server- und ESXi-Hostnamen korrekt auflösen kann. NTP ist optional, da öffentliche NTP-Server standardmäßig verwendet werden.
Datennetzwerksubnetz	1	Die Datenschnittstelle der Dienst-Engines, auch als Dienst-Engines bezeichnet, stellt eine Verbindung zu diesem Netzwerk her. Konfigurieren Sie einen Pool von IP-Adressen für die Dienst-Engines. Die virtuellen IPs (VIPs) des Lastausgleichsdiensts werden von diesem Netzwerk zugewiesen.
NSX Advanced Load Balancer-Controller-IPs	1 oder 4	Wenn Sie den NSX Advanced Load Balancer-Controller als einzelnen Knoten bereitstellen, ist eine statische IP-Adresse für seine Verwaltungsschnittstelle erforderlich. Für einen Cluster mit 3 Knoten sind 4 IP-Adressen erforderlich. Eine für jede Controller-VM und eine für die Cluster-VIP. Diese IPs müssen aus dem Subnetz des Verwaltungsnetzwerks stammen.
VIP-IPAM-Bereich	-	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Die IPs müssen aus dem Subnetz des Datennetzwerks stammen. Sie müssen für jeden Supervisor-Cluster einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Ports und Protokolle

Diese Tabelle enthält die Protokolle und Ports, die für die Verwaltung der IP-Konnektivität zwischen dem NSX Advanced Load Balancer, vCenter Server und anderen vSphere IaaS control plane-Komponenten erforderlich sind.

Quelle	Ziel	Protokoll und Ports
NSX Advanced Load Balancer-Controller	NSX Advanced Load Balancer-Controller (im Cluster)	TCP 22 (SSH) TCP 443 (HTTPS) TCP 8443 (HTTPS)
Dienst-Engine	Dienst-Engine in HA	TCP 9001 für VMware, LSC und NSX-T Cloud
Dienst-Engine	NSX Advanced Load Balancer-Controller	TCP 22 (SSH) TCP 8443 (HTTPS) UDP 123 (NTP)
NSX Advanced Load Balancer-Controller	vCenter Server, ESXi, NSX-T Manager	TCP 443 (HTTPS)
Supervisor-Steuerungsebenenknoten (AKO)	NSX Advanced Load Balancer-Controller	TCP 443 (HTTPS)

Weitere Informationen zu Ports und Protokollen für den NSX Advanced Load Balancer finden Sie unter https://ports.esp.vmware.com/home/NSX-Advanced-Load-Balancer.

Voraussetzungen für zonalen Supervisor mit NSX

Informieren Sie sich über die Voraussetzungen für das Aktivieren eines Supervisors auf drei vSphere-Clustern, die vSphere-Zonen zugeordnet sind, unter Verwendung des NSX-Netzwerk-Stacks.

Neben diesen Anforderungen finden Sie im Leitfaden zum NSX-Referenz-Design weitere Informationen zu den Best Practices für die Bereitstellung von NSX.

Platzierung von vSphere-Zonen über physische Sites hinweg

Sie können vSphere-Zonen auf verschiedene physische Sites verteilen, solange die Latenz zwischen den Sites 100 ms nicht überschreitet. Sie können beispielsweise die vSphere-Zonen auf zwei physische Sites verteilen: eine vSphere-Zone auf der ersten Site und zwei vSphere-Zonen auf der zweiten Site.

Mindestberechnungsanforderungen für ein Verwaltungs- und Edge-Cluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8	Klein	2	21 GB	290 GB
ESXi-Hosts 8	2 ESXi-Hosts	8	64 GB pro Host	Nicht anwendbar
NSX Manager	Mittel	6	24 GB	300 GB

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
NSX Edge 1	Groß	8	32 GB	200 GB
NSX Edge 2	Groß	8	32 GB	200 GB

Hinweis Stellen Sie sicher, dass alle ESXi-Hosts, die an dem vSphere-Cluster teilnehmen, auf dem Sie vSphere IaaS control plane konfigurieren möchten, als NSX-Transportknoten vorbereitet sind. Weitere Informationen finden Sie unter https://kb.vmware.com/s/article/95820 und Vorbereiten von ESXi-Hosts als Transportknoten in der Dokumentation zu NSX.

Mindestberechnungsanforderungen für Arbeitslastdomänencluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vSphere-Cluster	■ 3 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf jedem vSphere-Cluster. vSphere DRS muss im vollautomatisierten Modus ausgeführt werden. ■ Unabhängige Speicher- und Netzwerkkonfiguration für jeden vSphere-Cluster.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8	Bei jedem vSphere-Cluster: ■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 4 ESXi-Hosts pro Cluster mit mindestens 2 physischen Netzwerkkarten. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die den Clustern beitreten, in Kleinbuchstaben geschrieben sind. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB

Netzwerkanforderungen

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 8-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren.

Überprüfen Sie in der VMware-Produkt-Interoperabilitätstabelle, welche NSX-Versionen unterstützt werden.

Tabelle 5-8. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Layer-2-Gerät	1	Das Verwaltungsnetzwerk, das den Supervisor-Datenverkehr verarbeitet, muss sich auf demselben Layer-2-Gerät befinden. Pro Host muss mindestens eine physische Netzwerkkarte, die den Verwaltungsdatenverkehr verarbeitet, mit demselben Layer-2-Gerät verbunden sein.
Physische Netzwerk-MTU	1500	Die MTU-Größe muss für jede vSphere Distributed Switch-Portgruppe mindestens 1500 betragen.
Physische Netzwerkkarte	Mindestens 2 physische Netzwerkkarten pro Host, wenn vSAN verwendet wird	Um Antrea-CNI zu verwenden und eine optimale NSX-Leistung zu erzielen, muss jede physische Netzwerkkarte auf jedem teilnehmenden ESXi-Host die GENEVE-Kapselung unterstützen und diese aktivieren.

Tabelle 5-9. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
Latenz	100 ms	Die maximal empfohlene Latenz zwischen jedem Cluster, der Teil einer vSphere-Zone ist, die in einem Supervisor verbunden ist.
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.

Tabelle 5-9. Allgemeine Netzwerkanforderungen (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)
Image-Registrierung	1	Zugriff auf eine Registrierung für den Dienst.

Tabelle 5-10. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.

Tabelle 5-10. Anforderungen an das Verwaltungsnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Verwaltungsnetzwerk-Subnetz	1	Das Subnetz, das für den Verwaltungsdatenverkehr zwischen ESXi-Hosts und vCenter Server, NSX-Appliances und der Kubernetes-Steuerungsebene verwendet wird. Das Subnetz muss folgende Größe haben: ■ 1 IP-Adresse pro VMkernel-Adapter eines Hosts. ■ 1 IP-Adresse für die vCenter Server-Appliance. ■ 1 oder 4 IP-Adressen für NSX Manager. 4 IP-Adressen bei NSX Manager-Clustering von 3 Knoten und 1 virtuellen IP (VIP). ■ 5 IP-Adressen für die Kubernetes-Steuerungsebene. 1 für jeden der 3 Knoten, 1 für virtuelle IP, 1 für fortlaufendes Cluster-Upgrade. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.
Verwaltungsnetzwerk-VLAN	1	VLAN-ID des Verwaltungsnetzwerk-Subnetzes.

Tabelle 5-11. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
CIDR-Bereich für vSphere Pods	/23 private IP-Adressen	Ein privater-CIDR-Bereich, der IP-Adressen für vSphere-Pods bereitstellt. Diese Adressen werden auch für die Tanzu Kubernetes Grid-Clusterknoten verwendet. Sie müssen einen eindeutigen vSphere Pod-CIDR-Bereich für jeden Cluster angeben. **Hinweis** Der CIDR-Bereich für vSphere Pods und der CIDR-Bereich für die Kubernetes-Dienstadressen dürfen sich nicht überlappen.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.
Egress-CIDR-Bereich	/27 statische IP-Adressen	Eine CIDR-Anmerkung zur Ermittlung der Egress-IP für Kubernetes-Dienste. Für jeden Namespace im Supervisor wird nur eine Egress-IP-Adresse zugewiesen. Die Egress-IP ist die Adresse, die externe Entitäten für die Kommunikation mit den Diensten im Namespace verwenden. Die Anzahl der Egress-IP-Adressen beschränkt die Anzahl der Egress-Richtlinien, die der Supervisor haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. Beispielsweise `10.174.4.96/27` **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.

