Facultade de Informática
UNIVERSIDADE DA CORUÑA

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

PESCI: Plataforma de Entrega de Servicios Cloud para Investigación

Estudante: Amaro Castro Faci
Dirección: Outro Nome Completo

A Coruña, agosto de 2020.

Resumen

El Cloud Computing es un modelo que permite acceder a un conjunto de recursos, como por ejemplo, redes, almacenamiento, aplicaciones, servicios y potencia de cálculo, que pueden ser aprovisionados bajo demanda de una forma rápida sencilla, reduciendo el coste del servicio para el usuario y el esfuerzo en cuanto a gestión 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, que se traducen en máquinas virtuales, del tamaño que necesiten para realizar tareas que no serían posibles en dispositivos convencionales.
Actualmente el servicio está activo pero de forma limitada y no abierta a todos los usuarios del CITIC debido que no existe una plataforma que habilite el acceso a cada usuario a sus recursos, que gestione la autenticación de los usuarios, y que automatice y simplifique los procesos de aprovisionamiento de recursos. Hasta ahora, las tareas de aprovisionamiento y gestión de usuarios se realizan bajo petición previa al administrador del sistema y se ejecutan de forma manual lo cual consume más tiempo, recursos, y aumenta los riesgos sobre el servicio.
El objetivo principal de este proyecto es desplegar un servicio Cloud en el CITIC usando como base las herramientas que ya se encuentran sobre la infraestructura, donde cada usuario pueda tener su propio espacio donde poder gestionar sus recursos según sus necesidades y de forma simple y dinámica, que simplifique la gestión de usuarios integrando el sistema de autenticación de la UDC en el servicio, y establecer un sistema de valoración de los recursos para poder controlar y limitar la cantidad de recursos que un usuario puede aprovisionar para mejorar el uso y funcionamiento del servicio. En definitiva, el objetivo es desplegar un entorno Cloud que haga la infraestructura física más eficiente y que entregue todo su potencial.

Palabras clave:

- Cloud Computing
- CITIC
- Máquina virtual
- Aprovisionamiento de recursos
- Bajo demanda
- Control de recursos

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

Según *National Institute of Standards and Technology* (NIST), el Cloud Computing es un «modelo de recursos configurables y compartidos, accesibles através de la red bajo demanda y desde cualquier lugar en cualquier momento»[3]. Las principales características de este modelo son:

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

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

para ofrecerlo de forma abierta a todos los usuarios. Este permite aprovisionar recursos de un servidor en forma de máquinas virtuales con unas especificaciones determinas por el usuario para realizar tareas que precisan gran capacidad de cómputo, de almacenamiento, o de red.

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

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

1.1 Motivación

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

1.2 Objetivos

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

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

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 de este 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) [4]. Está formado por 5 nodos *Lenovo NeXtScale nx360 M5* y 3 nodos *Dell EMC PowerEdge R740*. Ambos componentes dan flexibilidad en cuanto a la escalabilidad y ofrecen gran rendimiento de cómputo.

Especificaciones principales de los nodos:

- Lenovo NeXtScale nx360 M5:
  - CPU: Dos Intel Xeon E5-2650
  - Memoria: 128 GB
  - Tarjeta gráfica: Tesla M60

  Más información: https://lenovopress.com/tips1195-nextscale-nx360-m5-e5-2600-v3
- Dell EMC PowerEdge R740:
  - CPU: Dos Xeon Gold 6146
  - Memoria: 384 GB

- Tarjeta gráfica: Tesla P40

Más información: https://www.dell.com/es-es/work/shop/servidores-almacenamiento-y-redes/smart-value-poweredge-r740-server-standard/spd/poweredge-r740/per7400m

El almacenamiento está colocado físicamente en la misma ubicación que los hosts pero en su abstracción lógica este es independiente y está separado de cada nodo. Está conformado por 13 discos duros SSD de 3.84 TB de capacidad, obteniendo así una capacidad total de casi 50 TB, pero que utilizan la configuración de almacenamiento RAID 5[Pal. 4] lo cual permite conseguir mayor integridad de los datos, tolerancia a fallos y ancho de banda, reduciendo la cantidad de almacenamiento utilizable a 34 TB. Estos discos forman un *pool* de almacenamiento que se divide en cinco LUNs (*Logical Storage Unit*)[Pal. 4] de 2 TB cada una, representadas en el sistema de virtualización como cinco *datastores* diferentes que utilizan el sistema de archivos VMFS el cual optimiza el almacenamiento de máquinas virtuales.
Los discos duros físicos están colocados en una misma cabina y son accesibles por todos los nodos a través de dos switches para aportar redundancia. Para ello, las cabinas incorporan dos controladores con conexión SFP+[Pal. 4] que se conectan a cada switch mediante dos puertos que aportan conectividad 10 Gb y, además, incorporan otros dos puertos con conectividad 1 Gb para la gestión de los discos. Estas conexiones utilizan los protocolos de red Ethernet y iSCSI, formando así, junto con el resto de componentes descritos, la estructura de una SAN [2.2].

La gestión del almacenamiento se realiza en la capa física, el nivel más bajo por lo que la configuración de cada LUN que utilizan las máquinas virtuales desplegadas se tiene que hacer antes de conocer los requisitos necesarios de lo que se vaya a desplegar en la capa software. Esto provoca que si se quiere desplegar una máquina virtual con más capacidad de almacenamiento o con una estructura RAID diferente haya que crear una nueva LUN que se adapte a los requisitos. Esta gestión hace que el uso de recursos de almacenamiento no sea el óptimo ya que no permite ajustar de forma precisa y rápida cada configuración a los requisitos necesarios generando así mayor coste.

2.2 Software

Actualmente el servicio está basado en el software de la empresa VMware, uno de los principales proveedores de software de virtualización, siendo **VMware vSphere** el software desplegado sobre la infraestructura. Este producto de VMware es el encargado de virtualizar parte de la infraestructura física y de proporcionar las herramientas necesarias para gestionarla. Sus principales componentes son los siguientes:

- **ESXi**: Hipervisor propio de VMware, de tipo 1 o *bare metal*[Pal. 4]. No requiere de sistema operativo para funcionar ya que funciona directamente sobre el hardware físico[5]. Este hipervisor está instalado en cada uno de los ocho nodos que forman la infraestructura.

- **VMware vCenter Server**: servicio que actúa como un administrador central para todas las máquinas virtuales y hosts. Normalmente, los servicios descritos en este apartado están disponibles para una instancia de vCenter Server agrupados en un *Platform Services Controller* (PSC) [Pal. 4]

- **vCenter Single Sign-On**: es un servicio de autenticación. Permite que los usuarios solo se tengan que autenticar una vez cuando acceden al entorno de la infraestructura a través de vSphere Client, en lugar de tener que autenticarse varias veces en cada componente. Cuando el usuario se autentica por primera vez, este recibe un token que le permitirá autenticarse en el resto de componentes sin volver a introducir sus credenciales. También se encarga de la administración perfiles de usuarios y de los dominios de autenticación (esto permite usar directorios de usuarios externos).

- **vSphere Web Client** y **vSphere Client**: interfaz que permite conectarse a una instancia de vCenter Server para gestionar la infraestructura.

- **vSphere Auto Deploy**: herramienta que permite desplegar gran cantidad de nodos físicos de forma automátizada.

- **vSphere Update Manager**: permite gestionar de forma centralizada y automatizada las actualizaciones para hosts ESXi, para el hardware de las máquinas virtuales y para actualizar e instalar software de terceros en los hosts ESXi. Este componente se ejecuta desde VMware vCenter Server como un servicio y requiere conexión a la red externa para obtener las actualizaciones.

- **vSphere Web Client**: interfaz web que permite acceder a vCenter Server de forma remota.

- **vMotion**: permite la migración de máquinas virtuales de un host a otro de forma transparente y sin detener su ejecución. Permite planificar las migraciones de máquinas virtuales entre hosts que pueden pertenecer a distintos clusters.

- **Storage vMotion**: permite migrar los discos y configuración de una máquina virtual de un *datastore* a otro sin interrumpir el servicio.

- **vSphere High Availability (HA)**: provee alta disponibilidad para las máquinas virtuales. En caso de que una máquina virtual deje de estar activa, este servicio intenta

encender la máquina virtual en otro host dentro del mismo cluster automáticamente. Este solo actúa en caso de fallo de un host, mientras que vMotion solo actúa cuando las migraciones se hacen entre hosts activos. Proporciona escalabilidad gracias a su modelo Maestro - Esclavo, fiabilidad gracias a que no tiene dependencias con otros servicios y a que se puede comunicar con las máquinas a través de varios caminos, y usabilidad gracias a que tiene una interfaz sencilla.

- **vSphere Distributed Resource Scheduler (DRS)** y **vSphere Distributed Power Management (DPM)**: DRS genera recomendaciones sobre donde se debería desplegar una máquina virtual cuando se está creando, utiliza vMotion para mover las máquinas virtuales a través de los hosts de un cluster para maximizar el rendimiento o durante tareas de mantenimiento en un host. DPM se encarga de gestionar el consumo de energía de cada host según el rendimiento necesario.

- **Storage DRS**: balancea la carga de almacenamiento y las operaciones I/O entre los diferentes datastores disponibles.

- **vSphere Fault Tolerance**: crea una copia de todos los archivos y discos de cada máquina virtual sincronizados con los originales. Esto junto con vSphere HA y vSphere DRS, proporciona recuperación ante fallos y disponibilidad continua de las máquinas virtuales de forma automática en caso de que la máquina virtual esté inactiva, sin perdida de datos y sin pérdida de conexiones de que había activas. Este servicio está orientado a proteger aquellas tareas que requieren un alto rendimiento o que son críticas.

- **vSphere Distributed Switch (vDS)**: habilita switches virtuales que se encargan de gestionar el tráfico de los hosts ESXi. Se utiliza para administrar la configuración de los puertos de cada host ESXi de forma centralizada, permitiendo crear redes y establecer políticas sobre el tráfico desde un único lugar. Desde aquí se establecen las conexiones de cada tipo de tráfico con las tarjetas de red físicas de cada host ESXi.

