FACULTAD DE INGENIERÍA

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE MÁSTER EN TECNOLOGÍAS DE LA INFORMACIÓN MENCIÓN EN REDES DE COMUNICACIONES

Tema: Estudio para la optimización del centro de datos para un ISP de la ciudad de Quito a través de una solución de SDN.

AUTOR: CRISTHIAN ANDRES CARRERA GUERRERO

Quito, 2021 - MAYO

AUTORÍA

Yo, Cristhian Andrés Carrera Guerrero, portador de la cédula de ciudadanía No 1723442321, declaro bajo juramento que la presente investigación es de total responsabilidad del autor, y que se ha respetado las diferentes fuentes de información realizando las citas correspondientes. Esta investigación no contiene plagio alguno y es resultado de un trabajo serio desarrollado en su totalidad por mi persona.

_____

Cristhian Andrés Carrera Guerrero

DEDICATORIA

Este tiempo de estudios y trabajo va dedicado a Dios, mis padres y mi familia.

A mi esposa Gaby por ser mi apoyo y ayuda, mi hija Sophy mi mayor bendición, son el motor que me impulsa a seguir adelante y nunca rendirme.

Cristhian Carrera

Tabla de Contenidos

Índice de figuras

Índice de tablas

RESUMEN

Este caso de estudio en el presente proyecto se basa en establecer una propuesta de diseño y solución para optimizar la infraestructura de red de un centro de Datos atreves de una solución con Redes Definidas por Software (SDN), donde se debe tener una sola plataforma de administración, donde la solución deba ser escalable, considerando que los equipos deben tener puertos con densidades de velocidad que van hasta los 100 Gbps.

Se realiza una comparativa mediante diferentes marcas en el mercado de IT, las mismas que se encuentran en el cuadrante de Gartner, las marcas que se comparan son Cisco, Huawei, Arista, las cuales cumplen con diferentes características que definirán la opción más adecuada.

2. JUSTIFICACIÓN

Actualmente el constante cambio de las necesidades de las empresas debido al crecimiento y giro del negocio de las mismas ha generado que estas deban someterse a una transformación digital, estos cambios producen que se busque redes de alto rendimiento, integrales, y bajo una administración centralizada para cubrir todas las necesidades y exigencias tanto de los administradores de red como de los usuarios.

Debido a esto, un esquema de gestión de recursos para los administradores de centros de datos es el centro de datos definido por software; lo que esto significa es que mediante una única interfaz de software se puede configurar fácilmente los requerimientos de un Centros de Datos de alto nivel con una menor intervención humana. Es decir, un solo administrador de red puede ejecutar una serie de comandos para asignar máquinas virtuales, redes y almacenamiento virtuales a un nuevo cliente con garantías de nivel de servicio especificadas y dicha tarea puede ser completada en unos minutos.

Por lo que mediante este estudio se desea realizar la optimización del Centro de Datos de un proveedor de servicios de internet (ISP) de Quito mediante una solución de redes definidas por software (SDN) que permitirá ofrecer servicios de calidad, crecimiento rápido, cambios o implementaciones con el menor impacto y siempre a la vanguardia de la tecnología moderna.

Adicional, esta tecnología, ofrece flexibilidad y un control de costos, porque, a diferencia de como ocurre en una estructura física fija, los recursos pueden ampliarse cada vez que sea necesario.

Una vez aplicadas las acciones de optimización, se requiere minimizar la complejidad en las configuraciones ejecutadas por la interfaz de línea de comandos (CLI), lo influiría en un menor error por el factor humano y aumentará la operabilidad de los administradores.

3. PLANTEAMIENTO DEL PROBLEMA

Los enfoques tradicionales para realizar procesos manuales, limitado rendimiento y poca escalabilidad de la red actual en un Centro de Datos ya no son adecuados ni sostenibles, ya que el equipamiento que se maneja en dichos centros crece día a día, lo cual impide al administrador de la red mantener la adecuada gestión para cubrir el despliegue de nuevo equipamiento o despliegue de servicios. (Villarrubia, 2018)

Durante el despliegue o mantenimiento de los equipos, es posible que se produzcan muchos errores por el factor humano, o incluso que las mismas tareas afecten a la operabilidad del equipo de IT para cubrir otros temas, ya que se puede requerir mayor tiempo y esfuerzo por no contar con una administración centralizada de la red del Data Center (DC). (Adolfo Manaure, 2015)

Adicionalmente la operación común consiste en aplicar cambios, ejecuciones de comandos a través de la interfaz de línea de comandos (CLI) de cada equipo, tarea con la cual no es posible predecir el impacto que producirán dichos cambios a menos que se tenga un amplio conocimiento de toda la red, lo cual no suele ser así, SDN permite visualizar gráficamente el efecto que podrían producir los cambios a los equipos de la red, aun sin ser ejecutado, para que en base a ello el administrador pueda tomar la mejor decisión. (Tejedor, 2016)

4. OBJETIVOS

a) Objetivo general

Elaborar un estudio para una propuesta de diseño y prueba de concepto para la optimización del Centro de Datos de un ISP de Quito, enfocado en tres puntos esenciales: alta disponibilidad, plataforma común de administración y rendimiento escalable.

b) Objetivos específicos

Analizar la tecnología SDN para un Centro de Datos y la sistemática de la arquitectura que permita tener una visión centralizada e integral de las aplicaciones, además de una supervisión del estado de estas en tiempo real en todos los entornos virtuales y físicos.

Realizar el levantamiento de información de la infraestructura actual, problemas generados, sus limitantes y cómo influyen directamente en toda la red, mediante el análisis de las conexiones, topología y características del actual modelo del Centro de Datos.

Definir las características y componentes a nivel de hardware mínimas a ser consideradas para la implementación de una nueva arquitectura de Centros de Datos basada en SDN.

Establecer las políticas de configuración interna que funcionarán en las plantillas para la implementación de la solución basada en SDN.

Diseñar una topología de red basada en la tecnología SDN que permita la centralización, control y la optimización de la infraestructura de red para un Centros de Datos por medio del análisis de la solución ACI.

Explicar las ventajas del uso de una arquitectura de red basada en SDN, mediante la verificación de cumplimiento de políticas y la medición de parámetros de rendimiento de la red para su posterior puesta en producción.

CAPÍTULO 1 FUNDAMENTACIÓN TEÓRICA

1.1 Red tradicional en Centros de datos

Un centro de datos contiene un conjunto de hardware y software heterogéneo creado para alojar servicios y almacenar datos de una empresa. Como resultado los administradores de IT se enfrentan a grandes desafíos cuando se trata de la administración, operación y gestión con el fin de tener éxito en su labor, además de monitorear dichos recursos de manera óptima y adecuada, entendiendo que las fallas físicas y lógicas afectan al usuario y su percepción del servicio.

1.1.1 Características de una red tradicional en Centro de datos.

La infraestructura de red de un Centro de Datos para grandes empresas está basada en un modelo jerárquico de 3 capas como se indica en la figura 1; este modelo de red es conocido como arquitectura de 3 niveles o también conocido como ethernet clásico; utiliza las redes de área local (LAN) de arquitectura jerárquica virtual (VLAN) para la segmentación de redes, un protocolo de capa dos del modelo OSI llamado spanning tree para control de ruta infinita y enrutamiento como la mejor manera de entregar paquetes, siendo aún el modelo predominante en los Centros de Datos. (Mallick, 2019)

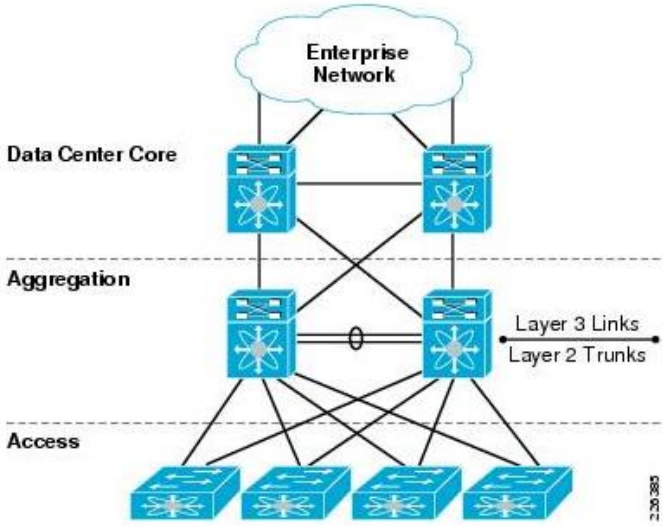


Figura 1. Modelo Jerárquico en la red de un Centro de Datos
Fuente: (Mallick, 2019)

De manera más específica un Data Center tradicional tiene un árbol de tres capas y raíces múltiples en la arquitectura, como se muestra en la Figura 2 (adaptado de la figura por

Cisco). Por lo general, consta de capa de núcleo, agregación y borde, la conmutación se realiza de arriba hacia abajo.

Los enlaces ascendentes de los conmutadores en la capa núcleo conectan el Centro de Datos a Internet, los conmutadores en el núcleo y las capas de agregación se interconectan para construir lógicamente gráficos bipartitos con enlaces 10G, y los servidores están conectados directamente a los conmutadores en la capa de borde con enlaces 1G.

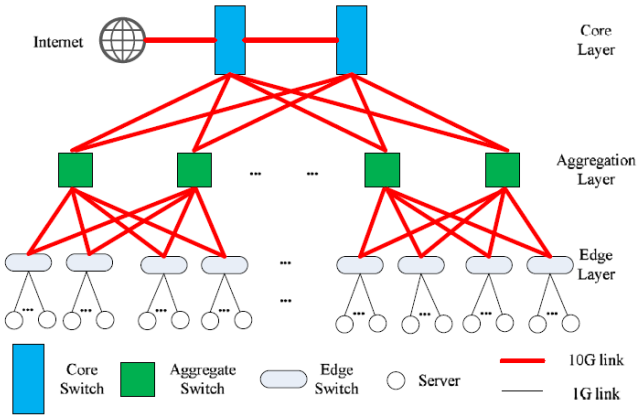


Figura 2. Arquitectura de red de un Centro de Datos tradicional
Fuente: (Chen et al., 2016)

1.1.2 Protocolo Spanning Tree.

El protocolo de árbol de expansión (STP) está definido por IEEE 802.1D. Evita lazos de capa 2 cuando los conmutadores están interconectados a través de múltiples rutas. STP está implementado mediante conmutadores que intercambian mensajes BPDU con otros conmutadores para detectar lazos, este algoritmo garantiza que exista una y solo una ruta activa entre dos dispositivos de red. (Bruno Anthony, 2017)

Como indica Bruno Anthony (2017), los puertos del conmutador STP entran en los siguientes estados:

Bloqueo: un puerto que provocaría un bucle de conmutación si estuviera activo. Los datos BPDU todavía se reciben en estado de bloqueo, pero evita el uso de caminos que podrían provocar un bucle.

Escucha: el conmutador procesa los paquetes de tipo BPDU y espera una posible nueva información antes de cambiar de estado.

Aprendizaje: aunque el puerto aún no reenvía tramas, sí aprende direcciones de origen de las tramas recibidas y las agrega a la base de datos de filtrado. Esto llena la tabla de direcciones MAC, pero no reenvía tramas.

Reenvío: un puerto que recibe y envía datos, en funcionamiento normal. STP aun monitorea los paquetes BPDU entrantes que indicarían que debería volver al estado de bloqueo para evitar un lazo.

Inhabilitado: no forma parte estrictamente de STP. Un administrador de red puede desactivar manualmente un puerto.

Los conmutadores Cisco admiten tres tipos de STP:

- Per VLAN Spanning Tree Plus (PVST+),
- Rapid-PVST+,
- Multiple Spanning Tree (MST).

PVST+ Per VLAN Spanning Tree Plus (PVST+) proporciona la misma funcionalidad que PVST utilizando tecnología de enlace 802.1Q en lugar de ISL.

PVST+ es una mejora del Especificación 802.1Q y no es compatible con dispositivos que no son de la marca Cisco. PVST+ se basa en IEEE 802.1D y agrega características patentadas por dicho vendor como BackboneFast, UplinkFast y PortFast.

Rapid_PVST + se basa en el estándar Rapid STP (RSTP) IEEE 802.1W. RSTP (IEEE 802.1w) incluye de forma nativa la mayoría de las mejoras patentadas de Cisco a la extensión 802.1D, como BackboneFast y UplinkFast. Rapid-PVST + tiene estas características únicas:

Utiliza Bridge Protocol Data Unit (BPDU) versión 2, que es compatible con versiones anteriores el 802.1D STP, que a su vez usa BPDU versión 0.

Todos los conmutadores generan BPDU y envían a todos los puertos cada 2 segundos, mientras que en 802.1D STP solo el puente raíz envía las BPDU de configuración.

Funciones de puerto: raíz, designado, alternativo y de respaldo.

Estados del puerto: descartar, aprender y reenvío.

Entre los tipos de puerto se tiene: el de borde (PortFast), punto a punto y puerto compartido.

Rapid-PVST utiliza RSTP para proporcionar una convergencia más rápida. Cuando cualquier puerto RSTP recibe tramas 802.1D de legado BPDU, recurre al STP legado y los beneficios inherentes de convergencia rápida de 802.1W se pierde cuando interactúa con puentes de legado.

MST está definido por IEEE 802.1S. Se basa en el árbol de expansión de múltiples instancias protocolo (MISTP). MISTP (802.1S) es un estándar IEEE que permite que varias VLAN sean mapeado a un número reducido de instancias de árbol de expansión. Esto es posible porque la mayoría las redes no necesitan más que unas pocas topologías lógicas. Cada instancia maneja múltiples VLAN que tienen la misma topología de capa 2.

Este tipo de modelo permite tener una gran cantidad de puertos en la capa de acceso, sin embargo, el problema con este diseño de red tradicional es el alto costo y que no es determinista; es decir que la comunicación va siempre desde la capa de agregación hacia el núcleo generándose múltiples saltos frecuentemente a través del Backplane con exceso de suscripción, que puede conllevar a un mayor tiempo en la conmutación de paquetes. Considerando el drástico incremento de servicios en la nube, los Centros de Datos deben poder interconectarse con cientos de miles o incluso millones de servidores, que proporcionan suficiente ancho de banda para garantizar la calidad de los servicios en la nube, además deben ser flexibles, confiables y tener una alta densidad para garantizar que las diversas aplicaciones se ejecutan de manera constante y eficiente. (Mallick, 2019)

1.1.3 Problemas en los centros de datos tradicionales.

Como lo indica Chen (2016), existen varios problemas con las redes tradicionales ya que una vez que se establece la infraestructura de red jerárquica, realizar cambios es difícil, añadir un nuevo rack de equipos no solo significa otro switch de acceso, sino posiblemente otro switch de agregación o incluso más puertos en switch de Core, además, la visibilidad del tráfico es baja y la depuración puede considerarse un desafío.

Difícil administración: debido a que están integrados verticalmente, cada uno de los dispositivos tiene su propio firmware, el mismo que está en su memoria, por lo tanto,

realizar actualizaciones, cambios de red implica una configuración manual en cada equipo lo que aumenta la complejidad y da cabida a posibles errores.

Poca flexibilidad: esto debido a que el número de puerto de los conmutadores de agregación determinan el número máximo de servidores admitidos en las arquitecturas en forma de árbol de múltiples raíces. Ahora si existiera una necesidad de incrementar los servidores y si los puertos de los conmutadores están al tope de su capacidad, los conmutadores actuales deben reemplazarse por otros nuevos con más puertos. Sin embargo, este tipo de implementación incremental consume mucho tiempo y es costoso.

Cableado complejo: una vez que la infraestructura de un Centro de datos tradicional se expanda a un tamaño grande, la cantidad de cables puede ser enorme convirtiéndose en una tarea pesada y compleja a medida que aumentan los servidores. El cableado y el sistema de refrigeración se enfrentarán a un gran desafío.

La capacidad de sus enlaces es limitada: este es un problema común presentado y que es conocido como sobresuscripción que generalmente ocurre cuando se usa un Centro de Datos tradicional para reducir el costo de operación; un ejemplo esto puede ser: ocho enlaces descendentes de un switch de acceso se pueden enrutar a un solo enlace ascendente, por lo que el ancho de banda de un servidor es realmente limitado cuando las cargas de trabajo incrementan llegando a tener picos de consumo, los conmutadores centrales pueden convertirse en cuellos de botella, lo que hace que el rendimiento del Centro de Datos tradicional se degrade abruptamente y corra el riesgo de sufrir una caída. (Chen et al., 2016)

Los Centros de Datos modernos deben evitar las desventajas de los tradicionales y poder hacer uso de la capacidad completa de sus enlaces, tener buena escalabilidad, alta utilización, cableado sencillo y bajo costo para proporcionar servicios en la nube de alta calidad.

1.2 Red Definida por software (SDN)

La red definida por software (SDN) es una arquitectura emergente que es dinámica, manejable, rentable y adaptable, lo que la hace ideal para la naturaleza dinámica de gran ancho de banda de las aplicaciones actuales, la *Open Network Foundation* (ONF) es un consorcio sin ánimo de lucro enfocada en la transformación de la infraestructura de red, dedicada al desarrollo y estandarización de SDN. (ONF, 2020)

Como se establece en Ali et al. (2019), el objetivo de la arquitectura SDN es el control de tráfico de datos de manera centralizada, siendo su rasgo más característico la eliminación del modelo de integración vertical, dando a lugar una separación del plano de control y el plano de datos del núcleo de la red. Todas las tareas de control se realizan por medio de un controlador programable centralizado. El controlador y su software también se definen como el Sistema Operativo de Red (NOS).

Red Definida por Software es una arquitectura en la que se establece la separación física del plano de control del plano de datos; es una tecnología que hace posible que una red sea flexible y fácil de administrar, controlándola de manera central e inteligente usando aplicaciones de software. Mediante el uso de una aplicación los administradores de IT pueden configurar o administrar una red sin la necesidad de cambiar la configuración física del equipo, es decir se pueden hacer cambios sin la necesidad de enchufar o desenchufar continuamente cables y dispositivos. SDN es interoperable, lo que significa que debería funcionar con cualquier tipo de enrutador, conmutador sin importar la marca o fabricante, posee un controlador el mismo que es el elemento central de esta arquitectura además debe tener un API (*Application Programming Interface*) que es la interface que permite la transmisión de la información entre el controlador y los dispositivos de red individuales sean estos conmutadores, enrutadores firewalls entre otros.

Dándole un enfoque más simple a la idea fundamente detrás de SDN como se aprecia en la figura 3, el plano de control determina la ruta que los paquetes deben seguir a través de la red (mediante un protocolo de enrutamiento como BGP, OSPF o RIP) manteniendo una tabla de enrutamiento que incluye cualquier información auxiliar necesaria para seleccionar la mejor ruta en un momento dado, mientras que el plano de datos mantiene una tabla de reenvío que está optimizada para el procesamiento rápido de paquetes el cual se encarga de tomar decisiones de reenvío en cada conmutador paquete por paquete.

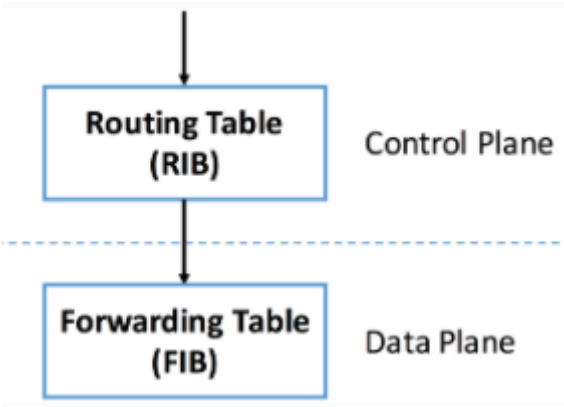


Figura 3. Plano de control (Tabla de enrutamiento RIB), plano de Datos (Tabla de reenvío FIB)
Fuente: (ONF, 2020)

Es necesario recalcar que, en principio, la desagregación significa que un operador de red podría comprar su plano de control a un proveedor X y su plano de datos al proveedor Y; una consecuencia natural de la desagregación es que los componentes del plano de datos (es decir, los conmutadores) se convierten en dispositivos de reenvío de paquetes básicos, comúnmente denominados conmutadores de caja blanca, en la que toda la inteligencia esta implementada en software y ejecutándose en el plano de control; la primera interface que soportaba la desagregación fue llamada OpenFlow. (Peterson Larry, Cascone Carmelo, O'Connor Brian, 2020)

1.2.1 Características de las redes definidas por software.

Como se indica en Ali et al (2019), una arquitectura SDN cumple con las siguientes características:

- Directamente Programable: el control de la red se puede programar directamente porque está desacoplado de las funciones de reenvío.
- Ágil: la abstracción del control del reenvío permite a los administradores ajustar dinámicamente el flujo de tráfico en toda la red para satisfacer las necesidades cambiantes.
- Gestión Centralizada: la inteligencia de la red está (lógicamente) centralizada en controladores SDN basados en software que mantienen una visión global de la red, que para las aplicaciones y los motores de políticas aparece como un único conmutador lógico.

- Configurado programáticamente: permite a los administradores de red configurar, administrar, proteger y optimizar los recursos de red muy rápidamente a través de programas SDN dinámicos y automatizados, que pueden escribir ellos mismos porque los programas no dependen de software propietario.

- Abierto y basado en estándares de vendor neutral: cuando se implementa a través de estándares abiertos, SDN simplifica el diseño y la operación de la red porque los controladores proporcionan las instrucciones en lugar de múltiples dispositivos y protocolos específicos del proveedor.

1.2.2 Arquitectura SDN.

Se basa en la división del plano de control y el plano de reenvío de datos la cual se denomina desagregación ya que estas dos se puede obtener por separado en lugar de implementarlas juntas como un sistema integrado, como se puede apreciar en la figura 4. Este tipo de arquitectura permite a las aplicaciones obtener más información sobre el estado de toda la red desde el controlador (o Controller), el cual llevará la gestión de la red, a diferencia de las redes tradicionales donde la red solo reconoce la aplicación. (Universidad Publica de Navarra, 2015)

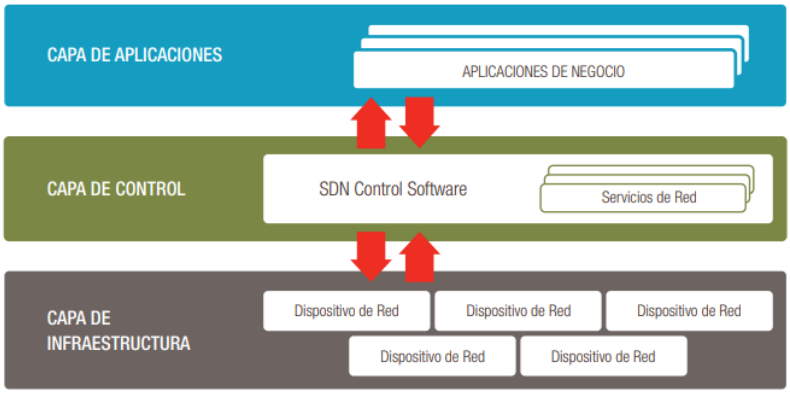


Figura 4. Arquitectura SDN de tres capas
Fuente: (Spera, 2013)

1.2.2.1 Capa de Aplicaciones.

Las aplicaciones SDN son programas que comunican comportamientos y recursos necesarios con el controlador SDN a través de interfaces de programación de aplicaciones (API). Además, las aplicaciones pueden crear una vista abstracta de la red recopilando información del controlador para la toma de decisiones. Estas aplicaciones podrían incluir

administración de redes, análisis o aplicaciones comerciales que se utilizan en grandes centros de datos. (Studios, 2015)

1.2.2.2 Capa de Control.

Como se menciona en Spera (2013), el controlador SDN es una entidad lógica que recibe instrucciones o requisitos de la capa de aplicación SDN y los transmite a los componentes de red. El controlador también extrae información sobre la red de los dispositivos de hardware y se comunica con las aplicaciones SDN con una vista abstracta de la red, que incluye estadísticas y eventos sobre lo que está sucediendo.

1.2.2.3 Capa de Infraestructura.

Los dispositivos de red SDN controlan las capacidades de reenvío y procesamiento de datos para la red. Esto incluye el reenvío y procesamiento de la ruta de datos. (Spera, 2013)

Las API de la arquitectura SDN a menudo se denominan interfaces hacia el norte y hacia el sur, que definen la comunicación entre las aplicaciones, los controladores y los sistemas de red. Una interfaz hacia el norte se define como la conexión entre el controlador y las aplicaciones, mientras que la interfaz hacia el sur es la conexión entre el controlador y el hardware de red físico. Como SDN es una superposición de red virtual, estos elementos no tienen que estar ubicados físicamente en el mismo lugar. (Studios, 2015)

1.2.2.4 Redes virtuales.

De acuerdo con Peterson Larry, Cascone Carmelo, O'Connor Brian, (2020), las redes virtuales, incluidas las redes privadas virtuales (VPN) y las redes de área local virtual (VLAN), han sido parte de Internet durante años. Históricamente, las VLAN han demostrado su utilidad dentro de las empresas, donde se utilizan para aislar diferentes grupos organizativos, como departamentos o laboratorios, dando a cada uno de ellos la apariencia de tener su propia LAN privada.

En el contexto de SDN, la idea es hacer que las redes virtuales similares a VLAN sean fáciles de usar, para que puedan configurarse, administrarse y destruirse mediante programación (es decir, sin que un administrador de sistemas tenga que configurar manualmente las etiquetas de VLAN en conmutadores de red). Al hacerlo, las redes virtuales se vuelven comunes, proporcionando una manera de aislar de forma segura todo

tipo de aplicaciones y actividades computacionales, no solo para separar grupos organizacionales. Además, dado que los entornos informáticos actuales se basan principalmente en máquinas virtuales (VM), estas redes virtuales conectan máquinas virtuales y no solo servidores físicos. Por tanto, tiene sentido que los sistemas de gestión de redes virtuales más utilizados estén estrechamente acoplados con los sistemas de gestión de VM.

Este uso de redes virtuales es bastante similar a lo que ocurre en los centros de datos en la nube, donde es importante aislar el tráfico de red de diferentes inquilinos de la nube, sin embargo, el número actual de VLANS disponibles para este propósito no son suficientes (1 - 4096) para soportar todas las redes privadas que se pueden tener ya sea en la nube o en una empresa. Para solucionar este problema se introdujo el estándar de redes de área local virtual extendida (VXLAN).

1.2.2.5 Spine-Leaf.

Como se señala en Eclassvirtual (2020), una estructura de conmutación de centro de datos es una red diseñada comúnmente con una topología llamada leaf-spine. La idea básica se ilustra en la figura 5. Cada rack tiene un conmutador Top-of-Rack (Torr) que interconecta los servidores en ese rack; se los conoce como conmutador leaf que forman parte de la fábrica. (Se suele tener dos conmutadores ToR de este tipo por bastidor por motivos de redundancia, pero la figura muestra solo uno por motivos de simplicidad). Cada conmutador Leaf se conecta a un conmutador conocido como Spine cumpliendo los siguientes requisitos: 1. Debe existir múltiples rutas o conexiones entre conmutadores en cada rack, y 2. Que cada ruta de rack a rack sea de dos saltos (es decir que se conectaran cada Leaf de cada rack por medio de un Spine).

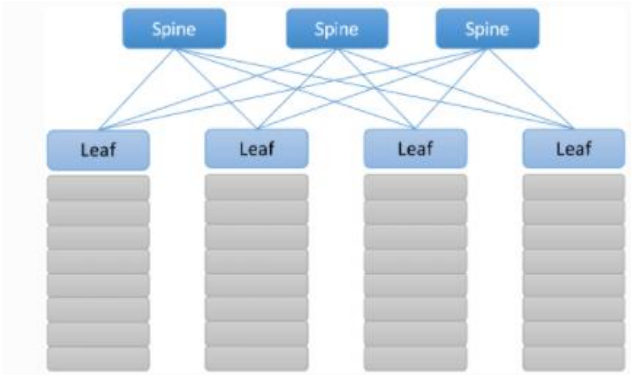


Figura 5. Ejemplo de una estructura de conmutación leaf-spine común en los centros de datos. Fuente: (Cisco Systems, 2018)

Como se aprecia en la figura 5 en la capa superior se tiene conmutadores Spine que son para interconexión y la capa de abajo tiene conmutadores Leaf que son para acceso es decir conexión con equipos finales tales como servidores, firewalls, routers, balanceadores entre otros. La arquitectura Spine-Leaf es un modelo de 2 niveles (tiers) en la que todos los dispositivos tienen el mismo número de segmentos y contienen una latencia constante y predecible en la difusión de información, la red utiliza enrutamiento de capa 3 y todas las rutas se configuran en un estado activo mediante el uso de rutas múltiples de igual costo (ECMP). Esto permite que todas las conexiones se utilicen simultáneamente mientras se mantiene estabilidad y evita lasos dentro de la red.

1.2.2.6 Redes de área amplia.

Aplicación de ingeniería de tráfico en enlaces de área amplia entre centros de datos. La red de área extendida definida por software (SD-WAN) simplifica las redes de las sucursales y optimiza el rendimiento de las aplicaciones en Internet y la WAN híbrida, provisionando un enrutamiento avanzado y optimización de la WAN.

1.2.2.7 Redes de acceso.

Las redes de acceso que implementan la última milla que conecta hogares, empresas y dispositivos móviles a Internet son otra oportunidad para aplicar los principios de SDN. Las tecnologías de red de acceso de ejemplo incluyen las redes ópticas pasivas (PON), conocidas coloquialmente como FTTH (fibra hasta el hogar), y la red de acceso de radio (RAN) en el corazón de la red celular 4G / 5G.

1.2.2.8 Telemetría de red.

Telemetría de red en banda (INT): La idea de INT es programar la canalización de reenvío para recopilar el estado de la red a medida que se procesan los paquetes (es decir, "en banda"). Esto contrasta con la supervisión convencional realizada por el plano de control leyendo varios contadores fijos (por ejemplo, paquetes recibidos / transmitidos) o muestreando subconjuntos de paquetes (por ejemplo, sFlow).

En el enfoque INT, las “instrucciones” de telemetría se codifican en campos de encabezado de paquete y luego se procesan mediante conmutadores de red. Estas instrucciones le dicen a un dispositivo con capacidad INT qué estado recopilar y luego cómo escribir ese estado en el paquete a medida que transita por la red. Las fuentes de tráfico INT pueden añadir las instrucciones en paquetes de datos normales o en paquetes

de sondeo especiales. De manera similar, los sumideros de tráfico INT recuperan e informan los resultados recopilados de estas instrucciones, lo que permite que los sumideros de tráfico monitoreen el estado exacto del plano de datos que los paquetes observaron (experimentaron) mientras se reenvían.

