Facultade de Informática
UNIVERSIDADE DA CORUÑA

TRABALLO FIN DE GRAO
GRAO EN ENXEÑARÍA INFORMÁTICA
MENCIÓN EN TECNOLOXÍAS DA INFORMACIÓN

PESCI: Plataforma de Entrega de Servicios Cloud para Investigación

Estudante: Amaro Castro Faci
Dirección: Outro Nome Completo

A Coruña, agosto de 2020.

Resumen

El Cloud Computing es un modelo que permite acceder a un conjunto de recursos, por ejemplo, redes, almacenamiento y cómputo, que pueden ser aprovisionados bajo demanda de forma automatizada y dinámica, reduciendo el coste del servicio para el usuario y el esfuerzo en cuanto a la administración de los recursos. El Centro de Investigación en Tecnoloxías da Información e as Comunicacións (CITIC) de la Universidade da Coruña cuenta con una infraestructura ideada para ofrecer un servicio de Cloud Computing a la comunidad universitaria. Este servicio consiste en que los usuarios pueden aprovisionar un conjunto de recursos del tamaño que requieran para realizar tareas que no serían posibles en dispositivos convencionales. Actualmente ese servicio está activo pero de forma limitada y no abierta a todos los usuarios del CITIC debido a que no existe una plataforma que permita gestionar los perfiles de usuario y su autenticación, ni un portal de acceso para aprovisionar recursos y gestionarlos de forma automatizada. En la actualidad, las tareas de aprovisionamiento y gestión de usuarios se realizan bajo petición previa al administrador del sistema que las ejecuta de forma manual lo cual produce gran coste en tiempo y recursos y aumenta los riesgos del servicio.
El objetivo principal de este proyecto es desplegar un servicio Cloud en el CITIC usando como base las herramientas que ya se encuentran sobre la infraestructura, donde cada usuario pueda tener su un espacio donde gestionar la obtención de recursos y que permita reducir las tareas de administración al integrar el sistema de autenticación de la UDC y al automatizar los procesos de aprovisionamiento, permitiendo establecer unos límites y controlar los recursos que utiliza cada usuario. Todo esto con el fin de mejorar la eficiencia de la infraestructura.
Para evitar problemas en el entorno de producción, el proyecto se desarrollará en un entorno de pruebas para mostrar las funcionalidades y características de la solución implementada pero que tendrán menor rendimiento que el entorno real por contar con recursos reducidos. El proceso se realizará siguiendo la metodología incremental Scrum en la cual primero se analizarán las diferentes alternativas disponibles y posteriormente se irán desplegando componentes para añadir nuevas funcionalidades que completen los objetivos del proyecto.

Palabras clave:

- Cloud Computing
- CITIC
- Máquina virtual
- Aprovisionamiento de recursos

- Bajo demanda
- Control de recursos

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1
Introducción

Según *National Institute of Standards and Technology* (NIST), el Cloud Computing es un «modelo de recursos configurables y compartidos, accesibles através de la red bajo demanda y desde cualquier lugar en cualquier momento»[2]. Las principales características de este modelo son:

- *Autoservicio bajo demanda*: El usuario puede aprovisionar recursos según sus necesidades y de forma automática sin requerir ninguna interacción humana con el proveedor del servicio.
- *Acceso por red*: El servicio está disponible para los usuarios a través de red de forma remota.
- *Almacén de recursos*: Los recursos son accesibles por múltiples usuarios simultáneamente, y todos ellos acceden a la misma instancia del software que gestiona el servicio, siendo así un servicio de *multi-tenant* [4]. Estos se pueden gestionar de forma dinámica y permiten conocer su ubicación física a un nivel de abstracción alto.
- *Elasticidad*: Los recursos se pueden aprovisionar o liberar de forma elástica, es decir, que se pueden escalar de forma rápida según las necesidades del usuario.
- *Servicio medido*: El sistema Cloud es capaz de aportar información sobre los recursos que el cliente tiene aprovisionados, que pueden ser almacenamiento, ancho de banda, procesamiento, y usuarios activos.

El Centro de Investigación en Tecnoloxías da Información e as Comunicacións (CITIC) de la Universidade da Coruña tiene en sus instalaciones una infraestructura construida para ofrecer un servicio Cloud al personal que trabaja allí y que así tengan acceso a hardware que no está disponible en dispositivos convencionales. Actualmente, esta infraestructura ya tiene instalado un software de la empresa VMware específico para crear y gestionar entornos

virtuales, por lo que el servicio ya está activo pero no cuenta con las herramientas suficientes para ofrecerlo de forma abierta a todos los usuarios. Este permite aprovisionar recursos de un servidor en forma de máquinas virtuales con unas especificaciones determinas por el usuario para realizar tareas que precisan gran capacidad de cómputo, de almacenamiento, o de red.

Inicialmente, el sistema cuenta con una plataforma, a la que los usuarios no tienen acceso debido a la falta de perfiles de usuario, para obtener recursos de la infraestructura física bajo demanda. Esto tiene que ser realizado por el personal encargado de recibir sus peticiones y de activar máquinas virtuales solicitadas, un proceso no automático y lento. Aunque actualmente si que es posible la creación de un perfil para cada usuario, esto no es viable ya que tampoco dispondrían de un espacio propio dentro del servicio si no que tendrían visibilidad y acceso, dependiendo de sus permisos, a los recursos de otros usuarios, a parte de que la interfaz es compleja y poco intuitiva, difícil de manejar para un usuario que no sea administrador del servicio. Por esto, el servicio no cumple con las características que define el NIST[1] para un servicio de Cloud Computing, especialmente en lo que se refiere al *Autoservicio bajo demanda*, *Elasticidad*, y *Servicio medido*, así que, usando como base esta definición, es necesario desplegar un portal donde los usuarios puedan acceder, usando sus perfiles de la UDC para facilitar la administración. En este portal el usuario puede aprovisionar recursos en forma de máquinas virtuales, modificarlos como necesite, y monitorizarlos. Esto implica que los usuarios podrían aprovisionar gran cantidad de recursos que luego podrían ser infrautilizados no pudiendo ser aprovechados por otros usuarios, por esto también es necesario implementar un sistema que permita a los administradores medir, valorar y limitar de alguna forma la cantidad de recursos aprovisionados por un mismo usuario.

Estas mejoras consiguen optimizar el uso de la infraestructura y aumentar su eficiencia debido a la automatización de gran parte de las operaciones que se repiten constantemente como el aprovisionamiento, gestión de usuarios, y creación de máquinas virtuales. Así se consigue un servicio más dinámico, útil y fácil de administrar y gestionar.

1.1 Motivación

La motivación para realizar este proyecto se basa en mejorar el servicio Cloud del CITIC para que aquellos usuarios que necesiten equipos de grandes prestaciones para sus tareas puedan conseguirlos de una forma sencilla y ágil al mismo tiempo que se mejora la gestión interna del servicio, y así reducir sus costes e incidencias a largo plazo. En definitiva, hacer que esta herramienta sea eficiente, útil y capaz de dar servicio a todos sus usuarios.

1.2 Objetivos

El objetivo general de este proyecto es crear un servicio piloto desplegando una herramienta sobre el sistema actual para hacerlo más eficiente y sacar el máximo potencial de toda la infraestructura y recursos administrativos que se encuentran disponibles tanto en el CITIC como en la UDC. Este servicio debe ser útil, ágil y accesible. Los objetivos concretos se pueden resumir en los siguientes:

- Centralizar y mejorar la gestión de usuarios integrando el sistema de autenticación de la UDC y así facilitar el acceso.
- Desplegar un portal de acceso para los usuarios que simplifique la gestión y aprovisionamiento de sus recursos.
- Implementar un sistema de valoración del servicio que permita limitar y controlar la cantidad de recursos que un usuario puede aprovisionar, y así evitar tener recursos ociosos.
- Documentar las soluciones desplegadas en el sistema para facilitar la transmisión de conocimiento a largo plazo.

1.3 Organización

La documentación de este proyecto se divide en cinco capítulos. El primero es 2.Estado de los recursos y en él se describe el hardware y el software que forman la infraestructura situada en el CITIC, la situación actual de la tecnología que se quiere implementar, las alternativas encontradas en el mercado y la descripción y componentes de la solución elegida. Posteriormente, en capítulo 3.Planificación se describen las tareas y los costes de la realización del proyecto en base a la solución elegida en el capítulo anterior. Una vez expuestas las tareas del proyecto, en el capítulo 4.Metodología se describen conceptos referidos a la infraestructura y arquitectura propios de la solución que se va a implementar, los requisitos físicos y servicios que la infraestructura debe proveer antes de realizar la implementación y, finalmente, la instalación y funcionalidades de los componentes de la solución dentro de un entorno de pruebas necesarios para cumplir los objetivos del proyecto.

Capítulo 2

Estado de los recursos

Con el fin de contextualizar los recursos que se utilizarán en este trabajo, en este capítulo se expone la situación actual de toda la infraestructura en lo relacionado al software que está en funcionamiento, a los recursos físicos de los que se compone, y al estado actual de las herramientas que rodean a dichos recursos.

2.1 Infraestructura

La infraestructura física donde se planea desplegar el servicio de virtualización, se encuentra localizada en el edificio del CITIC de la UDC, dentro de un rack alojado en su Centro de Proceso de Datos (CPD) [3].

2.1.1 Cómputo

La forman 5 nodos *Lenovo NeXtScale nx360 M5* cada uno con dos procesadores Intel Xeon E5-2650, 128 GB de memoria RAM y una tarjeta gráfica Tesla M60, y 3 nodos *Dell EMC PowerEdge R740* cada uno con dos procesadores Xeon Gold 6146, 384 GB de memoria RAM y una tarjeta gráfica Tesla P40. Todos ellos aportan flexibilidad en cuanto a la escalabilidad de la infraestructura y ofrecen gran rendimiento de cómputo.

2.1.2 Almacenamiento

El almacenamiento está colocado físicamente en la misma ubicación que los hosts pero en su abstracción lógica este es independiente y está separado de cada nodo. Está conformado por 13 discos duros SSD de 3.84 TB de capacidad, obteniendo así una cantidad total de casi 50 TB pero la capacidad útil es de 34 TB ya que se utiliza la configuración de almacenamiento RAID 5[Pal. 4] para aporta mayor integridad de los datos, mayor tolerancia a fallos y mayor ancho de banda. Los discos duros están colocados en una misma cabina donde forman un *pool* de almacenamiento que se divide en cinco LUNs (*Logical Storage Unit*)[Pal. 4] de 2 TB cada una,

representadas en el software de virtualización como cinco *datastores* y que emplean el sistema de archivos VMFS propio de la compañía VMware y el cual optimiza el almacenamiento de máquinas virtuales. La configuración y gestión de este sistema se tiene que realizar al nivel de la capa física, por lo tanto si se quiere hacer un despliegue en el sistema de virtualización con una configuración de almacenamiento diferente a la existente, como por ejemplo un sistema RAID con diferentes características, sería necesario modificar la configuración del sistema físico pudiendo generar un gran coste de tiempo. Esto no permite ajustar de forma precisa y rápida las configuraciones que se necesitan para realizar despliegues sobre la infraestructura.

2.1.3 Red

El sistema de almacenamiento forma una SAN, para ello las conexiones que se implementan entre los nodos y las cabinas donde se encuentran los discos duros son de tipo 10 Gbit. Para soportar esta conexión, cada cabina incorpora dos controladores SFP+[Pal. 4]. Además, las cabinas de almacenamiento incorporan otros dos puertos de 1 Gbit para la administración de los discos. En esta estructura se utilizan los protocolos de red Ethernet y iSCSI. Finalmente, para mantener la disponibilidad del acceso al sistema de almacenamiento y las comunicaciones entre los nodos, cada uno de ellos se conecta a dos switches *trunk* estableciendo rutas redundantes. Igual que con el sistema de almacenamiento, si se requieren realizar modificaciones sobre la red para adaptarse a los requisitos de un determinado despliegue habría que hacerlas directamente sobre la red física. Esto puede generar problemas en la conectividad del entorno a parte de generar gran coste de tiempo.

2.2 Software

Actualmente, el software desplegado sobre la infraestructura está formado por los productos de la compañía VMware, uno de los principales proveedores de software de virtualización, siendo **VMware vSphere**, versión 6.7, el principal componente ya que se utiliza para virtualizar parte de la infraestructura física y proporcionar las herramientas necesarias para gestionarla, sus principales componentes internos se describen a continuación. En cada nodo físico está instalado el **hipervisor ESXi**. Este es un hipervisor de tipo 1 o *baremetal*, es decir, funciona directamente sobre el hardware sin necesidad de un sistema operativo adicional. Sobre los nodos corre el servicio **VMware vCenter Server** que actúa como centro de administración de todas las máquinas virtuales (VMs) y nodos que forman la infraestructura y, además, contiene una instancia embebida de **Platform Services Controller** (PSC) punto que centraliza el acceso a distintos servicios como APIs de VMware vCenter Server, servidor de licencias o el servicio de autenticación **vCenter Single Sign-On**, este último se utiliza para gestionar la autenticación de los usuarios registrados en VMware vCenter Server. El acceso

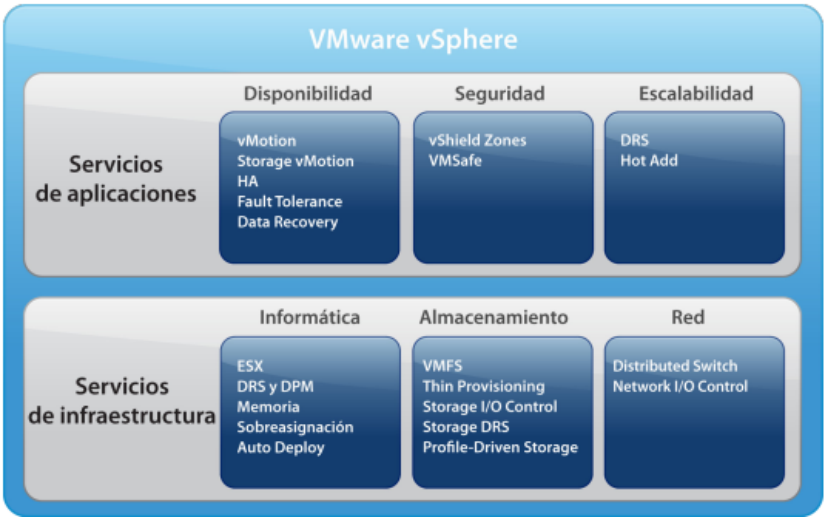

Figura 2.1: Componentes de VMware vSphere[1]

e interfaz de VMware vCenter Server se realiza a través de la página **vSphere Web Client** donde el usuario puede autenticarse y gestionar las VMs y nodos que forman el entorno y el resto de servicios de VMware vSphere. Además, incorpora **vSphere Update Manager** que permite gestionar las actualizaciones de software de los componentes de VMware vSphere. Para administrar las conexiones de las VMs, vSphere utiliza **vSphere Distributed Switch** (vDS), un switch virtual que gestiona el tráfico de cada VM permitiendo indicar que interfaces físicas de cada nodo físico, configurar sus puertos, establecer políticas y crear subredes de forma centralizada. Finalmente, existen varios servicios de gran importancia que se encargan de mantener la disponibilidad de las VMs desplegadas sobre la infraestructura:

- **vMotion** y **Storage vMotion**: el primero se encarga de migrar VMs de un nodo a otro de forma transparente y sin detener su ejecución, permitiendo planificar las migraciones. El segundo servicio se encarga de migrar los discos y configuración de una VM de un *datastore* a otro sin interrumpir el servicio.

