HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Vũ Thị Thanh Tú

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

HÀ NỘI - 2025

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Vũ Thị Thanh Tú

NGHIÊN CỨU GIẢI PHÁP VXLAN TRONG CÁC TRUNG TÂM DỮ LIỆU

CHUYÊN NGÀNH : KỸ THUẬT VIỄN THÔNG

MÃ SỐ: 8.52.02.08 (Kỹ thuật Viễn thông)

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC

TS. VŨ TUẤN LÂM

HÀ NỘI - 2025

LỜI CẢM ƠN

Trong suốt quá trình học tập và rèn luyện tại Học viện Công nghệ Bưu chính Viễn thông, em đã nhận được rất nhiều sự quan tâm, chỉ dẫn và hỗ trợ tận tình từ quý Thầy Cô, đặc biệt là các Thầy Cô trong Khoa Viễn thông 1 và Khoa Đào tạo Sau Đại học. Em xin được bày tỏ lòng biết ơn chân thành và sâu sắc tới Ban Giám đốc Học viện, Ban Chủ nhiệm khoa cùng toàn thể quý Thầy Cô đã không ngừng truyền đạt cho em những kiến thức chuyên môn quý báu, cũng như phương pháp tư duy và làm việc khoa học, giúp em có nền tảng vững chắc để hoàn thành chương trình học và thực hiện đề án tốt nghiệp.

Em xin đặc biệt gửi lời cảm ơn chân thành tới Thầy TS. Vũ Tuấn Lâm, người đã trực tiếp hướng dẫn và đồng hành cùng em trong suốt quá trình thực hiện đề án tốt nghiệp với đề tài "Nghiên cứu giải pháp VXLAN trong các trung tâm dữ liệu". Với sự nhiệt tình, tận tâm và chuyên môn sâu rộng, Thầy không chỉ giúp em tiếp cận bài toán một cách bài bản mà còn truyền cảm hứng để em có thêm động lực nghiên cứu và phát triển trong lĩnh vực mạng máy tính và công nghệ dữ liệu hiện đại.

Bên cạnh đó, em cũng xin gửi lời cảm ơn đến bạn bè, người thân và các thầy cô trong suốt quá trình học tập đã luôn đồng hành, chia sẻ và động viên em vượt qua những khó khăn trong học tập và nghiên cứu.

Mặc dù em đã cố gắng nỗ lực hết mình trong quá trình thực hiện đề án, song với năng lực và kinh nghiệm còn hạn chế, đề án này khó tránh khỏi những thiếu sót. Em rất mong nhận được sự góp ý chân thành từ quý Thầy Cô để em có thể hoàn thiện hơn trong những công trình nghiên cứu tiếp theo.

Em xin trân trọng cảm ơn!

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi.

Các số liệu, kết quả nêu trong đề án tốt nghiệp là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Hà Nội, tháng 06 năm 2025

Tác giả đề án

Vũ Thị Thanh Tú

DANH MỤC HÌNH ẢNH

Hình 1. 1 Kiến trúc mạng ba lớp truyền thống giới hạn phạm vi migration VM động 2
Hình 1.2 VXLAN ảo hóa toàn bộ mạng thành một "switch ảo Lớp 2" lớn 3
Hình 1. 3 Định dạng gói tin VXLAN 4
Hình 1. 4 VXLAN Tunnel Endpoint (VTEP) 6
Hình 1. 5 VXLAN Multicast Group trong mạng truyền dẫn 8
Hình 1. 6 Mặt phẳng điều khiển VXLAN Multicast 9
Hình 1. 7 VXLAN High Availability 12
Hình 1. 8 Tenant Routed Multicast (TRM) 14
Hình 1. 9 Luồng chuyển tiếp gói tin VXLAN 15
Hình 1. 10 Đường hầm multicast của một phân đoạn VXLAN thông qua mạng IP 16
Hình 1. 11 Khám phá VTEP từ xa và học địa chỉ 17
Hình 2.1 Giao thức Spanning -tree 25
Hình 2.2 Topology minh họa STP 26
Hình 2.3 BPDU được gửi giữa các switch 27
Hình 2.4 Bầu chọn root bridge 28
Hình 2. 5 Đường đi ngắn nhất đến Root port 29
Hình 2.6 Vòng lặp đã được ngăn chặn thông qua giao thức STP 30
Hình 2.7 Cấu trúc liên kết vật lý và logic của vPC 31
Hình 2. 8 Các thành phần trong vPC 33
Hình 2. 9 Mô hình mạng ba lớp 34
Hình 2. 10 Mô hình Clos 3 lớp 36
Hình 2. 11 Mô hình Spine-Leaf thường được sử dụng trong DC hiện đại 36
Hình 2. 12 Mạng underlay điển hình 38
Hình 2. 13 Cấu trúc liên kết mạng Overlay 39
Hình 2. 14 Mạng Overlay trong trung tâm dữ liệu 40
Hình 2. 15 Tổng quan về EVPN 41
Hình 2. 16 Tích hợp EVPN-VXLAN 43
Hình 2. 17 Xử lý lưu lượng BUM trong VXLAN-EVPN 45

Hình 2. 18 Xử lí lưu lượng unicast 50

Hình 3. 1 Một số sản phẩm dòng Cisco Nexus 51

Hình 3. 2 Giao diện Cisco Nexus Dashboard 52

Hình 3. 3 Kiểm tra địa chỉ IP của server 57

Hình 3. 4 Tình trạng mất gói tin 100% do fabric chưa được cấu hình 58

Hình 3. 5 Topology 59

Hình 3. 6 Xác minh kết nối giữa các server và kết nối tới tài nguyên bên ngoài 60

Hình 3. 7 Các server đã có thể kết nối tới tài nguyên bên ngoài 63

Hình 3. 8 Topology trong NDFC sau khi thực hiện cấu hình 64

DANH MỤC BẢNG

Bảng 1. 1 Một số trường quan trọng trong định dạng gói tin VXLAN 5

Bảng 1. 2 Bảng so sánh VLAN và VXLAN 20

Bảng 3. 1 Các thông tin để tạo external fabric 61

Bảng 3. 2 Thông tin cấu hình liên kết 61

Bảng 3. 3 Thông số cấu hình các kết nối của external router 62

Bảng 3. 4 Thông số cấu hình interface loopback của external router 63

Bảng 3. 5 Thông tin cấu hình để thực hiện quảng bá BGP 63

THUẬT NGỮ VIẾT TẮT

Thuật ngữ viết tắt	Tiếng Anh	Tiếng Việt
ARP	Address Resolution Protocol	Giao thức phân giải địa chỉ
AS	Autonomous System	Miền/ hệ thống tự trị
BUM	Broadcast, Unknown unicast, and Multicast	Lưu lượng Broadcast, unicast, multicast
DC	Data Center	Trung tâm dữ liệu
eBGP	External Border Gateway Protocol	Giao thức BGP ngoại miền
ECMP	Equal-Cost Multi Path	Cân bằng tải đa đường
EVPN	Ethernet Virtual Protocol Netwotk	Giao thức mạng Ethernet ảo
F&L	Flood and Learn	Cơ chế làm đầy và học
HA	High Availability	Tính khả dụng cao
iBGP	Internal Border Gateway Protocol	Giao thức BGP nội miền
IGMP	Internet Group Management Protocol	Giao thức quản lý nhóm Internet
IP	Internet Protocol	Giao thức Internet
LACP	Link Aggregation Control Protocol	Giao thức gom link
MAC	Media Access Control	Địa chỉ vật lý module mạng
MAN	Metropolitan Area Network	Mạng đô thị
MP-BGP EVPN	Multipoint - BGP VPN	Đa điểm – BGP VPN
NIC	Network Interface Card	Thẻ giao diện mạng
NLRI	Network Layer Reachability Information	Khả năng tiếp cận thông tin lớp mạng

Thuật ngữ viết tắt	Tiếng Anh	Tiếng Việt
OSPF	Open Shortest Path First	Giao thức OSPF
OTV	Overlay Transport Virtual	Lớp phủ ảo hóa tầng vận chuyển
PBDU	Protocol Bridge Data Unit	Khối đơn vị dữ liệu giao thức cầu nối
PIM	Protocol Independent Multicast	Giao thức dựa trên Multicast
POD	Point of delivery	Điểm nhận
RR	Route Reflector	Ánh xạ tuyến đường
SDN	Software Defined Network	Mạng định nghĩa bằng phần mềm
STP	Spanning Tree Protocol	Giao thức cây bao trùm mở rộng
TCO	The Total Cost of Ownership	Tổng chi phí sở hữu
UDP	User Datagram Protocol	Giao thức dữ liệu người dùng
VLAN	Virtual Local Area Network	Mạng LAN ảo
VLAN-ID	VLAN-Identifier	Định danh mạng LAN
VM	Virtual Machine	Máy ảo
VNI	VXLAN network identifier	Mã định danh mạng VXLAN
VPC	Virtual Port Channel	Kênh cổng ảo
VRF	Virtual Routing Forwarding	Định tuyến chuyển mạch ảo
VSI	Virtual Switch Instance	Chuyển mạch ảo
VTEP	VXLAN Tunnel EndPoint	Đường hầm VXLAN đầu cuối
VXLAN	Virtual Extensible LAN	Mạng LAN ảo mở rộng
WAN	Wide Area Network	Mạng diện rộng

MỤC LỤC

LỜI CẢM ƠN i

LỜI CAM ĐOAN ii

DANH MỤC HÌNH ẢNH iii

DANH MỤC BẢNG v

THUẬT NGỮ VIẾT TẮT vi

MỤC LỤC viii

MỞ ĐẦU x

CHƯƠNG 1: TỔNG QUAN VỀ CÔNG NGHỆ VXLAN 1

1.1 Giới thiệu chung về VXLAN 1

1.1.1 Khái niệm VXLAN 1

1.1.2 Nguyên nhân ra đời của VXLAN 1

1.1.3 Đóng gói VXLAN và định dạng gói tin 3

1.1.4 VXLAN Tunnel Endpoint (VTEP) 5

1.1.5 Virtual Network Identifier (VNI) 6

1.1.6 Mặt phẳng điều khiển VXLAN 7

1.1.7 VXLAN Gateway 10

1.1.8 VXLAN High Availability 11

1.1.9 Multicast định tuyến theo tenant của VXLAN (VXLAN Tenant Routed Multicast - TRM) 13

1.2 Chuyển tiếp gói tin VXLAN 15

1.2.1 Lưu lượng Unicast (Unicast Traffic) 15

1.2.2 Lưu lượng BUM (BUM Traffic) 16

1.2.3 Khám phá VTEP từ xa và Học địa chỉ (Remote VTEP Discovery and Address learning) 17

1.3. Ưu điểm và hạn chế của VXLAN 19

1.3.1 So sánh VLAN và VXLAN 19

1.3.2 Ưu điểm của VXLAN 20

1.3.3 Hạn chế của VXLAN 20

1.4. Kết luận chương 1 ..... **21**

CHƯƠNG 2: ỨNG DỤNG VXLAN TRONG CÁC TRUNG TÂM DỮ LIỆU **22**

2.1 Tổng quan về trung tâm dữ liệu **22**

2.1.1 Khái niệm trung tâm dữ liệu *22*

2.1.2 Sự phát triển của trung tâm dữ liệu *22*

2.1.3 Các hạng mục chính của trung tâm dữ liệu *23*

2.2. Một số giao thức được sử dụng phổ biến trong các trung tâm dữ liệu **24**

2.2.1 Spanning-tree (STP) *24*

2.2.2 Virtual Port Channel (vPC) *30*

2.3. Sự phát triển của thiết kế mạng trong trung tâm dữ liệu **33**

2.3.1 Mô hình ba lớp *33*

2.3.2 Mô hình Clos *35*

2.4. Ứng dụng VXLAN trong các trung tâm dữ liệu **37**

2.4.1 Ảo hoá mạng, overlay và underlay *37*

2.4.2 Công nghệ EVPN-VXLAN *40*

2.5. Kết luận chương 2 ..... **50**

CHƯƠNG 3: XÂY DỰNG MÔ PHỎNG VXLAN TRONG MẠNG TRUNG TÂM DỮ LIỆU VỚI CISCO NEXUS DASHBOARD FABRIC **51**

3.1. Tổng quan về dòng switch Cisco Nexus và Cisco Nexus Dashboard **51**

3.1.1 Switch Cisco Nexus *51*

3.1.2 Cisco Nexus Dashboard *52*

3.2. Triển khai mô phỏng VXLAN trong mạng trung tâm dữ liệu với Cisco Nexus Dashboard **54**

3.2.1 Mô hình mô phỏng *54*

3.2.2 Lựa chọn công nghệ và kịch bản mô phỏng *55*

3.2.3 Các bước thực hiện *57*

3.2.4 Nhận xét *64*

3.3. Kết luận chương 3 ..... **64**

KẾT LUẬN **66**

TÀI LIỆU THAM CHIẾU **68**

MỞ ĐẦU

Ngày nay, sự phát triển nhanh chóng và cấp bách của các công nghệ mới như điện toán đám mây, dữ liệu lớn và AI đang thúc đẩy các trung tâm dữ liệu áp dụng công nghệ ảo hóa và tăng cường ảo hóa máy chủ để cung cấp dịch vụ đám mây. Thị trường ảo hóa trung tâm dữ liệu toàn cầu được định giá ở mức 8,0 tỷ đô la Mỹ vào năm 2023 và dự kiến sẽ tăng lên 28,9 tỷ đô la Mỹ vào năm 2032, với tốc độ tăng trưởng kép hàng năm (CAGR) là 15,42% trong giai đoạn dự báo 2024–2032 [1]. Do vậy, nhu cầu chuyển đổi dữ liệu không giới hạn qua các máy chủ ảo hóa trên các mạng lớp 2 ngày càng trở nên cấp thiết.

Các liên kết VLAN truyền thống đã được chứng minh là không đủ để đáp ứng các yêu cầu của trung tâm dữ liệu đám mây về quy mô lớn và tính linh hoạt. Do đó, VXLAN đã xuất hiện và trở thành một phần quan trọng của kiến trúc mạng trung tâm dữ liệu hiện đại.

VXLAN hỗ trợ 16 triệu phân đoạn, nhiều hơn gấp nhiều lần 4.094 VLAN [2], cho phép cô lập liên kết giữa nhiều bên thuê và cho phép các trung tâm dữ liệu xử lý lượng lớn lưu lượng trong môi trường đám mây với đủ liên kết và dung lượng.

Hơn nữa, VXLAN đảm bảo cấu trúc mạng nhất quán trên toàn bộ hệ thống, giúp giảm độ phức tạp của mạng và cải thiện hiệu suất mạng. Sử dụng VXLAN, mạng trung tâm dữ liệu có thể hỗ trợ triển khai dịch vụ đám mây quy mô lớn và đáp ứng nhu cầu của trung tâm dữ liệu đám mây.

VXLAN, kết hợp với EVPN, có thể tăng đáng kể khả năng mở rộng mạng trung tâm dữ liệu [3]. EVPN-VXLAN thậm chí cho phép mở rộng các mạng giống hệt nhau trên nhiều trung tâm dữ liệu thông qua một lớp phủ duy nhất, giúp chúng hoạt động như một thể thống nhất.

Đề án bao gồm 3 chương

Chương 1: Tổng quan về công nghệ VXLAN

Chương này giới thiệu về công nghệ VXLAN, bắt đầu bằng việc phân tích nguyên nhân ra đời của nó, liên quan đến những hạn chế của VLAN truyền thống trong môi trường mạng quy mô lớn và ảo hóa. Tiếp theo, chương trình bày chi tiết về định dạng gói tin VXLAN, giải thích cấu trúc các trường và cách VXLAN đóng gói các frame Ethernet layer 2 vào gói tin UDP/IP layer 3. Cuối cùng, chương này mô tả các thành phần và nguyên lý hoạt động cơ bản của VXLAN, bao gồm các khái niệm như VTEP

(VXLAN Tunnel Endpoint), VNI (VXLAN Network Identifier) và cách các gói tin được định tuyến và chuyển mạch trong mạng VXLAN overlay.

Chương 2: Ứng dụng VXLAN trong các trung tâm dữ liệu

Chương này tập trung vào việc ứng dụng công nghệ VXLAN trong các trung tâm dữ liệu. Mở đầu bằng tổng quan về trung tâm dữ liệu, chương này đề cập đến kiến trúc, các yêu cầu và thách thức trong môi trường này. Sau đó, giới thiệu một số giao thức được sử dụng phổ biến trong các trung tâm dữ liệu. Điểm nhấn của chương là phần trình bày về công nghệ ảo hóa mạng trong các trung tâm dữ liệu, trong đó VXLAN được xem xét như một giải pháp quan trọng để mở rộng khả năng ảo hóa layer 2 qua hạ tầng layer 3, tăng tính linh hoạt và khả năng mở rộng cho mạng.

Chương 3: Xây dựng mô phỏng VXLAN trong mạng trung tâm dữ liệu với Cisco Nexus Dashboard Fabric

Chương này đi vào khía cạnh ứng dụng, tập trung vào việc xây dựng mô phỏng mạng trung tâm dữ liệu sử dụng công nghệ VXLAN với một nền tảng cụ thể là Cisco Nexus Dashboard Fabric. Chương bắt đầu bằng việc giới thiệu về dòng switch Cisco Nexus - một dòng sản phẩm phổ biến trong các trung tâm dữ liệu hiện đại và giải pháp Cisco Nexus Dashboard Fabric. Phần trọng tâm sẽ là mô tả quá trình xây dựng mô hình mạng trung tâm dữ liệu cụ thể và các bước mô phỏng việc triển khai công nghệ VXLAN trên nền tảng Cisco Nexus Dashboard Fabric.

CHƯƠNG 1: TỔNG QUAN VỀ CÔNG NGHỆ VXLAN

1.1 Giới thiệu chung về VXLAN

1.1.1 Khái niệm VXLAN

VXLAN (Virtual Extensible LAN - Mạng LAN ảo mở rộng), là một công nghệ ảo hóa mạng được sử dụng rộng rãi trên các mạng Lớp 2 lớn. VXLAN thiết lập một đường hầm logic giữa các thiết bị mạng nguồn và đích, thông qua đó nó sử dụng cơ chế đóng gói MAC-trong-UDP cho các gói tin [4].

Cụ thể, VXLAN đóng gói các khung Ethernet ban đầu được gửi bởi một máy ảo vào các gói UDP. Sau đó, nó đóng gói các gói UDP này với tiêu đề IP và tiêu đề Ethernet của mạng vật lý làm tiêu đề bên ngoài, cho phép các gói tin này được định tuyến trên mạng như các gói IP thông thường. Điều này giải phóng các VM trên mạng Lớp 2 khỏi những hạn chế về cấu trúc của mạng Lớp 2 và Lớp 3.

1.1.2 Nguyên nhân ra đời của VXLAN

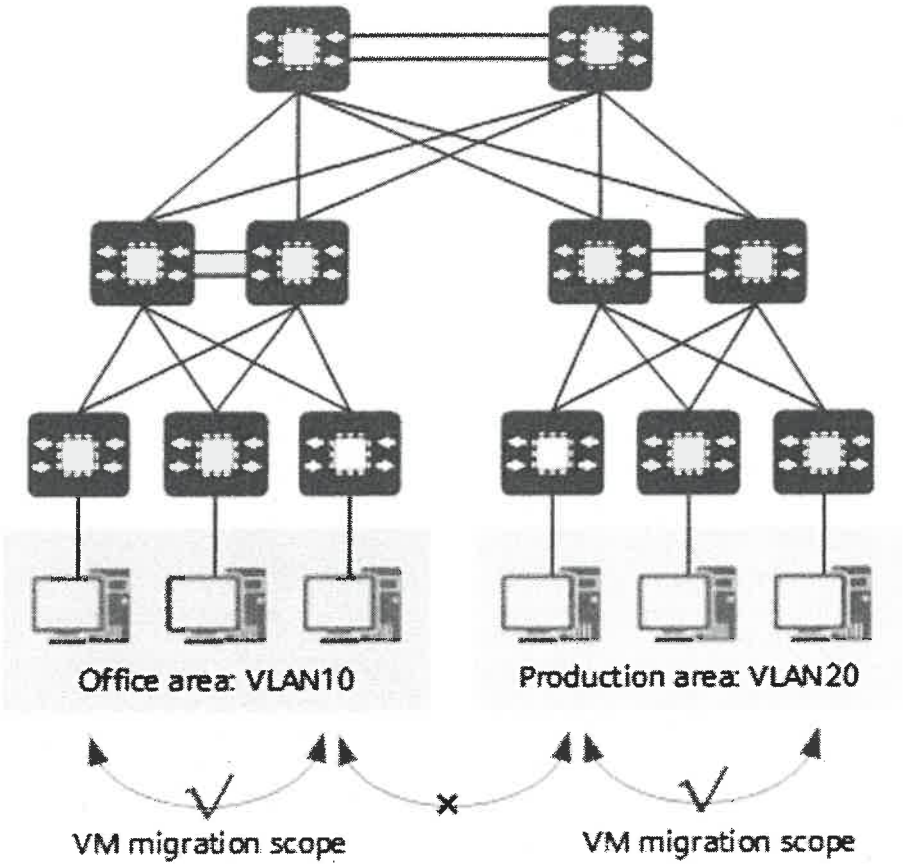
Theo xu hướng ảo hóa máy chủ, việc migration máy ảo động diễn ra, điều này đòi hỏi địa chỉ IP và địa chỉ MAC không được thay đổi trước và sau khi migration. Bên cạnh đó, ảo hóa máy chủ cũng dẫn đến sự gia tăng nhanh chóng về số lượng tenant, điều này giúp mạng cần được cô lập một cách hiệu quả [4].

1.1.2.1 Migration VM động

Ảo hóa máy chủ truyền thống hoạt động bằng cách ảo hóa một máy chủ vật lý thành nhiều máy chủ logic được gọi là máy ảo (VM). Ảo hóa máy chủ là một cách hiệu quả để cung cấp tài nguyên hiệu quả hơn đồng thời cải thiện hiệu suất và tiết kiệm chi phí [5]. Những ưu điểm này giúp nó được sử dụng rộng rãi.

Kể từ khi ảo hóa máy chủ được áp dụng rộng rãi, việc migration VM động ngày càng trở nên phổ biến. Để đảm bảo tính liên tục của dịch vụ trong quá trình migration VM, địa chỉ IP và trạng thái hoạt động của VM (ví dụ: trạng thái phiên TCP) phải không thay đổi. Do đó, VM chỉ có thể được di chuyển động trong cùng một miền Lớp 2.

Như hình bên dưới, kiến trúc mạng ba lớp truyền thống giới hạn phạm vi di chuyển VM động. VM chỉ có thể migration trong một phạm vi giới hạn, hạn chế rất nhiều ứng dụng.



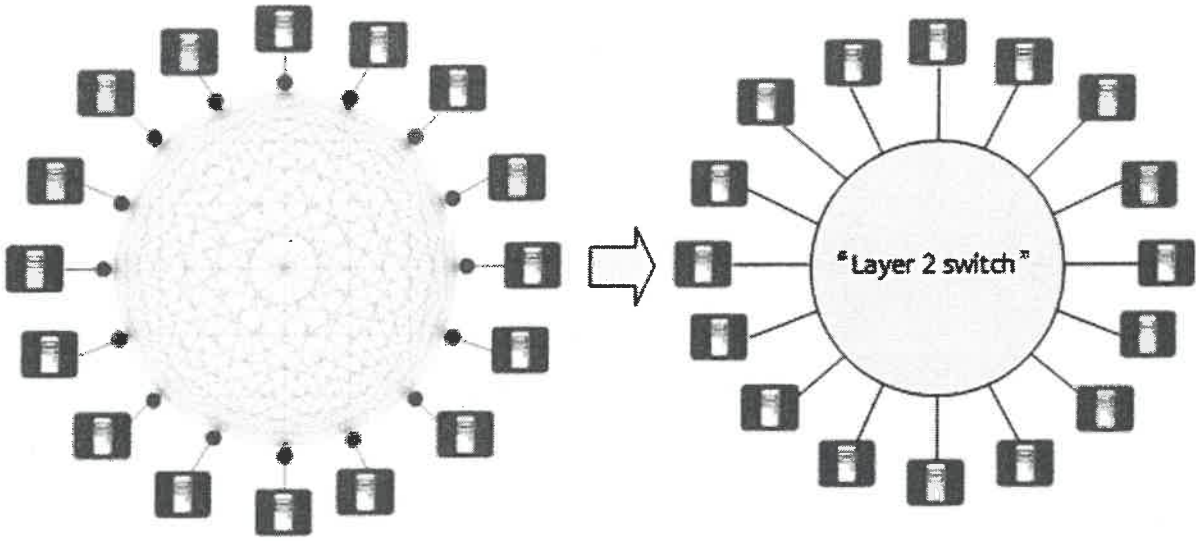
Hình 1. 1 Kiến trúc mạng ba lớp truyền thống giới hạn phạm vi migration VM động

Để cho phép di chuyển VM mượt mà trên phạm vi rộng hoặc thậm chí giữa các khu vực, tất cả các máy chủ liên quan phải được triển khai trong một miền Lớp 2 lớn.

Một switch Lớp 2 có thể hỗ trợ truyền thông Lớp 2 giữa các máy chủ được kết nối với switch. Khi một máy chủ được migration từ một cổng của switch Lớp 2 sang một cổng khác, địa chỉ IP của máy chủ có thể không thay đổi. Điều này đáp ứng các yêu cầu cho việc migration VM động. Chính khái niệm này đã truyền cảm hứng cho việc thiết kế VXLAN.

VXLAN cung cấp một phương pháp để tạo một đường hầm ảo trên mạng IP để chuyển tiếp dữ liệu người dùng một cách trong suốt khi cần giao tiếp giữa một nút nguồn và nút đích trên mạng IP. Bất kỳ hai nút nào cũng có thể giao tiếp thông qua đường hầm VXLAN, bất kể cấu trúc mạng bên dưới và các chi tiết khác. Đối với máy chủ, VXLAN ảo hóa toàn bộ mạng cơ sở hạ tầng thành một "switch ảo Lớp 2" lớn, với tất cả các máy chủ được kết nối với switch này. Các máy chủ không cần biết dữ liệu được chuyển tiếp như thế nào bên trong "switch lớn" này.

Tương tự như cách một máy chủ vật lý hoạt động khi được chuyển từ cổng này sang cổng khác của một switch vật lý, một VM cũng không cần thay đổi địa chỉ IP của nó khi nó được di chuyển từ một cổng của "switch ảo Lớp 2" sang một cổng khác.



Hình 1.2 VXLAN ảo hóa toàn bộ mạng thành một "switch ảo Lớp 2" lớn

1.1.2.2 Sự gia tăng mạnh số lượng tenant làm tăng nhu cầu cô lập mạng

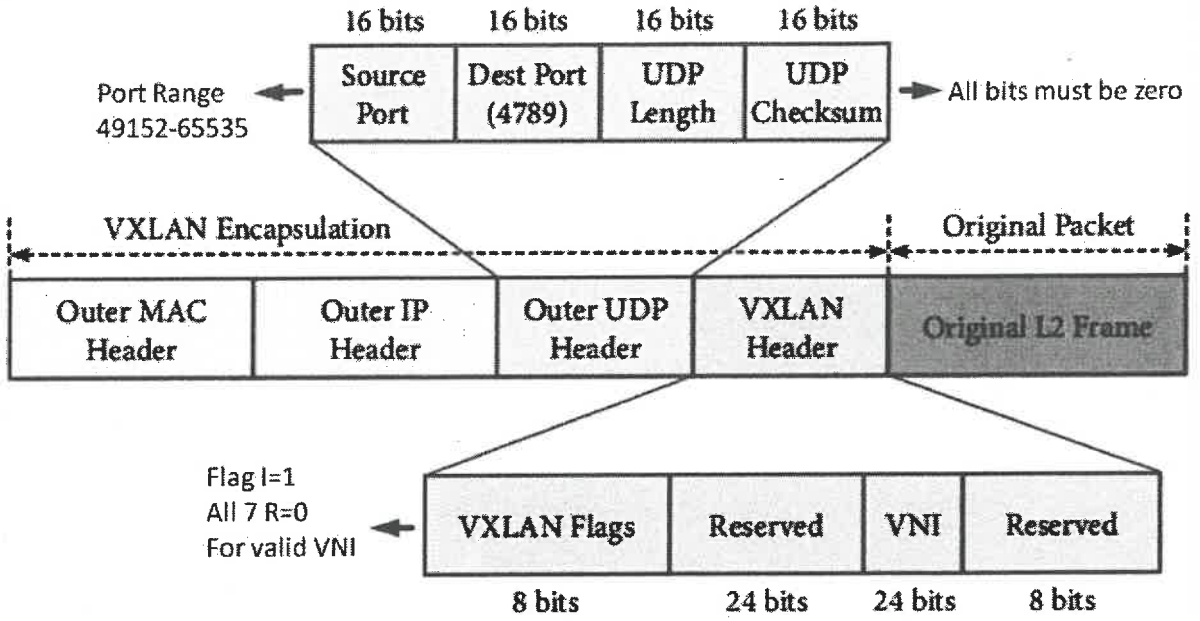
Theo tiêu chuẩn, mạng VLAN truyền thống hỗ trợ tối đa khoảng 4000 VLAN. Sau ảo hóa máy chủ, một máy chủ vật lý chứa nhiều VM (máy ảo), mỗi VM có địa chỉ IP và địa chỉ MAC độc lập. Điều này tương đương với việc số lượng máy chủ tăng lên gấp bội. Ví dụ, các đám mây công cộng hoặc các trung tâm dữ liệu đám mây ảo hóa lớn khác cần chứa hàng vạn tenant hoặc thậm chí nhiều hơn. Trong trường hợp này, VLAN không thể đáp ứng các yêu cầu này.