La idea se ilustra en la Figura 6, en la cual se muestra un paquete de ejemplo que atraviesa una ruta desde el conmutador de origen S1 al conmutador de destino S5 a través del conmutador de tránsito S2. Los metadatos INT agregados por cada conmutador a lo largo de la ruta indican qué datos se recopilarán para el paquete y registran los datos correspondientes para cada conmutador.

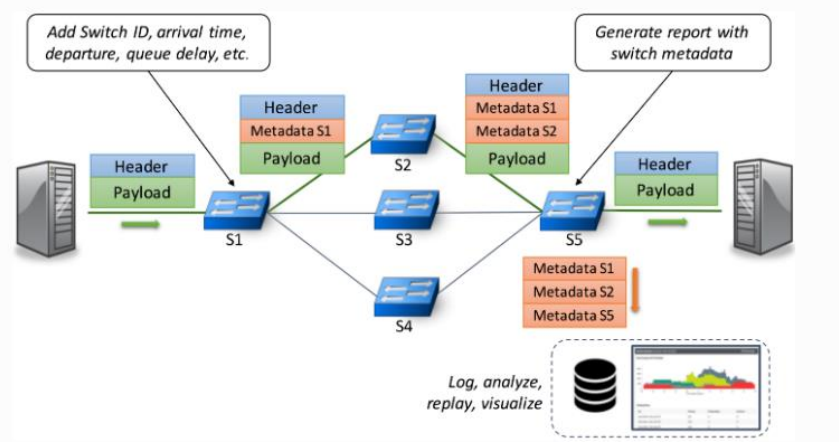


Figura 6. Ilustración de la telemetría de red en banda (INT), con cada paquete recolectando datos de medición a medida que atraviesa la red.
Fuente: (Peterson Larry, Cascone Carmelo, O'Connor Brian, 2020)

1.2.3 Interfaz de aplicación de programaciones (API).

En general, una API abstrae gran parte de los detalles operativos de bajo nivel que existen debajo de la API del software de capa superior.

Por ejemplo, el software de capa superior puede usar una llamada de función API como SetPortSpeed (24, 10), que se usa para configurar el puerto número 24 como un puerto Ethernet de 10 Gb. Cuando la API recibe esta llamada de función, debe escribir en varios registros del chip para que el puerto 24 esté habilitado como un puerto de 10 GbE. Usar una llamada a la función API es más fácil que escribir registros directamente y también hace que el software de capa superior sea mucho más fácil de crear y comprender.

Los dispositivos y la mayoría de los sistemas operativos de red se escriben utilizando estas funciones API llamadas en lugar de acceder a los registros directamente. Este mismo

concepto también se extiende al de capas superiores en stack de software. Por ejemplo, una aplicación de capa superior puede requerir para configurar una cierta cantidad de ancho de banda desde el punto A al B en el centro de datos realizando una llamada de función al sistema operativo de la red como SetBandwidthPath (A, B, 10) que a su vez puede generar otras llamadas de función API en capas API inferiores como setPortSpeed (24, 10) en un chip del conmutador dado. En estas situaciones, una API en dirección norte generalmente se refiere a una API que recibe comandos de una aplicación de capa superior y una API en dirección sur generalmente se refiere a una API que envía llamadas de función al software de capa inferior. Una API abierta se refiere a un conjunto de llamadas de funciones de API comunes acordadas por múltiples proveedores de hardware y software en la industria para asegurar la interoperabilidad entre capas de software.

1.2.4 VXLAN.

VXLAN inicialmente fue propuesto por Cisco y VMware, pero en la actualidad fue adoptada por varias empresas de tecnología. Es una técnica para hacer túneles virtuales en redes de capa 2 a través de redes físicas de capa 3 en grandes centros de datos que necesitan mantener varios inquilinos; para tunelizar estos paquetes, se encapsulan utilizando etiquetas VXLAN especiales.

La red de área local extensible virtual (VXLAN) es una de las tecnologías de virtualización de red sobre la capa 3, es una extensión de la red de área local virtual (VLAN); VXLAN encapsula una trama Ethernet de capa 2 en un paquete UDP y transmite el paquete a través de una red de capa 3 y que, en teoría, puede crear hasta 16 millones de VXLAN en un dominio (a diferencia de 4094 para la VLAN tradicional).

Como se muestra en la Figura 7, VXLAN es esencialmente una tecnología de tunelización. Establece un túnel lógico en la red IP entre los dispositivos de red de origen y destino para encapsular los paquetes del lado del usuario y reenviarlos a través del túnel, los servidores están conectados a diferentes puertos de dispositivos de red en la red VXLAN del centro de datos, que puede considerarse como un conmutador virtual de Capa 2.

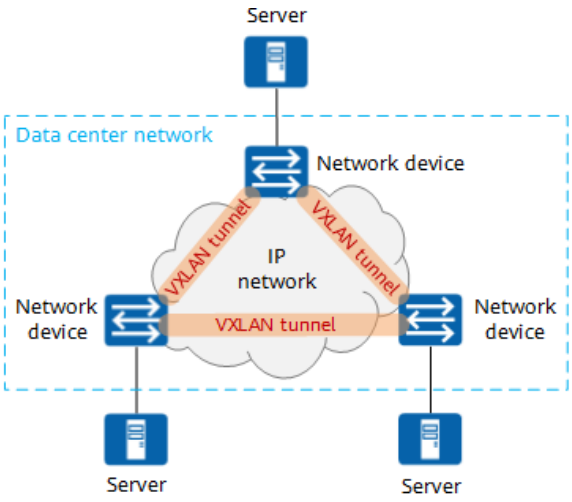


Figura 7. Ilustración de funcionamiento de VXLAN
Fuente: (Huawei, 2019)

Los servidores dentro de un Centro de Datos suelen tener tramas TCP/IP, que están encapsuladas con información de encabezado de capa de red, en muchos casos, este encabezado también contiene una etiqueta de VLAN, que el huésped puede usar para identificar diferentes departamentos dentro de su organización.

Desde el punto de vista del servidor o VM, es enviar y recibir este tipo de tramas a través de una red de capa 2. El detalle es cómo admitir múltiples inquilinos como este en una gran red del centro de datos de capa 3. Aquí es donde entran en juego los protocolos de tunelización como VXLAN. El protocolo VXLAN proporciona estas características encapsulando estos marcos con una etiqueta VXLAN junto con un encabezado de protocolo de datagramas de usuario (UDP) como se muestra en la figura 8 donde la trama original que contiene direcciones MAC y una etiqueta VLAN se encapsulan mediante un encabezado VXLAN y un encabezado UDP. Toda esta combinación luego se enruta a través de la red de capa 3 utilizando direcciones IP; utilizando únicamente etiquetas VXLAN para cada inquilino, cada inquilino puede usar su propio grupo de direcciones MAC mientras que el protocolo VXLAN proporciona aislamiento entre estos diferentes dominios de red de capa 2.

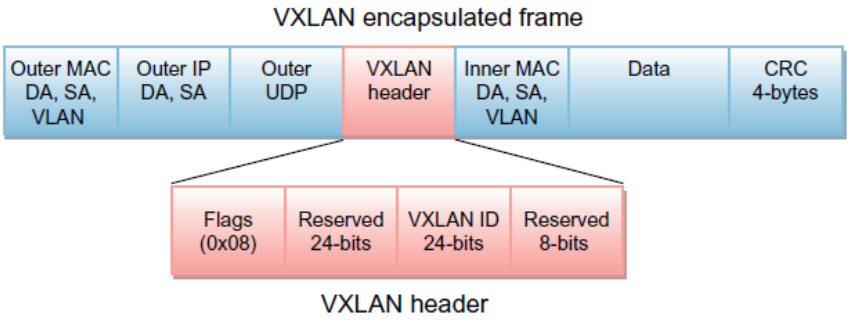


Figura 8. Trama Encapsulada de VXLAN
Fuente: (Lee, 2014)

1.2.5 VTEP.

VTEP es un dispositivo de borde en una red VXLAN. Es el punto inicial o final de un túnel VXLAN, que encapsula y des encapsula las tramas de datos como se aprecia en la siguiente figura 9, donde un VTEP de origen encapsula las tramas de datos originales enviadas por el servidor de origen en paquetes VXLAN y transmite los paquetes VXLAN al VTEP de destino en la red IP, mientras que el VTEP de destino des encapsula los paquetes VXLAN de las tramas de datos originales y las envía al servidor de destino.

Un dispositivo VTEP también descubre los VTEP remotos para sus segmentos VXLAN y aprende asignaciones de direcciones MAC remotas a VTEP a través de su interfaz IP. Los componentes funcionales de los VTEP y la topología lógica que se crea para la conectividad de Capa 2 a través de la red IP de transporte. (HUAWEI TECHNOLOGIES CO., 2020)

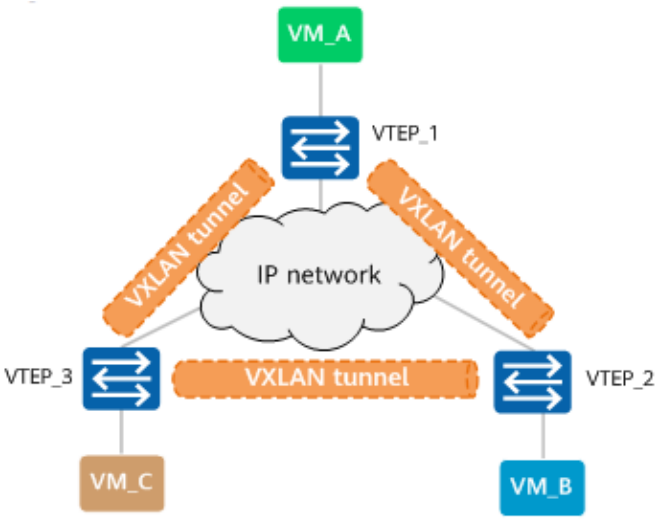


Figura 9. VXLAN Tunnel Endpoint (VTEP)
Fuente: (Huawei, 2019)

Los segmentos de VXLAN son independientes de la topología de red subyacente; a la inversa, la red IP subyacente entre los VTEP es independiente de la superposición de VXLAN, además enruta los paquetes encapsulados según el encabezado de la dirección IP externa, que tiene el VTEP de inicio como dirección IP de origen y el VTEP de terminación como dirección IP de destino. (Firas Ahmed, 2020)

1.2.6 VNI.

Un identificador de red virtual (VNI) es un valor que señala una red virtual específica en el plano de datos. Consta de 24 bits del encabezado VXLAN, por lo que puede admitir hasta 16 millones de segmentos de red individuales. (Los valores de VNI válidos son de 4096 a 16.777.215). Hay dos ámbitos de VNI principales:

VNI de ámbito de toda la red: el mismo valor es utilizado para identificar una red virtual de Capa 3 en todos los dispositivos de borde de la red. Este alcance de red es útil en entornos como centro de datos donde las redes pueden ser aprovisionadas automáticamente por sistemas de orquestación central.

VNI asignados localmente: el identificador tiene un significado local para el dispositivo de borde de red que anuncia la ruta; en este caso, el impacto de la escala de la red virtual se determina por nodo frente a una red. (Firas Ahmed, 2020)

1.3 Soluciones SDN para Centros de Datos

En esta sección se analiza las tecnologías SDN para centros de datos y la sistemática que estos utilizarán para tener una gestión centralizada y validación del estado de la red en tiempo real.

1.3.1 Cisco ACI.

Como se indica en GBM (2019), Cisco tiene su sede en California, EE. UU., con un estimado de 45.000 clientes empresariales en este mercado, su solución infraestructura centrada en aplicaciones (ACI) es una arquitectura integral con automatización centralizada y perfiles de aplicaciones basados en políticas. ACI ofrece flexibilidad de software con la escalabilidad del rendimiento de hardware.

Cisco ACI (Infraestructura Centrada en Aplicaciones) ofrece un modelo operativo de transformación que posibilita las aplicaciones en la nube y los Data Centers de última generación.

1.3.1.1 Las características más significativas.

- **Cisco APIC:** Es el único punto de gestión de toda la Red (física y virtual).
- **Escalabilidad:** Los nuevos nodos se añaden automáticamente al Fabric.
- **Seguridad:** Todo el Fabric actúa como un Router + Firewall.
- **Gestión sencilla:** Actualizaciones, identificación avanzada de problemas, diagnóstico

1.3.1.2 Constitución de la plataforma.

- Switches de la serie Nexus 9000, hardware de red físico basado en arquitectura de columna, se tiene Nexus 9500 tipo chasis y Nexus 9300 fijos.
- Gestión centralizada de políticas y Cisco Application Policy Infrastructure Controller (APIC) utilizado para la gestión centralizada de la fabric de ACI, constituido por 3 o más controladores para proveer alta disponibilidad.
- Un switch virtual de aplicaciones (Application Virtual Switch o AVS) para el perímetro de la red virtual.
- Innovaciones de software y hardware.
- Infraestructura física y virtual integradas.
- Un ecosistema abierto de proveedores de red, almacenamiento, gestión y organización.

El futuro de las redes con ACI tiene que ver con proporcionar una red que se pueda implementar, supervisar y gestionar de una forma que admita DevOps y rápidos cambios de aplicaciones. ACI lo hace mediante la reducción de la complejidad y un marco común de políticas que permite automatizar el aprovisionamiento y la gestión de recursos.

El diseño de ACI de Cisco está constituido en conformidad con Open Source y Open APIs. Dentro de ACI todo se representa como un objeto, y cada objeto se puede manipular

utilizando REST API (POST, GET, UPDATE, DELETE) y Python, permitiendo así varios tipos de orquestación, automatización y disposición de recursos de networking a otras plataformas tales como OpenStack, Puppet, Ansible, etc.

1.3.1.3 Consideraciones.

- Cisco tiene productos sólidos y una base instalada grande y global. La cartera del proveedor ofrece una gran variedad de funciones y cubre casi todos los escenarios de uso, incluido el enrutamiento avanzado y la conmutación de latencia ultra baja.

- Cisco ofrece niveles crecientes de análisis y automatización se alinea con los requisitos emergentes de los clientes para ofrecer una red más autónoma y autorreparable.

- Network Insights de Cisco mejora las actividades operativas del segundo día, como la resolución de problemas, la generación de informes y la eliminación de errores, y se integra con la infraestructura centrada en aplicaciones (ACI) y los controladores DCNM. (Lerner, 2020)

1.3.2 Huawei DCN.

Huawei tiene su sede en Shenzhen, China, con más de 7.000 clientes empresariales en este mercado, su solución de SDN de Huawei para Centros de Datos está basada en los principios de intelectualidad, simplicidad, banda ultra ancha, apertura y seguridad optimizada gracias al controlador Agile Controller. Esta solución también es utilizada por el analizador de redes de DPC FabricInsight para la implementación del análisis predictivo y la detección automática de desviaciones. Posteriormente se convertirá en la base para la creación de un sistema automático y autorregulado de circuito cerrado que resultará aún más conveniente para el usuario.

1.3.2.1 Características.

- **Flexibilidad:** alta escalabilidad de servicios y canales en la nube para transferencia de Big Data.

- **Simplificación de procesos:** la solución en la nube para la red de Huawei acelera 10 veces la activación de los servicios en la nube.

- **Apertura:** simplificación de la computación en la nube debido a la conexión perfecta con las principales plataformas en la nube.

1.3.2.2 Constitución de la plataforma.

Automatización de IDN

Como se indica en TADVISER (2018), Agile Controller permite completar automáticamente la planificación de la red teniendo en cuenta los propósitos del servicio y las intenciones del usuario que aumenta la calidad del servicio.

Un clúster puede administrar 3 mil nodos en el nivel de acceso y se puede escalar hasta 100 mil servidores. Además, el grupo de controladores que trabaja en el modo activo y en espera en proporción de uno a uno aumenta la disponibilidad del sistema y garantiza la ausencia de interrupciones en el servicio en la conmutación automática entre los nodos activos y de reserva.

Análisis predictivo e identificación de desviaciones

El analizador FabricInsight establece el modelo de comportamiento de la red sobre la base del análisis de Big Data y la tecnología de aprendizaje automático, controla continuamente el estado de la red para respaldar el mantenimiento de diagnóstico y la identificación de desviaciones de la red en segundos de lectura y también conecta automáticamente las desviaciones de la red (por ejemplo, sobrecargas) con los servicios marcados que ayudan a los clientes a revelar y corregir errores rápidamente.

Utilizada una plataforma con capacidad de 400 Gbps para un núcleo de conmutación con las mejores propiedades de utilización de la industria. El conmutador CE12800 admite 36 tarjetas de interfaz de red de 400 gigabits e interacción de extremo a extremo de 400 gigabits.

La solución Huawei Intent-Driven Networking para CloudFabric, que es la piedra angular del concepto de IDN, abre el acceso a soluciones y productos avanzados para proveedores de servicios, grandes corporaciones y medianas empresas, lo que permite revelar las intenciones comerciales y los propósitos de los clientes y lograr el no-touch y despliegue intelectual y funcionamiento de la red.

Huawei Cloud Fabric proporciona productos de red de varias capas para proporcionar servicios de ventanilla única para los usuarios y para simplificar la construcción de la red del centro de datos del cliente. Aprovechando los conmutadores centrales insignia de la serie Huawei CE12800, que tienen las especificaciones más altas del mundo, y las series CE7800 / CE6800 / CE5800 de alto rendimiento para conmutadores Top of de rack (TOR). (TADVISER, 2018)

1.3.2.3 Consideraciones.

- El precio promedio del proveedor es el más bajo comparado con Cisco, Aruba, Dell, H3C, Arista.

- El proveedor tiene una sólida cartera de conmutadores y gestión asociada para satisfacer las necesidades de la mayoría de los clientes en este mercado. En particular, la red de centro de datos inteligente y sin pérdidas AI Fabric de Huawei se adapta bien a las necesidades de la informática de alto rendimiento, ya que requiere cero pérdidas de paquetes y redes de baja latencia que admitan RoCE, como AI / ML y memoria no volátil expresada a través de fabric (NVME-oF) cargas de trabajo. (Lerner, 2020)

1.3.3 Arista Software Defined Cloud Networking SDCN.

Como se establece en Arista (2016), Arista tiene su sede en California, Estados Unidos, con más de 4.000 clientes empresariales, su solución Software Defined Cloud Networking está constituido por una red de 2 capas basadas en conmutadores Leaf/spine, donde la administración puede ser inband o outband y principalmente está basada en un esquema ZTP (Zero touch provisioning) la cual permite tener un despliega automático.

1.3.3.1 Características.

- **Sin protocolos propietarios:** Arista cree en estándares abiertos, no se requieren protocolos propietarios y bloqueos de proveedores para construir redes de escalamiento horizontal muy grande.

- **Menos niveles es mejor que más niveles:** Los diseños con menos niveles (por ejemplo, un diseño de leaf/spine de 2 niveles en lugar de 3 niveles) reducen el costo, la complejidad, el cableado y la energía/calor. Los diseños de red Spines de

un solo nivel no utilizan ningún puerto para interconectar niveles de conmutadores, por lo que ofrecen el menor costo por puerto utilizable.

- **Los diseños deben ser ágiles y permitir flexibilidad en las velocidades de los puertos:** El punto de inflexión cuando la mayoría de los servidores/nodos de cómputo se conectan a 1000Mb a 10G. Esto, a su vez, impulsa el requisito de que los enlaces ascendentes de la red migren de 10G a 40G y a 100G. Los interruptores de Arista y los diseños de referencia permiten esa flexibilidad.

- **Funciones y sistema operativo consistentes:** todos los conmutadores Arista utilizan el mismo Arista EOS. No hay diferencia en la plataforma, los trenes de software o el sistema operativo. Es la misma imagen binaria en todos los conmutadores.

- **Interoperabilidad:** los conmutadores y diseños de Arista pueden interoperar con otros proveedores de redes sin un bloqueo de propiedad.

1.3.3.4 Constitución de la plataforma.

Utiliza un controlador llamado Central EOS, para controlar toda la red que está basada en VXLAN.

Los conmutadores Leaf se conectan a los conmutadores SPINE e intercambian información mediante el protocolo BGP.

Provee una rápida recuperación ante falla de enlaces o nodos, además provee una simplicidad para el cambio, control o mantenimiento de la red.

Provee descubrimiento automático de máquinas virtuales para las plataformas vmware y Open stack, así como auto aprovisionamiento de VLANS para vsphere y Open Stack.

Auto Binding: Provee una asignación automática de VLAN hacia VXLAN en los conmutadores, actuando como un túnel virtual de punto final.

Eliminación de IP multicast para los casos de inundación por tormenta de broadcast, tramas desconocidas y tráfico multicas.

Las tecnologías SDN emergentes complementan los conmutadores del centro de datos al automatizar las políticas de red y el aprovisionamiento dentro de un ecosistema de

infraestructura de nube integrado más amplio. Arista define la combinación de tecnologías SDN y el Sistema Operativo Extensible Arista (Arista EOS®) como Software Defined Cloud Networking (SDCN). (Arista, 2016)

CAPÍTULO 2 LEVANTAMIENTO DE LA INFORMACIÓN SOBRE LA INFRAESTRUCTURA ACTUAL

Una vez descrito los conceptos necesarios para el entendimiento general de la problemática se realizará el levantamiento de la información de la estructura actual del Centro de Datos; se realiza el levantamiento de información únicamente de la red del Centro de datos, sin tomar en cuenta equipos finales de clientes o servicios asociados a estos ya sean firewall, servidores, red SAN, entre otros, para lo cual se utiliza diagramas de referencia del departamento de Networking y al no tener documentación de las conexiones, equipamiento instalado se revisó en cada uno de los equipos para obtener la información necesaria, además por motivos de acuerdos de confidencialidad no se nombra al ISP donde se realizó este levantamiento de información.

2.1 Topología física y escenario de la red actual

2.1.1 Capa de Core.

La infraestructura de red del Centro de Datos consiste en la interconexión de varios equipos ya sean de red como router, switch, de seguridad como firewall, DDOS, servidores entre otros.

La conectividad a nivel LAN de los equipos de red en el Centro de Datos utilizan la arquitectura jerárquica de 3 capas de Cisco. Teniendo en el Core de la arquitectura switch Cisco Nexus 7000 los mismos que se encuentran conectados a switch Nexus 5000 en la capa de distribución y que a su vez a los switch Nexus 5000 se encuentran conectados como extender switch Nexus 2000 para la capa de acceso como se puede observar en la figura 10, donde se representa la conexión a nivel LAN de los equipos de red del Centro de Datos ubicado en la ciudad de Quito.

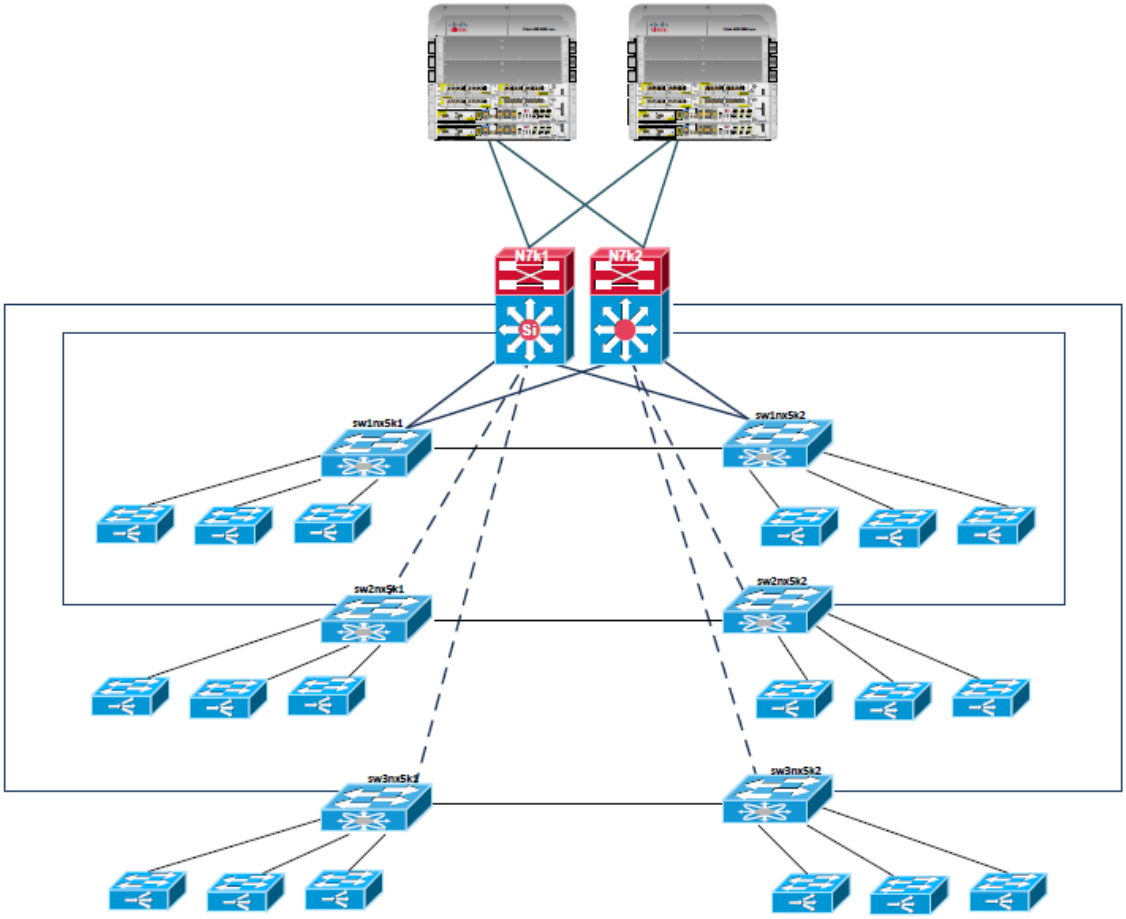


Figura 10. Diagrama de la Red LAN del Centro de Datos de la empresa

A nivel de capa de Core, se utilizan dos Routers Cisco Nexus C7010, interconectados entre sí mediante un virtual port channel (VPC) domain (peer-link), proporcionando redundancia en esta capa; ambos router se encuentran activos, el par VPC secundario deshabilita todas las conexiones VPC conectados a este y el tráfico fluye por el VPC primario como se observa en la figura 11.

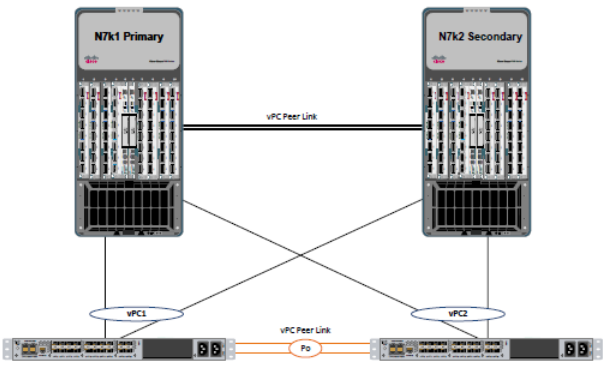


Figura 11. VPC Nexus 7000 y Nexus 5000

Al ser el cisco Nexus 7010 un equipo tipo chasis posee en sus ranuras módulos con interfaces de 10G por lo que sus enlaces de subida, así como los enlaces a los switch de agregación se realizó por medio de enlaces agregados VPC de 20Gbps, con esto se proporciona redundancia a nivel de conectividad.

A nivel de enrutamiento se utilizan VRF (Virtual Routing and Forwarding) la cual es una tecnología que permite tener múltiples routers virtuales en un solo router físico. Cada uno de estos routers virtuales tendrá su propia tabla de enrutamiento, proporcionando las rutas de cada red totalmente independiente y separada una de otra.

Para cada cliente se le asigna una VLAN, en los enlaces troncales se limitarán las VLANS para controlar que cliente pasan por cada puerto para limitar el tráfico unicast flooding, además de seguridad, adicional a esto se maneja un protocolo de redundancia que es HSRP, donde el Nexus 7010 primario se le asignará una mejor prioridad para que sea el router que actúe como activo y el segundo Nexus 7010 estará como standby con una prioridad menor que el primero.

Adicional se tiene en la capa de Core una conexión L2 punto a punto por medio de fibra óptica para la interconexión con el Centro de Datos de la ciudad de Guayaquil, esta conexión es usada para contingencia, aplicaciones de replicación y entre otras.

El protocolo que se está utilizando para evitar lazos de capa 2 es MST en el cual el root se encuentra ubicado en Guayaquil y la comunicación es por el enlace de capa 2 indicado.

2.1.2 Capa de distribución y acceso.

La capa de distribución está constituida por 6 switch cisco Nexus 5548 los mismos que están conectadores en pares entre sí y estos a su vez tienen conexiones a cada uno de los Nexus 7000, además a cada Nexus 5000 se conectan switch Cisco Nexus 2248TP o Nexus 2224TP como extender por medio de una VPC por redundancia; en estas dos capas no se realiza una gran distinción ya que los dispositivos finales se conectan a los switch Nexus 5000 cuando la capacidad requerida supera el 1G y por lo tanto se requiere conexiones en 10G, mientras que para capacidades menores se opta por conexiones a los switch extender Nexus 2000, de igual manera que en la capa de Core existen conexiones redundantes, es decir hacia los switch Nexus C7010, para brindar alta disponibilidad y minimizar puntos de falla.

En estas capas la interconexión es por medio de enlaces troncales (trunk) o de acceso, es decir solo son utilizados como enlaces de capa 2 donde no se realizará ningún tipo de enrutamiento.

Los puertos donde se conectan clientes o están destinados se ha optado por bloquear tramas unicast, multicast, para evitar flooding en la red, además también se ha desactivado el protocolo de descubrimiento de Cisco (CDP), a nivel de spanning-tree se usar Guard root el cual designa que un puerto solo puede enviar o reenviar paquetes de tipo BPDU, pero no puede recibir dichos paquetes, con lo que se previene que un puerto sea root ya que en un puerto de cliente no tiene sentido que lleguen este tipo de paquetes.

2.2 Características de los equipos de la red actual

2.2.1 Cisco Nexus 7010 Switch.

Es un equipo de gran tamaño y capacidad es de tipo chasis por lo cual tiene capacidad de ampliación como se puede observar en la figura 12, consta de 10 ranuras para ser integradas con diferentes tipos de tarjetas de línea.

Figura 12. Cisco Nexus 7010
Fuente: (Cisco, 2013d)

En la tabla 1 se muestran las diferentes especificaciones del equipo como por ejemplo la parte física como unidades de rack necesarias para la instalación, numero de ranuras, fuentes, dimensiones, capacidad entre otros.

