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: Antonio Daniel López Rivas
Jose Carlos Dafonte Váquez

A Coruña, setembro de 2020.

Resumen

El Cloud Computing es un modelo que permite acceder a un conjunto de recursos, como por ejemplo redes, almacenamiento y cómputo, los cuales 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 consiste en desplegar un servicio Cloud en el CITIC, usando como base la infraestructura y herramientas ya existentes. El servicio debe proveer un sistema de autenticación para que cada usuario pueda acceder con sus credenciales de la UDC a una plataforma, la cual le permita aprovisionar y gestionar recursos de forma automatizada. Además, también debe automatizar la gestión de todos los componentes de la infraestructura, incluyendo la gestión de perfiles de usuarios, el control de la cantidad de recursos disponibles para los usuarios y el despliegue de aplicaciones, con el fin de liberar a los administradores de las tareas más redundantes y repetitivas, para así obtener el máximo rendimiento de la infraestructura disponible en el CITIC. Con el objetivo de 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
- Virtualización
- SDDC
- Aprovisionamiento

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1
Introducción

Según *National Institute of Standards and Technology* (NIST), «Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources»[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*.
- *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 servicio Cloud es capaz de obtener y abstraer información acerca del consumo de recursos para monitorizarlos, controlarlos e informar al usuario y al proveedor.

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 virtualización la empresa VMware, que no cuenta con las herramientas suficientes para ser accedido por todos los usuarios. Este servicio de virtualización permite aprovisionar recursos de un conjunto de servidores en forma de máquinas virtuales, con unas

especificaciones establecidas por el usuario, para realizar tareas que requieren gran capacidad de cómputo, de almacenamiento o de red.

El sistema cuenta con una plataforma de autenticación pero los usuarios a los que está destinado el servicio no tienen acceso, ya que no existe una herramienta que permita gestionar los perfiles de usuario ya existentes. El único modo de habilitar el acceso consiste en que el administrador cree una perfil de forma manual dentro del servicio. También carece de una plataforma donde cada usuario solo tenga acceso a sus recursos, en la vista actual pueden tener visibilidad y acceso a los recursos de otros usuarios dependiendo de los permisos que se hayan asignado a la cuenta. Además, en la plataforma actual la forma de aprovisionamiento de recursos mediante la creación de máquinas virtuales es compleja por tener una interfaz poco intuitiva y difícil de manejar para un usuario que no es administrador del sistema, a parte de que todo el proceso debe realizarse manualmente, esto implica que la monitorización, control y medición de los recursos que aprovisiona cada usuario sean también complejas. La falta de automatización y simplicidad en el sistema provoca que el administrador tenga que gestionar todo el entorno de forma manual, tanto los perfiles de usuarios como los recursos y su configuración, lo cual genera un gran coste y aumenta los riesgos de la infraestructura.

Como se puede observar, el servicio no cumple con las características que define el NIST para un servicio de Cloud Computing, especialmente en lo que se refiere al aprovisionamiento bajo demanda, a la elasticidad y a la monitorización y control de los recursos. Por ello, usando como base la definición de servicio Cloud Computing, en este proyecto se desplegará un conjunto de servicios sobre la infraestructura del CITIC que juntos permitan habilitar un servicio al que los usuarios puedan acceder autenticándose con sus credenciales de la UDC, aprovisionar recursos de red, almacenamiento y cómputo, y que realice mediciones y monitorización sobre los recursos que cada usuario posee. Además, para facilitar las tareas de administración, el servicio debe automatizar las operaciones de obtención de recursos y permitir al administrador limitar la cantidad de recursos que un usuario puede aprovisionar para evitar que estos sean infrautilizados. Con estas mejoras se busca construir un servicio que sea útil, dinámico, sencillo de administrar, que optimice el uso de los recursos de la infraestructura y que aumente su eficiencia gracias a la automatización de tareas y al aprovechamiento de elementos que ya se encuentran disponibles.

1.1 Motivación

La motivación para realizar este proyecto es crear un servicio Cloud en el CITIC para proporcionar recursos a aquellos usuarios que necesiten equipos de grandes prestaciones y que los puedan conseguir de una forma sencilla y ágil, a la vez que se mejora la gestión interna del servicio para así reducir los costes e incidencias de la infraestructura 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.
- Limitar y controlar la cantidad de recursos que un usuario puede aprovisionar y así evitar tener recursos ociosos.
- Automatizar las tareas de administración y configuración de la infraestructura.
- 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 hosts 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 hosts 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 host. 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 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 formando un *pool* de almacenamiento que se divide en cinco LUNs (Logical Storage Unit) 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 realizar un despliegue en el sistema de virtualización que requiera 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 generando un gran coste de tiempo. Por lo tanto, este sistema de almacenamiento no permite ajustar de forma precisa, rápida y bajo demanda la configuración de almacenamiento que un usuario requiera para sus aplicaciones.

2.1.3 Red

El sistema de almacenamiento forma una SAN, para ello las conexiones que se implementan entre los hosts y las cabinas donde se encuentran los discos duros son de tipo 10 Gbit. Para soportar esta conexión, cada cabina incorpora dos controladores con conectores de tipo SFP+. 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. Para mantener la disponibilidad del acceso al sistema de almacenamiento y las comunicaciones entre los hosts, 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 host físico está instalado el **hipervisor ESXi** de tipo baremetal. Sobre los hosts corre una VM que alberga el servicio **VMware vCenter Server** el cual actúa como centro de administración de todas las máquinas virtuales (VMs) y hosts 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 e interfaz

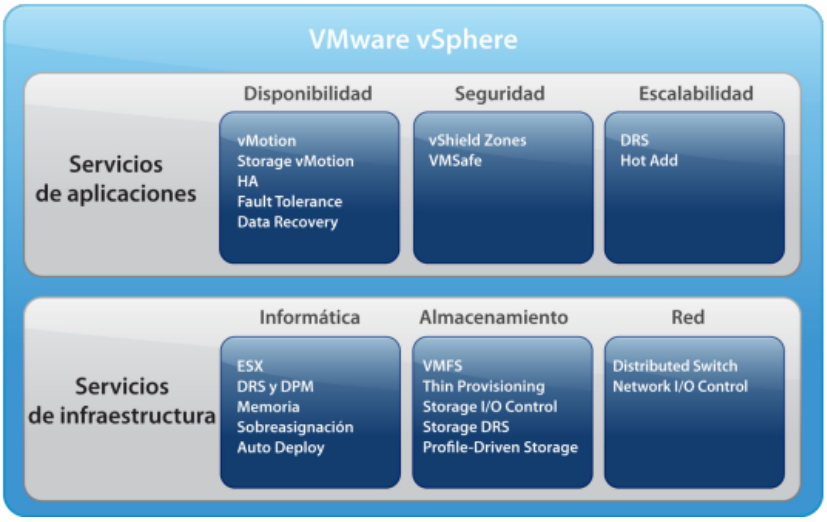

Figura 2.1: Componentes de VMware vSphere[1]

de VMware vCenter Server se realiza a través de la página **vSphere Web Client** donde el usuario puede autenticarse y gestionar las VMs y hosts 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 host 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 host a otro de forma transparente y sin detener su ejecución, permitiendo planificar operaciones de mantenimiento. 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 host del entorno. A diferencia de vMotion, este solo actúa en caso de que la VM o el host 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 host. 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 host diferente. Este servicio está orientado a proteger aquellas tareas que requieren un alto rendimiento o que son críticas.

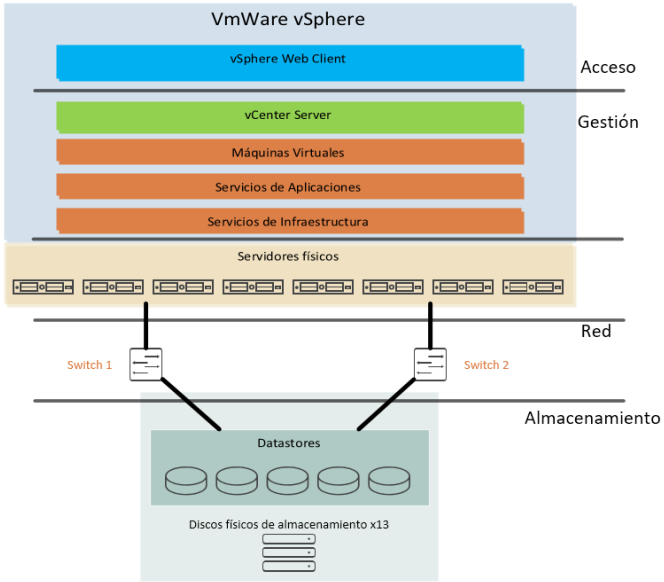

Figura 2.2: Componentes físicos y software que forman la infraestructura actual.