VXLAN đáp ứng các yêu cầu này như thế nào? VXLAN thêm một định danh mạng VXLAN 24-bit (VNI) tương đương với ID VLAN vào tiêu đề VXLAN. Về lý thuyết, sẽ có tối đa 16 triệu đoạn VXLAN được hỗ trợ, đáp ứng các yêu cầu về nhận dạng và cách ly cho một số lượng lớn tenant.

1.1.3 Đóng gói VXLAN và định dạng gói tin

VXLAN là một giải pháp hỗ trợ môi trường multi tenant linh hoạt, quy mô lớn trên cơ sở hạ tầng vật lý chung được chia sẻ. Giao thức vận chuyển trên mạng trung tâm dữ liệu vật lý là IP cộng với UDP [2].

VXLAN định nghĩa một lược đồ đóng gói MAC-trong-UDP, trong đó khung Lớp 2 gốc được thêm tiêu đề VXLAN và sau đó được đặt trong một gói UDP-IP. Với cách đóng gói MAC-trong-UDP này, VXLAN tạo đường hầm cho mạng Lớp 2 qua mạng Lớp 3. Định dạng gói VXLAN được hiển thị trong Hình 1.3.



Hình 1. 3 Định dạng gói tin VXLAN

Chi tiết mô tả về một số trường quan trọng được thể hiện ở bảng 1.1

Trường	Mô tả
VXLAN header	VXLAN Flags: Đây là trường trong tiêu đề VXLAN chứa các cờ (flags) điều khiển. Trường này bao gồm 8 bit, trong đó bit thứ 5 (I flags) được thiết lập là 1 để chỉ ra rằng đó là một frame có VNI có giá trị. 7 bit còn lại dùng để dự trữ được thiết lập là 0 hết
	VNI: Định danh mạng VXLAN, được sử dụng để xác định một phân đoạn VXLAN. Trường VNI có 24 bit và có thể xác định tối đa 16 triệu tenant. Một tenant có thể có một hoặc nhiều VNI. Các tenant có VNI khác nhau không thể giao tiếp trực tiếp với nhau ở Lớp 2.
	Reserved: Hai trường này được đặt là 0. Tổng 2 trường là 32 bit, được đặt trong trạng thái lưu trữ, có thể dùng cho sau này. Hiện tại chưa có nhu cầu dùng tới
Outer UDP header	DestPort: IANA đã chỉ định cổng 4789 cho VXLAN, cổng này sẽ được sử dụng theo mặc định. Tuy nhiên, cổng này có thể được cấu hình để đảm bảo khả năng tương tác với các triển khai cũ hơn.
	SourcePort: Nên tính toán cổng nguồn bằng cách sử dụng hàm băm của các trường gói tin bên trong, chẳng hạn như tiêu đề của khung Ethernet bên trong. Phải nằm trong phạm vi cổng nguồn động/riêng

Trường	Mô tả
	tư (49152-65535). Phương pháp này giúp cân bằng tải lưu lượng từ điểm cuối đến điểm cuối trên lớp phủ VXLAN.
	UDP Checksum: Theo mặc định, tổng kiểm tra UDP nên được truyền đi là giá trị không (0). Nếu tổng kiểm tra khác không được sử dụng, nó phải được tính toán chính xác trên toàn bộ gói tin. Điểm cuối nhận có thể xác minh tổng kiểm tra này và loại bỏ gói tin nếu quá trình xác minh thất bại.
Outer IP header	Source IP Address: Chỉ ra địa chỉ IP của VTEP (Điểm cuối đường hầm VXLAN) nơi điểm cuối giao tiếp được đặt.
	Destination IP Address: Có thể là địa chỉ unicast hoặc multicast. Địa chỉ unicast đại diện cho địa chỉ IP của VTEP đích, trong khi địa chỉ multicast được sử dụng trong các tình huống liên quan đến nhiều VTEP.
Outer MAC header	Destinantion MAC address: Có thể là địa chỉ của VTEP đích hoặc một router Lớp 3 trung gian.

Bảng 1. 1 Một số trường quan trọng trong định dạng gói tin VXLAN

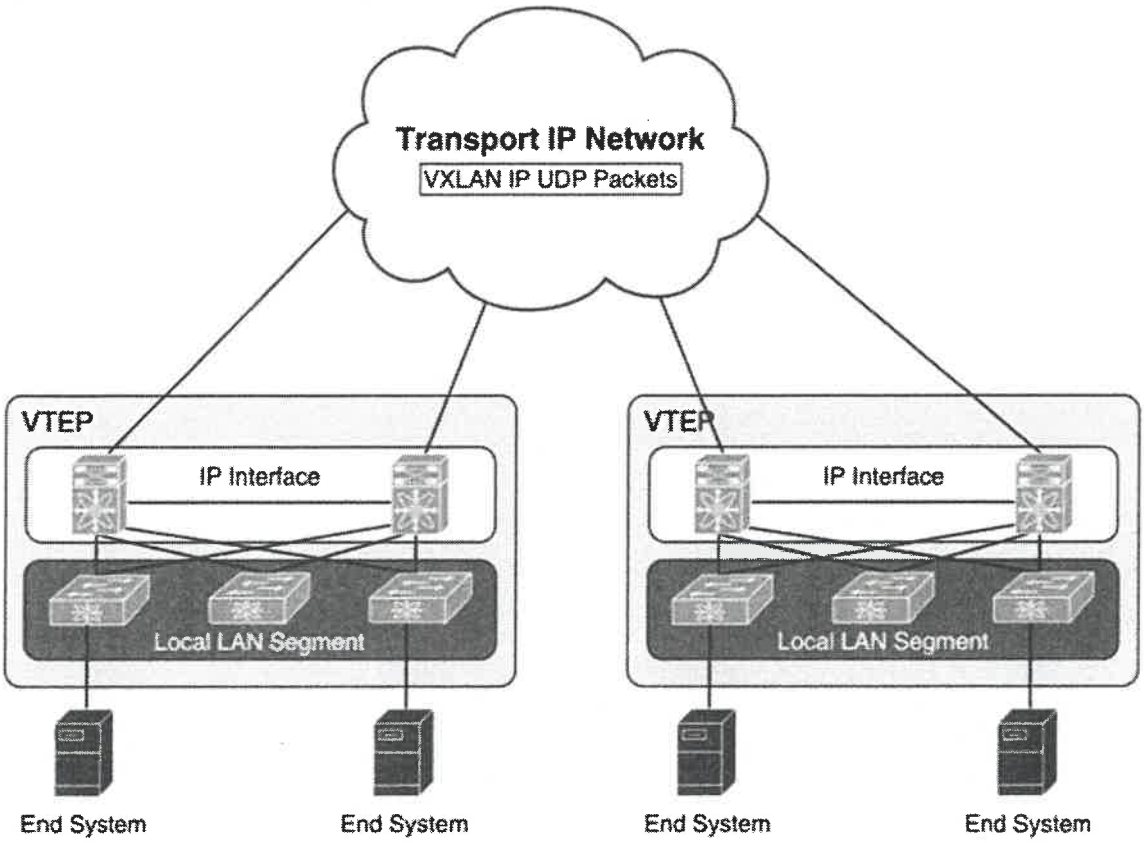
1.1.4 VXLAN Tunnel Endpoint (VTEP)

VTEP là phần tử phần cứng hoặc phần mềm ở biên mạng chịu trách nhiệm khởi tạo đường hầm VXLAN và thực hiện đóng gói và giải đóng gói VXLAN. Mỗi VTEP có hai giao diện: một là giao diện chuyển mạch trên phân đoạn LAN và hai là giao diện IP đến mạng IP vận chuyển.

VLAN cơ sở hạ tầng là một địa chỉ IP duy nhất xác định thiết bị VTEP trên mạng IP vận chuyển. Thiết bị VTEP sử dụng địa chỉ IP này để đóng gói các khung Ethernet và truyền các gói được đóng gói đến mạng vận chuyển thông qua giao diện IP.

Thiết bị VTEP cũng khám phá các VTEP từ xa cho các phân đoạn VXLAN của nó và học các ánh xạ địa chỉ MAC từ xa đến VTEP thông qua giao diện IP của nó. Các thành phần chức năng của VTEP và cấu trúc liên kết logic được tạo ra cho kết nối Lớp 2 trên mạng IP vận chuyển được hiển thị trong Hình 1.4.

Các phân đoạn VXLAN độc lập với cấu trúc liên kết mạng bên dưới; ngược lại, mạng IP bên dưới giữa các VTEP độc lập với lớp phủ VXLAN. Nó định tuyến các gói được đóng gói dựa trên tiêu đề địa chỉ IP bên ngoài, có VTEP khởi tạo làm địa chỉ IP nguồn và VTEP kết thúc làm địa chỉ IP đích.



Hình 1. 4 VXLAN Tunnel Endpoint (VTEP)

1.1.5 Virtual Network Identifier (VNI)

Định danh mạng ảo (VNI) là một giá trị xác định một mạng ảo cụ thể trong mặt phẳng dữ liệu. Thông thường, đó là giá trị 24 bit nằm trong tiêu đề VXLAN, có thể hỗ trợ tối đa 16 triệu phân đoạn mạng riêng lẻ. Các giá trị VNI hợp lệ từ 4096 đến 16.777.215.

Có hai phạm vi VNI chính [2]:

a) *VNI có phạm vi toàn mạng*

Giá trị giống nhau được sử dụng để xác định mạng ảo Lớp 3 cụ thể trên tất cả các thiết bị biên mạng. Phạm vi mạng này hữu ích trong các môi trường như trong trung tâm dữ liệu, nơi các mạng có thể được cung cấp tự động bởi các hệ thống điều phối trung tâm.

Việc có một VNI thống nhất cho mỗi VPN là một cách tiếp cận đơn giản, đồng thời giúp dễ dàng vận hành mạng (chẳng hạn như khắc phục sự cố). Nó cũng có nghĩa là các yêu cầu đơn giản hóa đối với các thiết bị biên mạng, cả thiết bị vật lý và thiết bị ảo. Một yêu cầu quan trọng đối với loại phương pháp này là phải có một số lượng lớn các giá trị định danh mạng, do phạm vi toàn mạng.

b) *VNI được gán cục bộ*

Trong một phương pháp thay thế được hỗ trợ theo RFC 4364, định danh có ý nghĩa cục bộ đối với thiết bị biên mạng quảng bá tuyến đường. Trong trường hợp này, tác động quy mô mạng ảo được xác định trên cơ sở từng nút so với cơ sở mạng.

Khi nó có phạm vi cục bộ và sử dụng ngữ nghĩa hiện có giống như nhãn MPLS VPN, các hành vi chuyển tiếp được chỉ định trong RFC 4364 có thể được sử dụng. Phạm vi này do đó cho phép ghép nối liền mạch một VPN trải dài cả lớp phủ mạng dựa trên IP và MPLS VPN.

Tình huống này có thể xảy ra, chẳng hạn, ở cạnh trung tâm dữ liệu, nơi mạng lớp phủ cấp vào MPLS VPN. Trong trường hợp này, định danh có thể được phân bổ động bởi thiết bị quảng bá.

Điều quan trọng là phải hỗ trợ cả hai trường hợp và, khi làm như vậy, đảm bảo rằng phạm vi của định danh phải rõ ràng và các giá trị không xung đột với nhau.

1.1.6 Mặt phẳng điều khiển VXLAN

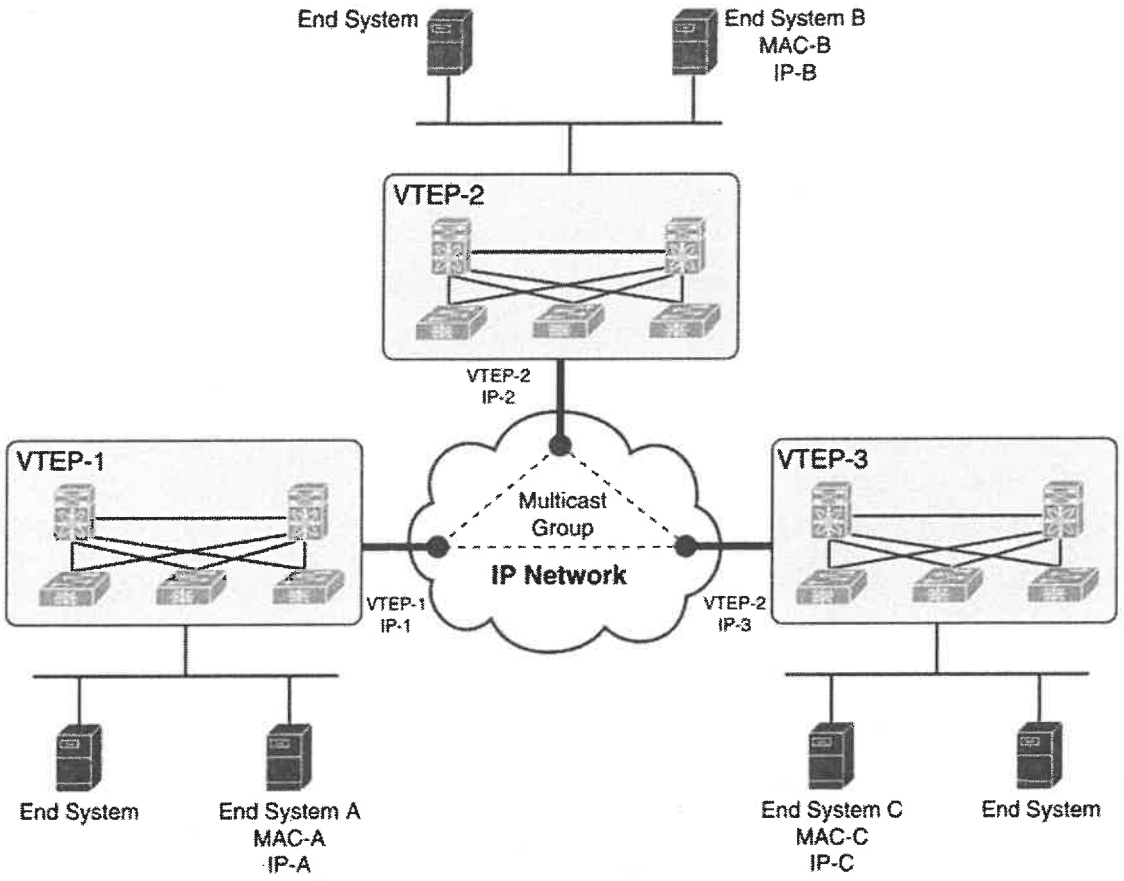
Hai mặt phẳng điều khiển (Control Plane) được áp dụng rộng rãi với VXLAN là *Mặt phẳng điều khiển dựa trên Multicast Flood and Learn của VXLAN* và *Mặt phẳng Điều khiển MPBGP EVPN của VXLAN* [2].

1.1.6.1 Mặt phẳng điều khiển dựa trên Multicast Flood and Learn của VXLAN

Các bộ chuyển mạch sử dụng các cơ chế flooding Lớp 2 hiện có và khả năng tự động học địa chỉ MAC để:

- Vận chuyển lưu lượng broadcast, unknown unicast và multicast (BUM)
- Khám phá các VTEP từ xa
- Học địa chỉ MAC của máy chủ từ xa và ánh xạ MAC-to-VTEP cho từng phân đoạn VXLAN

Multicast IP được sử dụng để giảm phạm vi flooding của tập hợp các máy chủ tham gia vào phân đoạn VXLAN. Mỗi phân đoạn VXLAN, hay VNID, được ánh xạ tới một nhóm multicast IP trong mạng IP truyền tải. Mỗi thiết bị VTEP được cấu hình độc lập và tham gia nhóm multicast này như một máy chủ IP thông qua giao thức IGMP. Các lệnh tham gia IGMP kích hoạt các lệnh tham gia và báo hiệu PIM thông qua mạng truyền tải cho nhóm multicast cụ thể. Cây phân phối multicast cho nhóm này được xây dựng thông qua mạng truyền tải dựa trên vị trí của các VTEP tham gia. Đường hầm multicast của một phân đoạn VXLAN thông qua mạng IP bên dưới được hiển thị trong hình.



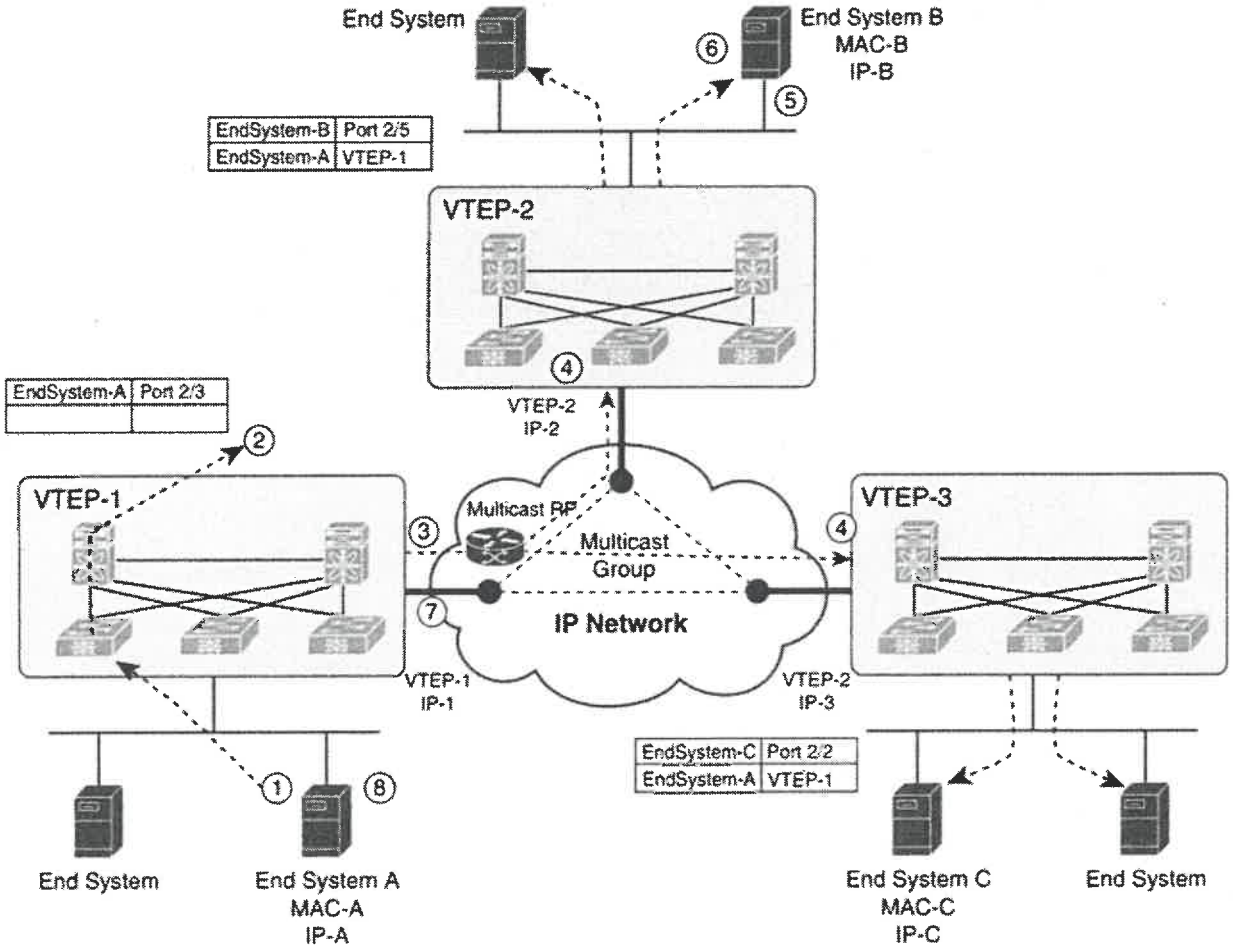
Hình 1. 5 VXLAN Multicast Group trong mạng truyền dẫn

Nhóm multicast được hiển thị trong Hình 1.5 được sử dụng để truyền lưu lượng broadcast, unknown unicast và multicast của VXLAN thông qua mạng IP, giới hạn việc flooding Lớp 2 chỉ đến những thiết bị có các hệ thống cuối tham gia vào cùng một phân đoạn VXLAN. Các VTEP giao tiếp với nhau thông qua lưu lượng được flood hoặc multicast trong nhóm multicast này.

Nếu hệ thống đầu đầu cuối A muốn liên lạc với hệ thống đầu đầu cuối B, nó sẽ thực hiện các bước sau:

1. Hệ thống đầu cuối A tạo một yêu cầu ARP để cố gắng khám phá địa chỉ MAC của Hệ thống đầu cuối B.
2. Khi yêu cầu ARP đến SW1, nó sẽ tra cứu trong bảng cục bộ của mình, và nếu không tìm thấy mục nào, nó sẽ đóng gói yêu cầu ARP qua VXLAN và gửi nó qua nhóm multicast được cấu hình cho VNI cụ thể.
3. RP multicast nhận gói tin và chuyển tiếp một bản sao đến mọi VTEP đã tham gia nhóm multicast.
4. Mỗi VTEP nhận và giải đóng gói gói tin VXLAN và học địa chỉ MAC của hệ thống A trỏ đến địa chỉ VTEP từ xa.

5. Mỗi VTEP chuyển tiếp yêu cầu ARP đến các đích cục bộ của nó.
6. Hệ thống đầu cuối B tạo phản hồi ARP. Khi SW2 VTEP2 nhận được nó, nó tra cứu trong bảng cục bộ của mình và tìm thấy một mục với thông tin rằng lưu lượng đích đến Hệ thống A phải được gửi đến địa chỉ VTEP1. VTEP2 đóng gói phản hồi ARP với tiêu đề VXLAN và unicast nó đến VTEP1.
7. VTEP1 nhận và giải đóng gói gói tin và chuyển nó đến Hệ thống đầu cuối A.
8. Khi thông tin địa chỉ MAC đã được học, các gói tin tiếp theo sẽ được gửi trực tiếp đến địa chỉ VTEP tương ứng.



Hình 1. 6 Mặt phẳng điều khiển VXLAN Multicast

1.1.6.2 Mặt phẳng điều khiển MP-BGP EVPN của VXLAN

Lớp phủ EVPN chỉ định các điều chỉnh cho giải pháp EVPN dựa trên MPLS BGP để nó được áp dụng như một lớp phủ ảo hóa mạng với việc đóng gói VXLAN, trong đó:

- Vai trò nút PE được mô tả trong BGP MPLS EVPN tương đương với thiết bị VTEP/edge ảo hóa mạng (NVE).
- Thông tin VTEP được phân phối thông qua BGP.

- Các VTEP sử dụng cơ chế học/phân phối mặt phẳng điều khiển thông qua BGP cho các địa chỉ MAC từ xa thay vì cơ chế học mặt phẳng dữ liệu.
- Lưu lượng dữ liệu broadcast, unknown unicast và multicast (BUM) được gửi bằng cách sử dụng một cây multicast chung.
- Một bộ phản xạ tuyến đường BGP (RR) được sử dụng để giảm số lượng phiên BGP full mesh giữa các VTEP xuống thành một phiên BGP duy nhất giữa một VTEP và RR.
- Lọc tuyến đường và phân phối tuyến đường có ràng buộc được sử dụng để đảm bảo rằng lưu lượng mặt phẳng điều khiển cho một overlay nhất định chỉ được phân phối đến các VTEP nằm trong instance overlay đó.
- Cơ chế di động máy chủ (MAC) đảm bảo rằng tất cả các VTEP trong instance overlay đều biết VTEP cụ thể được liên kết với địa chỉ MAC.
- Các định danh mạng ảo (VNI) là duy nhất trên toàn cầu trong overlay.

Giải pháp overlay EVPN cho VXLAN cũng có thể được điều chỉnh để cho phép nó được áp dụng như một overlay ảo hóa mạng với VXLAN cho việc phân đoạn lưu lượng Lớp 3. Các điều chỉnh cho VXLAN Lớp 3 tương tự như VXLAN Lớp 2, ngoại trừ những điểm sau:

- Các VTEP sử dụng cơ chế học/phân phối mặt phẳng điều khiển thông qua BGP cho các địa chỉ IP (thay vì địa chỉ MAC).
- Các instance định tuyến và chuyển tiếp ảo được ánh xạ tới VNI.
- Địa chỉ MAC đích bên trong tiêu đề VXLAN không thuộc về máy chủ mà thuộc về VTEP nhận thực hiện định tuyến payload VXLAN. Địa chỉ MAC này được phân phối thông qua thuộc tính BGP cùng với các tuyến EVPN.

1.1.7 VXLAN Gateway

VXLAN gateway được sử dụng để kết nối các phân đoạn VXLAN và VLAN truyền thống nhằm tạo ra một miền chuyển tiếp chung để các thiết bị thuê có thể tồn tại ở cả hai môi trường. Các loại cổng VXLAN bao gồm:

- Cổng Lớp 2: Cổng VXLAN Lớp 2 là một thiết bị đóng gói khung Ethernet cổ điển (CE) thành khung VXLAN và giải đóng gói khung VXLAN thành khung CE. Một thiết bị cổng cung cấp các lợi ích của VXLAN một cách trong suốt cho một thiết bị không hỗ trợ VXLAN; thiết bị đó có thể là một máy chủ vật lý hoặc một máy ảo. Các máy chủ vật lý hoặc máy ảo hoàn toàn không nhận biết được việc đóng gói VXLAN.

- Cổng VXLAN Lớp 3: Tương tự như định tuyến truyền thống giữa các VLAN khác nhau, một bộ định tuyến VXLAN là cần thiết cho việc giao tiếp giữa các thiết bị nằm trong các phân đoạn VXLAN khác nhau. Bộ định tuyến VXLAN dịch các khung từ một VNI sang một VNI khác. Tùy thuộc vào nguồn và đích, quá trình này có thể yêu cầu giải đóng gói và đóng gói lại một khung. Thiết bị Cisco Nexus hỗ trợ tất cả các tổ hợp giải đóng gói, định tuyến và đóng gói. Việc định tuyến cũng có thể được thực hiện trên các giao diện Lớp 3 native và các phân đoạn VXLAN.

Có thể bật định tuyến VXLAN ở lớp tổng hợp hoặc trên các nút tổng hợp của thiết bị switch. Lớp spine chỉ chuyển tiếp lưu lượng dựa trên IP và bỏ qua các gói tin đã được đóng gói. Để giúp mở rộng, một vài nút lá (một cặp lá biên) thực hiện định tuyến giữa các VNI. Một tập hợp các VNI có thể được nhóm thành một instance VRF (Virtual Routing and Forwarding - Định tuyến và Chuyển tiếp Ảo) (VRF thuê) để cho phép định tuyến giữa các VNI đó. Nếu cần bật định tuyến giữa một số lượng lớn các VNI, có thể cần chia các VNI giữa nhiều bộ định tuyến VXLAN. Mỗi bộ định tuyến chịu trách nhiệm cho một tập hợp các VNI và một mạng con tương ứng. Tính dự phòng được đảm bảo bằng FHRP (First Hop Redundancy Protocol - Giao thức Dự phòng Bước Nhảy Đầu Tiên).