2.2.2 Especificaciones.

Tabla 1. Características Nexus 7010

Item	Specification
Cisco Nexus 7000 10-Slot Switch	
Product compatibility	Supports all Cisco Nexus 7000 Series Supervisor and I/O modules Supports Fabric1 and Fabric2 modules
Max local Switching capacity	600 Gbps
Max inter-slot switching capacity	550 Gbps
Software compatibility	Cisco NX-OS Software Release 4.0 or later
Reliability and availability	OIR of all redundant components: Supervisor and fabric modules, power supplies, and fan trays
Performance	5.76 bps (IPv4 unicast) in combination with supervisor and fabric modules.
MIBs	Supports SNMPv3, v2c, and v1 (see Cisco NX-OS Software release notes for details about specific MIB support)
Network management	Cisco DCNM 4.0 or later
Programming interfaces	XML,Scriptable CLI, Cisco DCNM 4.0 web services
Physical specifications	Usable rack space: 21RU 10-slot chassis: 2 dedicated supervisor modules and 8 I/O modules 5 fabric module slots 3 power supply slots Dimensions (H x W x D): 36.5 x 17.3 x 33.1 in.(92.7x43.9x84.1 cm) Chassis depth including cable management and chassis doors is 38 in. (96.5 cm) Unit is rack mountable in a standard 19-inch EIA rack Weight Chassis only: 200 lb. (90 kg) Fabric Module: 4 lb (1.8 kg) System Fan Tray: 20 lb (9.1 kg) Fabric Fan Tray: 5 lb (2.3 kg) Supports 6-kW and 7.5-kW AC and DC power supplies
Environmental specifications	Airflow direction: Bottom front of chassis to top back
	Operating temperature: 32 to 104°F (0 to 40°C)
	Operational relative humidity: 5 to 90%, noncondensing
	Operating altitude: -500 to 13,123 ft. (agency certified 0 to 6500 ft.)
	Seismic: Zone 4 per GR63
	Floor loading: 190 lb. per sq. ft.
	Operational vibration
	GR63, Section 5.4.2
	ETS 300 019-1-3, Class 3.1, Section 5.5
	Storage altitude: 1000 to 30,000 ft.
	Storage temperature: -40 to 158°F (-40 to 70°C)
	Storage relative humidity: 5 to 95%, noncondensing
	Heat dissipation: Maximum 12,000W per chassis (actual dissipation will be lower, depending on the chassis configuration)

Fuente: (Cisco, 2013d)

2.2.3 Modulo N7K-M108X2-12L.

La tarjeta de línea que se usa en los Nexus 7010 es un módulo de 8 puertos de 10G como se observa en la siguiente figura 13.

Figura 13. Cisco modulon7k-M-108X2
Fuente: (Cisco, 2013a)

2.2.2.1 Especificaciones.

- Soportado en todas las versiones de Nexus 7000
- Soportado para Fabric tipo 1 y 2
- Soportado con supervisoras tipo 1, 2 y 2E
- Conectividad de 8 puertos de 10 Gigabit Ethernet mediante xceivers de tipo X2
- Entradas MAC 128k
- Rutas IPv4 mayor a 1M
- Rutas IPv6 mayor a 350K
- ACL permitidas 128

2.3 Resumen de conexiones

En este apartado se detalla un resumen de los puertos usados en los equipos de la infraestructura actual, el mismo que se puede ver en las siguientes tablas.

Tabla 2. Resumen equipamiento físico Nexus 7000

Características Equipo Nexus7010	
Número de Equipos	2
Número de slot	10
Módulos de 10G	3
Tipo de Xceivers	X2
Supervisoras	2
Módulos Fabric	3
Fuentes	2
Ventiladores	4

Tabla 3. Resumen equipamiento físico Nexus 5000

Características Equipo Nexus 5000 2000	
Número de Equipos	6
Número de puertos	48
Puertos de 10G	32
Supervisoras	2
Módulos Fabric	3
Fuentes	2
Ventiladores	2
Fex por cada N5k	6

Tabla 4. Resumen puertos 10G Nexus 7010

Conexiones equipo Nexus 7010 #1	
Total, puertos de 10G	24
Puertos usados de 10G	17
Puerto 10G de Usuario Final	4
Conexiones equipo Nexus 7010 #2	
Total, puertos de 10G	24
Puertos 10G usados de 10G	17
Puerto de Usuario Final	4

Tabla 5. Resumen puertos 10G y 1G Nexus 5000

Conexiones equipo Nexus 5000 #1	
Total, puertos de 10G	32
Puertos usados de 10G	23
Puerto 10G de Usuario Final	11
Total, Puertos 1G	192
Puertos 1G usados	78
Conexiones equipo Nexus 5000 #2	
Total, puertos de 10G	32
Puertos usados de 10G	22
Puerto 10G de Usuario Final	11
Total, Puertos 1G	192
Puertos 1G usados	78
Conexiones equipo Nexus 5000 #3	
Total, puertos de 10G	32
Puertos usados de 10G	19
Puerto 10G de Usuario Final	7
Total, Puertos 1G	264
Puertos 1G usados	160
Conexiones equipo Nexus 5000 #4	
Total, puertos de 10G	32
Puertos usados de 10G	19
Puerto 10G de Usuario Final	7
Total, Puertos 1G	264
Puertos 1G usados	160
Conexiones equipo Nexus 5000 #5	
Total, puertos de 10G	32
Puertos usados de 10G	12
Puerto 10G de Usuario Final	4
Total, Puertos 1G	240
Puertos 1G usados	101
Conexiones equipo Nexus 5000 #6	

Total, puertos de 10G	32
Puertos usados de 10G	12
Puerto 10G de Usuario Final	4
Total, Puertos 1G	240
Puertos 1G usados	101

Como se observa mediante las tablas 4 y 5 se establece que el número total de conexiones en puertos de 10 Gbps utilizadas netamente para servicios de clientes, es decir sin tomar en cuenta las conexiones entre dispositivos de backbone es de 26 puertos y para conexiones de hasta 1 Gbps es de 339.

Cabe recalcar que las conexiones que actualmente llegan a los Nexus 7010 son por medio de fibra óptica, en algunos casos fibra óptica monomodo y en otras multimodo sin que existan una norma o estándar que defina el uso de la misma y a su vez se utilizan transceivers para la conversión óptica eléctrica en el equipo; a parte de las conexiones hacia los Nexus 5000 existe conexiones hacia otros equipos como firewall y equipos hacia el Core MPLS, a estos equipos no se conectan ningún servidor o cliente; por el contrario, en los equipos Nexus 5000 hay conexiones en los puertos de 10G hacia servidores, clientes y se lo realiza mediante cableado de fibra óptica; la conexión de clientes a los FEX (extensores) se lo realiza con enlaces de 1Gbps en cobre ya se utilizando un transceiver o conexión directa al equipo final del cliente (CPE) por medio de cable de cobre de par trenzado (UTP).

CAPÍTULO 3 DESCRIPCIÓN DE LA TECNOLOGÍA SDN PARA CENTROS DE DATOS

En este capítulo se establece los requisitos y características mínimos a considerar para la nueva arquitectura y en base a esto se realizó la selección de la tecnología más adecuada, además se analizó a detalle la tecnología escogida que permitirá tener una visión centralizada, administración y supervisión en tiempo real de la red.

La solución debe cumplir los siguientes requisitos mínimos:

- Capacidad de administrar una colección de conmutadores físicos y/o virtuales como una única construcción, con automatización integrada a la que se puede acceder mediante API.
- Integración fácil con productos de fabricantes como Open Source por ejemplo Red Hat, Docker, así como también con plataformas como vmware, Microsoft
- Conmutadores con al menos 48 puertos de hasta 25 Gbps o más que admita protocolos L2/L3.
- Capacidad de conexión e integración de Fabric Extender.
- Multi-tenancy, múltiples entornos compartan la misma red física.
- Seguridad, solo se debe proveer comunicación que sea autorizada de forma explícita.
- Gestión Centralizada que permita tener visibilidad de toda la red, errores, problemas, topología y capacidad de generación de informes.
- Debe proveer una solución de interconexión entre Centros de Datos ubicados en localidades distintas.
- Capacidades de programación basadas en modelos y de estándares abiertos.

3.1 Selección de la propuesta para la solución

Para la selección de la tecnología más adecuada se realizó una comparación de 3 empresas de tecnología grandes como Cisco, Huawei, Arista y en base a los requerimientos

mínimos planteados se realiza la comparación como se muestra en la tabla 6, además de esto también se basó en los estudios e información provista por Gartner.

Tabla 6. Comparación de las diferentes tecnologías

	Cisco	Huawei	Arista
Automatización y Fabric programable	X	X	X
Integración con Open Source por ejemplo Red Hat, Docker, vmware, Microsoft	X	X	X
Conmutadores con al menos 48 puertos de 25 Gbps o más que admita protocolos L2/L3.	X	X	X
Capacidad de conexión e integración de Fabric Extender	X	-	X
Multi-tenancy, múltiples entornos compartan la misma red física.	X	X	X
Seguridad, solo se debe proveer comunicación que sea autorizada de forma explícita.	X	-	X
Gestión Centralizada que permita tener visibilidad de toda la red, errores, problemas, topología y capacidad de generación de informes.	X	-	-
Debe proveer una solución de interconexión entre Centros de Datos ubicados en localidades distintas.	X	X	-
Capacidades de programación basadas en modelos y de estándares abiertos.	X	-	X
Menor costo de implementación	-	X	-

Fuente: Propia, (Services, 2019)

Automatización y Fabric programable, en la empresa se tiene un sistema automatizado el mismo que hasta el momento no ha podido ser integrado infraestructura de red tradicional, por lo cual se necesita que la nueva tecnología sea capaz de integrarse a las plataformas de automatización de la empresa.

Integración con Open Source por ejemplo Red Hat, Docker, vmware, Microsoft, si bien en la empresa se utiliza en gran proporción la plataforma vmware se tiene también gran apertura para el uso y despliegue de plataformas open Source.

Conmutadores con al menos 48 puertos de 25 Gbps o más que admita protocolos L2/L3, las necesidades actuales en el Centro de Datos implican tener puerto de capacidades superiores a 1Gbps, además de equipos con gran cantidad de puertos.

Capacidad de conexión e integración de Fabric Extender, si bien las necesidades actuales es tener capacidades mayores a 1 Gbps, existen conexiones en cobre que su función es solo para gestión o administración de equipos en el cual su capacidad no supera los 100 Mbps o incluso para servicios de Housing donde el cliente trae su propio equipo que muchas veces tiene interfaces ethernet de 100 Mbps o 1 Gbps, por lo cual no es óptimo usar puertos que su capacidad es de 10 Gbps para este tipo de conexiones.

Multi-tenancy, múltiples entornos compartan la misma red física, la plataforma debe soportar distintos modos de operación, como desarrollo, pruebas, preproducción, producción sobre la misma red sin que exista afectación a otros clientes.

Seguridad, solo se debe proveer comunicación que sea autorizada de forma explícita, se requiere que poder controlar los flujos de comunicaciones para minimizar ataques o poder aislar la comunicación en caso de que algún equipo sea comprometido.

Gestión Centralizada que permita tener visibilidad de toda la red, errores, problemas, topología y capacidad de generación de informes, se requiere poder tener una visibilidad completa de la red con la solución SDN.

Debe proveer una solución de interconexión entre Centros de Datos ubicados en localidades distintas, la solución debe tener distintos tipos de topologías para poder tener continuidad del negocio ante cualquier evento no esperado.

Capacidades de programación basadas en modelos y de estándares abiertos, la solución debe ofrecer la posibilidad de programación con Python ya que hay una tendencia en el desarrollo con este tipo de programación en la empresa.

Menor costo de implementación existe la suficiente holgura presupuestaria para elegir cualquiera de las tecnologías por lo cual está no es una de las razones de peso para tomar una decisión sobre cuál sería la mejor opción.

En base a los resultados la mejor solución es la tecnología presentada por Cisco, además de esto también se tiene una fuerte relación de confianza con dicha empresa lo que hace que sea una razón de peso para escogerla, adicional a esto los ingenieros a cargo de la administración, soporte están muy familiarizados con la tecnología de Cisco por lo cual tienen un mayor expertise para trabajar con los productos de esta marca.

Además, en la figura 14 se puede ver el cuadrante mágico de Gartner 2020 para centros de datos el mismo que es llamado Cuadrante Mágico para Redes de Centros de Datos y Nube, Gartner define a los proveedores de centros de datos y redes en la nube como proveedores de infraestructura de pila completa que proporcionan software o hardware que puede cumplir con los requisitos de los centros de datos empresariales y redes en la nube.

Como se indica en (Hein, 2020) el "Cuadrante mágico" enumera 11 proveedores de centros de datos y redes en la nube diferentes, y destaca sus productos de hardware y software, sus fortalezas y debilidades. Luego, Gartner evaluará la integridad de la visión y las capacidades de ejecución de cada proveedor. Según estas métricas, los proveedores se pueden clasificar en una de cuatro categorías: líderes, retadores, visionarios y nichos de mercado. Los 11 proveedores de centros de datos y redes en la nube en el Cuadrante Mágico de este año son Arista Networks, Cisco, Cumulus Networks, Dell EMC, Extreme, H3C, HPE (Aruba), Huawei, Juniper Networks, NVIDIA-Mellanox Technologies y VMware.

Líderes. Normalmente, gigantes innovadores que se destacan tanto en visión como en ejecución.

Challengers. Ejecución fuerte pero baja visión.

Visionarios. Buena visión, pero baja ejecución.

Jugadores de nicho. Centrado en un segmento pequeño, lo que resulta en baja visión y ejecución.

Figura 14. Magic Quadrant for Data Center and Cloud Networking
Fuente: (Lerner, Andrew Zeng, 2020)

Cisco se encuentra en los líderes como se puede apreciar en el Cuadrante Mágico para Redes de Centros de Datos y Nube, esto debido a sus productos como conmutadores Nexus, su solución ACI y de más herramientas para la gestión y administración de la red.

Su hoja de ruta prevé formas de aumentar el nivel de automatización y análisis para que los clientes se adapten a la demanda de una red más autónoma y autorreparable.

3.2 Infraestructura centrada en Aplicaciones (ACI)

La infraestructura centrada en aplicaciones (ACI) es en esencia una infraestructura de red basada en políticas y definida por software. Esta arquitectura simplifica, personaliza, optimiza y acelera todo el ciclo de vida de implementación mediante segmentación de lista blanca, uso de VXLAN, políticas centralizadas de configuración entre otras.

3.2.1 Componentes.

El hardware de la estructura de la infraestructura centrada en aplicaciones (ACI) incluye una infraestructura de políticas de aplicaciones de controlador (APIC) (constituido por un

grupo de tres controladores comúnmente), uno o más conmutadores de hoja conocidos como leaf (por ejemplo, N9K-C93180LC-EX or N9K-C93180YC-FX) y uno o más conmutadores de columna o Spine (por ejemplo, 9336PQ, 9504, 9508 o 9516 conmutadores) que se conectarán a cada leaf. En la figura 15 se muestra cómo estos los componentes se interconectan para formar la estructura ACI.

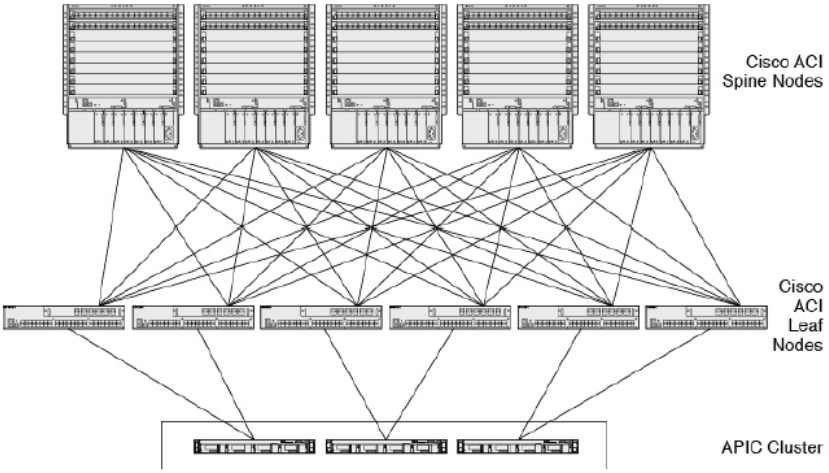


Figura 15. Interconexión Fabric ACI
Fuente: (Cisco, 2018)

3.2.1.1 Nexus 9300.

Como se indica en (Cisco, 2009) la serie 9300 ofrece una variedad de opciones de interfaz para migrar de manera transparente los centros de datos desde velocidades de 100 Mbps, 1 Gbps y 10 Gbps hasta 25 Gbps para servidores, y velocidad desde 10, 40 Gbps a 50 y 100 Gbps en la capa de agregación estos conmutadores son los dispositivos responsables de la mayor parte de la funcionalidad de la red: conmutación L2/L3 a velocidad de línea, compatible con operaciones VTEP para VXLAN, enrutamiento IGP protocolos como BGP, OSPF, EIGRP, multicast, gateways anycast, en la tabla 7 y 8 se muestra un ejemplo de la familia de conmutadores Nexus usados para spine y leaf respectivamente.

Tabla 7. Conmutadores Nexus 9000 utilizados como Spines

Nexus 9300 40/100 GE Switches				
	Nexus 9364C	Nexus 9336C-FX2	Nexus 9332C	Nexus 9364C-GX
Usage	ACI spine & NXOS(Leaf/spine), powered by Cloud Scale	ACI leaf & NXOS(Leaf/spine), powered by Cloud Scale	ACI spine & NXOS(Leaf/spine), powered by Cloud Scale	ACI (leaf/spine*) & NXOS(Leaf/spine), powered by Cloud Scale
Form factor	2 RU	1 RU	1 RU	2 RU
100 M/1 GE/10 GE copper ports	NA	NA	NA	NA
1/10 GE fiber ports	2	NA	2	0
40 GE ports	64	36	32	64
50 GE ports	NA	72	NA	128
100 GE ports	64	36	32	64
Latency (microseconds)	~1.3	~1	~1.3	~1.6
MACsec & Cloudsec	Yes on last 16 ports	Yes, all ports	Yes, on last 8 ports	NA

Fuente: (Cisco, 2018)

Tabla 8. Conmutadores Nexus 9000 utilizados como Spines

Nexus 9300 1/10/25GE Fiber Switches					
	93180YC-EX	93180YC-FX	93180YC-FX3S	93240YC-FX2	93360YC-FX2
Usage	ACI leaf, top of rack, FEX aggregation, powered by Cloud Scale	ACI leaf, top of rack, FEX aggregation, powered by Cloud Scale	ACI leaf, top of rack, FEX aggregation, powered by Cloud Scale	ACI leaf, top of rack, FEX aggregation, powered by Cloud Scale	ACI leaf, top of rack, FEX aggregation, powered by Cloud Scale
Form factor	1 RU	1 RU	1 RU	1.2	2 RU
Throughput (Tbps)	3.6	3.6	3.6	4.8	7.2
1/10 GE fiber ports	48	48	48	48	96
25 GE ports	48	48	48	48	96
40 GE ports	6	6	6	12	12
100 GE ports	6	6	6	12	12
Latency (microseconds)	~1	~1	~1	~1	~1
Buffer (MB)	40	40	40	40	40
Minimum software version	NXOS-703I4.2/ACI-N9KDK9-11.3	NXOS-703I7.1/ACI-N9KDK9-12.2A	NXOS-9.3.5 / ACI-N9KDK9-x.x	NXOS-703I7.3/ACI-N9KDK9-14.0	NXOS-9.3.1/ACI-14.1.2/4.1.2
Orderable	Yes	Yes	Yes	Yes	Yes
SKU	N9K-C93180YC-EX	N9K-C93180YC-FX	N9K-C93180YC-FX3S	N9K-C93240YC-FX2	N9K-C93360YC-FX2
Operating System	NX-OS, ACI	NX-OS, ACI	NX-OS, ACI	NX-OS, ACI	NX-OS, ACI

Fuente: (Cisco, 2018)

3.2.1.2 Controlador de infraestructura de Políticas de Aplicaciones (APIC).

El APIC es el cerebro de la solución de ACI, estos dispositivos se basan en servidores x86 UCS series C como se aprecia en la tabla 9 y siempre deben ser instalados en número impares esto es debido a que toda política se almacena en una base de datos. Esa base de datos divide los elementos de una política en fragmentos y distribuye copias de un fragmento en elementos impares que son los controladores de infraestructura centrados en la aplicación.

Los APIC's de Cisco permiten que las aplicaciones se conecten directamente a un grupo de recursos seguro, compartido y de alto rendimiento, que incluye funciones de red, informática y almacenamiento, permitiendo el acceso a éste por medio de una GUI, además por este controlador los conmutadores obtienen sus actualizaciones y parches de firmware, eliminando las operaciones por secuencia de comandos en cada dispositivo.

En si los APIC's son un sistema de políticas y control de red en clúster que proporciona administración de imágenes arranque y configuración de políticas para la estructura de ACI, conocen todo lo que pasa en el fabric y al ser este el punto central de ACI se pueden ver detalles como el estado de los dispositivos, configuraciones, diagramas, problemas y proporciona la habilidad de recuperación de errores en intervalos de tiempo realmente cortos. (Dagenhardt & Moreno, Jose Dufresne, 2012)

Tabla 9. Servidores APIC

Cisco APIC appliance				
	Cisco APIC appliance Medium configuration: M2		**Cisco APIC appliance Large configuration: L2**	
	Description	Default units	Description	Default units
Processor	1.90-GHz Intel® Xeon® processor E5-2609 v3 with 85 watts (W), 6 cores, 15-MB cache, DDR4, and 1600 MHz	2	2.40-GHz Intel Xeon processor E5-2620 v3 with 85W, 6 cores, 15-MB cache, DDR4, and 1866 MHz	2
Memory	16-GB DDR4 2133-MHz RDIMM PC4-17000, dual-rank x4 with 1.2V	6	16-GB DDR4 2133-MHz RDIMM PC4-17000 dualrank x4 with 1.2V	12
Hard Drive	600GB 12G SAS 10K RPM SFF HDD	2	1.2 TB 12G SAS 10K RPM SFF HDD	2

Fuente: (Cisco, 2018)

El APIC proporciona las siguientes funciones de control:

● Administrador de políticas: administra el repositorio de políticas distribuido responsable de la definición y la implementación de la configuración basada en políticas de Cisco ACI.

● Administrador de topología: mantiene actualizada la topología de Cisco ACI y la información de inventario.

● Observador: el subsistema de seguimiento de la APIC; sirve como un repositorio de datos para la información de rendimiento, salud y estado operativo de Cisco ACI

● Boot director: controla el arranque y las actualizaciones de firmware de los conmutadores de spine y leaf, así como los elementos APIC.

● Administrador de dispositivos: gestiona la formación y el control del clúster de dispositivos APIC.

● Administrador de máquinas virtuales (o VMM): actúa como un agente entre el repositorio de políticas y un hipervisor, es responsable de interactuar con los sistemas de administración del hipervisor, como vCenter de VMware

● Administrador de eventos: administra el repositorio de todos los eventos y fallas iniciadas desde la APIC y los nodos de la estructura.

● Inventario de dispositivo: administra el inventario y el estado del dispositivo APIC local.

3.2.1.3 Plano de Datos.

El reenvío a través de la estructura ACI está completamente encapsulado en VXLAN este protocolo desacopla los dominios de capa 2 de la infraestructura de red de capa 3 subyacente

VXLAN es utilizado en toda la fabric de ACI (conmutadores spine, leaf), e incluso dentro de varios elementos como vSwitch conectados al fabric a través de varios hipervisores. Sin embargo, las VLAN 802.1Q todavía están son usadas en el modelo de política ACI porque la vNIC real de cualquier carga de trabajo "hipervisada" y las de los servidores sin sistema operativo hoy en día no admiten la encapsulación nativa VXLAN. Por lo tanto, las redes 802.1Q siguen apareciendo en la política ACI y son métodos de reenvío válidos en la NIC de carga de trabajo. (Dagenhardt & Moreno, Jose Dufresne, 2012)

3.2.1.4 Plano de Control.

Cada Leaf o Spine utiliza un valor de longitud de tipo específico en un flujo de señalización del Protocolo de descubrimiento de enlace local (LLDP) para conectarse con el APIC y, por lo tanto, registrarse como una posible nueva adición a la estructura, la misma que es aceptada por el administrador de la plataforma no se permite la adición sin intervención humana.

Para el reenvío y accesibilidad a través de la fabric de ACI se utiliza el protocolo IS-IS; además se permiten varios protocolos de tipo IGP para el enrutamiento externo como I-BGP, OSPF y EIGRP

La fabric de ACI utiliza el protocolo Council of Oracle Protocol (COOP) para obtener la información de accesibilidad que se relacionan entre leaf y spine para rastrear los elementos adheridos al fabric.

OpFlex es otro nuevo protocolo de plano de control utilizado en ACI, este es un protocolo diseñado para comunicar la intención de la política, desde APIC, y el cumplimiento o el incumplimiento de un elemento de aplicación de políticas adjunto al fabric de ACI. (Dagenhardt & Moreno, Jose Dufresne, 2012)

3.2.2 Interfaz.

Existen 3 maneras de interactuar con ACI.

La primera es por medio de una interfaz de usuario HTML 5 en el propio APIC, el mismo que puede ser accedido por una URL en cualquier navegador.

La segunda es a través de la interfaz de línea de comandos (CLI) muy similar a la sintaxis del sistema operativo utilizado en los equipos Nexus (NX-OS).

La tercera es por medio de un de API REST que permite interactuar de manera interesante con el APIC y las tareas que se pueden hacer en éste, se controla modificando objetos mediante instrucciones como POST, PUT, GET, DELETE y GET, para lo cual dispone de un elemento llamado API inspector que está integrado en la GUI de APIC. Esto permite que cualquier persona con acceso a él determine rápidamente el formato de las llamadas a la API para realizar las mismas funciones que se crean en la GUI. (Dagenhardt & Moreno, Jose Dufresne, 2012)

3.2.3 Topologías

3.2.3.1 Modelo de Sitio Único

La fabric de ACI más fácil de implementar es la del modelo de sitio único, donde existe un único Fabric de ACI en una sola ubicación física. Por lo tanto, todos los elementos de la Fabric, APIC y dispositivos tales como leaf, spine se encuentran interconectados por pares de fibras de corto alcance, la capacidad de escalamiento que se tiene en este modelo es de hasta seis spines y 80 leaf, con los mismos tres APIC para control y administración.

Hay que tener en cuenta que en la fabric cada leaf se conecta a un spine y que no hay conexiones cruzadas entre spines o entre leaf (arquitectura spine-leaf) y que en caso de necesitar más allá de 80 conmutadores leaf una sola ubicación se puede escalar hasta 200 leaf como máximo con los mismos 6 conmutadores leaf, pero se requiere un grupo de cinco APIC. (Dagenhardt & Moreno, Jose Dufresne, 2012)

3.2.3.2 Modelo Multi-Pod

La fabric de ACI se pueden implementar en módulos, donde puede haber un plano de administración común con un único clúster APIC, pero planos de control individuales de MP-BGP y COOP, y planos de reenvío individuales entre los conmutadores leaf y spine de cada módulo. Esto permite la segmentación de planos de control dentro de fabric's muy grandes o para cumplir con un requisito de conformidad. Multi-Pod utiliza una red de tránsito IP separada usando MP-BGP EVPN entre los pods. Esta red de tránsito está conectada a través de cada spine en cada grupo a 40 o 100 Gbps, dependiendo de la estructura de spine en cada grupo como se muestra en la figura 16.

Para este modelo se puede tener más de tres grupos, pero hay que considerar que el numero recomendado de APIC's por grupo es de 3. También existe el requisito de tiempo de ida y vuelta (RTT) de 50 milisegundos, este tiempo de ida y vuelta se mide spine a spine través de la red de tránsito. (Dagenhardt & Moreno, Jose Dufresne, 2012)

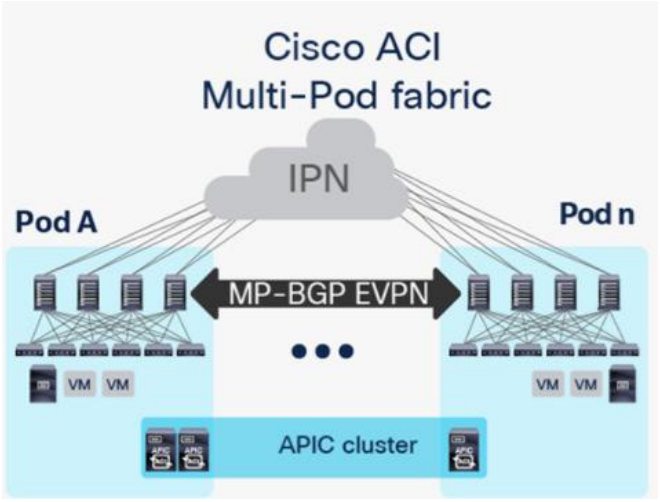


Figura 16. Topología Cisco ACI Multi-POD fabric
Fuente: (Cisco, 2019)

3.2.3.3 Modelo Multi-Site

Más allá del modelo Multi-Pod para conectar y operar múltiples estructuras ACI está el diseño Multi-Site. Esencialmente, Multi-Pod y Multi-Site comparten muchos atributos para la conectividad entre las estructuras y las ubicaciones. El uso de MP-BGP con red IP basada en EVPN para interconectar los Spine de cada fabric es el mismo entre las dos opciones. La demora de ida y vuelta y las velocidades del puerto de la red que conecta los pods y los sitios entre sí también son especificaciones compartidas para ambas opciones. Lo que es sorprendentemente diferente es en Multi-Site cada Pod utiliza un clúster de APIC separado e independiente para cada sitio, por lo cual para lograr consistencia entre ambos sitios se requiere emplear el uso del controlador multi sitio ACI como se aprecia en la figura 17. Este se convierte en el elemento que sincroniza las configuraciones de políticas entre las diversas estructuras ACI en la implementación de varios sitios. La principal diferencia entre Multi-Pod y Multi-Site es la separación de los planos de control y administración entre fabric: mientras que todos los pods en un diseño de Multi-Pod comparten el mismo Cisco ACI Clúster APIC, cada sitio en una arquitectura de sitios múltiples tiene su propio conjunto de APIC. (Cisco, 2019)

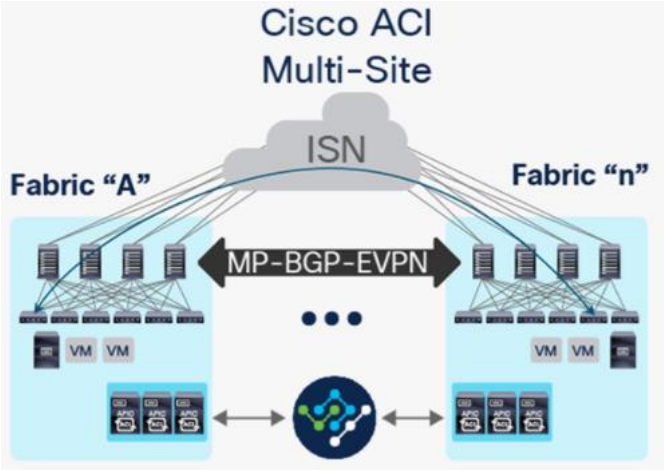


Figura 17. Topología Cisco ACI Multi-POD fabric
Fuente: (Cisco, 2019)

3.1 Modelo de Políticas de ACI

Como se indica en Cisco (2013b), la política de ACI es un modelo orientado a objetos basado en la teoría de promesa, que tiene como característica, el control escalable de objetos inteligentes y depende de estos objetos subyacentes para manejar los cambios de estado en la configuración iniciados por el sistema de control mismo como "cambios de estado deseados", esto se puede apreciar en la siguiente figura 18. De esta manera, los objetos son responsables de devolver las excepciones o fallas al sistema de control. Este método reduce la carga y la complejidad del sistema de control y logra una mayor escalabilidad.