- **Virtual Machine File System (VMFS)**: sistema de archivos de alto rendimiento nativo de VMware vSphere. Se utiliza para implementar los almacenes de datos de la infraestructura y está optimizado para el almacenamiento de máquinas virtuales.

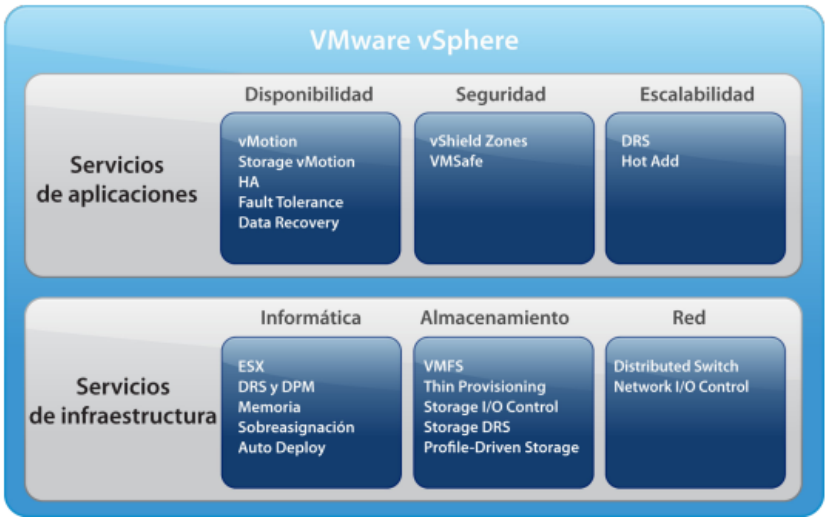


Figura 2.1: Elementos de la plataforma VMWare vSphere[2]

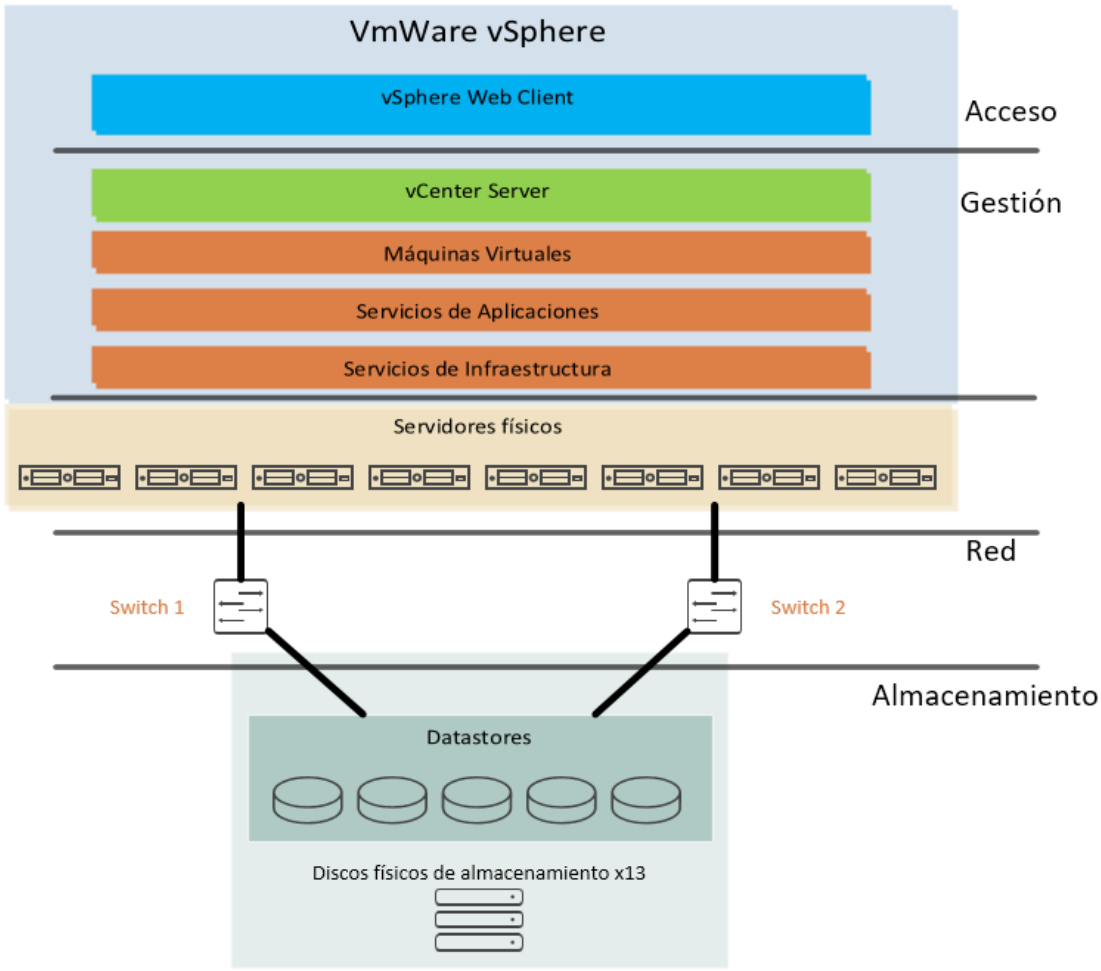


Figura 2.2: Esquema de los recuros software y hardware del entorno

2.2.1 Estado de la tecnología

En los últimos tiempos los servicios de *Infrastructure as a Service* (IaaS) se han extendido de forma considerable con la aparición de software que permite la gestión de un sistema de Cloud Computing, como pueden ser VMware Cloud Foundation (2011), OpenStack (2010), o Apache CloudStack (2012). Estas herramientas construyen una infraestructura virtual sobre un entorno físico estandarizado que les permite administrar y automatizar la escalabilidad, el sistema de almacenamiento, la disponibilidad del servicio, la red, y la seguridad del servicio, con lo que se consigue reducir el coste y el tiempo de gestión y configuración, mejorando la eficiencia de infraestructura física.

A la hora de alcanzar los objetivos descritos en este proyecto se nos plantea la duda de que solución software desplegar ya que actualmente existen tres principales alternativas en el mercado, VMware Cloud Foundation, OpenStack y Apache CloudStack. Cada una de ellas ofrece diferentes características con diferentes requisitos que se pueden adaptar mejor o peor al entorno de despliegue, pero después de comprobar esos aspectos tenemos claro que la solución elegida es VMware Cloud Foundation.

VMware Cloud Foundation

Esta solución virtualiza todas las capas de la infraestructura[Fig. 2.4] (red, computación y almacenamiento) combinando cuatro componentes principales, vSphere para gestionar el cómputo, vSAN para la gestión del almacenamiento, NSX para la gestión de la red, y vRealize para gestionar todas las operaciones que tienen lugar en el servicio, integrando todos los componentes para que la gestión de la infraestructura sea lo más simple posible. Este cojunto de herramientas convierten el CPD en un *Software Defined Datacenter* (SDDC), un entorno donde todas las partes físicas de la infraestructura pasan a estar controladas a través de software haciendo más flexible, independiente y menos costosa la configuración de estos componentes. Las principales características de VMware Cloud Foundation son:

- **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.

- **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.

- **Aprovisionamiento automatizado**: todos los componentes necesarios para formar un SDDC son desplegados automáticamente por VMware Cloud Foundation, incluyendo los recursos informáticos, los componentes de almacenamiento, los de red y los de administración.
- **Ciclo de vida automatizado**: automatiza las operaciones previas, iniciales y posterios de una plataforma de software para ofrecer una gestión más sencilla. Esto incluye su implementación, el aprovisionamiento de clústeres aislados bajo demanda y la instalación de actualizaciones y parches.

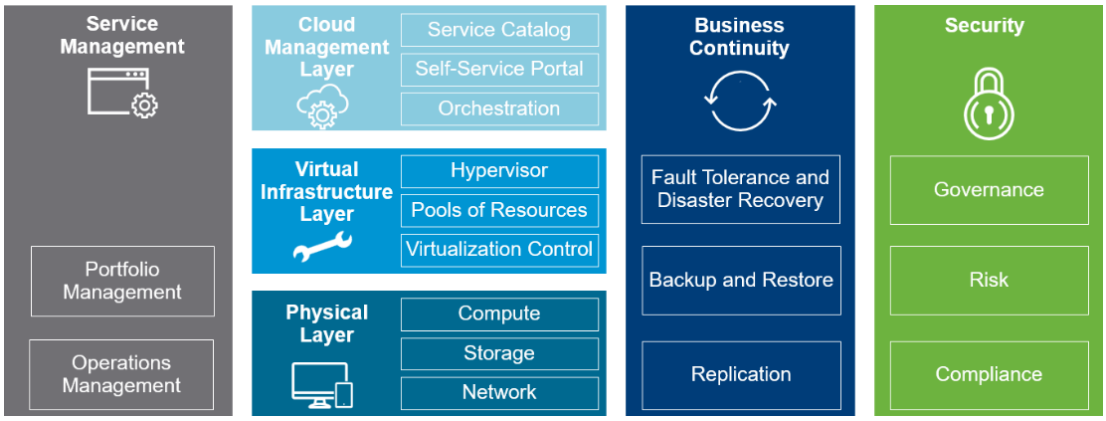


Figura 2.3: Partes virtualizadas en un SDDC.