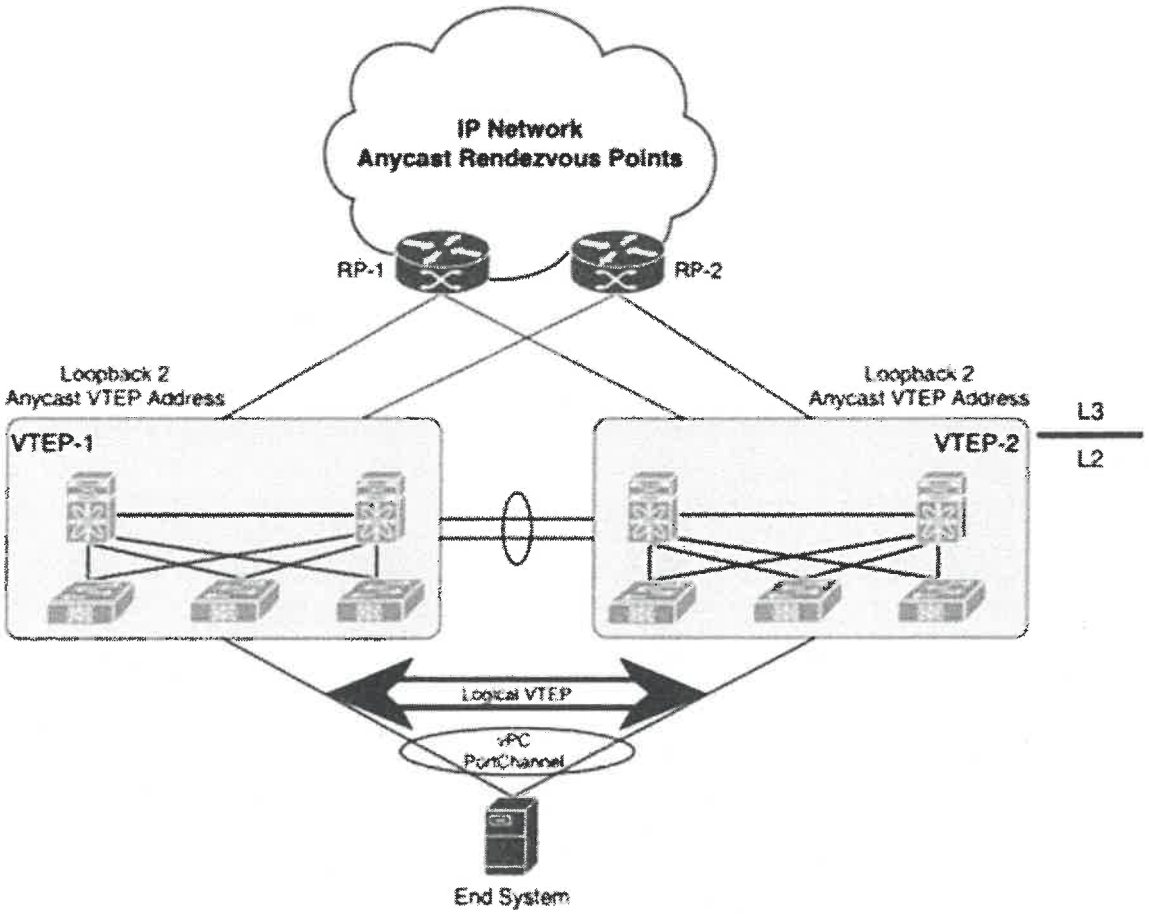
1.1.8 VXLAN High Availability

Để đạt được tính sẵn sàng cao, một cặp switch kênh cổng ảo (vPC) có thể được sử dụng như một thiết bị VTEP logic chia sẻ một địa chỉ VTEP anycast (được hiển thị trong Hình 1.7).

Các switch vPC cung cấp vPC cho kết nối host dự phòng trong khi vẫn chạy độc lập các giao thức Lớp 3 với các thiết bị upstream trong mạng underlay. Cả hai sẽ tham gia cùng một nhóm multicast cho cùng một VXLAN VNI và sử dụng cùng một địa chỉ VTEP anycast làm địa chỉ nguồn để gửi các gói tin được đóng gói VXLAN đến các thiết bị trong mạng underlay, bao gồm điểm tập trung multicast (rendezvous point) và các thiết bị VTEP từ xa. Hai switch VTEP vPC xuất hiện như một thực thể VTEP logic duy nhất.

Các peer vPC phải có các cấu hình giống hệt nhau sau:

- Ánh xạ nhất quán giữa VLAN và phân đoạn mạng ảo (VN-segment)
- Binding NVE nhất quán
- với cùng một địa chỉ IP phụ trên loopback (địa chỉ VTEP anycast)
- Ánh xạ VNI-to-group nhất quán.



Hình 1. 7 VXLAN High Availability

Đối với địa chỉ IP anycast, các switch VTEP vPC phải sử dụng một địa chỉ IP phụ trên giao diện loopback được binding với tunnel NVE VXLAN. Hai switch vPC cần có chính xác cùng một địa chỉ IP loopback phụ.

Cả hai thiết bị sẽ quảng bá địa chỉ VTEP anycast này trên mạng underlay để các thiết bị upstream học được tuyến /32 từ cả hai VTEP vPC và có thể cân bằng tải lưu lượng unicast được đóng gói VXLAN giữa chúng.

Trong trường hợp liên kết peer-link vPC bị lỗi, switch vPC thứ cấp đang hoạt động sẽ tắt giao diện loopback được binding với VXLAN NVE của nó. Việc này sẽ khiến switch vPC thứ cấp rút địa chỉ VTEP anycast khỏi quảng bá IGP của nó, để các thiết bị upstream trong mạng underlay bắt đầu gửi tất cả lưu lượng chỉ đến switch vPC chính. Mục đích của quá trình này là để tránh tình huống vPC active-active khi peer link bị down. Với cơ chế này, các thiết bị orphan (mồ côi) được kết nối với switch vPC thứ cấp sẽ không thể nhận được lưu lượng VXLAN khi liên kết peer-link vPC bị down.

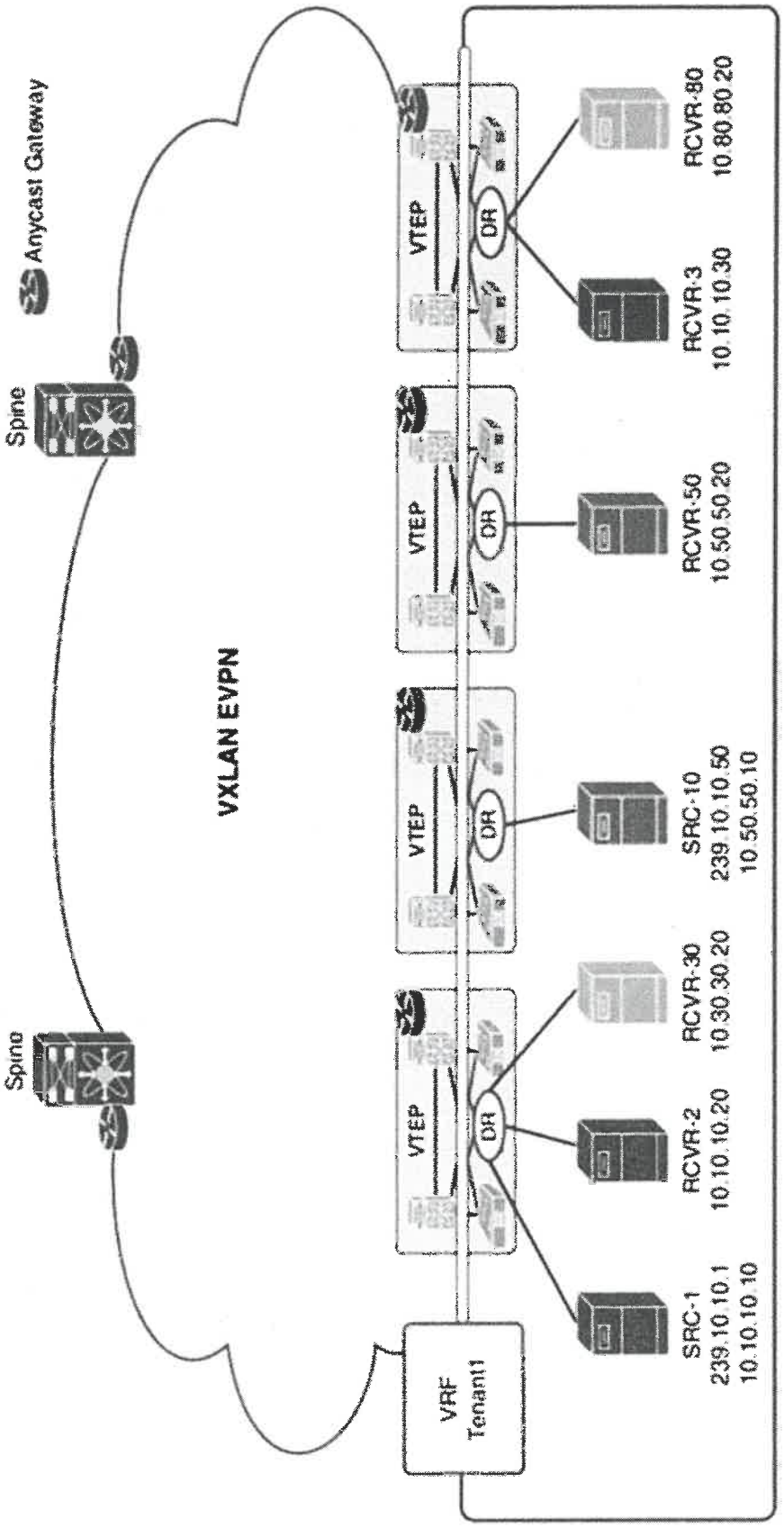
1.1.9 Multicast định tuyến theo tenant của VXLAN (VXLAN Tenant Routed Multicast - TRM)

Tenant Routed Multicast (TRM) mang lại hiệu quả của việc phân phối multicast cho các lớp phủ VXLAN. Nó dựa trên mặt phẳng điều khiển thế hệ tiếp theo (next-gen control plane - ngMVPN) dựa trên các tiêu chuẩn được mô tả trong IETF RFC 6513 và 6514. TRM cho phép phân phối lưu lượng multicast Lớp 3 của khách hàng trong một fabric đa tenant một cách hiệu quả và có khả năng phục hồi cao.

Trong khi BGP EVPN cung cấp một mặt phẳng điều khiển cho định tuyến unicast, như được hiển thị trong hình, ngMVPN cung cấp chức năng định tuyến multicast có khả năng mở rộng. Nó tuân theo cách tiếp cận "luôn định tuyến" (always route) trong đó mọi thiết bị biên (VTEP) có Cổng IP Anycast phân tán cho unicast trở thành bộ định tuyến được chỉ định (Designated Router - DR) cho multicast. Chuyển tiếp multicast bắc cầu (Bridged multicast forwarding) chỉ có trên các thiết bị biên (VTEP) nơi IGMP snooping tối ưu hóa việc chuyển tiếp multicast đến các receiver quan tâm. Tất cả lưu lượng multicast khác ngoài việc phân phối cục bộ đều được định tuyến hiệu quả.

Khi TRM được bật, việc chuyển tiếp multicast trong lớp underlay được tận dụng để nhân bản lưu lượng multicast đã được định tuyến và đóng gói VXLAN. Một cây phân phối Multicast mặc định (Default-MDT) được xây dựng cho mỗi VRF. Đây là sự bổ sung cho các nhóm multicast hiện có cho broadcast, unknown unicast và nhóm nhân bản multicast Lớp 2 của VNI. Các địa chỉ nhóm multicast riêng lẻ trong overlay được ánh xạ tới địa chỉ multicast tương ứng trong underlay để nhân bản và vận chuyển. Ưu điểm của việc sử dụng phương pháp dựa trên BGP là TRM có thể hoạt động như một điểm tập trung (Rendezvous Point - RP) overlay phân tán hoàn toàn, với sự hiện diện của RP trên mọi thiết bị biên (VTEP).

Một fabric trung tâm dữ liệu hỗ trợ multicast thường là một phần của mạng multicast tổng thể. Các nguồn multicast, receiver và thậm chí cả điểm tập trung multicast có thể nằm bên trong trung tâm dữ liệu nhưng cũng có thể nằm bên trong khuôn viên hoặc có thể truy cập từ bên ngoài qua WAN. TRM cho phép tích hợp liền mạch với các mạng multicast hiện có. Nó có thể tận dụng các điểm tập trung multicast bên ngoài fabric. Hơn nữa, TRM cho phép kết nối bên ngoài nhận biết tenant bằng cách sử dụng các giao diện vật lý hoặc subinterface Lớp 3.

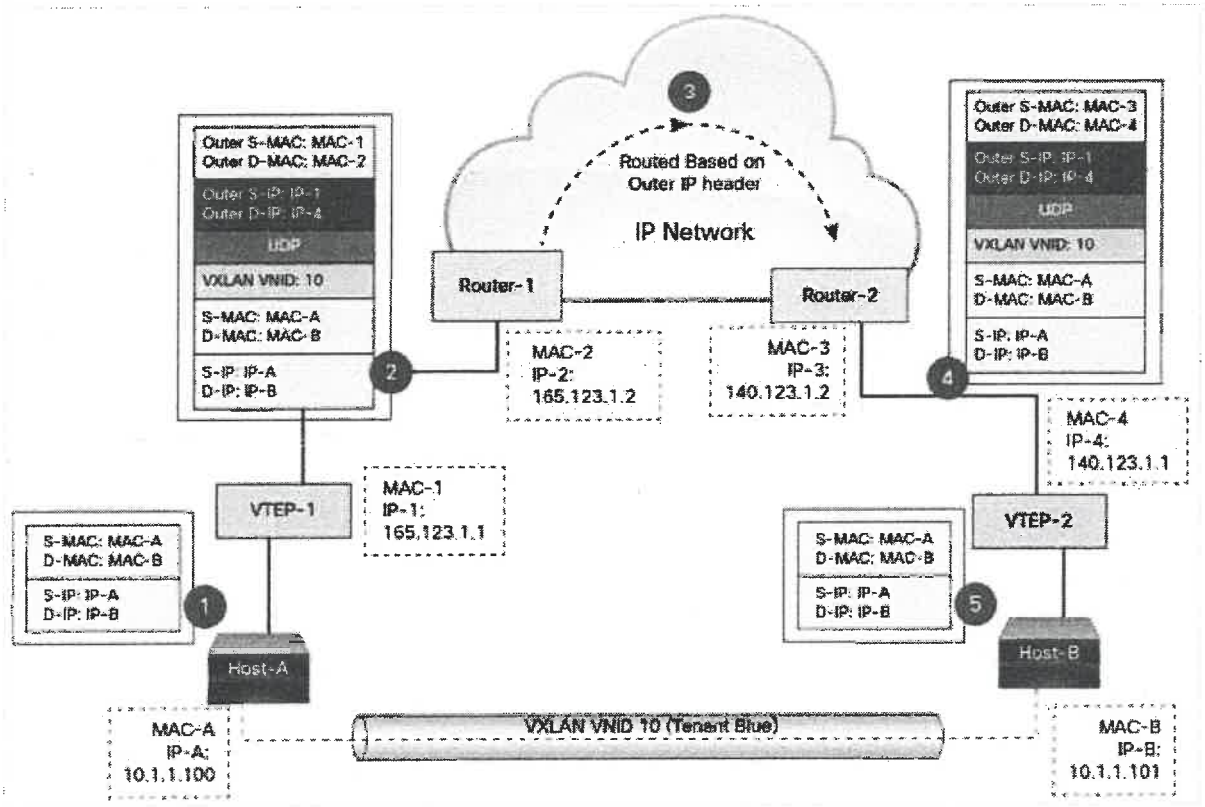


Hình 1. 8 Tenant Routed Multicast (TRM)

1.2 Chuyển tiếp gói tin VXLAN

1.2.1 Lưu lượng Unicast (Unicast Traffic)

VXLAN sử dụng các đường hầm không trạng thái (stateless tunnels) giữa các VTEP để truyền lưu lượng của mạng phủ Lớp 2 (overlay Layer 2) thông qua mạng truyền tải Lớp 3. Để làm rõ và chi tiết hơn, hình dưới đây minh họa về luồng chuyển tiếp gói tin VXLAN [6].



Hình 1. 9 Luồng chuyển tiếp gói tin VXLAN

Host-A và Host-B trong phân đoạn VXLAN 10 giao tiếp với nhau thông qua đường hầm VXLAN giữa VTEP-1 và VTEP-2. Ví dụ này giả định rằng việc học địa chỉ đã được thực hiện ở cả hai phía và các ánh xạ MAC-tới-VTEP tương ứng đã tồn tại trên cả hai VTEP.

Khi Host-A gửi lưu lượng đến Host-B, nó tạo các khung Ethernet với địa chỉ MAC-B của Host-B làm địa chỉ MAC đích và gửi chúng đến VTEP-1. VTEP-1 gán VNID dựa trên ánh xạ của VLAN sang VNID. VTEP-1, với ánh xạ MAC-B tới VTEP-2 trong bảng ánh xạ của nó, thực hiện đóng gói VXLAN trên các gói tin bằng cách thêm vào đó các header VXLAN, UDP và địa chỉ IP bên ngoài. Trong header địa chỉ IP bên ngoài, địa chỉ IP nguồn là địa chỉ IP của VTEP-1 và địa chỉ IP đích là địa chỉ IP của VTEP-2. Sau đó, VTEP-1 thực hiện tra cứu địa chỉ IP cho địa chỉ IP của VTEP-2 để xác định chặng

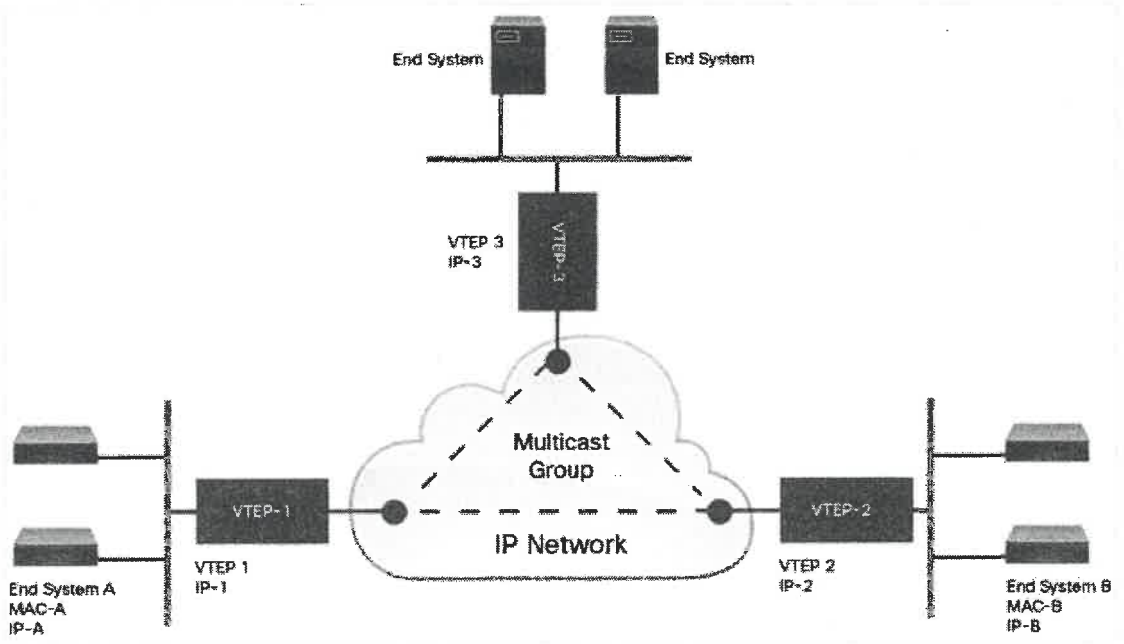
kế tiếp (next hop) trong mạng truyền tải và sau đó sử dụng địa chỉ MAC của thiết bị chặng kế tiếp để đóng gói thêm các gói tin trong một khung Ethernet để gửi đến thiết bị chặng kế tiếp.

Các gói tin được định tuyến về phía VTEP-2 thông qua mạng truyền tải dựa trên header địa chỉ IP bên ngoài của chúng, trong đó có địa chỉ IP của VTEP-2 làm địa chỉ đích. Sau khi VTEP-2 nhận được các gói tin, nó tra cứu VNID để xác định VLAN đích, sau đó loại bỏ các header Ethernet, IP, UDP và VXLAN bên ngoài, và chuyển tiếp các gói tin đến Host-B, dựa trên địa chỉ MAC đích ban đầu trong khung Ethernet.

1.2.2 Lưu lượng BUM (BUM Traffic)

Lưu lượng BUM là cách diễn đạt của lưu lượng Broadcast, Unknown unicast & Multicast. Đối với các loại lưu lượng này, IP multicast được sử dụng để giảm phạm vi flooding của tập hợp các host đang tham gia vào phân đoạn VXLAN.

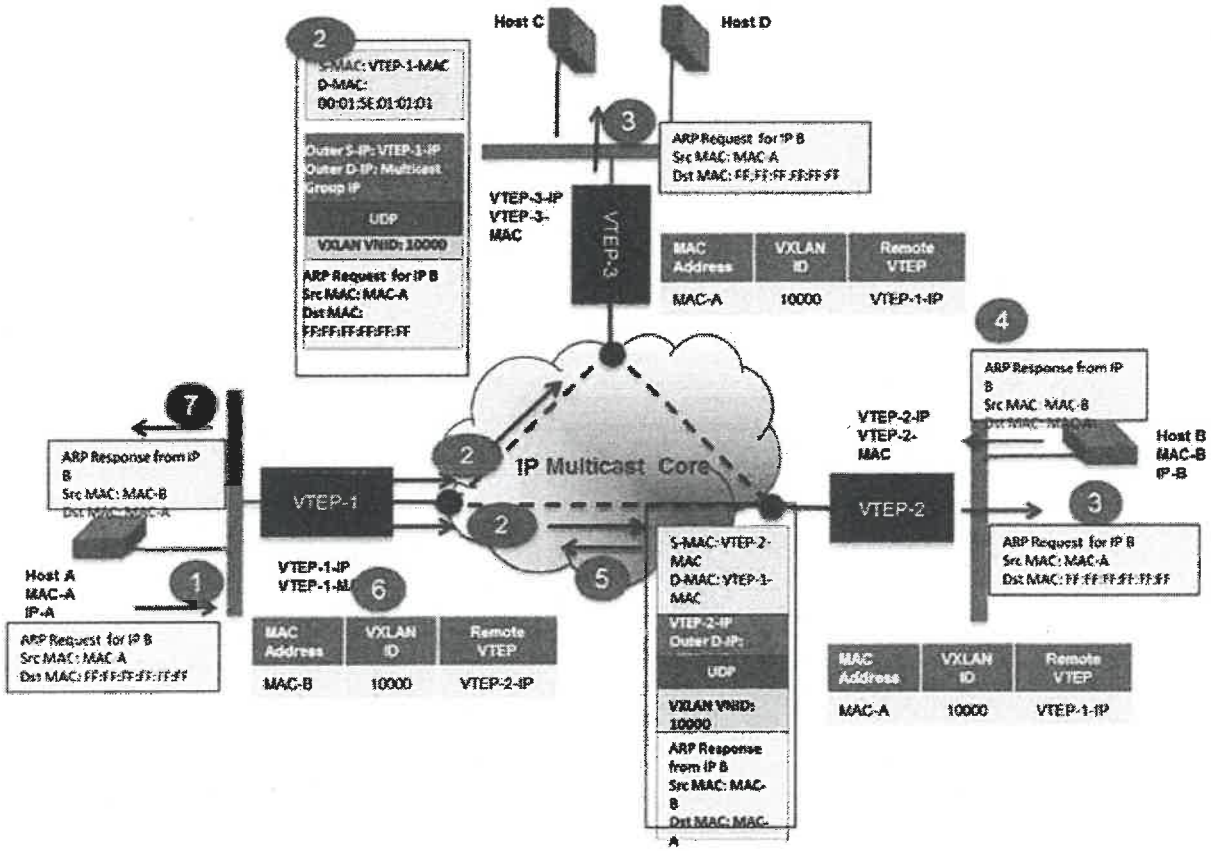
Mỗi phân đoạn VXLAN, hoặc VNID, được ánh xạ tới một nhóm IP multicast trong mạng IP truyền tải. Mỗi thiết bị VTEP được cấu hình độc lập và tham gia vào nhóm multicast này như một IP host thông qua giao thức IGMP. Việc tham gia IGMP kích hoạt việc tham gia và báo hiệu giao thức PIM thông qua mạng truyền tải cho nhóm multicast cụ thể đó. Cây phân phối multicast cho nhóm này được xây dựng thông qua mạng truyền tải dựa trên vị trí của các VTEP tham gia. Đường hầm multicast của một phân đoạn VXLAN thông qua mạng IP bên dưới được hiển thị trong hình.



Hình 1. 10 Đường hầm multicast của một phân đoạn VXLAN thông qua mạng IP

1.2.3 Khám phá VTEP từ xa và Học địa chỉ (Remote VTEP Discovery and Address learning)

Bất kỳ VTEP nào cũng phải biết tất cả các VTEP từ xa và học các ánh xạ MAC-tới-VTEP để có thể thực hiện đóng gói VXLAN. Việc triển khai VXLAN sử dụng các cơ chế flooding and learning mặt phẳng dữ liệu Lớp 2 cổ điển để khám phá VTEP từ xa và học địa chỉ.



Hình 1. 11 Khám phá VTEP từ xa và học địa chỉ

Trong hình này, phân đoạn VXLAN có VNID 10 và sử dụng nhóm multicast 239.1.1.1 qua mạng truyền tải. Nó có ba VTEP tham gia trong trung tâm dữ liệu. Giả sử rằng chưa có việc học địa chỉ nào được thực hiện giữa các vị trí. Hệ thống đầu cuối A (với IP-A, MAC-A) bắt đầu giao tiếp IP với Hệ thống đầu cuối B (với IP-B, MAC-B).

1. Hệ thống đầu cuối A gửi một yêu cầu Giao thức Phân giải Địa chỉ (Address Resolution Protocol - ARP) cho IP-B trên mạng VXLAN Lớp 2 của nó.
2. VTEP-1 nhận được yêu cầu ARP. Nó chưa có ánh xạ cho IP-B. VTEP-1 đóng gói yêu cầu ARP trong một gói tin IP multicast và chuyển tiếp nó đến nhóm multicast VXLAN. Gói tin multicast được đóng gói có địa chỉ IP của VTEP-1 làm địa chỉ IP nguồn và địa chỉ nhóm multicast VXLAN làm địa chỉ IP đích.

3. Gói tin IP multicast được phân phối đến tất cả các thành viên trong cây. VTEP-2 và VTEP-3 nhận được gói tin multicast được đóng gói vì chúng đã tham gia nhóm multicast VXLAN. Chúng giải đóng gói gói tin và kiểm tra VNID của nó trong header VXLAN. Nếu nó khớp với VNID phân đoạn VXLAN đã cấu hình của chúng, chúng sẽ chuyển tiếp yêu cầu ARP đến mạng VXLAN cục bộ của chúng. Chúng cũng học địa chỉ IP của VTEP-1 từ header địa chỉ IP bên ngoài và kiểm tra gói tin để học địa chỉ MAC của Hệ thống đầu cuối A, đặt ánh xạ này vào bảng cục bộ.
4. Hệ thống đầu cuối B nhận được yêu cầu ARP được chuyển tiếp bởi VTEP-2. Nó phản hồi bằng địa chỉ MAC của chính nó (MAC-B) và học ánh xạ IP-A-tới-MAC-A.
5. VTEP-2 nhận được phản hồi ARP của Hệ thống đầu cuối B có MAC-A làm địa chỉ MAC đích. Bây giờ nó đã biết về ánh xạ MAC-A-tới-IP-1. Nó có thể sử dụng đường hầm unicast để chuyển tiếp phản hồi ARP trở lại VTEP-1. Trong gói tin unicast được đóng gói, địa chỉ IP nguồn là IP-2 và địa chỉ IP đích là IP-1. Phản hồi ARP được đóng gói trong payload UDP.
6. VTEP-1 nhận được phản hồi ARP được đóng gói từ VTEP-2. Nó giải đóng gói và chuyển tiếp phản hồi ARP đến Hệ thống đầu cuối A. Nó cũng học địa chỉ IP của VTEP-2 từ header địa chỉ IP bên ngoài và kiểm tra gói tin ban đầu để học ánh xạ MAC-B-tới-IP-2.
7. Các gói tin IP tiếp theo giữa Hệ thống đầu cuối A và B được chuyển tiếp unicast, dựa trên thông tin ánh xạ trên VTEP-1 và VTEP-2, sử dụng đường hầm VXLAN giữa chúng.
8. VTEP-1 có thể tùy chọn thực hiện proxy ARP cho các yêu cầu ARP tiếp theo cho IP-B để giảm thiểu flooding qua mạng truyền tải.