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 estos servicios juntos 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.

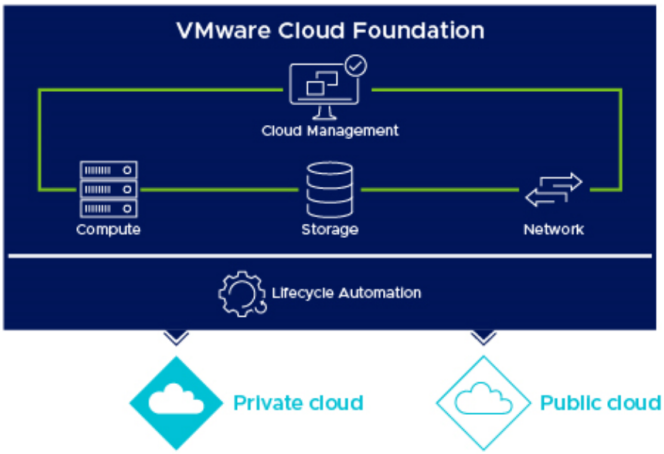

Figura 2.3: Resumen partes de VMare Cloud Foundation.

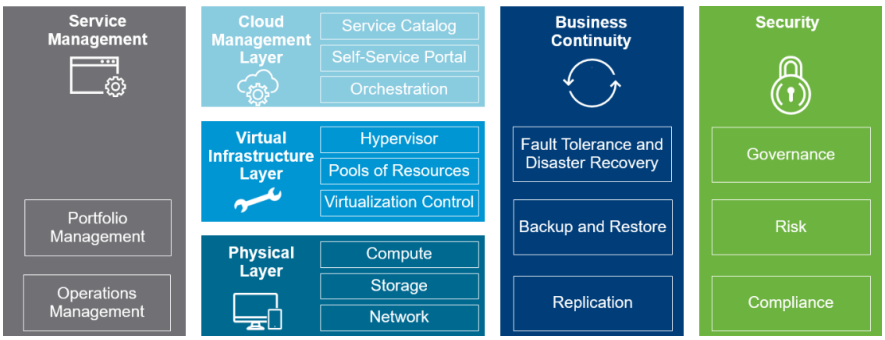

Figura 2.4: Elementos de un SDDC gestionado con VMware Cloud Foundation.

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.

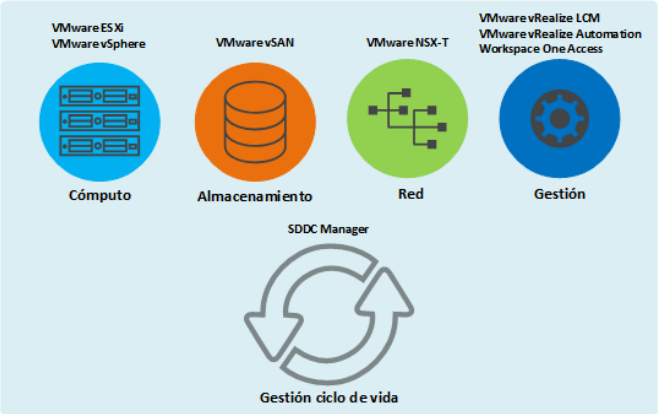

Figura 2.5: Partes de un SDDC y componentes de VCF que las implementan.

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 elemento 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 host 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 que se organizan en grupos ligados a un host (un máximo de cinco grupos por host). Los grupos pueden tener configuración *Hybrid* que combinia discos HDD y

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

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.

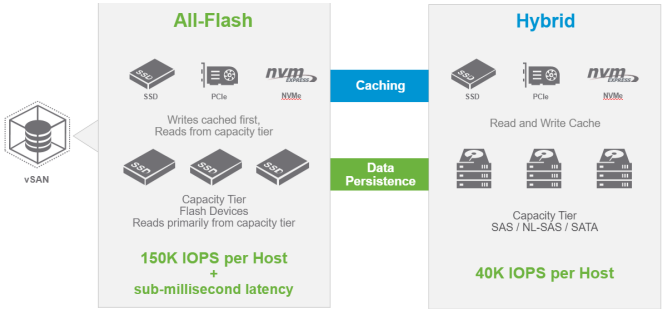

Figura 2.6: Configuración *All-Flash* y configuración *Hybrid* en vSAN

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 hosts 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 host 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.

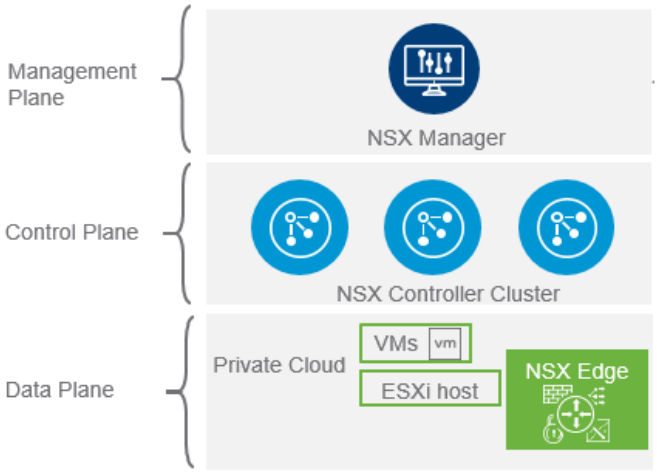

Figura 2.7: Componentes de VMware NSX-T y capas en las que se dividen

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 aprovisionar 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.

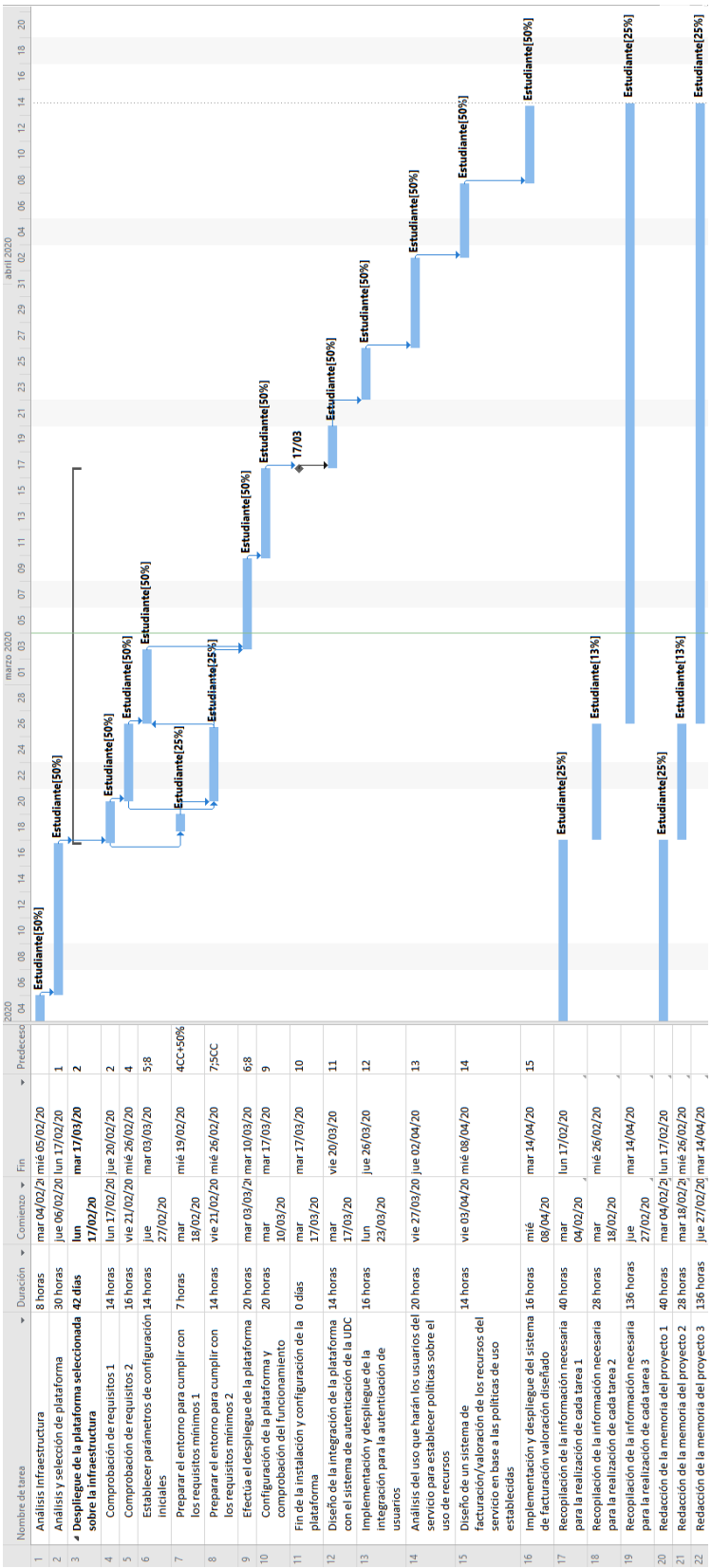

Figura 3.1: Diagrama de Grantt sobre la planificación del proyecto.