Tabelle 5-11. Anforderungen an das Arbeitslastnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Ingress-CIDR	/27 statische IP-Adressen	Ein privater CIDR-Bereich, der für Ingress-IP-Adressen verwendet wird. Mit Ingress können Sie Datenverkehrsrichtlinien auf Anforderungen anwenden, die in den Supervisor von externen Netzwerken eingehen. Die Anzahl der Ingress-IP-Adressen beschränkt die Anzahl der Ingresses, die der Cluster haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.
Namespace-Netzwerkbereich	1	Ein oder mehrere IP-CIDRs zum Erstellen von Subnetzen/Segmenten sowie zum Zuweisen von IP-Adressen an Arbeitslasten.
Namespace-Subnetzpräfix	1	Das Subnetzpräfix, das die Größe des für Namespace-Segmente reservierten Subnetzes angibt. Der Standardwert ist „28“.

Tabelle 5-12. NSX-Anforderungen

Komponente	Mindestanzahl	V
VLANs	3	Diese VLAN-IPs sind die IP-Adressen für die Tunnel-Endpoints (TEP). Die TEPs des ESXi-Hosts und die Edge-TEPs müssen routingfähig sein. VLAN-IP-Adressen sind für Folgendes erforderlich: ■ ESXi-Host-VTEP ■ Edge-VTEP mit statischer IP-Adresse ■ Tier-0-Gateway und Uplink für Transportknoten. **Hinweis** Der ESXi-Host-VTEP und der Edge-VTEP müssen eine MTU-Größe von mehr als 1600 aufweisen. ESXi-Hosts und NSX-T-Edge-Knoten fungieren als Tunnel-Endpoints, und jedem Host sowie jedem Edge-Knoten wird eine TEP-IP zugewiesen. Da die TEP-IPs für ESXi-Hosts einen Overlay-Tunnel mit TEP-IPs auf den Edge-Knoten erstellen, müssen die VLAN-IPs routingfähig sein. Ein zusätzliches VLAN ist erforderlich, um Nord-Süd-Konnektivität zum Tier-0-Gateway zu bieten. IP-Pools können über Cluster hinweg gemeinsam genutzt werden. Der Host-Overlay-IP-Pool/VLAN darf jedoch nicht mit einem Edge-Overlay-IP-Pool/VLAN gemeinsam genutzt werden. **Hinweis** Wenn Host-TEP und Edge-TEP verschiedene physische Netzwerkkarten verwenden, können sie dasselbe VLAN verwenden.
Tier-0-Uplink-IP	/24 private IP-Adressen	Das IP-Subnetz, das für den Tier-0-Uplink verwendet wird. Für die IP-Adresse des Tier-0-Uplinks gelten folgende Anforderungen: ■ 1 IP, wenn Sie keine Edge-Redundanz verwenden. ■ 4 IPs, wenn Sie BGP und Edge-Redundanz verwenden – 2 IP-Adressen pro Edge. ■ 3 IPs, wenn Sie statische Routen und Edge-Redundanz verwenden. Die IP, das Subnetz und das Gateway für die Edge-Verwaltung und die IP, das Subnetz und das Gateway für Uplink müssen eindeutig sein.

Voraussetzungen für zonalen Supervisor mit NSX und NSX Advanced Load Balancer

Informieren Sie sich über die Voraussetzungen für das Aktivieren eines Supervisors auf drei vSphere-Clustern, die vSphere-Zonen zugeordnet sind, unter Verwendung des NSX-Netzwerk-Stacks und NSX Advanced Load Balancer.

Platzierung von vSphere-Zonen über physische Sites hinweg

Sie können vSphere-Zonen auf verschiedene physische Sites verteilen, solange die Latenz zwischen den Sites 100 ms nicht überschreitet. Sie können beispielsweise die vSphere-Zonen auf zwei physische Sites verteilen: eine vSphere-Zone auf der ersten Site und zwei vSphere-Zonen auf der zweiten Site.

NSX-Bereitstellungsoptionen

Weitere Informationen zu den Best Practices für die Bereitstellung von NSX finden Sie im Leitfaden zum NSX Referenz-Design.

Mindestberechnungsanforderungen für ein Verwaltungs- und Edge-Cluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8	Klein	2	21 GB	290 GB
ESXi-Hosts 8	2 ESXi-Hosts	8	64 GB pro Host	Nicht anwendbar
NSX Manager	Mittel	6	24 GB	300 GB
NSX Edge 1	Groß	8	32 GB	200 GB
NSX Edge 2	Groß	8	32 GB	200 GB
Dienst-Engine-VMs	Pro Supervisor werden mindestens zwei Dienst-Engine-VMs bereitgestellt	1	2 GB	Nicht verfügbar

Hinweis Stellen Sie sicher, dass alle ESXi-Hosts, die an dem vSphere-Cluster teilnehmen, auf dem Sie vSphere IaaS control plane konfigurieren möchten, als NSX-Transportknoten vorbereitet sind. Weitere Informationen finden Sie unter https://kb.vmware.com/s/article/95820 und Vorbereiten von ESXi-Hosts als Transportknoten in der Dokumentation zu NSX.

Angeben der Systemkapazität des Controllers

Sie können die Systemkapazität des Controllers während der Bereitstellung angeben. Die Systemkapazität basiert auf Zuteilungen von Systemressourcen wie CPU, RAM und Festplatte. Die Menge an Ressourcen, die Sie zuteilen, wirkt sich auf die Leistung des Controllers aus.

Bereitstellungstyp	Knotenanzahl	Empfohlene Zuteilungen – CPU	Empfohlene Zuteilungen – Arbeitsspeicher	Empfohlene Zuteilungen – Festplatte
Demo/ Kundenbewertung	1	6	24 GB	128 GB

Bei Demobereitstellungen reicht ein einzelner Controller aus und wird für alle Aktivitäten und Workflows der Steuerungsebene sowie für Analysen verwendet.

In einer Produktionsbereitstellung wird ein Cluster mit drei Knoten empfohlen.

Weitere Informationen finden Sie unter NSX Advanced Load Balancer Controller-Größenanpassung.

Mindestberechnungsanforderungen für Arbeitslastdomänencluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vSphere-Cluster	■ 3 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf jedem vSphere-Cluster. vSphere DRS muss im vollautomatisierten Modus ausgeführt werden. ■ Unabhängige Speicher- und Netzwerkkonfiguration für jeden vSphere-Cluster.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8	Bei jedem vSphere-Cluster: ■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 4 ESXi-Hosts pro Cluster mit mindestens 2 physischen Netzwerkkarten. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die den Clustern beitreten, in Kleinbuchstaben geschrieben sind. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB

Netzwerkanforderungen

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 8-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren.

Überprüfen Sie in der VMware-Produkt-Interoperabilitätstabelle, welche NSX-Versionen unterstützt werden.

Tabelle 5-13. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Layer-2-Gerät	1	Das Verwaltungsnetzwerk, das den Supervisor-Datenverkehr verarbeitet, muss sich auf demselben Layer-2-Gerät befinden. Pro Host muss mindestens eine physische Netzwerkkarte, die den Verwaltungsdatenverkehr verarbeitet, mit demselben Layer-2-Gerät verbunden sein.
Physische Netzwerk-MTU	1700	Die MTU-Größe muss für jede vSphere Distributed Switch-Portgruppe mindestens 1700 betragen.
Physische Netzwerkkarte	Mindestens 2 physische Netzwerkkarten pro Host, wenn vSAN verwendet wird	Um Antrea-CNI zu verwenden und eine optimale NSX-Leistung zu erzielen, muss jede physische Netzwerkkarte auf jedem teilnehmenden ESXi-Host die GENEVE-Kapselung unterstützen und diese aktivieren.

Tabelle 5-14. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
Latenz	100 ms	Die maximal empfohlene Latenz zwischen jedem Cluster, der Teil einer vSphere-Zone ist, die in einem Supervisor verbunden ist.
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.

Tabelle 5-14. Allgemeine Netzwerkanforderungen (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)
Image-Registrierung	1	Zugriff auf eine Registrierung für den Dienst.

Tabelle 5-15. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.

Tabelle 5-15. Anforderungen an das Verwaltungsnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Verwaltungsnetzwerk-Subnetz	1	Das Subnetz, das für den Verwaltungsdatenverkehr zwischen ESXi-Hosts und vCenter Server, NSX-Appliances und der Kubernetes-Steuerungsebene verwendet wird. Das Subnetz muss folgende Größe haben: ■ 1 IP-Adresse pro VMkernel-Adapter eines Hosts. ■ 1 IP-Adresse für die vCenter Server-Appliance. ■ 1 oder 4 IP-Adressen für NSX Manager. 4 IP-Adressen bei NSX Manager-Clustering von 3 Knoten und 1 virtuellen IP (VIP). ■ 5 IP-Adressen für die Kubernetes-Steuerungsebene. 1 für jeden der 3 Knoten, 1 für virtuelle IP, 1 für fortlaufendes Cluster-Upgrade. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.
Verwaltungsnetzwerk-VLAN	1	VLAN-ID des Verwaltungsnetzwerk-Subnetzes.