Cuối cùng, mạng VXLAN cho phép mở rộng kết nối Lớp 2 qua một mạng Lớp 3 trung gian, đồng thời cung cấp phân đoạn mạng giống như VLAN, nhưng không có giới hạn về khả năng mở rộng của các VLAN truyền thống.

Điều quan trọng là các tiêu chuẩn VXLAN đã định nghĩa một VXLAN dựa trên multicast flood and learning mà không có mặt phẳng điều khiển (control plane), nói cách khác, VXLAN là một công nghệ đóng gói mặt phẳng dữ liệu (Data plane) cần có sự hỗ trợ từ một giao thức khác để dựa vào đó làm mặt phẳng điều khiển

1.3. Ưu điểm và hạn chế của VXLAN

1.3.1 So sánh VLAN và VXLAN

VLAN giúp tạo các mạng ảo trong một mạng LAN và nhóm các thiết bị thường xuyên giao tiếp với nhau lại với nhau. VXLAN là công nghệ ảo hóa mạng được phát triển để khắc phục những hạn chế của VLAN bằng cách cho phép một mạng duy nhất được sử dụng bởi nhiều tổ chức khác nhau.

VLAN hoạt động ở Layer 2 và phân đoạn mạng vật lý thành nhiều miền quảng bá (broadcast domain), trong khi VXLAN hoạt động ở Layer 2 trên nền Layer 3. Trong Layer 2, nó đóng gói các khung ethernet vào các gói tin UDP.

VLAN sử dụng bộ định danh 12 bit, cho phép tạo 4094 mạng qua Ethernet, trong khi VXLAN sử dụng bộ định danh 24 bit và có thể tạo ra tới 16 triệu mạng. VLAN sử dụng giao thức cây spanning tree (spanning tree protocol), giao thức này chặn một nửa số cổng, trong khi VXLAN cho phép sử dụng tất cả các cổng, giúp tăng hiệu quả. VXLAN được thiết kế để tương thích với cơ sở hạ tầng hiện có và có thể cùng tồn tại với các VLAN truyền thống.

Bảng tóm tắt so sánh về VLAN và VXLAN:

Tiêu chí so sánh	**VLAN (Virtual LAN)**	**VXLAN (Virtual eXtensible LAN)**
Mục đích chính	Tạo mạng ảo trong LAN, nhóm thiết bị liên lạc thường xuyên.	Khắc phục hạn chế của VLAN, cho phép nhiều tổ chức dùng chung mạng.
Lớp hoạt động	Layer 2	Layer 2 trên nền Layer 3
Đóng gói	Sử dụng tag 802.1Q trong khung Ethernet.	Đóng gói khung Ethernet vào gói tin UDP/IP.
Kích thước bộ định danh	12 bit	24 bit
Số lượng mạng tối đa	4094	Lên tới 16 triệu
Spanning Tree / Tận dụng cổng	Sử dụng Spanning Tree, có thể chặn cổng.	Cho phép dùng tất cả các cổng (tăng hiệu quả).

Tiêu chí so sánh	VLAN (Virtual LAN)	VXLAN (Virtual eXtensible LAN)
Tương thích	Công nghệ truyền thống.	Tương thích với hạ tầng hiện có, có thể cùng tồn tại với VLAN truyền thống.

Bảng 1. 2 Bảng so sánh VLAN và VXLAN

1.3.2 Ưu điểm của VXLAN

VXLAN có một loạt lợi ích, bao gồm [7]:

- *Khả năng mở rộng (Scalability):* VXLAN có khả năng mở rộng cao, cho phép tạo ra tới 16 triệu mạng cô lập. Điều này rất hữu ích cho các tổ chức và trung tâm dữ liệu, cho phép họ phục vụ nhiều đối tượng thuê (tenants).
- *Di chuyển máy ảo linh hoạt (Dynamic VM migration):* Việc di chuyển máy ảo từ máy chủ vật lý này sang máy chủ vật lý khác mà không làm gián đoạn dịch vụ hoặc người dùng không hề hay biết có thể thực hiện được thông qua VXLAN. Điều này rất quan trọng để duy trì sự liên tục của dịch vụ và tận dụng hiệu quả các tài nguyên sẵn có.
- *Dễ dàng quản lý và cấu hình:* Vì VXLAN là một mạng phần mềm (software network), nó có thể dễ dàng được quản lý và cấu hình bằng bộ điều khiển tập trung (centralized controller).
- *Quyền riêng tư và bảo mật*: Việc phân đoạn mạng (segmentation) cho phép tăng cường bảo mật và quyền riêng tư, nhờ đó một đối tượng thuê không thể xem lưu lượng truy cập của đối tượng thuê khác.
- *Mã hóa (Encryption):* VXLAN vốn dĩ không cung cấp mã hóa, nhưng các cơ chế mã hóa có thể được áp dụng cùng với VXLAN.

1.3.3 Hạn chế của VXLAN

Bên cạnh những lợi ích, VXLAN cũng có một số hạn chế cần lưu ý [7]

- *Yêu cầu tài nguyên mạng và xử lí lớn hơn:* Việc đóng gói Layer 2 vào bên trong các gói tin UDP có thể dẫn đến tăng chi phí.
- *Phức tạp*: VXLAN có thể trở nên phức tạp, đặc biệt khi xử lý nhiều phân đoạn (segments) và VTEP.

- *Tác động đến hiệu suất khi làm việc với các thiết bị mạng vật lý*: Các địa chỉ MAC mở rộng (Enhanced MAC addresses) và lưu lượng VXLAN có thể ảnh hưởng đến hiệu suất của các bộ chuyển mạch vật lý và các thiết bị mạng.

1.4. Kết luận chương 1

Chương 1 đã cung cấp một cái nhìn tổng quan về công nghệ VXLAN. VXLAN ra đời như một giải pháp đột phá để khắc phục những hạn chế cố hữu của VLAN truyền thống, đặc biệt là trong môi trường trung tâm dữ liệu hiện đại với nhu cầu về khả năng mở rộng mạng Layer 2 vượt trội và sự linh hoạt cho các máy ảo.

Chương này đã làm rõ các khái niệm cốt lõi của VXLAN, bao gồm cơ chế đóng gói khung Ethernet vào gói tin UDP/IP, vai trò quan trọng của VTEP (VXLAN Tunnel Endpoint) trong việc tạo đường hầm và xử lý gói tin, cũng như ý nghĩa của VNI (Virtual Network Identifier) trong việc phân biệt và cô lập lên tới 16 triệu mạng logic. Bên cạnh đó, cũng đã khám phá cách thức hoạt động của mặt phẳng điều khiển VXLAN, vai trò của VXLAN Gateway trong kết nối liên-VXLAN hoặc với mạng truyền thống, và các khía cạnh nâng cao như High Availability và Multicast định tuyến theo tenant (TRM).

Phần hoạt động của VXLAN đã đi sâu vào việc sử dụng IP Multicast để xử lý lưu lượng quảng bá, unicast và multicast, đồng thời xem xét trường hợp nhiều mạng logic cùng ánh xạ tới một nhóm multicast. Đặc biệt, chương đã giải thích chi tiết quá trình VTEP học địa chỉ MAC của máy ảo và cách nó xây dựng, duy trì bảng chuyển tiếp để định tuyến gói tin hiệu quả.

Cuối cùng, thông qua việc so sánh trực tiếp với VLAN đã làm nổi bật những ưu điểm của VXLAN về khả năng mở rộng, hỗ trợ di chuyển máy ảo linh hoạt và quản lý tập trung. Đồng thời, chương cũng đã thẳng thắn nhìn nhận một số hạn chế như chi phí tài nguyên bổ sung và sự phức tạp trong cấu hình, triển khai ở quy mô lớn.

Tóm lại, VXLAN là một công nghệ nền tảng không thể thiếu trong kiến trúc mạng trung tâm dữ liệu ngày nay, mang lại sự linh hoạt và khả năng mở rộng cần thiết cho môi trường ảo hóa và điện toán đám mây. Việc nắm vững các khái niệm và cơ chế hoạt động của VXLAN là bước đầu quan trọng để thiết kế, triển khai và quản lý hiệu quả hạ tầng mạng cho các ứng dụng hiện đại.

CHƯƠNG 2: ỨNG DỤNG VXLAN TRONG CÁC TRUNG TÂM DỮ LIỆU

2.1 Tổng quan về trung tâm dữ liệu

2.1.1 Khái niệm trung tâm dữ liệu

Theo Cisco [8], trung tâm dữ liệu là một cơ sở hạ tầng bao gồm các máy tính kết nối mạng, hệ thống lưu trữ và cơ sở hạ tầng điện toán mà các tổ chức sử dụng để tập hợp, xử lý, lưu trữ và phân phối một lượng lớn dữ liệu. Doanh nghiệp thường phụ thuộc nhiều vào các ứng dụng, dịch vụ và dữ liệu nằm trong trung tâm dữ liệu, khiến nó trở thành một tài sản quan trọng cho các hoạt động hàng ngày. Các thành phần chính của một trung tâm dữ liệu thường bao gồm: router (bộ định tuyến), firewall (tường lửa), switch (bộ chuyển mạch), hệ thống lưu trữ và bộ điều khiển phân phối ứng dụng.

Trước đây, các trung tâm dữ liệu là môi trường vật lý được kiểm soát nghiêm ngặt. Tuy nhiên, cơ sở hạ tầng hiện đại đã chuyển dịch từ các máy chủ vật lý sang môi trường ảo hóa, giúp triển khai ứng dụng và khối lượng công việc trên nhiều môi trường đa đám mây.

Trung tâm dữ liệu hiện đại có khả năng hỗ trợ nhiều loại khối lượng công việc, từ các ứng dụng doanh nghiệp truyền thống cho đến các dịch vụ gốc đám mây hiện đại. Các trung tâm dữ liệu doanh nghiệp ngày càng tích hợp các công cụ để bảo mật và bảo vệ cả tài nguyên đám mây và tài nguyên tại chỗ. Chúng được thiết kế nhằm đáp ứng nhu cầu ngày càng tăng của doanh nghiệp về tài nguyên điện toán, đồng thời tối ưu hiệu quả năng lượng và giảm chi phí vận hành.

Khi các doanh nghiệp chuyển sang điện toán đám mây và môi trường đa đám mây, các trung tâm dữ liệu truyền thống đang dần phát triển, làm mờ ranh giới giữa trung tâm dữ liệu của nhà cung cấp đám mây và trung tâm dữ liệu của doanh nghiệp.

2.1.2 Sự phát triển của trung tâm dữ liệu

Trung tâm dữ liệu lần đầu tiên xuất hiện vào đầu những năm 1940, khi phần cứng máy tính quá phức tạp để vận hành và bảo trì. Hệ thống máy tính ban đầu đòi hỏi nhiều thành phần lớn mà người vận hành phải kết nối với nhiều cáp. Chúng cũng tiêu thụ một lượng lớn điện và cần làm mát để tránh quá nóng. Để quản lý các máy tính này, được gọi là máy tính lớn, các công ty thường đặt tất cả phần cứng trong một phòng duy nhất, được gọi là trung tâm dữ liệu. Mỗi công ty đều đầu tư và bảo trì cơ sở trung tâm dữ liệu của riêng mình.

Theo thời gian, những đổi mới trong công nghệ phần cứng đã giảm yêu cầu về kích cỡ và điện năng của máy tính. Tuy nhiên, cùng với đó, các hệ thống CNTT trở nên phức tạp hơn [9], chẳng hạn như:

- Lượng dữ liệu do các công ty tạo ra và lưu trữ tăng lên theo cấp số nhân.
- Công nghệ ảo hóa tách phần mềm khỏi phần cứng cơ sở.
- Những đổi mới về kết nối mạng giúp chạy các ứng dụng trên phần cứng từ xa.

Thiết kế trung tâm dữ liệu hiện đại đã phát triển để quản lý tốt hơn sự phức tạp của CNTT. Các công ty đã sử dụng trung tâm dữ liệu để lưu trữ cơ sở hạ tầng vật lý ở một vị trí trung tâm mà họ có thể truy cập từ mọi nơi. Với sự xuất hiện của điện toán đám mây, các công ty bên thứ ba quản lý và duy trì trung tâm dữ liệu đồng thời cung cấp cơ sở hạ tầng như một dịch vụ cho các tổ chức khác.

2.1.3 Các hạng mục chính của trung tâm dữ liệu

Hầu hết cơ sở hạ tầng trung tâm dữ liệu doanh nghiệp chia thành ba hạng mục chính:

- Điện toán
- Lưu trữ
- Mạng

Ngoài ra, thiết bị trung tâm dữ liệu bao gồm cơ sở hạ tầng hỗ trợ như hệ thống điện, giúp thiết bị chính hoạt động hiệu quả.

2.1.3.1 Cơ sở hạ tầng điện toán

Tài nguyên điện toán bao gồm một số loại máy chủ có bộ nhớ trong, năng lực xử lý và các thông số kỹ thuật khác đa dạng

a) Máy chủ tủ mạng (Rack Server)

Máy chủ tủ mạng có thiết kế phẳng, hình chữ nhật và có thể xếp máy chủ vào tủ mạng hoặc kệ trong tủ máy chủ. Tủ có các tính năng đặc biệt như cửa lưới, kệ trượt và không gian cho các tài nguyên khác của trung tâm dữ liệu như cáp và quạt.

b) Máy chủ phiến (Blade Server)

Máy chủ phiến là một thiết bị mô-đun và có thể xếp nhiều máy chủ trong một khu vực nhỏ hơn. Bản thân máy chủ khá mỏng và thường chỉ có bộ nhớ, CPU, bộ điều khiển mạng tích hợp và một số ổ lưu trữ tích hợp, có thể trượt nhiều máy chủ vào một đơn vị lưu trữ được gọi là khung. Khung tạo điều kiện thuận lợi cho bất kỳ thành phần bổ sung nào mà các máy chủ bên trong nó yêu cầu. Máy chủ phiến chiếm ít không gian hơn máy

chủ tủ mạng và cung cấp tốc độ xử lý cao hơn, hệ thống dây ít và tiêu thụ điện năng thấp hơn.

2.1.3.2 Cơ sở hạ tầng lưu trữ

Sau đây là hai loại hệ thống lưu trữ trung tâm dữ liệu.

a) Thiết bị lưu trữ khối dữ liệu

Các thiết bị lưu trữ khối dữ liệu như ổ cứng và ổ đĩa thể rắn lưu trữ dữ liệu theo khối và cung cấp nhiều terabyte dung lượng dữ liệu. Mạng khu vực lưu trữ (SAN) là các đơn vị lưu trữ chứa một số ổ đĩa bên trong và hoạt động như các hệ thống lưu trữ khối dữ liệu lớn.

b) Thiết bị lưu trữ tệp

Các thiết bị lưu trữ tệp, chẳng hạn như thiết bị lưu trữ gắn vào mạng (NAS) có thể lưu trữ một lượng lớn tệp, có thể sử dụng chúng để tạo kho lưu trữ hình ảnh và video.

2.1.3.3 Cơ sở hạ tầng mạng

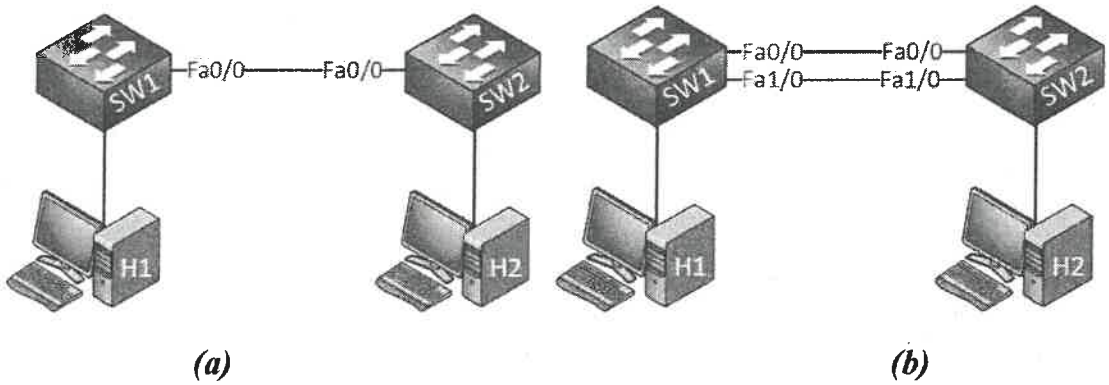
Một số lượng lớn các thiết bị kết nối mạng, như cáp, bộ chuyển mạch, bộ định tuyến và tường lửa kết nối các thành phần khác của trung tâm dữ liệu với nhau và với các vị trí của người dùng cuối. Chúng cung cấp khả năng di chuyển và kết nối dữ liệu trơn tru trên toàn hệ thống.

2.2. Một số giao thức được sử dụng phổ biến trong các trung tâm dữ liệu

2.2.1 Spanning-tree (STP)

STP, giao thức bridge IEEE 802.1D, là một giao thức quản lý liên kết Lớp 2 cung cấp tính năng dự phòng đường truyền đồng thời ngăn chặn các vòng lặp không mong muốn trong mạng. Để một mạng Ethernet Lớp 2 hoạt động bình thường, chỉ có thể tồn tại một đường dẫn hoạt động duy nhất giữa hai trạm bất kỳ. Hoạt động của STP là trong suốt đối với các trạm cuối, chúng không thể phát hiện liệu mình đang kết nối với một đoạn LAN đơn lẻ hay một mạng LAN chuyển mạch gồm nhiều đoạn.

2.2.1.1 Nguyên nhân cần sử dụng STP



(a) *(b)*

Hình 2.1 Giao thức Spanning-tree

Trong hình (a) trên, có hai thiết bị chuyển mạch (switch). Các thiết bị chuyển mạch này được kết nối bằng một sợi cáp duy nhất, do đó có một điểm lỗi đơn lẻ (single point of failure). Để loại bỏ điểm lỗi đơn lẻ này, thêm một sợi cáp khác

Với sợi cáp bổ sung, giờ đây kết nối đã có tính dự phòng như hình (b). Thật không may, tính dự phòng cũng mang đến vòng lặp (loops). Kịch bản như sau:

1. H1 gửi một yêu cầu ARP vì nó đang tìm địa chỉ MAC của H2. Một yêu cầu ARP là một khung quảng bá (broadcast frame).
2. SW1 sẽ chuyển tiếp khung quảng bá này trên tất cả các giao diện của nó, ngoại trừ giao diện mà nó đã nhận khung đó.
3. SW2 sẽ nhận được cả hai khung quảng bá.

Bây giờ, SW2 làm gì với những khung quảng bá đó?

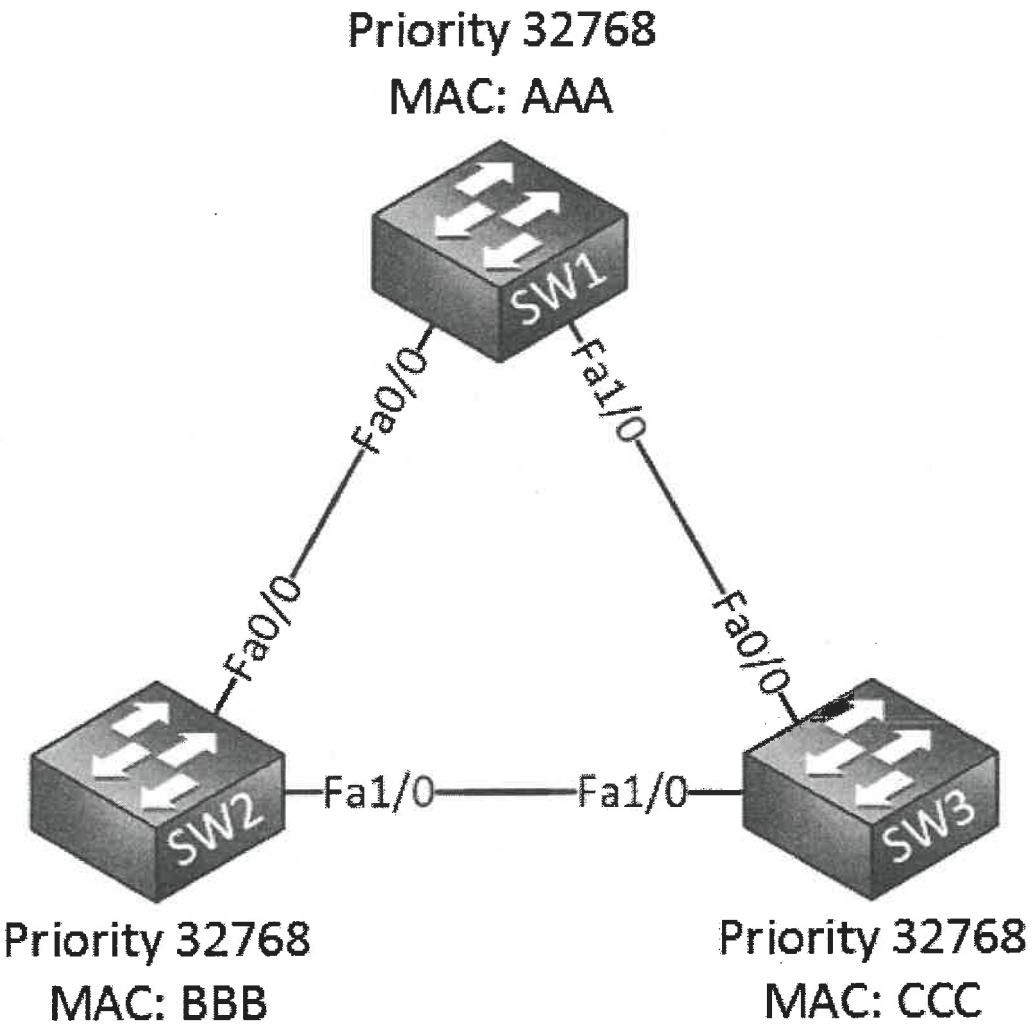
1. Nó sẽ chuyển tiếp khung đó từ mọi giao diện ngoại trừ giao diện mà nó đã nhận khung.
2. Điều này có nghĩa là khung đã được nhận trên giao diện Fa0/0 sẽ được chuyển tiếp trên Giao diện Fa1/0.
3. Khung đã được nhận trên Giao diện Fa1/0 sẽ được chuyển tiếp trên Giao diện Fa0/0.

Ở đây sẽ xuất hiện một vòng lặp. Cả hai switch sẽ tiếp tục chuyển tiếp lặp đi lặp lại cho đến khi một trong những điều sau xảy ra:

- Khắc phục vòng lặp bằng cách ngắt kết nối một trong các sợi cáp.
- Một trong các switch sẽ bị treo vì chúng bị quá tải lưu lượng truy cập.

Các khung Ethernet không có giá trị TTL (Thời gian sống), vì vậy chúng sẽ lặp lại mãi mãi. Bên cạnh các yêu cầu ARP, nhiều khung cũng được quảng bá. Ví dụ, bất cứ khi nào switch không biết về một địa chỉ MAC đích, nó sẽ gửi tràn (flooded).

2.2.1.2 Cách hoạt động của STP



Hình 2.2 Topology minh họa STP

Có ba thiết bị chuyển mạch (switch), và thêm tính dự phòng bằng cách kết nối các switch theo hình tam giác, điều này cũng có nghĩa là có một vòng lặp ở đây. Để đơn giản hóa, các địa chỉ MAC được rút gọn cho từng switch

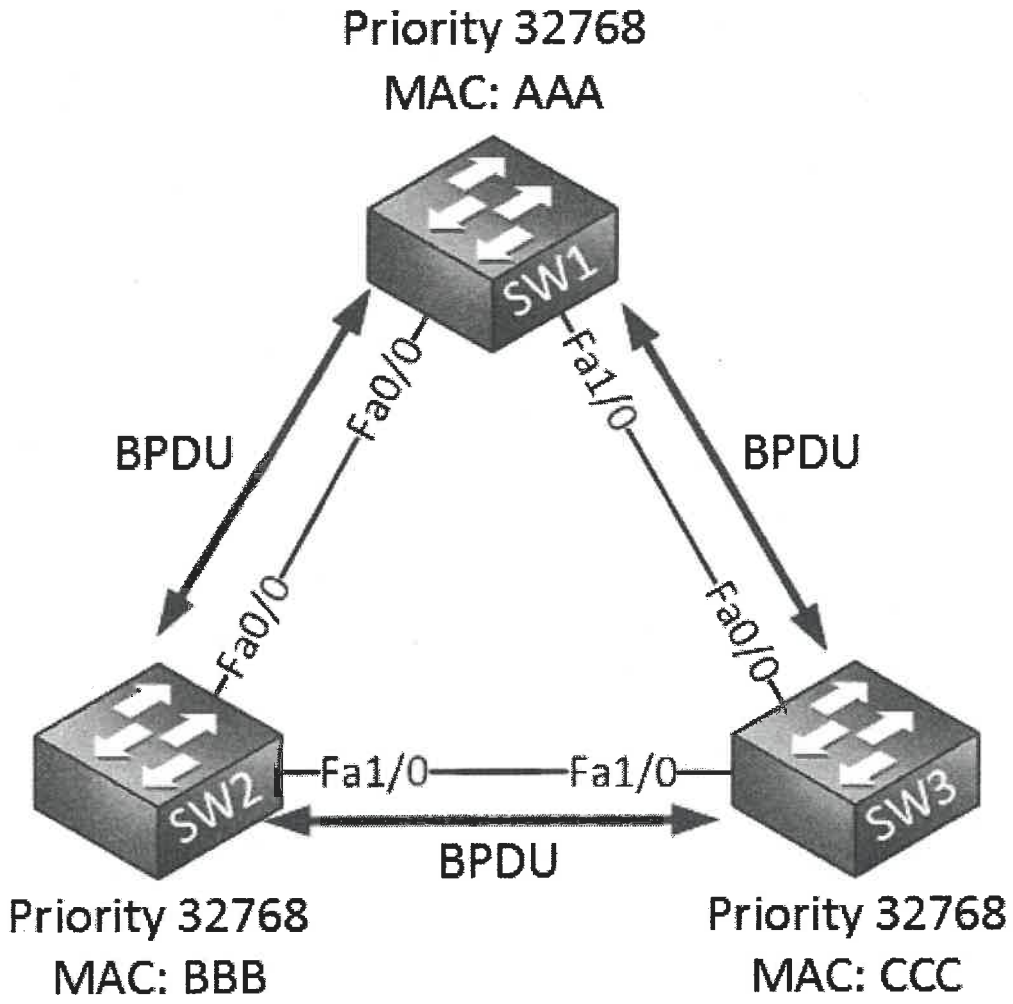
- SW1 có địa chỉ MAC: AAA
- SW2 có địa chỉ MAC: BBB
- SW3 có địa chỉ MAC: CCC

Vì giao thức spanning tree được kích hoạt, tất cả các switch sẽ gửi cho nhau một khung đặc biệt gọi là BPDU (Bridge Protocol Data Unit - Đơn vị Dữ liệu Giao thức Bridge). Trong BPDU này, có hai thông tin mà spanning tree yêu cầu:

- Địa chỉ MAC

- Độ ưu tiên (Priority)

Địa chỉ MAC và độ ưu tiên cùng nhau tạo thành bridge ID (Định danh Bridge). BPDU được gửi giữa các switch như được hiển thị trong hình ảnh sau:



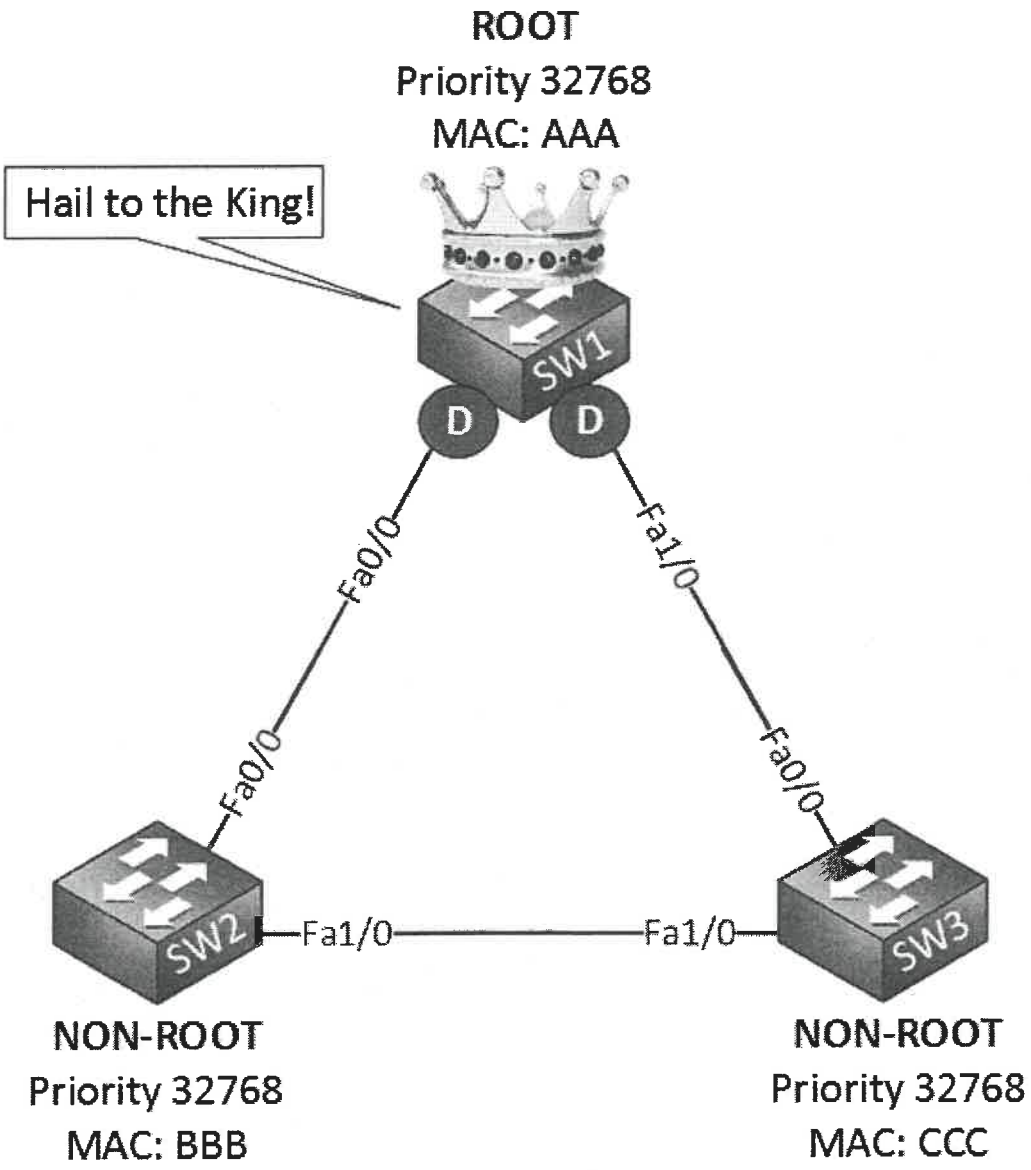
Hình 2.3 BPDU được gửi giữa các switch

Spanning tree yêu cầu bridge ID (Định danh Bridge) cho việc tính toán của nó. Đây là cách hoạt động:

Trước hết, spanning tree sẽ bầu chọn một root bridge (bridge gốc); root bridge này sẽ là bridge có "bridge ID" tốt nhất. Trong đó, switch có bridge ID thấp nhất là tốt nhất. Theo mặc định, độ ưu tiên (priority) là 32768, nhưng có thể thay đổi giá trị này nếu muốn.