- **vSphere High Availability** (HA): En caso de que una VM deje de estar activa, este servicio intenta encenderla de forma automática en otro nodo del entorno. A diferencia de vMotion, este solo actúa en caso de que la VM o el nodo donde se encuentra la VM sufra un fallo y esta pase a estar no disponible.

- **vSphere Distributed Resource Scheduler** (DRS), **vSphere Distributed Power Management** (DPM) y **Sotrage DRS**: vSphere DRS genera recomendaciones sobre donde se debería desplegar una máquina virtual durante su creación, utiliza vMotion para migrar las VMs y así maximizar el rendimiento o para manterner la VM activa durante

tareas de mantenimiento en un nodo. vSphere DPM se encarga de gestionar el consumo de energía de cada host según el rendimiento actual. Sotrage DRS se encarga de balancear la carga de almacenamiento y las operaciones de lectura y escritura entre los *datastores* disponibles.

- **vSphere Fault Tolerance**: gestiona una copia de todos los archivos y discos de cada VM sincronizada con los archivos originales. Este servicio usado con vSphere HA y vSphere DRS proporciona recuperación ante fallos automática y disponibilidad continua de las VMs, sin perdida de datos y sin pérdida de las conexiones establecidas. En caso de que una VM deje de estar disponible esta se reinicia en un nodo diferente. Este servicio está orientado a proteger aquellas tareas que requieren un alto rendimiento o que son críticas.

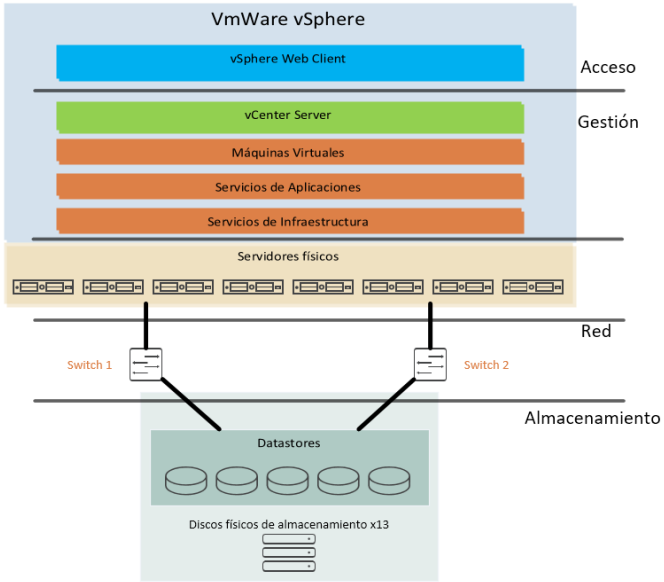

Figura 2.2: Componentes físicos y software que forman la infraestructura actual.

2.3 Estado de la tecnología

Con el desarrollo de las tecnologías web y la comercialización por parte de grandes empresas de su infraestructura los servicios *Infrastructure as a Service* (IaaS) han ganado una popularidad considerable, con ello también se han desarrollado herramientas software dedicadas a la gestión infraestructura para la implementación de sistemas Cloud Computing. Algunas de estos servicios son VMware Cloud Foundation (creado en 2011), OpenStack (creado en 2010) o Apache CloudStack (creado en 2012). Estas herramientas construyen una infraestructura virtual sobre un conjunto de recursos físicos estandarizados que permite separar la

administración de la capa física de la capa virtual, para simplificar y automatizar la gestión y escalabilidad de los recursos físicos y virtuales. Esto persigue reducir costes de gestión de la infraestructura y aumentar la disponibilidad del servicio, es decir, aumentar la eficiencia de la infraestructura física.

Si bien en el mercado existen varias alternativas que se pueden desplegar[1] sobre la infraestructura existente, finalmente, para cumplir los objetivos de este proyecto se ha escogido el producto **VMware Cloud Foundation** (VCF) ya que se integra perfectamente con los componentes de VMware ya instalados en la infraestructura y, por lo tanto, su mantenimiento a largo plazo es más sencillo. Desplegar un producto de una compañía diferente podría producir problemas de compatibilidad entre versiones a largo plazo, a pesar de que este se pueda integrar con el software VMware vSphere. Utilizando los productos de un mismo proveedor se asegura el soporte de las diferentes versiones del software instalado y la obtención del máximo rendimiento de cada componente. Para poder usar este software es necesaria la adquisición de licencias, estas se organizan por componente y por número de hosts sobre los que se va a instalar el producto. Aunque tienen un coste elevado, este producto aporta grandes beneficios en cuanto a la gestión del SDDC.

2.3.1 VMware Cloud Foundation

Esta solución de VMware virtualiza todas las capas de la infraestructura combinando cuatro de sus productos. Utiliza **VMware vSphere** para virtualizar y gestionar el cómputo, **VMware vSAN** para virtualizar y gestionar el almacenamiento, **VMware NSX-T** para la virtualización y gestión de la red, y **VMware vRealize** para gestionar las operaciones de la infraestructura virtual como el aprovisionamiento de recursos o la gestión de *logs* centralizada. Todos juntos, estos servicios convierten el CPD en un *Software Defined Datacenter* (SDDC), un entorno donde existe una infraestructura física que se abstrae en una capa virtual para separar la gestión de ambas y poder modificar la infraestructura virtual según las necesidades de los usuarios sin necesidad de modificar la configuración de la infraestructura física. Gracias a esa estructura obtiene las siguientes características:

- **Servicios software con integración nativa**: ofrece un conjunto de servicios software para el almacenamiento, red, seguridad y gestión de la cloud. Estos servicios se integran de forma nativa con la infraestructura minimizando las tareas de configuración y administración.
- **Escalabilidad y elasticidad de los recursos**: la capacidad de la infraestructura se puede modificar de forma sencilla gracias a la automatización del ciclo de vida de todos los elementos y al desacople entre las dos capas (la física y la virtual).

[1] OpenStack y Apache CloudStack entre otras.

- **Supervisión de los recursos**: ofrece supervisión de los recursos con reconocimiento de aplicaciones y solución de problemas, permitiendo conocer todos los eventos que tienen lugar en la infraestructura. También permite establecer políticas de seguridad en cuanto al acceso a los recursos y la red.
- **Aprovisionamiento automatizado**: permite la otención de recursos de forma automática incluyendo servicios de red, almacenamiento y cómputo. Los componentes de la infraestructura virtualizada se encargan de la reserva de los recursos y de todas las operaciones necesarias para llevarla acabo.
- **Ciclo de vida automatizado**: automatiza las operaciones previas, iniciales y posterios de los recursos de la plataforma para simplificar y coordinar su gestión. En estas tareas se incluye desde el despliegue de la plataforma y su implementación, el aprovisionamiento de nuevos recursos físicos y la instalación de actualizaciones para cada componente software.

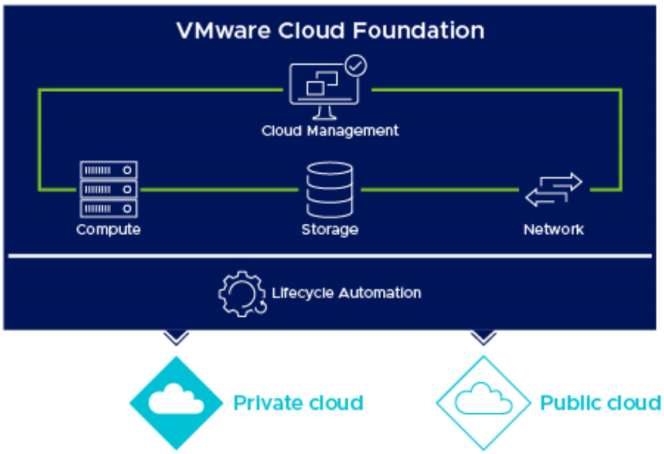

Figura 2.3: Resumen partes de VMare Cloud Foundation.

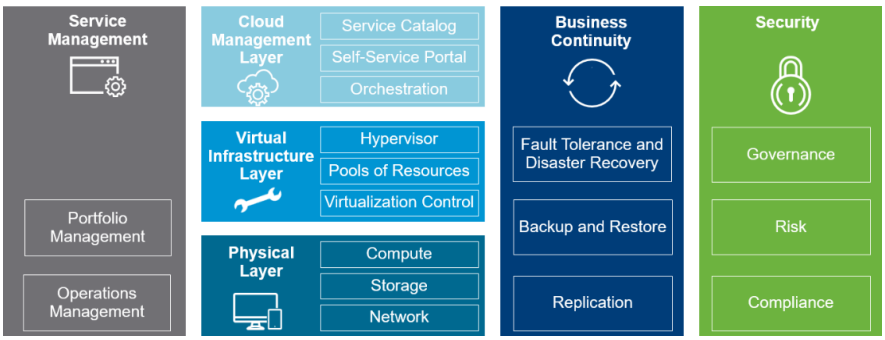

Figura 2.4: Elementos de un SDDC gestionado con VMware Cloud Foundation.

2.3.2 Componentes de VMware Cloud Foundation

Ya se ha visto que VCF está formado por cuatro productos principales de VMware. En este apartado se describirán las características de esos cuatro componentes más el servicio que los coordina[2]. Se utilizará la versión 4.0 de VMware Cloud Foundation lo cual implica que se implementarán las versiones[4] 4.0 de SDDC Manager, 7.0.0 de VMware vSphere, 7.0.0 de VMware vSAN, 3.0 de VMware NSX-T y 8.1 de VMware vRealize Suite.

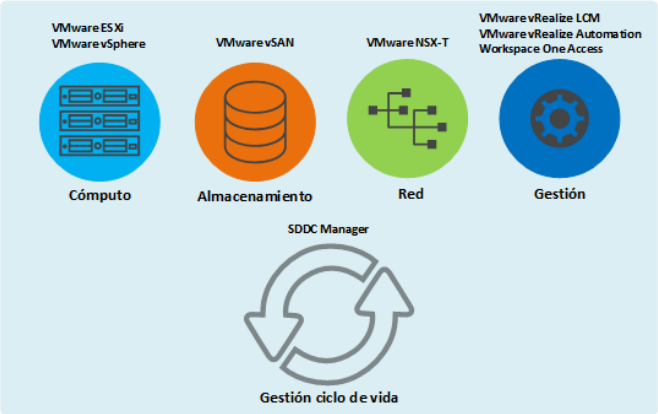

Figura 2.5: Partes de un SDDC y componentes de VCF que las implementan.

SDDC Manager

SDDC Manager se encarga de gestionar el ciclo de vida de todos los componentes de VCF, esto incluye el despliegue de cada uno, su configuración y la obtención e instalación de actualizaciones. Centraliza la gestión de las licencias y certificados de cada componente y administra el aprovisionamiento de nuevos recursos físicos para el SDDC y los ya existentes.

VMware vSAN

VMware vSAN virtualiza el almacenamiento del SDDC. Permite gestionar de forma centralizada desde la interfaz de vSphere Web Client el sistema de almacenamiento sin necesidad de tener que modificar la configuración física, como es el caso de las LUNs de la infraestructura actual. Trata todos los recursos de almacenamiento como un único elmento denominado *datastore* sobre el cual se pueden establecer políticas incluso a nivel de VM lo cual aporta gran flexibilidad. El acceso por parte de cada nodo al *datastore* se realiza con el protocolo IP a través de una subred dedicada al servicio. Con VMware vSAN, el *datastore* esta formado por discos de almacenamiento se organizan en grupos ligados a un nodo (un máximo de cinco grupos por nodo). Los grupos pueden tener configuración *Hybrid* que combinia discos HDD

[2]Las características del componente VMware vSphere son las mismas que las descritas en el punto 2.2

y SDD, o configuración *All-Flash* que solo utiliza SSD y por lo tanto tiene mayor rendimiento. Dentro de cada grupo existe un disco de caché y al menos un disco de capacidad donde se almacenan los datos persistentes[5]. En el modo *All-Flash*[3] la operación de lectura se realiza directamente sobre los discos de capacidad y la operación de escritura se hace sobre el disco caché que posteriormente escribe los datos en el disco de capacidad.

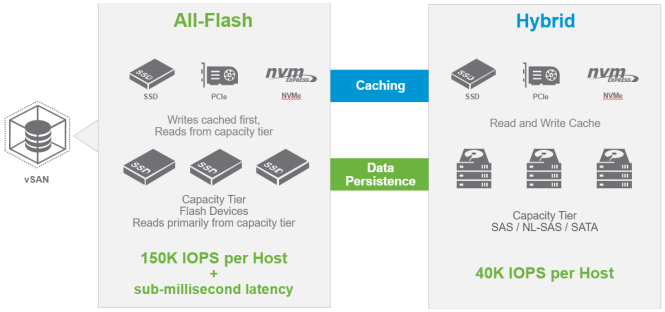

Figura 2.6: Configuración *All-Flash* y configuración *Hybrid* en vSAN

VMware NSX-T

VMware NSX-T virtualiza la red del SDDC. Abstrae los componentes físicos de la red para generar una red virtual desacoplada de la infraestructura física que se puede configurar sin modificar la red física, para ello aporta servicios de red virtualizados y la posibilidad de crear y extender subredes. Internamente está formado por varias instancias de **NSX-T Manager Appliance** que a su vez se compone de *NSX-T Manager* y **NSX-T Controller**. El primero es el punto de acceso a la configuración de VMware NSX-T y el que almacena y transmite la configuración establecida, el segundo controla las redes y servicios virtuales aportando la información y configuración necesarias para que gestionen el tráfico correctamente y obteniendo estadísticas sobre este. El control del tráfico y la monitorización de las conexiones se hace desde el componente **Transport Node** (TN) con la información que recibe de las instancias de NSX-T Controller. Existen dos tipos de TNs, **Hypervisor Transport Node** que son nodos con ESXi instalado y que están configurados para correr los servicios de VMware NSX-T, y **NSX-T Edge Node** que se trata de una *appliance* instalada en una VM o sobre un nodo físico para proveer un conjunto de servicios de red centralizados para las redes virtuales de VMware NSX-T.

[3]Solo se describe el modo *All-Flash* porque es la configuración recomendada por VMware.

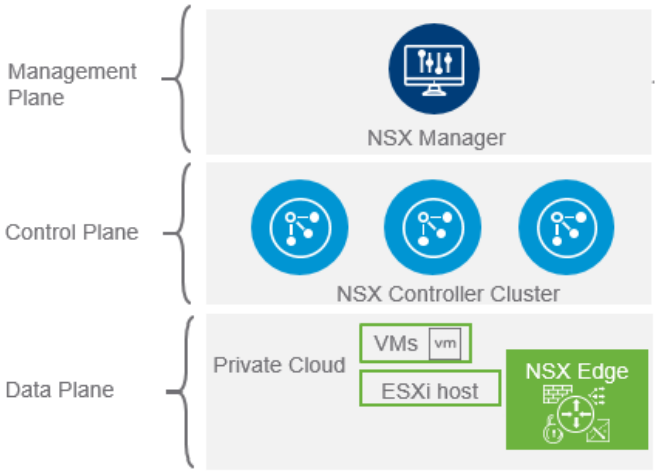

Figura 2.7: Componentes de VMware NSX-T y capas en las que se dividen

VMware vRealize Suite

VMware vRealize Suite agrupa un conjunto de productos que si bien no son obligatorios para desplegar VCF, aportan funcionalidades extra que completan la formación del SDDC. Los productos que se utilizarán en este proyecto son **vRealize Suite Lifecycle Manager** dedicado a gestionar el desliegue, actualizaciones, certificados y licencias de los productos que forman VMware vRealize, **Workspace One Access** dedicado a gestionar los usuarios y ser el punto de acceso centralizado de las aplicaciones de VMware vRealize y, finalmente, **vRealize Automation** el cual permite a los usuarios del SDDC diseñar y aprovisionar un conjunto de recursos de la infraestructura según sus necesidades y de forma automatizada mientras el administrador puede limitar la cantidad de recursos que se consumen.