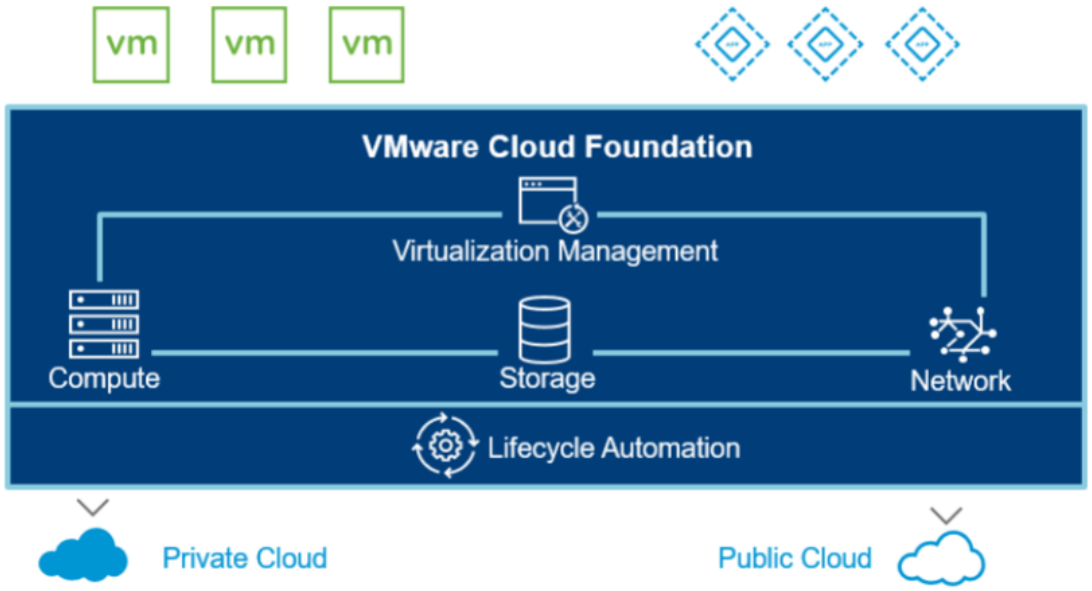


Figura 2.4: Cloud Foundation virtualiza toda la infraestructura.

VMware Cloud Foundation permite reducir el tiempo de mantenimiento ya que todo está controlado por el software que integra todos los componentes, automatizando gran parte de las operaciones y el ciclo de vida de todos los elementos desde su creación, como puede ser el control de versiones de cada elemento, los perfiles de usuario y las máquinas virtuales creadas, además de proporcionar una plataforma de acceso para que cada usuario pueda gestionar sus recursos. Además, su arquitectura se divide en entornos aislados administrados bajo demanda desde un entorno principal. Cada uno de esos entornos tiene un conjunto de recursos dedicados cuyas capacidades se pueden ajustar para proporcionar determinadas características, como por ejemplo alta disponibilidad, mayor capacidad de almacenamiento o mayor capacidad de cómputo.

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. A pesar de tener un coste elevado, teniendo en cuenta los beneficios que aporta en cuanto integración nativa con los componentes ya instalados y a que su mantenimiento es más sencillo, se ha elegido este paquete ya que es más rentable que otras opciones.
Si bien VMware ofrece plugins para conectar sus componentes con otras soluciones, como es el caso de OpenStack [6], estos no ofrecen el rendimiento que da la integración nativa, además, en caso de recibir actualizaciones, habría que actualizar cada componente de forma individual aumentando el riesgo de incompatibilidades con el resto de elementos del sistema, mientras que Cloud Foundation gestiona todo el ciclo de vida de cada actualización para cada componente, permitiendo comprobar si existe alguna incompatibilidad con el resto de versio-

nes antes de aplicar una actualización. En definitiva, VmWare Cloud Foundation simplifica el proceso de instalación, configuración, gestión, y mantenimiento, tanto para los usuarios como para el administrador del sistema.

Componentes de VMware Cloud Foundation [7]

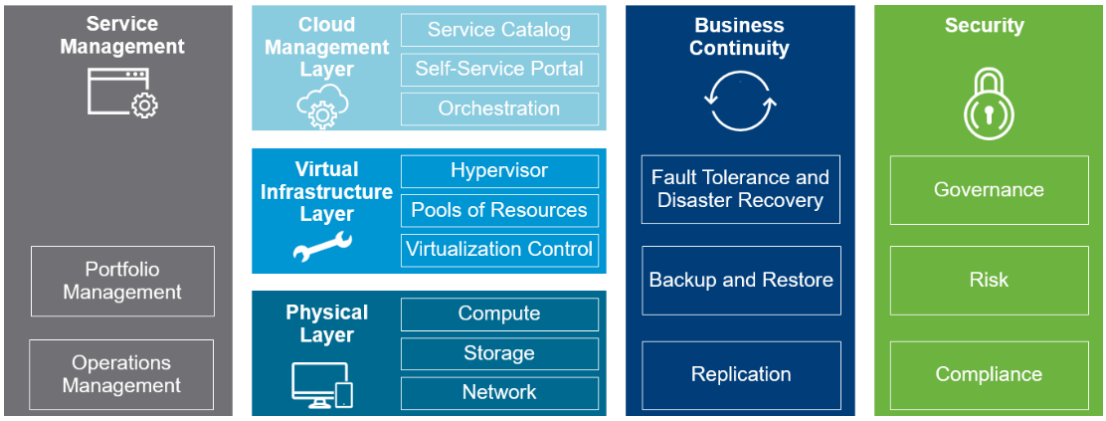


Figura 2.5: Partes virtualizadas en un SDDC.

El SDDC de VMware Cloud Foundation está dividido en varias capas y módulos con las cuales se gestiona todas las partes de la infraestructura [Fig. 2.5]. Cada una de estas capas incluye una serie de productos y servicios, con unas funciones concretas que permiten gestionar un entorno complejo de una forma más sencilla. Las partes de esta infraestructura son las siguientes:

- **Capa de Infraestructura física**: Es la base de la infraestructura y en ella residen los componentes físicos de almacenamiento, red y cómputo.
- **Capa de Infraestructura Virtual**: Se encuentra sobre la capa física y se encarga de realizar las tareas de acceso a los recursos físicos para controlar su aprovisionamiento.
- **Capa de Gestión Cloud**: Es la capa superior y se dedica a las tareas de consumo de los recursos, es decir, realiza peticiones a las capas inferiores para la obtención de los recursos.
- **Capa de Gestión de Servicios**: Esta capa se centra en el mantenimiento de la arquitectura gestionando el ciclo de vida, la monitorización, logs y alertas de los componentes.
- **Capa de Gestión de Operaciones**: La función de esta capa es la monitorización de la capa de infraestructura física, de la capa de infraestructura virtual y de los flujos de trabajo de los usuarios.

- **Capa de Continuidad del Servicio**: Contiene elementos que ayudan a que el servicio se mantenga activo y a evitar la pérdida de información crítica, proveyendo copias de seguridad, restauración y recuperación ante desastres.
- **Capa de Seguridad**: Se encarga de que todos los componentes de la infraestructura estén bien delimitados para establecer un nivel de seguridad.

En VMware Cloud Foundation existen diversos productos y servicios, algunos son requeridos para construir la instancia mínima de un SDDC y otros ofrecen servicios adicionales que incrementan el rendimiento de la infraestructura. A continuación se describen aquellos componentes necesarios para la instalación mínima de VMware Cloud Foundation y que se usarán en la implementación de este proyecto:

- **SDDC Manager**: gestiona el ciclo de vida de todos los componentes del sistema, incluyendo el proceso inicial de despliegue de Cloud Foundation, su configuración y aprovisionamiento, y las actualizaciones. Monitoriza los recursos físicos y lógicos de la infraestructura, facilita su configuración y permite añadir nuevos recursos cuando sea necesario.
- **vSphere**: ya está incluído en el servicio actual [2.2].
- **VMware vSAN**: componente clave que virtualiza el almacenamiento. Como ya se ha explicado, el almacenamiento del servicio actual está configurado con LUNs que se deben gestionar individualmente en una capa distinta a los componentes software, provocando que su configuración sea más compleja y costosa. El objetivo de VMware vSAN es gestionar de forma automatizada y desde un único lugar el sistema de almacenamiento de la infraestructura, tratando toda la capacidad y recursos de almacenamiento como un único elemento, eliminando así la necesidad de tener que crear LUNs aisladas, consiguiendo abstraer la configuración de almacenamiento de la capa física en la capa de software y permitiendo establecer políticas de almacenamiento desde cada máquina virtual para adecuarlo a las necesidades de cada una, sin tener que editar la configuración real del entorno físico. Así el rendimiento de los recursos de almacenamiento es más eficiente, flexible y su configuración se integra dentro del mismo servicio junto con la gestión del resto de componentes.
  En VMware vSAN, en lugar de tratar el almacenamiento de forma independiente este pasa a estar ligado a cada host, es decir, cada uno de los nodos tiene asignados hasta cinco grupos de discos. Estos grupos de discos pueden ser de tipo *Hybrid*, donde se combinan discos duros SSD y HDD, o *All-Flash*, donde todos los discos son de tipo SSD. Dentro de cada grupo, los discos se dividen en dos tipos con distintas funciones, el disco de caché y el disco de capacidad[8]:

- **Caché**: Hay uno en cada grupo. Realiza la función de memoria caché y se encarga de escribir los datos persistentes en los discos de capacidad.
- **Capacidad**: Puede haber hasta siete discos en cada grupo. Almacena los datos persistentes del entorno.

En cada grupo de discos la gestión de la lectura y escritura de datos se hace de la siguiente forma:

- **Lectura**: En el caso de la solución *Hybrid*, si el dato que se busca no está en el disco de caché entonces se busca en los discos de capacidad y después se incorpora al disco de caché. Con la solución *All-flash*, los datos se leen siempre directamente de los discos de capacidad sin que estos sean escritos en el disco de caché dejando a este completamente libre para las operaciones de escritura. Gracias a esto, la estructura *All-flash* ofrece mayor rendimiento respecto a la *Hybrid*.
- **Escritura**: Tanto en la solución *Hybrid* con en la *All-flash*, el host ESXi primero escribe en el disco de caché, este le responde con una confirmación de escritura y más tarde vSAN se encarga de escribir ese dato en los discos de capacidad cuando el disco de caché está casi completo o cuando el dato lleva un tiempo sin ser utilizado.