	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, formado por 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 y organizar el acceso 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 WD existentes. En este *workload domain* se genera un cluster de VMware vSphere donde se despliegan 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 controla sus redes virtuales, y VMware SDDC Manager permite administrar los aspectos que afectan a todo el SDDC como la instalación de otras aplicaciones de VMware, la creación de nuevos *workload domains* o la instalación de actualizaciones.

Virtual Infrastructure Domain (VI)

Este tipo de WD se crea manualmente y bajo demanda desde Management Domain para habilitar entornos cuyos recursos puedan ser usados por los usuarios. 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 que se van a realizar en el. Con cada WD se genera un nuevo cluster de VMware vSphere que agrupa los nuevos recursos, aunque parte de sus componentes que se despliegan se controlan desde el Management Domain:

- Una VM de VMware vCenter Server que se sitúa en el Management Domain.
- 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 WD desde VMware NSX-T Manager situado también en el Management Domain. Los usuarios despliegan sus aplicaciones sobre los recursos de este WD, de esta forma, las tareas administrativas y las de consumo se ejecutan desde entornos separados.

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 que contiene todos los componentes de gestión de toda la infraestructura, y al menos un VI Domain. Desde el Management Domain se administra toda la infraestructura del SDDC y cada VI Domain existente, los cuales son creados bajo demanda desde el Management Domain y sus recursos se establecen según su finalidad. Un Management Domain puede gestionar hasta un máximo de de catorce VI Domain.

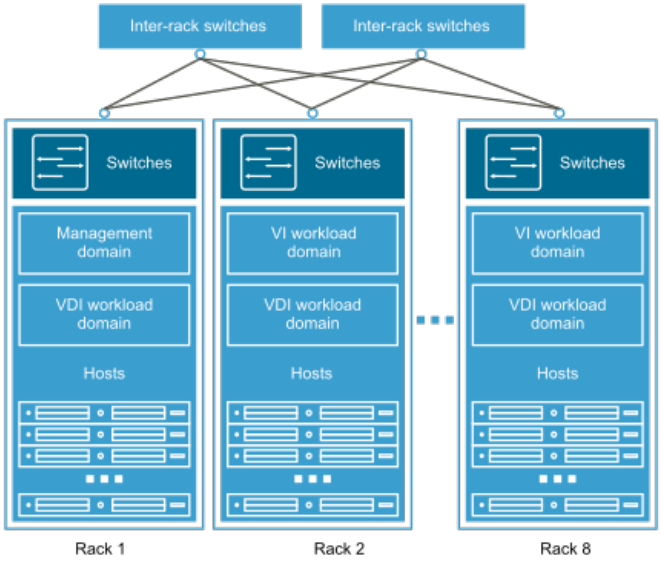

Figura 4.1: Esquema del modelo de arquitectura estándar.

Modelo consolidado

Este modelo está orientado a entornos de tamaño pequeño con menos de siete hosts. Se compone de un único WD que cumple las funciones de un Management Domain y de un VI Domain, es decir, en él se alojarán las instancias de los componentes dedicados a la gestión del SDDC[1] junto con las aplicaciones desplegadas por los usuarios. Así, a diferencia del modelo estándar, las tareas de administración, de aprovisionamiento de recursos y las realizadas por los usuarios, se ejecutan dentro de un mismo entorno. Dentro del cluster de VMware vSphere que se crea, las instancias pertenecientes a los componentes de administración del SDDC y las pertenecientes al trabajo de los usuarios, se colocan en *resource pools* separados. Un *resource pool* es un elemento de VMware vSphere que permite establecer unos límites de consumo de recursos sobre las instancias que se sitúan en su interior[7].

[1]Se despliega la misma cantidad de instancias que en el Management Domain.

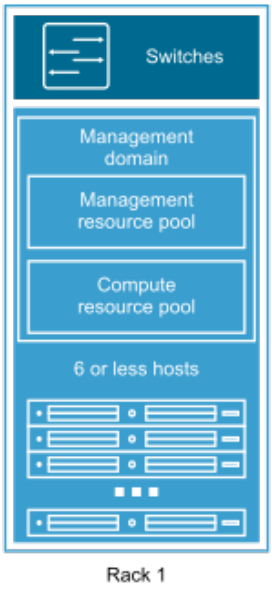

Figura 4.2: Esquema del modelo de arquitectura consolidado.

4.1.3 Clusters, zonas y distribución de un SDDC

Los recursos de un SDDC pueden estar distribuídos en diferentes localizaciones para proporcionar mayor disponibilidad y recuperación ante fallos. Estos recursos, se agrupan para formar una estructura que permite usar y gestionar los recursos disponibles de forma conjunta y dinámica.

Availability Zone, Region y Cluster

- Availability Zone (AZ): se le llama AZ a un conjunto de recursos físicos que forman una infraestructura independiente, es decir, cada una tiene una fuente de energía, un sistema de refrigeración, un sistema de seguridad y una red propias, no compartidas con otra AZ, para evitar la propagación de fallos hacia otras AZs. Cuando existen varias AZs estas se pueden usar de forma que cuando ocurre un fallo en una de ellas la carga de trabajo se distribuye a una segunda AZ para minimizar el tiempo de caída del servicio. Dentro de una AZ se alojan uno o más *workload domains*.

- Region: se le llama Region a un conjunto de AZs situadas en una misma ubicación, es decir, las AZs de una Region son independientes del resto pero se sitúan próximas entre si. Estas AZs deben tener al menos una latencia de 5 ms entre ellas. Una Region contiene al menos una AZ. Dentro de un SDDC puede existir varias Regions pero estas se sitúan en ubicaciones más distantes, la latencia debe ser de al menos 150 ms. Esta estructura permite ofrecer los servicios de un SDDC en diferentes ubicaciones a la vez

que se aumenta su disponibilidad y recuperación ante fallos.

- Cluster: un cluster de VMware vSphere es una agrupación de hosts. A las instancias que están desplegadas sobre ellos se les aplica una configuración de disponibilidad determinada con los componentes Vmware vSphere High Availability y VMware vSphere Distributed Resource Scheduler. Estos servicios permiten determinar como se restablecen las instancias situadas en un cluster cuando falla alguno de los hosts que lo forman. Un cluster está colocado dentro de un WD, por lo tanto sus recursos estarán limitados por el alcance del WD donde esté situado.

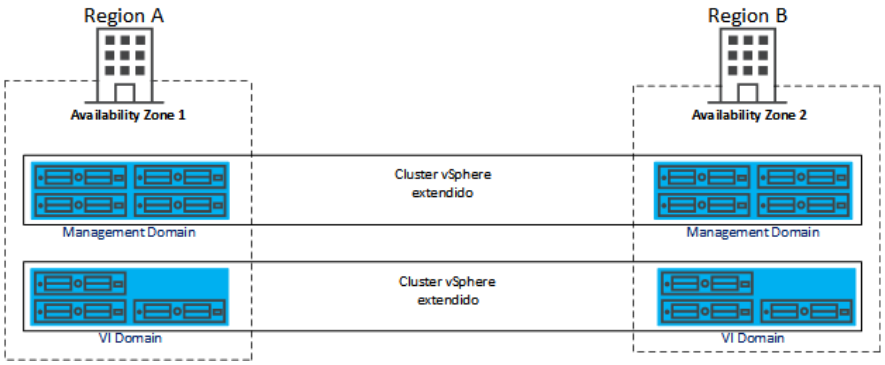

Figura 4.3: Ejemplo de un SDDC con dos Regions y una AZ en cada uno.