Capítulo 3

Planificación

En este capítulo se propone una planificación del proyecto con el fin de organizar su estructura y exponer sus costes temporales y económicos aproximados necesarios para su realización.

3.1 Tareas

Tarea 1. Analizar como está formada la infraestructura, que componentes hardware y software la componen y cual es la función de cada uno de ellos.
En cuanto a la parte física se comprueban las especificaciones concretas del hardware de cómputo, almacenamiento y red. También como están organizados y estructurados tanto el sistema de almacenamiento y la red de la infraestructura. En la parte de software, se detallan las funciones de los principales programas y servicios que están instalados en el entorno.

Tarea 2. Analizar y seleccionar una herramienta de las disponibles en el mercado que se adapte a las necesidades del servicio que se quiere construir y a las características de la infraestructura. La herramienta seleccionada debe permitir reducir el coste y la complejidad de los trabajos de mantenimiento y administración del servicio a la vez que el usuario final lo utiliza de forma sencilla. En este proceso también se debe tener en cuenta la compatibilidad y eficiencia de la nueva herramienta con los componentes ya existentes en el entorno.

Tarea 3. Tarea que agrupa las tareas dedicadas al proceso de configuración de la infraestructura, configuración de la herramienta seleccionada y su instalación. Estas son las tareas 4, 5, 6, 7, 8, 9 y 10.

Tareas 4, 5, 6, 7, 8, 9 y 10. Comprobación de requisitos, preparación del entorno, establecimiento de parámetros configuración, despliegue de la plataforma sobre la infraestructura existente y configuración de la plataforma después del despliegue. Antes de realizar la ins-

talación de la nueva herramienta es necesario comprobar sus requisitos necesarios para que las capacidades del servicio final se adapten a las necesidades de uso (tareas 4 y 5). También se deben establecer los parámetros de configuración iniciales que se van a aplicar a la nueva plataforma (tarea 6). Durante el proceso de comprobación de requisitos puede surgir la necesidad de realizar cambios sobre las capacidades de la infraestructura y la configuración de los componentes ya existentes en el entorno inicial para que este se adapte a los requisitos de la nueva plataforma (tareas 7 y 8). Una vez el entorno está preparado para la herramienta pueda ser instalada entonces se efectúa el despliegue (tarea 9), posteriormente se configura y se comprueba el funcionamiento del nuevo servicio (tarea 10).

Tarea 11. Fin de la instalación y configuración de la plataforma. Marca el final del despliegue y configuración del nuevo servicio en la infraestructura.

Tarea 12. Diseñar una integración de la nueva plataforma con el sistema de autenticación de la UDC para que los usuarios finales del servicio se puedan autenticar sin necesitar nuevas credenciales. Para ello es preciso comprobar el método de acceso al directorio de usuarios de la UDC y la forma de conectarlo con la plataforma desplegada para, posteriormente, realizar un diseño de la solución. Este proceso requiere realizar una solicitud de acceso a los servicios internos de la UDC.

Tarea 13. Implementación y despliegue de la integración para la autenticación de usuarios con sus credenciales de la UDC. Durante este proceso puede ser necesario realizar cambios sobre la configuración de perfiles de usuarios que está establecida en la plataforma.

Tarea 14. Análisis del uso que harán los usuarios del servicio para establecer políticas sobre el uso de recursos. Para realizar este cálculo, primero se debe analizar el uso previo al despliegue del nuevo servicio que los usuarios hacen de la infraestructura y, después, estimar el uso que pueden llegar a realizar una vez el servicio sea accesible. Hay que tener en cuenta la cantidad de usuarios que lo utilizan, que lo van a utilizar y la cantidad de recursos que se emplean y que se van a emplear. Una vez obtenida una estimación, se realiza un diseño de las políticas que se van a aplicar.

Tarea 15. Diseño de un sistema de facturación/valoración de los recursos del servicio en base a las políticas de uso establecidas. Basándose en las políticas establecidas en la tarea 14, se debe pensar como se pueden aplicar sobre el servicio. Esto puede ser a través de una herramienta externa, en ese caso sería necesario realizar un desarrollo, o integrando la configuración en los parámetros de configuración de la plataforma.
La intención de este sistema es limitar la cantidad de recursos que un usuario puede apro-

visionar permitiendo aumentar la eficiencia de los recursos físicos reduciendo la cantidad de recursos ociosos.

Tarea 16. Implementación y despliegue del sistema de facturación/valoración. Para implementar este sistema puede que sea necesario realizar el desarrollo de una herramienta si se determina que no es posible establecerlo a través de los parámetros de configuración de la plataforma.

Tarea 17, 18 y 19. Recopilación de la información necesaria para la realización de cada tarea. La información de apoyo se debe obtener de documentaciones, artículos, vídeos o libros de fuentes fiables como empresas desarrolladoras de los productos utilizados o expertos especializados. El objetivo la recopilación de información es obtener conocimiento sobre las herramientas con las que se está trabajando para luego tener una base que facilite la realización de las tareas descritas. Esto se realiza desde el comienzo del proyecto hasta su finalización para tener claros los conceptos que se desarrollan y para conocer los detalles del trabajo que hay que realizar en cada tarea.

Tarea 20, 21 y 22. Redacción de la memoria del proyecto. Se escribe un documento con todos los detalles de todas las tareas realizadas durante el proyecto, incluyendo los cambios realizados en la infraestructura, las configuraciones establecidas y como se lleva a cabo cada proceso del proyecto. Su objetivo es transmitir el conocimiento adquirido durante el proyecto sobre como realizar el despliegue de una plataforma de virtualización y los beneficios que esta puede tener.La escritura de este documento se realiza a la vez que completa cada tarea para detallar los pasos realizados en cada caso, por lo que su duración es igual a la duración total de todo el proyecto.

La duración total del proyecto se estima en 101 días. teniendo en cuenta que el estudiante trabaja durante 4 horas diarias. El coste mostrado se refiere al coste correspondiente al estudiante si trabaja por 25 €/hora[Fig. 3.2].

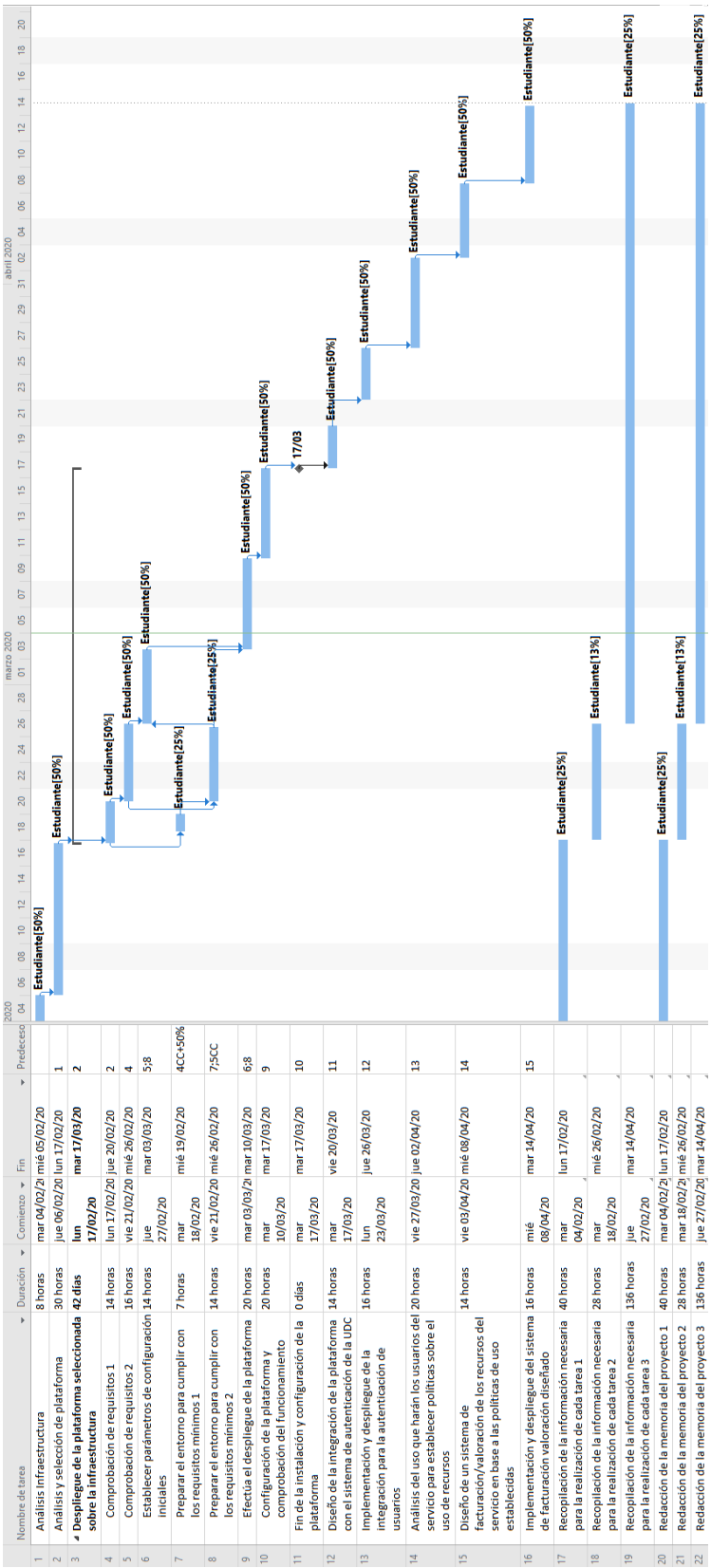

Figura 3.1: Diagrama de Grantt sobre la planificación del proyecto.

	Comienzo	Fin
Actual	mar 04/02/20	mar 14/04/20
Previsto	mar 04/02/20	mar 14/04/20
Real	NOD	NOD
Variación	0d	0d

	Duración	Trabajo	Costo
Actual	100,75d	201,25h	5.031,25 €
Previsto	100,75d	201,25h	5.031,25 €
Real	0d	0h	0,00 €
Restante	100,75d	201,25h	5.031,25 €

Figura 3.2: Estadísticas sobre la planificación del proyecto.

3.2 Costes

Los principales costes del proyecto son aquellos relacionados con los trabajadores que lo llevan a cabo y las licencias necesarias para cada componente de VMware Cloud Foundation en la infraestructura[1].
Cada componente de VMware Cloud Foundation requiere su propia licencia[6]. Estos componentes son SDDC Manager, VMware vSphere, VMware vCenter, VMware vSAN, VMware NSX for vSphere y VMware vRealize Log Insight. El precio de cada licencia dependerá del número de CPUs físicas sobre las que se va usar esta plataforma por lo que, como en la infraestructura hay un total de ocho hosts con dos CPUs cada uno, el precio por cada componente es el siguiente:

- **SDDC Manager**: 18.000€[2] por CPU y 6.500€ anuales de soporte por cada CPU. El precio total de la licencia es de 288.000€ y 104.000€ anuales de soporte por 16 CPUs.
- **VMware vSphere**: 4.000€[3] por CPU. El precio total de la licencia es de 64.000€ por 16 CPUs y el precio anual por las tareas de soporte es de 16.000€.
- **VMware vCenter**: 6.000€[4] por una licencia que permite usar VMware vCenter sobre todos los hosts del entorno. El precio anual por las tareas de soporte es de 1.500€.
- **VMware vSAN**: 4.000€[5] por CPU. El precio total de la licencia es de 64.000€ por 16 CPUs y el precio anual por las tareas de soporte es de 16.000€.

[1] Los componentes que se especifican son aquellos que son obligatorios para desplegar VMware Cloud Foundation.
[2] Para la edición *Advanced* de VMware Cloud Foundation.
[3] Para la edición *Standard* de VMware vSphere.
[4] Para la edición *Standard* de VMware vCenter
[5] Para la edición *Advanced* de VMware vSAN.

- **VMware NSX for vSphere**: 5.300€[6] por CPU. El precio total de la licencia es de 84.400€ por 16 CPUs y el precio anual por las tareas de soporte es de 21.100€.
- **VMware vRealize Log Insight**: 1.500€ por CPU. El precio total de la licencia es de 24.000€ por 16 CPUs y el precio anual por las tareas de soporte es de 6.000€.

El precio total de todas las licencias necesarias para el entorno, teniendo en cuenta que hay 16 CPUs, sería igual a 530.400€, y el precio total por las tareas de soporte sería igual a 164.600€ anuales.

En caso de que ya estén instalados algunos de los componentes entonces solo se requieren licencias para aquellos componentes que aún no están en el entorno. En el caso del entorno inicial, los componentes que ya están instalados son VMware vSphere, VMware vCenter Server. Esto hace que el coste real para implementar VMware Cloud Foundation en el entorno sea igual a 460.400€, ya que solo son necesarias licencias para los componentes SDDC Manager, VMware vSAN, VMware NSX for vSphere y VMware vRealize Log Insight. El coste total de la instalación y mantenimiento de la plataforma VMware Cloud Foundation sobre la infraestrructura del CITIC es el siguiente:

- **Licencias**: 460.400€ en total.
- **Soporte**: 164.600€ anuales.
- **Sueldo empleado**: 5.031,25€ en total.

[6]Para la edición *Advanced* de NSX.

Capítulo 4
Metodología

En este capítulo se describirá el desarrollo del proyecto y las funcionalidades más destacadas de la solución. Para ello se describirán varios conceptos necesarios para entender las partes y estructura de VMware Cloud Foundation, los requisitos para implementar VCF en un entorno real, y finalmente el despliegue del producto sobre un entorno de prueba para demostrar sus características.

4.1 Conceptos

En este apartado se describen algunos conceptos que se deben tener claros para entender la estructura y arquitectura de los componentes de VMware Cloud Foundation.

4.1.1 Workload Domain

Un *workload domain* (WD) representa un bloque de recursos dentro del SDDC, que son componentes de la infraestructura física, de la infraestructura virtual, y de seguridad. Los componentes virtuales controlan el acceso y la reserva de los recursos físicos, mientras que la capa de seguridad permite establecer organizar la entrada al WD. Cada WD contiene sus propias instancias de VMware ESXi, VMware vCenter Server, VMware NSX-T y VMware vSAN, pudiendo así gestionar los recursos de cada WD de forma independiente.

Management Domain

El *management domain* es el primer WD que se crea dentro del SDDC, y su función es la gestión de todos los componentes de VMware Cloud Foundation, tanto del propio *management domain* como del resto de *workload domains* existentes. En este *workload domain* se genera un cluster de VMware vSphere donde se despliegan las instancias de los siguientes componentes:

- Una VM de SDDC Manager.
- Una VM de VMware vCenter Server.
- Tres VMs de VMware NSX-T Manager Appliance.
- Dos VMs de VMware NSX-T Edge.

El administrador gestiona los recursos del *management domain* desde VMware vSphere Client, VMware NSX-T Manager (para gestionar sus redes virtuales) y VMware SDDC Manager para administrar los aspectos que afectan a todo el SDDC como puede ser la instalación de otras aplicaciones de VMware o la creación de nuevos *workload domains*.