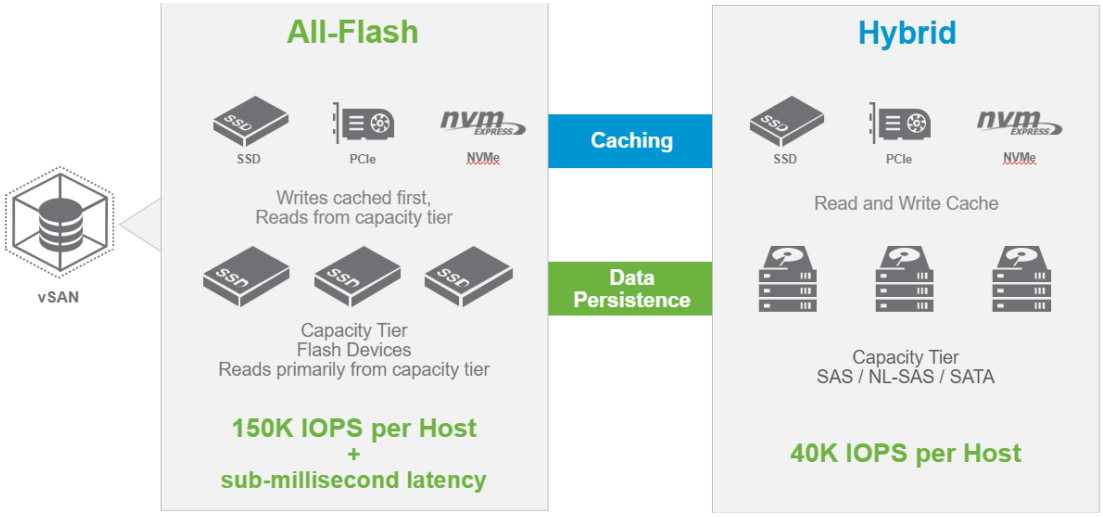


Figura 2.6: Almacenamiento All-Flash vs. Híbrido en vSAN

Los hosts acceden al *datastore* de vSAN mediante protocolo IP en una red accesible por todos los nodos.
Este componente permite reducir las tareas de gestión del almacenamiento físico ya que ya no es necesario hacer ajustes en la capa física para cumplir unos requisitos en la capa software.

- **VMware NSX-T**: otro de los componentes clave. Tiene un papel similar a vSAN, pero en este caso se encarga de virtualización de los componentes físicos de la red de la infraestructura, es decir, abstrae los componentes de la red para desacoplar la configuración de la red física y la red virtual del SDDC simplificando así las operaciones en la infraestructura física. Con esto, los administradores pueden configurar la red del SDDC sin tener en cuenta el equipamiento físico ya que trabajan con una red virtual independiente de la red física sobre la que funciona. Los componentes de VMware NSX-T se agrupan en tres capas/planos[9]:

  - **Management Plane**: desde este punto se gestiona la conectividad, seguridad y operaciones de los componentes del sistema a través de un inventario de objetos, proporciona una interfaz y una API para acceder acceder a los componentes de VMware NSX-T, almacena y transmite la configuración del sistema al *Control Plane* y obtiene información del sistema como estadísticas de uso. Estas funciones son implementadas por el componente **NSX-T Manager**.

  - **Control Plane**: este plano comunica al cada componente la configuración establecida desde *Management Plane*. Además, se encarga de obtener información sobre la topología de la red desde el *Data Plane*. Este plano se divide en dos partes:

    * *Central Control Plane* (CCP): implementado por un cluster de varias VM donde cada una contiene el componente **NSX-T Controller**, el cual es el encargado de controlar las redes virtuales existentes y de proveer de configuración a otros componentes como *logical switches*, *logical routers* y *edge nodes*. El componente **NSX-T Controller** está integrado, junto con el componente **NSX-T Manager**, en una única *appliance* denominada **NSX-T Manager Appliance**
    * *Local Control Plane* (LCP): este plano lo implementan los **transport nodes** y es responsable de monitorizar los enlaces locales y de reenviar entradas de configuración a los componentes del *Data Plane*.

  - **Data Plane**: este plano es responsable de retransmitir el tráfico por la red basándose en lass tablas y reglas generadas por el *Control Plane*. También informa al *Control Plane* sobre la topología de la red, recopila estadísticas a nivel de paquete y controla el estado de los enlaces para resolver posibles fallos. En definitiva, el *Data Plane* se centra en como gestionar los paquetes que circulan por la red. Está implementado por dos tipos de **transport nodes** (TN):

    * *Hypervisor Transpor Node*: son los hosts con el hipervisor ESXi y que utilizan VMware NSX-T para proveer servicios de red a las VM que funcionan sobre sobre el host.

    - *VMware NSX-T Edge Node*: también llamado *Edge Node*, se trata de una *appliance* instalado sobre un servidor físico o en forma de VM, que provee al entorno de una serie de servicios de red centralizados.

  VMware NSX-T inlcuye otros componentes:

    - **NSX-T Virtual Distributed Switch** (N-VDS): funciona sobre el hipervisor ESXi de cada host y es el encargado de retransmitir el tráfico entre las VM que se encuentran dentro de un mismo *transport node* y desde una VM a la red física de la infraestructura.
    - **Logical Router**: Integra dos componentes, **Distributed Router** (DR) proporciona conectividad entre las redes virtuales que se crean dentro del SDDC funciona de forma distribuída en los *transport nodes*), y **Service Router** (SR) da conectividad hacia redes redes externas al SSDC y ofrece una serie de servicios centralizados, NAT y DHCP entre otros.
    - **Logical Firewall**: NSX-T permite establecer reglas para gestionar el tráfico de entrada y de salida del SDDC. Estas reglas se pueden establecer a nivel de *Layer 2* o *Layer 3* de la red.
    - **Logical Load Balancer**: se encarga de distribuir el tráfico que recibe a los componentes correspondientes, permitiendo desacoplar el acceso a un servicio de su implementación. Esto permite que un servicio se pueda escalar y ofrecer alta disponibilidad.

- **VMware vRealize Automation**: se trata de un componente opcional que permite automatizar el despliegue de máquinas virtuales, procesos y aplicaciones reduciendo la complejidad y eliminando tareas manuales. Esto acelera la entrega del servicio gracias a que las tareas de aprovisionamiento y entrega de recursos y aplicaciones son más rápidas permitiendo establecer políticas de seguridad y control. Internamente, este servicio cuenta con los siguientes componentes [10] para cumplir con sus funciones:

    - **vRealize Automation Appliance**: contiene un portal donde los usuarios pueden acceder y gestionar y aprovisionar servicios cloud bajo demanda, un servicio de autenticación y una interfaz para gestionar este componente.
    - **IaaS Web Server**: realiza las tareas de administración de la infraestructura que se requieren desde el portal de vRealize Automation Appliance.
    - **Microsoft SQL Server**: se utiliza para almacenar información sobre los elementos de IaaS y sobre las máquinas virtuales controladas por vRealize Automation Appliance.

Decir como se puede implementar un sistema de facturación.*****
Como se puede conectar los usuarios de la UDC.******

Capítulo 3

Planificación

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

3.1 Tareas

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

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

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

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

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

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

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

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

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

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

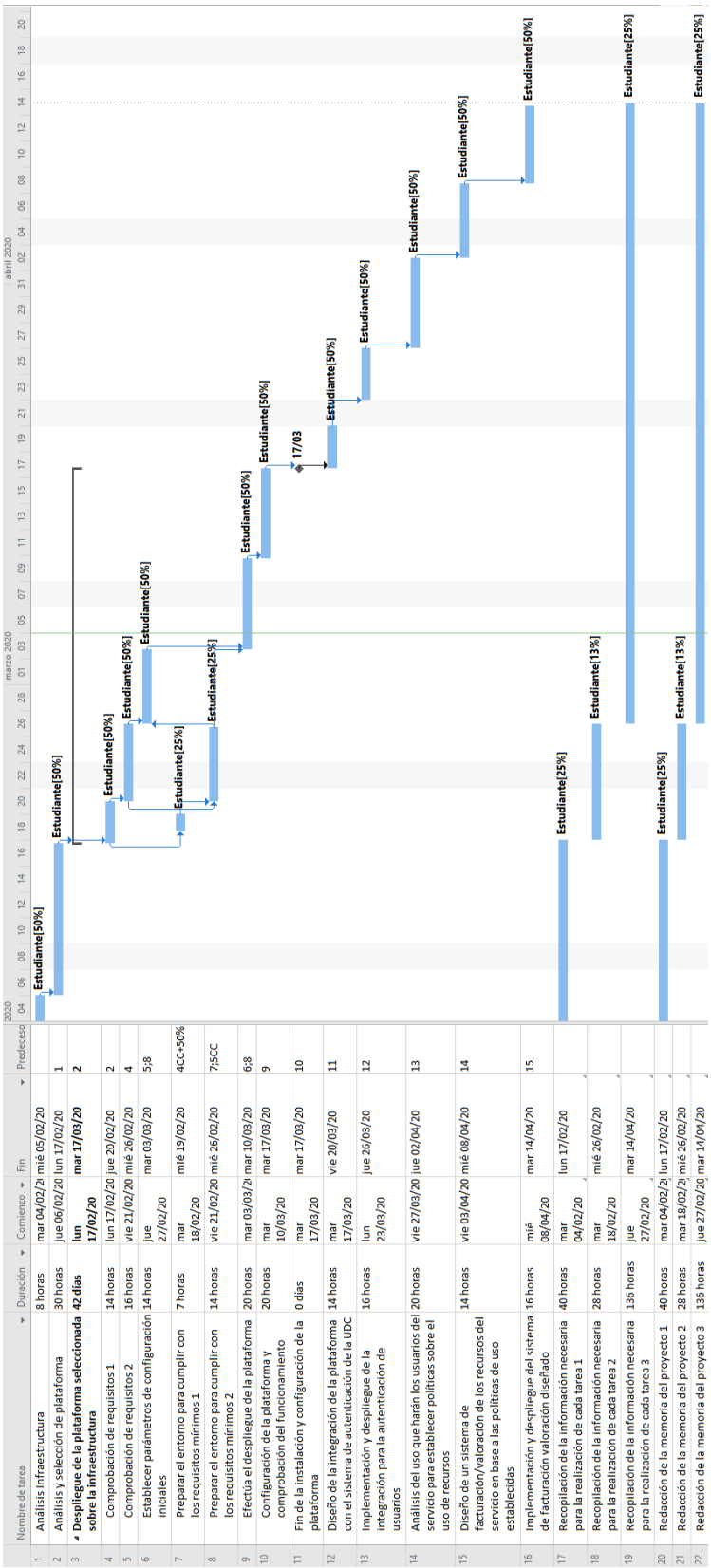
visionar permitiendo aumentar la eficiencia de los recursos físicos reduciendo la cantidad de recursos ociosos.