Tabelle 5-16. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
CIDR-Bereich für vSphere Pods	/23 private IP-Adressen	Ein privater-CIDR-Bereich, der IP-Adressen für vSphere-Pods bereitstellt. Diese Adressen werden auch für die Tanzu Kubernetes Grid-Clusterknoten verwendet. Sie müssen einen eindeutigen vSphere Pod-CIDR-Bereich für jeden Cluster angeben. **Hinweis** Der CIDR-Bereich für vSphere Pods und der CIDR-Bereich für die Kubernetes-Dienstadressen dürfen sich nicht überlappen.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.
Egress-CIDR-Bereich	/27 statische IP-Adressen	Eine CIDR-Anmerkung zur Ermittlung der Egress-IP für Kubernetes-Dienste. Für jeden Namespace im Supervisor wird nur eine Egress-IP-Adresse zugewiesen. Die Egress-IP ist die Adresse, die externe Entitäten für die Kommunikation mit den Diensten im Namespace verwenden. Die Anzahl der Egress-IP-Adressen beschränkt die Anzahl der Egress-Richtlinien, die der Supervisor haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. Beispielsweise `10.174.4.96/27` **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.

Tabelle 5-16. Anforderungen an das Arbeitslastnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Ingress-CIDR	/27 statische IP-Adressen	Ein privater CIDR-Bereich, der für Ingress-IP-Adressen verwendet wird. Mit Ingress können Sie Datenverkehrsrichtlinien auf Anforderungen anwenden, die in den Supervisor von externen Netzwerken eingehen. Die Anzahl der Ingress-IP-Adressen beschränkt die Anzahl der Ingresses, die der Cluster haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.
Namespace-Netzwerkbereich	1	Ein oder mehrere IP-CIDRs zum Erstellen von Subnetzen/Segmenten sowie zum Zuweisen von IP-Adressen an Arbeitslasten.
Namespace-Subnetzpräfix	1	Das Subnetzpräfix, das die Größe des für Namespace-Segmente reservierten Subnetzes angibt. Der Standardwert ist „28“.

Tabelle 5-17. NSX-Anforderungen

Komponente	Mindestanzahl	V
VLANs	3	Diese VLAN-IPs sind die IP-Adressen für die Tunnel-Endpoints (TEP). Die TEPs des ESXi-Hosts und die Edge-TEPs müssen routingfähig sein. VLAN-IP-Adressen sind für Folgendes erforderlich: ■ ESXi-Host-VTEP ■ Edge-VTEP mit statischer IP-Adresse ■ Tier-0-Gateway und Uplink für Transportknoten. **Hinweis** Der ESXi-Host-VTEP und der Edge-VTEP müssen eine MTU-Größe von mehr als 1600 aufweisen. ESXi-Hosts und NSX-T-Edge-Knoten fungieren als Tunnel-Endpoints, und jedem Host sowie jedem Edge-Knoten wird eine TEP-IP zugewiesen. Da die TEP-IPs für ESXi-Hosts einen Overlay-Tunnel mit TEP-IPs auf den Edge-Knoten erstellen, müssen die VLAN-IPs routingfähig sein. Ein zusätzliches VLAN ist erforderlich, um Nord-Süd-Konnektivität zum Tier-0-Gateway zu bieten. IP-Pools können über Cluster hinweg gemeinsam genutzt werden. Der Host-Overlay-IP-Pool/VLAN darf jedoch nicht mit einem Edge-Overlay-IP-Pool/VLAN gemeinsam genutzt werden. **Hinweis** Wenn Host-TEP und Edge-TEP verschiedene physische Netzwerkkarten verwenden, können sie dasselbe VLAN verwenden.
Tier-0-Uplink-IP	/24 private IP-Adressen	Das IP-Subnetz, das für den Tier-0-Uplink verwendet wird. Für die IP-Adresse des Tier-0-Uplinks gelten folgende Anforderungen: ■ 1 IP, wenn Sie keine Edge-Redundanz verwenden. ■ 4 IPs, wenn Sie BGP und Edge-Redundanz verwenden – 2 IP-Adressen pro Edge. ■ 3 IPs, wenn Sie statische Routen und Edge-Redundanz verwenden. Die IP, das Subnetz und das Gateway für die Edge-Verwaltung und die IP, das Subnetz und das Gateway für Uplink müssen eindeutig sein.

Tabelle 5-18. Netzwerkanforderungen für den Lastausgleichsdienst

NTP- und DNS-Server	1	Die DNS-Server-IP ist erforderlich, damit der NSX Advanced Load Balancer-Controller die vCenter Server- und ESXi-Hostnamen korrekt auflösen kann. NTP ist optional, da öffentliche NTP-Server standardmäßig verwendet werden.
Datennetzwerksubnetz	1	Die Datenschnittstelle der Dienst-Engines, auch als Dienst-Engines bezeichnet, stellt eine Verbindung zu diesem Netzwerk her. Konfigurieren Sie einen Pool von IP-Adressen für die Dienst-Engines. Die virtuellen IPs (VIPs) des Lastausgleichsdiensts werden von diesem Netzwerk zugewiesen.
NSX Advanced Load Balancer-Controller-IPs	1 oder 4	Wenn Sie den NSX Advanced Load Balancer-Controller als einzelnen Knoten bereitstellen, ist eine statische IP-Adresse für seine Verwaltungsschnittstelle erforderlich. Für einen Cluster mit 3 Knoten sind 4 IP-Adressen erforderlich. Eine für jede Controller-VM und eine für die Cluster-VIP. Diese IPs müssen aus dem Subnetz des Verwaltungsnetzwerks stammen.
VIP-IPAM-Bereich	-	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Die IPs müssen aus dem Subnetz des Datennetzwerks stammen. Sie müssen für jeden Supervisor-Cluster einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Ports und Protokolle

Diese Tabelle enthält die Protokolle und Ports, die für die Verwaltung der IP-Konnektivität zwischen dem NSX Advanced Load Balancer, vCenter Server und anderen vSphere IaaS control plane-Komponenten erforderlich sind.

Quelle	Ziel	Protokoll und Ports
NSX Advanced Load Balancer-Controller	NSX Advanced Load Balancer-Controller (im Cluster)	TCP 22 (SSH) TCP 443 (HTTPS) TCP 8443 (HTTPS)
Dienst-Engine	Dienst-Engine in HA	TCP 9001 für VMware, LSC und NSX-T Cloud
Dienst-Engine	NSX Advanced Load Balancer-Controller	TCP 22 (SSH) TCP 8443 (HTTPS) UDP 123 (NTP)
NSX Advanced Load Balancer-Controller	vCenter Server, ESXi, NSX-T Manager	TCP 443 (HTTPS)
Supervisor-Steuerungsebenenknoten (AKO)	NSX Advanced Load Balancer-Controller	TCP 443 (HTTPS)

Weitere Informationen zu Ports und Protokollen für den NSX Advanced Load Balancer finden Sie unter https://ports.esp.vmware.com/home/NSX-Advanced-Load-Balancer.

Voraussetzungen für die zonale Supervisor-Bereitstellung mit HAProxy-Lastausgleichsdienst

Informieren Sie sich über die Voraussetzungen zum Aktivieren eines Supervisors mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst auf drei vSphere-Clustern, die vSphere-Zonen zugeordnet sind.

Platzierung von vSphere-Zonen über physische Sites hinweg

Sie können vSphere-Zonen auf verschiedene physische Sites verteilen, solange die Latenz zwischen den Sites 100 ms nicht überschreitet. Sie können beispielsweise die vSphere-Zonen auf zwei physische Sites verteilen: eine vSphere-Zone auf der ersten Site und zwei vSphere-Zonen auf der zweiten Site.

Mindestanforderungen für Computing

In der Tabelle sind die Computing-Mindestanforderungen bei Aktivierung eines Supervisors mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst in drei vSphere-Zonen aufgeführt. Es wird empfohlen, die Verwaltungs- und die Arbeitslastdomäne zu trennen. Die Arbeitslastdomäne hostet die Supervisor, in der Sie Ihre Arbeitslasten ausführen. Die Verwaltungsdomäne hostet alle Verwaltungskomponenten wie vCenter Server.

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8.0	Klein	2	21 GB	290 GB
vSphere-Cluster	■ 3 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf jedem vSphere-Cluster. vSphere DRS muss sich im vollautomatischen oder teilweise automatisierten Modus befinden. ■ Unabhängige Speicher- und Netzwerkkonfiguration für jeden vSphere-Cluster.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8.0	Bei jedem vSphere-Cluster: ■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host ■ Mit vSAN: 4 ESXi-Hosts pro Cluster mit mindestens 2 physischen Netzwerkkarten. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die den Clustern beitreten, in Kleinbuchstaben geschrieben sind. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB

Mindestanforderungen für das Netzwerk

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 8-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren.