Virtual Infrastructure Domain (VI)

Este tipo de *workload domain* se crea manualmente y bajo demanda desde *management domain* para habilitar entornos con una finalidad diferente. Su configuración de hardware y lógica se especifican durante su proceso de creación, permitiendo indicar la cantidad de hosts, cantidad de almacenamiento, configuración de la red y políticas de rendimiento y disponibilidad, todo para satisfacer las necesidades del tipo de tareas para las que se crea. Con cada *workload domain* se genera un nuevo cluster de VMware vSphere que agrupa los nuevos recursos pero parte de sus componentes que se despliegan se controlan desde el *management domain*:

- Una VM de VMware vCenter Server.
- Tres VMs de VMware NSX-T Manager Appliance situadas en el *management domain*.
- Dos VMs de VMware NSX-T Edge.

El administrador gestiona los recursos del *VI domain* desde VMware vSphere Client y la instancia de SDDC Manager situada en el *management domain*, y gestiona las redes virtuales del *workload domain* desde VMware NSX-T Manager situado también en el *management domain*.

4.1.2 Arquitectura

La arquitectura de VMware Cloud Foundation tiene dos posibles modelos de despliegue dependiendo del número de hosts sobre los que se despliega VMware Cloud Foundation.

Modelo estándar

Este modelo está pensado para desplegar VMware Cloud Foundation en entornos de tamaño medio/grande con un mínimo de siete hosts. Está formado por un *management domain*

que se despliega en cuatro de los hosts y contiene todos los componentes de gestión de toda la infraestructura, desde este *workload domain* se administra la infraestructura del SDDC y cada *virtual infrastructure domain* existente. Además, este modelo contiene al menos un *virtual infrastructure domain*, creado bajo demanda y con capacidades establecidas según su finalidad que posteriormente se pueden pueden modificar, se despliega sobre al menos tres hosts. Un *management domain* puede gestionar un máximo de de catorce *virtual infrastructure domains*.

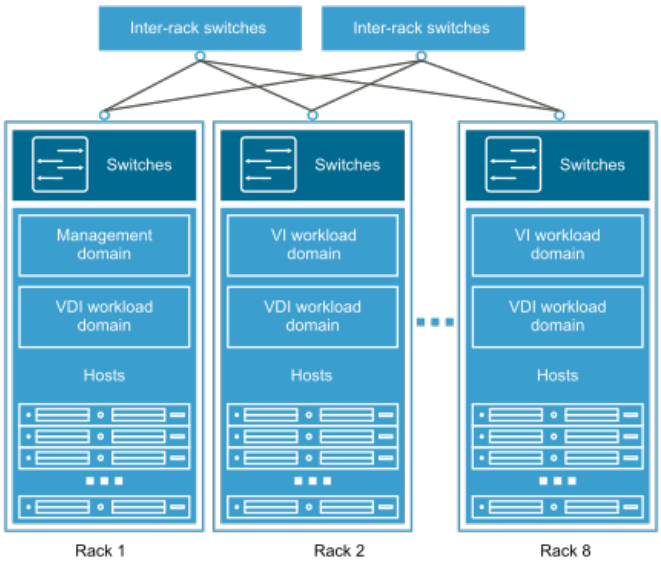

Figura 4.1: Esquema del modelo de arquitectura estándar.

Modelo consolidado

Este modelo está pensado para desplegar VMware Cloud Foundation en entornos de tamaño pequeño, normalmente cuando hay menos de siete hosts, aunque se puede también se puede utilizar sobre entornos más grandes de hasta 64 hosts. En este modelo los flujos de trabajo que corresponden al *virtual infrastructure domain* y al *management domain* en el despliegue estándar, están colocados dentro de un mismo *workload domain* en un único cluster pero aislados gracias a que cada uno se coloca dentro de un *resource pool* diferente, es decir, existe un cluster con varios *resource pool*. El modelo consolidado se convierte en un modelo estándar cuando se añade un *workload domain* al SDDC.[Fig. 4.2].

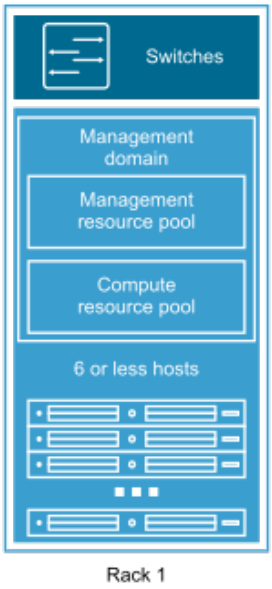

Figura 4.2: Esquema del modelo de arquitectura consolidado.

4.1.3 Clusters, zonas y distribución de un SDDC

AZ y Region

Los recursos de un SDDC pueden estar distribuídos en diferentes localizaciones para proporcionar alta disponibilidad y posibilidad de recuperación ante fallos. Los recursos de una o varias ubicaciones se agrupan para formar una estructura que permite usar y gestionar los recursos disponibles de forma conjunta y dinámica. Los componentes de esta estructura son:

- Availability Zone (AZ): conjunto de recursos físicos que forman una infraestructura independiente para evitar la propagación de fallos hacia otras AZs. Cuando existen varias AZ estas se pueden usar de forma que cuando ocurre un fallo en una de ellas la carga de trabajo se mueve a una segunda AZ para minimizar el tiempo de caída del servicio.

- Region: conjunto de AZs gestionadas que representa una instancia del SDDC. Las AZs que la forman deben tener una latencia de 5 ms como máximo mientras que la latencia entre Regions debe ser de al menos 100 ms. Esta estructura permite acercar el servicio a ubicaciones separadas por grandes distancias. La arquitectura del modelo consolidado solo soporta una Region con una AZ, mientras que el modelo estandar permite desplegar múltiples Regions con múltiples AZs.

Cluster y Resource Pool

Dentro de un *workload domain* pueden existir varios clusters. Un cluster es una agrupación de hosts a cuyos recursos se les puede aplicar una una configuración de disponibilidad determinada con los componentes VMware vSphere High Availability y VMware vSphere Distributed Resource Scheduler para establecer como se restablece el servicio en caso de fallos en alguno de los hosts. Un cluster puede estar extendido en más de una AZ para que si una de las AZs falla, las aplicaciones que corrían en ella pueden ser migradas a otra AZ mejorando la disponibilidad del servicio. En un WD se despliegan dos clusters:

- Management cluster: es el cluster que se crea al desplegar VMware Cloud Foundation. Contiene los componentes para administrar los recursos del WD.
- Shared Edge and Workload Cluster: después del management cluster, este es el primero que se crea. Su finalidad es alojar las aplicaciones y cargas de trabajo de los usuarios dentro de un WD. Además contiene instancias de Vmware NSX-T para proporcionar sevicios de red.

Dentro de un cluster se pueden crear resource pools. Un *resource pool* es una característica de VMware vSphere que permite abstraer un conjunto de recursos de un cluster estableciendo unos límites de capacidad que puede usar [7]. Usar resource pools permite agrupar las VMs con una finalidad similar y controlar la cantidad de recursos del WD que esas VMs pueden consumir.

4.2 Requisitos

En este apartado se describe aquello que debe complir la infraestructura física para que los componentes de VMware Cloud Foundation funcionen de forma adecuada y que la configuración y mantenimiento de los componentes físicos sea simple a la hora de expandir el entorno.

4.2.1 Cómputo

Hosts ESXi

Para realizar el despliegue del primer WD (el *management domain*) se requieren al menos cuatro[1] hosts ESXi con al menos un total de 256 GB de memoria RAM y un disco de arranque de 32 GB cada uno. Para cada WD adicional solo se requiere un mínimo de tres hosts y la cantidad de memoria RAM depende de la finalidad del WD, por lo tanto para implementar el

[1]Se reserva una cuarta parte de los recursos para que el *management domain* permanezca activo en caso de caída de alguno de los hosts.

modelo de arquitectura estándar se requieren al menos siete hosts ESXi. Cada uno de los hosts debe tener al menos dos interfaces de red físicas (NIC) que soporten al menos 10 Gbit/seg de velocidad.

4.2.2 Almacenamiento

En el *management domain* es obligatorio el uso de un *datastore* de VMware vSAN con al menos tres hosts con recursos de almacenamiento. Se debe aplicar la configuración All-Flash con discos SSD. Cada host debe tener un grupo de discos con al menos dos discos, uno de caché y otro de capacidad. El tamaño total de almacenamiento debe ser de 10TB y el tamaño total de caché debe ser de 1,2TB[2] (alrededor del 10% de la capacidad de almacenamiento). Para WD adicionales se puede utilizar almacenamiento NFS en lugar de un *datastore* de VMware vSAN, aunque la solución de VMware aporta mayor rendimiento y simplifica la administración de esta parte de la infraestructura física.

4.2.3 Red

Switch Top Of Rack

Los hosts están colocados en racks, en un rack puede haber hosts pertenecientes a distintos WD. Para favorecer la alta disponibilidad y tolerancia a fallos de la infraestructura física, un rack debe tener dos switches Top Of Rack (TOR) y cada host debe tener una interfaz conectada a cada switch TOR, una capa superior de switches conecta los diferentes racks entre si. Todas las conexiones de la red física deben soportar *Jumbo frames* (MTU hasta 9000 Bytes), etiquetado *Quality of Service* (QoS) de tráfico y las VLAN configuradas para las redes del SDDC[3]. Todos los switches TOR deben tener al menos dos interfaces 10 Gbit Ethernet como mínimo.

Servicios

En el SDDC se deben habilitar varios servicios requeridos por los componentes de VMware Cloud Foundation para su correcto funcionamiento.

- DNS: servidor de nombres para resolver todas las direcciones IP y *hostnames* de los componentes del SDDC.
- DHCP: servidor para asignar de forma automática una dirección IP a los hosts que forman el SDDC.

[2] La capacidad de los discos descrita es la necesaria para desplegar el *management domain* y un *workload domain* adicional.

[3] Para el *management domain* las subredes cuya VLAN debe ser configurada en la red física son la subred *management*, la subred para VMware vSAN, la subred para overlay y la subred para VMware vSphere vMotion.

- NTP: servidor de tiempo para sincronizar la hora de todos los componentes del SDDC.
- Router: se requiere para enrutar el tráfico que emiten todas las instancias del SDDC y para dar acceso a redes externas. Debe soportar enrutamiento dinámico BGP y debe tener configuradas las subredes y VLANS que se vayan a utilizar en la infraestructura.
- SMTP: servidor de correo utilizado por el componente VMware vRealize Automation.
- Active Directory: servidor de usuarios y grupos de usuarios que el SDDC utiliza como fuente para configurar el acceso a cada parte de la infraestructura virtual.
- Certificate Authority: se debe configurar una autoridad certificadora que genere certificados firmados para cada uno de los componentes de VMware Cloud Foundation. Permite establecer conexiones seguras cuando se accede a los componentes.

4.3 Prueba de concepto

Para no afectar al funcionamiento del servicio proporcionado por el CITIC y para mostrar y probar las capacidades de VMware Cloud Foundation, en lugar de utilizar un entorno real el proyecto se lleva a cabo en un entorno aislado de prestaciones reducidas. Siguiendo la metodología Scrum, primero se deplegará VMware Cloud Foundation con la herramienta VMware Lab Constructor (VLC)[4], que genera de forma automatizada una infraestructura física embebida basada en el diseño propuesto por VMware y sobre la cual posteriormente despliega los componentes base de VCF. Después se añadirá el componente que permite la gestión de usuarios y finalmente el servicio de aprovisionamiento. Una vez desplegados todos los elementos se realizará una demostración del servicio.

4.3.1 Preparación

Host ESXi

Como base para la instalación se utiliza un servidor físico con el hipervisor ESXi instalado que aunque no cumpla con alguno de requisitos mínimos de VMware Cloud Foundation, no aporta gran rendimiento pero si permite crear un entorno funcional a modo de prueba. Este host cuenta con una memoria RAM de 128 GB, una CPU de 28,8 GHz y un *datastore* con discos SSD con 2 TB de capacidad. Cuenta con dos interfaces físicas, una que conecta al host con el *datastore* y otra a la que se conectan dos redes, una llamada *Management Network* que permite acceder al host desde una VM para gestionarlo, y otra llamada *VM Network* donde se conectan todas las VMs generadas por VLC y de los servicios que dan soporte a los componentes de VMware Cloud Foundation.

[4]Se utiliza la versión 4.0.1 del instalador.

Servicios

Todos los servicios requeridos por VMware Cloud Foundation se despliegan sobre el mismo servidor en forma de VMs. Una de las VMs es Windows Server 2016 que contiene un servidor DNS, un servidor NTP, un servidor Active Directory, un servidor SMTP y ejerce también como Certificate Authority. Otra VM contiene el sistema operativo VyOS que funciona como un router virtual y como servidor DHCP. Una última VM con Windows 10[5] se requiere para ejecutar VLC y acceder al entorno embebido generado por VLC. El servidor DNS contiene los *hostnames* y sus respectivas direcciones IP de todas las VMs, tanto las que residen dentro del host físico como las que se alojan dentro del entorno embebido generado por la herramienta VLC. Este servidor DNS implementa un único dominio que se denomina *pesci.domain*. El servidor Active Directory proporciona almacena usuarios y grupos de usuarios requeridos para establecer roles y proporcionar acceso a los componentes y servicios de VMware Cloud Foundation. Se utiliza este servidor de usuarios en lugar del directorio real de la UDC para evitar posibles problemas del servicio. El router VyOS tiene configuradas todas las subredes y VLANs que VMware Cloud Foundation utiliza en la capa L3 de la infraestructura física y proporciona acceso a Internet, en las cuatro interfaces que conectan con las instancias de VMware NSX-T Edge utiliza enrutamiento dinámico BGP. El servidor DHCP asigna una dirección IP a cada TEP de cada host ESXi.

VMware Lab Constructor

VLC genera en el host ESXi cuatro VMs que representan cuatro hosts ESXi. posteriormente, dentro de estos hosts VLC inicia la creación del *management domain* de esta infraestructura embebida incluyendo todos los componentes de VMware Cloud Foundation. El diseño y configuración generados se describirá en las siguientes secciones.

[5] Se refiere a ella como *Jump Host*.

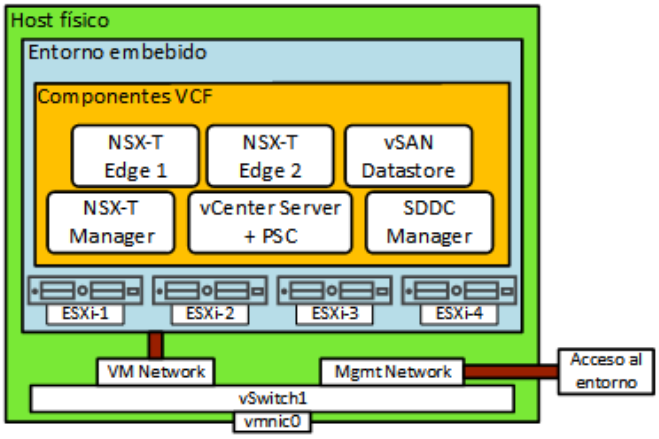

Figura 4.3: Muestra la estructura generada por el instalador VLC. Cuatro hosts ESXi embebidos con los componentes de VMware Cloud Foundation cuyo tráfico circula a través del *port group* VM Network.

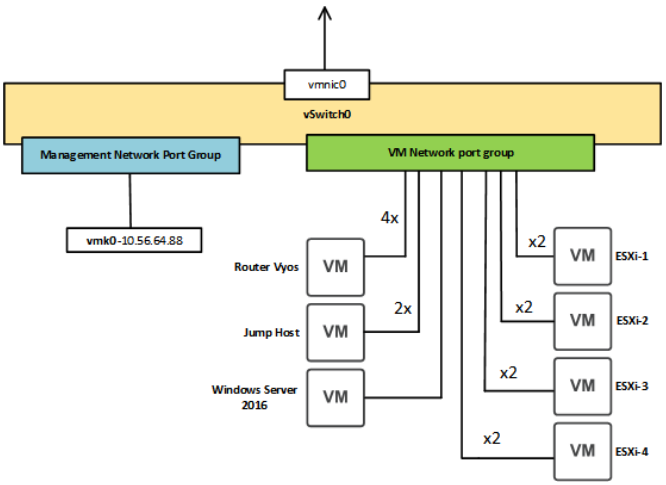

Figura 4.4: Máquinas virtuales en el host físico.