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

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

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

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



	Nombre de tarea	Duración	Comienzo	Fin	Predeceso
1	Análisis infraestructura	8 horas	mar 04/02/20	mié 05/02/20	
2	Análisis y selección de plataforma	30 horas	jue 06/02/20	lun 17/02/20	1
3	**Despliegue de la plataforma seleccionada sobre la infraestructura**	**42 días**	**lun 17/02/20**	**mar 17/03/20**	2
4	Comprobación de requisitos 1	14 horas	lun 17/02/20	jue 20/02/20	2
5	Comprobación de requisitos 2	16 horas	vie 21/02/20	mié 26/02/20	4
6	Establecer parámetros de configuración iniciales	14 horas	jue 27/02/20	mar 03/03/20	5;8
7	Preparar el entorno para cumplir con los requisitos mínimos 1	7 horas	mar 18/02/20	mié 19/02/20	4CC+50%
8	Preparar el entorno para cumplir con los requisitos mínimos 2	14 horas	vie 21/02/20	mié 26/02/20	7;5CC
9	Efectúa el despliegue de la plataforma	20 horas	mar 03/03/20	mar 10/03/20	6;8
10	Configuración de la plataforma y comprobación del funcionamiento	20 horas	mar 10/03/20	mar 17/03/20	9
11	Fin de la instalación y configuración de la plataforma	0 días	mar 17/03/20	mar 17/03/20	10
12	Diseño de la integración de la plataforma con el sistema de autenticación de la UDC	14 horas	mar 17/03/20	vie 20/03/20	11
13	Implementación y despliegue de la integración para la autenticación de usuarios	16 horas	lun 23/03/20	jue 26/03/20	12
14	Análisis del uso que harán los usuarios del servicio para establecer políticas sobre el uso de recursos	20 horas	vie 27/03/20	jue 02/04/20	13
15	Diseño de un sistema de facturación/valoración de los recursos del servicio en base a las políticas de uso establecidas	14 horas	vie 03/04/20	mié 08/04/20	14
16	Implementación y despliegue del sistema de facturación valoración diseñado	16 horas	mié 08/04/20	mar 14/04/20	15
17	Recopilación de la información necesaria para la realización de cada tarea 1	40 horas	mar 04/02/20	lun 17/02/20	
18	Recopilación de la información necesaria para la realización de cada tarea 2	28 horas	mar 18/02/20	mié 26/02/20	
19	Recopilación de la información necesaria para la realización de cada tarea 3	136 horas	jue 27/02/20	mar 14/04/20	
20	Redacción de la memoria del proyecto 1	40 horas	mar 04/02/20	lun 17/02/20	
21	Redacción de la memoria del proyecto 2	28 horas	mar 18/02/20	mié 26/02/20	
22	Redacción de la memoria del proyecto 3	136 horas	jue 27/02/20	mar 14/04/20	

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[11]. 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

4.1 Conceptos

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

4.1.1 Workload Domain

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

Management Domain

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

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

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

Virtual Infrastructure Domain (VI)

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

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

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

4.1.2 Arquitectura

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

Modelo estándar

Este modelo está pensado para desplegar VMware Cloud Foundation en entornos de tamaño medio/grande con un mínimo de siete hosts. Está formado por un *management domain* que se despliega en cuatro de los hosts y contiene todos los componentes de gestión de toda la infraestructura, desde este *workload domain* se administra la infraestructura del SDDC y cada *virtual infrastructure domain* existente. Además, este modelo contiene al menos un *virtual infrastructure domain*, creado bajo demanda y con capacidades establecidas según su finalidad que posteriormente se pueden pueden modificar, se despliega sobre al menos tres hosts. Un *management domain* puede gestionar un máximo de de catorce *virtual infrastructure domains*.

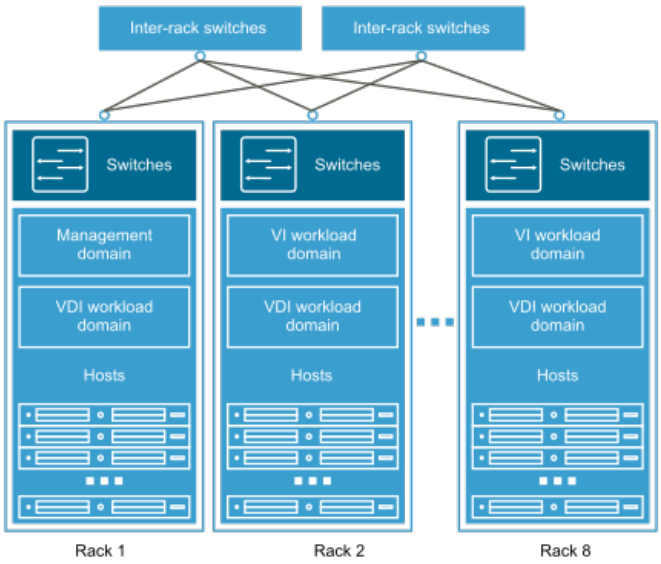


Figura 4.1: Esquema del modelo de arquitectura estándar.

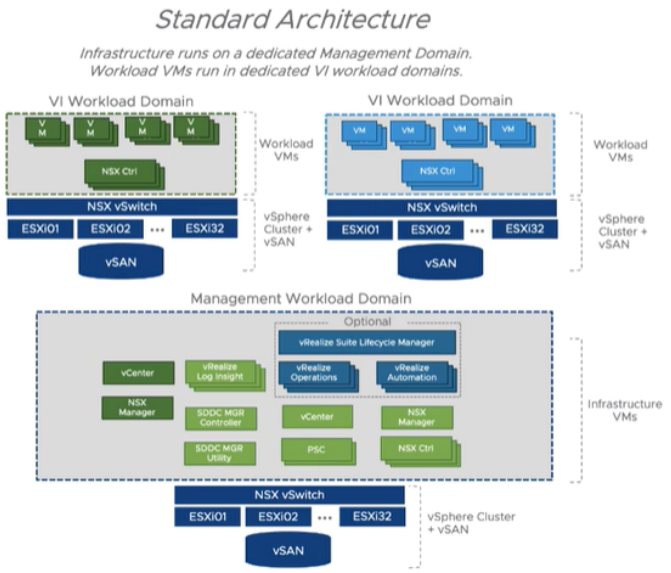


Figura 4.2: Estructura de los componentes en una arquitectura estándar.

Modelo consolidado

Este modelo está pensado para desplegar VMware Cloud Foundation en entornos de tamaño pequeño, normalmente cuando hay menos de siete hosts, aunque se puede también se puede utilizar sobre entornos más grandes de hasta 64 hosts. En este modelo los flujos de trabajo que corresponden al *virtual infrastructure domain* y al *management domain* en el despliegue estándar, están colocados dentro de un mismo *workload domain* en un único cluster

pero aislados gracias a que cada uno se coloca dentro de un *resource pool* diferente, es decir, existe un cluster con varios *resource pool*. El modelo consolidado se convierte en un modelo estándar cuando se añade un *workload domain* al SDDC.[Fig. 4.3].

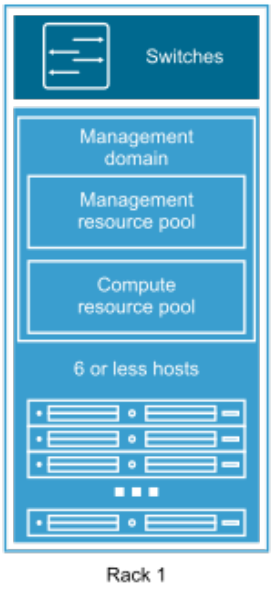


Figura 4.3: Esquema del modelo de arquitectura consolidado.

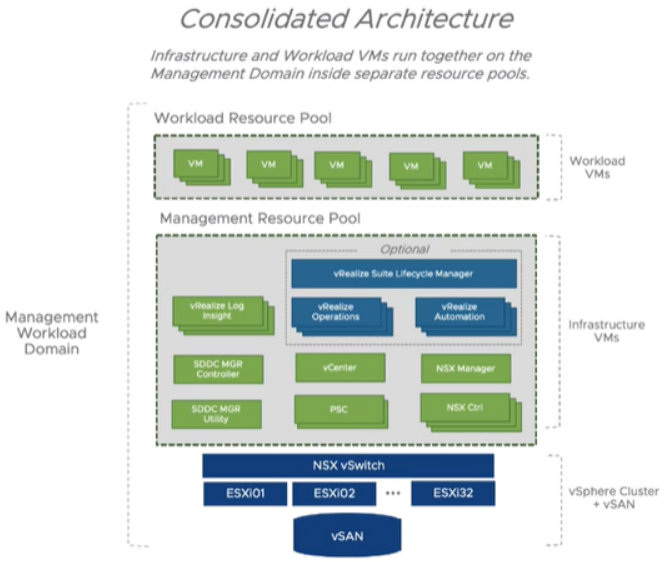


Figura 4.4: Estructura de los componentes de una arquitectura consolidado.