Tabelle 5-19. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Layer-2-Gerät	1	Das Verwaltungsnetzwerk, das den Supervisor-Datenverkehr verarbeitet, muss sich für alle Cluster, die zum Supervisor gehören, auf demselben Layer-2-Gerät befinden. Das primäre Arbeitslastnetzwerk muss sich ebenfalls auf demselben Layer-2-Gerät befinden.
Physische Netzwerk-MTU	1500	Die MTU-Größe muss für jede verteilte Portgruppe mindestens 1500 betragen.

Tabelle 5-20. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
Latenz	100 ms	Die maximal empfohlene Latenz zwischen jedem Cluster, der Teil einer vSphere-Zone ist, die in einem Supervisor verbunden ist.
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)

Tabelle 5-21. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.
Verwaltungsnetzwerk-Subnetz	1	Das Subnetz, das für den Verwaltungsdatenverkehr zwischen ESXi-Hosts und vCenter Server und der Kubernetes-Steuerungsebene verwendet wird. Das Subnetz muss folgende Größe haben: ■ 1 IP-Adresse pro VMkernel-Adapter eines Hosts. ■ 1 IP-Adresse für die vCenter Server-Appliance. ■ 5 IP-Adressen für die Kubernetes-Steuerungsebene. 1 für jeden der 3 Knoten, 1 für virtuelle IP, 1 für fortlaufendes Cluster-Upgrade. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.
Verwaltungsnetzwerk-VLAN	1	VLAN-ID des Verwaltungsnetzwerk-Subnetzes.

Tabelle 5-22. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
vSphere Distributed Switch	1	Alle Hosts aus allen drei vSphere-Clustern müssen mit einem VDS verbunden sein.
Arbeitslastnetzwerke	1	Mindestens eine verteilte Portgruppe muss auf dem VDS erstellt werden, den Sie als primäres Arbeitslastnetzwerk konfigurieren. Je nach der gewählten Topologie können Sie dieselbe verteilte Portgruppe verwenden wie für das Arbeitslastnetzwerk der Namespaces oder mehrere Portgruppen erstellen und sie als Arbeitslastnetzwerke konfigurieren. Arbeitslastnetzwerke müssen die folgenden Anforderungen erfüllen: ■ Routing-Fähigkeit zwischen einem beliebigen Arbeitslastnetzwerk und dem von HAProxy für die Zuteilung von virtuellen IPs verwendeten Netzwerk. ■ Keine Überschneidung von IP-Adressen über alle Arbeitslastnetzwerke innerhalb eines Supervisors hinweg. **Wichtig** Das Arbeitslastnetzwerk muss sich in einem anderen Subnetz als das Verwaltungsnetzwerk befinden.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Tabelle 5-23. Netzwerkanforderungen für den Lastausgleichsdienst

HAProxy-Lastausgleichsdienst	1	Eine Instanz des HAProxy-Lastausgleichsdiensts, die mit einer vCenter Server-Instanz konfiguriert ist. ■ Wenn dieselbe HAProxy-Instanz mehrere Supervisoren unterstützt, muss sie Datenverkehr zu und von allen Arbeitslastnetzwerken über alle Supervisoren weiterleiten können. ■ IP-Bereiche über Arbeitslastnetzwerke hinweg in allen Supervisoren, die der HAProxy unterstützt, dürfen sich nicht überschneiden. ■ Das Netzwerk, das HAProxy wurde die Zuteilung von virtuellen IPs verwendet, muss zu den Arbeitslastnetzwerken geroutet werden können, die in allen Supervisoren verwendet werden, mit denen HAProxy verbunden ist.
IP-Bereich des virtuellen Servers	1	Ein dedizierter IP-Bereich für virtuelle IPs. Die HAProxy-VM muss der einzige Besitzer dieses Bereichs für virtuelle IPs sein. Der Bereich darf sich mit keinem IP-Bereich überschneiden, der einem beliebigen Arbeitslastnetzwerk im Besitz eines beliebigen Supervisors zugewiesen ist. Der Bereich darf sich nicht im selben Subnetz wie das Verwaltungsnetzwerk befinden.

Voraussetzungen für die Cluster-Supervisor-Bereitstellung

6

Informieren Sie sich über die Voraussetzungen zum Aktivieren eines Supervisors auf einem einzelnen vSphere-Cluster, der einer vSphere-Zone zugeordnet ist.

Lesen Sie als Nächstes die folgenden Themen:

- Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX Advanced Load Balancer und VDS-Netzwerk
- Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX
- Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX und NSX Advanced Load Balancer
- Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst

Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX Advanced Load Balancer und VDS-Netzwerk

Informieren Sie sich über die Voraussetzungen zum Aktivieren eines Supervisors auf einem vSphere-Cluster mit vDS-Netzwerk und NSX Advanced Load Balancer, auch als Avi-Lastausgleichsdienst bezeichnet. vSphere IaaS control plane unterstützt mehrere Topologien: ein einzelnes vDS-Netzwerk für die Avi-Dienst-Engine und die Load Balancer-Dienste, einen vDS für die Avi Management Plane und einen weiteren vDS für den NSX Advanced Load Balancer.

Arbeitslastnetzwerke

Zum Konfigurieren eines Supervisors mit dem vDS-Netzwerk-Stack müssen Sie alle Hosts aus dem Cluster mit einem vDS verbinden. Je nach der von Ihnen für den Supervisor implementierten Topologie erstellen Sie eine oder mehrere verteilte Portgruppen. Sie bestimmen die Portgruppen als Arbeitslastnetzwerke für vSphere-Namespaces. Arbeitslastnetzwerke bieten eine Verbindung zu den Knoten von Tanzu Kubernetes Grid-Clustern und zu den VMs der Supervisor Control Plane. Das Arbeitslastnetzwerk, das Konnektivität zu den Kubernetes-Steuerungsebenen-VMs bereitstellt, wird als „primäres Arbeitslastnetzwerk“ bezeichnet. Jeder Supervisor muss über ein primäres Arbeitslastnetzwerk verfügen. Sie müssen eine der verteilten Portgruppen als primäres Arbeitslastnetzwerk für den Supervisor festlegen.

Die Kubernetes-Steuerungsebenen-VMs im Supervisor verwenden drei IP-Adressen aus dem IP-Adressbereich, der dem primären Arbeitslastnetzwerk zugewiesen ist. Jeder Knoten eines Tanzu Kubernetes Grid-Clusters verfügt über eine eigene, aus dem Adressbereich des Arbeitslastnetzwerks zugewiesene IP-Adresse. Das Arbeitslastnetzwerk ist mit dem Namespace konfiguriert ist, in dem der Tanzu Kubernetes Grid-Cluster ausgeführt wird.

Netzwerkanforderungen

Der NSX Advanced Load Balancer erfordert zwei routingfähige Subnetze:

- Das Verwaltungsnetzwerk. Der NSX Advanced Load Balancer-Controller, auch als Controller bezeichnet, befindet sich im Verwaltungsnetzwerk. Das Verwaltungsnetzwerk stellt dem Controller Konnektivität zum vCenter Server, zu den ESXi-Hosts und den Supervisor-Knoten der Steuerungsebene bereit. In diesem Netzwerk befindet sich die Verwaltungsschnittstelle der AVI-Dienst-Engine. Für dieses Netzwerk sind ein vDS und eine verteilte Portgruppe erforderlich.
- Das Datennetzwerk. Die Datenschnittstelle der AVI-Dienst-Engines, auch als Dienst-Engines bezeichnet, stellt eine Verbindung zu diesem Netzwerk her. Die virtuellen IPs (VIPs) des Lastausgleichsdiensts werden von diesem Netzwerk zugewiesen. Für dieses Netzwerk sind ein vDS und verteilte Portgruppen erforderlich. Sie müssen den vDS und die verteilen Portgruppen konfigurieren, bevor Sie den Lastausgleichsdienst installieren.

Zuteilung der IP-Adressen

Der Controller und die Dienst-Engine sind mit dem Verwaltungsnetzwerk verbunden. Wenn Sie den virtuellen NSX Advanced Load Balancer installieren und konfigurieren, stellen Sie für jede Controller-VM eine statische, routingfähige IP-Adresse bereit.

Die Dienst-Engines können DHCP verwenden. Wenn DHCP nicht verfügbar ist, können Sie einen Pool von IP-Adressen für die Dienst-Engines konfigurieren.