En la figura anterior se describe el esquema de un SDDC compuesto de dos Regions situadas en dos ubicaciones diferentes. Dentro de cada Region existe una AZ, AZ1 y AZ2. Cada una de las AZ contiene dos WD, un Management Domain desde donde se administra el SDDC, y un VI Domain donde se realizan las operaciones del SDDC. Como se mencionaba anteriormente, las instancias situadas en una AZ pueden migrarse de una ubicación a otra en caso de fallo de los recursos físicos. Para ello, el WD donde se encuentran esas instancias debe estar extendido en las dos AZ. En la imagen, el Management Domain está formado por ocho hosts situados en dos AZs distintas, estos hosts están agrupados dentro del mismo cluster de VMware vSphere, por lo tanto, los componentes cuyas instancias estén situadas en este cluster, se podrán migrar entre los 8 hosts dependiendo de la configuración que se establezca en los servicios de disponibilidad de VMware vSphere y VMware vSAN. Así, cuando los hosts de AZ1 sufren una caída, la AZ2 seguiría activa por lo que los componentes podrían seguir funcionando en esta segunda AZ. Lo mismo sucedería con el VI Domain[2].

[2]Se puede encontrar una descripción más detallada de esta estructura en el siguiente enlace https://docs.vmware.com/en/VMware-Validated-Design/6.0/introducing-vmware-validated-design/GUID-661B1CE3-1F74-4E00-80F3-0F5EA39528CD.html

4.2 Requisitos

En este apartado se describe aquello que debe cumplir 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[3] hosts ESXi con al menos un 128 GB de memoria RAM y un disco de arranque de 32 GB cada uno[4]. 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 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, este necesita al menos tres hosts con recursos de almacenamiento para funcionar[5]. Se debe aplicar la configuración All-Flash con discos SSD. Basándose en los perfiles que VMware establece para su producto vSAN Ready Node[8], cada host debe tener al menos un grupo de dos discos donde la cantidad de almacenamiento para la capa de capacidad debe ser de 4 TB y para la capa de caché de 200 GB. VMware vSAN soporta discos con adaptadores SAS, SATA o SCSI y estos pueden estar configurados en modo *pass-through* o RAID 0. En cuanto a esto último, es preferible que los discos se configuren en modo *pass-through* ya que permite que estos se puedan gestionar de forma independiente sin tener que apagar los hosts cuando es necesario retirar o añadir discos. 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.

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

[4] Según la configuración establecida para el producto vSAN ReadyNode [8]

[5] VMware vSAN requiere un mínimo de tres hosts mientras que el Management Domain requiere un mínimo de cuatro hosts.

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 uno de ellos, 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 el etiquetado VLAN configurado para las subredes del SDDC[6].

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.
- 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.

[6]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.

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)[7], 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. Este host se utiliza para desplegar los componentes de VMware Cloud Foundation para crear un pequeño SDDC embebido para probar sus funciones. Este host cuenta con una memoria RAM de 192 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.

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[8] se requiere para ejecutar VLC y acceder al entorno embebido generado por VLC. El servidor DNS contiene el nombre y su respectiva dirección que un componente de VCF utilizará para que sus instancias se puedan comunicar con otras. Este servidor DNS implementa un único dominio que se denomina *pesci.domain*. El servidor Active Directory proporciona un almacén de usuarios y grupos de usuarios a los cuales se les configura un rol dentro de cada compo-

[7] Se utiliza la versión 4.0.1 del instalador.
[8] Se refiere a ella como *Jump Host*.

nente de SDDC. Se utiliza este repositorio 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 se utiliza para asignar una dirección IP a las interfaces Tunnel EndPoint (TEP)[9] 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.

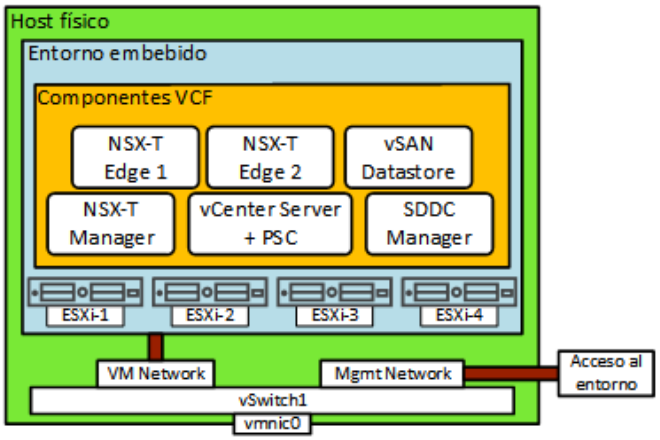

Figura 4.4: Muestra la estructura generada por el instalador VLC. Cuatro hosts ESXi embebidos con los componentes de VMware Cloud Foundation cuyo tráfico circula a través del *port group* VM Network.

[9]Más adelante se describirá la función de este elemento

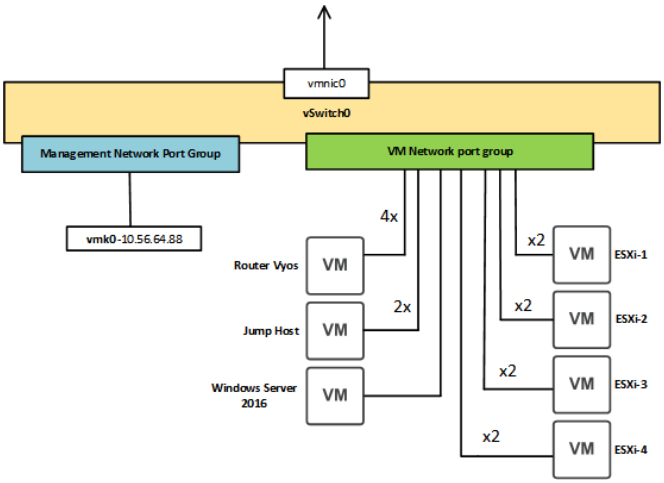

Figura 4.5: Máquinas virtuales en el host físico.

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.

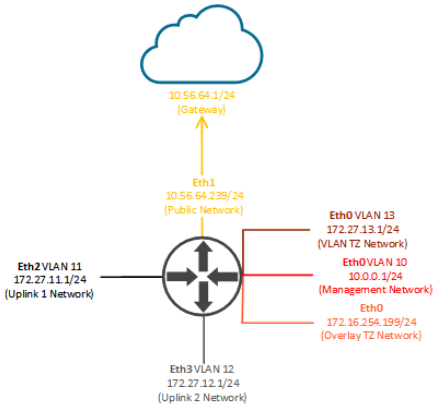

Figura 4.6: Interfaces del router Vyos.

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 9000 Bytes ya que la mayoría de componentes de VCF utilizan paquetes de red *jumbo frame*. 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.

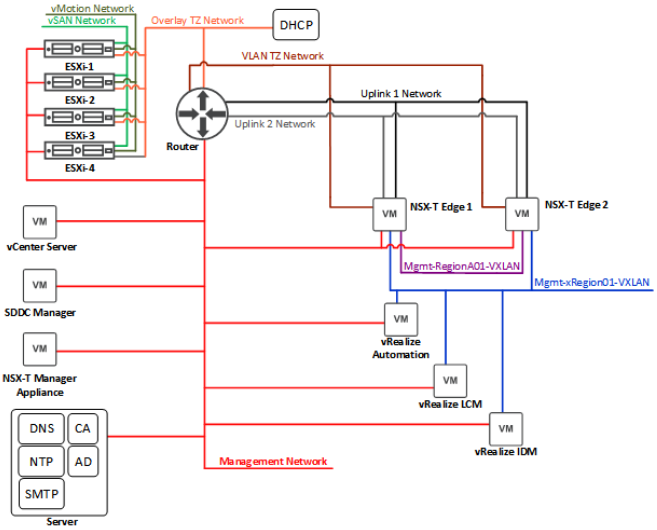

Figura 4.7: Topología de las redes del entorno desplegado.

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.

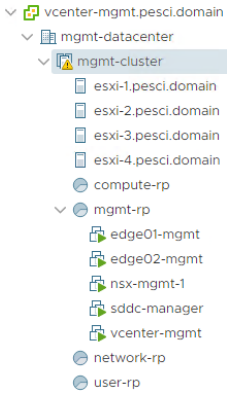

Figura 4.8: Dominio y cluster vSphere del *management domain*.

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. Debido a la limitada cantidad de recursos que ofrece el host físico donde se realiza el despliegue, 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 de 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 también se pueden establecer reglas para determinar el orden de encendido de las VMs pertenecientes a un mismo grupo.