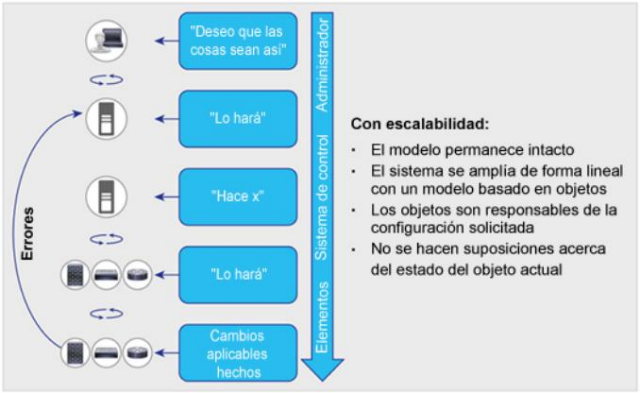


Figura 18. Enfoque de la teoría de la promesa para un control del sistema a gran escala
Fuente: (Cisco, 2013b)

3.1.1 Modelo de objetos.

Como menciono anteriormente, APIC funciona como un motor de políticas orientado a objetos. Esto permite que el administrador de la red defina los estados deseados de entramado, pero dejando la implementación al propio controlador APIC, la figura 19 proporciona una descripción general del modelo de políticas en ACI y sus construcciones lógicas, como Tenan, VRF, Bridge Domain o EPG.

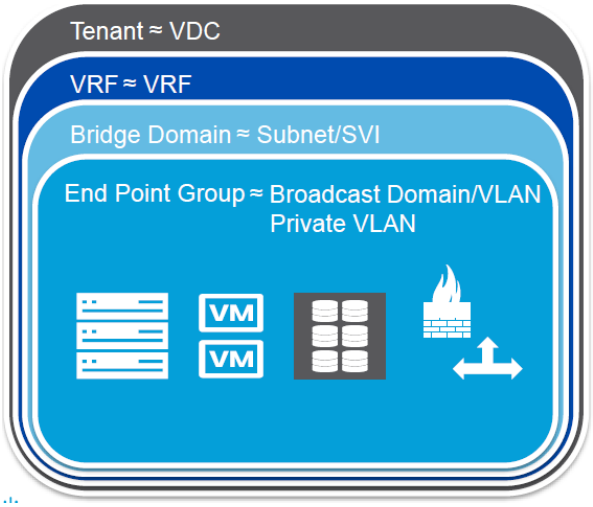


Figura 19. Modelo de objetos en ACI
Fuente: (He, 2018)

Como se describe en Cisco (2013b), en el nivel más alto, el modelo de objetos de ACI se construye en base a grupos de arrendatarios (tenans), lo que permite que la gestión de la infraestructura de red y de los flujos de datos se maneje por separado. Los tenans pueden ser clientes, unidades empresariales o grupos, por ejemplo, una empresa puede ser un tenan para toda la organización y un proveedor de nube puede tener clientes que usen uno o varios tenans para representar a sus organizaciones.

Los tenans se pueden dividir aún más en contextos, algo relacionado directamente con el reenvío y routing virtuales (VRF) o los espacios de IP independientes, cada tenan puede tener uno o varios contextos.

Los contextos proporcionan una forma de separar aún más los requisitos de un cliente y debido a que los contextos usan instancias de reenvío distintas, el direccionamiento IP puede ser el mismo en contextos independientes.

En el contexto, el modelo proporciona una serie de objetos que definen la aplicación, estos objetos son: End Points (EP), End Point Groups (EPG), y las políticas que definen la relación entre ellos. Las políticas son más que un simple conjunto de listas de control de acceso (ACL) si no también una recopilación de: filtros de entrada y de salida, configuración de calidad del tráfico, reglas de marcas y reglas de redirección.

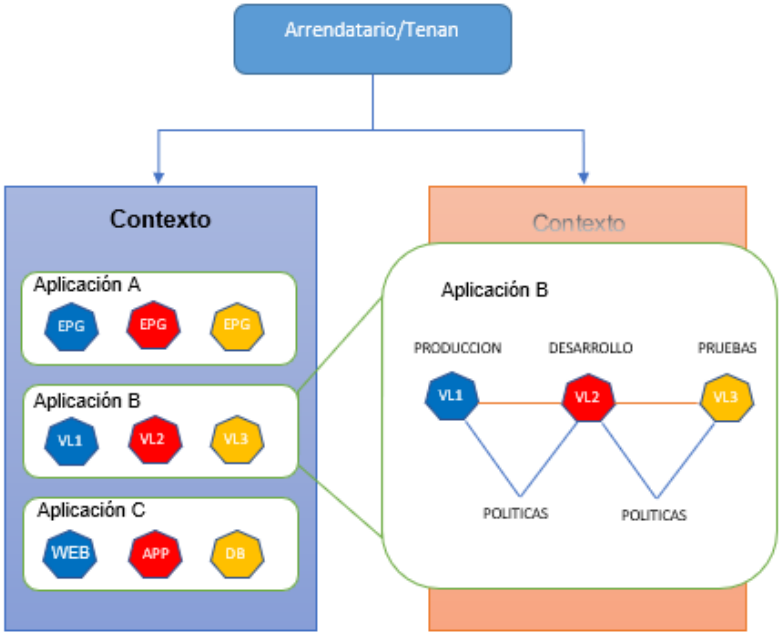


Figura 20. Modelo de objetos lógicos

En la figura 20 se observa un arrendatario/tenan con dos contextos y los perfiles de aplicación (AP) que crean dichos contextos. Los EPG que se muestran son grupos de terminales que crean un nivel de la aplicación. Por ejemplo, el perfil de la aplicación B que se muestra ampliada en el lado derecho de la figura 20, puede estar compuesta varios EPG como VL1 producción (azul), una VL2 para desarrollo (rojo) y una VL3 pruebas (amarillo). La combinación de los EPG y las políticas que definen su interacción es un perfil de red de aplicación en el modelo de ACI.

3.3.1.1 Endpoint Group (EPG).

Una entidad lógica que contiene una colección de puntos finales de red físicos o virtuales proporcionando una agrupación lógica de objetos que requieren políticas similares como ejemplo se puede tener terminales los mismos que incluyen servidores, máquinas virtuales, almacenamiento o clientes en Internet como se aprecia en la figura 21. (Cisco Public, 2016)

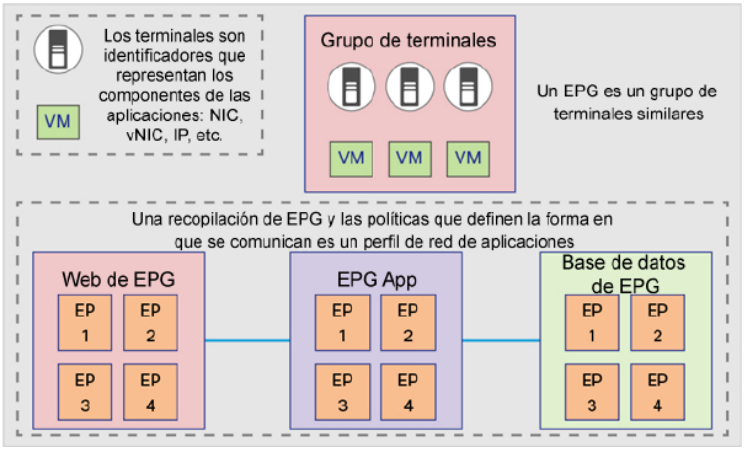


Figura 21. Relación de grupos de terminales
Fuente: (Cisco, 2013c)

Como se indica en Cisco (2013c), los EPG también se utilizan para representar redes externas, servicios de red y almacenamiento de red, es decir son recopilaciones de uno o varios terminales que proporcionan una función similar, además son utilizados para definir los elementos a los que se aplica una política. En la figura 21 se puede ver una recopilación de los EPG, políticas y su forma de interactuar en un perfil de aplicación.

Dentro del fabric de red, la política se aplica entre los EPG, donde se establece la forma en que los EPG se comunican entre sí algunos ejemplos de su uso son:

- EPG creado por VLAN de red tradicionales: todos los terminales conectados a una VLAN determinada ubicada en un EPG
- EPG creado por una LAN virtual ampliable (VXLAN): igual que las VLAN, excepto en que usan VXLAN
- EPG asignado a un grupo de puertos de VMware
- EPG definido por la IP o subred
- EPG definido por nombre de DNS o rangos de DNS

3.3.1.2 Application Programming Interface (API).

Una aplicación de interfaz de programación es un software intermediario que permite que dos aplicaciones se comuniquen entre sí. Cada vez que usan más en aplicaciones como Facebook.

3.3.1.3 Application profile (AP).

Define las políticas, servicios y las relaciones entre los grupos de puntos finales (EPG), un EPG origen y otro EPG destino se relacionarán mediante una aplicación de política en caso de que exista la necesidad de que estos se comuniquen entre sí ya sea para calidad de servicio (QoS), control de acceso, inserción de servicios entre otros; para lo cual hacen uso de reglas o filtros de entrada, salida como permitir, denegar, redireccionar, registro, copias y marcas.

3.3.1.4 Bridge Domain (BD).

Un dominio de puente es un conjunto de puertos lógicos que comparten las mismas características de difusión o inundación. Al igual que una LAN virtual (VLAN), los dominios de puente abarcan varios dispositivos. (Cisco Public, 2016)

Un BD debe estar vinculado a una VRF (llamada también contexto o red privada) en donde se definen las direcciones de red a usar, dichas subredes se pueden definir en uno o más DB que harán referencia a la VRF, además en un BD se define el espacio de direcciones MAC de capa 2, por lo cual las siguientes opciones de configuración se deben considerar para determinar el comportamiento del DB:

- El uso de hardware proxy or unknown unicast flooding: hardware proxy es la opción por default, en el cual si no se conoce el destino del paquete es enviado al Spine, si el Spine no conoce también el paquete entonces este es descartado.

- Habilitar o no ARP flooding: cuando se encuentra habilitado el tráfico de ARP se inundará dentro de la fabric de ACI como en las redes tradicionales, por el contrario, si esta opción esta desactivada, la fabric intentará utilizar tramas unicast para enviar el tráfico ARP al destino. Esta opción solo es aplicable cuando la opción unicast routing está habilitada en el DB, si está estuviera des habilitada el tráfico ARP siempre será inundada.

- Habilitar o no unicast routing: Al estar esta opción habilitada permite configurar una dirección de red en la que la fabric de ACI proporciona la función de puerta de enlace predeterminada para enrutar hacia este Gateway el tráfico sin embargo el aprendizaje de IP no depende de tener una subred configurada.

- Definición de una subred, en esta opción se configura la dirección IP como SVI (Gateway) para el bridge domain.

3.3.1.5 Tenan.

Un Tenan es un contenedor lógico para políticas de aplicación que permite a un administrador ejercer un control de acceso basado en el dominio, representa una unidad de aislamiento desde una perspectiva política, pero no representa una red privada. Los Tenan pueden representar a un cliente en una configuración de proveedor de servicios, una organización o dominio en una configuración empresarial, o simplemente una agrupación conveniente de políticas. (Aci et al., 2019)

3.3.1.6 Attachable Entity Profile (AEP).

Un perfil de entidad adjunta (AEP) representa un grupo de entidades externas con requisitos de política de infraestructura similares. Las políticas de infraestructura consisten en políticas de interfaz física que configuran varias opciones de protocolo, como el Protocolo de descubrimiento de Cisco (CDP), el Protocolo de descubrimiento de capa de enlace (LLDP) o el Protocolo de control de agregación de enlaces (LACP). (Systems, 2019)

3.3.1.7 Contratos.

Los contratos establecen los permisos de entrada, salida, denegaciones, reglas y políticas de calidad de servicio, redirección. Permiten realizar definiciones simples y complejas para determinar la manera en la que un EPG se comunicará con otro. Los contratos se aplican entre EPG mediante una relación proveedor- consumidor, es decir un EPG proporcionará el contrato (proveedor) y otro EPG consumirá dicho contrato (consumidor) como se aprecia en la figura 22.

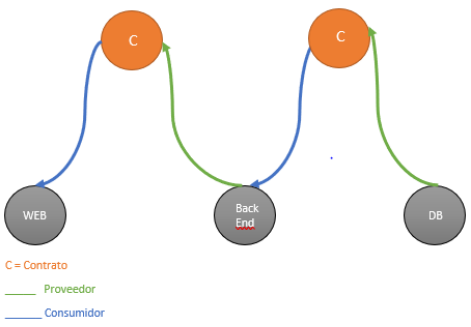


Figura 22. Contratos

El modelo de proveedor-consumidor es útil para diversos fines. Ofrece una forma natural de añadir una "protección" o "membrana" a un nivel de aplicación que imponga la forma en que dicho nivel interactúe con otras partes de una aplicación. Por ejemplo, un servidor web puede ofrecer HTTP y HTTPS, por lo que el servidor web se puede incluir en un contrato que solo permita estos servicios. Además, el modelo de proveedor-consumidor de contrato fomenta la seguridad al permitir actualizaciones de políticas sencillas y uniformes en un solo objeto de política en lugar de en diversos enlaces que pueda representar un contrato. Los contratos también ofrecen sencillez al permitir que las políticas se definan una vez y se reutilicen muchas veces como se aprecia en la figura 23. (Cisco, 2013b)

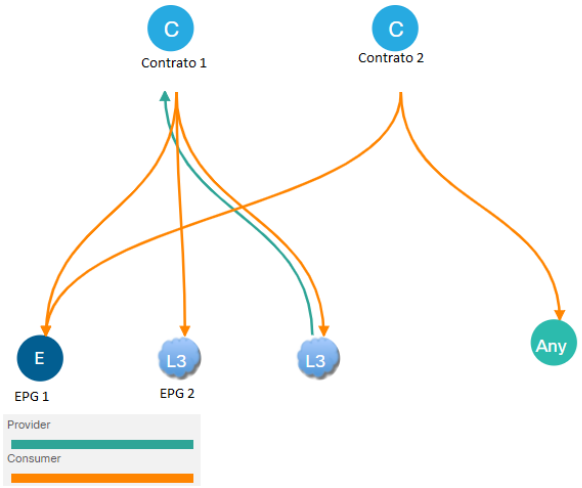


Figura 23. Reutilización de contratos

3.2 Propuesta del diseño

3.2.1 Definición de la topología física.

En esta sección se define la topología, conexiones y el diagrama de red física de la solución propuesta.

3.2.1.1 Topología.

La topología del diseño será de tipo spine-leaf en la cual se utilizará equipos que no sean de tipo chasis debido a su tamaño, cada Leaf ira conectado a cada Spine y no existirán conexiones entre leaf debido a que este modelo está constituido de esta manera.

Para los spine se propone el uso de 2 switch cisco Nexus N9K-C9364C por redundancia, tienen un tamaño de 2 UR; soporta hasta 12.84 Tbps de capacidad y 4.3 Bpps. Pose 64

puertos de velocidades de 40/100 Gbps con formato de interfaces QSFP28 y de 10 Gbps con interfaces en formato SFP+.

Además, permite hasta 1000 VRF, un máximo de 4096 VLANS, 32 enlaces en un solo enlace agregado, 64 enlaces agregados, 256 vteps.

Para los Leaf se propone el uso de 2 switch Cisco Nexus C93180YC-FX; además cada Leaf debe tener al menos 2 fabric extender de 48 puertos para las conexiones de clientes especiales con puertos de cobre por ejemplo para el servicio de housing donde el cliente suele traer su propio switch y por ende no se tiene administración del mismo, o en casos donde el cliente trae su propio servidor virtualizado en el cual estos se comportan como un virtual switch, adicional a esto cada host de cliente deberá conectarse en VPC hacia un puerto de cada Fex asociado a un Leaf por redundancia y alta disponibilidad; estos switch leaf son de 1 UR, tienen 48 puertos ópticos con capacidades de 1 hasta 25 Gbps para la conexión de usuario final y 6 puertos de 10 hasta 100 Gbps para la conexión hacia los Spine en formato QSFP28, capacidad de 3.6 Tbps y 1.2 Bpps, además soporta conexiones de 16 a 32 Gbps para Fibra canal.

Además, soporta 512000 entradas mac, conexión de máximo 16 Fabric extender por switch, 4096 VLANS, hasta 16000 VRF, 512 enlaces agregados, máximo 32 enlaces en un enlace agregado.

Los dos modelos tanto para el spine como para el leaf soportan software de ACI, se propone el uso de 3 servidores APIC para la administración y gestión ya que es un ambiente pequeño y seria innecesario el uso de 4 o más APIC para la administración.

La cantidad de puertos de switch Leaf está asociada con la cantidad de puertos utilizados de 10G que actualmente están siendo usados en la infraestructura tradicional y en base a la sección 2.3 con 48 puertos de 10G por cada leaf se puede suplir la necesidad actual para migrar los clientes o servicios que sean de tipo multi tenan o plataformas que sirvan a varios clientes tales como host/servidores para máquinas virtuales, firewalls que actualmente tienen conexiones en 10 Gbps en la infraestructura tradicional, estos irán conectados ahora a la infraestructura de ACI y además se tendrá puertos disponibles para futuro crecimiento; hasta el momento no se contempla la migración total de todos los clientes de la infraestructura tradicional, es decir todos clientes conectados a los extensores ya que esto implica un análisis por cada cliente y sus servicios para mejorar el

diseño actual que tenga cada uno de estos y en base a eso proporcionar una nueva solución en la plataforma de ACI lo cual no está contemplado en el alcance de este estudio.

3.2.1.2 Conexión.

Cada switch Spine tendrá una conexión a cada switch leaf en 100 Gbps, es decir se tendrá 2 conexiones que salen de cada spine por redundancia, como las distancias entre estos equipos son cortas se usará Fibra Óptica multimodo con transceivers de tipo SR que sirven con este tipo de fibra óptica, además cada switch Spine tendrá una conexión WAN en 10 Gbps hacia el Core de la red hacia equipos Cisco ASR 9000, para conexión de Multisite, la misma que no forma parte del alcance de este estudio, con lo cual se tiene 2 conexiones redundantes hacia la capa de Core, que también irán conectadas a distintos equipos para tener redundancia a nivel de conexión y equipo físico.

Cada Switch Leaf tendrá una conexión hacia cada Spine en 100 Gbps, esto con el propósito de proveer redundancia en caso de falla de un Spine, de igual manera en caso de falla de uno de sus enlaces de 100Gbps, no se propone añadir más conexiones redundantes entre estos equipos ya que están en una misma localidad y problemas de fibra por corte o agentes externos no se va a dar, lo más probable que puede darse es una falla de una de las conexiones pero no varias a la vez; como las distancias entre estos equipos son cortas se usará Fibra Óptica multimodo, además tendrá una conexión a 10 Gbps hacia cada switch Nexus 7010 en una VPC para temas de migración y comunicación hacia la infraestructura de red tradicional como un enlace capa 2 (L2Out) estático, de la misma manera las conexiones están dentro de una misma localidad y se establecerá como una VPC por temas de redundancia y alta disponibilidad; además se tendrá una conexión WAN desde cada leaf por redundancia hacia los equipos de Core para proveer una conectividad en capa 3 (L3Out) con el fin de conectar a redes externas a la Fabric ACI, esta conexión y la utilidad de la misma no forman parte del alcance de este estudio.

Cada APIC tendrá una conexión hacia cada Leaf, es decir serán 2 conexiones por cada APIC, por redundancia y además debido a la limitante de puertos por cada servidor, dichas conexiones serán para la gestión, configuración, comunicación y administración de los equipos (Spine y Leaf), esta conexión será mediante Fibra Óptica Multimodo e interfaces de tipo 10G-SR, debido a que la distancia de conexión entre los equipos es corta.

3.2.1.3 Diagrama de red.

En la figura 24 se puede observar el diagrama de red físico propuesto para la solución SDN, donde las líneas más gruesas y resaltadas corresponden a las conexiones del alcance de este estudio.

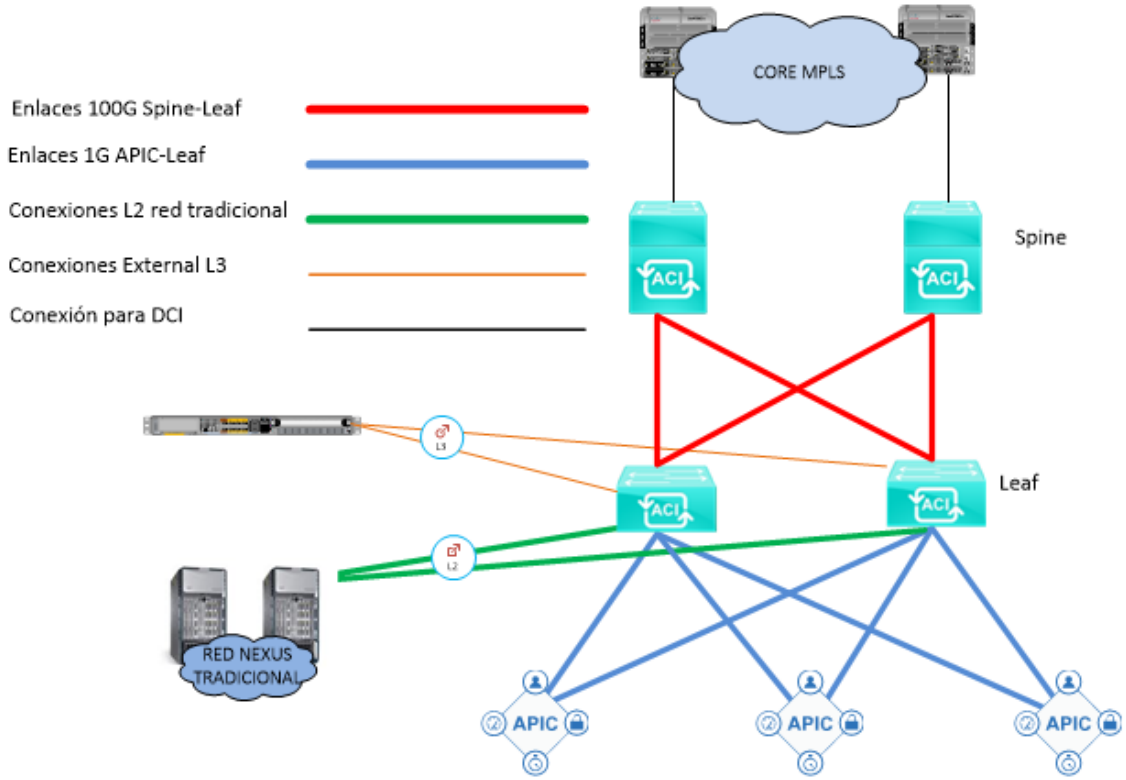


Figura 24. Diagrama de la solución SDN

3.2.2 Definición de políticas y parámetros lógicos de configuración.

En esta sección se proponen las políticas y parámetros lógicos a considerar para la solución.

3.2.2.1 Aprovisionamiento.

Cabe recalcar que durante el despliegue del fabric de ACI es fundamental considerar que no es posible cambiar el rango de las direcciones IP de la infraestructura es decir las IPS de TEP, ya que al inicio de la configuración se solicitará el ingreso de un rango de direcciones TEP, el mismo que es utilizado para los nodos Spine y Leaf dentro de la Fabric, por lo cual se recomienda asignar un bloque de direcciones único que no vaya a superponerse con ninguna otra red enrutada ni durante el despliegue ni a futuro para no

tener conflicto de direccionamiento cuando se utilice la integración de Vcenter al ACI (VMM).

3.2.2.2 Tenans.

Se propone el uso de un solo Tenan para clientes pequeños, es decir clientes con pocos servicios o que comparten la misma infraestructura virtual (servidores/host) para evitar el abuso en la configuración de contratos que puede generar mayor complejidad de la necesaria, se debe tener en cuenta que cada tenan es independiente y no tendrá comunicación con otros salvo por un contrato.

3.2.2.3 Bridge Domains.

Es bridge domain, es un objeto de reenvío de capa 2 dentro de la fabric de ACI que se utilizada para restringir el tráfico de broadcast y multicast, por lo cual se creará uno por cada VLAN/cliente, para el escenario provisto el gateway se mantendrá fuera de la Fabric de ACI por lo cual solo proporcionará servicios de capa 2 a los dispositivos que se encuentren en un EPG, por lo cual las siguientes características deben tener los siguientes atributos:

Unicast routing, deben estar deshabilitado (direcciones IP no se aprenderán) ya que no es necesario.

ARP Flooding, debe estar activado, ya que esto habilita las solicitudes ARP, para que sean enviados (inundados) a través de la infraestructura de ACI hacia la red tradicional donde se encuentra su puerta de enlace.

L2 Unknown Multicast, unicast Flooding, debe estar como flood, estas solicitudes de tramas de multicast y unicast desconocidas de capa 2 que se originan en los dispositivos conectados a ACI deben poder alcanzar a la red tradicional u otros puntos que formen parte del mismo segmento de red y que pueden estar conectados en la red tradicional.

Para los casos en las que se conecten firewalls que funcionen con alta disponibilidad (HA) (modo activo/standby) o se utilice clúster de servidores se recomienda usar la siguiente configuración ya que va a existir dispositivos donde su dirección IP se podría mover a otro equipo lo que significa que la dirección IP estará asociada a una nueva dirección MAC.

L2 Unknown Unicast: Hardware Proxy.

Unicast Routing: habilitado, pese a que no se vaya a usar una subred y que la puerta de enlace predeterminada este en la red tradicional se habilita con fines de ayuda para resolución de problemas por ejemplo uso de tracert.

ARP Flooding: habilitado.

Subnet Configured: solo si el gateway estará en el fabric de ACI caso contrario no es necesario.

3.2.2.4 Administración de la solución

Para la administración de la solución se puede realizar por medio de una red fuera de banda (Out-band) como una red dentro de banda (In-Band), para la red Out-band que es el método de administración por default se utilizan los enlaces dedicados que proporcionan los APIC, mientras que para la administración de In-Band se requiere conectividad por medio de la infraestructura de ACI que se conecta hacia la red tradicional donde el Gateway esta fuera de la Fabric de ACI.

Se recomienda mantener la red Out-Band por default para los temas de administración de la solución, ya que ante alguna falla de la Fabric de ACI se podrá ingresar sin inconvenientes al Cluster de APIC y validar el inconveniente, de igual manera al estar los leaf y spine conectadas a una red fuera de banda se puede acceder por la interfaz línea de comandos para temas de resolución de problemas.

Además, al tener la empresa un sistema de monitoreo se necesita también tener una red In-Band con fines de almacenamiento de historial de tráfico, visibilidad de alertas por el Centro de Operaciones de Red (NOC) y su sistema de gestión de alarmas, por lo cual se recomienda tener una red dentro de banda.

3.4.2.5 Identificación de los conmutadores

Durante la asociación de los nodos al Fabric de ACI se requiere de un identificador o ID que está compuesto de un número entre 101 - 4000, por lo cual se recomienda identificar los conmutadores en base a la función que estos desempeñaran ya sea Spine o Leaf y además para una fácil identificación es recomendable que parte de la identificación sea el rack donde se encuentra instalado.

Para los Spine se recomienda que el primer dígitos sean 1 mientras que para los Leaf se recomienda que sea el 2, el segundo digito se puede hacer referencia al POD para este

caso se tendrá un solo POD por lo cual seria 1, el tercer y cuarto digito puede asociarlo al rack en el cual está instalado el equipo, por ejemplo al referirse al primer Spine que se encuentra instalado en el rack 15 la identificación o el ID de nodo 1115, en cambio para el primer Leaf ubicado en el mismo rack 15 el ID de nodo quedaría 2115.

3.4.2.6 Vlan Pool

Un Pool de VLANS se puede configurar de 2 formas dentro de la Fabric de ACI, el primero es de manera estática y el segundo de manera dinámica, se recomienda la configuración de manera estática para equipos finales de clientes como servidores, firewalls, conexiones WAN mientras que para máquinas virtuales o servicios automatizados se recomienda configurarlos de manera dinámica, además se recomienda agregar las VLANS de una en una o en rangos pequeños para una manipulación más granular, opciones como el rol se mantendrán en externo ya que por lo general será para conexiones externas al fabric.

3.4.2.7 Mis-Cabling Protocol

En la fabric de ACI no se utiliza Spanning Tree Protocol (STP), por lo cual los conmutadores no participan en ningún dominio de STP y evidentemente no se generan paquetes de tipo BPDU, si se diera el caso de un cableado incorrecto que posiblemente genere un lazo dentro de la fabric de ACI se detecta mediante LLDP, sin embargo, a manera de dar una protección adicional se recomienda habilitar el protocolo de Cableado incorrecto (MCP) indicando que al detectar un lazo el puerto caiga.