Mindestanforderungen für Computing

In der Tabelle sind die Computing-Mindestanforderungen für VDS-Netzwerke mit NSX Advanced Load Balancer aufgelistet. Es wird empfohlen, die Verwaltungs- und die Arbeitslastdomäne zu trennen. Die Arbeitslastdomäne hostet den Supervisor, in dem Sie Arbeitslasten ausführen. In der Verwaltungsdomäne werden alle Verwaltungskomponenten gehostet, z. B. vCenter Server

Tabelle 6-1. Mindestanforderungen für Computing

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server8.0	Klein	2	21 GB	290 GB
ESXi-Hosts 8.0	■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 4 ESXi-Hosts pro Cluster mit mindestens 2 physischen Netzwerkkarten. Die Hosts müssen in einem Cluster mit aktiviertem vSphere DRS und HA verbunden sein. vSphere DRS muss sich im vollautomatischen oder teilweise automatisierten Modus befinden. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die dem Cluster beitreten, Kleinbuchstaben verwenden. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB
NSX Advanced Load Balancer-Controller	Enterprise Für Produktionsumgebungen wird empfohlen, einen Cluster mit 3 AVI-Controller-VMs zu installieren. Für HA sind mindestens 2 Dienst-Engine-VMs erforderlich.	4 (Klein) 8 (Mittel) 24 (Groß)	12 GB 24 GB 128 GB	128 GB 128 GB 128 GB
Dienst-Engine	Für HA sind mindestens 2 Dienst-Engine-VMs erforderlich.	1	2 GB	15 GB

Mindestanforderungen für das Netzwerk

In der Tabelle sind die Mindestnetzwerkanforderungen für vSphere-Netzwerke mit NSX Advanced Load Balancer aufgelistet.

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 7-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren. NSX Advanced Load Balancer-Dienste bieten derzeit keine Unterstützung für IPv6.

Tabelle 6-2. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Physische Netzwerk-MTU	1500	Die MTU-Größe muss für jede verteilte Portgruppe mindestens 1500 betragen.

Tabelle 6-3. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs) **Hinweis** Die DHCP-Konfiguration für Arbeitslastnetzwerke wird mit Supervisor-Dienste auf einem mit dem VDS-Stack konfigurierten Supervisor nicht unterstützt. Um Supervisor-Dienste zu verwenden, konfigurieren Sie Arbeitslastnetzwerke mit statischen IP-Adressen. Sie können DHCP weiterhin für das Verwaltungsnetzwerk verwenden.

Tabelle 6-4. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.
Verwaltungsnetzwerk-Subnetz	1	Der NSX Advanced Load Balancer-Controller, auch als Controller bezeichnet, befindet sich im Verwaltungsnetzwerk. Hier wird auch die Verwaltungsschnittstelle der Dienst-Engine verbunden. Der Controller muss mit den vCenter Server- und ESXi-Verwaltungs-IPs aus diesem Netzwerk verbunden sein. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.

Tabelle 6-5. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
vSphere Distributed Switch	1	Alle Hosts aus dem vSphere-Cluster müssen mit einem VDS verbunden sein.
Arbeitslastnetzwerke	1	Mindestens eine verteilte Portgruppe muss auf dem VDS erstellt werden, den Sie als primäres Arbeitslastnetzwerk konfigurieren. Je nach der gewählten Topologie können Sie dieselbe verteilte Portgruppe verwenden wie für das Arbeitslastnetzwerk der Namespaces oder mehrere Portgruppen erstellen und sie als Arbeitslastnetzwerke konfigurieren. Arbeitslastnetzwerke müssen die folgenden Anforderungen erfüllen: ■ Routing-Fähigkeit zwischen einem beliebigen Arbeitslastnetzwerk und dem vom NSX Advanced Load Balancer für die Zuteilung von virtuellen IPs verwendeten Netzwerk. ■ Keine Überschneidung von IP-Adressen über alle Arbeitslastnetzwerke innerhalb eines Supervisors hinweg.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Tabelle 6-6. Netzwerkanforderungen für den Lastausgleichsdienst

NTP- und DNS-Server	1	Die DNS-Server-IP ist erforderlich, damit der NSX Advanced Load Balancer-Controller die vCenter Server- und ESXi-Hostnamen korrekt auflösen kann. NTP ist optional, da öffentliche NTP-Server standardmäßig verwendet werden.
Datennetzwerksubnetz	1	Die Datenschnittstelle der NSX Advanced Load Balancer-Dienst-Engines, auch als Dienst-Engines bezeichnet, stellt eine Verbindung zu diesem Netzwerk her. Konfigurieren Sie einen Pool von IP-Adressen für die Dienst-Engines. Die virtuellen IPs (VIPs) des Lastausgleichsdiensts werden von diesem Netzwerk zugewiesen.
NSX Advanced Load Balancer-Controller-IPs	1 oder 4	Wenn Sie den NSX Advanced Load Balancer-Controller als einzelnen Knoten bereitstellen, ist eine statische IP-Adresse für seine Verwaltungsschnittstelle erforderlich. Für einen Cluster mit 3 Knoten sind 4 IP-Adressen erforderlich. Eine für jede NSX Advanced Load Balancer-Controller-VM und eine für die Cluster-VIP. Diese IPs müssen aus dem Subnetz des Verwaltungsnetzwerks stammen.
VIP-IPAM-Bereich	-	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Die IPs müssen aus dem Subnetz des Datennetzwerks stammen. Sie müssen für jeden Supervisor-Cluster einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Ports und Protokolle

Diese Tabelle enthält die Protokolle und Ports, die für die Verwaltung der IP-Konnektivität zwischen dem NSX Advanced Load Balancer, vCenter und anderen vSphere IaaS control plane-Komponenten erforderlich sind.

Quelle	Ziel	Protokoll und Ports
NSX Advanced Load Balancer-Controller	NSX Advanced Load Balancer-Controller (im Cluster)	TCP 22 (SSH) TCP 443 (HTTPS) TCP 8443 (HTTPS)
Dienst-Engine	Dienst-Engine in HA	TCP 9001 für VMware, LSC und NSX-T Cloud
Dienst-Engine	NSX Advanced Load Balancer-Controller	TCP 22 (SSH) TCP 8443 (HTTPS) UDP 123 (NTP)
AVI-Controller	vCenter Server, ESXi, NSX-T Manager	TCP 443 (HTTPS)
Supervisor-Steuerungsebenenknoten (AKO)	NSX Advanced Load Balancer-Controller	TCP 443 (HTTPS)

Weitere Informationen zu Ports und Protokollen für den NSX Advanced Load Balancer finden Sie unter https://ports.esp.vmware.com/home/NSX-Advanced-Load-Balancer.

Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX

Überprüfen Sie die Systemanforderungen zum Konfigurieren von vSphere IaaS control plane auf einem vSphere-Cluster mithilfe des NSX-Netzwerk-Stacks. Wenn Sie einen vSphere-Cluster als Supervisor aktivieren, wird automatisch eine vSphere-Zone für den Supervisor erstellt.

Neben diesen Anforderungen finden Sie im Leitfaden zum NSX-Referenz-Design weitere Informationen zu den Best Practices für die Bereitstellung von NSX.

Mindestberechnungsanforderungen für den Verwaltungs- und Edge-Cluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8	Klein	2	21 GB	290 GB
ESXi-Hosts 8	2 ESXi-Hosts	8	64 GB pro Host	Nicht anwendbar
NSX Manager	Mittel	6	24 GB	300 GB
NSX Edge 1	Groß	8	32 GB	200 GB
NSX Edge 2	Groß	8	32 GB	200 GB

Hinweis Stellen Sie sicher, dass alle ESXi-Hosts, die an dem vSphere-Cluster teilnehmen, auf dem Sie vSphere IaaS control plane konfigurieren möchten, als NSX-Transportknoten vorbereitet sind. Weitere Informationen finden Sie unter https://kb.vmware.com/s/article/95820 und Vorbereiten von ESXi-Hosts als Transportknoten in der Dokumentation zu NSX.

Mindestberechnungsanforderungen für den Arbeitslastdomänencluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vSphere-Cluster	■ 1 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf dem vSphere-Cluster. vSphere DRS muss im vollautomatisierten Modus ausgeführt werden.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8	■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 4 ESXi-Hosts mit mindestens zwei physischen Netzwerkkarten. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die den Clustern beitreten, in Kleinbuchstaben geschrieben sind. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB

Netzwerkanforderungen

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 8-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren.

Überprüfen Sie in der VMware-Produkt-Interoperabilitätstabelle, welche NSX-Versionen unterstützt werden.

Tabelle 6-7. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Physische Netzwerk-MTU	1500	Die MTU-Größe muss für jede vSphere Distributed Switch-Portgruppe mindestens 1500 betragen.
Physische Netzwerkkarte	Mindestens 2 physische Netzwerkkarten pro Host, wenn vSAN verwendet wird	Um Antrea-CNI zu verwenden und eine optimale NSX-Leistung zu erzielen, muss jede physische Netzwerkkarte auf jedem teilnehmenden ESXi-Host die GENEVE-Kapselung unterstützen und diese aktivieren.

Tabelle 6-8. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)
Image-Registrierung	1	Zugriff auf eine Registrierung für den Dienst.