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 group* 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*. Además, el vDS está configurado sobre los cuatro hosts por lo tanto todos tienen acceso a todos los *port groups*:

- **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. Tiene el nombre *sddc-vds01-mgmt*, a él se conectan las VMs *vcenter-mgmt*, *sddc-manager*, *nsx-mgmt-1*,*edge01-mgmt* y *edge02-mgmt*. Utiliza la subred con IP 10.0.0.0, máscara de red 255.255.255.0, VLAN 10 y 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. Tiene el nombre *sddc-vds01-vmotion* y utiliza la subred con IP 10.0.4.0, máscara de red 255.255.255.0, VLAN 10 y 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. Tiene el nombre *sddc-vds01-vsan* y utiliza la subred con IP 10.0.8.0, máscara de red 255.255.255.0, VLAN 10 y 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 ambos hay configuradas las dos VMs 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*.

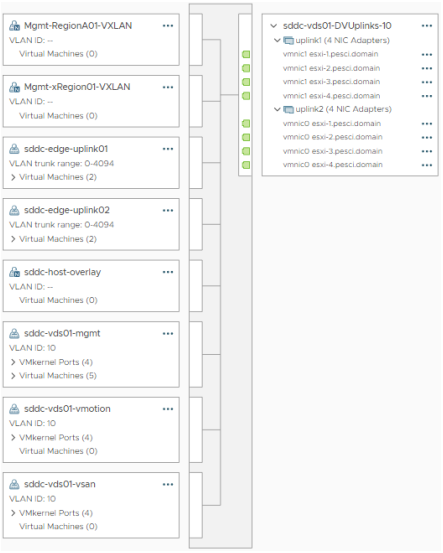

Figura 4.9: Contenido de vSphere Distributed Switch *sddc-vds01*.

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 cada interfaz *uplink* se mapea con 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 VMware 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 del SDDC y dan acceso a otras redes externas, se elimina la dependencia con el estado de VMware vCenter Server permitiendo que la comunicación continúe aunque VMware 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.

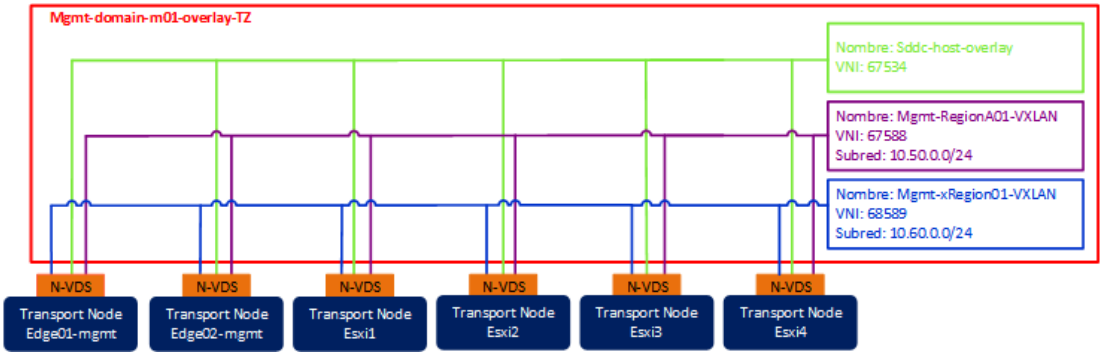

Figura 4.10: *Segments* de la *transport zone mgmt-domain-m01-overlay-tz*

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.

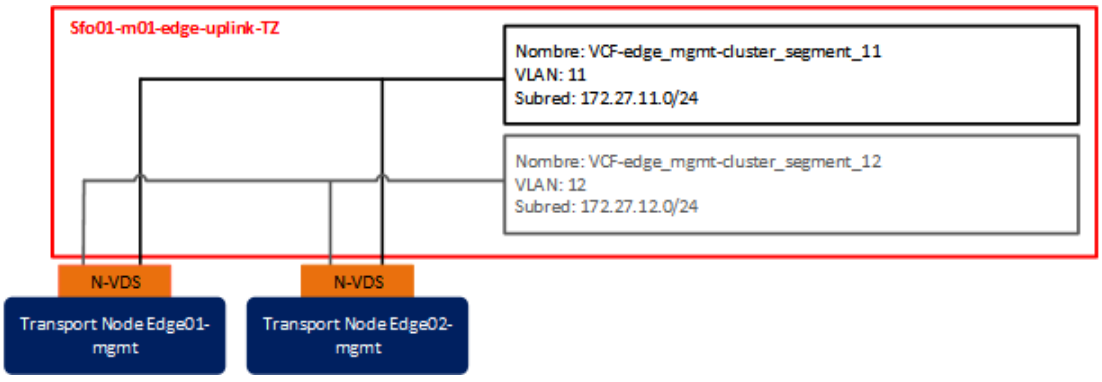

Figura 4.11: *Segments* de la *transport zone sfo01-m01-edge-uplink-tz*

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).

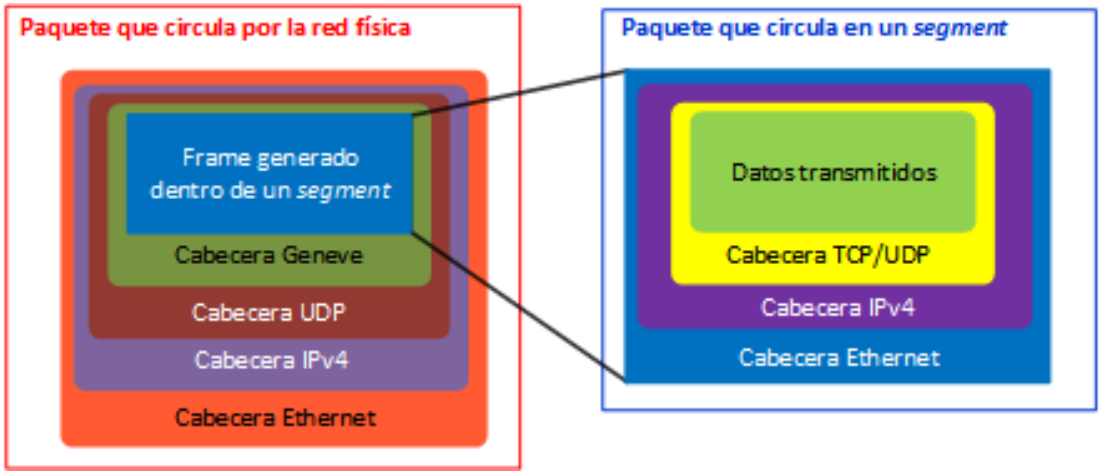

Figura 4.12: Cabeceras de un paquete de red encapsulado con Geneve.

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*[10] 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*.

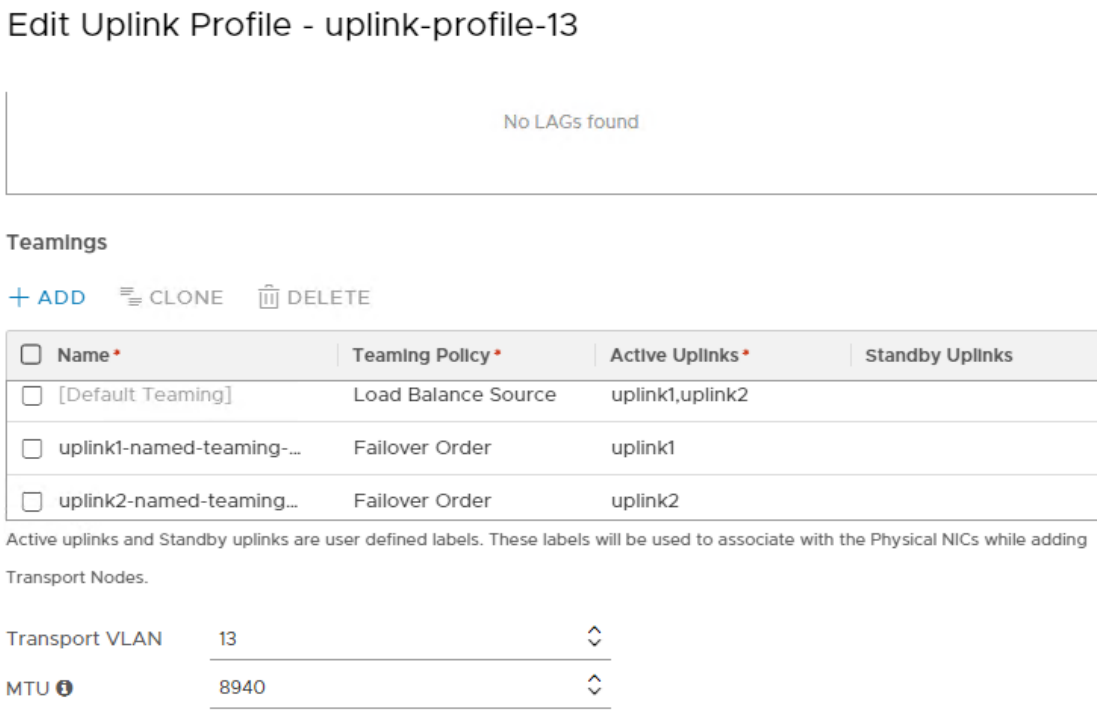

Figura 4.13: *Uplink Policy* configurada para la *transport zone sfo01-m01-dge-uplink-tz*.