3.4.2.7 Nomenclatura de los objetos

Es recomendable establecer una nomenclatura de los objetos que sea fácil de identificar respecto a la función que realizan con el fin de que la administración sea mucho más fácil, hay que tener en cuenta que en ACI una vez creado un objeto con un nombre este no puede ser cambiado salvo que sea eliminado para volverlo a crear, además ACI permite poner descripciones a los objetos el cual puede ser de mucha utilidad para brindar información adicional del objeto; como ejemplo de la nomenclatura se recomienda:

Para los grupos de políticas:

Nombre o parámetro que identifica al equipo remoto seguido del guion bajo

- EquipoA_VPC (Virtual Port Channel)

- EquipoB_Po (Port Channel)

Para los perfiles de interfaces en los leaf:

- LfIdNodo_IntProf (Lf=Leaf ID=numero 101-4000 InfProf=Interface Profile)

Para los extenders:

- FEX0918_IntPro (Fex=Extender, FilaRack=0918, InfProf=Interface Profile)
- Eth1_1 (Puerto 1/1, relacionado a la nomenclatura de las interfaces en el conmutador)

Perfiles de Aplicación:

ClienteX_App (App=Application profile)

- Vl100_EPG (EPG del Cliente en la VLAN 100)

Bridge Domain:

- Vl100_DB (Bridge Domain del cliente 100)

3.3 Ventajas y limitaciones de la solución ACI

- La arquitectura de ACI permite una visión general de toda la plataforma de manera centralizada en tiempo real de sus entornos físicos y virtuales.
- La solución permite tener un control detallado de las aplicaciones, tenans, lo que implica una mayor seguridad para los servicios.
- La arquitectura es flexible y escalable debido a las buenas prestaciones de hardware lo que conlleva a que se pueda incrementar o cambiar un nodo sin que exista un impacto a las aplicaciones o usuarios que estén en el fabric.
- Permite la integración con APIS abiertas, así como códigos abiertos como Python, convirtiéndola en una plataforma programable con un aprovisionamiento acelerado y automatizable.
- Dentro de sus limitaciones cabe indicar que no está soportado la creación de VPC con los fabric extender.

- No esta soportado realizar Q-in-Q sobre los puertos de un extender.
- Al ser una tecnología nueva en crecimiento y desarrollo se generarán vulnerabilidades de seguridad.

CAPÍTULO 4 IMPLEMENTACIÓN Y EVALUACIÓN DEL DISEÑO DE LA RED

En este capítulo se establecerá la topología de la red SDN a utilizar, además se indicarán los parámetros a considerar para la implementación, así como también se establecerán las políticas y parámetros de medición a considerar.

4.1 Implementación de la solución SDN

Para la implementación de la solución se instalan en parejas un leaf y un spine en un rack y el otro par de leaf y spine en otro rack; para los servidores se instalan 2 en un rack y el tercero en el otro rack, en resumen, en el rack 1 quedarán 1 leaf, 1 spine y 2 servidores, en el segundo rack queda 1 leaf, 1 spine y 1 servidor.

4.1.1 Configuración APIC's.

En primera instancia para la implementación de la solución en base a ACI de Cisco, se levantará el cluster de APIC's, para lo cual se realiza la siguientes conexiones: en cada uno de los servidores APIC hay 2 puertos que tienen una identificación a la izquierda llamada Fabric los cuales van conectados a los primeros puertos de los switch Leaf, por redundancia deben ser conectados a leaf diferentes, además se tienen 2 puertos con nombres eth1 y eth2 los cuales van conectados a la red fuera de banda (Out-of-Band Management) que servirá para la gestión y administración de la plataforma, también existe un puerto llamado CIMC que servirá de conexión hacia el servidor para ver el estado físico de sus componentes, actualización de su firmware entre otros estos puertos se visualizan de mejor manera en la figura 25:

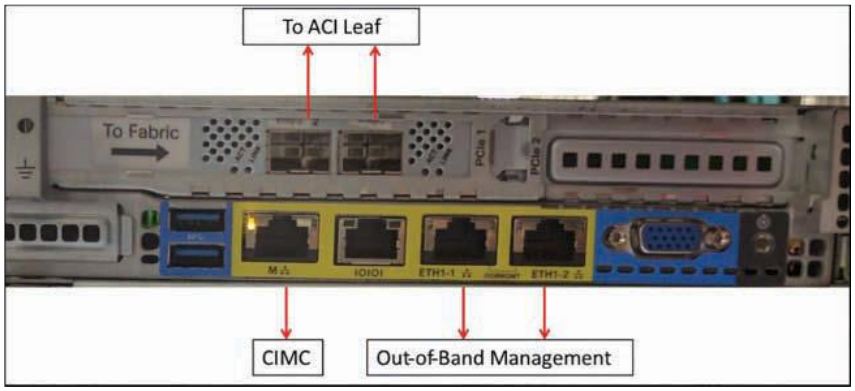


Figura 25. Puertos del servidos APIC
Fuente: (Dagenhardt & Moreno, Jose Dufresne, 2012)

Una vez realizadas las conexiones adecuadas, se conecta teclado y pantalla a cada servidor para realizar el cambio de la IP de CIMC de la siguiente manera cuando el servidor se esté encendiendo y se encuentre en su proceso de POST (power-on selftest) se presiona F8 para acceder a las opciones de configuración del CIMC y se llenará la información como se aprecia en la figura 26:

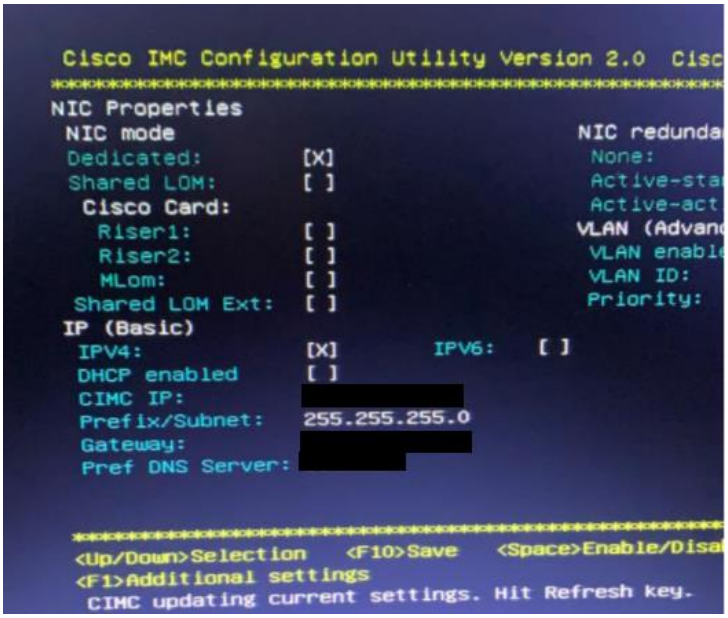


Figura 26. Configuración CIMC

Una vez terminada la configuración se guardará presionando F10, posterior a esto se reinicia el servidor y se podrá ingresar vía navegador web a la consola virtual de cada servidor.

Al iniciar el servidor se solicita ingresar la información del nombre, cantidad de APIC a usar y parámetros requeridos por el script del servidor, estos parámetros se detallan en la tabla 10 y se deben ingresar en cada servidor para poder levantar cada controlador y por ende el cluster de APIC's.

Tabla 10. Parámetros de configuración levantar el cluster de APIC's

PARAMETROS	DETALLE	COMENTARIOS
Fabric Name	FABRIC-UIO1	Nombre del fabric UIO. Defaul: ACI Fabric
Fabric ID	1	ID del fabric.
Cantidad APICS	3	
POD ID	1	ID del POD: 1 (solo se tiene un POD)
Controller name APIC 1	APIC1	Nombre pare el server APIC1

Controller name APIC 2	APIC2	Nombre pare el server APIC2
Controller name APIC 3	APIC3	Nombre pare el server APIC3
TEP address	10.181.0.0/16	Rango red /16 interno para cada fabric
Infra VLAN	3967	valor default en cada POD
BD Multicast Address	default	Red multicast interna de cada POD. Default: 255.0.0.0/15
Enable IPV6	no	Si/no
Out Of band Configuration		
IP MGMT APIC 1	172.16.24.30/24	IP de administracion asignada al APIC 1.
IP MGMT APIC 2	172.16.24.31/24	IP de administracion asignada al APIC 2.
IP MGMT APIC 3	172.16.24.32/24	IP de administracion asignada al APIC 3.
MGMT Gateway	172.16.24.1	IP del gateway de red MGMT
Default MGMT user	admin	cuenta embebida de administracion APIC
Default MGMT Password	--	Password incial del user admin
IP CIMC APIC 1	172.16.24.36/24	IP CICM (gestion chasis) asiganada al APIC 1
IP CIMC APIC 2	172.16.24.37/24	IP CICM (gestion chasis) asiganada al APIC 2
IP CIMC APIC 3	172.16.24.38/24	IP CICM (gestion chasis) asiganada al APIC 3
CIMC Gateway	172.16.24.1	IP del gateway de red CICM
Default CICM User	admin	cuenta embebida de administracion CIMC
Default CICM Password	--	Password incial del user admin

Una vez configurado se puede ingresar por un navegador web mediante la IP de administración asignada a dicho APIC, donde se solicita el user ID y el password como se puede observar en la figura 27:

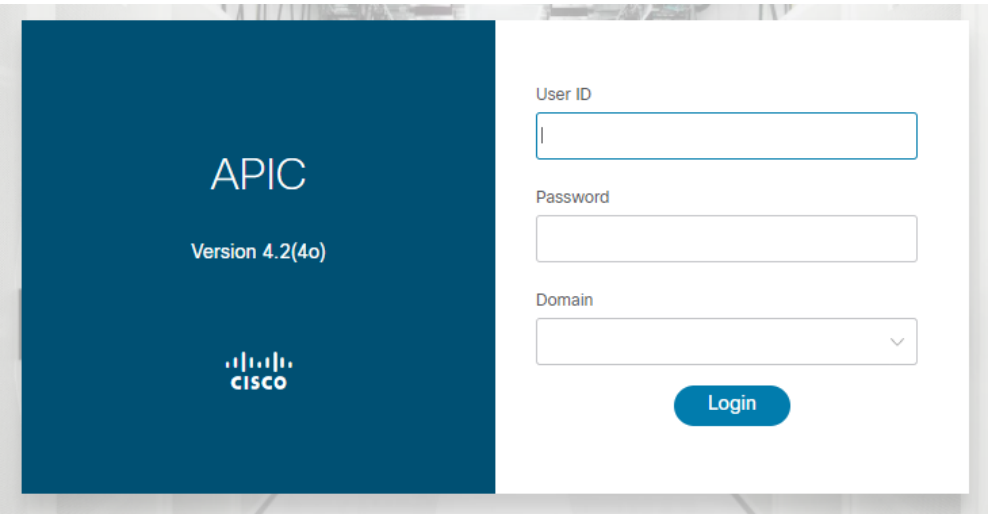


Figura 27. Pantalla de ingreso APIC

Una vez ingresado al APIC se despliega el Dashboard como se aprecia en la figura 28, en la cual se puede apreciar la salud del sistema, un contador de las fallas que existen detalladas por criticidad, salud de los nodos, tenans y estatus de los controladores.

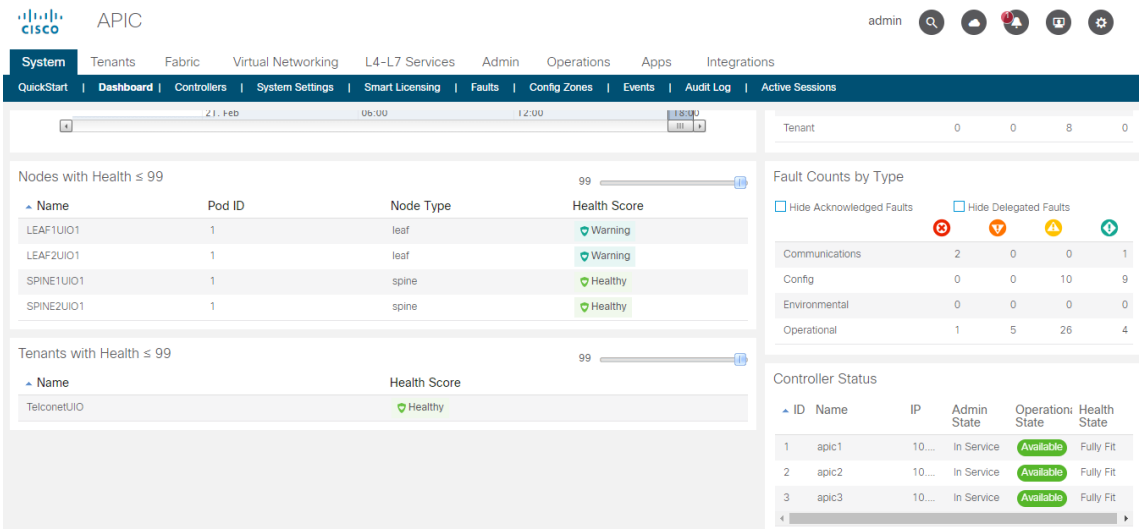


Figura 28. Dashboard de ACI

Durante el despliegue de los APIC's se valida que el status de los 3 controladores sea Fully Fit y que este como disponible como se aprecia en la figura 29, con lo cual se puede establecer que el despliegue del cluster ha sido exitoso y se procede con la configuración de la solución, sin que se presente alertas respecto al tema.

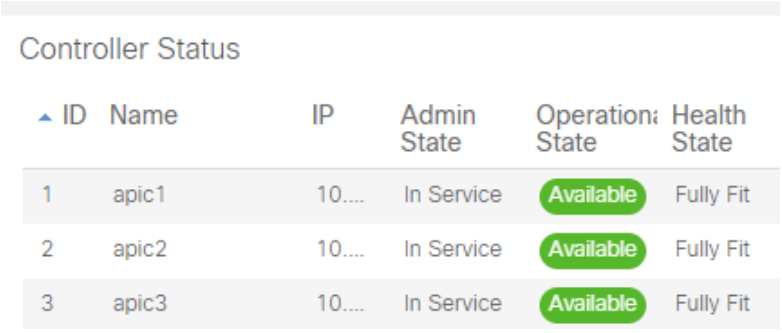


Figura 29. Estado de los 3 controladores de ACI

Para la configuración básica inicial se desplegará un menú como se observa en la figura 30, donde registrará cada uno de los nodos (switch detectados por el fabric por LLDP) al fabric, configuración de BGP para los spine, IPS de administración para los switch, configuración DNS, NTP.

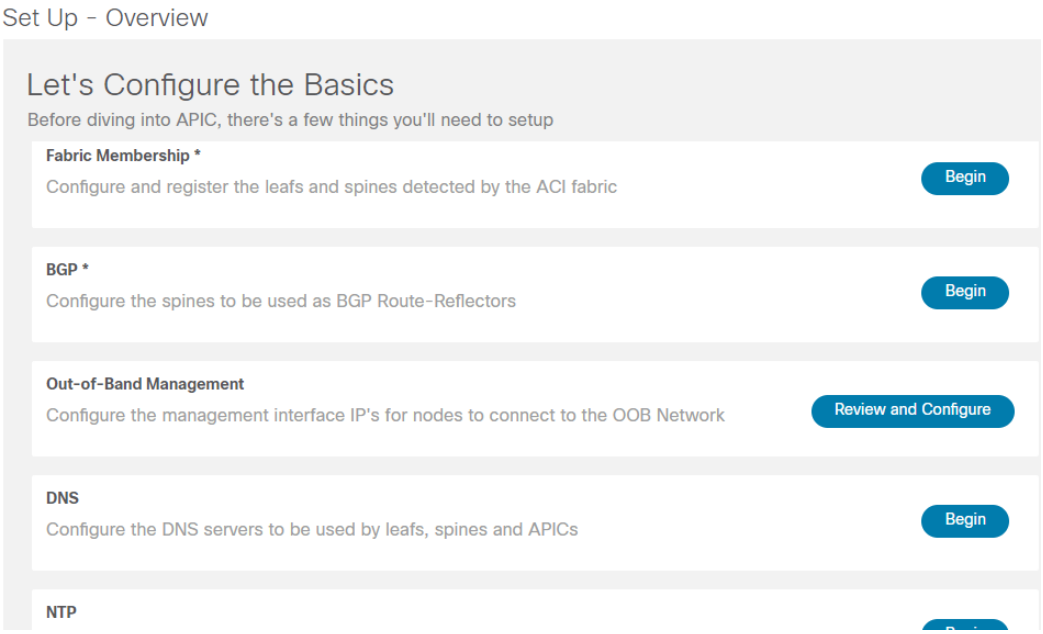


Figura 30. Pantalla de configuración básica inicial

En la primera opción se configura y registrará cada nodo (leaf y spine) al fabric de ACI como se aprecia en la figura 31 y 32 respectivamente, donde se debe agregar un ID a cada conmutador, un nombre que lo identifique y se debe definir la función que tendrá ya sea como leaf o spine.

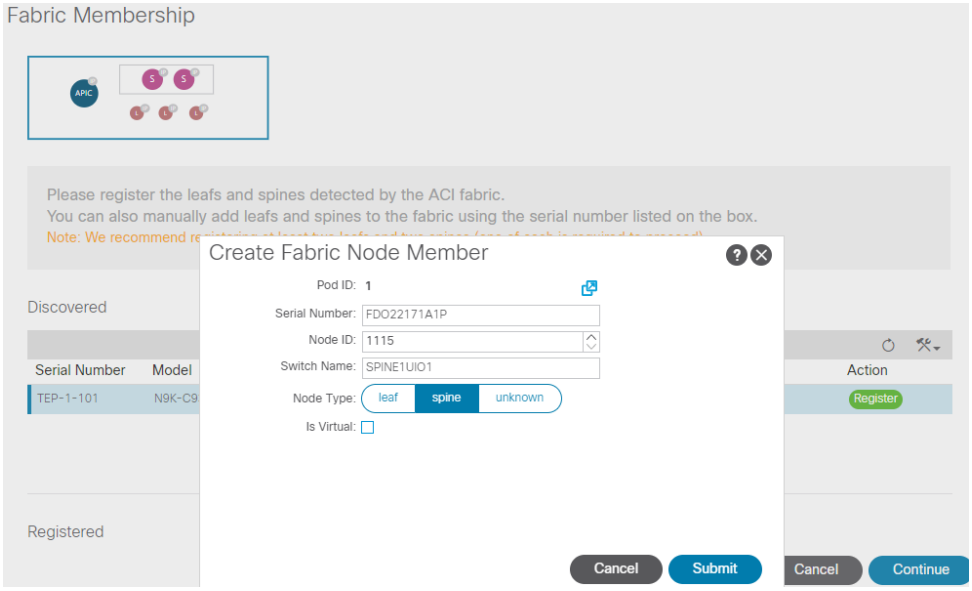


Figura 31. Registro de un conmutador Spine a la Fabric de ACI

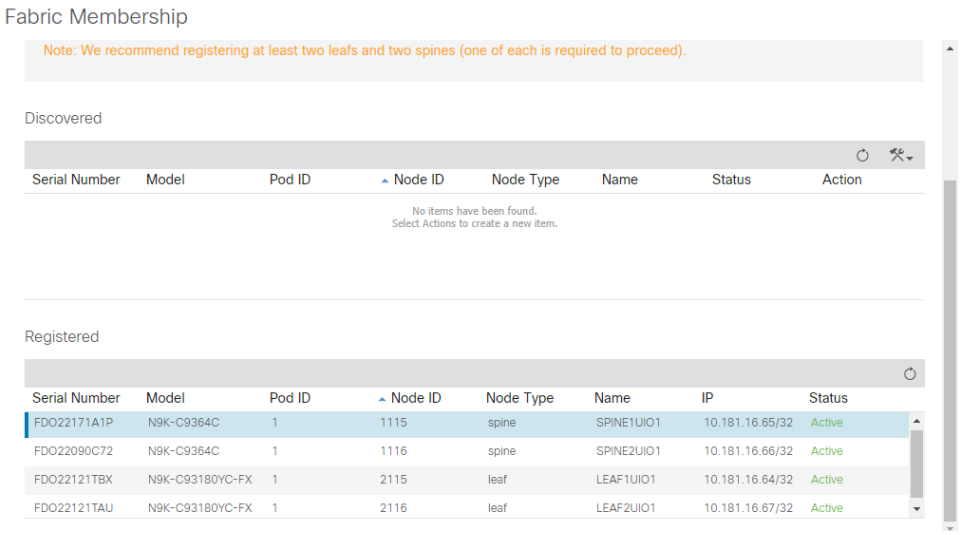


Figura 32. Nodos (Spine/Leaf) registrados a la fabric de ACI

Se considera ahora la configuración de BGP, se requiere establecer un sistema autónomo (ASN) que se utiliza en procesos internos (MP-BGP), que se usa en la distribución de rutas externas dentro de ACI, BGP es configurado en los SPINE como reflectores de ruta (route reflectors) figura 33.

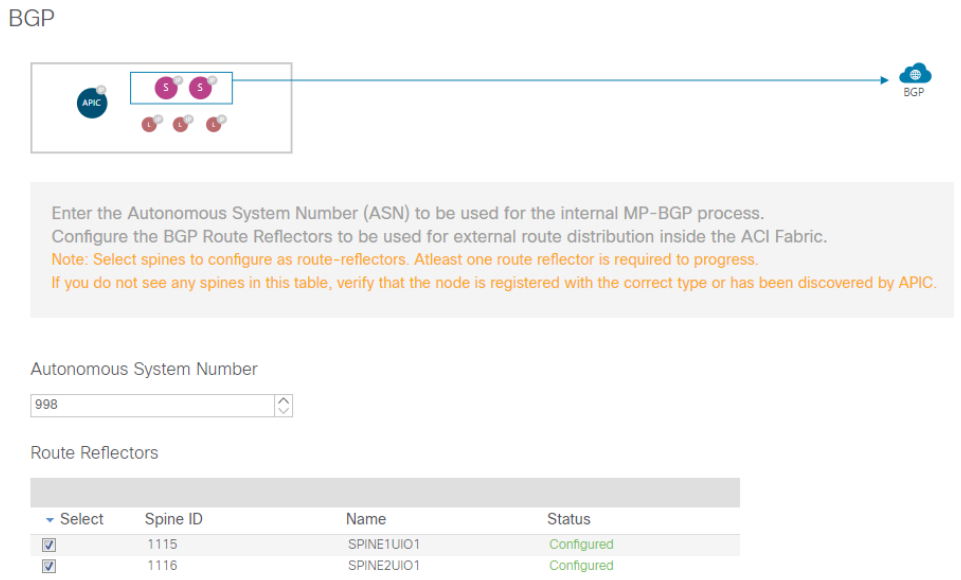


Figura 33. Configuración de BGP para los SPINE

Se requiere también la configuración de las IPS usadas para los servicios de DNS para la resolución de nombres y NTP para la sincronización de la zona horaria y la hora como se aprecia en las figuras 34 y 35 respectivamente, la comunicación hacia los servidores DNS y NTP se la realizará por medio de la conexión fuera de banda (OOB)

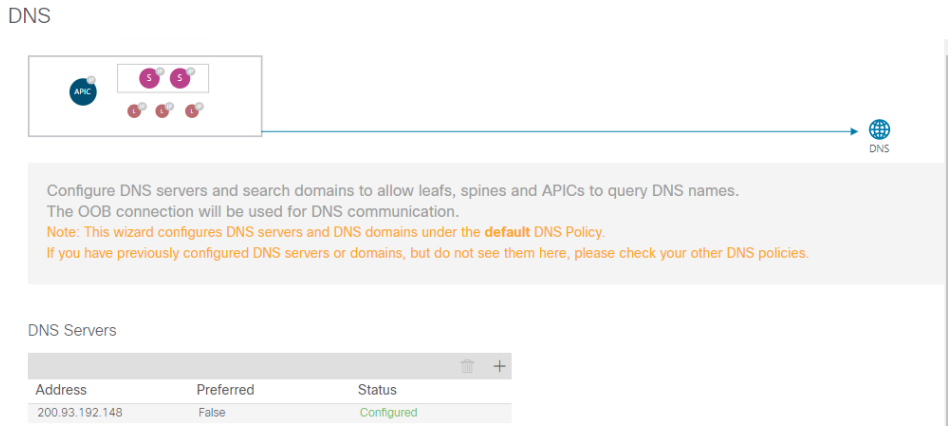


Figura 34. Configuración DNS

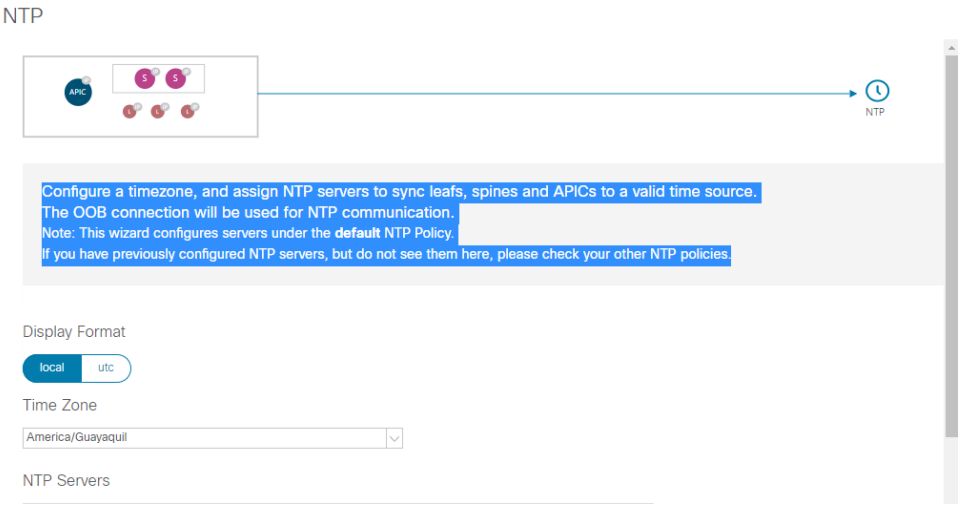


Figura 35. Configuración NTP

Se requiere también configurar las IPS de administración fuera de banda para los conmutadores con el fin de acceder a los mismos de manera tradicional, es decir por la interfaz de línea de comandos (CLI), esto en caso de que se requiere realizar algún tipo de resolución de problemas o en caso de que se pierda comunicación mediante el Fabric de ACI, en la figura 36 se aprecia las direcciones IP de cada uno de los elementos que confirman la fabric de ACI

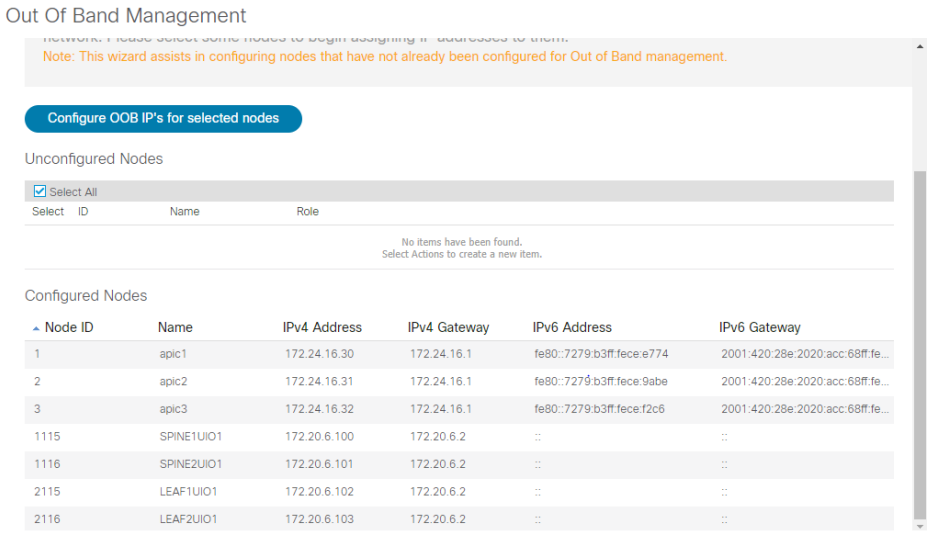


Figura 36. Configuración de IPS OOB para los dispositivos conectados al Fabric

Una vez concluida la configuración básica, se muestra un resumen de lo configurado como se puede observar en la figura 37, donde se muestra los miembros del Fabric, estado de la configuración de BGP, OOB entre otros.

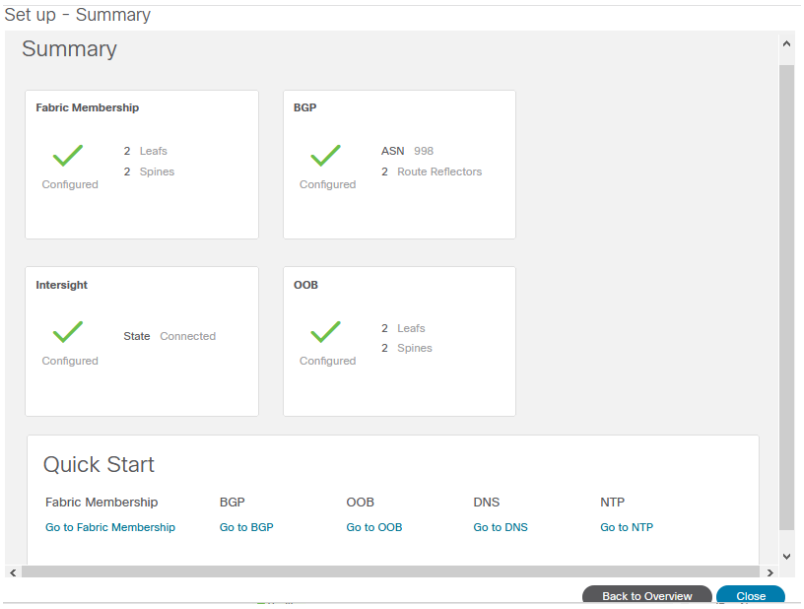


Figura 37. Resumen de la configuración básica aplicada.

Ahora, considérese verificar lo configurado hasta el momento y como ACI muestra aquello, en la figura 38 se muestra la barra de herramientas de la cual se selecciona el icono de fabric donde se muestran 3 opciones.

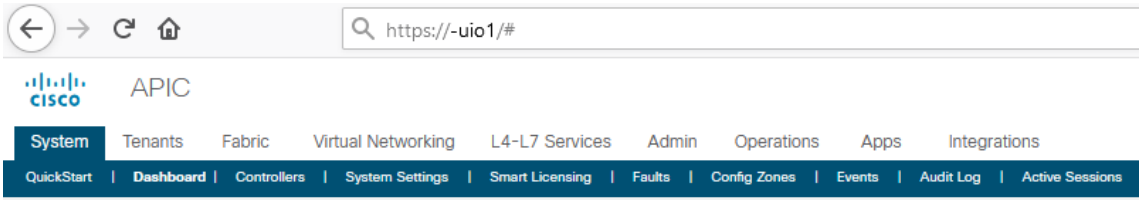


Figura 38. Barra de herramientas de ACI

Dentro de la pestaña de Fabric se escoge la opción de Inventory, donde nuevamente se ve un resumen del fabric además se muestra la topología de la solución, los puntos finales configurados, interfaces usadas entre otros como se aprecia en la figura 39, además también se tiene la opción para configuración de políticas de la fabric y para la configuración de políticas sobre los puertos.