4.1.3 Clusters, zonas y distribución de un SDDC

AZ y Region

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

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

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

Cluster y Resource Pool

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

- Management cluster: es el cluster que se crea al desplegar VMware Cloud Foundation. Contiene los componentes para administrar los recursos del WD.

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

Dentro de un cluster se pueden crear resource pools. Un *resource pool* es una carcaterística de VMware vSphere que permite abstraer un conjunto de recursos de un cluster estableciendo unos límites de capacidad que puede usar [12]. Usar resource pools permite agrupar las VMs

con una finalidad similar y controlar la cantidad de recursos del WD que esas VMs pueden consumir.

4.2 Requisitos

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

4.2.1 Cómputo

Hosts ESXi

Para realizar el despliegue del primer WD (el *management domain*) se requieren al menos cuatro[1] hosts ESXi con al menos un total de 256 GB de memoria RAM y un disco de arranque de 32 GB cada uno. Para cada WD adicional solo se requiere un mínimo de tres hosts y la cantidad de memoria RAM depende de la finalidad del WD, por lo tanto para implementar el modelo de arquitectura estándar se requieren al menos siete hosts ESXi. Cada uno de los hosts debe tener al menos dos interfaces de red físicas (NIC) que soporten al menos 10 Gbit/seg de velocidad.

4.2.2 Almacenamiento

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

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

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

4.2.3 Red

Switch Top Of Rack

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

Servicios

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

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

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

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. La instalación de VMware Cloud Foundation se realiza con la herramienta VMware Lab Constructor (VLC)[4], que genera de forma automatizada una infraestructura embebida basada en el diseño de VMware Cloud Foundation propuesto por VMware sobre la cual despliega cada uno de los componentes para simular un entorno real.

4.3.1 Preparación

Host ESXi

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

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 Cetificate Authority. Otra VM contiene el sistema operativo VyOS que funciona como un router virtual y como servidor DHCP. Una última VM con Windows 10[5] se requiere para ejecutar VLC y acceder al entorno embebido generado por VLC. El servidor DNS contiene los *hostnames* y sus respectivas direcciones IP de todas las VMs, tanto las que residen dentro del host físico como las que se alojan dentro del entorno embebido generado por la herramienta VLC. Este servidor DNS implementa un único dominio que se denomina *pesci.domain*. El servidor Active Directory proporciona almacena usuarios y grupos de usuarios requeridos para establecer roles y proporcionar acceso a los componentes y servicios de VMware Cloud

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

Foundation. Se utiliza este servidor de usuarios en lugar del directorio real de la UDC para evitar posibles problemas del servicio. El router VyOS tiene configuradas todas las subredes y VLANs que VMware Cloud Foundation utiliza en la capa L3 de la infraestructura física y proporciona acceso a Internet, en las cuatro interfaces que conectan con las instancias de VMware NSX-T Edge utiliza enrutamiento dinámico BGP. El servidor DHCP asigna una dirección IP a cada TEP de cada host ESXi.

VMware Lab Constructor

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

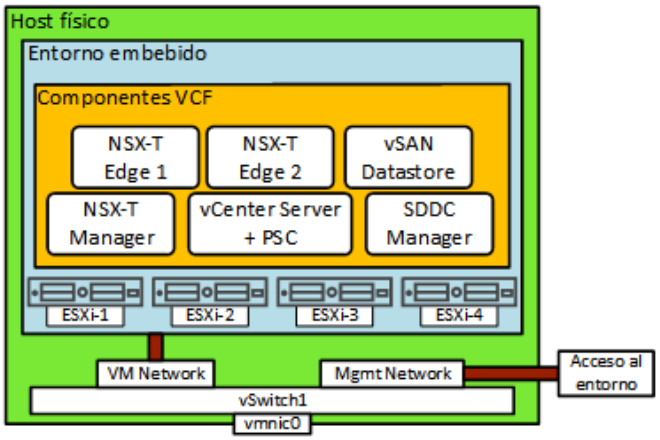


Figura 4.5: Muestra la estructura generada por el instalador VLC. Cuatro hosts ESXi embebidos con los componentes de VMware Cloud Foundation cuyo tráfico circula a través del *port group* VM Network.

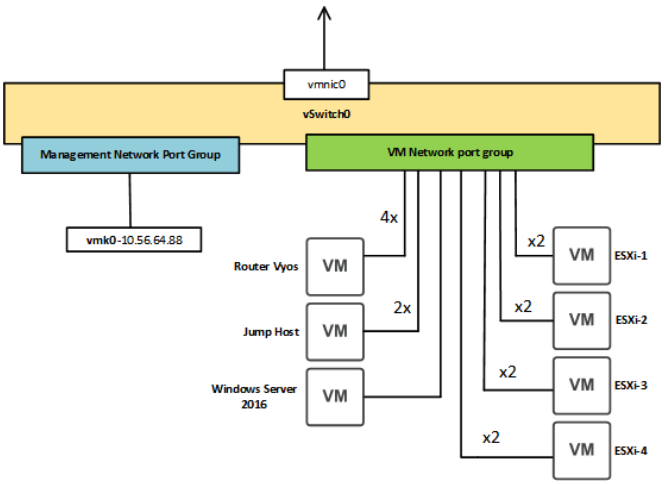


Figura 4.6: Muestra las VMs que están funcionando sobre el host físico y que representan los componentes de la infraestructura física de un SDDC real, junto con el número de interfaces que se utilizan en cada una. El router VyOS, Jump Host y Windows Server 2016 se configuran antes del despliegue de VMware 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 que se indica. Se utiliza la interfaz vmnic0 del host como salida del tráfico generado por el vSwitch0.

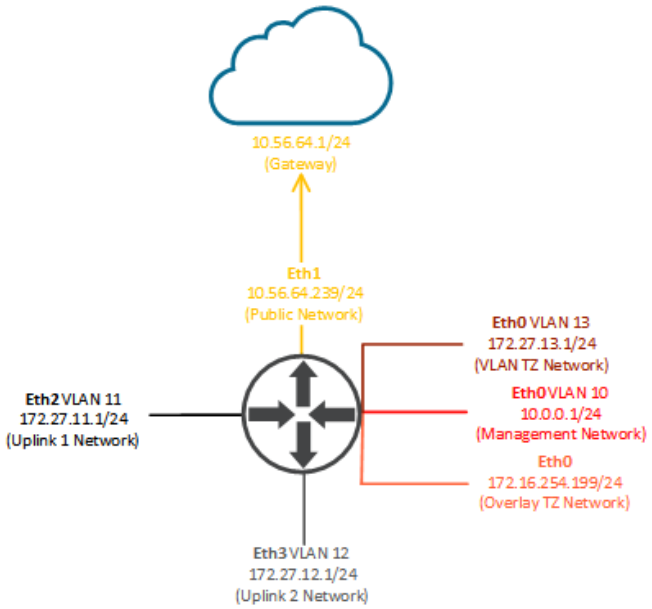


Figura 4.7: Muestra la configuración del router VyOS. Cada una de las interfaces se debe configurar antes del despliegue de VCF. Todas usan MTU de 8940 Bytes. En las interfaces Eth2 y Eth3 el router utiliza enrutamiento dinámico BGP donde el AS local es 65001 y el AS remoto es AS 65003, configurado para anunciar a sus vecinos la red 10.0.0.0/24 Management Network. Las direcciones configuradas como *neighbour* son: 172.27.11.2, 172.27.11.3, 172.27.12.2 y 172.27.12.3. En la dirección IP 172.27.254.199 de la interfaz eth0, el router proporciona un servidor DHCP que asigna direcciones IP en el rango 172.16.254.0 - 172.16.254.100.

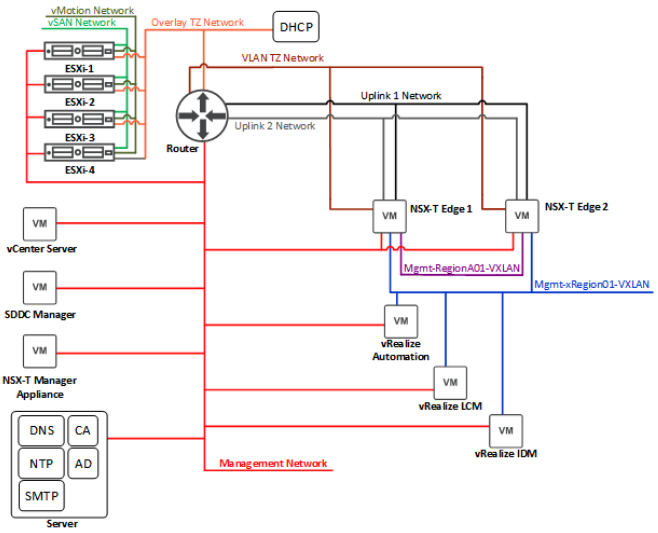


Figura 4.8: Muestra todos los componentes de VMware Cloud Foundation desplegados por VLC, como se conectan con los distintos servicios de red y a que redes se conectan. Las redes Mgmt-xRegion01-VXLAN y Mgmt-Region01A-VXLAN se corresponden a redes virtuales gestionadas por VMware NSX-T que no requieren 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 y servicios localizados en los hosts ESXi que forman parte de su dominio. En VMware Cloud Foundation se utiliza una instancia de VMware vCenter Server para controlar un *workload domain* para aislar el flujo de trabajo y gestión de cada uno, simplifica la escalabilidad del SDDC, la gestión de actualizaciones para los componentes es más sencilla, permite determinar roles específicos y responsabilidades y permite aislar las redes de otras instancias de vCenter Server. Además, para gestionar vSpehere SSO Domain, VMware vCenter Server contiene embebido el componente PSC con todos los servicios necesarios.En caso de que existan varios *Workload Domain* se puede habilitar el modo *Enhanced Linked Mode* para poder gestionar todas las instancias de vCenter Server de forma centralizada desde un único vSphere Client. Por lo anterior, en el *management domain* se despliega una instancia de VMware vCenter Server que incluye un cluster de VMware vSphere.