Trong Hình 2.3, SW1 sẽ trở thành root bridge. Độ ưu tiên và địa chỉ MAC tạo nên bridge ID. Vì độ ưu tiên trên tất cả các switch là như nhau, nên địa chỉ MAC sẽ là yếu tố quyết định. SW1 có địa chỉ MAC thấp nhất, do đó có bridge ID tốt nhất và sẽ trở

thành root bridge. Các cổng trên root bridge luôn là designated (cổng chỉ định), có nghĩa là chúng ở trạng thái forwarding (chuyển tiếp).



Hình 2.4 Bầu chọn root bridge

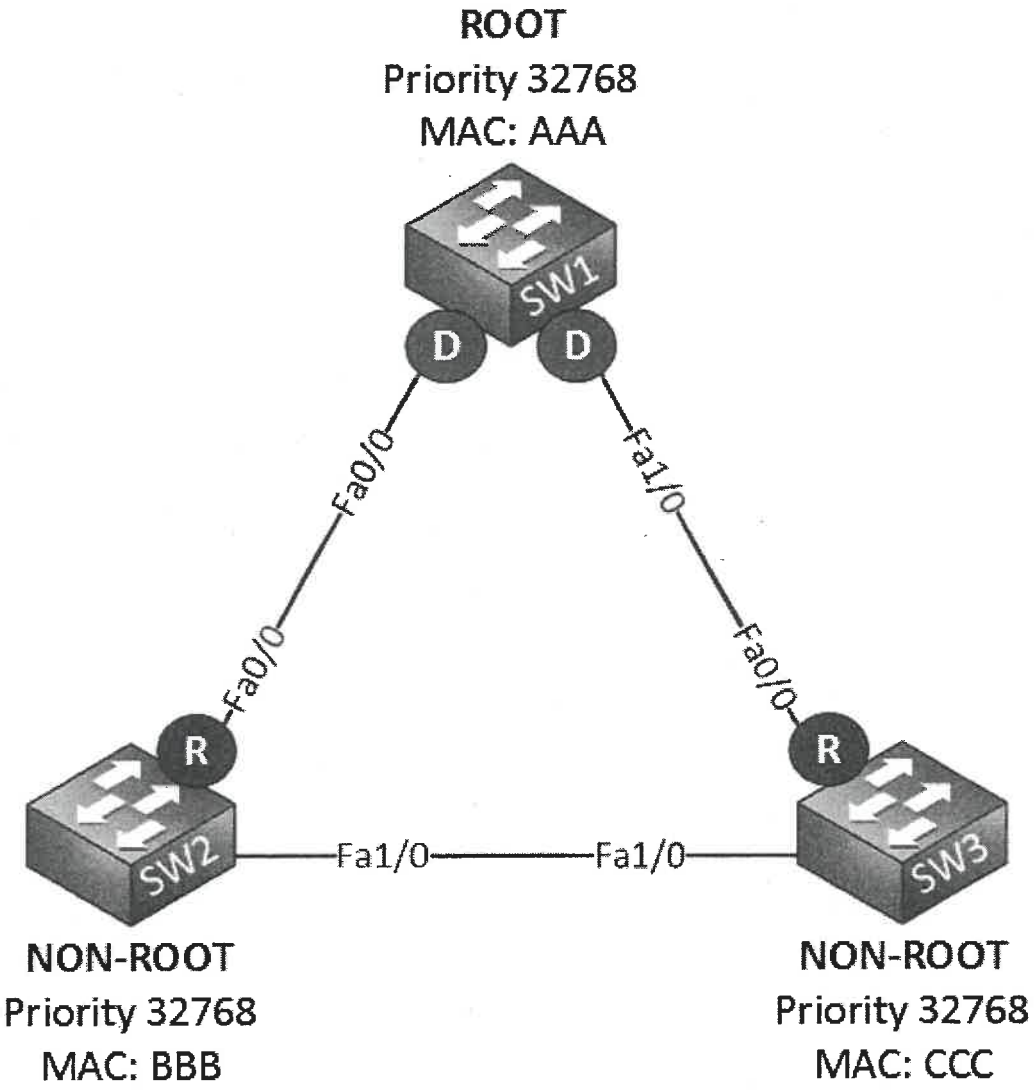
Ở hình trên đã thể hiện SW1 đã được bầu chọn làm root bridge và chữ "D" trên các giao diện là viết tắt của designated (chỉ định).

Bước tiếp theo đối với tất cả các non-root bridge (nghĩa là mọi switch không phải là root) sẽ phải tìm đường đi ngắn nhất đến root bridge. Đường đi ngắn nhất đến root bridge được gọi là root port (cổng gốc).

Hình 2.5 đã thể hiện chữ "R" cho "root port" trên SW2 và SW3. Giao diện Fa0/0 là đường đi ngắn nhất để đến root bridge. Hình minh họa đã giữ mọi thứ đơn giản, nhưng "đường đi ngắn nhất" trong spanning tree có nghĩa là nó sẽ thực sự xem xét tốc độ của

giao diện. Mỗi giao diện có một chi phí (cost) nhất định, và đường đi với chi phí thấp nhất sẽ được sử dụng. Đây là tổng quan về các giao diện và chi phí của chúng:

- Tốc độ 10 Mbit có cost 100
- Tốc độ 100 Mbit có cost 19
- Tốc độ 1000 Mbit có cost 4

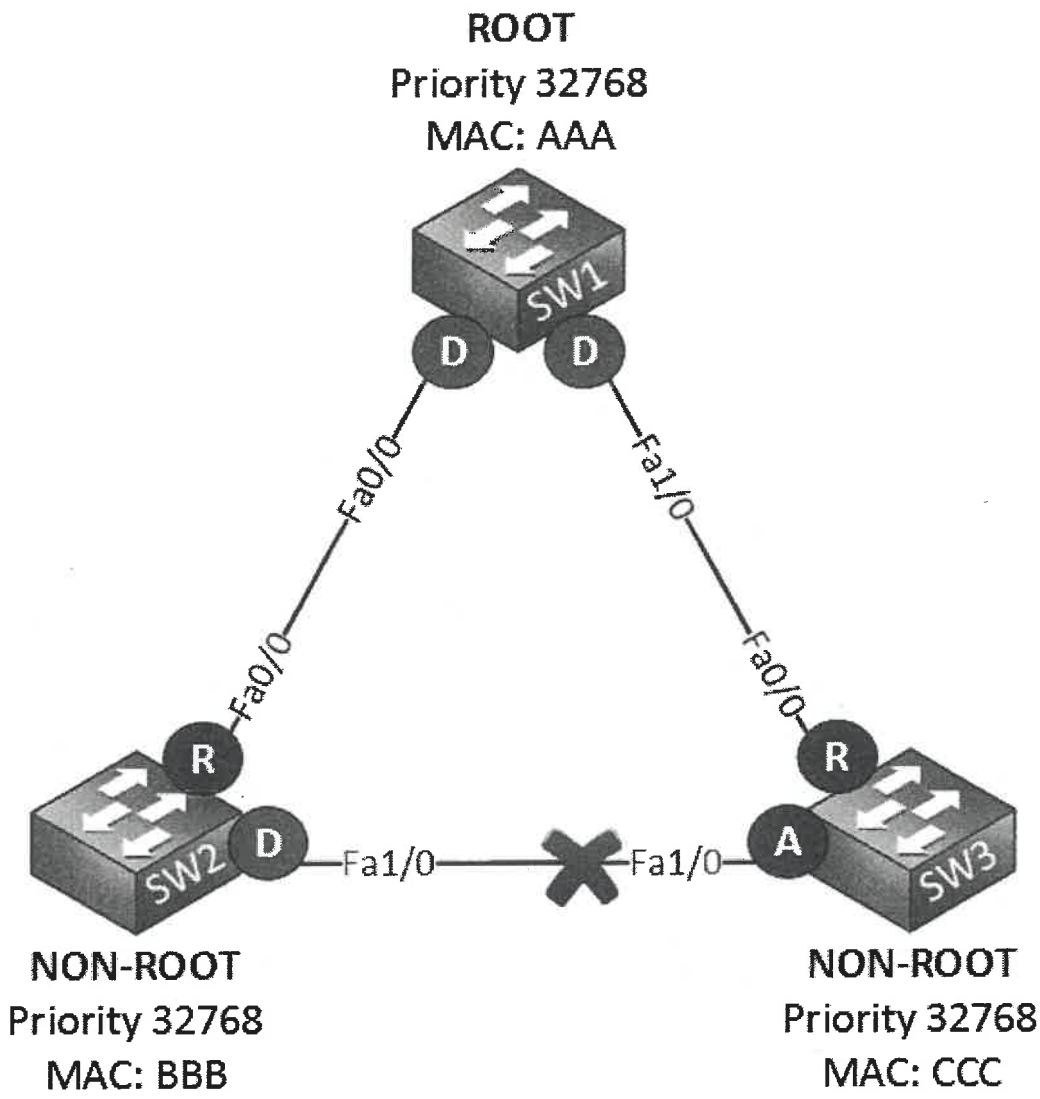


Hình 2. 5 Đường đi ngắn nhất đến Root port

Hiện tại, đã có các cổng designated trên root bridge và các cổng root trên các non-root bridge, tuy nhiên, vẫn có một vòng lặp, vì vậy cần tắt một cổng giữa SW2 và SW3 để phá vỡ vòng lặp đó. Xem xét lại bridge ID tốt nhất:

Bridge ID = Độ ưu tiên + địa chỉ MAC.

Thấp hơn là tốt hơn, cả hai switch có cùng độ ưu tiên, nhưng địa chỉ MAC của SW2 thấp hơn, điều đó có nghĩa là SW3 sẽ phải chặn cổng của mình, phá vỡ vòng lặp một cách hiệu quả.

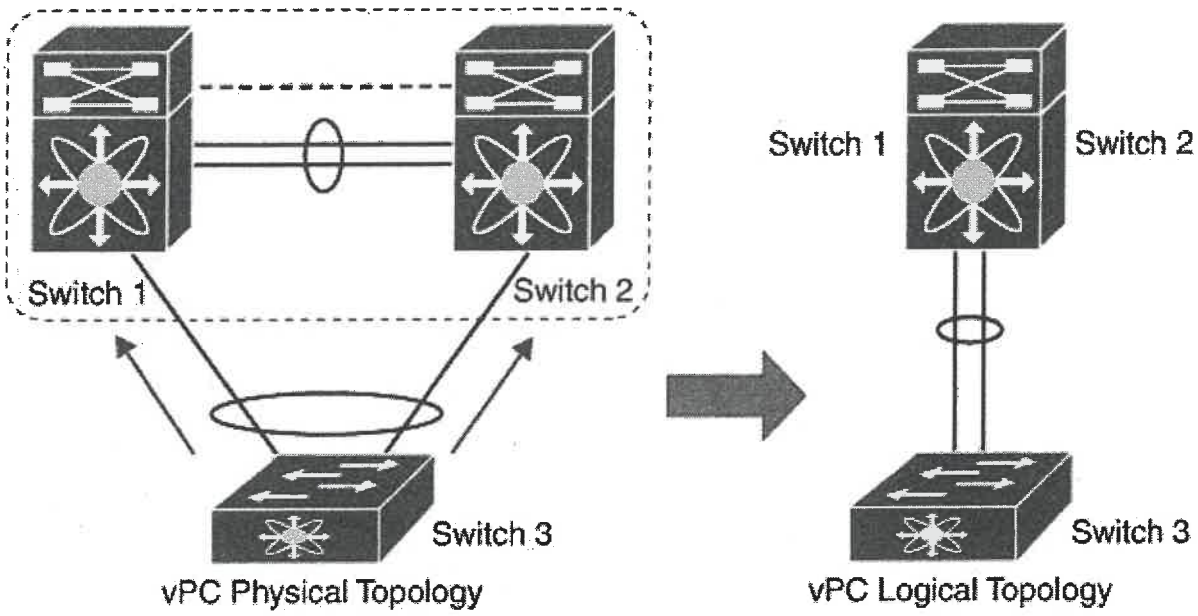


Hình 2.6 Vòng lặp đã được ngăn chặn thông qua giao thức STP

2.2.2 Virtual Port Channel (vPC)

2.2.2.1 Nguyên lý hoạt động của vPC

Virtual Port Channel (vPC) cho phép các liên kết được kết nối vật lý đến hai thiết bị khác nhau xuất hiện như một kênh cổng (port channel) duy nhất đối với một thiết bị thứ ba. Thiết bị thứ ba này có thể là một switch, máy chủ hoặc bất kỳ thiết bị mạng nào khác hỗ trợ kênh cổng. Một vPC có thể cung cấp đa đường lớp 2, cho phép tạo ra tính dự phòng và tăng băng thông hai chiều bằng cách kích hoạt nhiều đường dẫn song song giữa các nút và cho phép cân bằng tải lưu lượng. Chỉ có thể sử dụng các kênh cổng lớp 2 trong vPC, cấu hình các kênh cổng bằng cách sử dụng LACP hoặc cấu hình tĩnh không dùng giao thức.



Hình 2.7 Cấu trúc liên kết vật lý và logic của vPC

vPC cung cấp các lợi ích kỹ thuật sau [10]:

- Cho phép một thiết bị duy nhất sử dụng một kênh cổng trải rộng trên hai thiết bị upstream (thiết bị phía trên/phía nguồn).
- Loại bỏ các cổng bị chặn bởi Giao thức Spanning Tree (STP).
- Cung cấp một cấu trúc liên kết không có vòng lặp.
- Sử dụng tất cả băng thông uplink có sẵn.
- Cung cấp khả năng hội tụ nhanh nếu liên kết hoặc thiết bị gặp sự cố.
- Cung cấp khả năng phục hồi ở cấp độ liên kết.
- Đảm bảo tính sẵn sàng cao.

2.2.2.2 Thành phần kiến trúc của vPC

Kiến trúc vPC bao gồm các thành phần sau:

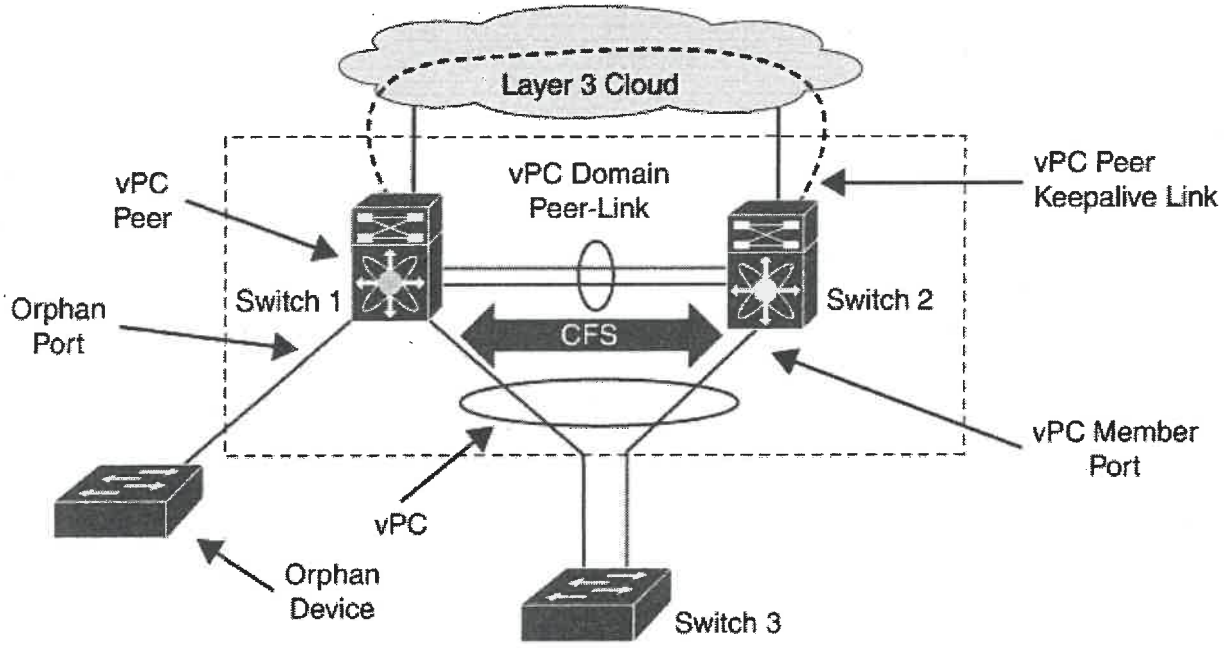
- *vPC*: Kênh cổng kết hợp giữa các thiết bị vPC peer và thiết bị downstream.
- *Thiết bị vPC peer*: Mô hình virtual Port Channel bao gồm hai thiết bị switch trong một cặp, cho phép các thiết bị khác kết nối tới hai switch này. Hai thiết bị này được gọi là vPC Peer và được kết nối với nhau thông qua vPC Peer Link.
- *Liên kết peer-link vPC*: Liên kết peer-link vPC mang lưu lượng vPC thiết yếu giữa các thiết bị chuyển mạch vPC peer và được sử dụng để đồng bộ hóa trạng thái giữa các thiết bị vPC peer. Liên kết peer-link là một kênh cổng và nên bao gồm ít nhất hai liên kết Ethernet 10 Gigabit chuyên dụng được kết cuối trên hai

mô-đun I/O khác nhau, nếu có thể, để đảm bảo tính sẵn sàng cao. Các giao diện có băng thông cao hơn (chẳng hạn như Ethernet 25 Gigabit, Ethernet 40 Gigabit, Ethernet 100 Gigabit, v.v.) cũng có thể được sử dụng để tạo thành kênh cổng.

- *Liên kết peer-keepalive vPC*: Liên kết peer-keepalive giám sát trạng thái hoạt động của các thiết bị vPC peer. Liên kết peer-keepalive gửi các thông điệp keepalive định kỳ, có thể cấu hình giữa các thiết bị vPC peer. Không có lưu lượng dữ liệu hoặc đồng bộ hóa nào di chuyển qua liên kết peer-keepalive vPC; lưu lượng duy nhất trên liên kết này là một thông điệp chỉ ra rằng thiết bị chuyển mạch nguồn đang hoạt động và chạy vPC.
- *Cổng thành viên vPC*: Một cổng được gán cho một nhóm kênh vPC. Các cổng này tạo thành kênh cổng ảo và được chia đều giữa các vPC peer.
- *Cổng vPC máy chủ*: Một giao diện máy chủ fabric extender thuộc về một vPC.
- *Cổng mồ côi (orphan port):* Một cổng không phải vPC, còn được gọi là cổng bị bỏ rơi, là một cổng không thuộc vPC.
- *Thiết bị mồ côi (orphan device*): Một thiết bị mồ côi là một thiết bị được kết nối với một miền vPC bằng các liên kết thông thường thay vì kết nối thông qua vPC.
- *Miền vPC*: Miền vPC bao gồm cả hai thiết bị vPC peer, liên kết peer-keepalive vPC và tất cả các kênh cổng trong vPC được kết nối với các thiết bị downstream.. Mỗi miền vPC có một số phiên bản vPC duy nhất được chia sẻ giữa hai thiết bị. Chỉ có hai thiết bị có thể là một phần của cùng một miền vPC, nhưng có thể có nhiều miền vPC trên một thiết bị duy nhất. ID miền có thể là bất kỳ giá trị nào từ 1 đến 1000 và cùng một giá trị phải được cấu hình trên cả hai thiết bị chuyển mạch tạo thành cặp vPC. Các thiết bị vPC peer sử dụng ID miền vPC để tự động gán một địa chỉ MAC hệ thống vPC duy nhất. Mỗi miền vPC có một địa chỉ MAC duy nhất được sử dụng làm định danh duy nhất cho hoạt động cụ thể liên quan đến vPC.
- *Dịch vụ Fabric Cisco*: Dịch vụ Fabric Cisco (CFS) là một cơ chế truyền tải trạng thái đáng tin cậy được sử dụng để đồng bộ hóa các hành động của các thiết bị vPC peer. CFS mang các thông điệp và gói tin cho nhiều tính năng liên quan đến vPC, chẳng hạn như STP và IGMP. Thông tin được mang trong các đơn vị dữ liệu giao thức (PDUs) CFS/CFS over Ethernet (CFSoE). Khi bật tính năng vPC, thiết bị sẽ tự động bật CFSoE và không cần phải cấu hình bất kỳ điều gì. Việc phân phối CFSoE cho vPC không cần khả năng phân phối qua IP hoặc các vùng CFS. Các thông điệp CFS cung cấp một bản sao cấu hình trên thiết bị vPC peer cục bộ cho thiết bị vPC peer từ xa. Tất cả các địa chỉ MAC cho các VLAN được

cấu hình trên cả hai thiết bị đều được đồng bộ hóa giữa các thiết bị vPC peer bằng giao thức CFSoE. Thiết bị vPC chính đồng bộ hóa trạng thái STP trên thiết bị vPC phụ bằng cách sử dụng Cisco Fabric Services over Ethernet (CFSoE).

- *VLAN vPC*: Các VLAN được phép trên vPC được gọi là VLAN vPC. Các VLAN này cũng phải được phép trên vPC peer-link.
- *VLAN không phải vPC*: Bất kỳ VLAN STP nào không được truyền qua vPC peer-link.



Hình 2. 8 Các thành phần trong vPC

2.3. Sự phát triển của thiết kế mạng trong trung tâm dữ liệu

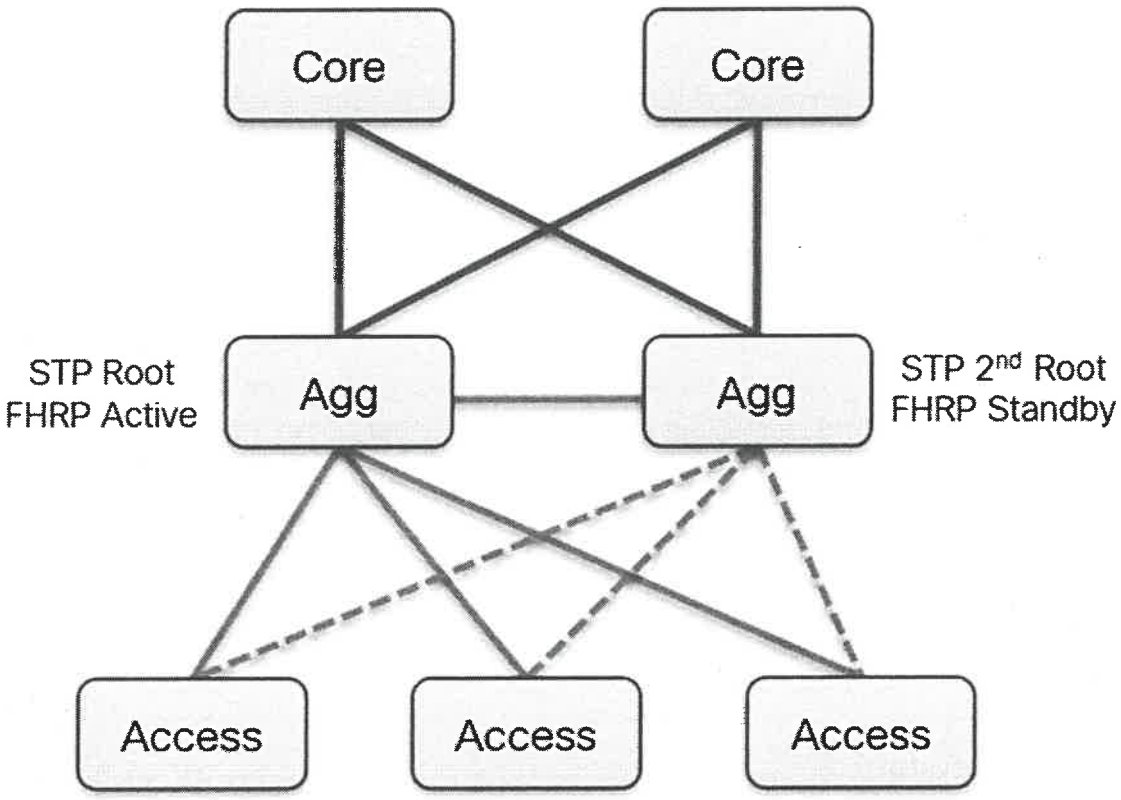
Phần này cung cấp cái nhìn tổng quan về các công nghệ mới trong trung tâm dữ liệu, cách chúng hỗ trợ các nguyên tắc kiến trúc đã được đề ra trước đó và cách chúng ảnh hưởng đến thiết kế và triển khai cơ sở hạ tầng [11].

2.3.1 Mô hình ba lớp

Theo mô hình này, mạng được chia thành ba lớp: Lõi (core), gom (Agg) và truy cập (Access), các thiết bị mạng ở mỗi lớp này đóng vai trò khác nhau và có thông số kỹ thuật khác nhau

Các mạng dựa trên cấu trúc liên kết (topology) kiểu này là các mạng có phân đoạn lưu lượng ở cấp độ VLAN. Mỗi VLAN tạo thành một miền quảng bá (broadcast domain) riêng biệt và được ánh xạ một-một với một mạng con (subnet) dành riêng cho nó. Ví dụ, có thể có VLAN 20 cho nhân viên phòng tài chính, được ánh xạ tới mạng con

192.168.20.0/24. Nếu nhân viên của phòng này ở các vị trí thực tế khác nhau, việc mở rộng VLAN giữa các switch lớp truy cập (access layer) là điều tất yếu.



Hình 2. 9 Mô hình mạng ba lớp

Các VLAN như vậy tất nhiên sẽ trải dài qua các switch lớp AGG, vốn đóng vai trò là cổng vào mặc định (default gateway). Ngoài ra cũng sử dụng các giao thức bổ sung như STP (Spanning Tree Protocol), VRRP (Virtual Router Redundancy Protocol) hoặc HSRP (Hot Standby Router Protocol). Vấn đề mà thiết kế này mang lại là: một sự cố ở một phần của mạng sẽ lan truyền đến nhiều thiết bị vì đã mở rộng các miền quảng bá.