En la imagen anterior se muestran las VMs que están funcionando sobre el host físico y que representan los componentes de la infraestructura física de un SDDC real, junto con el número de interfaces que se utilizan en cada una. Cada host ESXi generado por VLC cuenta con dos interfaces de red. El router VyOS, Jump Host y Windows Server 2016 se configuran antes del despliegue de VMware Cloud Foundation con VLC y se comunican con el entorno generado por VLC a través del *port group* VM Network. El *port group* Management Network se utiliza para acceder a la configuración del host físico a través de la dirección IP que se indica. Se utiliza la interfaz vmnic0 del host como salida del tráfico generado por el vSwitch0.

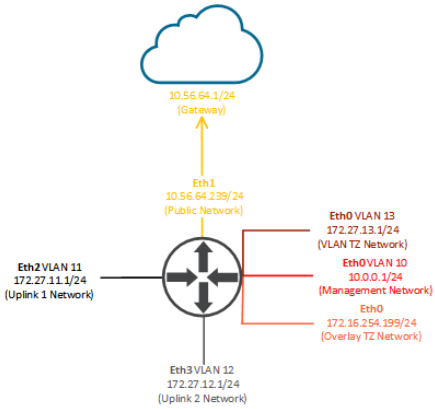

Figura 4.5: Interfaces del router Vyos.

En la imagen anterior se muestra la configuración del router VyOS. Cada una de las interfaces se debe configurar antes del despliegue de VCF. Todas usan MTU de 8940 Bytes. En las interfaces Eth2 y Eth3 el router utiliza enrutamiento dinámico BGP donde el AS local es 65001 y el AS remoto es AS 65003, configurado para anunciar a sus vecinos la red 10.0.0.0/24 Management Network. Las direcciones configuradas como *neighbour* son: 172.27.11.2, 172.27.11.3, 172.27.12.2 y 172.27.12.3. En la dirección IP 172.27.254.199 de la interfaz eth0, el router proporciona un servidor DHCP que asigna direcciones IP en el rango 172.16.254.0 - 172.16.254.100.

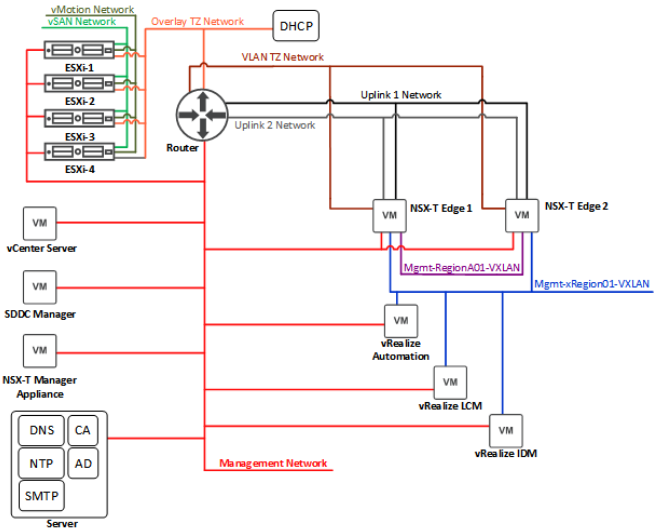

Figura 4.6: Topología de las redes del entorno desplegado.

En la imagen anterior se muestran todos los componentes de VMware Cloud Foundation desplegados por VLC y los desplegados posteriormente para completar los objetivos del proyecto, como se conectan con los distintos servicios de red y a que redes se conectan. Las

redes Mgmt-xRegion01-VXLAN y Mgmt-Region01A-VXLAN se corresponden a redes virtuales gestionadas por VMware NSX-T que no requieren ninguna configuración adicional en la capa 3 de la infraestructura física (esto se verá con detalle en el apartado de diseño de VMWare NSX-T).

4.3.2 Diseño y configuración del Management Domain

Diseño de VMware vCenter Server

El componente VMware vCenter Server es el punto de acceso y de control de todas las máquinas virtuales localizados en los hosts ESXi que forman parte de su dominio. En el entorno desplegado se utiliza una instancia de VMware vCenter Server para controlar el *management domain*, se denomina *vcenter-mgmt*. Esta instancia de vCenter Server contiene un dominio que con un cluster vSphere formado por los cuatro hosts ESXi desplegados por VLC, estos se denominan respectivamente *esxi-1*, *esxi-2*, *esxi-3* y *esxi-4*. En vCenter Server se gestionan los recursos de las VMs de cada componente, se monitorizan los recursos, permite la creación y asignación de roles, permisos y usuarios, aisla las redes que usan los recursos que controla de otras instancias de vCenter Server, permite gestionar los grupos de discos de almacenamiento de cada host ESXi que forman el *datastore* de VMware vSAN, administrar las redes a las que se conecta cada componente, en definitiva, VMware vCenter Server es el punto desde el cual se controlan los recursos que utiliza cada componente. Además, incluye el componente PSC que controla el dominio de autenticación de VMware vSphere SSO Domain denominado *local*. Desde vCenter Server también se controlan las características de alta disponibilidad y recuperación ante fallos de VMware vSphere como se verá a continuación. El acceso a vCenter Server se hace a través del componente web vSphere Client.

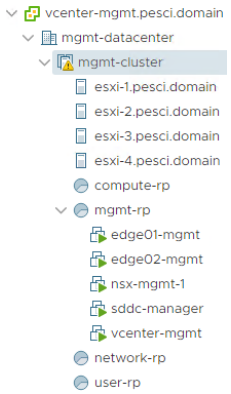

Figura 4.7: Dominio y cluster vSphere del *management domain*.

En la imagen anterior se muestra el dominio (*vcenter-mgmt.pesci.domain*) de la instancia de vCenter Server y el cluster vSphere (*mgmt-cluster*) donde se alojan los componentes del

management domain. Incluye cuatro hosts ESXi y cuatro *resource pools*, uno de ellos contiene las VMs de los componentes dedicados a este *management domain*.

Diseño almacenamiento VMware vSAN

El almacenamiento del *management domain* desplegado, está implementado con VMwware vSAN. Los cuatro hosts ESXi contienen cuatro grupos de discos cada uno con configuración All-Flash. Como hay cuatro hosts participantes, soporta el fallo de un host lo cual permite dejar hosts fuera de servicio para tareas de mantenimiento. Esto es posible gracias a que con FTT (*Failures-To-Tolerate*) igual a 1 se mantiene la redundancia de los datos almacenados en el *datasotore*, en uno de los hosts. Cada grupo de discos cuenta con cuatro discos uno de ellos para caché, 16 discos en total. Para hacer disponible este servicio de almacenamiento, todos los hosts deben estar conectados a la subred generada para VMware vSAN y utilizar una VLAN para separar su tráfico.

Diseño cluster VMware vSphere

Dentro de un *workload domain* pueden existir varios clusters vSphere con diferentes características según su finalidad. Los hosts ESXi que lo forman pueden ser de diferentes tamaños teniendo en cuenta que se pueden usar menos hosts ESXi de mayor capacidad o más hosts con menores prestaciones, el coste de cada host ESXi, el uso que se le va a dar al cluster y las características máximas y mínimas del cluster vSphere. Para el *management domain* se utiliza un único cluster vSphere con de 4 hosts de los cuales se reserva un host para proveer redundancia. Todos los hosts ESXi cuenta con 64GB de memoria RAM menos uno que tiene 32 GB, y 19.9GHz de CPU. Dentro del cluster hay que configurar los servicios vSphere HA y vSphere DRS para proteger los componentes del SDDC. La configuración que se establece en el *management domain* es la siguiente:

- **vSphere High Availability**: en este servicio la propiedad *Admission Control Policy* permite establecer la cantidad recursos reservados en caso de fallo y como se establece el cálculo de esos recursos. En el *management domain* se configura para el fallo de al menos un host y reserva de recursos según un porcentaje, reservando así el 25% de la CPU y el 30% de la memoria RAM ya que funciona mejor cuando las VM usan mucha CPU y memoria. La otra propiedad que se debe habilitar para el correcto funcionamiento del servicio es *VM and Application Monitoring*, que se encarga de reiniciar las VM en caso de caída.

- **vSphere DRS**: este servicio permite migrar VMs de un host ESXi a otro dentro del mismo cluster vSphere para equilibrar la carga de trabajo y mantener las VMs activas en caso de caída de alguno de los hosts. se activa usando la opción por defecto *Fully*

Automated ya que aporta el mejor balance entre consumo de recursos y migraciones de VM innecesarias. Adicionalmente se pueden establecer reglas para determinar reglas de orden de encendido sobre grupos de VM.

Diseño de red para el cluster vSphere

Si bien en VMware Cloud Foundation existe VMware NSX-T, un componente dedicado únicamente a la administración de la red del SDDC, es desde VMware vSphere dónde se encuentran los elementos para establecer redes que separen cada tipo de tráfico de los componentes del SDDC. Estas redes se configuran en base a los siguientes aspectos:

- Separar el tráfico de cada servicio para mejorar la eficiencia de la red y la seguridad. Así se puede ajustar las características de cada red, como el ancho de banda o la latencia, a las necesidades de cada servicio.
- Utilizar un único vSphere Distributed Switch por cluster donde se añade un *port groups* por cada servicio.
- Las NICs físicas de cada host ESXi conectados a un mismo vSphere Distributed Switch están conectadas también a la misma red física.

Para el *management domain* del SDDC se crea un único vSphere Distributed Switch llamado *sddc-vds01* con la siguiente configuración:

- Se establece un MTU igual 9000 Bytes para permitir el tráfico de *jumbo frames* ya que son requeridos por algunos de los servicios.
- Se habilita el servicio *Network I/O* que permite establecer un nivel de prioridad a cada tipo de tráfico. Esto se realiza estableciendo limites de ancho de banda, políticas de balanceo de carga y reserva de recursos para un tipo de tráfico asociado a un servicio. Por cada tipo de tráfico hay cuatro aspectos que se pueden configurar que son *Shares* (indica el % de ancho de banda que se le da a un tipo de tráfico, el tipo de tráfico que tenga un mayor valor en *Shares* tendrá más prioridad a la hora de usar los recursos), *Reservation* (indica el valor de ancho de banda que se reserva para el tipo de tráfico) y *Limit* (establece un valor máximo para el ancho de banda de un tipo de tráfico). En el *management domain* los tipos de tráfico más relevantes que se deben configurar son los siguientes:
  - *Management Traffic*: el valor *Shares* se establece al 50% (*Normal*) lo cual le da mayor prioridad que el resto de tipos. El resto de valores no se modifican.

  - *vSphere vMotion Traffic*: el valor *Shares* se establece al 25% (*Low*) ya que durante el estado normal del entorno este tipo de tráfico no es muy importante. El resto de valores no se modifican.
  - *vSAN Traffic*: el valor *Shares* se establece al 100% (*High*) para garantizar que este servicio recibe la cantidad de ancho de banda que necesita. El resto de valores no se modifican.
  - *Virtual Machine Traffic*: el valor *Shares* se establece al 100% (*High*) para garantizar que las VMs siempre tienen acceso a la red ya que son una parte importante del SDDC. El resto de valores no se modifican.

- Para detectar errores de compatibilidad entre la configuración del vSphere Distributed Switch y la red física se habilita el servicio *Health Check*. Este se encarga de comprobar si la configuración de cada VLAN y MTU se adapta a la configuración de la capa física.

- Como puertos de salida *Uplink* se configuran las interfaces físicas *vmnic0* y *vmnic1*. Como vDS es un componente distribuído, en cada host se usarán ambas interfaces de red como *uplinks*.

En este vSpehere Distributed Switch para el *management domain* se configuran los siguientes *port groups*, que son de tipo *Distributed port group* y de tipo *Uplink port group*:

- **Management port group**: es un *Distributed port group* que comunica a todos los hosts ESXi entre si y transmite el tráfico entre los diferentes componentes de VMware Cloud Foundation, es decir, por este *port group* circulan los comandos de configuración y gestión que los componentes del SDDC se envían entre ellos. Con el nombre *sddc-vds01-mgmt*, en él están configurados los cuatro hosts ESXi y las VMs *vcenter-mgmt*, *sddc-manager*, *nsx-mgmt-1*,*edge01-mgmt* y *edge02-mgmt* bajo la subred con IP 10.0.0.0, con máscara de red 255.255.255.0, con VLAN 10 y con MTU igual a 1500 Bytes. Esta red debe ser configurada también en la infraestructura física.

- **vMotion port group**: es un *Distributed port group* que está dedicado al tráfico del componente vSphere vMotion para realizar las migraciones de máquinas virtuales de un host a otro. Con el nombre *sddc-vds01-vmotion*, en él están configurados los 4 hosts bajo la subred con IP 10.0.4.0, con máscara de red 255.255.255.0, con VLAN 10 y con MTU igual a 8940 Bytes.

- **vSAN port group**: es un *Distributed port group* que está dedicado al servicio de almacenamiento VMware vSAN y por él los hosts acceden al almacenamiento del SDDC. Con el nombre *sddc-vds01-vsan*, en él están configurados los 4 hosts bajo la subred con IP 10.0.8.0, con máscara de red 255.255.255.0, con VLAN 10 y con MTU igual a 8940 Bytes.

- **Edge Uplink port group**: es un *Distributed port group* dedicado a las conexiones del component NSX-T Edge que se dedica a dar acceso a determinados servicios y para proporcionar a otros *workload domain* conexión con la red externa. Están gestionados por VMware NSX-T ya que dan servicio a sus componentes. En el entorno existen dos *port groups* para proporcionar redundancia y alta dispobilidad, uno llamado *sddc-edge-uplink01* cuyas instancias están configuradas bajo la red con IP 172.27.11.0 y con máscara de red 255.255.255.0, y otro llamado *sddc-edge-uplink02* cuyas instancias están configuradas bajo la red con IP 172.27.12.0 y máscara de red 255.255.255.0. Ambos *port groups* están configurados como VLAN Trunk (por ellos puede circular tráfico de cualquier VLAN) y tienen un MTU de 8940 Bytes. En los dos están configuradas dos VM llamadas *edge01-mgmt* y *edge02-mgmt*. Estas dos redes también se deben configurar en la infraestructura física.

- **Uplink port group**: se trata de un *Uplink port group* al que se le asignan las NICs físicas de cada host para establecer políticas sobre el tráfico que se dirige desde los hosts y VMs hacia fuera del vSphere Distributed Switch. Con el nombre *sddc-vds01-DVUplinks-10*, en él están configuradas las dos NICs físicas de cada host, cada una en una interfaz *uplink*.

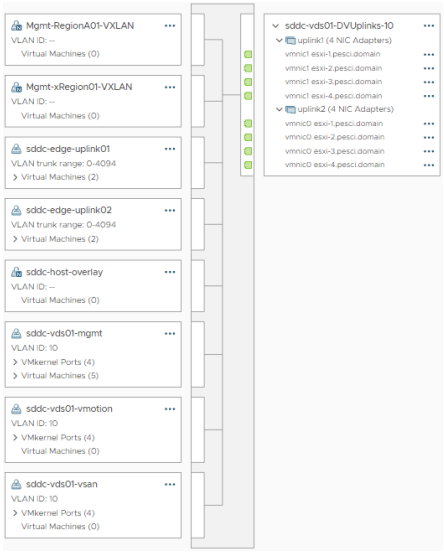

Figura 4.8: Contenido de vSphere Distributed Switch *sddc-vds01*.