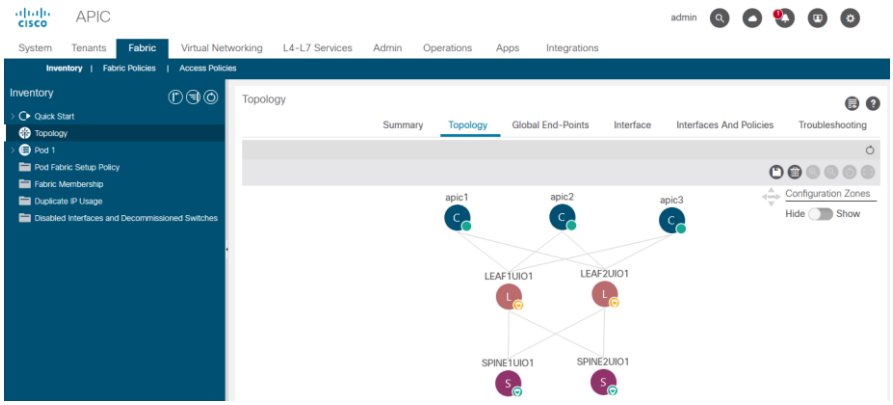


Figura 39. Topología de la Fabric de ACI implementada

4.1.2 Conexión a la red tradicional

Para la conexión a la red tradicional para propósitos de migración, es decir la conexión hacia los Nexus 7010 se tiene que crear políticas de acceso como son VLAN Pool, Dominio físico, Dominio externo L2, y un AEP que se encarga de enlazar la parte física con la parte lógica.

4.1.2.1 Pool de Vlans

En la barra de herramientas se elige Fabric, luego Access Policies, luego Pools y finalmente VLANS en la cual se da click derecho del mouse para seleccionar crear VLAN Pool como se aprecia en la figura 40, en el campo nombre se agrega un nombre representativo para el VLAN Pool, en la parte de encap block se selecciona en el icono de añadir (+) como se aprecia en la figura 41, para desplegar una nueva ventana donde en

la parte de range se irá agregando la VLAN requerida, las opciones allocation mode (inherit alloctmode from parent) y role (external) se mantendrán por defecto como se muestra en la figura 42, en base a las recomendaciones proporcionadas en el item 4.1.2.1.

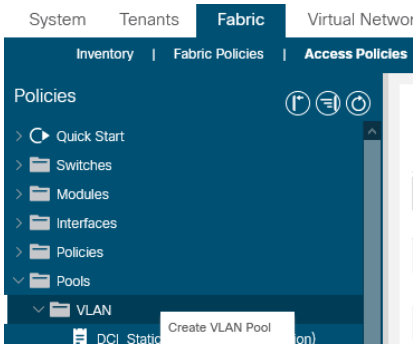


Figura 40. Ejemplo para crear un Pools de VLANS

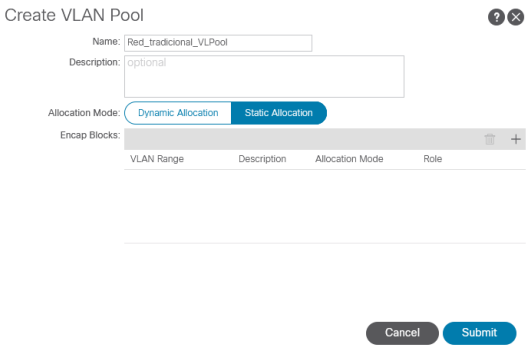


Figura 41. Creación del Pool de VLANS

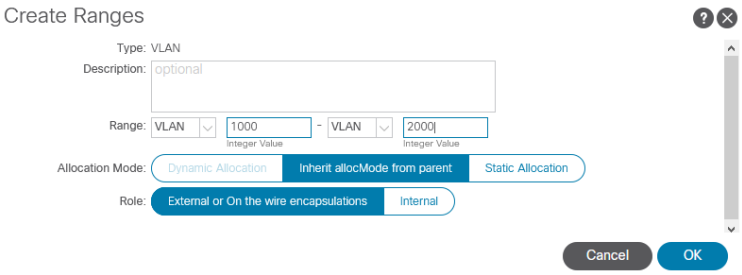


Figura 42. Ejemplo de creación de un rango de VLANS de la 1000 a la 2000

4.1.2.2 AEP

Para la creación del AEP en barra de herramientas se selecciona Fabric, después Access Policies, luego Policies, luego global y finalmente Attachable Access Entity Profiles se da click derecho con el mouse y se selecciona crear como se aprecia en la figura 43.

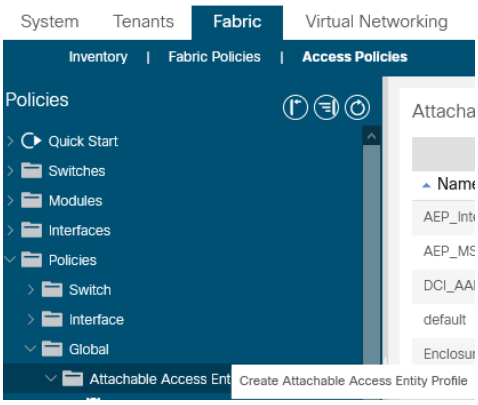


Figura 43. Creación AEP

Se despliega una pantalla en la que solo se pondrá el nombre y los demás parámetros se los deja por defecto ya que se los irá agregando conforme se los cree como son los dominios, EPG como se aprecia en la figura 44, se selecciona siguiente, se despliega la opción para asociar las interfaces al AEP como se puede ver en la figura 45, no se seleccionará ninguna interfaz se deja por defecto con la opción ninguno ya que luego se realizará las asociaciones requeridas a dicho AEP, finalmente se selecciona la opción finalizar para crearlo.

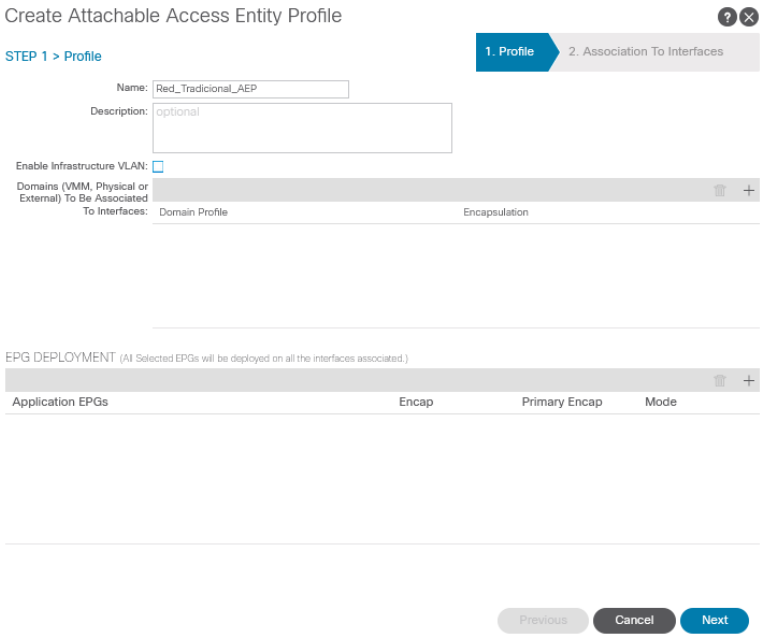


Figura 44. Configuración de un AEP

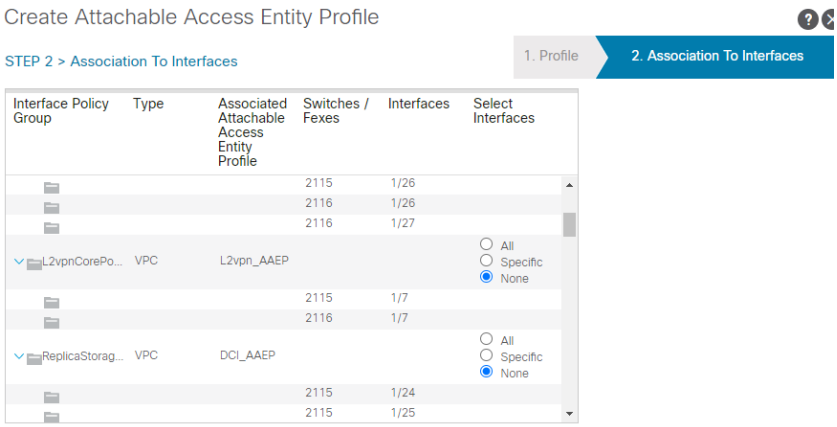


Figura 45. Asociación de las interfaces a un AEP

4.1.2.3 Dominio externo L2

Para la creación del dominio L2 externo, se selecciona en la barra de herramientas fabric, luego Access Policies, posteriormente se selecciona External Bridged Domain, se da click derecho con el mouse sobre esta opción y se selecciona crear como se observa en la figura 46.

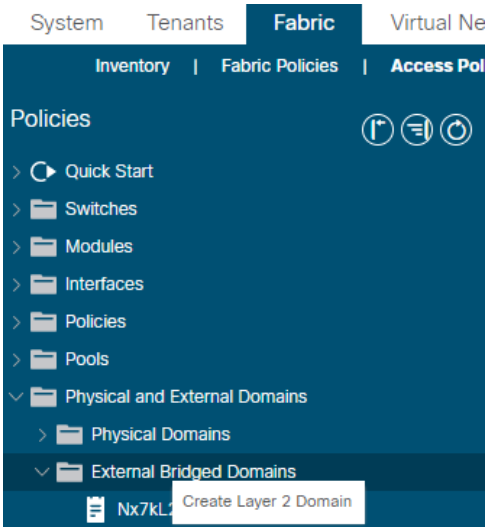


Figura 46. Creación de un dominio Externo L2

Una vez seleccionada la opción de creación se despliega una ventana para agregar el nombre, se selecciona el AEP, así como también el VLAN Pool a usar en este dominio.

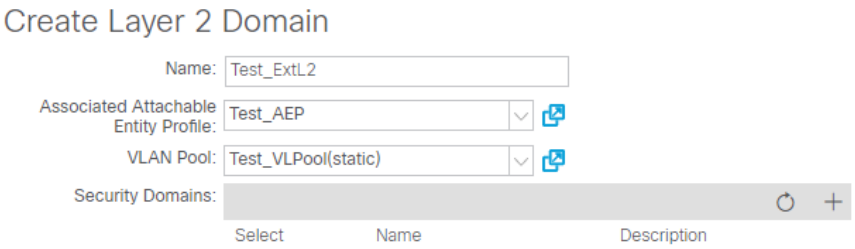


Figura 47. Creación de un dominio externo L2 y la asociación con los elementos anteriormente creados

4.1.2.4 Política de interfaz

Para la creación de las políticas de interfaz se va a ir a la barra de herramientas del fabric, posterior a esto se escoge la opción Access Policies, después se selecciona interfaz, en la cual se desplegará los diferentes tipos de políticas que se pueden configurar en las interfaces como por ejemplo la negociación del puerto (Link Level), CDP, LLDP, Port Channel entre otros como se puede observar en la figura 48.

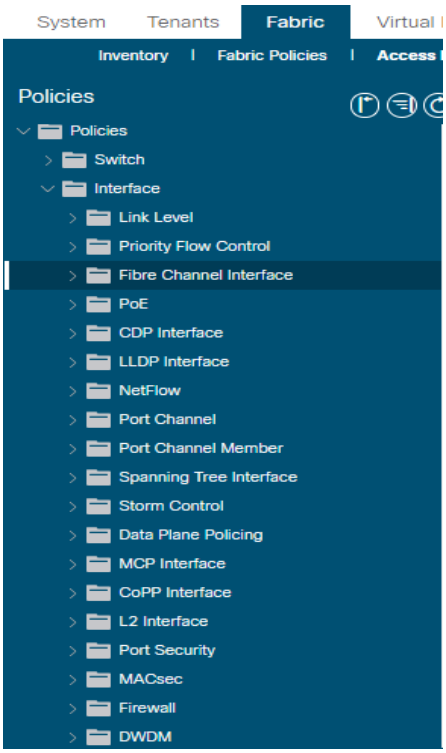


Figura 48. Plantillas de políticas de la interfaz

Para crear cada una de estas plantillas se debe dar click derecho con el mouse sobre cada una de las que se desea y seleccionar crear; para la negociación se crea una política de link Level en la cual se establece la velocidad de manera manual por lo cual la negociación

se la deja en off y se selecciona la velocidad requerida para este caso 10Gbps, no se indica ninguna descripción o alias ya que en el nombre se puede apreciar de la política que se trata, los demás parámetros como Link debounce Interval (tiempo en el cual la interface se considerará caída y no pasará tráfico) y para la corrección de errores (FEC) se mantendrán por defecto como se establece en la figura 49.

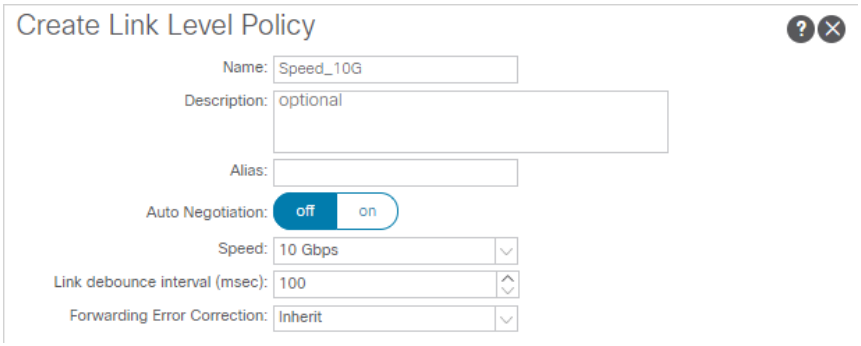


Figura 49. Creación de una política para la negociación de un puerto físico.

Para la creación de una política de CDP solo se tienen dos opciones enable y disable por lo cual se puede crear dos políticas una de enable y otro de disable como se representa en la figura 50.

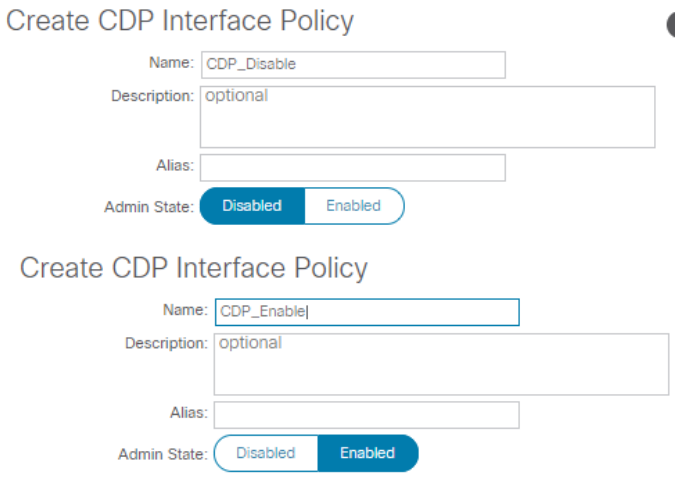


Figura 50. Políticas de CDP superior disable, inferior enable

Para la creación de una política de Port Channel se selecciona el modo de negociación de tipo LACP Active, para que se establezca el enlace agregado cuando lleguen paquetes de tipo LACP, además se mantiene todos los demás parámetros por defecto, como se representa en la figura 51.

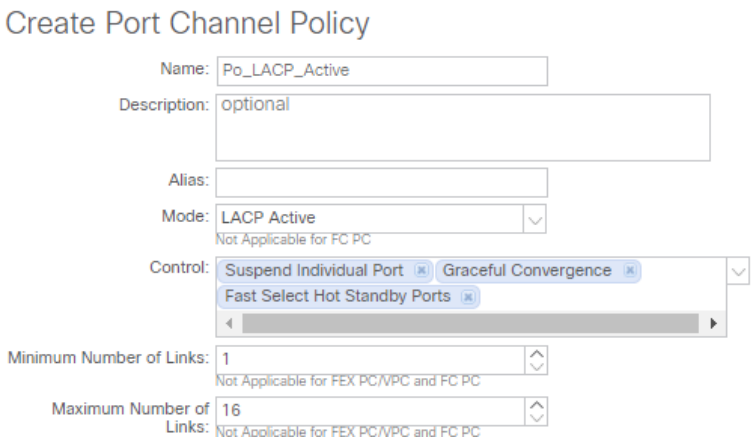


Figura 51. Creación de una Política de Port Channel modo LACP Active

4.1.2.5 Grupo de Políticas de interfaz

Para la creación del grupo de políticas de interfaz se selecciona en la barra de herramientas Fabric, luego Access Policies, luego Interfaces, posterior Leaf Interfaces y finalmente la pestaña de Policy Groups como se aprecia en la siguiente figura 52.

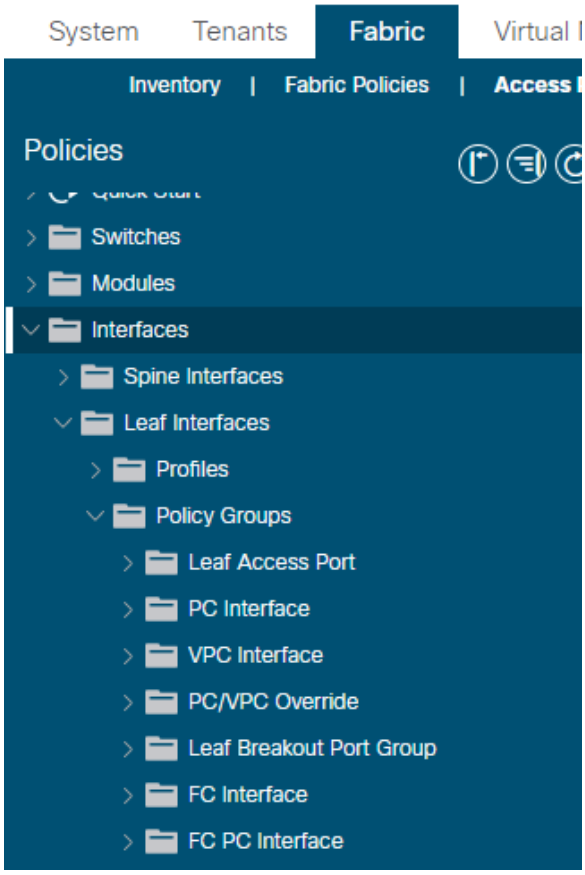


Figura 52. Grupos de Políticas de interfaz

Por otra parte para la conexión hacia la red tradicional como se ha mencionado anteriormente se lo realiza por medio de una VPC, para los clientes puede ser con una Port Channel o un puerto de acceso; para la creación del grupo de políticas para una VPC se da click derecho del mouse sobre VPC Interface y se selecciona crear; se despliega una nueva ventana como la que se aprecia en la figura 53 donde se coloca el nombre y se va seleccionando las políticas de interfaz creadas anteriormente como el link level, CDP, Port Channel y sobre todo en esta parte se asocia el AEP.

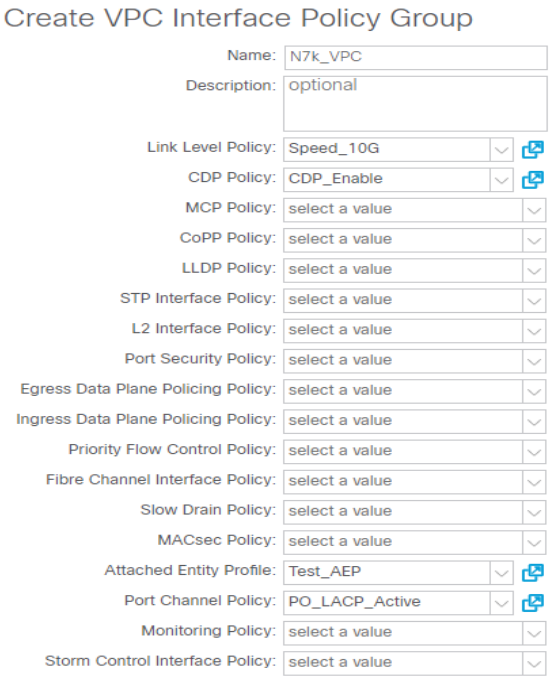


Figura 53. Creación de un grupo de políticas de interfaz VPC

Para la creación de un grupo de políticas de Interfaz para una Port Channel, se da click derecho del mouse sobre PC Interface, se selecciona crear y se asocia las políticas de interface creadas con anterioridad como se muestra en la figura 54.

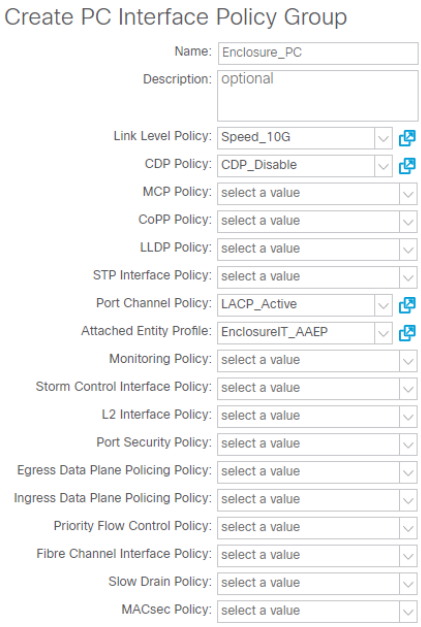


Figura 54. Creación del grupo de políticas para una Port Channel

Una vez creados los grupos de políticas estas deben ser asociadas a las interfaces físicas de los conmutadores leaf, para lo cual se debe seleccionar Fabric de las opciones de la barra de herramientas, seleccionar interface, luego seleccionar leaf interfaces y elegir la pestaña profile como se muestra en la figura 55, finalmente dar click derecho del mouse y seleccionar crear.

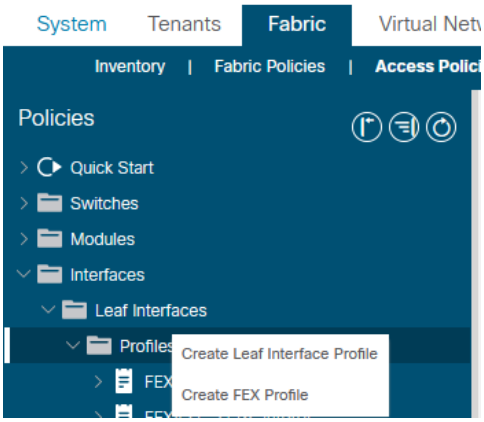


Figura 55. Creación de un perfil de interfaz para los leaf

A continuación se despliega una pestaña, en la cual se agrega el nombre el mismo que está basado en las recomendaciones del apartado 3.4.2.5 y 3.4.2.7; para asociar las interfaces que pertenecen a este perfil se selecciona el signo de añadir (+) que se encuentra al final de la opción con nombre interface selector, con lo cual se desplegará otra pestaña

en la que se agrega el nombre que hará referencia a la interfaz por ejemplo Eth1_7 se hace referencia al puerto 7, en interface ID va el identificativo del puerto tal y como se mostraría en el conmutador para este ejemplo 1/7 finalmente en interface Policy Group se asocia el grupo de políticas que se creó para este caso N7K_VPC, finalmente se da click en submit para crearlo como se aprecia en la figura 56.

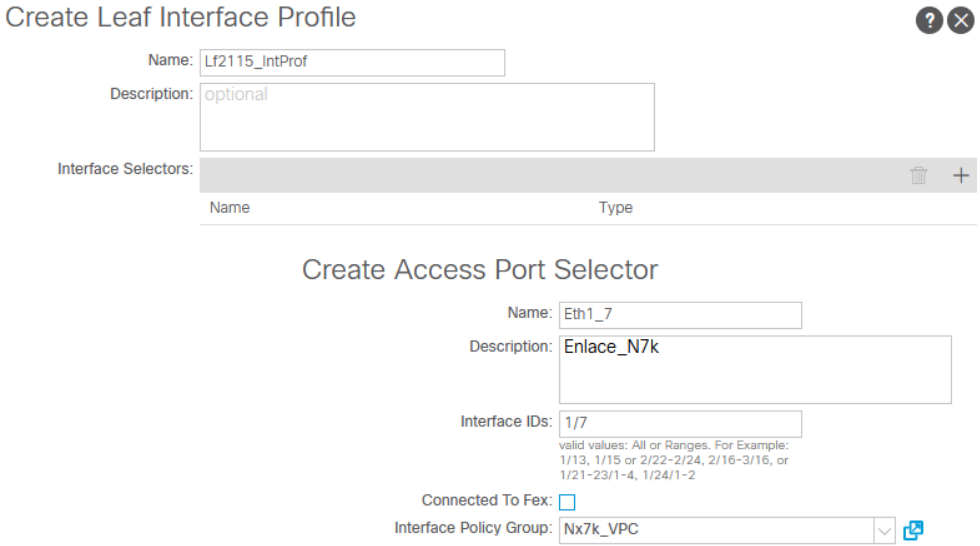


Figura 56. Creación de un perfil de interfaz y creación del selector del puerto.

Finalmente, en la figura 57 se aprecia como se asocia el perfil de interfaz creado en el leaf, para lo cual se escoge la opción de Fabric de la barra de herramientas, posterior a esto switches, luego Leaf switches, después se selecciona Profiles donde se da click derecho con el mouse para seleccionar Create Leaf Profile.

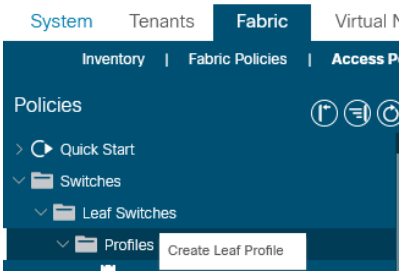


Figura 57. Creación de un perfil del Leaf parte 1

Avanzando en la creación de un perfil para el leaf y como se puede apreciar en la figura 58, se despliega una pantalla para crear el perfil en la que se coloca el nombre asociado a

cada Leaf, posterior a esto se selecciona siguiente donde se escoge el perfil de interface que se creó en el anterior paso, para este ejemplo Lf2115_IntProf.

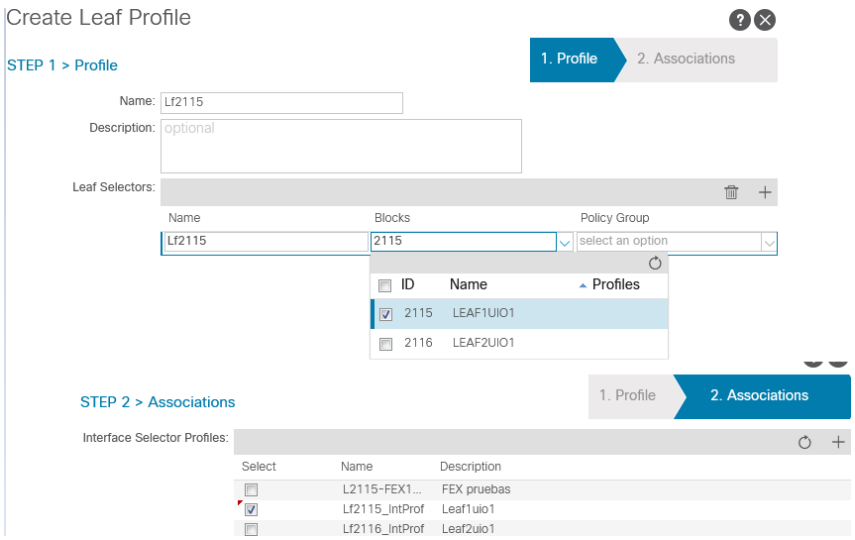


Figura 58. Creación de un perfil del Leaf parte 2

4.1.2.6 Tenan

Para la creación del Tenan se selecciona la pestaña de Tenants de la barra de herramientas, se mostrarán todos los tenans creados, en la parte derecha de la pantalla debajo de All Tenans se muestra un icono de herramientas, se lo despliega para seleccionar Create Tenan, se despliega una nueva ventana en la que se colocará únicamente el nombre que tendrá el Tenan y se selecciona submit, como se observa en la figura 59.

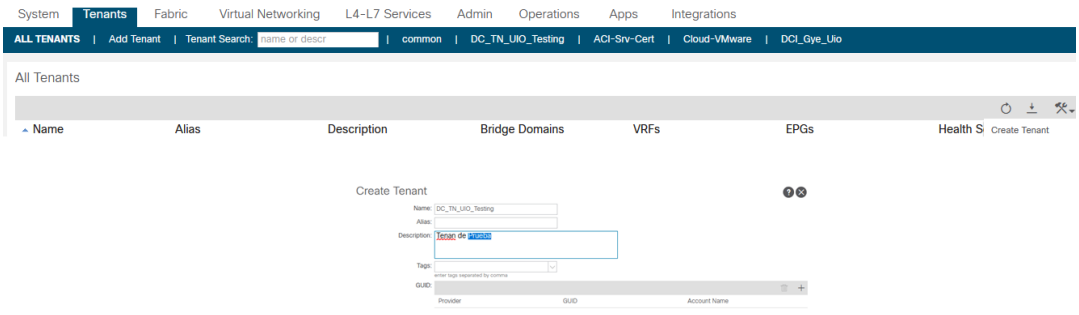


Figura 59. Creación de un Tenan

Como se puede observar en la figura 60, dentro del Tenan creado se despliega una nueva ventana, donde se debe crear el Perfil de aplicación para lo cual se escoge la opción

Application Profile y se da click derecho con el mouse para seleccionar Create Application Profile.

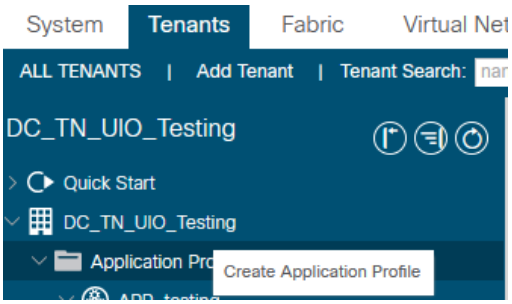


Figura 60. Creación de un perfil de aplicación parte 1

Posterior a lo indicado en el párrafo anterior, se despliega una nueva pestaña en la cual solo se agrega el nombre como se observa en la figura 61, parámetros del EPG se los seleccionará posteriormente.