Một vấn đề khác là lượng lưu lượng bắc-nam (north-south traffic) tăng lên, điều này không tốt. Ước tính rằng trong các mạng campus, và đặc biệt là trong các mạng trung tâm dữ liệu (data center), khoảng 80% lưu lượng là lưu lượng đông-tây (east-west traffic) (giao tiếp giữa các máy khách, máy chủ đến máy chủ, sao chép dữ liệu, di chuyển máy ảo bằng VMotion, v.v.). Cấu trúc mạng ba lớp bắt buộc phải có một lượng lớn lưu lượng bắc-nam bổ sung cho các giao tiếp vốn dĩ là lưu lượng đông-tây.

Vấn đề cuối cùng phải đối mặt ở đây là sự hiện diện của STP. Mặc dù STP rất quan trọng để duy trì một mạng không có vòng lặp và cung cấp tính dự phòng cho liên kết, sự hiện diện của nó cũng mang lại một số vấn đề:

- *Sử dụng băng thông không hiệu quả*: Để ngăn chặn vòng lặp, STP phải đặt một số cổng vào trạng thái chặn (blocking). Điều này có nghĩa là các liên kết vật lý đó không được sử dụng để chuyển tiếp dữ liệu, dẫn đến lãng phí băng thông tiềm năng. Trong một mạng có nhiều đường dự phòng, một phần đáng kể tài nguyên mạng có thể không được sử dụng.
- *Thời gian hội tụ chậm* (đối với các phiên bản STP cũ): Giao thức STP gốc (IEEE 802.1D) có thời gian hội tụ chậm. Điều này có nghĩa là khi có sự thay đổi trong cấu trúc mạng (ví dụ: một liên kết bị lỗi hoặc một switch mới được thêm vào), mạng có thể mất từ 30 đến 50 giây để ổn định lại. Trong thời gian này, một số phần của mạng có thể không truy cập được. Các phiên bản cải tiến như RSTP (Rapid Spanning Tree Protocol) và MSTP (Multiple Spanning Tree Protocol) đã cải thiện đáng kể thời gian hội tụ, nhưng quá trình tính toán lại cấu trúc liên kết vẫn cần thiết.
- *Tải CPU trên thiết bị chuyển mạch*: Trong các mạng lớn hoặc phức tạp, đặc biệt khi sử dụng các biến thể STP như PVST+ (Per-VLAN Spanning Tree Plus) – chạy một thực thể STP riêng cho mỗi VLAN – các switch phải xử lý một số lượng lớn các gói tin BPDU (Bridge Protocol Data Unit) và thực hiện nhiều tính toán STP. Điều này có thể gây áp lực đáng kể lên CPU của switch, ảnh hưởng đến hiệu suất tổng thể.
- *Độ phức tạp trong thiết kế và quản lý*: Việc cấu hình và khắc phục sự cố STP, đặc biệt trong các mạng lớn với nhiều VLAN và các phiên bản STP khác nhau, có thể phức tạp. Việc xác định root bridge, root port, designated port và các cổng bị chặn đòi hỏi sự hiểu biết sâu về giao thức.

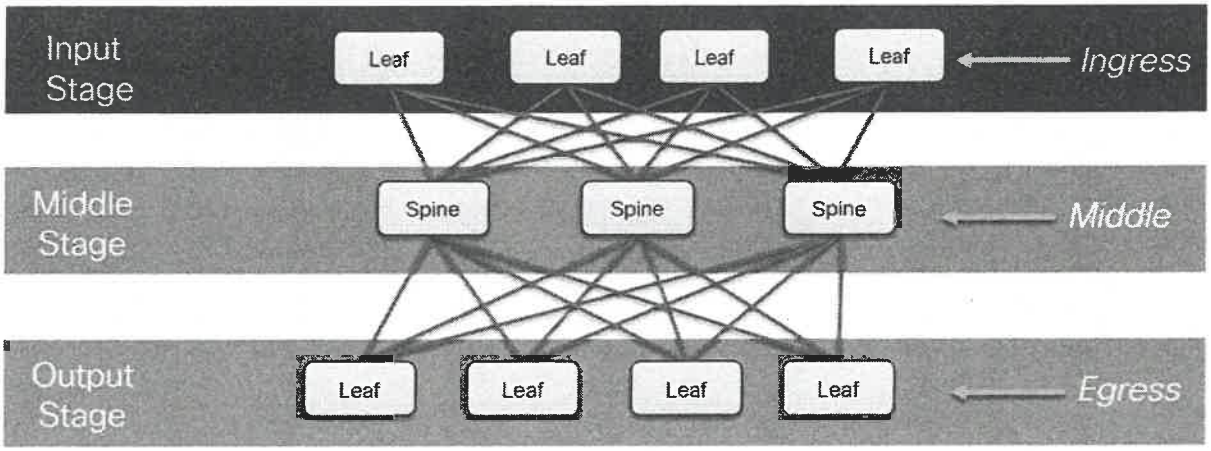
2.3.2 Mô hình Clos

Mạng Clos được đặt theo tên của nhà nghiên cứu Charles Clos thuộc Bell Labs, người lần đầu tiên đề xuất thiết kế mạng của mình vào năm 1952. Ông đã chứng minh rằng mô hình của mình có thể giải quyết các thách thức về độ tin cậy và chi phí.

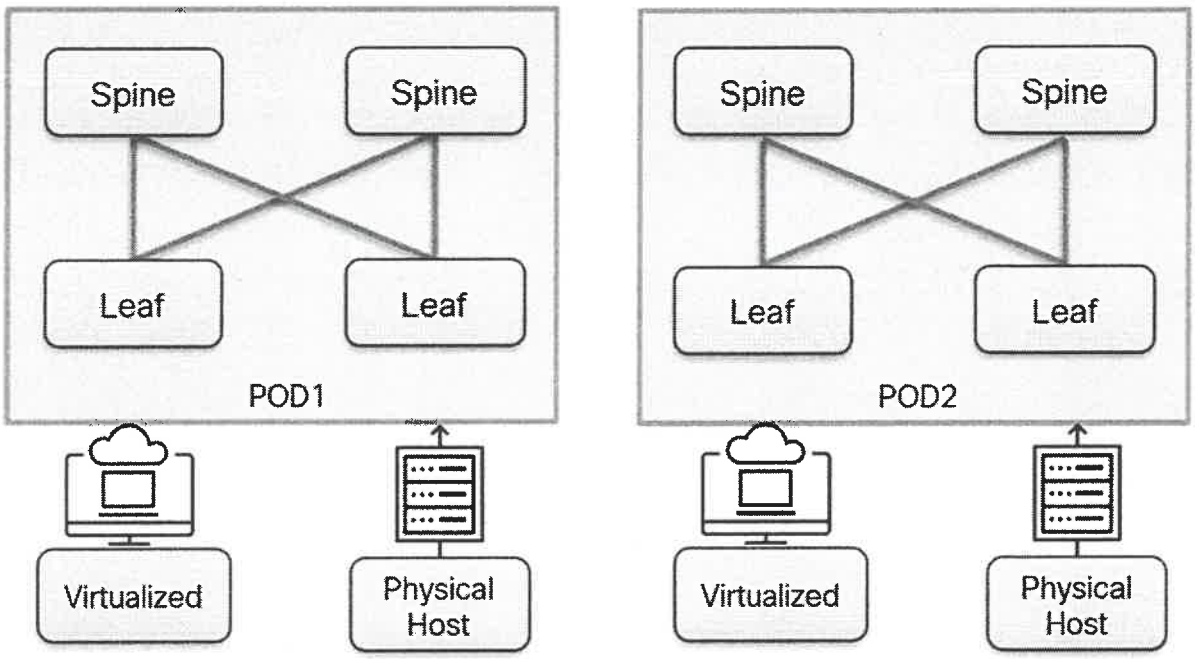
Clos đã đưa ra một thiết kế mạng để cung cấp khả năng kết nối không chặn (nonblocking), mọi-điểm-đến-mọi-điểm (any-to-any) nhằm giảm thiểu số lượng điểm chéo (crosspoints) cần thiết để hỗ trợ chuyển mạch.

Clos đã sử dụng lý thuyết toán học để chứng minh rằng có thể đạt được kết nối không chặn trong một mảng chuyển mạch, ngày nay được gọi là fabric (kết cấu chuyển mạch). Để đạt được kết nối này, các thiết bị chuyển mạch được tổ chức thành một kiến trúc ba tầng bao gồm một tầng giữa được kẹp giữa tầng vào (ingress - đầu vào) và tầng ra (egress - đầu ra). Trong thiết kế này, mỗi switch đầu vào kết nối với mỗi switch tầng

giữa, và mỗi switch tầng giữa lại kết nối với mỗi switch đầu ra. Kết quả là một cấu trúc liên kết không chặn đòi hỏi ít điểm chéo hơn so với các mạng chuyển mạch thông thường hơn vào thời điểm đó.



Hình 2. 10 Mô hình Clos 3 lớp



Hình 2. 11 Mô hình Spine-Leaf thường được sử dụng trong DC hiện đại

Trong nhiều trung tâm dữ liệu ngày nay, kiến trúc Clos được triển khai theo mô hình Spine - Leaf, trong đó lớp xương sống (spine layer) đại diện cho các thiết bị chuyển mạch ở tầng giữa và lớp lá (leaf layer) đại diện cho các thiết bị chuyển mạch ở cả tầng vào (ingress) và tầng ra (egress).

Ưu điểm của mô hình Leaf-Spine

- Ưu điểm đầu tiên của Spine-Leaf là cung cấp nhiều tuyến đường giữa các Leaf switch. Mô hình Spine-Leaf thường được triển khai với các liên kết Layer 3. Tất

cả các link đều được sử dụng để cân bằng tải nhờ giao thức ECMP do các kết nối sử dụng các cổng có băng thông bằng nhau và có chính xác 2 bước nhảy giữa các Leaf switch. Với kiến trúc này, bất kể máy chủ nào được kết nối với máy chủ nào, lưu lượng truy cập của nó luôn phải vượt qua cùng một số thiết bị để đến máy chủ khác (trừ khi máy chủ khác nằm trên cùng một Leaf). Cách tiếp cận này giữ độ trễ ở mức có thể dự đoán được vì một lưu lượng chỉ phải đi tới một Spine Switch và một Leaf Switch khác để đến đích.

- Tiếp theo là tính dự phòng cao. Nếu 1 thiết bị Spine bị lỗi, nó chỉ giảm 1 phần nhỏ hiệu suất của mạng mà không ảnh hưởng đến dịch vụ. Nếu 1 Leaf switch bị lỗi, nó chỉ ảnh hưởng đến các máy chủ đang kết nối tới Leaf switch đó.
- Một ưu điểm khác tính mở rộng cao. Nếu cần thêm băng thông, chỉ cần thêm Spine switch, nếu cần thêm nhiều máy chủ, chỉ cần thêm Leaf switch mà không phải thiết kế lại toàn bộ hệ thống.

2.4. Ứng dụng VXLAN trong các trung tâm dữ liệu

2.4.1 Ảo hoá mạng, overlay và underlay

2.4.1.1 Khái niệm ảo hoá mạng

Ảo hóa mạng là quá trình chuyển đổi mạng vốn phụ thuộc vào phần cứng thành mạng dựa trên phần mềm. Giống như mọi hình thức ảo hóa CNTT, mục tiêu cơ bản của ảo hóa mạng là tạo ra một lớp trừu tượng giữa phần cứng vật lý và các ứng dụng và dịch vụ sử dụng phần cứng đó.

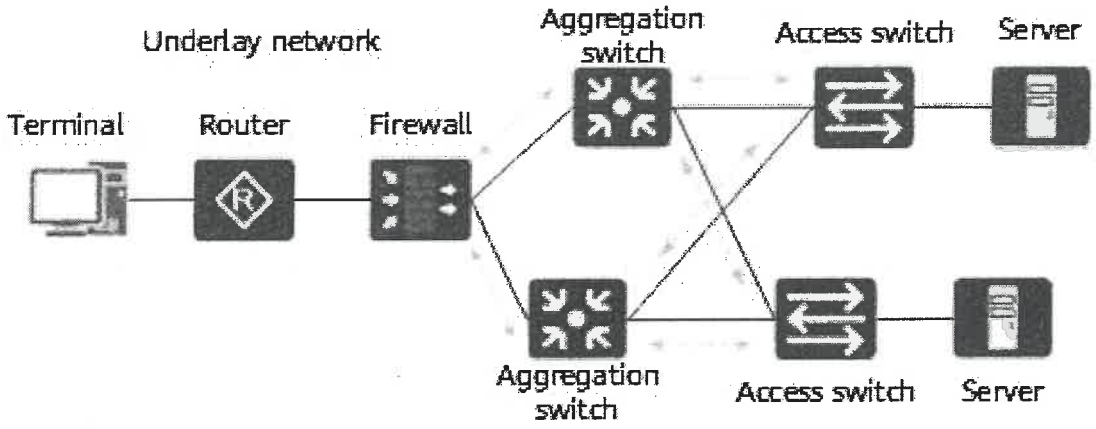
Cụ thể hơn, ảo hóa mạng cho phép các chức năng mạng, tài nguyên phần cứng và tài nguyên phần mềm được phân phối độc lập với phần cứng—như một mạng ảo. Nó có thể được sử dụng để hợp nhất nhiều mạng vật lý, chia nhỏ một mạng như vậy hoặc kết nối các máy ảo (VM) với nhau.

Một trong những cách tiếp cận quan trọng để thực hiện ảo hóa mạng là thông qua việc phân tách thành hai lớp kiến trúc chính: Overlay (mạng ảo) và Underlay (hạ tầng vật lý).

2.4.1.2 Underlay

Như tên gọi, mạng underlay là hạ tầng vật lý cơ bản của các mạng overlay.

Như hình dưới đây, một mạng underlay là một mạng vật lý bao gồm nhiều loại thiết bị và chịu trách nhiệm truyền tải các gói dữ liệu trong mạng.



Hình 2. 12 Mạng underlay điển hình

Trên một mạng underlay, các thiết bị như bộ chuyển mạch, bộ định tuyến, bộ cân bằng tải và tường lửa có thể được kết nối với nhau. Tuy nhiên, các giao thức định tuyến phải được sử dụng để đảm bảo kết nối IP giữa các thiết bị này.

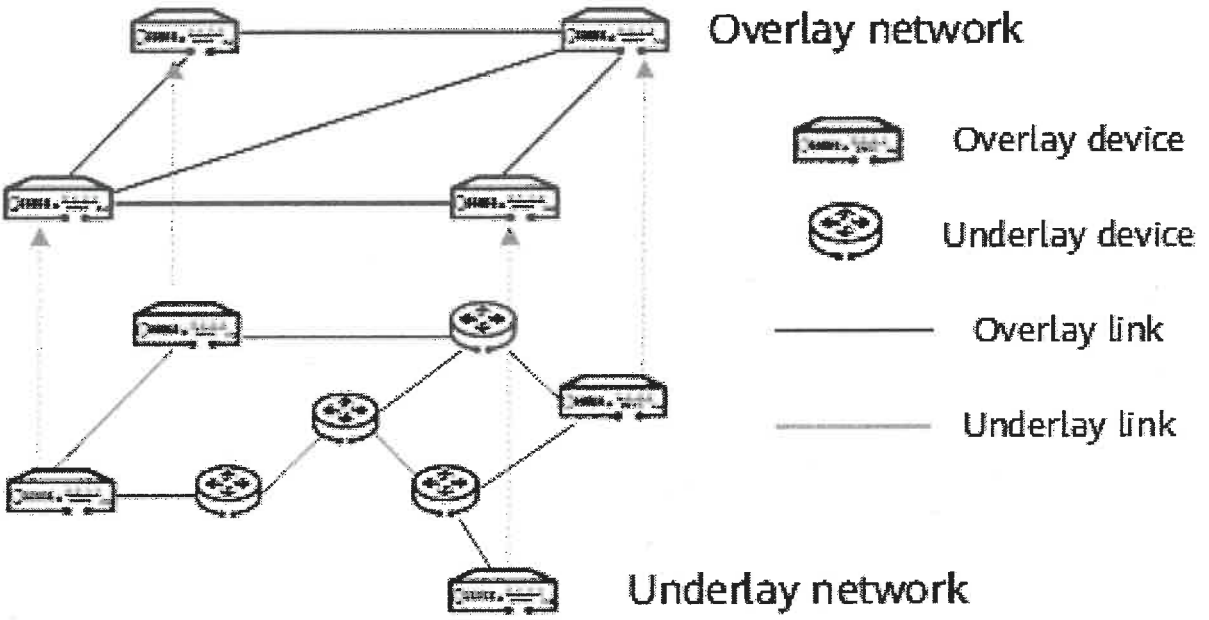
Mạng underlay có thể là mạng Lớp 2 hoặc Lớp 3. Một ví dụ điển hình của mạng Lớp 2 là mạng Ethernet, nơi các VLAN được tạo ra. Internet là một mạng Lớp 3 điển hình. Giao thức Open Shortest Path First (OSPF) hoặc Intermediate System to Intermediate System (IS-IS) được sử dụng cho định tuyến bên trong một hệ thống tự trị (AS), trong khi Border Gateway Protocol (BGP) được sử dụng để truyền tải và kết nối giữa các AS. Khi công nghệ phát triển, các mạng underlay cũng có thể được xây dựng bằng cách sử dụng công nghệ Multiprotocol Label Switching (MPLS), một công nghệ mạng diện rộng (WAN) chạy ở Lớp 2 và Lớp 3.

Tuy nhiên, các mạng underlay truyền thống dựa trên phần cứng có những hạn chế sau:

- Phần cứng chuyển tiếp các gói dữ liệu dựa trên địa chỉ IP đích. Do đó, việc chuyển tiếp gói tin phụ thuộc nhiều vào đường truyền vật lý.
- Khi các dịch vụ được thêm vào hoặc thay đổi, các kết nối mạng hiện có cần được sửa đổi. Việc cấu hình lại tốn thời gian.
- Tính bảo mật của các mạng riêng thường được đảm bảo trên Internet.
- Việc phân chia và phân đoạn mạng phức tạp và không thể đạt được việc phân bổ tài nguyên mạng theo yêu cầu.
- Việc chuyển tiếp đa đường phức tạp và nhiều mạng underlay không thể được tích hợp để thực hiện cân bằng tải.

2.4.1.3 Overlay

Để loại bỏ những hạn chế của mạng underlay, các mạng ảo, giao thức ảo có thể được tạo ra trên mạng underlay bằng cách sử dụng các công nghệ ảo hóa mạng.



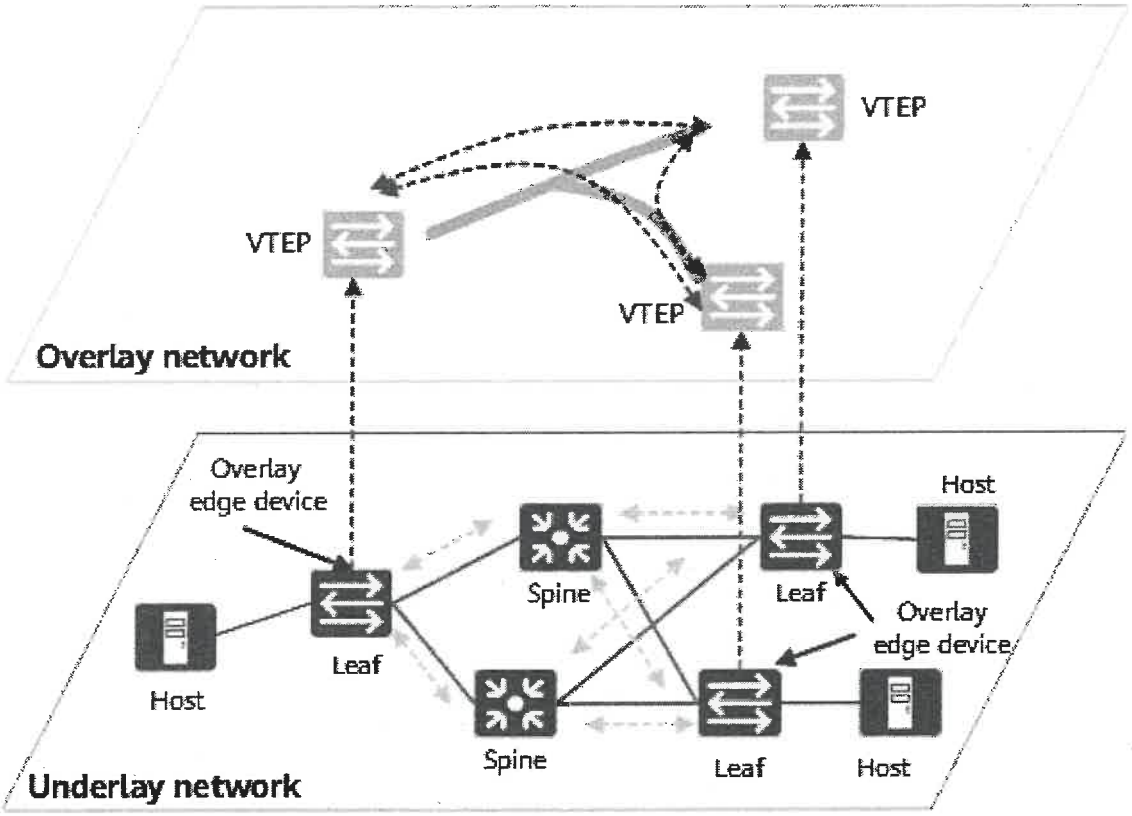
Hình 2. 13 Cấu trúc liên kết mạng Overlay

Các thiết bị trên mạng overlay được kết nối với nhau bằng các liên kết logic khi cần thiết, tạo thành các cấu trúc liên kết overlay.

Các đường hầm được thiết lập giữa các thiết bị mạng overlay được kết nối với nhau. Khi gửi một gói dữ liệu, thiết bị nguồn thêm một tiêu đề IP mới và một tiêu đề đường hầm vào gói dữ liệu và che chắn tiêu đề IP bên trong. Gói dữ liệu sau đó được chuyển tiếp dựa trên tiêu đề IP bên ngoài. Khi gói dữ liệu được nhận bởi thiết bị đích, thiết bị này sẽ loại bỏ tiêu đề IP bên ngoài và tiêu đề đường hầm để lấy lại gói dữ liệu ban đầu. Trong quá trình này, mạng overlay không nhận biết được mạng underlay.

Các mạng overlay hỗ trợ nhiều giao thức và tiêu chuẩn ảo hóa mạng, bao gồm VXLAN (Virtual Extensible LAN), NVGRE (Network Virtualization using Generic Routing Encapsulation), SST (single spanning tree), GRE (Generic Routing Encapsulation), Network Virtualization over Layer 3 (NV03) và EVPN (Ethernet Virtual Private Network).

Khi kiến trúc trung tâm dữ liệu phát triển, hầu hết các trung tâm dữ liệu sử dụng kiến trúc spine-leaf để xây dựng mạng underlay và sử dụng công nghệ VXLAN để xây dựng các mạng overlay kết nối. Các gói dịch vụ được truyền trên các mạng overlay VXLAN, vốn tách biệt khỏi các lớp truyền tải vật lý.



Hình 2. 14 Mạng Overlay trong trung tâm dữ liệu

Các nút leaf và spine được kết nối hoàn toàn (fully meshed) sao cho các đường dẫn ECMP luôn sẵn sàng để đảm bảo tính sẵn sàng cao của mạng.

Các nút leaf đóng vai trò là các nút truy cập để kết nối các thiết bị mạng khác nhau trên mạng underlay với mạng VXLAN. Các nút leaf cũng là các thiết bị biên của mạng overlay và đóng vai trò là các điểm cuối đường hầm VXLAN (VTEP).

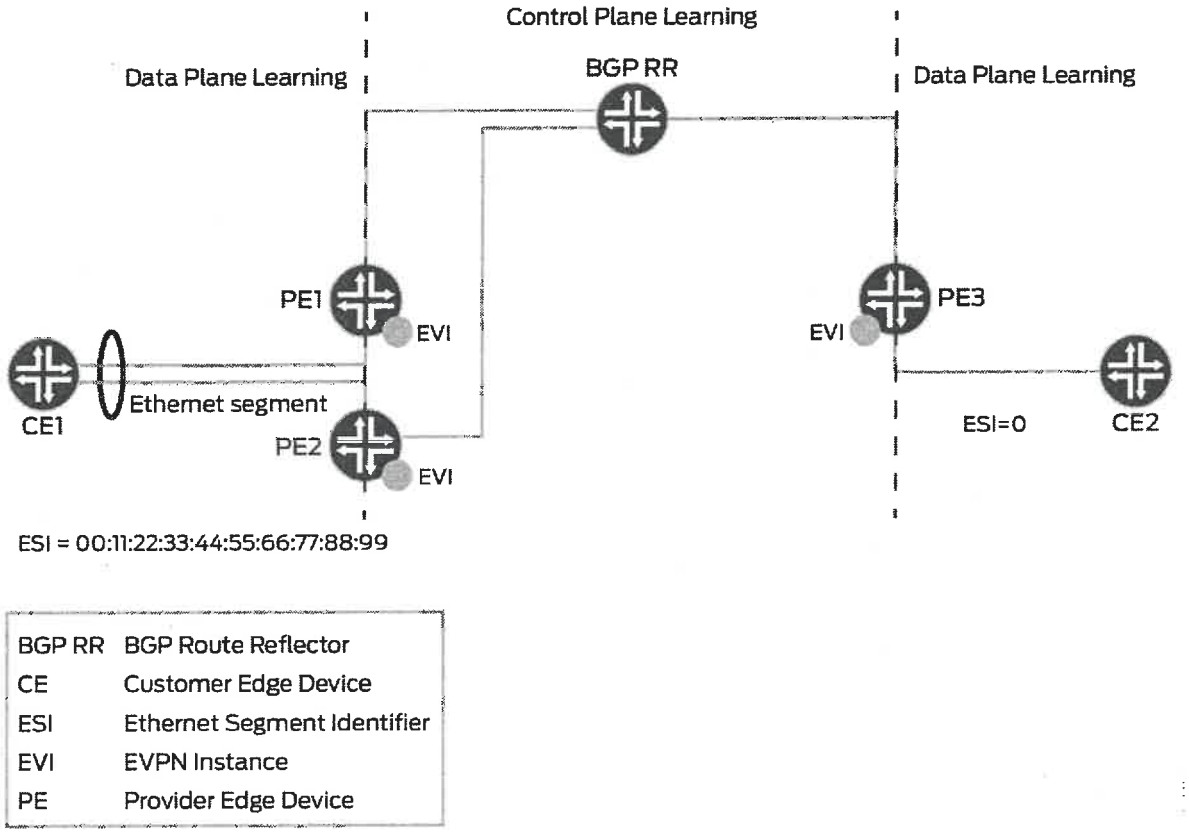
Các nút spine là các nút lõi của mạng trung tâm dữ liệu. Chúng cung cấp khả năng chuyển tiếp tốc độ cao và kết nối các nút leaf thông qua các giao diện tốc độ cao.

2.4.2 Công nghệ EVPN-VXLAN

2.4.2.1 Giới thiệu chung về EVPN

EVPN là một công nghệ mới dựa trên tiêu chuẩn, cung cấp kết nối bắc cầu đa điểm ảo (virtual multi-point bridged connectivity) giữa các miền Lớp 2 khác nhau qua mạng trục IP hoặc IP/MPLS. Tương tự như các công nghệ VPN khác, chẳng hạn như IP VPN và VPLS, các phiên bản EVPN (EVIs) được cấu hình trên các bộ định tuyến PE để duy trì sự tách biệt dịch vụ logic giữa các khách hàng. Các PE kết nối với các thiết bị CE, có thể là bộ định tuyến, bộ chuyển mạch hoặc máy chủ. Các bộ định tuyến PE sau đó trao đổi thông tin về khả năng tiếp cận (reachability) bằng cách sử dụng Giao thức cổng biên đa giao thức (Multi-Protocol BGP - MP-BGP) và lưu lượng được đóng gói sẽ được

chuyển tiếp giữa các PE. Bởi vì các yếu tố của kiến trúc này tương đồng với các công nghệ VPN khác, EVPN có thể được giới thiệu và tích hợp liền mạch vào các môi trường dịch vụ hiện có.