En la imagen anterior se muestran todos los *Distributed Port Groups* y *Uplink port group* que se alojan en el vSphere Distributed Switch (*sddc-vds01*) dedicado al *management domain*. En el *port group sddc-vds01-DVUplinks-10* se muestra como a cada interfaz *uplink* se mapea una interfaz física (vmnic) de cada host ESXi. Los *port groups mgmt-Region01A-VXLAN*, *mgmt-xRegion01-VXLAN* y *sddc-host-overlay* son generados y administrados por el componente VM-

ware NSX-T como se explicará más adelante. Cada *port group* informa de cuantas VMs y hosts ESXi tiene conectados.

La configuración que se aplica a cada *Distributed port group* descrito anteriormente es la siguiente:

- *Port binding*: permite indidcar como se gestionan los puertos de un *port group* cuando se añade o elimina una VM. Tiene dos opciones de configuración, la primera se denomina *Static Port Binding* y su función consiste en asignar un puerto dentro del *port group* a la VM que se conecta y solo se elimina cuando la VM es borrada. La segunda opción se denomina *Ephemeral Port Binding* y consiste en que el puerto se asigna a la VM cuando esta se enciende y se elimina cuando se apaga o elimina. Para los *port groups sddc-vds01-vsan* y *sddc-vds01-vmotion* se configura la opción *Static Port Binding* ya que así se asegura que las VMs se conectan siempre al mismo puerto lo cual permite mantener datos históricos y hacer monitoreo a nivel de puerto. Para los *port group sddc-vds01-mgmt*, *sddc-edge-uplink01* y *sddc-edge-uplink02* se configura la opción *Ephemeral Port Binding* ya que como el tráfico que circula por ellos es el que gestiona todos los componentes y da acceso a otras redes etonces se elimina la dependencia del estado de vCenter Server permitiendo que la comunicación continúe aunque vCenter Server no se encuentre operativo.

- *Load Balancing*: indica como se distribuye el tráfico de salida de cada VM/host que se encuentran en el *port group* entre las NICs físicas. Se selecciona *Route based on physical NIC load*, es decir, el tráfico de una VM se transmite por una única NIC por lo que si esa NIC física está saturada, se asignará otra NIC física a la VM.

- *Network failure detection*: esta opción permite establecer como debe determinar el *port group* que alguna de las NICs físicas está fuera de servicio. Se selecciona *Link status only* para que esto se determine según el estado que le transmite la NIC física, así se pueden detectar los fallos que ocurren en la red física.

- *Notify switches*: se habilita para permitir a los host enviar *frames* a los switches físicos para que estos conozcan la localización de las VM que están funcionando en cada host.

- *Failback*: permite determinar como se reactiva una NIC cuando esta se recupera de un fallo. Se habilita para establecer que la NIC se marcará como activa inmediatamente después de que se haya recuperado. Esta opción se debería desactivar en caso de que el estado de la NIC sea inestable.

- *Failover Order*: permite determinar que uplinks se deben utilizar, los que se seleccionan como *active* son los que se utilizarán por defecto, los que se seleccionan como *stand*

by se usarán cuando los uplinks marcados como *active* se encuentren desactivados. Se seleccionan las dos interfaces *uplink* disponibles en el estado *active*. Para el *port group sddc-edge-uplink01* se selecciona la interfaz *uplink1* como activa y se deja sin usar la interfaz *uplink2*, mientras que se configura de forma contraria en el *port group sddc-edge-uplink02*.

Diseño de la red del SDDC con VMware NSX-T

En un SDDC debe existir una red virtual, es decir, definida por software o también conocida como *Software-Defined Network*. Esta red al estar construída con componentes de software, se desacopla de la red física sobre la que funciona lo que hace posible que se pueda modificar sin necesidad de cambiar la configuración en la capa física, reduciendo así la complejidad de la red física y el tiempo dedicado a la gestión de la misma. Además, este tipo de arquitectura habilita la posibilidad de implementar múltiples configuraciones de red en tiempo reducido proporcionando elasticidad y flexibilidad a la hora de administrar los recursos, tanto para el administrador como para el usuario final. El componente encargado de crear, configurar y administrar la red virtualizada del SDDC es VMware NSX-T que a su vez contiene otros componentes entre los que se dividen distintas responsabilidades y funciones, ya descritos anteriormente.

Aunque se recomienda desplegar tres instancias de NSX-T Manager en el *management domain*, VLC solo despliega una única instancia llamada *nsx-mgmt-1* para mejorar el rendimiento del entorno. Además, VLC también genera dos instancias de NSX-T Edge que se denominan *edge01-mgmt* y *edge02-mgmt*. Estas VMs están conectadas al *port group sddc-vds01-mgmt* que les permite comunicarse entre ellas y con vCenter Server, además las instancias de NSX-T Edge también están conectadas a otros dos *Distributed port group* llamados *sddc-edge-uplink01* y *sddc-edge-uplink02*.

Los elementos que utiliza VMware NSX-T para crear una red independiente de la configuración de la red física son *Segment* y *Transport Zone*. Con estos componentes VMware NSX-T puede crear túneles que definen redes de capa 2 sin necesidad de realizar ningún cambio en la configuración de la red física.

- **Transport Zone** (TZ): define el alcance de la red virtual. Pueden ser de dos tipos distintos, basada en VLAN o basada en Overlay. Una TZ se puede asignar a varios TN que tendrán acceso a los *segments* que funcionen en esa TZ. Un TN se conecta a una TZ a través de un N-VDS los cuales pueden pueden estar conectados a varias TZ de tipo VLAN pero solo a una TZ de tipo Overlay al mismo tiempo.

- **Segment**: también llamado *Logical Switch*, representa un dominio de broadcast de capa 2 que forma parte de una *Transport Zone*. El tipo de tráfico puede ser VLAN u Overlay

dependiendo de como se haya configurado la *Transport Zone* de la que forma parte. Las VMs de cada TN se pueden conectar a los *Segments* situados en las *Transport Zones* a las que el host está conectado. Estas VMs se pueden comunicar con el resto de VMs conectadas al mismo *Segment*.

Para gestionar las conexiones de cada TN, tanto para los nodos NSX-T Edge como para los hosts ESXi, VMware NSX-T introduce el componente llamado **NSX-T Virtual Distributed Switch** (N-VDS). Cada TN del *management domain* posee un N-VDS, este elemento conecta sus interfaces a los *segments* que se configuran en cada TN. Para el *management domain* del entorno, VLC despliega dos *transport zones* diferentes.

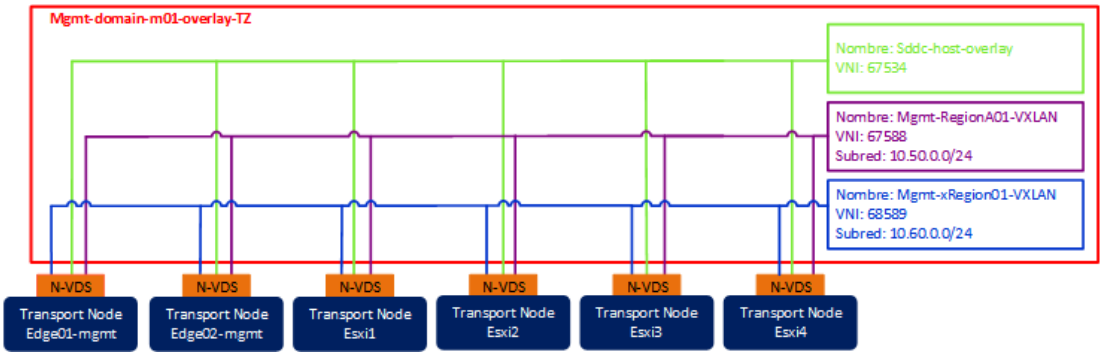

Figura 4.9: *Segments* de la *transport zone mgmt-domain-m01-overlay-tz*

En la imagen anterior se muestra la *transport zone* con el nombre *mgmt-domain-m01-overlay-tz* de tipo Overlay. Está extendida en los seis TNs que hay en el *management domain* y contiene tres *segments*. El *segment mgmt-xRegion01-VXLAN* se utiliza para desplegar aplicaciones que deben ser accesibles desde todas las *regions* que existan en el SDDC, es decir, la misma instancia de una aplicación está disponible desde varios puntos, así se reduce el consumo de recursos y aumenta la disponibilidad de esas aplicaciones ya que pueden migrar a distintas localizaciones según el estado de los recursos. El *segment mgmt-Region01A-VXLAN* tiene como finalidad alojar aplicaciones solo deben ser accesibles desde dentro de una misma *region*. En estos dos *segments* es donde se colocan los productos de VMware vRealize. Por último, el *segment sddc-host-overlay* es utilizado por los componentes de VMware NSX-T para comunicarse con y entre los diferentes TNs. VMware NSX-T genera en el vSphere vSwitch un *port group* por cada *segment* para poder conectar la VM de cada componente al *port group* que le corresponda.

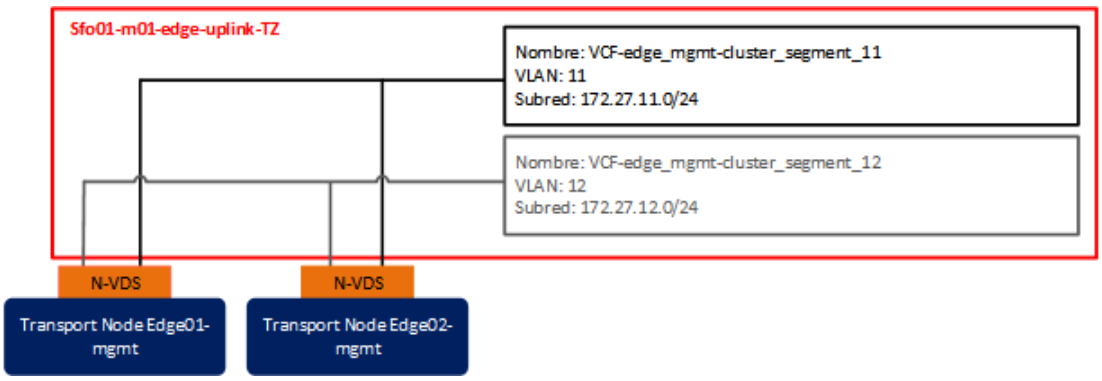

Figura 4.10: *Segments* de la *transport zone sfo01-m01-edge-uplink-tz*

En la imagen anterior se muestra la*transport zone* con el nombre *sfo01-m01-edge-uplink-tz* de tipo VLAN. Está extendida en los dos TNs *edge01-mgmt* y *edge02-mgmt* y contiene dos *segments*. Ambos *segments VCF-edge-mgmt-cluster-segment-11* y *VCF-edge-mgmt-cluster-segment-12* son utilizados por las instancias de NSX-T Edge para transmitir el tráfico que proviene de los *segments* donde se despliegan aplicaciones hacia la red externa (esto se explicará con más detalle). Estos *segments* utilizan los *port groups trunk* del vSphere Switch para transmitir su tráfico hacia las interfaces de red físicas de cada host.

Las TZ, tanto las basadas en Overlay como las basadas en VLAN, sirven para comunicar TNs que se encuentran en distintas partes de la infraestructura física (por ejemplo, en distintos racks) como si estuvieran situadas en el mismo dominio broadcast de capa 2 físico. Aquellas que usan Overlay utilizan el protocolo Geneve para crear un túnel entre los puntos de origen y destino por el cual circula el tráfico generado por los *segments* que pertenecen a esa TZ y que debe salir a la red física para alcanzar su destino. El protocolo Geneve añade una cabecera UDP a cada paquete Ethernet que generan los *segments* cuando este sale del TN donde se generó hacia un TN situado en otra red física. La nueva cabecera incorpora un identificador llamado VNI y es único para cada *segment* (cada *segment* tiene su propio identificador VNI).

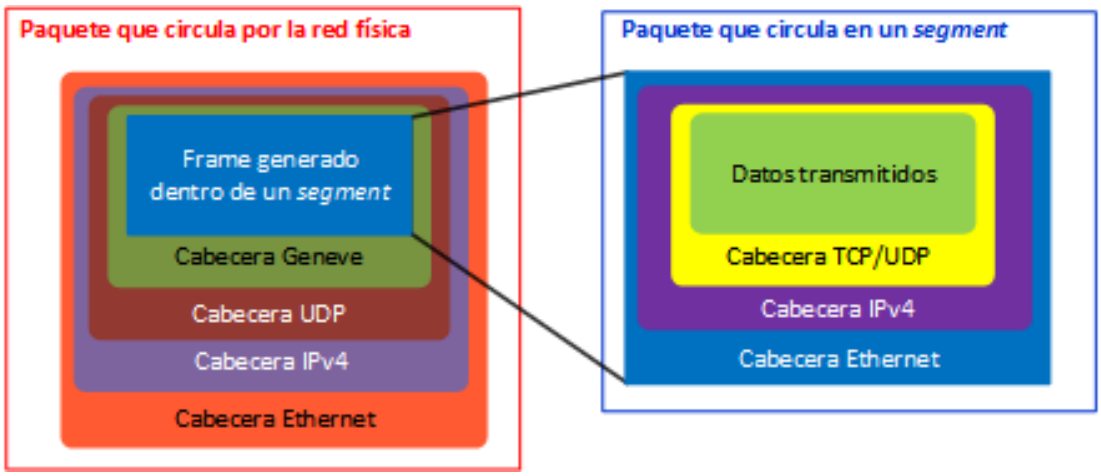

Figura 4.11: Cabeceras de un paquete de red encapsulado con Geneve.

En la imagen anterior se muestran las cabeceras que forman un paquete de red que pertenece a un *segment* cuando este sale de un TN. Cuando el paquete que circula por la red virtual (contiene los datos que se transmiten y la información de las VMs origen y destino) tiene salir a la infraestructura física para alcanzar el destino, el *transport node* añade la cabecera del protocolo Geneve con el identificador VNI correspondiente, una cabecera UDP con un puerto origen y destino establecidos por defecto, una cabecera IP con las direcciones IP origen y destino de los TN que se están comunicando, y una cabecera Ethernet donde se idican las direcciones MAC de los TN origen y destino. Así, dos VMs conectadas al mismo *segment* pero alojadas en TNs de dominios broadcast diferentes, pueden comunicarse de forma transparente como si estuvieran conectadas directamente una con la otra en la misma subred. En las TZ de tipo VLAN, tráfico de sus *segments* es encapsulado añadiendo un identificador de VLAN definido en una plantilla *Uplink Policy*[6] que se asigna a la TZ. Se aplica la misma VLAN para encapsular el tráfico de todos los *segments* de una misma TZ de tipo VLAN. Estas plantillas permiten establecer como debe el N-VDS de un TN tratar el tráfico de la *transport zone* a la que se asigna. En cada una se especifican varias *Teaming Policy*, la VLAN que debe usar el N-VDS cuando tiene que enviar el tráfico fuera del TN y el MTU de cada interfaz *uplinks*. Una *Teaming Policy* indica como el N-VDS utiliza los *uplinks* a nivel de *segment* para distribuir el tráfico y conseguir conexiones redundantes y balanceo de la carga, se especifica una *Teaming Policy* por defecto más otras adicionales. Se aplica una *Uplink Policy* a la TZ de tipo VLAN *sfo01-m01-dge-uplink-tz*.