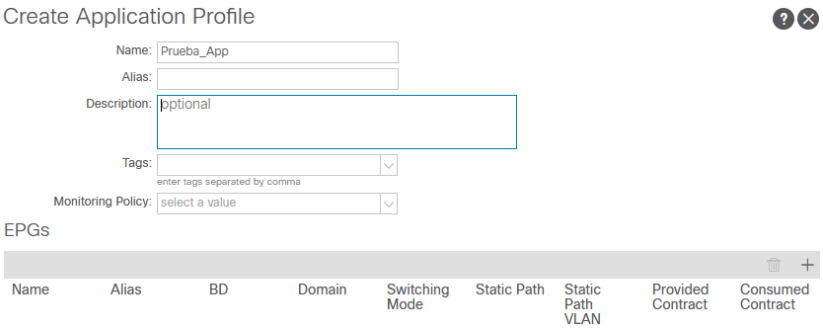


Figura 61. Creación de un perfil de Aplicación parte 2

A continuación dentro del tenan se crea la VRF para lo cual se selecciona networking, luego VRFs y se da click derecho del mouse para seleccionar crear como se muestra en la figura 62, se abrirá una nueva pestaña en la que agrega el nombre de la VRF como se aprecia en la figura 63 y se selecciona siguiente, los demás parámetros quedan por defecto ya que no se requiere configurar un tag, políticas adicionales, políticas de monitoreo, DNS, OSPF, BGP, EIGRP, debido a que se manejarán enlaces capa 2 dentro de la fabric de ACI.

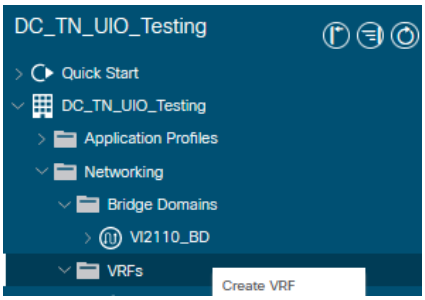


Figura 62. Creación de una VRF parte 1

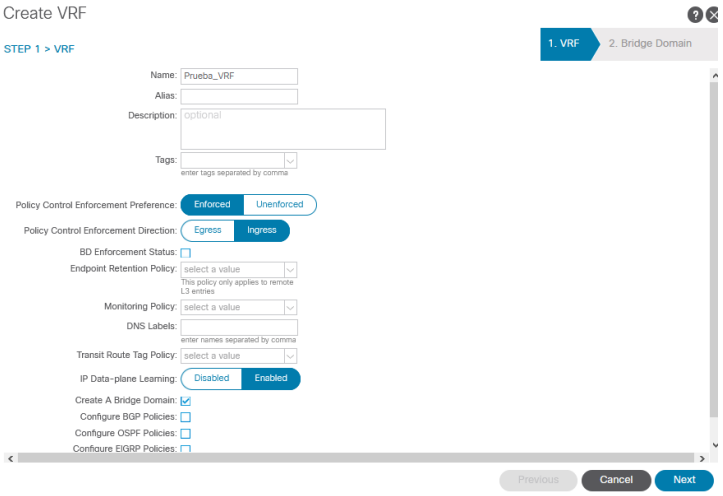


Figura 63. Creación de una Vrf parte 2

Dicho lo anterior, luego de seleccionar siguiente se despliega una nueva ventana figura 64, que solicita la creación de un Bridge Domain el mismo que está asociada a dicha VRF para lo cual solo se agrega el nombre y se selecciona finalizar, las demás opciones se mantendrán por defecto en base a las recomendaciones del apartado 3.3.14.

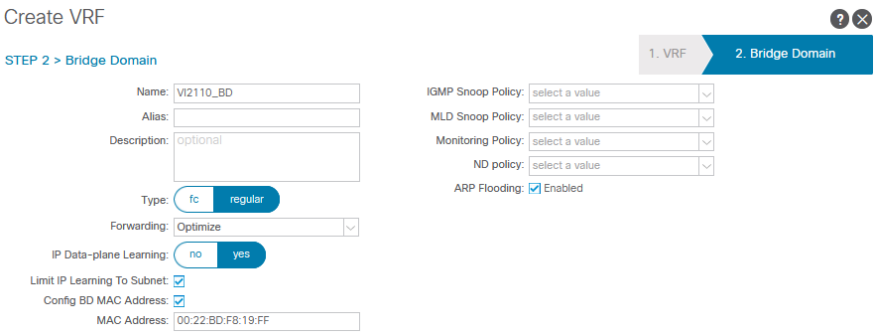


Figura 64. Creación de un Bridge Domain

Llegados a este punto luego de tener la VRF y el Bridge Domain se crea el EPG figura 64, para lo cual dentro del tenan se selección Application profile, posterior se selecciona Application EPG en el cual se da click derecho con el mouse y se selecciona crear.

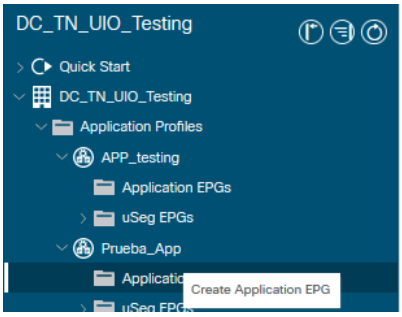


Figura 64. Creación de un EPG parte 1

A continuación, se despliega una nueva pantalla donde se agrega el nombre del EPG, y se debe escoger el Bridge Domain que se creó que en este caso es VL2110_EPG, parámetros como QoS, contratos, tag y demás se mantienen por defecto ya que no son necesarios para las aplicaciones que al momento se utilizará

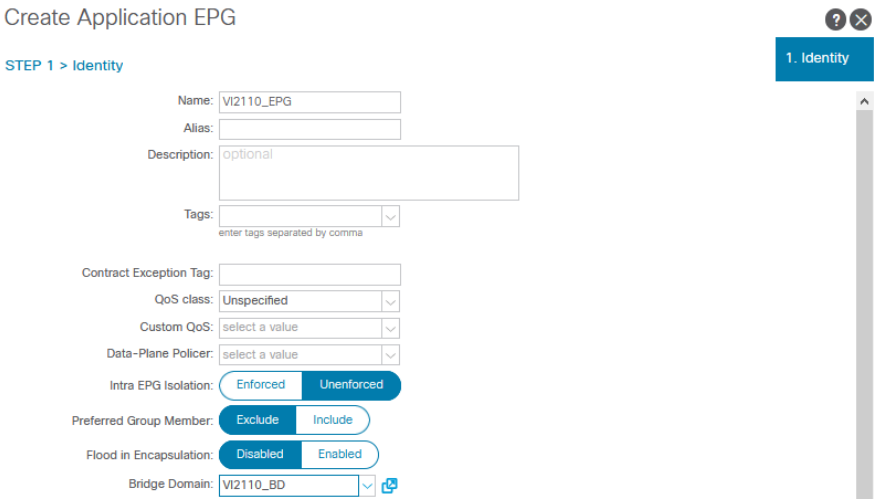


Figura 65. Creación de un EPG parte 2

4.1.2.7 EPG

Para la comunicación entre la parte lógica y física se asocia el dominio físico dentro del EPG, para lo cual se despliega las opciones del EPG y se selecciona Domains en el cual se da click derecho con el mouse y se selecciona Add L2 External Domain Association para generar la comunicación hacia la red tradicional; se abre una nueva ventana en la

que se escoge el dominio externo creado en el punto 4.1.2.3, posterior a esto se selecciona submit figura 66.

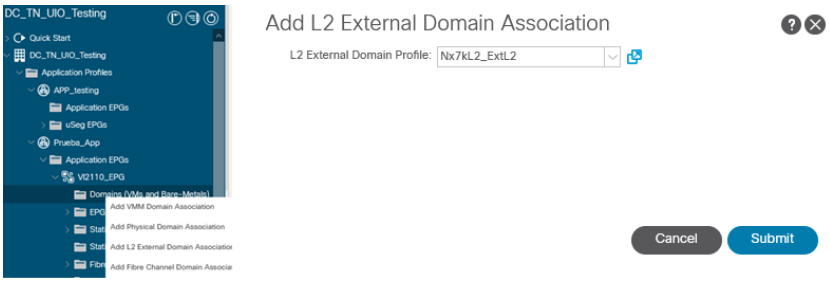


Figura 66. Asociación de un dominio físico externo

A continuación, se configura la VLAN del cliente por los puertos que debe ser propagada, para lo cual dentro del EPG se selecciona static Port, luego se da click derecho con el mouse y se escoge la opción Deploy static EPG on PC, VPC, or Interface como se aprecia en la figura 67, este paso debe repetirse por todos los puertos donde se requiere que sea propagada la VLAN.

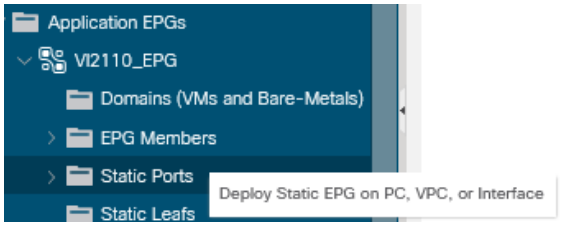


Figura 67. Asociación de la VLAN del cliente

Posterior a lo indicado se despliega una nueva ventana en la que se escoge el camino que debe cursar la VLAN, como la conexión hacia la red tradicional fue mediante una VPC se selecciona Virtual Port Channel, en el path se selecciona N7K_VPC, en port Encap se coloca la VLAN y el modo de operación de esta conexión es troncal por lo cual se selecciona trunk, y finalmente se aplica la configuración con submit como se aprecia en la figura 81.

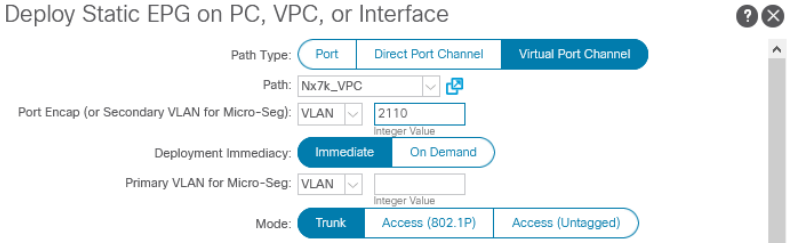


Figura 68. Creación del puerto Estático

4.1.2.7 Configuración opcional de los puertos

Se ha configurado los puertos de manera separada es decir cada elemento por separado sin embargo en ACI también es posible realizar la configuración de una manera más fácil, para lo cual se escoge fabric de la barra de herramientas, luego se selecciona Access Policies, posterior a esto se selecciona Interfaces And Policies como se muestra en la figura 69 donde se debe seleccionar el nodo en el cual se ejecuta la configuración.

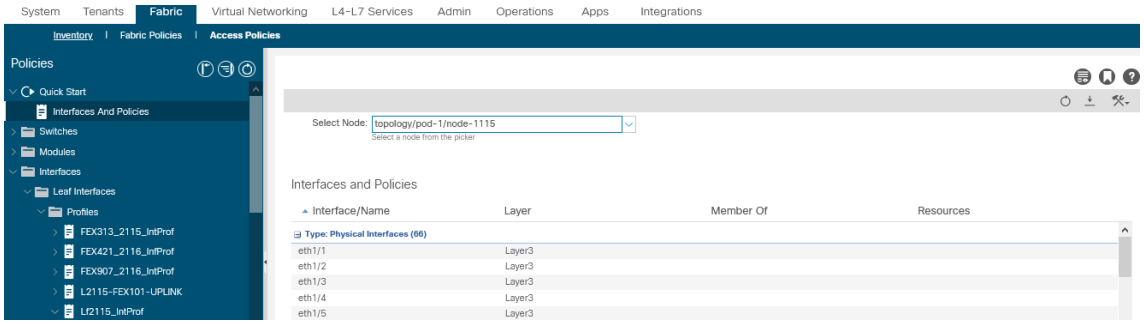


Figura 69. Configuración de un puerto parte 1

Luego como se aprecia en la figura 70, se selecciona el puerto y se escoge la opción de configuración (icono de herramientas), se desplegará una pantalla en la que se escogerá nuevamente el leaf donde se ejecuta la configuración, la interface y los perfiles de cada uno, se debe escoger el tipo de interface a configurar y el grupo de Políticas que tendrá dicha interface.

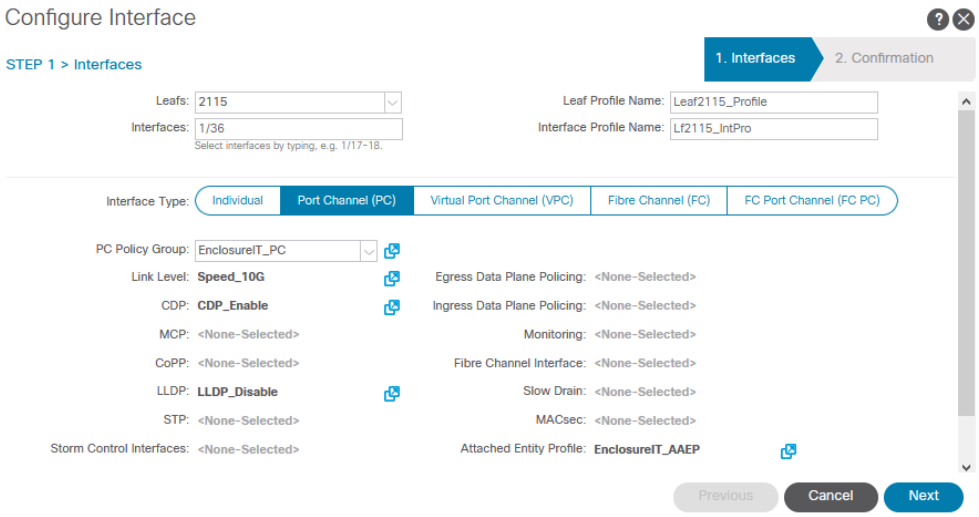


Figura 70. Configuración de un puerto parte 2

Como se pudo observar esta implementación requirió de la instalación y despliegue de la solución, en la cual se instaló, configuró y levanto el clúster del controlador, se realizó la asociación de los conmutadores Leaf y Spine hacia el fabric, se estableció parámetros y recomendaciones para la configuración de objetos, se realizó la comunicación entre la red tradicional y el Fabric de ACI.

4.2 Evaluación de resultados

4.2.1 Pruebas de funcionamiento de la red SDN.

En la figura 71 se observa que se realizó la consulta del estado del controlador por medio de CLI ingresando a uno de los 3 controladores, en el cual se puede ver los 3 controladores, su salud, direccionamiento de cada uno.

```
apic1# show controller
Fabric Name          : FABRIC UIO1
Operational Size     : 3
Cluster Size         : 3
Time Difference      : -25508356
Fabric Security Mode : PERMISSIVE

 ID    Pod   Address          In-Band IPv4     In-Band IPv6             OOB IPv4        OOB IPv6                        Version           Flags  Serial Number    Health
 ----  ----  ---------------  ---------------  -----------------------  --------------  ------------------------------  ----------------  -----  ---------------  ------------
 1*    1     10.181.0.1       10.108.25.5      fc00::1                  172.24.16.30    fe80::7279:b3ff:fece:e774       4.2(4o)           crva-  FCH2215V16C      fully-fit
 2     1     10.181.0.2       10.108.25.6      fc00::1                  172.24.16.31    fe80::7279:b3ff:fece:9abe       4.2(4o)           crva-  FCH2215V198      fully-fit
 3     1     10.181.0.3       10.108.25.7      fc00::1                  172.24.16.32    fe80::7279:b3ff:fece:f2c6       4.2(4o)           crva-  FCH2215V19A      fully-fit

Flags - c:Commissioned | r:Registered | v:Valid Certificate | a:Approved | f/s:Failover fail/success
(*)Current (~)Standby (+)AS
```

Figura 71. Consulta de los controladores por CLI

Adicional mediante el comando show running-config System de la figura 72, se observa el tamaño del clúster implementado y cada uno de los nodos asociados al fabric con su respectivo ID de nodo, serial y el nombre que lo identifica.

```
apic1# show running-config system
# Command: show running-config system
# Time: Sun Feb 28 11:03:40 2021
  system cluster-size 3
  system switch-id FDO22090C72 1116 SPINE2UIO1 pod 1
  system switch-id FDO22121TAU 2116 LEAF2UIO1 pod 1
  system switch-id FDO22121TBX 2115 LEAF1UIO1 pod 1
  system switch-id FDO22171A1P 1115 SPINE1UIO1 pod 1
  # system pod 1
  system pod 1 tep-pool 10.181.0.0/16
```

Figura 72. Salida de la configuración del POD1

Por otra parte, y de manera similar también se puede observar por medio de la GUI del clúster como se muestra en la figura 73, donde se valida el serial, nombre, Id de nodo, direccionamiento y estado de los conmutadores asociados al fabric.

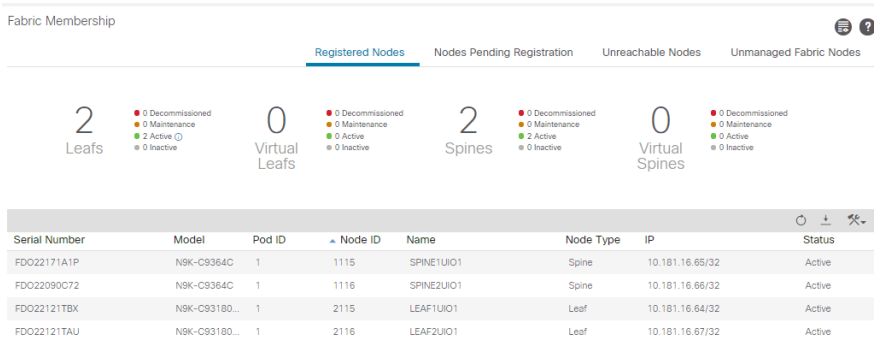


Figura 73. Miembros del Fabric de ACI

Para las pruebas de conectividad entre la red tradicional y ACI se establece el uso de la VLAN 2110 para pruebas con la subred privada 172.24.7.0/29, la misma que estará configurada en un equipo de Core de la Red como se muestra en la figura 74 y para la comunicación se lo hará mediante la conexión de capa 2 que se tiene entre los conmutadores Lead y los Nexus 7010.

```
Bundle-Ether9.2110 is up, line protocol is up
  Interface state transitions: 1
  Hardware is VLAN sub-interface(s), address is b026.802c.a681
  Description: Pruebas_ACI_CC
interface Bundle-Ether9.2110
 description Pruebas_ACI_CC
 mtu 1560
 vrf acifabric_test
 ipv4 address 172.24.7.1 255.255.255.248
```

Figura 74. Red de pruebas ACI – Red Tradicional

En el GUI de ACI en la VLAN 2110 creada en el Tenan de Pruebas al escoger la opción de Operacional, seleccionando Cliente End-Point se observa que la MAC de la interfaz del equipo de Core, se observa por cual interface se está aprendiendo dicha VLAN y como tal la VLAN que se está aprendiendo, como se observa en la figura 75.

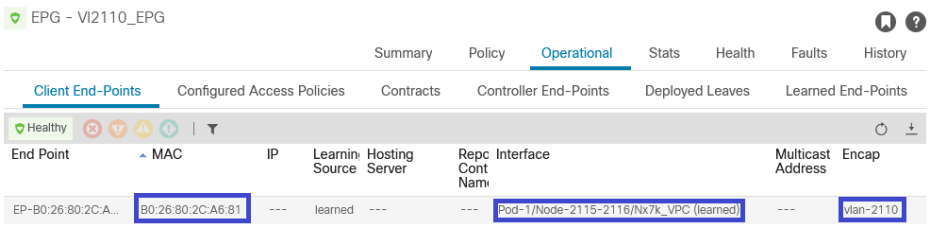


Figura 75. Pruebas de conectividad parte 1

Ahora se configura la IP 172.24.7.2/29 perteneciente a la subred de pruebas en el ACI dentro del Bridge Domain de pruebas para verificar conectividad de extremo a extremo por medio de ping, como se observa en la figura 76:

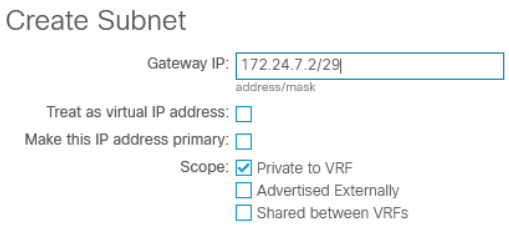


Figura 76. Configuración red de pruebas en ACI

Para probar la conectividad se realiza un ping desde el equipo de Core hacia la IP 172.24.7.2 que es la IP configurada en ACI, al inicio se tiene un paquete perdido que se da hasta que se realice el descubrimiento en L2 del camino, como se muestra en la figura 77.

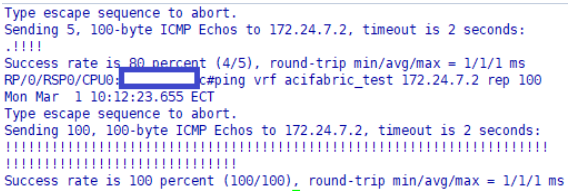


Figura 77. Pruebas de Conectividad parte 2

Como resultado tanto en ACI como en el equipo de Core a nivel de ARP se puede ver la MAC y la IP asociada que se aprende en ambos equipos, como se observa en la figura 78.

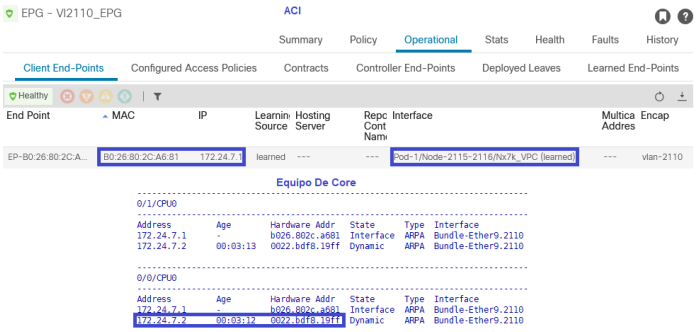


Figura 78. Pruebas de conectividad parte 3

A continuación se hará la prueba de conectividad entre un equipo de Cliente y la Fabric donde su Gateway estará en el ACI, como primer punto se debe crear la conexión entre el dispositivo y los leaf, para lo cual se debe a ir a Fabric, luego Interfaces, luego Policy Groups aquí en cambio se dará click derecho con el mouse y se selecciona crear y para lo cual se debe llenar los parámetros de Link Level a 1G, CDP enable, STP interface Policy

STP BPU Filter (para filtrar los paquetes BPDU que puedan llegar del cliente) y finamente se selecciona el AEP como se aprecia en la figura 79.

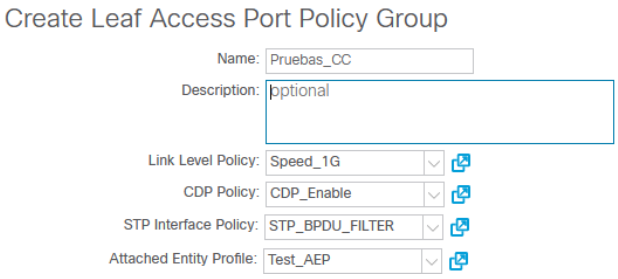


Figura 79. Configuración de políticas de grupo cliente de prueba

Luego se asocia el grupo de interfaz al puerto del Leaf en el cual está conectado el cliente de prueba, como se aprecia en la figura 80.

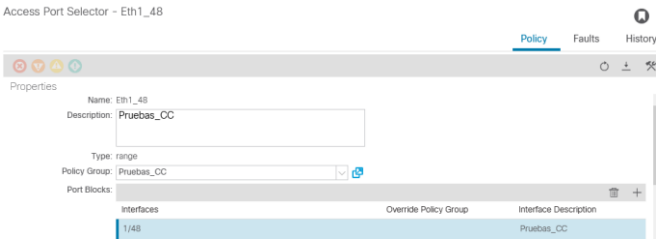


Figura 80. Asociación de la política de grupo Pruebas_CC al puerto 48

Ahora en el tenan de Pruebas en el EPG llamado VL2110_EPG se asocia el dominio físico, como se observa en la figura 81.

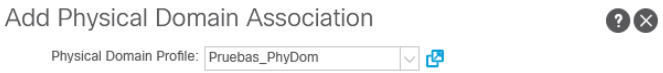


Figura 81. Asociación del dominio físico en el EPG de prueba

Finalmente se agrega el camino hacia el nuevo dispositivo en el static port del EPG con lo cual ya se puede hacer pruebas de conectividad, como se aprecia en la figura 82, donde se coloca el puerto 48 asociado a la VLAN 2110 del equipo Leaf2.

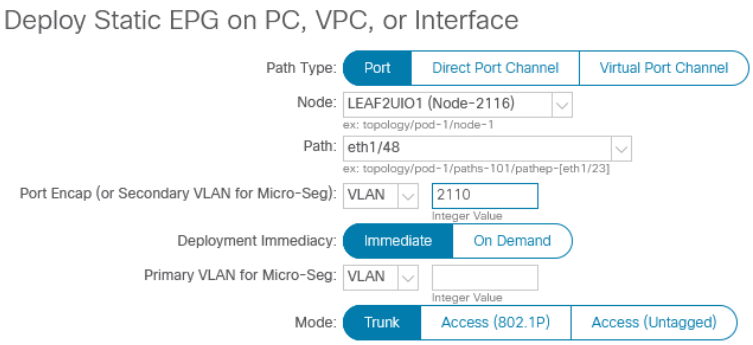


Figura 82. Configuración static port

Se realiza ping desde el equipo de pruebas del cliente con IP 172.24.7.4 hacia la IP que está en el ACI que es la 172.24.7.2 se observa que se tiene respuesta, además se realiza otra prueba de ping hacia la IP 172.24.7.1 que se encuentra en la red tradicional validando que también se tiene respuesta.

```
[redacted]#ping 172.24.7.2

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.24.7.2, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/2/9 ms
[redacted]#ping 172.24.7.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 172.24.7.1, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
[redacted]#sh ip ar
[redacted]#sh ip arp | i 2110
Internet  172.24.7.2             19   0022.bdf8.19ff  ARPA   Vlan2110
Internet  172.24.7.1              0   b026.802c.a681  ARPA   Vlan2110
Internet  172.24.7.4              -   001e.7a98.88c3  ARPA   Vlan2110
```

Figura 83. Pruebas de ping hacia la red tradicional y cliente

A nivel del GUI de ACI de igual modo se puede validar las direcciones MAC que se están aprendiendo y la interfaz por la cual se está aprendiendo dicha IP y VLAN, como se aprecia en la figura 84.

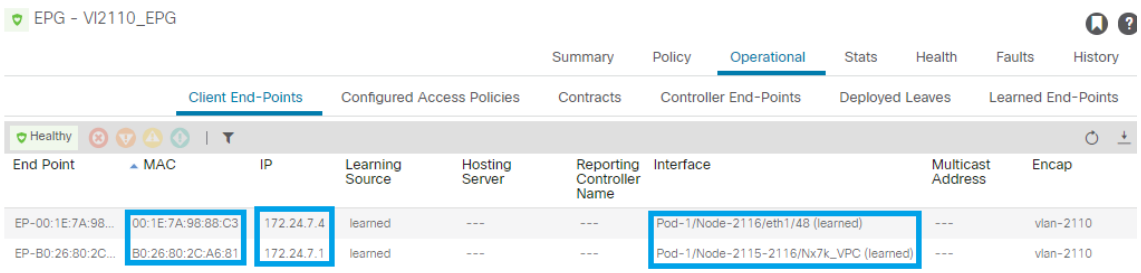
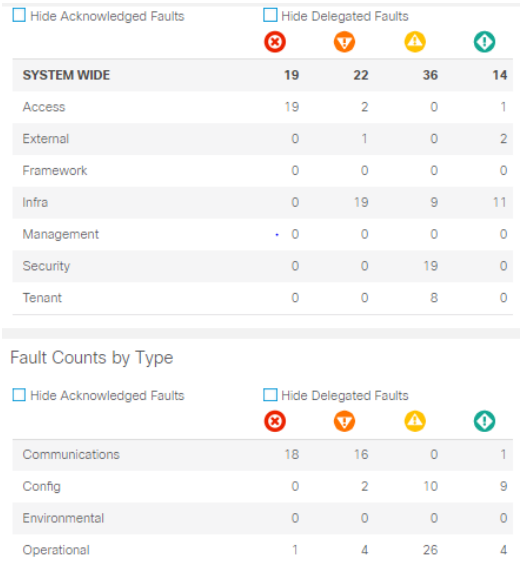


Figura 84. Validación IP, MAC, VLAN 2110 GUI ACI

Por medio ACI se puede ver las fallas que se han producido en la plataforma de manera clasificada en base a una prioridad como se aprecia en la figura 85 donde el color verde hace referencia a una falla de advertencia, color amarillo a una falla menor, color naranja una alarma mayor y de color rojo una alarma crítica.



Hide Acknowledged Faults Hide Delegated Faults

SYSTEM WIDE	**19**	**22**	**36**	**14**
Access	19	2	0	1
External	0	1	0	2
Framework	0	0	0	0
Infra	0	19	9	11
Management	0	0	0	0
Security	0	0	19	0
Tenant	0	0	8	0

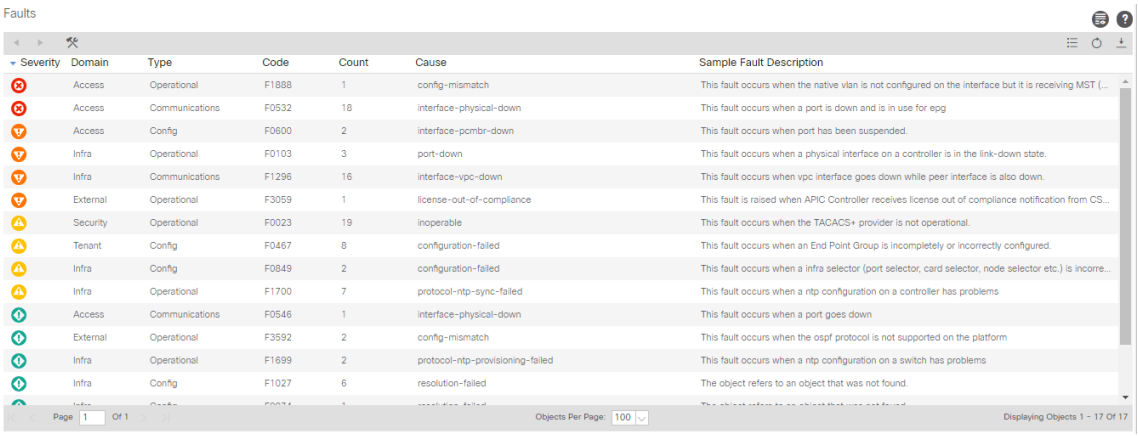
Fault Counts by Type

Hide Acknowledged Faults Hide Delegated Faults

Communications	18	16	0	1
Config	0	2	10	9
Environmental	0	0	0	0
Operational	1	4	26	4

Figura 85. Fallas desplegadas en el dashboard de los API

Además, también se puede ver con mayor detalle cada una de las fallas que se van produciendo dentro del Fabric de ACI figura 86, donde cada una está clasificada por color en base a su criticidad.