Hình 2. 15 Tổng quan về EVPN

Công nghệ EVPN cung cấp các cơ chế cho Kết nối liên Trung tâm Dữ liệu (Data Center Interconnect - DCI) thế hệ tiếp theo bằng cách thêm các quy trình mặt phẳng điều khiển mở rộng để trao đổi thông tin Lớp 2 (địa chỉ MAC) và Lớp 3 (địa chỉ IP) giữa các Bộ định tuyến biên trung tâm dữ liệu (Data Center Border Routers - DCBR) tham gia. Các tính năng này giúp giải quyết một số thách thức của DCI, chẳng hạn như tính di động liền mạch của máy ảo (VM) và định tuyến IP tối ưu. Tính di động liền mạch của VM đề cập đến thách thức của việc mở rộng Lớp 2 và duy trì kết nối khi VM di chuyển, còn định tuyến IP tối ưu đề cập đến thách thức hỗ trợ hành vi cổng mặc định cho lưu lượng đi ra của VM và tránh định tuyến tam giác cho lưu lượng đi vào của VM.

Công nghệ EVPN được nhà điều hành trung tâm dữ liệu sử dụng để cung cấp các dịch vụ đa thuê bao (multi-tenancy), linh hoạt và có khả năng phục hồi cao, có thể mở rộng theo yêu cầu. Sự linh hoạt và khả năng phục hồi này có thể yêu cầu sử dụng tài nguyên điện toán giữa các trung tâm dữ liệu vật lý khác nhau cho một dịch vụ duy nhất (mở rộng Lớp 2) và di chuyển VM.

EVPN hỗ trợ đa kết nối toàn hoạt động (all-active multihoming) cho phép một thiết bị CE kết nối với hai hoặc nhiều bộ định tuyến PE sao cho lưu lượng được chuyển tiếp bằng tất cả các liên kết giữa các thiết bị. Điều này cho phép CE cân bằng tải lưu lượng đến nhiều bộ định tuyến PE. Quan trọng hơn, nó cho phép một PE từ xa cân bằng tải lưu lượng đến các PE đa kết nối thông qua mạng lõi. Việc cân bằng tải các luồng lưu lượng giữa các trung tâm dữ liệu này được gọi là aliasing (bí danh). EVPN cũng có các cơ chế ngăn chặn vòng lặp của lưu lượng broadcast, unknown unicast và multicast (BUM) trong cấu trúc liên kết đa kết nối toàn hoạt động.

Đa kết nối cung cấp khả năng dự phòng trong trường hợp liên kết truy cập hoặc một trong các bộ định tuyến PE bị lỗi. Trong cả hai trường hợp, luồng lưu lượng từ CE đến PE sử dụng các liên kết hoạt động còn lại. Đối với lưu lượng theo hướng ngược lại, PE từ xa cập nhật bảng chuyển tiếp của nó để gửi lưu lượng đến các PE hoạt động còn lại được kết nối với phân đoạn Ethernet đa kết nối. EVPN cung cấp một cơ chế hội tụ nhanh để thời gian thực hiện điều chỉnh này không phụ thuộc vào số lượng địa chỉ MAC mà PE đã học được.

Mặt phẳng điều khiển MP-BGP của EVPN cho phép di chuyển động các máy ảo đang hoạt động từ trung tâm dữ liệu này sang trung tâm dữ liệu khác, còn được gọi là di chuyển VM (VM motion). Sau khi một VM được chuyển đến máy chủ/trình ảo hóa đích, nó sẽ truyền một Gratuitous ARP để cập nhật bảng chuyển tiếp Lớp 2 của PE tại trung tâm dữ liệu đích. Sau đó, PE truyền một bản cập nhật tuyến MAC đến tất cả các PE từ xa, từ đó cập nhật bảng chuyển tiếp của chúng. Bằng cách này, EVPN theo dõi sự di chuyển của VM, còn được gọi là tính di động MAC. EVPN cũng có các cơ chế để phát hiện và dừng hiện tượng MAC flapping (địa chỉ MAC thay đổi liên tục giữa các cổng).

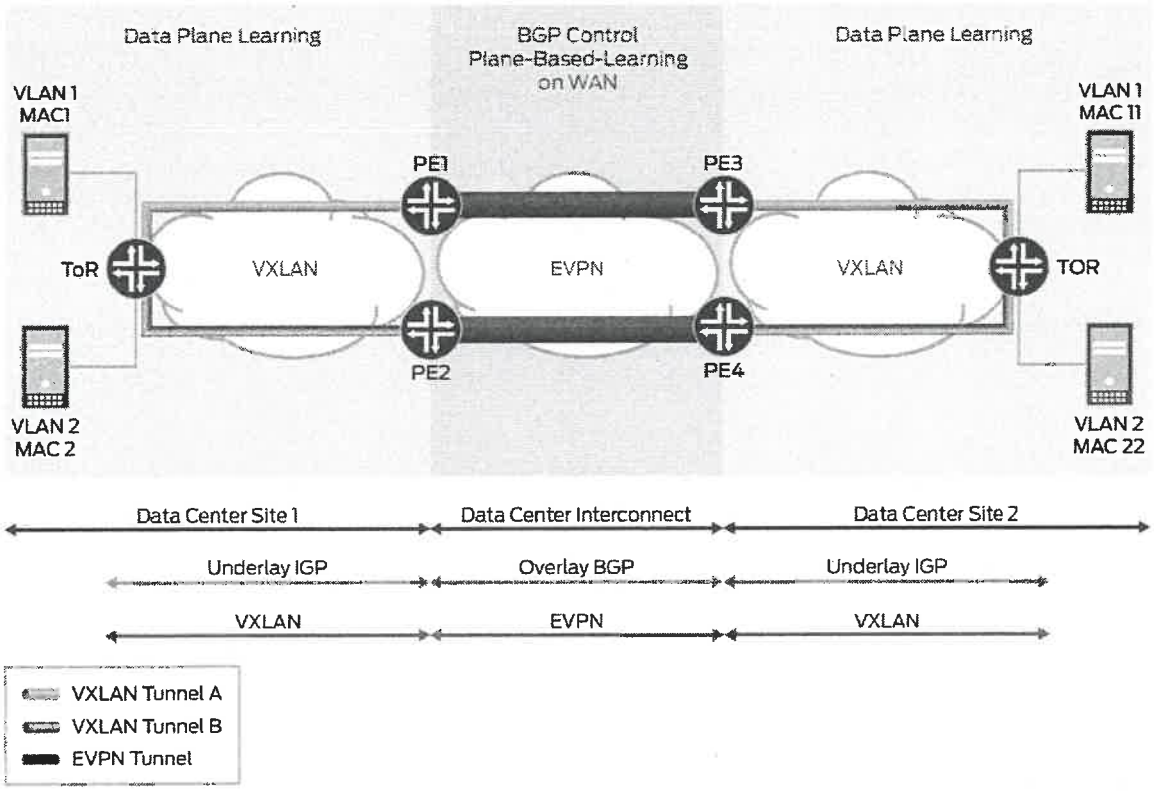
Công nghệ EVPN là một công nghệ giới thiệu khái niệm định tuyến địa chỉ MAC bằng MP-BGP qua lõi MPLS. Một số lợi ích quan trọng của việc sử dụng EVPN bao gồm:

- Khả năng có thiết bị biên đa kết nối kép hoạt động (dual active multihomed edge device).
- Cung cấp cân bằng tải trên các liên kết kép hoạt động (dual-active links).
- Cung cấp tính di động của địa chỉ MAC.
- Cung cấp đa thuê bao.
- Cung cấp aliasing.

2.4.2.2 Tích hợp EVPN-VXLAN

VXLAN định nghĩa một sơ đồ tạo đường hầm để phủ mạng Lớp 2 lên trên mạng Lớp 3. Nó cho phép chuyển tiếp tối ưu các khung Ethernet với sự hỗ trợ đa đường cho lưu lượng unicast và multicast bằng cách sử dụng đóng gói UDP/IP để tạo đường hầm, và chủ yếu được sử dụng cho kết nối trong nội bộ trung tâm dữ liệu.

Mặt khác, một đặc điểm độc đáo của EVPN là việc học địa chỉ MAC giữa các thiết bị PE xảy ra trong mặt phẳng điều khiển. Một địa chỉ MAC mới được phát hiện từ thiết bị CE sẽ được PE cục bộ quảng bá, sử dụng MP-BGP, đến tất cả các thiết bị PE từ xa. Phương pháp này khác với các giải pháp VPN Lớp 2 hiện có như VPLS, vốn học bằng cách làm tràn (flooding) lưu lượng unicast không xác định trong mặt phẳng dữ liệu. Phương pháp học MAC dựa trên mặt phẳng điều khiển này là yếu tố hỗ trợ chính cho nhiều tính năng hữu ích do EVPN cung cấp.



Hình 2. 16 Tích hợp EVPN-VXLAN

Bởi vì việc học MAC được xử lý trong mặt phẳng điều khiển, điều này mang lại cho EVPN sự linh hoạt để hỗ trợ các công nghệ đóng gói mặt phẳng dữ liệu khác nhau giữa các PE. Điều này quan trọng vì không phải mạng trục nào cũng có thể đang chạy MPLS, đặc biệt là trong các mạng doanh nghiệp.

Hiện nay có rất nhiều sự quan tâm đến EVPN vì nó giải quyết nhiều thách thức mà các nhà khai thác mạng phải đối mặt khi xây dựng các trung tâm dữ liệu để cung cấp dịch vụ đám mây và ảo hóa. Ứng dụng chính của EVPN là Kết nối Liên Trung tâm Dữ liệu (DCI), khả năng mở rộng kết nối Lớp 2 giữa các trung tâm dữ liệu khác nhau được triển khai để cải thiện hiệu suất phân phối lưu lượng ứng dụng đến người dùng cuối và phục hồi sau thảm họa.

Mặc dù có nhiều công nghệ DCI khác nhau, EVPN có lợi thế bổ sung so với các công nghệ MPLS khác nhờ các tính năng độc đáo của nó, chẳng hạn như dự phòng chủ active-active redundancy, aliasing (bí danh) và rút MAC hàng loạt (mass MAC withdrawal). Kết quả là, để cung cấp giải pháp cho DCI, VXLAN được tích hợp với EVPN.

Mỗi mạng VXLAN, được kết nối với lõi MPLS hoặc IP, chạy một phiên bản độc lập của mặt phẳng điều khiển IGP. Mỗi thiết bị PE tham gia vào phiên bản mặt phẳng điều khiển IGP của mạng VXLAN của nó. Ở đây, mỗi khách hàng là một trung tâm dữ liệu nên nó có bộ định tuyến ảo riêng cho VXLAN underlay.

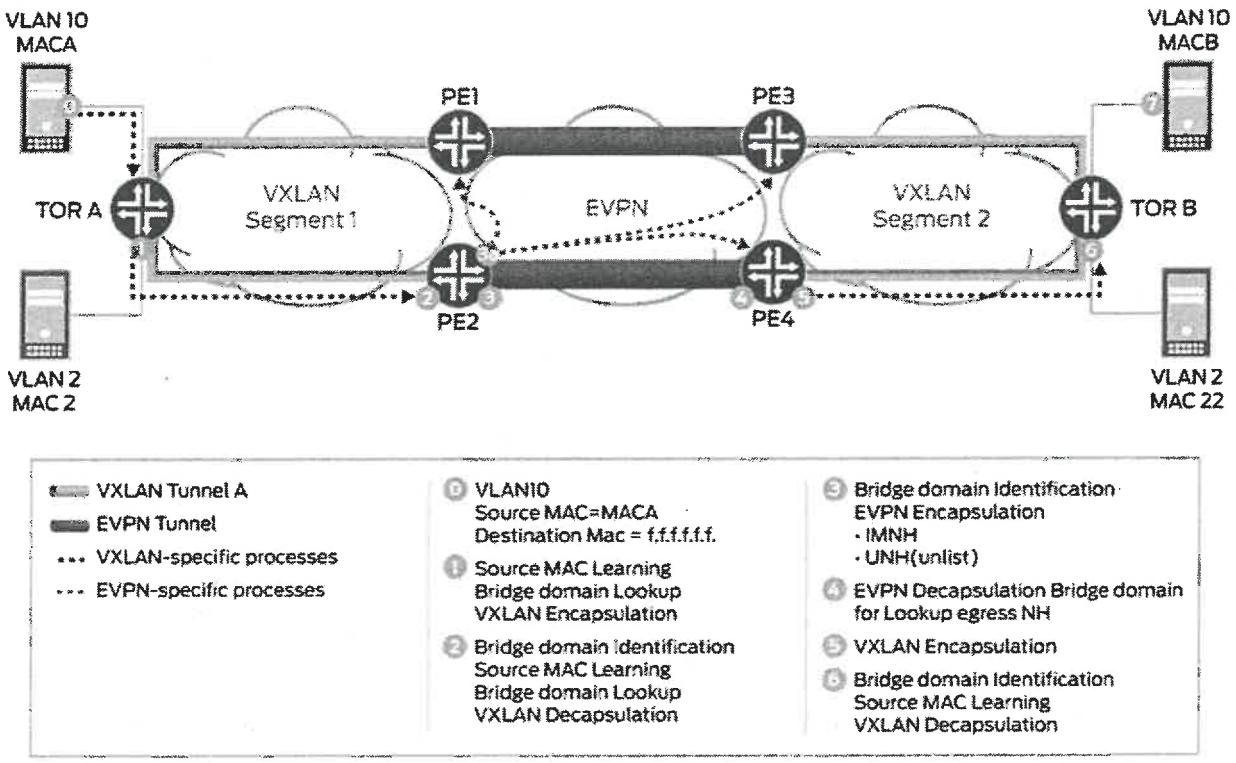
Mỗi nút PE có thể kết thúc việc đóng gói mặt phẳng dữ liệu VXLAN nơi mỗi VNI hoặc VSID được ánh xạ tới một miền cầu nối (bridge domain). Bộ định tuyến PE thực hiện việc học mặt phẳng dữ liệu trên lưu lượng nhận được từ mạng VXLAN.

Mỗi nút PE triển khai EVPN để phân phối các địa chỉ MAC của máy khách đã học được qua đường hầm VXLAN vào BGP. Mỗi nút PE đóng gói các khung VXLAN hoặc Ethernet bằng MPLS khi gửi các gói qua lõi MPLS và bằng tiêu đề đường hầm VXLAN khi gửi các gói qua mạng VXLAN.

2.4.2.3 Đường đi của gói tin EVPN-VXLAN

a) Xử lí lưu lượng BUM (Broadcast, Unknown Unicast, Multicast)

Lưu lượng BUM từ VXLAN đến EVPN, đi từ phân đoạn VXLAN 1 đến phân đoạn VXLAN 2 qua đám mây EVPN được xử lý như sau:



Hình 2. 17 Xử lý lưu lượng BUM trong VXLAN-EVPN

0 - Khi khởi động, Máy chủ A muốn gửi lưu lượng đến Máy chủ B. Do Máy chủ A không có liên kết ARP (ARP binding) cho Máy chủ B trong bảng ARP của mình, Máy chủ A xây dựng một yêu cầu ARP broadcast và gửi đi.

Nội dung của các gói tin ARP như sau:

VLAN ID = VLAN 10

MAC nguồn = MAC1

MAC đích = ff.ff.ff.ff.ff.ff

Địa chỉ IP nguồn = Địa chỉ IP của Máy chủ A hoặc địa chỉ IP của VM

Địa chỉ IP đích = Địa chỉ IP của Máy chủ B

Loại Ether của gói tin = 0x0806 Một khung Lớp 2 được gửi đến thiết bị chuyển mạch TOR A, đã được kích hoạt VXLAN.

1 - Khung yêu cầu ARP (broadcast) được nhận bởi thiết bị chuyển mạch TOR A. TOR A là nơi khởi tạo và kết thúc của VTEP VXLAN cho VNI 1000. VTEP cho VXLAN 1000 là một phần của miền broadcast cho VLAN 10 của Máy chủ A. Sau khi nhận khung, TOR A thực hiện xử lý đầu vào (ingress processing), bao gồm phân loại gói tin đầu vào. Dựa trên VLAN đến trong gói tin, TOR A phân loại gói tin vào một trong các IFL (Giao diện Logic Đầu vào) dưới một cổng nhất định. Họ của IFL này là

họ cầu nối (bridge family). Dựa trên họ cầu nối IFL, ID miền cầu nối (bridge domain ID) được xác định.

Sau khi miền cầu nối được xác định, TOR A học địa chỉ MAC nguồn của khung đến để MAC A có thể truy cập được thông qua IFL này. Vì khung này là khung broadcast, TOR A cần gửi khung đến tất cả các thành viên của miền broadcast (trừ thành viên mà khung đã được nhận). Một trong các thành viên của miền broadcast là VTEP cho VNI 1000. Để gửi khung trên phân đoạn VXLAN, TOR A hoàn tất xử lý next hop BUM của VXLAN trên khung. Next hop đẩy tiêu đề VXLAN vào.

Nội dung của tiêu đề VXLAN như sau:

Địa chỉ MAC nguồn = Địa chỉ MAC hoặc giao diện địa chỉ IP nguồn

Địa chỉ MAC đích = Địa chỉ MAC Multicast

Địa chỉ IP nguồn = 10.10.10.1

Địa chỉ IP đích = Địa chỉ nhóm Multicast (226.0.39.16)

Cổng UDP nguồn = Được tính toán dựa trên giá trị băm trên tiêu đề khung đến

Cổng UDP đích = 4789 (cổng phổ biến cho đường hầm VXLAN) Sau khi xây dựng khung đóng gói VXLAN, TOR A gửi khung đến Bộ định tuyến PE2.

2—Bộ định tuyến PE2 nhận khung VXLAN và xác định khung đó là khung VXLAN bằng cách xem cổng UDP đích nổi tiếng. VNI ID của khung VXLAN này được sử dụng để nhận dạng miền cầu nối. Sau khi bộ định tuyến PE2 xác định miền cầu nối, PE2 hoàn tất việc học MAC cho MAC nguồn bên trong với địa chỉ IP nguồn bên ngoài (ánh xạ MACA tới 10.10.10.1). Sau khi hoàn tất ánh xạ, quá trình xử lý next hop giải mã gói VXLAN sẽ loại bỏ tiêu đề VXLAN để kết thúc đường hầm VXLAN.

3A—Sau khi hoàn tất việc học MAC, MAC nguồn đã học được (MAC1 tới IP nguồn bên ngoài) được gửi đến L2ALD. Tuyến MAC này được L2ALD gửi đến RPD để mặt phẳng điều khiển học MAC này thông qua quảng bá tuyến MAC BGP đến các BGP peer. Sau khi các bộ định tuyến BGP peer nhận được quảng bá tuyến MAC, các bộ định tuyến sẽ cài đặt thông tin khả năng tiếp cận MAC này (MACA, MPLS LABEL L1) vào bảng miền cầu nối.

3—Miền cầu nối đã cho trỏ đến tuyến next hop multicast để chuyển tiếp gói tin qua đám mây EVPN. Next hop này đẩy nhãn dịch vụ (nhãn MPLS multicast được liên kết với VNI cho mỗi ID peer, miền cầu nối, nhãn là ID peer và VNI ID). Gói tin MPLS được hình thành và gửi qua đám mây MPLS.

4—Bộ định tuyến PE4 nhận khung dưới dạng gói tin MPLS. Tại đây, PE4 xác định miền cầu nối bằng cách tra cứu nhãn MPLS L1 trong bảng mpls.0. Tra cứu MPLS trỏ đến next hop bảng cho next hop miền cầu nối. Sau khi miền cầu nối được xác định và gói tin được xác định là gói tin broadcast, next hop flood tổng hợp BUM được thực thi. Next hop tổng hợp BUM cũng trỏ đến next hop VXLAN (được sử dụng để xây dựng gói tin multicast VXLAN).

5 - Next hop VXLAN chứa thông tin để xây dựng tiêu đề VXLAN.

Thông tin tiêu đề VXLAN như sau:

Địa chỉ MAC nguồn = Địa chỉ MAC hoặc giao diện địa chỉ IP nguồn

Địa chỉ MAC đích = Địa chỉ MAC Multicast

Địa chỉ IP nguồn = 11.10.10.1

Địa chỉ IP đích = Địa chỉ nhóm Multicast (226.0.39.16)

Cổng UDP nguồn = Được tính toán dựa trên giá trị băm (hash) trên tiêu đề khung đến

Cổng UDP đích = 4789 (cổng phổ biến cho đường hầm VXLAN)

6 - Việc xử lý khung cho bước này giống như Bước 1. Sau khi tiêu đề VXLAN bị loại bỏ, khung được chuyển tiếp đến tuyến flood CE được liên kết với miền broadcast, và gói tin được chuyển tiếp dưới dạng khung Lớp 2.

7 - Máy chủ B nhận gói tin yêu cầu ARP và gửi phản hồi ARP đến Máy chủ A.

b) Xử lí lưu lượng Unicast Traffic

Giả sử rằng cả việc học MAC ở mặt phẳng dữ liệu và mặt phẳng điều khiển đã diễn ra, lưu lượng unicast VXLAN sang EVPN (phản hồi ARP) từ Máy chủ B được xử lý như sau:

8 – Máy chủ B tạo một phản hồi ARP. Nội dung của các gói ARP như sau:

- ID VLAN = VLAN 10
- Địa chỉ MAC nguồn = MACB (Địa chỉ MAC giao diện Máy chủ B)
- Địa chỉ MAC đích = MACA
- Địa chỉ IP nguồn = Địa chỉ IP của Máy chủ B hoặc địa chỉ IP của Máy ảo
- Địa chỉ IP đích = Địa chỉ IP của Máy chủ A

Gói ARP được chuyển tiếp đến TOR B.

9 - Sau khi nhận được khung, switch TOR B phân loại khung đến. Khung được phân loại vào một IFL (Incoming Frame List - Danh sách khung đến) trên giao diện nhận. Dựa trên họ IFL, miền bridge (bridge domain) liên kết với IFL được xác định. Trên miền bridge đã cho, TOR B học địa chỉ MAC nguồn. Sau khi TOR B hoàn tất việc tra cứu địa chỉ MAC đích của miền bridge (MACA), việc tra cứu này sẽ cung cấp bước nhảy tiếp theo (next hop) unicast VXLAN. Bước nhảy tiếp theo này chứa tất cả thông tin cần thiết để hình thành tiêu đề VXLAN.

Nội dung của bước nhảy tiếp theo được yêu cầu để chuyển tiếp gói tin như sau:

- Địa chỉ MAC nguồn = Địa chỉ MAC của địa chỉ IP nguồn
- Địa chỉ MAC đích = Địa chỉ MAC của bước nhảy tiếp theo
- Địa chỉ IP nguồn = 11.10.10.2
- Địa chỉ IP đích = 11.10.10.1 (là kết quả của quá trình học MAC)
- Cổng UDP nguồn = Được tính toán dựa trên hàm băm (hash) trên tiêu đề khung đến
- Cổng UDP đích = 4789 (cổng phổ biến cho đường hầm VXLAN)

10 - Router PE nhận khung được đóng gói VXLAN4. PE4 xác định khung bằng cách hoàn tất việc tra cứu bằng cách sử dụng địa chỉ IP đích và cổng UDP đích. Việc tra cứu này dẫn đến việc giải đóng gói VXLAN. Bước nhảy tiếp theo của việc giải đóng gói cũng lưu trữ địa chỉ IP nguồn bên ngoài.

Bước nhảy tiếp theo được thực hiện dựa trên ID VNI 1000. Việc tra cứu này dẫn đến miền bride.

10A - Router PE hoàn tất việc ánh xạ MAC nguồn tới địa chỉ IP nguồn và L2ALD nhận được thông báo học MAC. MAC này được gửi đến RPD để phân phối cho các router PE khác thông qua tuyến quảng bá BGP-EVPN MAC. Mặt phẳng điều khiển BGP phân phối khả năng tiếp cận MAC này đến tất cả các router PE khác. Việc tra cứu MAC đích (MAC1) được thực hiện trong bảng địa chỉ MAC của miền bridge. Việc tra cứu này dẫn đến một bước nhảy tiếp theo unicast (EVPN NH).

11 - Bước nhảy tiếp theo unicast EVPN được thực thi. Bước nhảy tiếp theo này chứa một nhãn dịch vụ MPLS unicast. Nhãn này được phân phối thông qua mặt phẳng điều khiển MP-BGP. Peer phân bổ nhãn MPLS này. Việc phân bổ nhãn này có thể dựa trên PE (PE, VLAN) hoặc trên cơ sở địa chỉ MAC. Dựa trên thông tin trong bước nhảy tiếp theo, gói MPLS được hình thành và chuyển tiếp trên mạng MPLS.

12 - Router PE2 nhận được khung. Khung được xác định là một gói MPLS. Việc tra cứu nhãn MPLS được thực hiện trong bảng MPLS.0. Việc tra cứu này dẫn đến bước nhảy tiếp theo và miền bridge.

Việc tra cứu MAC đích (MAC1) được thực hiện trong bảng MAC của miền bridge. Việc tra cứu này dẫn đến một bước nhảy tiếp theo unicast VXLAN.

13 - Bước nhảy tiếp theo unicast VXLAN chứa tất cả thông tin để xây dựng tiêu đề gói tin được đóng gói VXLAN. Tiêu đề VXLAN được áp đặt lên gói tin.

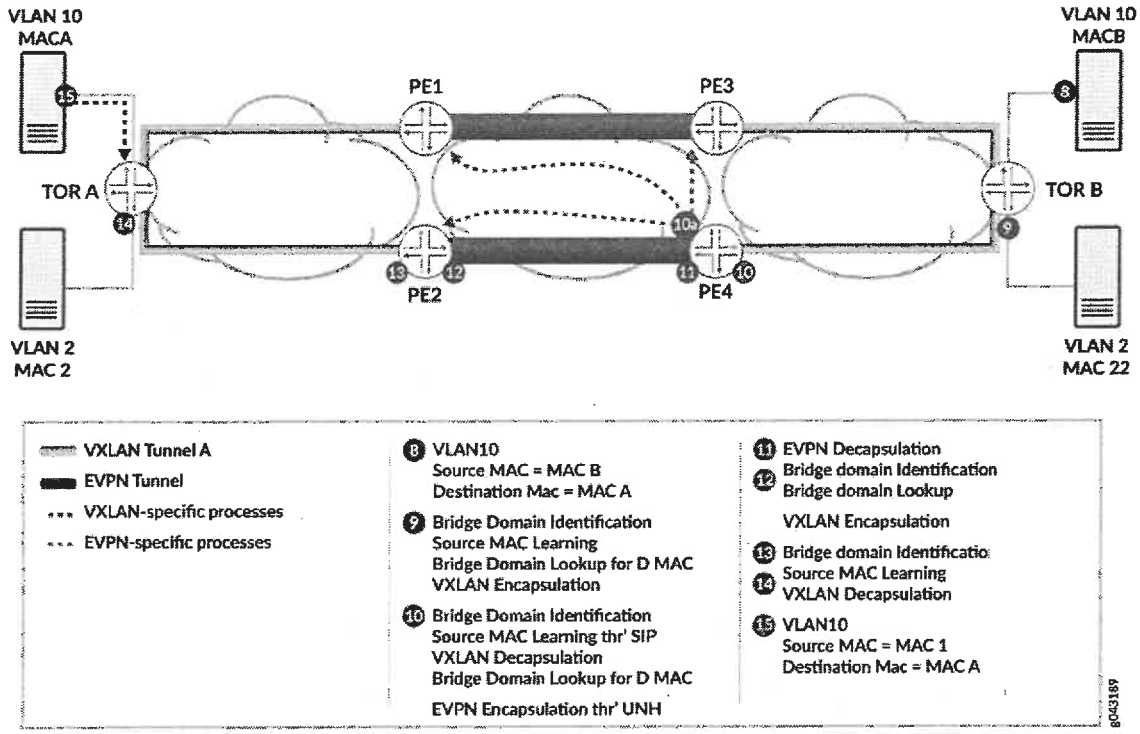
Nội dung của tiêu đề bước nhảy tiếp theo đóng gói VXLAN như sau:

- Địa chỉ MAC nguồn = Địa chỉ MAC của giao diện địa chỉ IP nguồn
- Địa chỉ MAC đích = Địa chỉ MAC của bước nhảy tiếp theo
- Địa chỉ IP nguồn = 10.10.10.2
- Địa chỉ IP đích = 10.10.10.1 (là kết quả của quá trình học MAC)
- Cổng UDP nguồn = Được tính toán dựa trên hàm băm (hash) trên tiêu đề khung đến
- Cổng UDP đích = 4789 (cổng phổ biến cho đường hầm VXLAN)

14- Khung được đóng gói VXLAN được nhận bởi switch TOR A. TOR A xác định khung bằng cách thực hiện tra cứu bằng địa chỉ IP đích và cổng UDP đích. Việc tra cứu này dẫn đến việc giải đóng gói VXLAN. Bước nhảy tiếp theo được giải đóng gói cũng lưu trữ địa chỉ IP nguồn bên ngoài.