[6] A las TZ de tipo Overlay no se les aplica ninguna *Uplink Policy*.

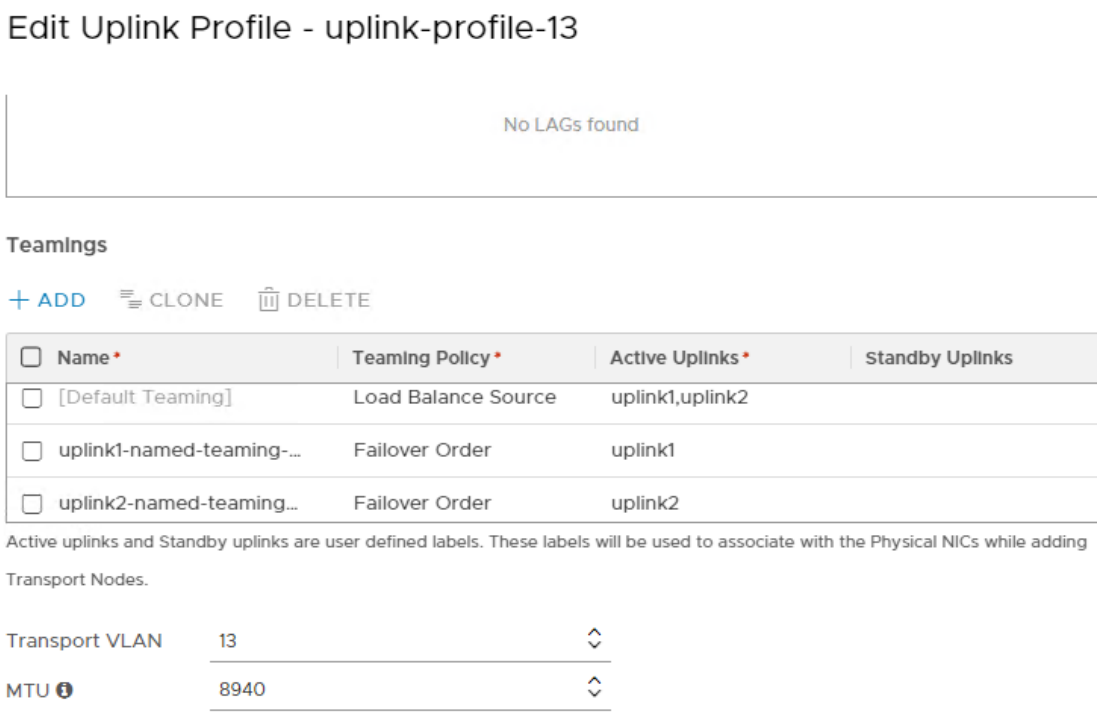

Figura 4.12: *Uplink Policy* configurada para la *transport zone sfo01-m01-dge-uplink-tz*.

En la imagen anteriro se muestra la *Uplink Policy* (*uplink-profile-13*) configurada para la TZ *sfo01-m01-dge-uplink-tz*. Se establece la VLAN 13 como Transport VLAN, utilizado para encapsular el tráfico saliente hacia la red física, y MTU de 8940 Bytes para los uplinks. Hay definidas tres *Teaming Policies*, una por defecto (*Default teaming*) que utiliza *Load Balance Source* para hacer un mapeo uno a uno entre las interfaces de cada VM y uno de los *uplinks* (todo el tráfico correspondiente a esa interfaz se envía y recibe por el mismo *uplink*), y dos adicionales *uplink1-named-teaming-policy* y *uplink2-named-teaming-policy* que utilizan *Failover Order* donde se establece un *uplink* como activo por donde se transmite todo el tráfico y otros de reserva que se usan en caso de que el *uplink* activo falle (en una de las políticas todo el tráfico se reenvía por *uplink1* y en la otra por *uplink2*).

A los *segments VCF-edge-mgmt-cluster-segment-11* y *VCF-edge-mgmt-cluster-segment-12* pertenecientes a la TZ *sfo01-m01-dge-uplink-tz* se les asigna la *Teaming Policy uplink1-named-teaming-policy* y *uplink2-named-teaming-policy* respectivamente. De esta forma se consigue que el tráfico que circula por cada uno de ellos solo utilice un único *uplink*.

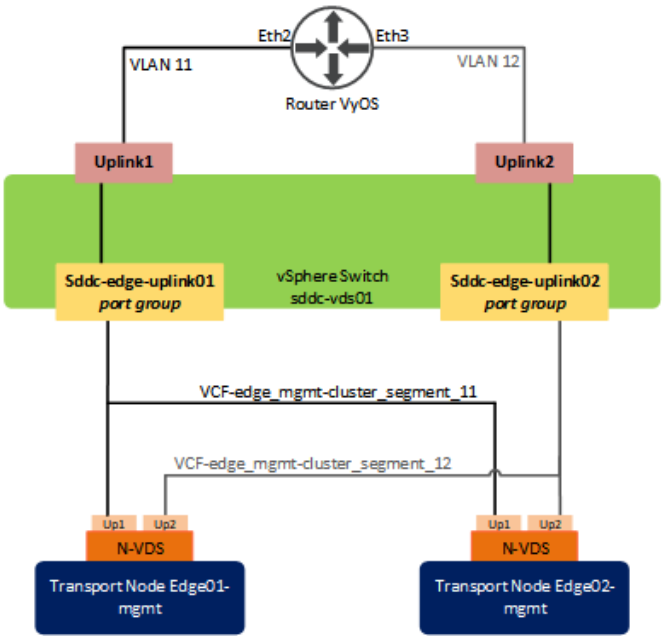

Figura 4.13: Topología de red de las insterfaces *uplink*.

En la imagen anterior se muestra la topología que forman las instancias de NSX-T Edge para generar la redundancia y alta disponibilidad en las rutas hacia la red externa para las aplicaciones cuya red está gestionada por VMware NSX-T, desde el punto de vista de VMware vSphere. Cada TN Edge posee dos interfaces *uplink* que durante su configuración se mapea cada una con uno de los *trunk port groups* de vSphere Switch. Las interfaces *uplink* que se indican en la *Teaming Policy* se refieren a las de la instancia de NSX-T Edge, no las de vSphere vSwitch, por lo tanto cada uno de los *segments* utiliza solo uno de los *uplinks* que combinado con la configuración de los *port groups* de vSphere vSwitch establecen las rutas que se muestran en la imagen.

Esta encapsulación, tanto VLAN como Overlay, tiene lugar cuando los paquetes salen de la interfaz de una VM y entran en el N-VDS del TN. Para ello, cada TN tiene dispositivo llamado *Tunnel End Point* (TEP) al que se le asigna una dirección IP utilizada para enviar y recibir el tráfico entre VMs que se encuentran en el mismo *segment* pero se alojan en TNs situados en redes L2 diferentes[7]. Los TNs que son hosts ESXi obtienen su dirección TEP de un servidor DHCP[8] mientras que los que son instancias de NSX-T Edge la dirección IP se asigna de forma manual. El TEP de cada TN tiene dos direcciones IP asignadas puesto que cada uno tiene dos interfaces de red, *esxi-1* tiene las direcciones 172.16.254.10 y 172.16.254.11, *esxi-2*

[7]En el entorno desplegado todos los TN se encuentran dentro de la misma red física. Esto implica que no se genere tráfico con los TEPs ya que todas las TZs funcionan sobre un único dominio broadcast.

[8]El servidor DHCP hace que se simplifique el proceso de configuración de un nuevo host ESXi ya que le asigna una dirección IP de forma automática.

tiene las direcciones 172.16.254.12 y 172.16.254.13, *esxi-3* tiene las direcciones 172.16.254.14 y 172.16.254.15, *esxi-4* tiene las direcciones 172.16.254.16 y 172.16.254.17, *edge01-mgmt* tiene las direcciones 172.27.13.2 y 172.27.13.3, y *edge02-mgmt* tiene las direcciones 172.27.13.4 y 172.27.13.5.

Al crear un *segment* dentro de una TZ, se configura un modo de replicación que indica como se retransmite el tráfico Broadcast, Multicast y Unknown Unicast propio del *segment* cuando este tiene que viajar a un TN que está en una ubicación distinta en el medio físico. El modo de replicación que se utiliza en todos los *segments* es *Two-Tier Hierarchical Mode.*

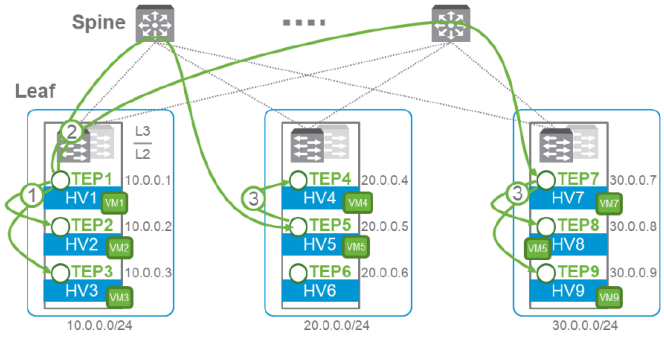

Figura 4.14: Modo de replicación *Two-Tier Hierarchical.*

En la imagen anterior se muestra un ejemplo del funcionamiento del modo de replicación *Two-Tier Hierarchical.* (1) Cuando el TN HV1 envía tráfico BUM a la red primero lo hace a los TNs que están en su mismo dominio, (2) después envía una copia de ese tráfico a un único TN de cada dominio broadcast donde exista el *segment* propietario del tráfico y, finalmente, (3) el TN de cada localización lo retransmite al resto de TNs de su dominio que lo requieran. De este modo se reduce el número de paquetes que el TN origen debe enviar a través de la red física.

El objetivo es crear nuevas subredes, es decir *segments* que se expanden por los distintos TNs adheridos a la *transport zone* correspondiente, y conectarlas a un router virtual para al final formar subredes distribuidas con servicios de red también distribuidos y virtualizados, todo gestionado desde VMware NSX-T y sin tener que configurar la red física. Para completar esto, VMware NSX-T introduce routers virtuales que se encuentran embebidos y distribuídos dentro del hypervisor ESXi de cada TN y que proporcionan enrutamiento entre *segments* y servicios de red distribuídos. Permiten definir un *gateway* para cada *segment* a través del cual las VMs conectadas pueden acceder a la red externa y los servicios de red. La herramienta VLC despliega para el *management domain* un modelo de enrutamiento de doble capa (*Two Tier Routing*) donde se utilizan dos routers virtuales, *Tier-0* (*mgmt-domain-tier0-gateway*) dedicado a gestionar el acceso a la red externa a través del router VyOS con conexiones redundantes, y *Tier-1* (*mgmt-domain-tier1-gateway*)que gestiona el enrutamiento entre *segments* y

proporciona servicios de red a las VMs.

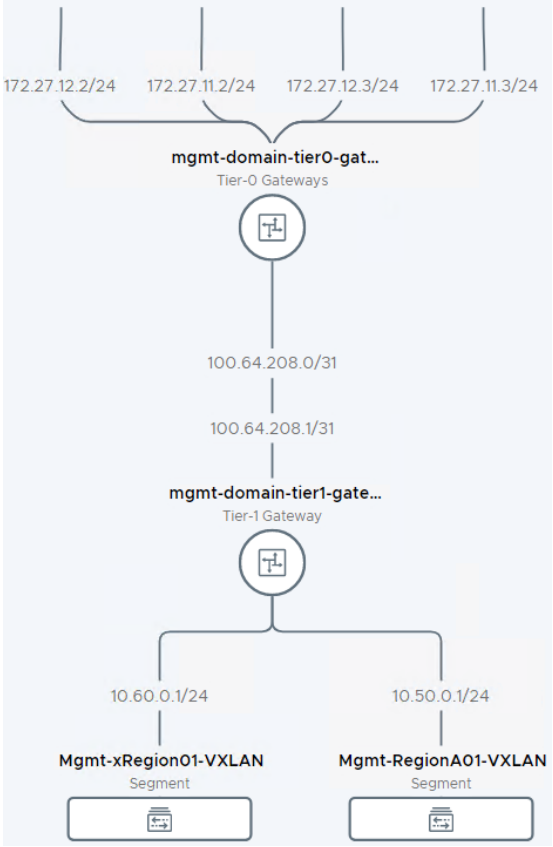

Figura 4.15: Topología de los routers virtuales de *Tier-1* y *Tier-0*.

En la imagen superior se muestra la topología del modelo de dos routers virtuales y los *segments* a los que se conecta cada uno desde el punto de vista de NSX-T. El router de Tier-1 enruta el tráfico entre los *segments* a los que está conectado y hacia el router de Tier-0 que se encarga de transmitir el tráfico hacia la red física.

Un router router lógico está formado por dos componentes:

- **Distributed Router** (DR): que gestiona el enrutamiento y se a subredes a través de sus interfaces lógicas. Estas subredes pueden ser *segments* u otro DR. A cada interfaz se le asigna una dirección MAC y una dirección IP que representa el *gateway* de la subred. Este componente está distribuído en todos los TN, tanto hosts ESXi como instancias de NSX-T Edge manteniendo la misma configuración (interfaces,tablas de enrutamiento, etc.). Su función es redirigir el tráfico que recibe entre las interfaces que tiene disponibles, es decir, enruta el tráfico entre los diferentes *segments* a los que está conectado el router virtual Tiene una interfaz llamada *Internal transit Link* conectada a la red *Internal Transit Network* que se utiliza para conectar todos los DR y SR de un Tier distribuídos

en los TN.

- **Service Router** (SR): proporciona servicios de red de forma centralizada (NAT, DHCP, Load Balancer, VPN, Gateway Firewall y Bridging L2) y proporciona acceso a la red externa. Este componente no está distribuído entre los diferentes TN, solo se se encuentra distribuido en las instancias de NSX-T Edge. Los servicios que proporciona solo se entregan a los recursos cuya red está gestionada por VMware NSX-T. Posee la interfaz Internal Transit Link para comunicarse con el resto de DR y SR pertenecientes al mismo Tier, dos interfaces *External Interface* que se conectan a los segments que dan acceso a la red externa [9], la interfaz *Router Link*[10] que conecta el SR de *Tier-0* con el de *Tier-1*, y la interfaz *Internal transit* que se habilita cuando se activa la opción *Inter SR iBGP*[11] y que comunica los SR de las dos instancias de NSX-T Edge.

[9] La *External Interface* solo existe en *Tier-0* ya que es el router que se comunica con el dispositivo físico.
[10] Router Link Network utiliza por defecto la subred 100.64.0.0/16.
[11] Inter SR iBGP Network utiliza por defecto la subred 169.254.0.0/24.