Faults

Severity	Domain	Type	Code	Count	Cause	Sample Fault Description
	Access	Operational	F1888	1	config-mismatch	This fault occurs when the native vlan is not configured on the interface but it is receiving MST (...
	Access	Communications	F0532	18	interface-physical-down	This fault occurs when a port is down and is in use for epg
	Access	Config	F0600	2	interface-pcmbr-down	This fault occurs when port has been suspended.
	Infra	Operational	F0103	3	port-down	This fault occurs when a physical interface on a controller is in the link-down state.
	Infra	Communications	F1296	16	interface-vpc-down	This fault occurs when vpc interface goes down while peer interface is also down.
	External	Operational	F3059	1	license-out-of-compliance	This fault is raised when APIC Controller receives license out of compliance notification from CS...
	Security	Operational	F0023	19	inoperable	This fault occurs when the TACACS+ provider is not operational.
	Tenant	Config	F0467	8	configuration-failed	This fault occurs when an End Point Group is incompletely or incorrectly configured.
	Infra	Config	F0849	2	configuration-failed	This fault occurs when a infra selector (port selector, card selector, node selector etc.) is incorre...
	Infra	Operational	F1700	7	protocol-ntp-sync-failed	This fault occurs when a ntp configuration on a controller has problems
	Access	Communications	F0546	1	interface-physical-down	This fault occurs when a port goes down
	External	Operational	F3592	2	config-mismatch	This fault occurs when the ospf protocol is not supported on the platform
	Infra	Operational	F1699	2	protocol-ntp-provisioning-failed	This fault occurs when a ntp configuration on a switch has problems
	Infra	Config	F1027	6	resolution-failed	The object refers to an object that was not found.

Page 1 Of 1 Objects Per Page: 100 Displaying Objects 1 - 17 Of 17

Figura 86. Detalle de todas las fallas producidas en la Fabric

Además, seleccionando e ingresando a cada una de las fallas se despliega con mayor detalla cual es la razón del error generado e incluso una pestaña de troubleshooting donde

se indica también una explicación del error y la posible solución a la falla, un ejemplo de lo indicado se tiene en la figura 87.

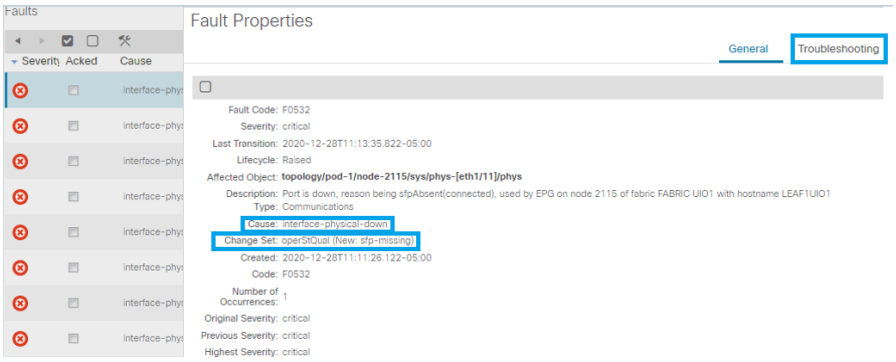


Figura 87. Detalles de una falla producida en la Fabric de ACI

4.2.2 Identificación de mejoras.

Si bien el trabajo no establece la creación de un script pienso que la generación de estos puede ayudar en la automatización para objetos claves de la operación de una manera mucho más eficiente o talvez como ejemplos para la creación de aplicaciones externas que usarán la API; ya que cada vez que se ejecuta alguna tarea en el GUI de ACI ya sea esta una consulta, la creación, eliminación, modificación de un objeto dentro de ACI se envía un mensaje interno de la acción que fue ejecutada y mediante un módulo de ACI llamado API inspector se observar ver el mensaje generado.

Se puede tener una aplicación interesante usando la infraestructura de ACI, como es el doble etiquetado de VLANS conocido como Q-in-Q, es decir en la cabecera tendrá 2 VLANS para etiquetar el tráfico y cuando entre a la fabric estas 2 etiquetas (VLANS) son procesados de manera independiente y son restaurados cuando salen de la Fabric, esta aplicación puede ser usada para realizar replicación entre 2 sitios, ya que un cliente al conectarse por medio de este túnel estaría extendiendo su conexión LAN lo que le permitirá hacer una replicación de una DB por ejemplo, además el cliente puede pasar las VLANS que desee, puede ser una por aplicación si lo requiere mientras que para el proveedor es transparente y solo verá una VLAN, el uso de esta aplicación tiene la limitante que no está soportado en puertos de Fex por lo cual se incrementaría el uso de los puertos en los Leaf lo que conlleva a la adquisición de más conmutadores.

CONCLUSIONES

Mediante este estudio y prueba de concepto se logró verificar que la solución SDN de ACI permite tener una visión central de toda la arquitectura de red desde el controlador, en él se pudo ver el estado de los equipos, configuraciones, fallas en tiempo real, alarmas, además permite tener una mejor gestión de la red, más rápida y oportuna resolución de problemas, además de ser una gran ayuda al momento de hacer actualizaciones en los equipos y como tal en toda la solución de red.

Mediante el levantamiento de información realizado se pudo evidenciar que la capacidad de los enlaces está limitada a conexiones de 10 Gbps y que físicamente es imposible crecer a capacidades de 100 Gbps debido a que el hardware actual no está diseñado para ofrecer estas velocidades y si se lo hace seria mediante enlaces agregados que evidentemente implica mayor número de cables, más interfaces, mayor cantidad de slots para las conexiones y evidentemente equipos de mayor tamaño como el actual cisco Nexus 7010, además se hace uso del protocolo MST para evitar lazos de capa 2 y que para servicios de housing implica un gran problema cuando el cliente conecta su equipo a la red del Centro de Datos.

Con el estudio realizado se pudo determinar las características y componentes de hardware que deben ser consideradas para la implementación de la arquitectura basada en Redes Definidas por software para Centros de Datos en la cual se debe tener conmutadores con gran densidad de puertos de 1,10 y hasta 100 Gbps como se establece en el análisis en el apartado 3.1.

De acuerdo con el análisis realizado se definieron y se recomendaron varias políticas, parámetros lógicos de configuración que a la vez pueden funcionar como plantillas para la implementación de servicios para la solución de SDN ACI de Cisco, sobre todo para escenarios donde luego de la implementación se requiere realizar la migración hacia la nueva tecnología por medio de un enlace L2 Externo.

Se realizó el diseño de la topología a usar en la solución SDN en la cual su arquitectura está basada en conmutadores Leaf y Spine, que a su vez tienen conexión a la red tradicional para temas de migración de servicios hacia la nueva infraestructura y se comprobó que toda la solución es administrada por un cluster de controlador constituido

por 3 servidores llamados APIC y en el cual se tiene administración, gestión y control de toda la Fabric de ACI.

Con el despliegue de la solución ACI de Cisco para centros de Datos se puede observar que la solución de SDN provee varias ventajas como se indica en el punto 3.2 y sobre todo de manera más explícita en el punto 3.5 al ser automática, escalable, simple de operar además la posibilidad de integrarse con varias herramientas para cambiar la forma en la que se gestionan y administran las redes permitiendo con un solo controlador central automatizar tareas y procedimientos habituales, abriendo un sinfín de posibilidades para nuevos servicios, implementaciones, una mejor gestión de la red, más rápida y oportuna resolución de problemas, además se pudo verificar que es una gran ayuda al momento de hacer actualizaciones en los equipos y como tal en toda la solución de red ya que el sistema operativo se lo carga solo en el controlador y este se encarga de subirlos a cada equipo y demás tareas al momento de la actualización.

Finalmente es valioso comentar que el despliegue inicial de la solución de ACI fue bastante rápido e intuitivo, no tomo más allá de 20 minutos por cada controlador APIC ya que prácticamente se configura todo el clúster por medio de scripts preconfigurados solo con el ingreso de ciertos parámetros que al inicio el script mismo va solicitando, como nombres, ips, entre otros, además el agregar o realizar un cambio de conmutador leaf es bastante fácil ya que luego del despliegue se tuvo una falla en uno de los leaf y al cambiarlo fue una tarea sin mayor complicación y que no afecto en nada la operación de los otros equipos de la red.

RECOMENDACIONES

Posterior a la ejecución del presente estudio, se recomienda implementar scripts para la implementación con API’s que permitan validar las políticas de red y resolución de problemas.

Se sugiere realizar un análisis de la integración de ACI con máquinas virtuales y la aplicación de la microsegmentación simulando por ejemplo un virus en una maquina Virutal y como ésta puede ser aislada.

Hay que tomar en consideración que este estudio está enfocado en servicios multiplataforma y de acceso masivo sin embargo se recomienda realizar un análisis de clientes en base a sus necesidades para conectarlos a la nueva infraestructura de red con ACI.

BIBLIOGRAFÍA

Mallick, A. (2019). *The Traditional Network Infrastructure Model and Issues Associated with it*. https://www.pluribusnetworks.com/blog/traditional-network-infrastructure-model-and-problems-associated-with-it/

Aci, C., Aci, C., Aci, C., Aci, C., Services, C., Leaf, A. C. I. R., Pod, A. C. I. V., Services, C., Fabric, A. C. I., Clouds, P., & Platform, G. C. (2019). *Cisco Services for Cisco Application Centric Infrastructure*.

Adolfo Manaure. (2015). *Fallas en el data center: 95% son causadas por errores humanos*. Marzo. https://thestandardcio.com/2015/03/25/data-center-95-son-causadas-por-errores-humanos/

Ali, T. E., Abdala, M. A., & Morad, A. H. (2019). Sdn implementation in data center network. *Journal of Communications*, *14*(3), 223–228. https://doi.org/10.12720/jcm.14.3.223-228

Arista. (2016). *Software Defined Cloud Networking Introduction*. 1–8.

Bruno Anthony, J. S. (2017). *CCDA Cisco*. https://b-ok.cc/?signAll=1

Chen, T., Gao, X., & Chen, G. (2016). The features, hardware, and architectures of data center networks: A survey. *Journal of Parallel and Distributed Computing*, *96*, 45–74. https://doi.org/10.1016/j.jpdc.2016.05.009

Cisco. (2009). *Cisco Nexus 9300 Series Switches*. *Figure 1*, 1–23.

Cisco. (2013a). *Cisco Nexus 7000 M2-Series 2-Port 100 Gigabit Ethernet Module with XL Option*. *Figure 1*, 1–7.

Cisco. (2013b). *Principios de la infraestructura centrada en aplicaciones*. *figura 1*, 1–8.

Cisco. (2013c). *Principios de la infraestructura centrada en aplicaciones*.

Cisco. (2014). *Cisco Application Policy Infrastructure Controller Enterprise Module*. 1–2.

Cisco. (2018). *Using Cisco ACI in Telecom Data Centers to Enhance Automation , Service Chaining , Scalability , Operational Simplification , Troubleshooting , and provide Consistent Policy across any location*. https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-740717.html

Cisco. (2019). *Cisco ACI Multi-Site Architecture*. 1–111.

Cisco, P. (2013d). *Cisco Nexus 7000 Series Switches Environment*. *Figure 1*, 1–11.

Cisco Public. (2016). Cisco Application Centric Infrastructure Fundamentals. *Manual*. http://www.cisco.com/c/dam/en/us/products/collateral/switches/nexus-9000-series-switches/at-a-glance-c45-730001.pdf

Cisco Systems. (2018). *Cisco Application Centric Infrastructure - Cisco*. *6387*, 1–137. https://www.cisco.com/c/en_uk/solutions/data-center-virtualization/application-centric-infrastructure/index.html

Dagenhardt, F., & Moreno, Jose Dufresne, B. (2012). Deploying ACI. In *Ciscopress*. http://www.eltexampreparation.com/sites/default/files/The Mini Guide To IELTS_0.pdf

Eclassvirtual. (2020.). *Arquitectura de topologías de redes CCNA 200-301*. https://eclassvirtual.com/arquitectura-de-topologias-de-redes-ccna-200-301/

Firas Ahmed, S. M. (2020). *Implementing Data Center Overlay Protocols*. Marzo. https://www.ciscopress.com/articles/article.asp?p=2999385&seqNum=3

GBM. (n.d.). *Simplifique y acelere radicalmente la integración de las aplicaciones*. https://www.gbm.net/es/cisco-aci

He, W. T. (2018). Introduction to ACI. *Cisco Live*.

Hein, D. (2020). *Gartner's 2020 Magic Quadrant for Data Center and Cloud Networking: Key Takeaways*. 1. https://solutionsreview.com/network-monitoring/gartners-2020-magic-quadrant-for-data-center-and-cloud-networking-key-takeaways/

Huawei. (2019). *What Is VXLAN*. https://support.huawei.com/enterprise/en/doc/EDOC1100086966

HUAWEI TECHNOLOGIES CO., L. (2020). *What Is VXLAN?* Junio. https://support.huawei.com/enterprise/en/doc/EDOC1100086966#EN-US_TOPIC_0259820545

Lee, G. (2014). Cloud Networking. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).

Lerner, Andrew Zeng, E. F. J. (2020). *Magic Quadrant for Data Center and Cloud Networking*. https://www.gartner.com/doc/reprints?id=1-1Z4QHTAB&ct=200529&st=sb

Lerner, A. (2020). *Cisco data center and cloud networking recognitions*. https://www.networkworld.com/article/3586800/cisco-data-center-and-cloud-networking-recognitions.html?page=2

Mallick, A. (2019). *The Traditional Network Infrastructure Model and Issues Associated with it*. https://www.pluribusnetworks.com/blog/traditional-network-infrastructure-model-and-problems-associated-with-it/

ONF. (2020). *opennetworking*. https://opennetworking.org/sdn-definition/

Peterson Larry, Cascone Carmelo, O'Connor Brian, V. T. (2020). *Open Networking Foundation*. sdn.systemsapproach.org

Services, I. (2019). *Comparación de redes de Cisco con la competencia*.

Spera, C. (2013). *Software Defined Network: el futuro de las arquitecturas de red*. 42–45.

Studios, Sd. (2015). *Understanding the SDN Architecture - SDN Control Plane & SDN Data Plane*. Marzo. https://www.sdxcentral.com/networking/sdn/definitions/inside-sdn-architecture/#:~:text=A software-defined network (SDN,plane of the networking stack

System, C. (2019). *Cisco Nexus 9332C and 9364C Fixed Spine Switches*. https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/datasheet-c78-739886.pdf

Systems, C. (2019). *Cisco Application Centric Infrastructure Fundamentals, Releases 2.x and 3.x*. Julio. https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/aci-fundamentals/b_ACI-Fundamentals/b_ACI-Fundamentals_appendix_01111.html

TADVISER. (2018). *Huawei Cloud Fabric*. https://tadviser.com/index.php/Product:Huawei_Cloud_Fabric

Tejedor, R. J. M. (2016). *Qué es... SDN (Software-Defined Networking)*. https://www.ramonmillan.com/tutoriales/softwaredefinednetworking.php

Universidad Publica de Navarra. (2015). *Arquitectura tradicional en el data center: limitaciones*. https://www.tlm.unavarra.es/~daniel/docencia/rng/rng15_16/slides/Tema1-19-ArquitecturaEnDCs.pdf

Villarrubia, C. (2018). *El fin del data center tradicional*. Agosto. https://www.datacenterdynamics.com/es/features/el-fin-del-data-center-tradicional/

ANEXOS

Hoja de Datos de los servidores APIC

Cisco APIC Medium and Large configurations	
Physical dimensions (H x W x D)	1 Rack Unit (1RU): 1.7 x 16.9 x 28.5 in. (4.32 x 43 x 72.4 cm)
Temperature: Operating	32 to 104°F (0 to 40°C) (operating, at sea level, with no fan fail and no CPU throttling, and with turbo mode)
Temperature: Nonoperating	–40 to 158°F (–40 to 70°C)
Humidity: Operating	10 to 90% noncondensing
Humidity: Nonoperating	5 to 93% noncondensing
Altitude: Operating	0 to 10,000 ft (0 to 3000m); maximum ambient temperature decreases by 1°C per 300m
Altitude: Nonoperating	0 to 40,000 ft (12,000m)

Table 1. Cisco APIC Sizes

Cisco APIC configuration	Part number	Description
Medium	APIC-M2	APIC with medium-size CPU, hard drive, and memory configurations (up to 1000 edge ports)
Medium	APIC-M3*	APIC with medium-size CPU, hard drive, and memory configurations (up to 1200 edge ports)
Large	APIC-L2	APIC with large CPU, hard drive, and memory configurations (more than 1000 edge ports)
Large	APIC-L3*	APIC with large CPU, hard drive, and memory configurations (more than 1200 edge ports)
Medium cluster	APIC-CLUSTER-M2	Cluster of 3 APIC devices with medium-size CPU, hard drive, and memory configurations (up to 1000 edge ports)
Medium cluster	APIC-CLUSTER-M3*	Cluster of 3 APIC devices with medium-size CPU, hard drive, and memory configurations (up to 1200 edge ports)
Large cluster	APIC-CLUSTER-L2	Cluster of 3 Cisco APIC devices with large CPU, hard drive, and memory configurations (more than 1000 edge ports)
Large cluster	APIC-CLUSTER-L3*	Cluster of 3 Cisco APIC devices with large CPU, hard drive, and memory configurations (more than 1200 edge ports)

Cisco APIC configuration	Part number	Description
XS Cluster	APIC-CLUSTER-XS	1 M3* APIC with medium-size CPU, hard drive, memory and 2 Virtual APICs. XS Cluster is only available as part of mini ACI fabric bundle part number - ACI-C9332-VAPIC-B1
Medium (spare)	APIC-M2=	APIC with medium-size CPU, hard drive, and memory configurations (up to 1000 edge ports)
Medium (spare)	APIC-M3=*	APIC with medium-size CPU, hard drive, and memory configurations (up to 1200 edge ports)
Large (spare)	APIC-L2=	APIC with large CPU, hard drive, and memory configurations (more than 1000 edge ports)
Large (spare)	APIC-L3=**	APIC with large CPU, hard drive, and memory configurations (more than 1200 edge ports)

	Cisco APIC appliance Medium configuration: M2		Cisco APIC appliance Large configuration: L2	
	Description	Default units	Description	Default units
Processor	1.90-GHz Intel® Xeon® processor E5-2609 v3 with 85 watts (W), 6 cores, 15-MB cache, DDR4, and 1600 MHz	2	2.40-GHz Intel Xeon processor E5-2620 v3 with 85W, 6 cores, 15-MB cache, DDR4, and 1866 MHz	2
Memory	16-GB DDR4 2133-MHz RDIMM PC4-17000, dual-rank x4 with 1.2V	4	16-GB DDR4 2133-MHz RDIMM PC4-17000 dual-rank x4 with 1.2V	8
Hard Drive	600GB 12G SAS 10K RPM SFF HDD	2	1.2 TB 12G SAS 10K RPM SFF HDD	2
PCI Express (PCIe) slots	Cisco UCS® Virtual Interface Card (VIC) 1225 dual-port 10-Gbps Enhanced Small Form-Factor Pluggable (SFP+) Converged Network Adapter (CNA) Or Cisco UCS VIC 1225T dual-port 10GBASE-T CNA	1	Cisco UCS VIC 1225 dual-port 10-Gbps SFP+ CNA Or Cisco UCS VIC 1225T dual-port 10GBASE-T CNA	1
Power supply	770W power supply	1	770W power supply	1

Fuente. (Cisco, 2014)

Hoja de datos Conmutador SPINE

Product overview

Based on Cisco® Cloud Scale technology, this platform supports cost-effective, ultra-high-density cloud-scale deployments, an increased number of endpoints, and cloud services with wire-rate security and telemetry. The platform is built on modern system-architecture designed to provide high performance and meet the evolving needs of highly scalable data centers and growing enterprises.

The product is designed to support innovative technologies such as Media Access Control Security (MACsec), Virtual Extensible LAN (VXLAN), tunnel endpoint VTEP-to¬-VTEP overlay encryption, CloudSec and Streaming Statistics Export (SSX)[1]. MACsec is a security technology that allows traffic encryption at the physical layer and provides secure server, border leaf, and leaf-to-spine connectivity. SSX is hardware-based, consisting of a module that reads statistics from the ASIC and sends them to a remote server for analysis. Through this application, users can better understand network performance without any impact on the switch control plane or CPU.

Cisco provides two modes of operation for Cisco Nexus® 9000 Series Switches. Organizations can use Cisco NX-OS Software to deploy the switches in standard Cisco Nexus switch environments (NX-OS mode). Organizations can also deploy the infrastructure that is ready to support the Cisco Application Centric Infrastructure (Cisco ACI™) platform to take full advantage of an automated, policy-based, systems-management approach (Cisco ACI mode).

Switch models

The Cisco Nexus 9364C Spine Switch is a 2-Rack-Unit (2RU) spine switch that supports 12.84 Tbps of bandwidth and 4.3 bpps across 64 fixed 40/100G QSFP28 ports and 2 fixed 1/10G SFP+ ports (Figure 1). Breakout cables are not supported. The last 16 ports marked in green are capable of wire-rate MACsec encryption.[1] The switch can operate in Cisco ACI Spine or NX-OS mode.

Item	Specifications
Maximum number of IPv4 Longest Prefix Match (LPM) routes	• Default: 7000 • LPM heavy *: 262,000
Maximum number of IPv4 host entries	• Default: 96,000 • LPM heavy *: 262,000
Maximum number of MAC address entries	92,000
Number of multicast routes	• Default: 8000 • LPM heavy *: 32,768
Number of Interior Gateway Management Protocol (IGMP) snooping groups	8000
Number of Access Control List (ACL) entries*	• Per slice of the forwarding engine: • 4000 ingress • 2000 egress • Maximum: 16,000 ingress • 8000 egress • Shipping: 14,328 ingress • 7160 egress
Maximum number of VLANs	4096**
Maximum number of Virtual Routing and Forwarding (VRF) instances	1000
Maximum number of links in a port channel	32
Maximum number of Equal-Cost Multipath (ECMP) paths	64
Maximum number of ECMP groups	1024
Maximum number of port channels	64
Number of active SPAN sessions	4
Maximum number of Rapid Per-VLAN Spanning Tree (RPVST) instances	3967
Maximum number of Hot-Standby Router Protocol (HSRP) groups	490
Maximum number of Multiple Spanning Tree (MST) instances	64
Maximum number of VTEPs	256
Maximum number of static Network Address Translation (NAT) entries	1023
Maximum number of dynamic NAT entries	1023
Maximum number of static twice NAT entries	768

Model	Cisco Nexus 9364C
Physical	• 64-port 40/100G QSFP28 ports and 2-port 1/10G SFP+ ports • Buffer: 40MB • System memory: 32GB • SSD: 128GB • USB: 1 port • RS-232 serial console ports: 1 • Management ports: 2 (1 x 10/100/1000BASE-T and 1 x 1-Gbps SFP) • Broadwell-DE CPU: 4 cores
Power and cooling	• Power: 1200W AC, 930W DC [3] or 1200W HVAC/HVDC • Input voltage: 100 to 240V * AC or -40V to -72V DC (min-max), -48V to -60V DC (nominal) *Supports input voltage of 100–120V for a max output of 800W, 200–240V for a max output of 1200W. PSU redundancy is not supported when used in 100–120V • Hot-swappable, dual fan trays with redundant fans • Frequency: 50 to 60 Hz (AC) • Efficiency: 90% or greater (20 to 100% load) • Port-side intake or port-side exhaust options • Typical power: 429W (AC) • Maximum power: 1245W (AC)
Environmental	• Physical (H x W x D): 3.38 x 17.37 x 22.27 in. (8.59 x 44.13 x 56.58 cm) • Weight: 36.9 lb (16.74kg) with power supplies and fans, 27.4 lb (12.43kg) without power supplies and fans • Operating temperature: 32 to 104° F (0 to 40° C) • Nonoperating (storage) temperature: -40 to 158° F (-40 to 70° C) • Humidity: 5 to 90% (noncondensing) • Altitude: 0 to 13,123 ft (0 to 4000m) • RoHS compliance: Yes
Acoustics	• Fan speed at 40%: 76.7 dBA • Fan speed at 70%: 88.7 dBA • Fan speed at 100%: 97.4 dBA

Fuente. (System, 2019)

Hoja de Datos conmutador Leaf

Switch models

Table 1. Cisco Nexus 9300-FX Series Switches

Model	Description
Cisco Nexus 93180YC-FX	48 x 1/10/25-Gbps fiber ports and 6 x 40/100-Gbps QSFP28 ports
Cisco Nexus 93108TC-FX	48 x 100M/1/10GBASE-T ports and 6 x 40/100-Gbps QSFP28 ports
Cisco Nexus 9348GC-FXP	48 x 100M/1G BASE-T ports, 4 x 1/10/25-Gbps SFP28 ports and 2 x 40/100-Gbps QSFP28 ports

The Cisco Nexus 93180YC-FX Switch (Figure 1) is a 1RU switch with latency of less than 1 microsecond that supports 3.6 Tbps of bandwidth and 1.2 bpps. The 48 downlink ports on the 93180YC-FX are capable of supporting 1-, 10-, or 25-Gbps Ethernet or as 16-, 32-Gbps Fibre Channel ports[1], creating a point of convergence for primary storage, compute servers, and back-end storage resources at the top of rack.

The uplink can support up to six 40- and 100-Gbps ports, or a combination of 1-, 10-, 25-, 40, 50-, and 100-Gbps connectivity, offering flexible migration options.

The switch has IEEE compliant, FC-FEC and RS-FEC enabled for 25-Gbps support. All ports support wire-rate MACsec encryption[2]. Please see the Licensing guide section to enable features on the platform.

Figure 1.

Cisco Nexus 93180YC-FX Switch

Product specifications

The Cisco Nexus 9300-FX series offer industry-leading density and performance with flexible port configurations that can support existing copper and fiber cabling (Tables 2).

Table 2. Cisco Nexus 9300-FX Series Switch specifications

Feature	Cisco Nexus 93180YC-FX	Cisco Nexus 93108TC-FX	Cisco Nexus 9348GC-FXP
Ports	48 x 1/10/25-Gbps and 6 x 40/100-Gbps QSFP28 ports	48 x 100M/1/10GBASE-T and 6 x 40/100-Gbps QSFP28 ports	48 x 1-GBASE-T ports, 4 x 1/10/25-Gbps SFP28 ports and 2 x 40/100 QSFP28 ports
Downlink supported speeds	1/10/25-Gbps Ethernet 16/32-Gbps Fibre Channel	100-Mbps and 1/10-Gbps speeds	100-Mbps and 1-Gbps speeds
CPU	6 cores	4 cores	4 cores
System memory	Up to 32 GB	24 GB	24 GB
SSD drive	128 GB	128 GB	128 GB
System buffer	40 MB	40 MB	40 MB
Management ports	1 RJ-45 port L1 and L2 ports are unused	2 ports: 1 RJ-45 and 1 SFP+	2 ports: 1 RJ-45 and 1 SFP+
USB ports	1	1	1
RS-232 serial ports	1	1	1
Power supplies (up to 2)	500W AC, 930W DC, or 1200W HVAC/HVDC	500W AC, 930W DC, or 1200W HVAC/HVDC	350W AC, 440W DC
Typical power (AC/DC)*	260W	276W	178W
Maximum power (AC/DC)*	425W	460W	287W
Input voltage (AC)	100 to 240V	100 to 240V	100 to 240V
Input voltage (High-Voltage AC [HVAC])	200 to 277V	200 to 277V	90 to 305V

Input voltage (High-Voltage DC [HVDC])	-240 to -380V	-240 to -380V	192 to 400V
Frequency (AC)	50 to 60 Hz	50 to 60 Hz	50 to 60 Hz
Fans	4	4	3
Airflow	Port-side intake and exhaust	Port-side intake and exhaust	Port-side intake and exhaust
Physical dimensions (H x W x D)	1.72 x 17.3 x 22.5 in. (4.4 x 43.9 x 57.1 cm)	1.72 x 17.3 x 22.5 in. (4.4 x 43.9 x 57.1 cm)	1.72 x 17.3 x 19.7 in. (4.4 x 43.9 x 49.9 cm)
Acoustics	57 dBA at 40% fan speed, 68.9 dBA at 70% fan speed, and 77.4 dBA at 100% fan speed	64.2 dBA at 40% fan speed, 68.9 dBA at 70% fan speed, and 77.8 dBA at 100% fan speed	67.5 dBA at 50% fan speed, 73.2 dBA at 70% fan speed, and 81.6 dBA at 100% fan speed
RoHS compliance	Yes	Yes	Yes
MTBF	238,470 hours	319,790 hours	257,860 hours

Minimum ACI image	ACI-N9KDK9-12.2A	ACI-N9KDK9-12.2A	ACI-N9KDK9-13.0
Minimum NX-OS image	NXOS-703I7.1	NXOS-703I7.1	NXOS-703I7.1

* **Typical and maximum power values are based on input drawn from the power circuit. The power supply value (for example, 500W AC power supp NXA-PAC-500W-PI) is based on the output rating to the inside of the switch**

Table 3 lists the performance and scalability specifications for the Cisco Nexus 9300-FX series switches. (Check the software releas for feature support information.)

Table 3. Hardware performance and scalability specifications*

Item	Cisco Nexus 9300-FX Series Switches
Maximum number of Longest Prefix Match (LPM) routes**	1,792,000
Maximum number of IP host entries**	1,792,000
Maximum number of MAC address entries**	512,000
Maximum number of multicast routes	128,000
Number of Interior Gateway Management Protocol (IGMP) snooping groups	Shipping: 8,000 Maximum: 32,000

Maximum number of Cisco Nexus 2000 Series Fabric Extenders per switch	16
Maximum number of Access Control List (ACL) entries	Single-slice forwarding engine: 5000 ingress 2000 egress
Maximum number of VLANs	4096**
Number of Virtual Routing and Forwarding (VRF) instances	Shipping: 1,000 Maximum: 16,000
Maximum number of ECMP paths	64
Maximum number of port channels	512
Maximum number of links in a port channel	32
Number of active SPAN sessions	4
Maximum number of VLAN's in Rapid per-VLAN Spanning Tree (RPVST) instances	3,967
Maximum number of Hot-Standby Router Protocol (HSRP) groups	490
Number of Network Address Translation (NAT) entries	1,023
Maximum number of Multiple Spanning Tree (MST) instances	64
Flow-table size used for Cisco Tetration Analytics platform***	64,000
Number of Queues	8

* More templates and greater scalability are on the roadmap. Refer to the Cisco Nexus 9000 Series Verified Scalability Guide documentation for the latest exact scalability values validated for specific software

** 127 VLANs out of 4096 are reserved

*** Raw capacity of flow table

Fuente. (Cisco, 2009)