Tabelle 6-9. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.
Verwaltungsnetzwerk-Subnetz	1	Das Subnetz, das für den Verwaltungsdatenverkehr zwischen ESXi-Hosts und vCenter Server, NSX-Appliances und der Kubernetes-Steuerungsebene verwendet wird. Das Subnetz muss folgende Größe haben: ■ 1 IP-Adresse pro VMkernel-Adapter eines Hosts. ■ 1 IP-Adresse für die vCenter Server-Appliance. ■ 1 oder 4 IP-Adressen für NSX Manager. 4 IP-Adressen bei NSX Manager-Clustering von 3 Knoten und 1 virtuellen IP (VIP). ■ 5 IP-Adressen für die Kubernetes-Steuerungsebene. 1 für jeden der 3 Knoten, 1 für virtuelle IP, 1 für fortlaufendes Cluster-Upgrade. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.
Verwaltungsnetzwerk-VLAN	1	VLAN-ID des Verwaltungsnetzwerk-Subnetzes.

Tabelle 6-10. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
CIDR-Bereich für vSphere Pods	/23 private IP-Adressen	Ein privater-CIDR-Bereich, der IP-Adressen für vSphere-Pods bereitstellt. Diese Adressen werden auch für die Tanzu Kubernetes Grid-Clusterknoten verwendet. Sie müssen einen eindeutigen vSphere Pod-CIDR-Bereich für jeden Cluster angeben. **Hinweis** Der CIDR-Bereich für vSphere Pods und der CIDR-Bereich für die Kubernetes-Dienstadressen dürfen sich nicht überlappen.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.
Egress-CIDR-Bereich	/27 statische IP-Adressen	Eine CIDR-Anmerkung zur Ermittlung der Egress-IP für Kubernetes-Dienste. Für jeden Namespace im Supervisor wird nur eine Egress-IP-Adresse zugewiesen. Die Egress-IP ist die Adresse, die externe Entitäten für die Kommunikation mit den Diensten im Namespace verwenden. Die Anzahl der Egress-IP-Adressen beschränkt die Anzahl der Egress-Richtlinien, die der Supervisor haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. Beispielsweise `10.174.4.96/27` **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.

Tabelle 6-10. Anforderungen an das Arbeitslastnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Ingress-CIDR	/27 statische IP-Adressen	Ein privater CIDR-Bereich, der für Ingress-IP-Adressen verwendet wird. Mit Ingress können Sie Datenverkehrsrichtlinien auf Anforderungen anwenden, die in den Supervisor von externen Netzwerken eingehen. Die Anzahl der Ingress-IP-Adressen beschränkt die Anzahl der Ingresses, die der Cluster haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.
Namespace-Netzwerkbereich	1	Ein oder mehrere IP-CIDRs zum Erstellen von Subnetzen/Segmenten sowie zum Zuweisen von IP-Adressen an Arbeitslasten.
Namespace-Subnetzpräfix	1	Das Subnetzpräfix, das die Größe des für Namespace-Segmente reservierten Subnetzes angibt. Der Standardwert ist „28“.

Tabelle 6-11. NSX-Anforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
VLANs	3	Diese VLAN-IPs sind die IP-Adressen für die Tunnel-Endpoints (TEP). Die TEPs des ESXi-Hosts und die Edge-TEPs müssen routingfähig sein. VLAN-IP-Adressen sind für Folgendes erforderlich: ■ ESXi-Host-VTEP ■ Edge-VTEP mit statischer IP-Adresse ■ Tier-0-Gateway und Uplink für Transportknoten. **Hinweis** Der ESXi-Host-VTEP und der Edge-VTEP müssen eine MTU-Größe von mehr als 1600 aufweisen. ESXi-Hosts und NSX-T-Edge-Knoten fungieren als Tunnel-Endpoints, und jedem Host sowie jedem Edge-Knoten wird eine TEP-IP zugewiesen. Da die TEP-IPs für ESXi-Hosts einen Overlay-Tunnel mit TEP-IPs auf den Edge-Knoten erstellen, müssen die VLAN-IPs routingfähig sein. Ein zusätzliches VLAN ist erforderlich, um Nord-Süd-Konnektivität zum Tier-0-Gateway zu bieten. IP-Pools können über Cluster hinweg gemeinsam genutzt werden. Der Host-Overlay-IP-Pool/VLAN darf jedoch nicht mit einem Edge-Overlay-IP-Pool/VLAN gemeinsam genutzt werden. **Hinweis** Wenn Host-TEP und Edge-TEP verschiedene physische Netzwerkkarten verwenden, können sie dasselbe VLAN verwenden.
Tier-0-Uplink-IP	/24 private IP-Adressen	Das IP-Subnetz, das für den Tier-0-Uplink verwendet wird. Für die IP-Adresse des Tier-0-Uplinks gelten folgende Anforderungen: ■ 1 IP, wenn Sie keine Edge-Redundanz verwenden. ■ 4 IPs, wenn Sie BGP und Edge-Redundanz verwenden – 2 IP-Adressen pro Edge. ■ 3 IPs, wenn Sie statische Routen und Edge-Redundanz verwenden. Die IP, das Subnetz und das Gateway für die Edge-Verwaltung und die IP, das Subnetz und das Gateway für Uplink müssen eindeutig sein.

Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit NSX und NSX Advanced Load Balancer

Überprüfen Sie die Systemanforderungen zum Konfigurieren von vSphere IaaS control plane auf einem vSphere-Cluster mithilfe des NSX-Netzwerk-Stacks. Wenn Sie einen vSphere-Cluster als Supervisor aktivieren, wird automatisch eine vSphere-Zone für den Supervisor erstellt.

NSX-Bereitstellungsoptionen

Weitere Informationen zu den Best Practices für die Bereitstellung von NSX finden Sie im Leitfaden zum NSX Referenz-Design.

Mindestberechnungsanforderungen für den Verwaltungs- und Edge-Cluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8	Klein	2	21 GB	290 GB
ESXi-Hosts 8	2 ESXi-Hosts	8	64 GB pro Host	Nicht anwendbar
NSX Manager	Mittel	6	24 GB	300 GB
NSX Edge 1	Groß	8	32 GB	200 GB
NSX Edge 2	Groß	8	32 GB	200 GB
Dienst-Engine-VMs	Pro Supervisor werden mindestens zwei Dienst-Engine-VMs bereitgestellt	1	2 GB	Nicht verfügbar

Hinweis Stellen Sie sicher, dass alle ESXi-Hosts, die an dem vSphere-Cluster teilnehmen, auf dem Sie vSphere IaaS control plane konfigurieren möchten, als NSX-Transportknoten vorbereitet sind. Weitere Informationen finden Sie unter https://kb.vmware.com/s/article/95820 und Vorbereiten von ESXi-Hosts als Transportknoten in der Dokumentation zu NSX.

Angeben der Systemkapazität des Controllers

Sie können die Systemkapazität des Controllers während der Bereitstellung angeben. Die Systemkapazität basiert auf Zuteilungen von Systemressourcen wie CPU, RAM und Festplatte. Die Menge an Ressourcen, die Sie zuteilen, wirkt sich auf die Leistung des Controllers aus.

Bereitstellungstyp	Knotenanzahl	Empfohlene Zuteilungen – CPU	Empfohlene Zuteilungen – Arbeitsspeicher	Empfohlene Zuteilungen – Festplatte
Demo/ Kundenbewertung	1	6	24 GB	128 GB

Bei Demobereitstellungen reicht ein einzelner Controller aus und wird für alle Aktivitäten und Workflows der Steuerungsebene sowie für Analysen verwendet.

In einer Produktionsbereitstellung wird ein Cluster mit drei Knoten empfohlen.

Weitere Informationen finden Sie unter NSX Advanced Load Balancer Controller-Größenanpassung.

Mindestberechnungsanforderungen für den Arbeitslastdomänencluster

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vSphere-Cluster	■ 1 vSphere-Cluster ■ Aktivierung von vSphere DRS und HA auf dem vSphere-Cluster. vSphere DRS muss im vollautomatisierten Modus ausgeführt werden.	Nicht anwendbar	Nicht anwendbar	Nicht anwendbar
ESXi-Hosts 8	■ Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. ■ Mit vSAN: 4 ESXi-Hosts mit mindestens zwei physischen Netzwerkkarten. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die den Clustern beitreten, in Kleinbuchstaben geschrieben sind. Andernfalls tritt bei der Aktivierung des Supervisors eventuell ein Fehler auf.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB

Netzwerkanforderungen

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 8-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren.

Überprüfen Sie in der VMware-Produkt-Interoperabilitätstabelle, welche NSX-Versionen unterstützt werden.