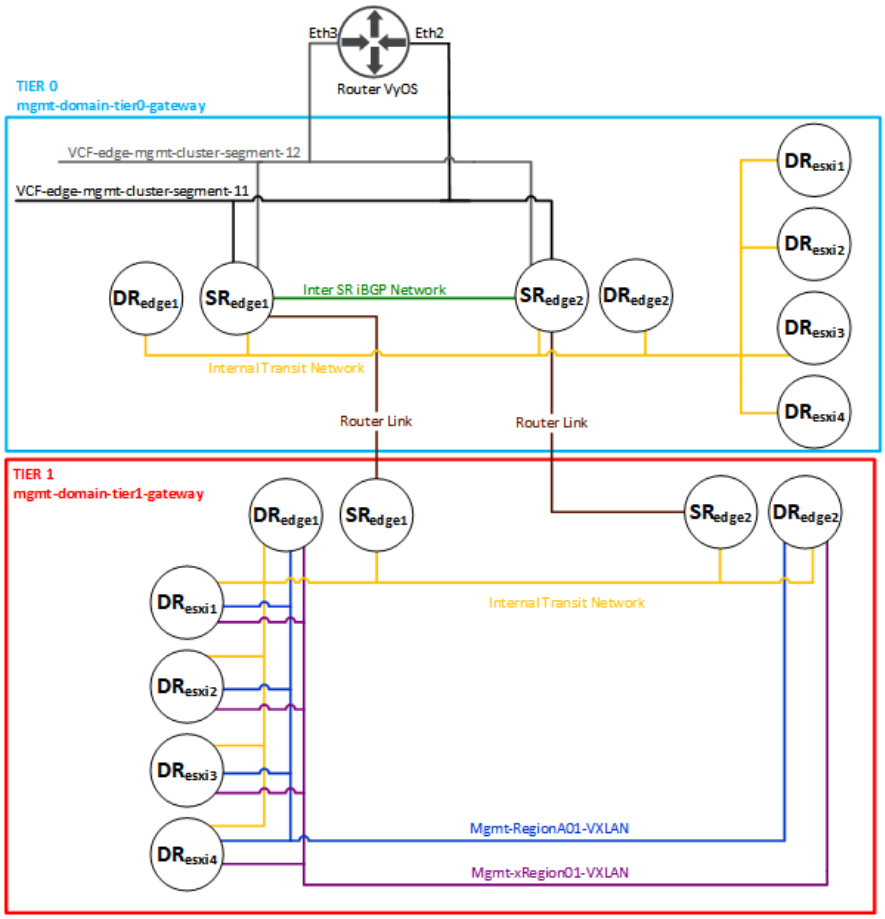

Figura 4.16: Estructura interna de los routers virtuales de *Tier-0* y de *Tier-1*.

En la imagen superior se muestran los componentes internos de cada router virtual y como estos están distribuidos por todos los TNs. El router *mgmt-domain-tier0-gateway* tiene el componente SR distribuído en dos TN que son las instancias de NSX-T Edge, y el componente DR distribuido en todos los TNs. El SR situado en *edge01-mgmt* está conectado al *segment VCF-edge_mgmt-edge-cluster-segment-11* y usa la dirección IP 172.27.11.2, y otra interfaz conectada al *segment VCF-edge_mgmt-edge-cluster-segment-12* donde usa la IP 172.27.12.2, mientras que el SR de *Tier-0* situado en *edge02-mgmt* se conecta a los mismos *segments* pero usando las direcciones 172.27.11.3 y 172.27.12.3 respectivamente. El router *mgmt-domain-tier1-gateway* tiene el componente SR distribuido en las dos instancias de NSX-T Edge y el componente DR distribuido por en todos los TNs. Ambos SR están conectados con los SR de *Tier-0* para transmitir el tráfico hacia la red física. Los DR de *Tier-1* están conectados a los *segments Mgmt-RegionA01-VXLAN* y *Mgmt-xRegion01-VXLAN* para proporcionar enrutamiento y servicios a las VMs situadas en ellos.

La razón de tener dos capas de enrutamiento se debe a la configuración del router *mgmt-*

domain-tier0-gateway. En este router se utiliza BGP para comunicarse con el router físico a través de las *external interfaces* (se establece 65003 como *autonomous system* local)[12], la opción *Inter SR iBGP* establecer una comunicación mediante BGP entre las instancias distribuídas del componente SR de *Tier-0* para que se pueda seguir transmitiendo el tráfico en caso de que alguna de las interfaces de una instancia de NSX-T Edge esté fuera de servicio, se activa el protocolo ECMP para balancear el tráfico hacia la red física entre los caminos disponibles ya que en conjunto, en el router de *Tier-0* existen cuatro rutas posibles para alcanzar el router VyOS. Un aspecto importante es la configuración de la disponibilidad de un router virtual ya que determina el modo en el que se va a ejecutar. En el router de *Tier-0* se selecciona el modo *active-active* el cual implica que las dos instancias de SR funcionan ambas de forma activa, es decir, el tráfico que reciba cada una será encaminado por una subred diferente proporcionando así mayor ancho de banda, mayor disponibilidad y mayor escalabilidad. Esto esto último no es compatible con servicios de red centralizados, por ello es necesario desplegar al menos un router lógico de *Tier-1* (*mgmt-domain-tier1-gateway*) configurado con el modo *active-stanby* el cual solo se utiliza una de las dos instancias del SR de *Tier-1* por lo tanto el tráfico dirigido a este router siempre será recogido en un único punto. Así, en el router *mgmt-domain-tier1-gateway* es donde se pueden desplegar servicios de red como NAT, Load Blancing, DNS y VPN, para los *segments* que están conectados.

4.3.3 Operaciones de la Arquitectura

El entorno ya está configurado para funcionar como un SDDC, a partir de este punto ya no es necesario realizar ninguna modificación en la infraestructura física ya que todas las tareas que se deben realizar están dentro del alcance de los componentes de VMware Cloud Foundation. Para finalizar la construcción del SDDC y habilitar un servicio donde los usuarios puedan aprovisionar recursos bajo demanda, se instalarán sobre el entorno desplegado las aplicaciones Workspace ONE Access [13] (WSA) y VMware vRealize Automation (vRA). La primera permite al administrador conectar con el servidor de usuarios Active Directory y gestionarlos para proveer un servicio de autenticación centralizado a múltiples aplicaciones como VMware vRealize Automation. La segunda aplicación permite a los usuarios aprovisionar recursos de forma automatizada desde un catálogo de recursos. VMware vRealize Suite Licfecycle Manager (vRSLCM) es el componente que permite administrar vRA y WSA, su instalación y actualizaciones, las contraseñas de administrador y sus certificados, para ello necesita comunicarse con VMware vCenter Server. Se desplegará una instancia de cada componente en el *management domain* creado anteriormente y estarán conectadas al *segment*/subred *Mgmt-*

[12] El uso de BGP simplifica la configuración de nuevas rutas cuando se añaden componentes al entorno, y que no se pierda la conectividad en caso de caída de alguna de las interfaces. Este protocolo también se configura en las interfaces del router Vyos

[13] VMware vRealize Identity Manager

xRegion01-VXLAN.

Workspace One Access

Los usuarios que necesiten acceder a vRA deben estar registrados en el directorio de Workspace One Access. Este componente centraliza el acceso de todos los productos de VMware vRealize. Cuando se despliega se debe configurar un Active Directory que en el caso del entorno está situado en la VM con Windows Server 2016. Dentro del Active Directory existen grupos de seguridad y perfiles de usuario, un perfil de usuario contiene información como nombre, apellidos, dirección e-mail, nombre de usuario y contraseña[14], y este puede formar parte de varios grupos de seguridad. Una vez configurado, cada aplicación se conectará a WSA y se podrán asignar roles para los grupos de seguridad y usuarios estableciendo así un nivel de acceso. Además, cada usuario registrado tendrá disponible un catálogo de aplicaciones en el portal de WSA cuyo administrador establecerá que aplicaciones están habilitadas para cada usuario o grupo, eso sí, para que el usuario pueda acceder a ella previamente se debe establecer un rol para ese usuario dentro de la aplicación.

Figura 4.17: Muestra los usuarios definidos en el Active Directory sincronizados en Workspace One Access.

En la Figura 4.17 se muestran los dos usuarios definidos en el Active Directory y dos usuarios que se corresponden a los perfiles de administración de WSA, no se utilizarán grupos de seguridad para reducir la complejidad pero su configuración en las aplicaciones de VMware es igual que para los perfiles de usuario. En un entorno real existen usuarios que controlan a otros usuarios y establecen su nivel de acceso, a parte de los perfiles de administrador de cada aplicación. Para el entorno se define el perfil *adminuser* que será el encargado de gestionar el acceso de dos usuarios (*baseuser1* y *baseuser2*) que serán los que consuman a las aplicaciones desplegadas (vRSLCM y vRA). El primero tendrá acceso y permisos de edición en las aplicaciones vRSLCM y vRA, mientras que los dos usuarios base solo podrán acceder a vRA y dentro de este el usuario admin definirá que servicios están habilitados para cada uno.

VMware vRealize Automation

El punto a través del cual los usuarios pueden aprovisionar sus recursos es vRealize Automation. Este producto provee el servicio cloud. Internamente vRA se divide en varios servicios que permiten gestionar los diferentes aspectos de la cloud. Para centrarse en los objetivos de

[14]Se pueden configurar más campos pero los que se describen son los obligatorios a la hora de crear un usuario.

este proyecto solo se hace referencia a dos de esos servicios, el primero es Cloud Assembly el cual permite administrar la infraestructura disponible controlar el uso que se hace de esos recursos, y el segundo es Service Broker, utilizado por los usuarios para aprovisionar los recursos desde un catálogo de plantillas. La obtención de los recursos por parte del usuario se hace desplegando una serie de plantillas llamadas Blueprints diseñadas previamente, en donde se define un conjunto de VMs y recursos de red y de almacenamiento incluyendo otros aspectos como la configuración de cada uno de los recursos, como redes de la infraestructura que se utilizan, cantidad de almacenamiento, o la ubicación del despliegue en la infraestructura. Son ficheros de código con extensión *.yaml* donde se indican etiquetas, aunque también se pueden diseñar con un editor gráfico. Estas plantillas están relacionadas con proyectos, una plantilla pertenece a uno o varios proyectos donde existe un coordinador de proyecto que se encarga de diseñar Blueprints y de administrar los usuarios miembros de ese proyecto. Los proyectos de vRA permiten limitar los recursos para que un conjunto de usuarios pueda desplegar los componentes definidos en las Blueprints disponibles, como la cantidad de memoria RAM, cantidad de instancias que se pueden desplegar y cantidad de almacenamiento, también aquellas redes que se pueden utilizar. Desde el punto de vista de vRA, la infraestructura se divide en Cloud Zones, las cuales son conjuntos de recursos situados en distintos proveedores Cloud que pueden ser públicos como AWS o Azure, o privados que solo pueden ser clusters vSphere. En el caso del entorno desplegado solo se tendrá una única Cloud Zone de tipo vSphere. En cada Cloud Zone se define como se deben distribuir los recursos aprovisionados sobre la infraestructura. Finalmente será el administrador de la infraestructura el que se encargue de proveer los recursos, administrar los proyectos disponibles, gestionar los coordinadores de cada proyecto y controlar y limitar el uso de los recursos.

Apéndices

Glosario

Tenencia múltiple : principio de la arquitectura de software donde una aplicación se sirve a varios clientes desde una misma instancia.

SDDC : *Software Defined DataCenter* es un modelo de infraestructura donde se virtualiza la abstracción, gestión y automatización de todos los recursos y servicios de un centro de datos.

Hipervisor baremetal : software instalado sobre el hardware de un servidor que permite instalar aplicaciones que funcionan sobre entornos virtuales directamente sobre el hardware.

Máquina virtual : software que emula un conjunto de recursos físicos para ejecutar otro software de forma aislada.

Datastore : contenedores que VMware vSphere utiliza para el almacenamiento archivos en un único lugar o a través de una red. Suelen utilizarse para almacenar ficheros de máquinas virtuales y puedent tener el formato VMFS, NFS o NAS.

Modelo multi-tenant : es un modelo de desarrollo donde una misma aplicación se entrega a distintos usuarios sin hacer un desarrollo específico para cada uno de ellos.

RAID 5 : Es un conjunto de discos duros que funciona como una única unidad de almacenamiento para aumentar el rendimiento y la eficiencia. RAID 5 necesita como mínimo tres discos duros, y distribuye la información de paridad en todos los discos (esta información permite recuperar datos corruptos a partir del resto de información no perdida).

Almacén de datos

LUN : *Logical Unit Number* es un identificador que agrupa un agrupa un conjunto o subconjunto de almacenamiento físico o virtual. Puede asignarse a un disco completo o solo a una parte.

Controlador SFP+ : módulo transceptor óptico que se utiliza en las telecomunicaciones y aplicaciones de transmisión de datos. Soportan Sonet, canal de Fibra y Gigabit Ethernet.

SAN : *Storage Area Network* es una red dedicada al almacenamiento, de alta velocidad con canal de Fibra o iSCSI, con equipos de conexión dedicados (p.e. switches) y con dispositivos de almacenamiento (discos duros).

VMFS : sistema de archivos de alto rendimiento nativo de VMware vSphere. Se utiliza para implementar los almacenes de datos y está optimizado para el almacenamiento de máquinas virtuales.

Platform Services Controller (PSC) : componente de la infraestructura que agrupa los servicios de infraestructura de un entorno vSphere. Estos servicios son la concesión de licencias, administración de certificados y la autenticación con vCenter Single Sign-On.

Cluster : Conjunto de dos o más Hosts para aprovisionar recursos.

Servicio LBT : servicio que se encarga de balancear el tráfico que entra en cada interfaz de un switch.

vCPU

Jumbo Frame : son los paquetes que se transmiten por una red y cuyo MTU es mayor a 1500.

VTEP : *VXLAN Tunnel End Point* es un componente del protocolo VXLAN cuya función es encapsular y desencapsular las tramas correspondientes a una VXLAN. Este componente se encuentra al principio y al final del camino que sigue una trama.

NIC : *Network Interface Controller* es un componente físico que conecta el host con una red.

Log : registro que muestra información sobre un evento que afecta a un proceso en particular.

CPD : Centro de Procesamiento de Datos es un espacio donde se encuentran los recursos necesarios para procesar información.

Pool de recursos : representa una partición de recursos disponibles que no se crean ni se eliminan bajo demanda.

CoS : *Class of Service* es un campo de la cabecera de un paquete Ethernet que determina su prioridad cuando se utiliza etiquetado VLAN. Es un protocolo QoS de capa 2.

DSCP : *Differentiated Services Code Point* es campo de la cabecera IP que forma parte del protocolo QoS de capa 3 *DiffServ*, y que sirve para clasificar el tráfico según servicios.

VLAN trunk : enlace que permite comunicar distintas VLANs.

ARP : protocolo responsable de obtener la dirección MAC que corresponde a una dirección IP.

Rack : armario metálico destinado a alojar servidores físicos.

DHCP helper address : elemento que permite retransmitir el tráfico broadcast de un servidor DHCP por múltiples redes.

Bibliografía

[1] "Vmware vsphere entreprise edition datasheet." [Online]. Available: https://www.vmware.com/files/es/pdf/VMware-vSphere-Entreprise-Edition-Datasheet.pdf

[2] T. G. Peter Mell, "The NIST Definition of Cloud Computing." [En línea]. Disponible en: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

[3] CITC, "Centro de Procesado de Datos." [En línea]. Disponible en: https://www.citic.udc.es/instalacion/centro-de-despliegue.html

[4] VmWare, "Cloud foundation components." [En línea]. Disponible en: https://docs.vmware.com/en/VMware-Cloud-Foundation/4.0/rn/VMware-Cloud-Foundation-40-Release-Notes.html#swversions

[5] V. vSAN, "vsan disk groups and data storage architecture: Hybrid or all-flash." [En línea]. Disponible en: https://youtu.be/PDcLgV37FP4?list=PLjwkgfjHppDux1XhPB8pW3vS43Aglfq2c

[6] V. C. Foundation, "Vmware software licenses." [En línea]. Disponible en: https://docs.vmware.com/en/VMware-Cloud-Foundation/3.9/com.vmware.vcf.planprep.doc_39/GUID-202ECBCF-2CAA-4167-BA54-4EE1169D312C.html

[7] VMware, "Managing resource pools." [En línea]. Disponible en: https://docs.vmware.com/en/VMware-vSphere/6.7/com.vmware.vsphere.resmgmt.doc/GUID-60077B40-66FF-4625-934A-641703ED7601.html