Diseño cluster VMware vSphere

Dentro de un *workload domain* se pueden crear varios clusters vSphere con diferentes características. Para crear cluster y establecer sus características hay que tener en cuenta que se pueden usar menos hosts ESXi de mayor capacidad o más hosts EXi 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 de un cluster vSphere. Para el *management domain* se utiliza un único cluster vSphere con un mínimo de 4 hosts de los cuales se reserva un host para proveer redundancia. En caso de que el *management domain* esté extendido en dos AZ entoces se requieren 4 hosts en cada AZ para proporcionar redundancia y disponibilidad en caso de caída de una de las AZ. Dentro del cluster hay que configurar los servicios vSphere HA y vSphere DRS para proteger los componentes del SDDC. La configuración que se establece en el *management domain* es la siguiente:

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

- **vSphere DRS**: se activa usando la opción por defecto, *Fully Automated* ya que aporta el mejor balance entre consumo de recursos y migraciones de VM innecesarias. Adicionalmente se pueden establecer reglas para determinar reglas de orden de encendido sobre grupos de VM. En caso de que exista más de una AZ, se deben crear grupos de VM y de hosts de cada AZ para luego implementar reglas de afinidad para que las VM de una AZ no sean migradas a otra AZ ya que esto puede afectar al rendimiento de la VM.

Diseño de red para el cluster vSphere

Si bien en VMware Cloud Foundation existe VMware NSX-T, un componente dedicado únicamente a la administración de la red del SDDC, es desde VMware vSphere dónde se crean los elementos para formar las redes básicas que conectan los principales componentes del SDDC y para entregar y configurar sus servicios. Estas redes se configuran en base a los siguientes aspectos:

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

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

- Se establece un MTU igual 9000 Bytes para permitir el tráfico de *jumbo frames* ya que son requeridos por algunos de los servicios.
- Se habilita el servicio *Network I/O* que permite establecer un nivel de prioridad a cada tipo de tráfico. Esto se realiza estableciendo limites de ancho de banda, políticas de balanceo de carga y reserva de recursos para un tipo de tráfico asociado a un servicio. Por cada tipo de tráfico hay cuatro aspectos que se pueden configurar que son *Shares* (indica el % de ancho de banda que se le da a un tipo de tráfico, el tipo de tráfico que tenga un mayor valor en *Shares* tendrá más prioridad a la hora de usar los recursos), *Reservation* (indica el valor de ancho de banda que se reserva para el tipo de tráfico) y *Limit* (establece un valor máximo para el ancho de banda de un tipo de tráfico). En el *management domain* los tipos de tráfico más relevantes que se deben configurar son los siguientes:

  - *Management Traffic*: el valor *Shares* se establece al 50% (*Normal*) lo cual le da mayor prioridad que el resto de tipos. El resto de valores no se modifican.
  - *vSphere vMotion Traffic*: el valor *Shares* se establece al 25% (*Low*) ya que durante el estado normal del entorno este tipo de tráfico no es muy importante. El resto de valores no se modifican.
  - *vSAN Traffic*: el valor *Shares* se establece al 100% (*High*) para garantizar que este servicio recibe la cantidad de ancho de banda que necesita. El resto de valores no se modifican.
  - *Virtual Machine Traffic*: el valor *Shares* se establece al 100% (*High*) para garantizar que las VMs siempre tienen acceso a la red ya que son una parte importante del SDDC. El resto de valores no se modifican.

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

- Como puertos de salida *Uplink* se configuran las dos interfaces físicas *vmnic0* y *vmnic1* de cada host.

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

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

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

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

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

- **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*.

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

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

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

- *Network failure detection*: esta opción permite establecer como debe determinar el *port group* que alguna de las NICs físicas está fuera de servicio. Se selecciona *Link status only* para que esto se determine según el estado que le transmite la NIC física, así se pueden detectar los fallos que ocurren en la red física.
- *Notify switches*: si está habilitada, permite a los host enviar *frames* a los switches físicos para que estos conozcan la localización de las VM que están funcionando en cada host. Se activa al seleccionar *Yes*.
- *Failback*: permite determinar como se reactiva una NIC cuando esta se recupera de un fallo. Al seleccionar *Yes* se establece 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 en cada estado de la conexión que puede ser *active* (cuando los uplinks seleccionados están activos), *stand by* (cuando alguno de los uplinks *active* está incativo). 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.

VMware NSX-T despliega en el *management domain* despliega tres instancias de NSX-T Manager aunque en el entorno solo se despliega una llamada *nsx-mgmt-1* para mejorar el rendimiento, y dos instancias de NSX-T Edge que en el entorno se denominan *edge01-mgmt* y *edge02-mgmt*, cada conjunto de instancias de cada componente forman un cluster donde cada VM está protegida por las funcionalidades vSphere HA y vSphere DRS para proveer alta disponibilidad del servicio y migrar las VMs a otra ubicación en caso de caída de una

AZ o de un host. Estas VMs están conectadas al *Distributed port group* llamado *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*. Si existe más de una AZ, varios de los *distributed port groups* se deben extender al resto de AZs para que en caso de que la primera AZ falle sus VMs se puedan migrar a otra AZ y sigan teniendo conectividad. Los *port groups* que deberían estar extendidos en todas las AZ son el *port group sddc-vds01-mgmt* de cada AZ[6], los *port group sddc-edge-uplink01*, *sddc-edge-uplink02* y *port group Edge Overlay*.

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 se pueden conectar a la vez a varias TZ de tipo VLAN pero solo a una TZ de tipo Overlay.

- **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*.

En el *management domain* del entorno existen dos *transport zones* diferentes:

- *mgmt-domain-m01-overlay-tz*:
  - Tipo: Overlay
  - Transport Nodes: *edge01-mgmt*, *edge02-mgmt*, *esxi1*, *esxi2*, *esxi3* y *esxi4*.
  - Segments:
    - *mgmt-xRegion01-VXLAN*:
      - VNI: 68589
      - Subred: 10.60.0.0/24
      - VMs: *vrlcm* (10.60.0.60), *vridm* (10.60.0.30)

[6] Cada AZ tiene su propio *Management port group*, entonces en cada AZ debe ser accesible el *Management port group* del resto de AZs.

        - Descripción: se usa para desplegar las aplicaciones (algunos productos de VMware vRealize Suite lo utilizan) que deben ser accesibles desde todas las *Regions* del SDDC, por lo tanto este *segment* debe estar extendido en todo el entorno para que las VMs que se alojen en él puedan mantener la misma configuración de red independientemente del lugar físico donde se encuentren.
    - *mgmt-Region01A-VXLAN*:
        - VNI: 67588
        - Subred: 10.50.0.0/24
        - Descripción: su finalidad es alojar aplicaciones que sean accesibles desde una misma *Region*. El alcance de estas aplicaciones está limitado a una *Region* por lo tanto este *segment* solo está extendido dentro de la misma.
    - *sddc-host-overlay*:
        - VNI: 67534
        - Descripción: *segment* usado por los componentes de VMware NSX-T para comunicarse entre los diferentes hosts ESXi.

- *sfo01-m01-edge-uplink-tz*:
    - Tipo: VLAN
    - Transport Nodes: *edge01-mgmt* y *edge02-mgmt*.
    - Segments:
        - *VCF-edge-mgmt-cluster-segment-11*:
            - VLAN: 11
            - VMs: *edge01-mgmt* (172.27.11.2/24) y *edge02-mgmt* (172.27.11.3/24).
            - Descripción: usado para transmitir el tráfico saliente hacia la red física.
        - *VCF-edge-mgmt-cluster-segment-12*:
            - VLAN: 12
            - VMs: *edge01-mgmt* (172.27.12.2/24) y *edge02-mgmt* (172.27.12.3/24).
            - Descripción: usado para transmitir el tráfico saliente hacia la red física.

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. El primer tipo de TZ utiliza el protocolo Geneve para crear un túnel entre el origen y el destino, el cual encapsula en paquetes UDP el tráfico de L2 que generan las redes lógicas de VMware NSX-T añadiendo un identificador llamado VNI que indica a que *segment* pertenece. Una TZ de

tipo VLAN utiliza encapsula el tráfico saliente a la red física añadiendo un identificador de VLAN que es el mismo para todos los *segments* que pertenecen a la misma TZ VLAN. Esta encapsulación 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. Los TNs que son hosts ESXi obtienen su dirección TEP de un servidor DHCP[7] mientras que los que son instancias de NSX-T Edge la dirección IP se asigna de forma manual. El TEP de cada TN tiene las siguientes direcciones IP:

- *esxi-1*: 172.16.254.10, 172.16.254.11
- *esxi-2*: 172.16.254.12, 172.16.254.13
- *esxi-3*: 172.16.254.14, 172.16.254.15
- *esxi-4*: 172.16.254.16, 172.16.254.17
- *edge01-mgmt*: 172.27.13.2, 172.27.13.3
- *edge02-mgmt*: 172.27.13.4, 172.27.13.5

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.

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*. Este consiste en que cuando desde un TN se envía algún tipo de tráfico BUM de un *segment* cuyos TNs están distribuídos en distintos puntos de la infraestructura física, el TEP del TN origen detecta que debe retransmitir el tráfico a una red externa, en la red externa se selecciona un TN que recibe el tráfico y lo reenvía al resto de TNs dentro del mismo dominio de red que deben recibir ese tráfico, así se reduce el número de paquetes que el TN origen debe enviar. El tráfico solo se enviará a los TN que contienen VMs que forman parte del *segment* que lo genera. Es el componente NSX-T Controller quien se encarga de actualizar e indicar a los TNs toda la información para que esta comunicación sea correcta.