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

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

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*.

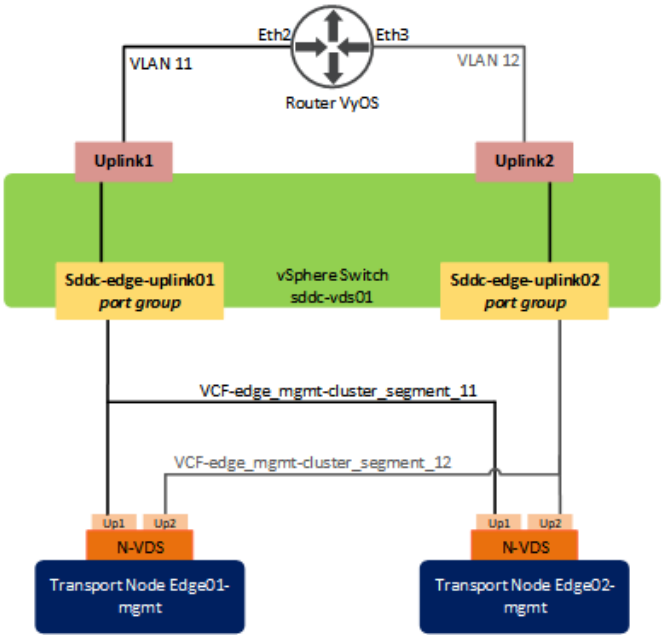

Figura 4.14: Topología de red de las insterfaces *uplink*.

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[11]. Los TNs que son hosts ESXi obtienen su dirección TEP de un servidor DHCP[12] 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* 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*.

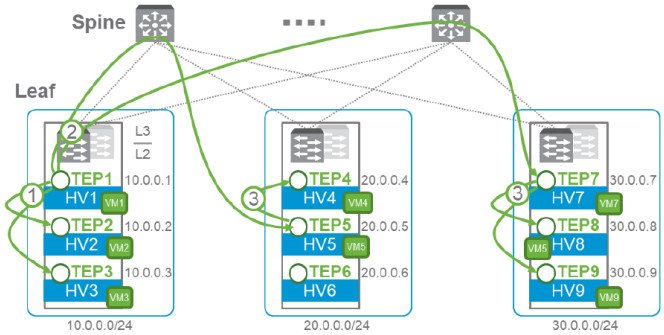

Figura 4.15: Modo de replicación *Two-Tier Hierarchical*.

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

[11] 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.

[12] 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.

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.

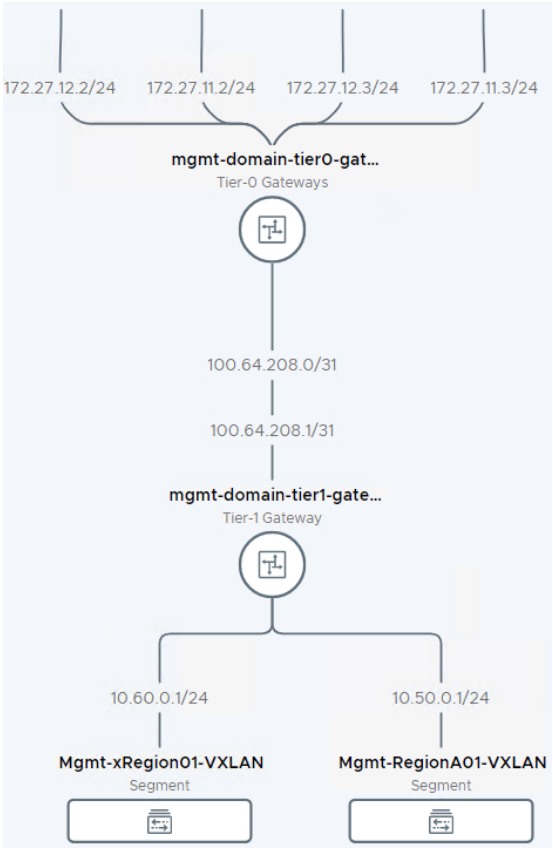

Figura 4.16: Topología de los routers virtuales de *Tier-1* y *Tier-0*.

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 [13], la interfaz *Router Link*[14] 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*[15] y que comunica los SR de las dos instancias de NSX-T Edge.

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

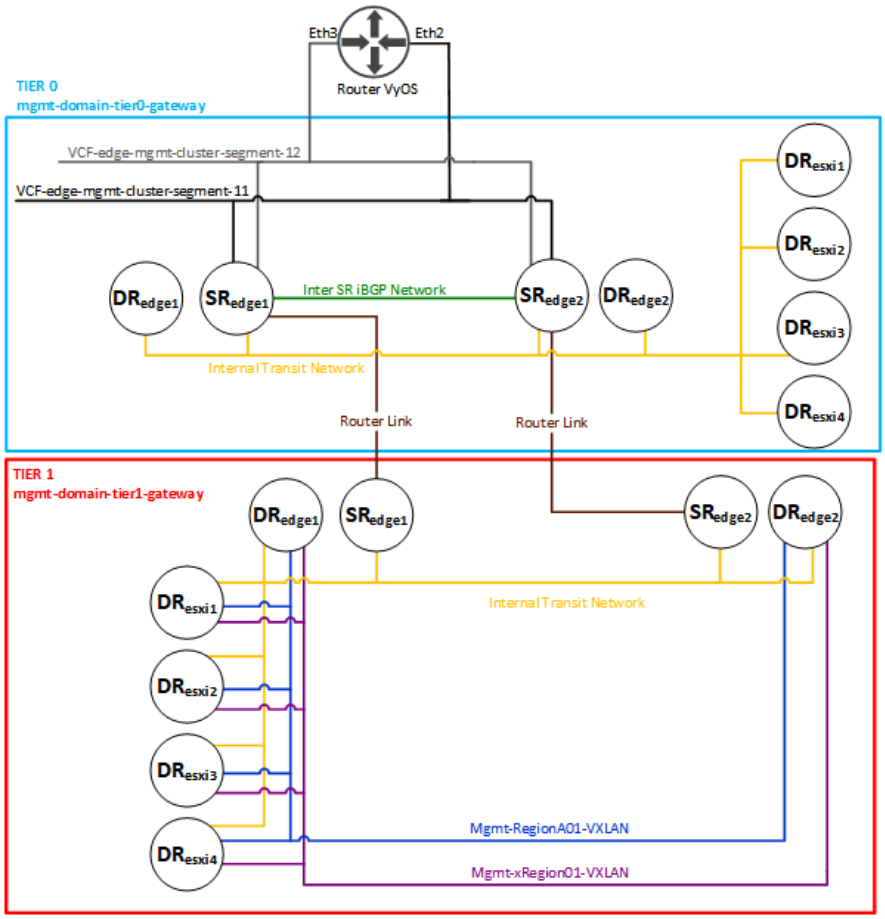

Figura 4.17: Estructura interna de los routers virtuales de *Tier-0* y de *Tier-1*.

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)[16], 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 [17] (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-*

[16] 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

[17] 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[18], 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 4.18: Muestra los usuarios definidos en el Active Directory sincronizados en Workspace One Access.

En la Figura 4.19 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.

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

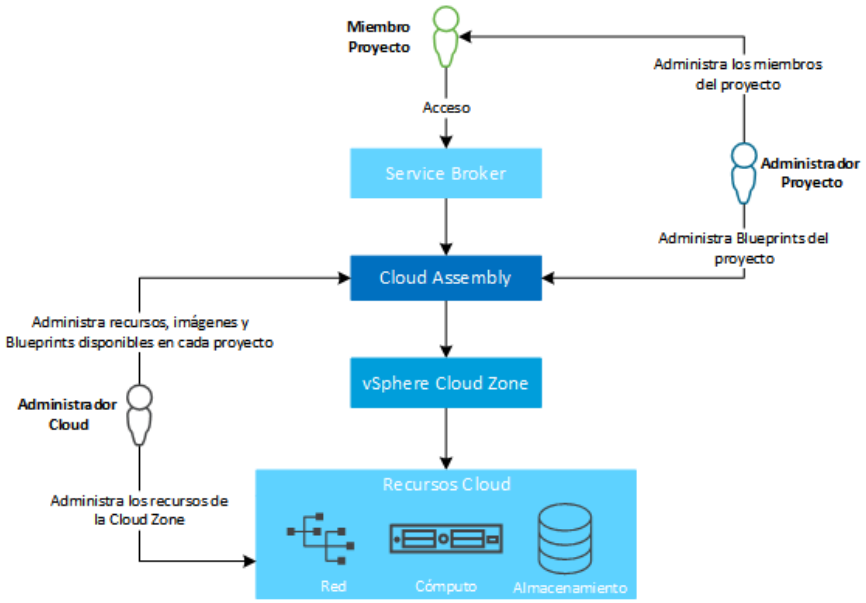

Figura 4.19: Componentes de VMware vRealize Automation y tareas que realiza cada rol de usuario.