Tabelle 6-12. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Physische Netzwerk-MTU	1700	Die MTU-Größe muss für jede vSphere Distributed Switch-Portgruppe mindestens 1700 betragen.
Physische Netzwerkkarte	Mindestens 2 physische Netzwerkkarten pro Host, wenn vSAN verwendet wird	Um Antrea-CNI zu verwenden und eine optimale NSX-Leistung zu erzielen, muss jede physische Netzwerkkarte auf jedem teilnehmenden ESXi-Host die GENEVE-Kapselung unterstützen und diese aktivieren.

Tabelle 6-13. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)
Image-Registrierung	1	Zugriff auf eine Registrierung für den Dienst.

Tabelle 6-14. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.
Verwaltungsnetzwerk-Subnetz	1	Das Subnetz, das für den Verwaltungsdatenverkehr zwischen ESXi-Hosts und vCenter Server, NSX-Appliances und der Kubernetes-Steuerungsebene verwendet wird. Das Subnetz muss folgende Größe haben: ■ 1 IP-Adresse pro VMkernel-Adapter eines Hosts. ■ 1 IP-Adresse für die vCenter Server-Appliance. ■ 1 oder 4 IP-Adressen für NSX Manager. 4 IP-Adressen bei NSX Manager-Clustering von 3 Knoten und 1 virtuellen IP (VIP). ■ 5 IP-Adressen für die Kubernetes-Steuerungsebene. 1 für jeden der 3 Knoten, 1 für virtuelle IP, 1 für fortlaufendes Cluster-Upgrade. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.
Verwaltungsnetzwerk-VLAN	1	VLAN-ID des Verwaltungsnetzwerk-Subnetzes.

Tabelle 6-15. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
CIDR-Bereich für vSphere Pods	/23 private IP-Adressen	Ein privater-CIDR-Bereich, der IP-Adressen für vSphere-Pods bereitstellt. Diese Adressen werden auch für die Tanzu Kubernetes Grid-Clusterknoten verwendet. Sie müssen einen eindeutigen vSphere Pod-CIDR-Bereich für jeden Cluster angeben. **Hinweis** Der CIDR-Bereich für vSphere Pods und der CIDR-Bereich für die Kubernetes-Dienstadressen dürfen sich nicht überlappen.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.
Egress-CIDR-Bereich	/27 statische IP-Adressen	Eine CIDR-Anmerkung zur Ermittlung der Egress-IP für Kubernetes-Dienste. Für jeden Namespace im Supervisor wird nur eine Egress-IP-Adresse zugewiesen. Die Egress-IP ist die Adresse, die externe Entitäten für die Kommunikation mit den Diensten im Namespace verwenden. Die Anzahl der Egress-IP-Adressen beschränkt die Anzahl der Egress-Richtlinien, die der Supervisor haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. Beispielsweise `10.174.4.96/27` **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.

Tabelle 6-15. Anforderungen an das Arbeitslastnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Ingress-CIDR	/27 statische IP-Adressen	Ein privater CIDR-Bereich, der für Ingress-IP-Adressen verwendet wird. Mit Ingress können Sie Datenverkehrsrichtlinien auf Anforderungen anwenden, die in den Supervisor von externen Netzwerken eingehen. Die Anzahl der Ingress-IP-Adressen beschränkt die Anzahl der Ingresses, die der Cluster haben kann. Die Mindestanzahl ist ein CIDR von/27 oder mehr. **Hinweis** Egress-IP-Adressen und Ingress-IP-Adressen dürfen sich nicht überlappen.
Namespace-Netzwerkbereich	1	Ein oder mehrere IP-CIDRs zum Erstellen von Subnetzen/Segmenten sowie zum Zuweisen von IP-Adressen an Arbeitslasten.
Namespace-Subnetzpräfix	1	Das Subnetzpräfix, das die Größe des für Namespace-Segmente reservierten Subnetzes angibt. Der Standardwert ist „28“.

Tabelle 6-16. NSX-Anforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
VLANs	3	Diese VLAN-IPs sind die IP-Adressen für die Tunnel-Endpoints (TEP). Die TEPs des ESXi-Hosts und die Edge-TEPs müssen routingfähig sein. VLAN-IP-Adressen sind für Folgendes erforderlich: ■ ESXi-Host-VTEP ■ Edge-VTEP mit statischer IP-Adresse ■ Tier-0-Gateway und Uplink für Transportknoten. **Hinweis** Der ESXi-Host-VTEP und der Edge-VTEP müssen eine MTU-Größe von mehr als 1600 aufweisen. ESXi-Hosts und NSX-T-Edge-Knoten fungieren als Tunnel-Endpoints, und jedem Host sowie jedem Edge-Knoten wird eine TEP-IP zugewiesen. Da die TEP-IPs für ESXi-Hosts einen Overlay-Tunnel mit TEP-IPs auf den Edge-Knoten erstellen, müssen die VLAN-IPs routingfähig sein. Ein zusätzliches VLAN ist erforderlich, um Nord-Süd-Konnektivität zum Tier-0-Gateway zu bieten. IP-Pools können über Cluster hinweg gemeinsam genutzt werden. Der Host-Overlay-IP-Pool/VLAN darf jedoch nicht mit einem Edge-Overlay-IP-Pool/VLAN gemeinsam genutzt werden. **Hinweis** Wenn Host-TEP und Edge-TEP verschiedene physische Netzwerkkarten verwenden, können sie dasselbe VLAN verwenden.
Tier-0-Uplink-IP	/24 private IP-Adressen	Das IP-Subnetz, das für den Tier-0-Uplink verwendet wird. Für die IP-Adresse des Tier-0-Uplinks gelten folgende Anforderungen: ■ 1 IP, wenn Sie keine Edge-Redundanz verwenden. ■ 4 IPs, wenn Sie BGP und Edge-Redundanz verwenden – 2 IP-Adressen pro Edge. ■ 3 IPs, wenn Sie statische Routen und Edge-Redundanz verwenden. Die IP, das Subnetz und das Gateway für die Edge-Verwaltung und die IP, das Subnetz und das Gateway für Uplink müssen eindeutig sein.

Tabelle 6-17. Netzwerkanforderungen für den Lastausgleichsdienst

NTP- und DNS-Server	1	Die DNS-Server-IP ist erforderlich, damit der NSX Advanced Load Balancer-Controller die vCenter Server- und ESXi-Hostnamen korrekt auflösen kann. NTP ist optional, da öffentliche NTP-Server standardmäßig verwendet werden.
Datennetzwerksubnetz	1	Die Datenschnittstelle der NSX Advanced Load Balancer-Dienst-Engines, auch als Dienst-Engines bezeichnet, stellt eine Verbindung zu diesem Netzwerk her. Konfigurieren Sie einen Pool von IP-Adressen für die Dienst-Engines. Die virtuellen IPs (VIPs) des Lastausgleichsdiensts werden von diesem Netzwerk zugewiesen.
NSX Advanced Load Balancer-Controller-IPs	1 oder 4	Wenn Sie den NSX Advanced Load Balancer-Controller als einzelnen Knoten bereitstellen, ist eine statische IP-Adresse für seine Verwaltungsschnittstelle erforderlich. Für einen Cluster mit 3 Knoten sind 4 IP-Adressen erforderlich. Eine für jede NSX Advanced Load Balancer-Controller-VM und eine für die Cluster-VIP. Diese IPs müssen aus dem Subnetz des Verwaltungsnetzwerks stammen.
VIP-IPAM-Bereich	-	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Die IPs müssen aus dem Subnetz des Datennetzwerks stammen. Sie müssen für jeden Supervisor-Cluster einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Ports und Protokolle

Diese Tabelle enthält die Protokolle und Ports, die für die Verwaltung der IP-Konnektivität zwischen dem NSX Advanced Load Balancer, vCenter und anderen vSphere IaaS control plane-Komponenten erforderlich sind.

Quelle	Ziel	Protokoll und Ports
NSX Advanced Load Balancer-Controller	NSX Advanced Load Balancer-Controller (im Cluster)	TCP 22 (SSH) TCP 443 (HTTPS) TCP 8443 (HTTPS)
Dienst-Engine	Dienst-Engine in HA	TCP 9001 für VMware, LSC und NSX-T Cloud
Dienst-Engine	NSX Advanced Load Balancer-Controller	TCP 22 (SSH) TCP 8443 (HTTPS) UDP 123 (NTP)
AVI-Controller	vCenter Server, ESXi, NSX-T Manager	TCP 443 (HTTPS)
Supervisor-Steuerungsebenenknoten (AKO)	NSX Advanced Load Balancer-Controller	TCP 443 (HTTPS)

Weitere Informationen zu Ports und Protokollen für den NSX Advanced Load Balancer finden Sie unter https://ports.esp.vmware.com/home/NSX-Advanced-Load-Balancer.