En cuanto al enrutamiento lógico, VMware NSX-T define routers virtuales que se encuentran embebidos dentro del hypervisor de cada host ESXi, igual que cada *segment*. Esta

[7] 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.

característica permite definir *gateways* para los *segments* creados sin necesidad de modificar la configuración de la red física. En el *management domain* se usa el modelo de enrutamiento de doble capa (*Two Tier Routing*) que utiliza dos routers de dos tipos distintos lo que permite separar las tareas de administración de la infraestructura lógica de las de gestión de la infraestructura del usuario. Los routers desplegados son un *Tier-0* llamado *mgmt-domain-tier0-gateway* y otro router *Tier-1* llamado *mgmt-domain-tier1-gateway*. Cada router lógico está formado por los siguientes elementos:

- **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 *Tier-0*.

- **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 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.

Las interfaces de cada router lógico son de los siguientes tipos:

- *External Interface*: se refiere a las interfaces que conectan con el dispositivo de la red física, normalmente un router.

- *Internal Transit Link*: está interfaz conecta a todos los DR de *Tier-0* distribuídos en cada TN con los SR formando una única red.

- *RouterLink Interface*: interfaz que conecta un router lógico de *Tier-1* con uno de *Tier-0* a través de una subred generada por defecto (100.64.0.0/16).

Los routers lógicos creados para el *management domain* en el entorno tienen la siguiente configuración:

- **mgmt-domain-tier0-gateway**: su función consiste en proporcionar acceso a la red física.

  - *External Interfaces*[8]: Dos que se conectan al *segment VCF-edge_mgmt-edge-cluster-segment-11* (172.27.11.2 y 172.27.11.3) y dos que se conectan al *segment VCF-edge_mgmt-edge-cluster-segment-12* (172.27.12.2 y 172.27.12.3).
  - *Internal Transit Link*: Usa la subred 169.254.0.0/24.
  - *RouterLink Interface*: Tiene la dirección 100.64.192.0/31.
  - Configuración: se utiliza BGP para comunicarse con el router físico a través de las *external interfaces* (en la configuración de BGP se establece 65003 como el *autonomous system* local)[9], se activa *Inter SR iBGP* para establecer una comunicación mediante BGP entre las instancias distribuídas del componente SR de *mgmt-domain-tier0-gateway*[10] 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 (existen cinco rutas posibles para alcanzar el router físico).
  - Configuración HA: determina el modo en el que se va a ejecutar el router de *Tier-0*. En este caso se selecciona el modo *active-active* el cual implica que las instancias de SR (una en cada nodo NSX-T Edge) funcionan ambas de forma activa, el tráfico que reciba cada una será encaminado por una subred diferente. Es por esto que el router lógico de *Tier-0* no puede ofrecer servicios centralizados.

- **mgmt-domain-tier1-gateway**: router de *Tier-1* dedicado a gestionar el enrutamiento de las VMs de aplcaciones no dedicadas a la administración del SDDC.

  - *Internal Transit Link*: Usa la subred 164.254.0.0/24.
  - *RouterLink Interface*: Tiene la dirección 100.64.192.1/31.
  - *Segment Interfaces*: conectado al *segment mgmt-Region01A-VXLAN* con la dirección IP 10.50.0.1, y al *segment mgmt-xRegion01-VXLAN* con la dirección IP 10.60.0.1.
  - Configuración HA: determina el modo en el que se va a ejecutar el router de *Tier-1*. En este caso se selecciona el modo *active-stanby* el cual solo utiliza una de las dos instancias del SR de *Tier-1* (cada una en un nodo de NSX-T Edge) por lo tanto el tráfico que reciba solo será encaminado por un único punto. Esto permite activar servicios centralizados, por lo tanto será *mgmt-domain-tier1-gateway* y no *mgmt-domain-tier0-gateway* el encargado de proporcionarlos.

[8] Estas direcciones corresponden a las interfaces de los TN NSX-T Edge que conectan con las interfaces Uplink sobre dos *segments* de tipo VLAN.

[9] 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. El protocolo BGP también debe estar configurado en el router físico.

[10] Cada una se encuentra en una instancia de NSX-T Edge respectivamente.

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 entorno posee un N-VDS, este elemento conecta sus interfaces a los *segments* que se configuran en cada TN y establece un mapeo con las interfaces *uplink* que se utilizan para dirigir el tráfico de cada TZ hacia el exterior del TN. En el caso de las instancias de NSX-T Edge, los dos *uplinks* están mapeados con las NICs físicas de cada host ESXi a través de los dos *distributed port groups* de vSphere vDS (*sddc-edge-uplink01* y *sddc-edge-uplink02*) a los que están ancladas ambas VMs, es decir, un *uplink* está mapeado con una NIC física del host ESXi. En las TZ de tipo VLAN, se utilizan plantillas *Uplink Policy* para indicar como debe el N-VDS tratar el tráfico de la *transport zone* a la que se asigna. En cada *Uplink Policy* se especifican varias *Teaming Policy*, el identificador VLAN que debe usar el N-VDS cuando tiene que enviar el tráfico fuera del TN y el MTU de los *uplinks*. Una *Teaming Policy* indica como el N-VDS utiliza los *uplinks* para conseguir conexiones redundantes y balanceo de la carga, en una *Uplink Policy* se especifica una *Teaming Policy* por defecto y otras adicionales. Para la TZ de tipo VLAN *sfo01-m01-dge-uplink-tz* se especifica la siguiente *Uplink Policy*:

- Nombre: *uplink-profile-13*
- Transport VLAN: 13
- MTU: 8940
- Teaming Policy: se especifican tres, una por defecto y dos adicionales:
  - *Default teaming*: *Load Balance Source*, hace un mapeo uno a uno entre cada interfaz virtual de cada VM y uno de los *uplinks* del N-VDS, así todo el tráfico correspondiente a esa interfaz se envía y recibe por el mismo *uplink*.
  - *uplink1-named-teaming-policy*: *Failover Order*, se establece un *uplink*, *uplink1* en este caso, como activo que se utiliza para enviar todo el tráfico, y una lista de *uplinks* ordenados que se utilizan en caso de que el primero no esté disponible, vacía para esta *Teaming policy*.
  - *uplink2-named-teaming-policy*: *Failover Order*, donde el *uplink* activo es *uplink2* y la lista de *uplinks* de reserva está vacía.

A los *segments* de esta TZ, *VCF-edge-mgmt-cluster-segment-11* y *VCF-edge-mgmt-cluster-segment-12* se les asignan las *Teaming Policy uplink1-named-teaming-policy* y *uplink2-named-teaming-policy* respectivamente. Con esta configuración el tráfico de cada *segment* circula por una única NIC física del host ESXi. Esto, junto con la configuración *Failover Order* establecida para los *port groups sddc-edge-uplink01* y *sddc-edge-uplink02* en el vSphere vDS,se consigue que el tráfico de salida hacia la red física perteneciente a los componentes de VMware NSX-T y

todas las aplicaciones cuya red gestiona VMware NSX-T, sea distribuído por dos redes distintas proporcionando redundancia y disponibilidad del servicio en caso de que ocurra una caída de alguna de las conexiones.

Estas conexiones están gestionadas por un N-VDS dentro de cada instancia. Este switch lógico utiliza tres interfaces que se conectan a las diferentes redes lógicas. Para aquellas redes lógicas que requieren salida al medio físico ya sea para comunicarse con otros TN o para acceder a la red externa, el N-VDS utiliza finalmente el switch VDS de VMware vSphere que conecta con las interfaces físicas del host ESXi donde corre la instancia de NSX-T Edge.

la utilizan tres interfaces para Estas interfaces son *eth0* que se dedica a la red *Management*, *fp-eth0* y *fp-eth1* que ambas se dedican a la conexión con cada uno de los *segments* Uplink.

****

N

4.3.3 Operaciones de la Arquitectura[1]

En este apartado se define como se gestionan en VMware Cloud Foundation las tareas de administración de todas las partes de la infraestructura. Estas tareas se agrupan en la gestión del ciclo de vida y la recopilación de información sobre el estado de cada componente existente.

Gestión del Ciclo de Vida

Elementos que se encargan de administrar el ciclo de vida de los componentes:

- **vRealize Suite Licfecycle Manager**: componente utilizado para desplegar, actualizar y configurar, de forma automatizada, los productos vRealize Operations, vRealize Log Insight, vRealize Automation y vRealize Business Cloud. De este componente se despliega una única instancia en una AVN accesible desde cada *region* por todas las instancias de VMware vCenter Server. Se debe registrar su nombre de dominio en el servidor DNS para hacerla accesible.

Apéndices

Glosario

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

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

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

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

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

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

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

Almacén de datos

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

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

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

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

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

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

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

vCPU

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

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

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

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

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

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

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

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

VLAN trunk : enlace que permite comunicar distintas VLANs.

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

Rack : armario metálico destinado a alojar servidores físicos.

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

Bibliografía

[1] VMware, "Operations management design." [Online]. Available: https://docs.vmware.com/en/VMware-Validated-Design/5.1/sddc-architecture-and-design/GUID-CBA6012B-0B91-4831-A5AD-521340AD72D4.html

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

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

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

[5] VMware, "VMware ESXi." [En línea]. Disponible en: https://www.vmware.com/latam/products/esxi-and-esx.html

[6] "Vmware integrated openstack." [En línea]. Disponible en: https://www.vmware.com/es/products/openstack.html

[7] VmWare, "Cloud foundation components." [En línea]. Disponible en: https://docs.vmware.com/en/VMware-Cloud-Foundation/3.0/com.vmware.vcf.ovdeploy.doc_30/GUID-07411CF9-AD3F-43EA-A348-A89940C2D4A2.html

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

[9] P. Cerda, "Nsx: Arquitectura y componentes." [En línea]. Disponible en: https://patriciocerda.com/nsx-arquitectura-y-componentes/

[10] VMware, "vrealize automation architecture." [En línea]. Disponible en: https://docs.vmware.com/en/VMware-Validated-Design/4.3/com.vmware.vvd.sddc-design.doc/GUID-73A3C12D-5F2E-4CE1-82D0-136218771065.html

[11] 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

[12] 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