Internamente vRA se divide en varios servicios que permiten gestionar los diferentes aspectos de la cloud. Para centrarse en los objetivos de 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 des-

plegado 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.

Capítulo 5

Aplicación de la solución

En este capítulo se detalla como sería la implementación de la solución desplegada en el entorno de pruebas sobre la infraestructura disponible en el CITIC.

5.1 Arquitectura del entorno

Cuando se despliegan los componentes de VMware Cloud Foundation se crea el primer *workload domain* que es el Management Domain. Este WD se despliega inicialmente sobre cuatro de los hosts pero como el entorno cuenta con ocho hosts, todavía quedarían otros cuatro hosts sin aprovisionar por lo tanto es necesario pensar que arquitectura se va a implementar. Existen dos posibilidades, el modelo estándar y el modelo consolidado, y para ambas el entorno cuenta con los recursos suficientes. El modelo estándar está pensado para entornos grandes con un mínimo de 7 hosts. Para el caso de la infraestructura del CITIC primero se crearía el Management Domain con cuatro hosts y después se añadiría un Virtual Infrastructure Domain con los cuatro hosts restantes, así, de esta forma, las operaciones del SDDC estarían separadas ya que el Management Domain se dedicaría a dar soporte a sus propios componentes y al resto de Workload Domains, permitiendo gestionar las actualizaciones, monitorizar y resolver conflictos, gestionar la seguridad y administrar las operaciones, mientras el VI Domain contendría los recursos que los usuarios aprovisionan desde el componente VMware vRealize Automation. Con el modelo consolidado solo existiría un único *workload domain* que contendría ocho hosts de la infraestructura. Este sería un Management Domain donde se comparten los mismos recursos para los componentes que gestionan el SDDC como en el modelo estándar y para las operaciones de aprovisionamiento de recursos, aunque la capacidad de uso de recursos de cada componente se podría limitar colocándolos en *resorce pools*. Viendo las diferencias entre ambos, el primer modelo proporciona mayor aislamiento y seguridad de los recursos ya que los dedicados a la administración están separados de los dedicados a las operaciones de los usuarios de la plataforma Cloud. Al mismo tiempo, esto limita la cantidad

de recursos disponibles y reduce la disponibilidad del servicio ya que cada WD está limitado por el número de hosts que lo forman. En cambio, con el modelo consolidado se reduce la capacidad de aislamiento de cada flujo de trabajo pero todos los hosts comparten las operaciones de administración y de los usuarios aumentando así la disponibilidad del servicio y la cantidad de recursos. El modelo recomendado por VMware en este caso es el modelo estándar porque tiene mejores medidas de seguridad y los recursos se administran de una forma más sencilla. En el momento de implementar esta solución es necesario decidir cuantos hosts y cuales de los disponibles se van a asignar a cada *workload domain* ya que no todos aportan la misma cantidad de recursos. El el objetivo es balancear los recursos entre los WD para proporcionar suficientes recursos y que sus operaciones se ejecuten correctamente. Todos los hosts y componentes de almacenamiento de la infraestructura están situados en una misma ubicación, dentro del CITIC, por lo tanto el entorno de producción estaría formado por una única *region* con solo una una AZ en su interior la cual agruparía todos los componentes.

5.2 Cumplimiento de requisitos

En esta sección se detallará si los recursos físicos ya disponibles en la infraestructura son suficientes para implementar VMware Cloud Foundation o si por el contrario sería necesario aumentarlos. La infraestructura estará formada por dos *workload domains*, un Management Domain y un VI Domain, el primero requiere al menos cuatro hosts mientras que el segundo debe contener al menos tres[1], por lo tanto hay suficientes hosts, ocho, para implementarlo. Aparte, los requisitos físicos descritos para los hosts del Management Domain se aplican también a los hosts del VI Domain. Teniendo en cuenta esto, un host con los requisitos mínimos debería contar con dos interfaces de red, un grupo de discos con dos discos duros con 4 TB para capacidad y 200 GB de caché y 128 GB de memoria RAM. Los hosts disponibles cumplen con todos los puntos pero se disponen de suficientes discos duros, hay trece discos disponibles pero se necesitan al menos dieciséis[2]. El ancho de banda de las conexiones de red existentes también cumple con el mínimo, 10 Gbit.

5.3 Diseño y configuración del VI Domain

Una vez desplegado el MD en la infraestructura cuya configuración sería similar a la descrita en el capítulo anterior, se debería desplegar un VI domain con los cuatro hosts restantes en el entorno de producción. En esta sección se describirán aquellos aspectos más relevantes el diseño y configuración de ese segundo *workload domain* una vez se decida instalar el

[1]Tres es la cantidad mínima de hosts para implementar VMware vSAN que será el tipo de almacenamiento que se utilice en el VI Domain.

[2]Dos discos cada uno de los ocho hosts.

servicio Cloud en las máquinas del CITIC.

Los componentes de este WD, como el Management Domain, también deben tener acceso al servidor DNS, al servidor DHCP y al servidor NTP para su correcto funcionamiento y para que todos puedan estar sincronizados entre si. Además, también se debe proveer acceso al directorio de usuarios de la UDC para establecer roles y habilitar usuarios que se puedan conectar a los componentes que gestionan este VI Domain, y al router o switches de la capa física para que los componentes puedan acceder a la red externa.

5.3.1 Diseño de los componentes

Si bien el diseño de los componentes que VMware Cloud Foundation despliega para controlar un VI Domain es muy similar al que se ha descrito en el capítulo anterior es necesario resumirlo y destacar aquellas diferencias.

La instancia del componente VMware vCenter Server que controla el VI Domain estaría alojada dentro del Management Domain y sería el encargado de controlar los cuatro hosts pertenecientes al WD. Además, sería necesario activar la opción *Enhanced Link Mode*, al igual que en la instancia de VMware vCenter que controla el Management Domain, para que ambas instancias compartan sus respectivos PSCs y así formen un único dominio de autenticación SSO con el que poder gestionar ambas desde una misma interfaz de vSphere Web Client.

El almacenamiento para un VI Domain puede ser de distintos tipos, SAN, NAS o VMware vSAN, incluso se pueden combinar, pero para para la realización de este proyecto solo se tiene en cuenta la configuración de almacenamiento con VMware vSAN ya que reduce la complejidad de administración de la infraestructura. Este VI Domain debería tener su propio VMware vSAN *datastore* diferente del utilizado por el Management Domain. El *datastore* debería tener el tamaño suficiente para soportar las operaciones de aprovisionamiento y despliegue de recursos que realicen los usuarios. Además, la configuración de FTT debería ser igual a 1 para soportar la caída de alguno de los hosts. Finalmente, también requiere de una subred dedicada para proporcionar el servicio de almacenamiento mediante el protocolo IP como se describe para el Management Domain.

Dentro del VI Domain se crearía un cluster de VMware vSphere donde se incluyen los cuatro hosts que forman el WD. Su configuración de vSphere High Availability y de vSphere DRS sería la misma para el Management Domain, es decir, para el primer servicio se configura la reserva del 25% de CPU y el 30% de memoria RAM del WD y se activa la propiedad *VM and Application Monitoring*, y para el segundo servicio se configura la opción *Fully Automated*. La red de este cluster estaría formada por un único vSphere Distributed Switch que contendría un *Management Port Group*, un *vMotion Port Group*, un *vSAN Port Group*, dos *Edge Uplink Port Groups* (todos ellos de tipo *Distributed Port Group*) y dos *Uplink Port Groups* que dan salida al tráfico hacia la red física. Las funciones y configuración de las propiedades para cada *port*

group son las mismas que las descritas en el Management Domain para los mismos *port groups* a excepción de la propiedad *Port Binding* que se establecería como *Static* para todos porque la opción *Emphemeral* ya no es necesaria al estar vCenter Server en otro WD, eliminándose así la dependencia con su estado. Además, las subredes que se deberían configurar tendrían que ser diferentes ya que sus servicios serían dedicados a este WD, excepto para el *Management Port Group* que se debería conectar a la misma subred que el *Management Port Group* del Management Domain y así poder comunicarse y controlar el VI Domain. La configuración de los servicios *Network I/O* y *Health Check* de este vDS es la misma que en el Management Domain. El tamaño del MTU también debe ser de 9000 Bytes.