Voraussetzungen für die Cluster-Supervisor-Bereitstellung mit VDS-Netzwerk und HAProxy-Lastausgleichsdienst

Überprüfen Sie die Systemanforderungen zum Einrichten eines vSphere-Clusters als Supervisor mit dem VDS-Netzwerk-Stack und dem HAProxy-Lastausgleichsdienst. Wenn Sie einen vSphere-Cluster als Supervisor aktivieren, wird automatisch eine vSphere-Zone für den Supervisor erstellt.

Mindestanforderungen für Computing

Es wird empfohlen, die Verwaltungs- und die Arbeitslastdomäne zu trennen. Die Arbeitslastdomäne hostet den Supervisor, in dem Sie Arbeitslasten ausführen. Die Verwaltungsdomäne hostet alle Verwaltungskomponenten wie vCenter Server.

System	Mindestbereitstellungsgröße	CPU	Arbeitsspeicher	Speicher
vCenter Server 8.0	Klein	2	21 GB	290 GB
ESXi-Hosts 8.0	Ohne vSAN: 3 ESXi-Hosts mit 1 statischen IP-Adresse pro Host. Mit vSAN: 4 ESXi-Hosts mit mindestens zwei physischen Netzwerkkarten. Die Hosts müssen in einem Cluster mit aktiviertem vSphere DRS und HA verbunden sein. vSphere DRS muss sich im vollautomatischen oder teilweise automatisierten Modus befinden. **Hinweis** Achten Sie darauf, dass die Namen der Hosts, die dem Cluster beitreten, Kleinbuchstaben verwenden. Andernfalls schlägt die Aktivierung des Clusters für die Arbeitslastverwaltung fehl.	8	64 GB pro Host	Nicht anwendbar
Kubernetes-Steuerungsebenen-VMs	3	4	16 GB	16 GB

Mindestanforderungen für das Netzwerk

Hinweis Sie können weder IPv6-Cluster mit einem vSphere 8-Supervisor erstellen noch IPv6-Cluster mit Tanzu Mission Control registrieren.

Tabelle 6-18. Anforderungen an das physische Netzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Physische Netzwerk-MTU	1500	Die MTU-Größe muss für jede verteilte Portgruppe mindestens 1500 betragen.

Tabelle 6-19. Allgemeine Netzwerkanforderungen

Komponente	Mindestanzahl	Erforderliche Konfiguration
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Für das Verwaltungsnetzwerk werden alle IP-Adressen, wie IP-Adressen von Steuerungsebenen-VMs, eine Floating-IP-Adresse, DNS-Server, DNS, Suchdomänen und NTP-Server, automatisch vom DHCP-Server erfasst. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs) **Hinweis** Die DHCP-Konfiguration für Arbeitslastnetzwerke wird mit Supervisor-Dienste auf einem mit dem VDS-Stack konfigurierten Supervisor nicht unterstützt. Um Supervisor-Dienste zu verwenden, konfigurieren Sie Arbeitslastnetzwerke mit statischen IP-Adressen. Sie können DHCP weiterhin für das Verwaltungsnetzwerk verwenden.

Tabelle 6-20. Anforderungen an das Verwaltungsnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
Statische IPs für Kubernetes Control Plane-VMs	Block mit 5	Ein Block mit 5 aufeinanderfolgenden statischen IP-Adressen, die vom Verwaltungsnetzwerk den Kubernetes-Steuerungsebenen-VMs im Supervisor zugewiesen werden.
Verwaltungsdatenverkehr-Netzwerk	1	Ein Verwaltungsnetzwerk, das zu den ESXi-Hosts, vCenter Server, dem Supervisor und einem Lastausgleichsdienst geroutet werden kann.

Tabelle 6-20. Anforderungen an das Verwaltungsnetzwerk (Fortsetzung)

Komponente	Mindestanzahl	Erforderliche Konfiguration
Verwaltungsnetzwerk-Subnetz	1	Das Subnetz, das für den Verwaltungsdatenverkehr zwischen ESXi-Hosts und vCenter Server und der Kubernetes-Steuerungsebene verwendet wird. Das Subnetz muss folgende Größe haben: ■ 1 IP-Adresse pro VMkernel-Adapter eines Hosts. ■ 1 IP-Adresse für die vCenter Server-Appliance. ■ 5 IP-Adressen für die Kubernetes-Steuerungsebene. 1 für jeden der 3 Knoten, 1 für virtuelle IP, 1 für fortlaufendes Cluster-Upgrade. **Hinweis** Das Verwaltungsnetzwerk und das Arbeitslastnetzwerk müssen sich in unterschiedlichen Subnetzen befinden. Das Zuweisen desselben Subnetzes zu den Verwaltungs- und Arbeitslastnetzwerken wird nicht unterstützt und kann zu Systemfehlern und Problemen führen.
Verwaltungsnetzwerk-VLAN	1	VLAN-ID des Verwaltungsnetzwerk-Subnetzes.

Tabelle 6-21. Anforderungen an das Arbeitslastnetzwerk

Komponente	Mindestanzahl	Erforderliche Konfiguration
vSphere Distributed Switch	1	Alle Hosts aus dem vSphere-Cluster müssen mit einem VDS verbunden sein.
Arbeitslastnetzwerke	1	Mindestens eine verteilte Portgruppe muss auf dem VDS erstellt werden, den Sie als primäres Arbeitslastnetzwerk konfigurieren. Je nach der gewählten Topologie können Sie dieselbe verteilte Portgruppe verwenden wie für das Arbeitslastnetzwerk der Namespaces oder mehrere Portgruppen erstellen und sie als Arbeitslastnetzwerke konfigurieren. Arbeitslastnetzwerke müssen die folgenden Anforderungen erfüllen: ■ Routing-Fähigkeit zwischen einem beliebigen Arbeitslastnetzwerk und dem von HAProxy für die Zuteilung von virtuellen IPs verwendeten Netzwerk. ■ Keine Überschneidung von IP-Adressen über alle Arbeitslastnetzwerke innerhalb eines Supervisors hinweg. **Wichtig** Das Arbeitslastnetzwerk muss sich in einem anderen Subnetz als das Verwaltungsnetzwerk befinden.
CIDR-Bereich für Kubernetes-Dienste	/16 private IP-Adressen	Ein privater CIDR-Bereich für die Zuweisung von IP-Adressen zu Kubernetes-Diensten. Sie müssen für jeden Supervisor einen eindeutigen CIDR-Bereich für Kubernetes-Dienste angeben.

Tabelle 6-22. Netzwerkanforderungen für den Lastausgleichsdienst

HAProxy-Lastausgleichsdienst	1	Eine Instanz des HAProxy-Lastausgleichsdiensts, die mit einer vCenter Server-Instanz konfiguriert ist. ■ Wenn dieselbe HAProxy-Instanz mehrere Supervisoren unterstützt, muss sie Datenverkehr zu und von allen Arbeitslastnetzwerken über alle Supervisoren weiterleiten können. ■ IP-Bereiche über Arbeitslastnetzwerke hinweg in allen Supervisoren, die der HAProxy unterstützt, dürfen sich nicht überschneiden. ■ Das Netzwerk, das HAProxy wurde die Zuteilung von virtuellen IPs verwendet, muss zu den Arbeitslastnetzwerken geroutet werden können, die in allen Supervisoren verwendet werden, mit denen HAProxy verbunden ist.
IP-Bereich des virtuellen Servers	1	Ein dedizierter IP-Bereich für virtuelle IPs. Die HAProxy-VM muss der einzige Besitzer dieses Bereichs für virtuelle IPs sein. Der Bereich darf sich mit keinem IP-Bereich überschneiden, der einem beliebigen Arbeitslastnetzwerk im Besitz eines beliebigen Supervisors zugewiesen ist. Der Bereich darf sich nicht im selben Subnetz wie das Verwaltungsnetzwerk befinden.

Komponente	Mindestanzahl	Erforderliche Konfiguration
NTP- und DNS-Server	1	Ein DNS-Server und ein NTP-Server, die in Verbindung mit vCenter Server verwendet werden können. **Hinweis** Konfigurieren Sie NTP auf allen ESXi-Hosts und in vCenter Server.
DHCP-Server	1	Optional. Konfigurieren Sie einen DHCP-Server, um automatisch IP-Adressen für die Verwaltungs- und Arbeitslastnetzwerke sowie Floating-IP-Adressen abzurufen. Der DHCP-Server muss Clientbezeichner unterstützen und kompatible DNS-Server, DNS-Suchdomänen und einen NTP-Server bereitstellen. Die DHCP-Konfiguration wird vom Supervisor verwendet. Lastausgleichsdienste benötigen möglicherweise statische IP-Adressen für die Verwaltung. DHCP-Bereiche sollten sich nicht mit diesen statischen IPs überlappen. DHCP wird nicht für virtuelle IPs verwendet. (VIPs)