En cuanto al componente VMware NSX-T, se desplegarían tres instancias de VMware NSX-T Manager Applicance en el Management Domain y dos instancias de Vmware NSX-T Edge dentro del propio VI Domain. Su configuración sería la misma que la descrita para el Management Domain, una TZ de tipo VLAN con su *Uplink Policy* correspondiente y con dos *segments* conectados a los dos *Edge Uplink Port Groups* del vDS que serán los que utilicen las instancias de VMware NSX-T Edge para dirigir el tráfico hacia el dispositivo de red físico, y otra TZ de tipo Overlay pero esta vez con un *segment* para comunicar los componentes de VMware NSX-T y adicionalmente tantos *segments* se quieran crear para que sean usados por los usuarios del servicio Cloud. Estos *segments* formarían parte de una topología con dos routers virtuales, uno de *Tier-1* y otro de *Tier-0* donde los *segments* de tipo VLAN se conectarían a *Tier-0* y los creados para entregar el servicio Cloud, al *Tier-1* donde además se podrían proporcionar otros servicios como DNS o DHCP. En la red física, este WD también requiere la configuración de BGP y ECMP para el correcto funcionamiento de los componentes de VMware NSX-T y el aprovechamiento de todas las funcionalidades que ofrece una red definida por software.

Apéndices

Notas

En este documento se utilizan términos en inglés ya que forman parte del campo que se está tratando o por ser su nombre original y por lo tanto están reconocidos.

Cuando en el documento se menciona

capa 2 o

capa 3 se está haciendo referencia a las capas establecidas por el Modelo OSI, la capa de enlace de datos y la capa de red respectivamente.

Lista de acrónimos

API *Application Programming Interface*

AS *Autonomous System*

AZ *Availability Zone*

BGP *Border Gateway Protocol*

BUM *Broadcast, Unknown Unicast, Multiacast*

CA *Certificate Authority*

CITIC *Centro de Invesitgación en Tecnoloxías da Información e as Comunicacións*

CPD *Centro de Procesamiento de Datos*

DHCP *Dynamic Host Configuration ProtocoL*

DNS *Domain Name Server*

DPM *vSphere Distributed Power Management*

DR *Distributed Router*

DRS *vSphere Distributed Resources Scheduler*

FTT *Failures To Tolerate*

HA *vSphere High Availability*

HDD *Hard Disk Drive*

IP *Internet Protocol*

iSCSI *Internet Small Computer System Interface*

LUN *Logical Unit Number*

MD *Management Domain*

MTU *Maximum Transmission Unit*

NAT *Network Address Translation*

NFS *Network File System*

NIST *National Institute of Standards and Technology*

NIC *Network Interface Card*

NTP *Network Time Protocol*

N-VDS *NSX-T Virtual Distributed Switch*

PSC *Platform Services Controller*

QoS *Quality of Service*

RAID *Redundant Array of Independent Disks*

SAN *Storage Area Network*

SDDC *Software Defined Data Center*

SFP *Small Form-factor Pluggable Transceptor*

SMTP *Simple Mail Transfer Protocol*

SSD *Solid-State Drive*

SR *Service Router*

TB *TeraByte*

TEP *Tunnel End Point*

ToR *Switch Top of Rack*

TN *Transport Node*

TZ *Transport Zone*

UDC *Universidade da Coruña*

UDP *User Datagram Protocol*

VCF *VMware Cloud Foundation*

VLAN *Virtual Local Area Network*

VLC *VMware Lab Constructor*

vDS *vSphere Distribute Switch*

VI *Virtual Infrastructure Domain*

VMFS *Virtual Machine File System*

VM *Virtual Machine*

VNI *Virtual Network Identifier*

vRA *VMware vRealize Automation*

vRSLCM *VMware vRealize Lifecycle Manager*

WD *Workload Domain*

WSA *Workspace One Access*

Glosario

Appliance : archivo que contiene una máquina virtual con un sistema operativo con el propósito de entregar una única aplicación preconfigurada.

BGP [9]: protocolo de enrutamiento que se utiliza para el intercambio de rutas entre Autonomous Systems (AS) de forma dinámica y así evitar configurarlas manualmente.

BUM : se refiere al tráfico de red Broadcast, Unknown unicast y Multicast. El primero es tráfico que se transmite a todos los dispositivos disponibles en la red, Unknown Unicast es tráfico enviado a un único destinatario para el que no se conoce su dirección MAC dentro de una misma VLAN y Multicast es tráfico que se envía a los dispositivos que pertenecen a un grupo dentro de una red.

Cluster [10]: agrupación de recursos de múltiples hosts que se gestionan como una única colección.

Controlador SFP+ : interfaz modular que permite conectar cables de fibra óptica a un dispositivo.

CPD : luegar donde se sitúan un conjunto recursos con gran capacidad de cómputo necesarios para procesar información, normalmente en grandes cantidades.

Datastore : dentro de VMware vSphere, un datastore es un contenedor lógico que abstrae los componentes físicos de almacenamiento y provee un modelo uniforme para almacenar máquinas virtuales, plantillas o imágenes ISO.

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

Host : servidor físico en el que se ejecuta el hipervisor.

IaaS [2]: servicio Cloud en el que se provee capacidad de aprovisionamiento de recursos de cómputo, almacenamiento y red, sobre los cuales se puede desplegar software.

iSCSI : estándar que implementa el protocolo de transporte SCSI para transmitir datos entre dispositivos.

Jumbo Frame [11]: paquetes de red cuyo MTU es mayor que el valor definido en el estándar Ethernet, 1500.

LUN : identifica una colección de dispositivos de almacenamiento que se presentan como un único volumen.

Máquina virtual : máquina que se ejecuta en un entorno virtualizado con hardware virtual detro de un hipervisor.

NIC : componente físico que conecta un dispositivo a una red y permite compartir sus recursos.

Pool de almacenamiento : agrupación de volúmenes de almacenamiento que se administran de forma conjunta.

Port Group : puertos que se añaden en el componente vSphere Distributed Switch y que agrupan las conexiones de múltiples máquinas virtuales sobre las cuales se pueden establecer una configuración determinada.

QoS : medida de rendimiento que se asigna a un servicio en la red. Los componentes de VMware utilizan el campo Differentiated Services Code Point (DSCP) en la cabecera de capa 3, y el campo Class of Service (CoS) en la cabecera de capa 2 para indicar la prioridad del tráfico.

Rack : armario metálico destinado a alojar servidores físicos.

RAID 5 : 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 de paridad en todos los discos para poder recuperar datos corruptos.

Red Overlay [12]: abstracción de una red sobre una red física implementada por un conjunto de nodos situados en diferentes localizaciones y conectados entre si.

SAN : red dedicada a proveer acceso a los dispositivos de almacenamiento.

SDDC [13]: Software-Defined Datacenter es un modelo de arquitectura de infraestructura para virtualizar los recursos de cómputo, almacenamiento y red.

UDP : protocolo de red de la capa de transporte que permite enviar paquetes sin establecer previamente una conexión.

VLAN : método para aislar múltiples dominios de broadcast sobre una misma red física.

VLAN trunk : enlace que permite la circulación del tráfico de diferentes redes VLAN.

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

[8] ——, "vsan all flash hardware guidance (af-4 series)," 2020. [En línea]. Disponible en: https://www.vmware.com/resources/compatibility/vsan_profile.html?locale=en

[9] P. Traina, "Bgp-4 protocol analysis," RFC 1774, DDN Network Information Center, Tech. Rep., 1995.

[10] V. Infrastructure, "Resource management with vmware drs," *VMware Whitepaper*, vol. 13, 2006.

[11] E. Alliance and B. Kohl, "Ethernet jumbo frames," 2009.

[12] D. Gifford, K. L. Johnson, M. F. Kaashoek, and J. W. O'Toole Jr, "Overcast: Reliable multicasting with an overlay network," in *Proceedings of USENIX Symposium on OSDI*, 2000.

[13] V. Törhönen, "Designing a software-defined datacenter," 2013. [En línea]. Disponible en: http://urn.fi/URN:NBN:fi:tty-201405261235