Việc tra cứu tiếp theo được thực hiện dựa trên VNI ID 1000. Việc tra cứu này dẫn đến miền bridge. TOR A hoàn tất việc học địa chỉ MAC nguồn (MAC2) sang địa chỉ IP nguồn (10.10.10.2). TOR A tra cứu địa chỉ MAC đích (MAC1) trong bảng địa chỉ MAC của miền bridge. Việc tra cứu này dẫn đến một bước nhảy tiếp theo unicast có thông tin về giao diện egress.

15 - Máy chủ A nhận được phản hồi ARP, và Máy chủ A và Máy chủ B đã sẵn sàng để giao tiếp.



Hình 2. 18 Xử lí lưu lượng unicast

2.5. Kết luận chương 2

Chương này đã trình bày một cái nhìn tổng quan về trung tâm dữ liệu, từ những khái niệm cơ bản, quá trình phát triển đến các hạng mục chính cấu thành nên một trung tâm dữ liệu hiện đại. Bên cạnh đó cũng đã tìm hiểu về một số giao thức mạng phổ biến được sử dụng, như Spanning-tree (STP) và Virtual Port Channel (vPC), cũng như sự phát triển trong thiết kế mạng lưới từ mô hình ba lớp truyền thống đến kiến trúc Clos tiên tiến hơn.

Một điểm nhấn quan trọng của chương này là việc khám phá ứng dụng của công nghệ VXLAN trong các trung tâm dữ liệu. Chương này đã đưa đến các thông tin về ảo hóa mạng, khái niệm overlay và underlay và đặc biệt là công nghệ EVPN-VXLAN như một giải pháp hiệu quả để xây dựng các mạng lớp 2 mở rộng, linh hoạt và hiệu suất cao trong môi trường trung tâm dữ liệu.

Qua các nội dung được trình bày, có thể thấy rằng công nghệ trung tâm dữ liệu đang không ngừng phát triển để đáp ứng nhu cầu ngày càng tăng về lưu trữ, xử lý và kết nối. Việc hiểu rõ các khái niệm, giao thức và kiến trúc nền tảng, cũng như nắm bắt các xu hướng công nghệ mới như VXLAN và EVPN là vô cùng quan trọng để có thể thiết kế, triển khai và vận hành hiệu quả các trung tâm dữ liệu trong tương lai. Sự kết hợp giữa các kiến trúc mạng tiên tiến và các công nghệ ảo hóa mạnh mẽ hứa hẹn sẽ tiếp tục định hình và tối ưu hóa cách thức hoạt động của các trung tâm dữ liệu, đóng vai trò là nền tảng vững chắc cho kỷ nguyên số.

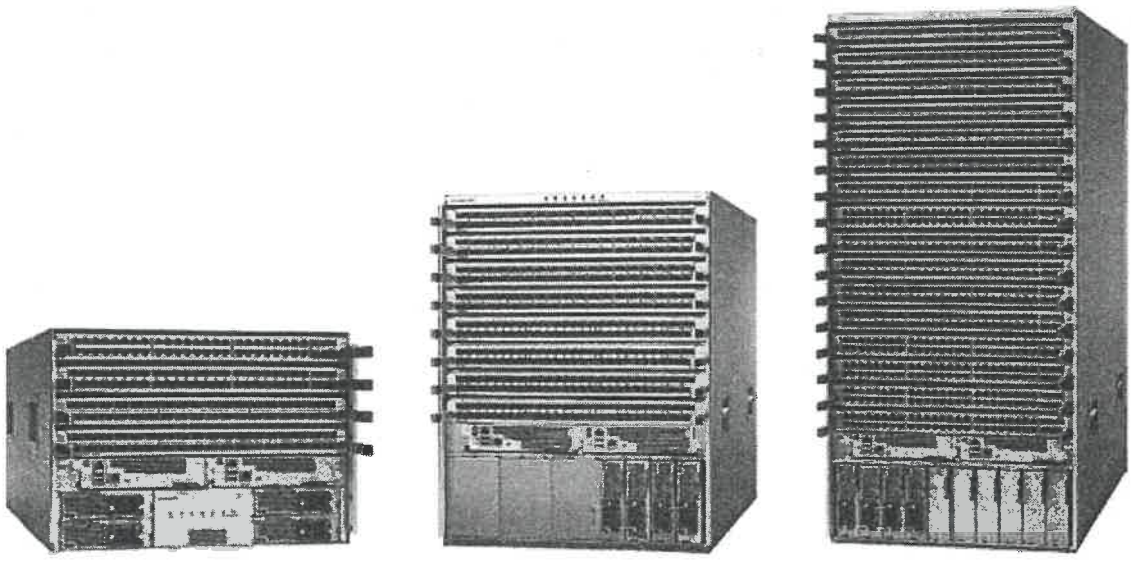
CHƯƠNG 3: XÂY DỰNG MÔ PHỎNG VXLAN TRONG MẠNG TRUNG TÂM DỮ LIỆU VỚI CISCO NEXUS DASHBOARD FABRIC

3.1. Tổng quan về dòng switch Cisco Nexus và Cisco Nexus Dashboard

3.1.1 Switch Cisco Nexus

Dòng sản phẩm Cisco Nexus được Cisco Systems thiết kế chuyên biệt để đáp ứng những yêu cầu khắt khe của các trung tâm dữ liệu hiện đại. Ra mắt lần đầu vào năm 2008, Cisco Nexus nhanh chóng khẳng định vị thế dẫn đầu nhờ vào sự kết hợp giữa hiệu suất vượt trội, khả năng mở rộng linh hoạt và các tính năng tiên tiến, phục vụ cho một loạt các nhu cầu từ doanh nghiệp nhỏ đến các nhà cung cấp dịch vụ đám mây quy mô lớn.

Các thiết bị chuyển mạch Cisco Nexus bao gồm cả dòng sản phẩm dạng mô-đun (modular) và cổng cố định (fixed-port), mang đến sự lựa chọn đa dạng để phù hợp với mọi quy mô và kiến trúc mạng. Điểm nổi bật của dòng Nexus là khả năng hỗ trợ các công nghệ mạng tiên tiến như VXLAN (Virtual Extensible LAN), EVPN (Ethernet VPN), và ACI (Application Centric Infrastructure), cho phép xây dựng các hạ tầng mạng ảo hóa, tự động hóa và có khả năng mở rộng cao.



Hình 3. 1 Một số sản phẩm dòng Cisco Nexus

Với hệ điều hành NX-OS mạnh mẽ, Cisco Nexus cung cấp một nền tảng vững chắc cho việc triển khai các dịch vụ mạng phức tạp, đồng thời đảm bảo tính sẵn sàng cao, độ trễ thấp và khả năng quản lý toàn diện. Dòng sản phẩm này không chỉ tối ưu hóa cho lưu lượng truyền tải lớn trong trung tâm dữ liệu mà còn hỗ trợ các kết nối tốc độ cao, từ 10 Gigabit Ethernet đến 800 Gigabit Ethernet.

3.1.2 Cisco Nexus Dashboard

Trong bối cảnh các trung tâm dữ liệu ngày càng trở nên phức tạp, phân tán và đa dạng, việc quản lý và vận hành hiệu quả hạ tầng mạng là một thách thức lớn đối với các tổ chức. Để giải quyết vấn đề này, Cisco đã cho ra đời Cisco Nexus Dashboard – một nền tảng quản lý và vận hành mạng trung tâm dữ liệu tập trung, mang tính cách mạng.

3.1.2.1 Cisco Nexus Dashboard là gì?

Cisco Nexus Dashboard là một bảng điều khiển (dashboard) hợp nhất, cung cấp một điểm truy cập duy nhất để giám sát, quản lý và tự động hóa các hoạt động mạng trên nhiều trung tâm dữ liệu và môi trường đám mây lai. Nó được thiết kế để đơn giản hóa các quy trình vận hành phức tạp, giảm thiểu lỗi do con người và tăng cường khả năng phục hồi của doanh nghiệp. Nền tảng này hoạt động như một trung tâm điều khiển, tích hợp các dịch vụ và ứng dụng quản lý mạng khác nhau của Cisco và các đối tác thứ ba.

Hình 3. 2 Giao diện Cisco Nexus Dashboard

3.1.2.2 Mục đích chính của Cisco Nexus Dashboard

Mục tiêu cốt lõi của Nexus Dashboard là cung cấp một trải nghiệm vận hành nhất quán và đơn giản hóa trên toàn bộ vòng đời của mạng trung tâm dữ liệu, từ khâu triển

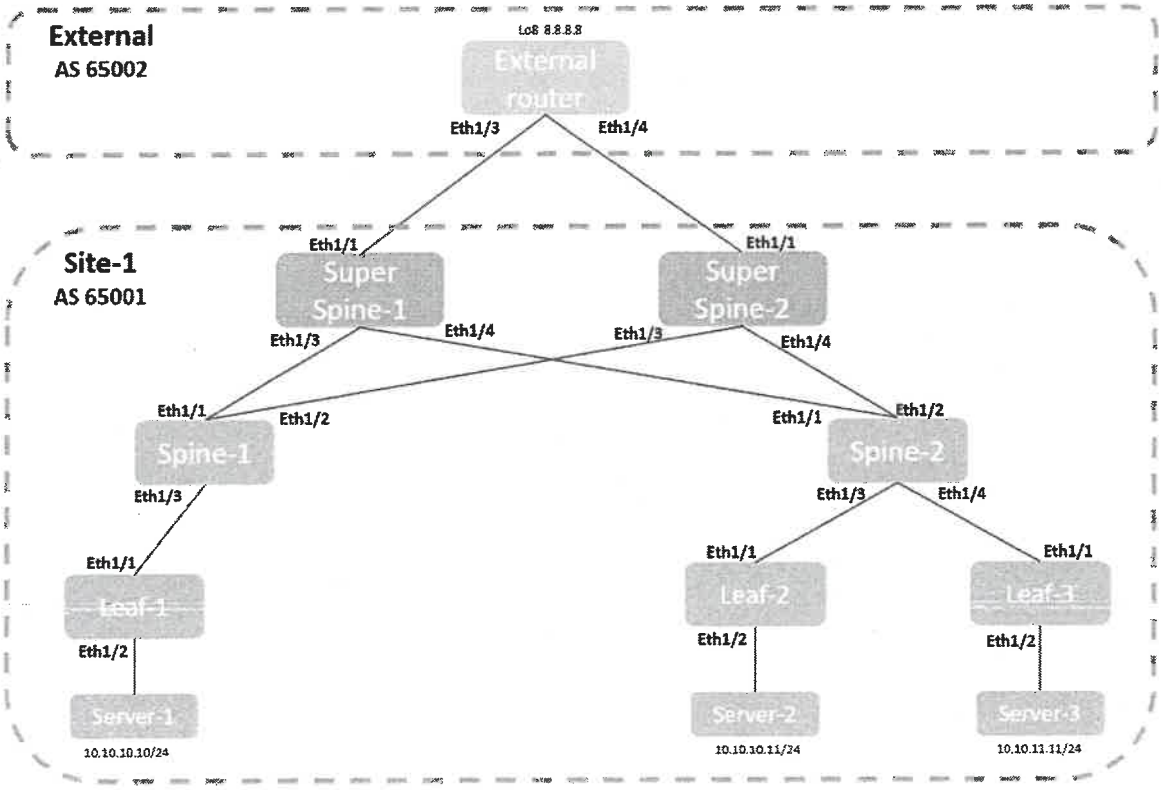
khai, giám sát, khắc phục sự cố cho đến tối ưu hóa. Nó giúp các đội ngũ vận hành mạng có được cái nhìn toàn diện và kiểm soát tập trung đối với hạ tầng mạng, bất kể quy mô hay vị trí địa lý.

3.1.2.3 Các tính năng và lợi ích chính

- *Quản lý tập trung và nhất quán*: Cung cấp một giao diện người dùng duy nhất để xem và quản lý cấu hình mạng, chính sách và trạng thái hoạt động trên nhiều fabric (VXLAN EVPN, ACI), nhiều địa điểm trung tâm dữ liệu và cả môi trường đám mây. Điều này giúp loại bỏ sự cần thiết phải làm việc với nhiều công cụ quản lý rời rạc.
- *Tự động hóa vòng đời mạng:* Hỗ trợ tự động hóa các tác vụ từ giai đoạn thiết kế, triển khai, thay đổi cấu hình cho đến nâng cấp phần mềm, giúp giảm thiểu thời gian triển khai và rủi ro vận hành.
- *Khả năng hiển thị và phân tích nâng cao:* Tích hợp các công cụ phân tích và giám sát mạnh mẽ (như Nexus Dashboard Insights) để cung cấp thông tin chi tiết về hiệu suất mạng, phát hiện sớm các vấn đề tiềm ẩn và đưa ra các khuyến nghị chủ động để khắc phục sự cố và tối ưu hóa.
- *Nền tảng mở và có khả năng mở rộng:* Được xây dựng trên kiến trúc microservices, cho phép dễ dàng tích hợp các ứng dụng và dịch vụ của Cisco cũng như của bên thứ ba, đồng thời có khả năng mở rộng linh hoạt để đáp ứng nhu cầu phát triển của doanh nghiệp.
- *Đơn giản hóa vận hành đa đám mây:* Mở rộng khả năng quản lý và vận hành nhất quán ra các môi trường đám mây công cộng, hỗ trợ các chiến lược đám mây lai.
- *Tăng cường tính sẵn sàng và khả năng phục hồi:* Giúp nhanh chóng xác định và giải quyết các sự cố mạng, giảm thiểu thời gian ngừng hoạt động và đảm bảo tính liên tục của hoạt động kinh doanh.

3.2. Triển khai mô phỏng VXLAN trong mạng trung tâm dữ liệu với Cisco Nexus Dashboard

3.2.1 Mô hình mô phỏng



Mô hình mô phỏng triển khai trong đề án áp dụng kiến trúc Clos ba lớp (3-tier Clos), bao gồm các lớp Leaf, Spine và Super Spine. Đây là kiến trúc phổ biến trong các trung tâm dữ liệu hiện đại, đặc biệt phù hợp với triển khai VXLAN BGP EVPN, nhờ vào tính khả mở, tính sẵn sàng cao và khả năng hỗ trợ lưu lượng east-west hiệu quả.

Cấu trúc chi tiết của mô hình:

- **Lớp Leaf**:
  - Bao gồm các thiết bị Leaf-1, Leaf-2, Leaf-3, mỗi thiết bị kết nối với một server thực tế (Server-1, Server-2, Server-3).
  - Đây là các VTEP (VXLAN Tunnel Endpoint) – nơi thực hiện chức năng đóng gói và giải đóng gói gói tin VXLAN.
- **Lớp Spine**:
  - Bao gồm hai thiết bị Spine-1 và Spine-2, có nhiệm vụ chuyển tiếp lưu lượng IP trong mạng underlay.

- Chúng hoạt động như Route Reflector hoặc iBGP Speaker, đảm bảo việc phân phối thông tin qua MP-BGP EVPN.

- **Lớp Super Spine**:
  - Super Spine-1 và Super Spine-2 đảm nhận vai trò Border Gateway Super Spine, giúp mở rộng cấu trúc mạng và đóng vai trò kết nối ra bên ngoài (external).
  - Chúng kết nối đến router bên ngoài (External Router – AS65002) thông qua giao thức eBGP.

- **Router ngoài (External)**:
  - Thiết bị External Router sử dụng Loopback interface (8.8.8.8) làm đại diện cho tài nguyên bên ngoài.
  - Có 2 kết nối tới Super Spine-1 và Super Spine-2.

3.2.2 Lựa chọn công nghệ và kịch bản mô phỏng

Trong phạm vi mô hình mô phỏng của đề án, việc lựa chọn công nghệ mạng phù hợp đóng vai trò quan trọng nhằm đảm bảo khả năng mở rộng, linh hoạt và vận hành hiệu quả của hệ thống. Thay vì sử dụng mô hình VLAN truyền thống, công nghệ VXLAN (Virtual Extensible LAN) kết hợp với EVPN (Ethernet VPN) đã được áp dụng trong kiến trúc mạng Leaf-Spine của mô hình. Giải pháp này không chỉ phản ánh xu hướng triển khai trong các trung tâm dữ liệu hiện đại, mà còn giúp nâng cao hiệu năng và khả năng quản lý mạng trong quá trình mô phỏng. Dưới đây là những lý do cốt lõi giải thích tại sao VXLAN BGP EVPN là sự lựa chọn hợp lý trong mô hình triển khai của đề án.

Mở rộng mạng ảo vượt trội

VLAN truyền thống chỉ hỗ trợ tối đa 4096 mạng ảo do giới hạn 12-bit VLAN ID, trong khi đó VXLAN sử dụng 24-bit VNI cho phép định danh tới 16.7 triệu mạng ảo. Mỗi VNI tương ứng với một không gian lớp 2 độc lập, đảm bảo khả năng phân vùng mạnh mẽ và dễ dàng mở rộng cho các môi trường multi-tenant phức tạp hoặc trung tâm dữ liệu lớn [3].

Khả năng mở rộng trung tâm dữ liệu và liên kết liên site

VXLAN cho phép mở rộng mạng giữa nhiều trung tâm dữ liệu khác nhau mà vẫn giữ được topology logic nhất quán. Nhờ lớp overlay đồng nhất, các mạng có thể mở rộng xuyên suốt giữa các site, phục vụ tốt cho các nhu cầu như mở rộng tài nguyên, sao

lưu, hoặc khôi phục thảm họa. Hơn nữa, với hỗ trợ hardware VTEP, độ trễ và độ phức tạp của quá trình đóng gói VXLAN được giảm thiểu đáng kể [3].

Giao tiếp liên subnet và di chuyển máy ảo linh hoạt

Không giống như Ethernet truyền thống cần router lớp 3 để giao tiếp liên mạng con, VXLAN hoạt động trên nền IP cho phép di chuyển máy ảo giữa các subnet mà không cần thay đổi địa chỉ IP. Điều này rất hữu ích trong môi trường trung tâm dữ liệu nơi yêu cầu về cân bằng tải, bảo trì hoặc mở rộng tài nguyên diễn ra thường xuyên [3].

Tương thích cao, dễ triển khai

VXLAN có thể hoạt động trên cơ sở hạ tầng mạng hiện có mà không yêu cầu thay đổi thiết bị lớn. Nó tương thích với các giao thức và thiết bị hiện hành như BGP, switch layer 3, và router, giúp quá trình triển khai và quản lý trở nên đơn giản hơn [3].

Độ tin cậy và hiệu suất cao

Với việc sử dụng MP-BGP EVPN làm control plane, VXLAN hỗ trợ định tuyến động, cân bằng tải và chọn đường đi tối ưu theo thời gian thực. Điều này đảm bảo hiệu suất mạng cao và khả năng khôi phục nhanh khi xảy ra lỗi liên kết [3].

Bảo mật và phân tách lưu lượng

Với khả năng phân chia mạng dựa trên VNI và VRF, VXLAN đảm bảo cách ly lưu lượng giữa các tenant, giúp bảo mật và quyền riêng tư được tăng cường. Đây là một yếu tố quan trọng trong môi trường multi-tenant hoặc hệ thống cung cấp dịch vụ [3].

Tích hợp NDFC tăng hiệu quả cấu hình và vận hành

Việc sử dụng Cisco Nexus Dashboard Fabric Controller (NDFC) trong mô phỏng càng làm nổi bật lợi thế của VXLAN. NDFC đơn giản hóa cấu hình VXLAN EVPN thông qua giao diện tập trung và khả năng tự động hóa mạnh mẽ. Người quản trị có thể triển khai mạng nhanh chóng, giám sát trạng thái cấu hình trực quan, và dễ dàng khắc phục sự cố nhờ vào các công cụ phân tích tích hợp.

Sự lựa chọn VXLAN thay vì VLAN không chỉ là một quyết định kỹ thuật mà còn phản ánh nhu cầu thực tế của các trung tâm dữ liệu hiện đại. Với khả năng mở rộng, bảo mật, tự động hóa và hiệu năng cao, VXLAN BGP EVPN là giải pháp phù hợp cho cả mô hình mô phỏng và triển khai thực tế.

Với mô phỏng này, sẽ sử dụng Nexus Dashboard Fabric Controller (NDFC) để thực hiện các thao tác sau:

- Tạo và cung cấp một fabric VXLAN EVPN 3 lớp Greenfield mới với các switch Super Spine.

- Tạo và cấu hình VRF và mạng trong một fabric VXLAN EVPN.
- Tạo và cung cấp fabric bên ngoài mới.
- Cấu hình kết nối liên Fabric giữa fabric VXLAN EVPN và fabric bên ngoài.

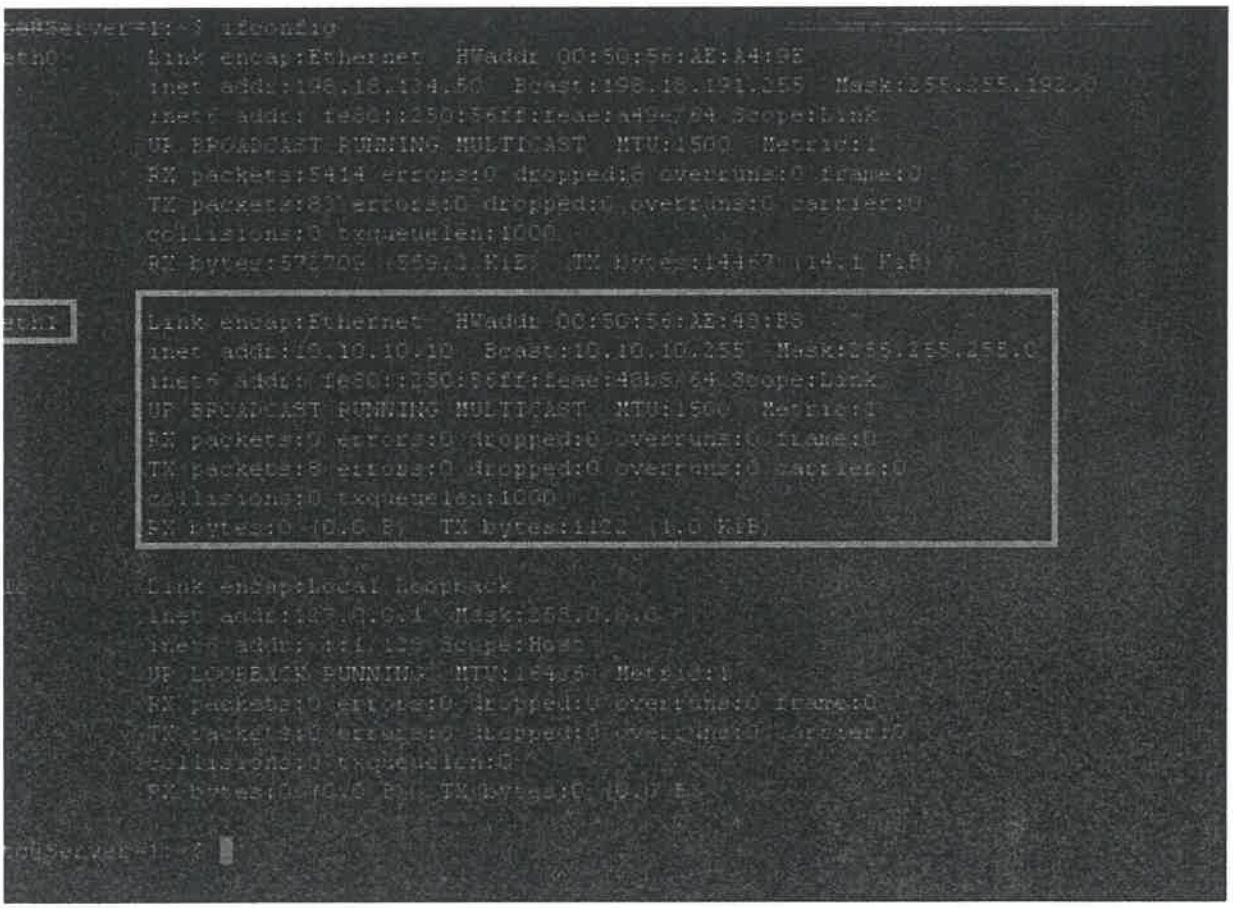
3.2.3 Các bước thực hiện

3.2.3.1 Xác minh cấu hình IP và kết nối của server

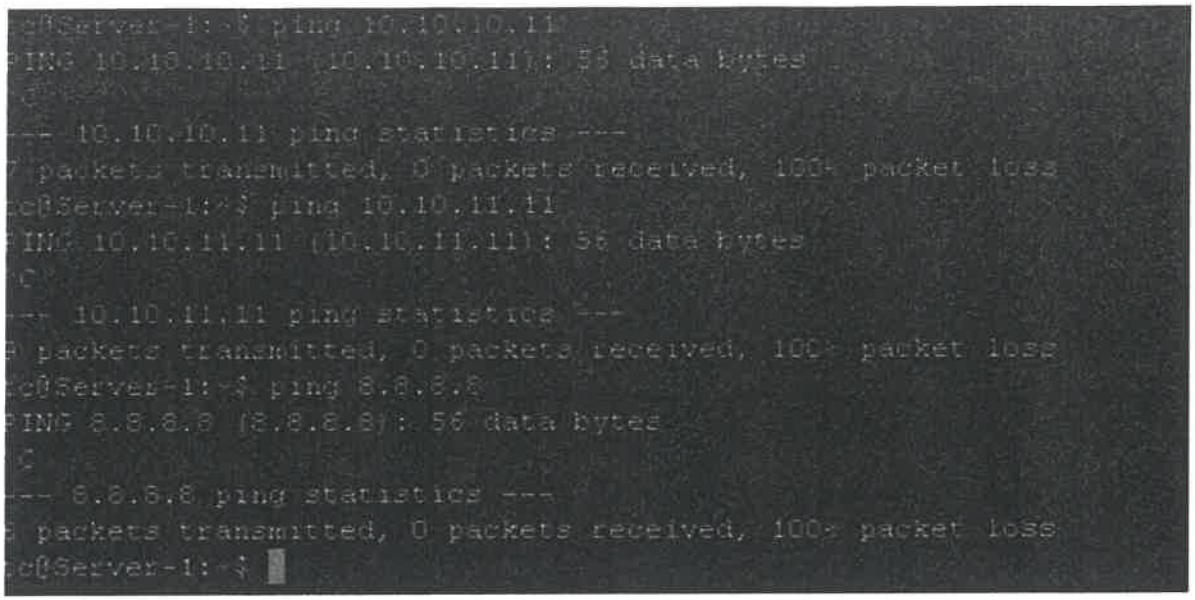
Mô phỏng này bao gồm 3 server được kết nối với fabric thông qua các leaf khác nhau. Đầu tiên, thực hiện cấu hình IP cho các server. Sau đó kiểm tra lại địa chỉ IP của giao diện Eth1 được nối với fabric.

Hiện tại, fabric chưa được cấu hình do đó các server chưa thể giao tiếp với nhau.

Để xác minh điều này, thực hiện ping 10.10.10.11 (server 2), 10.10.11.11 (server 3) và 8.8.8.8 (đại diện cho tài nguyên bên ngoài). Kết quả là, lệnh ping không thành công và xảy ra tình trạng mất gói tin 100%. Mục tiêu của mô phỏng này là đảm bảo kết nối hoạt động giữa máy chủ và tài nguyên bên ngoài.



Hình 3. 3 Kiểm tra địa chỉ IP của server



Hình 3. 4 Tình trạng mất gói tin 100% do fabric chưa được cấu hình

3.2.3.2 Tạo và cấu hình VXLAN EVPN Fabric trên Greenfield

Tạo một fabric rỗng và cấu hình VXLAN BGP EVPN trong site-1 bằng cách sử dụng Nexus Dashboard Fabric Controller (NDFC). NDFC sẽ xóa tất cả các cấu hình trên các switch ngoại trừ cấu hình quản lý trước khi đẩy các cấu hình VXLAN EVPN xuống các switch.

a) Tạo fabric

Thực hiện tạo Fabric mới, trong trường *Fabric Name*, nhập tên fabric. Với *Templete,* chọn *Easy_Fabric*, nó sẽ hỗ trợ triển khai VXLAN EVPN với các switch Nexus 9000. Trong *General Parameters,* trong trường *BGP ASN,* nhập **65001** để cấu hình AS 65001 cho phần site-1.

b) Thêm switch và gán vai trò

Chọn các hộp kiểm cho tất cả các thiết bị và gán vai trò (set rules) cho các loại thiết bị tương ứng.

Để xem trước cấu hình, trong cửa sổ *Deploy Configuration*, nhấp vào liên kết trong cột *Pending Config*. Có thể xem trước *Pending Config* cho mỗi switch, cũng như so sánh *Side-to-Side Comparison* từ góc độ thay đổi so với cấu hình hiện tại.

Đóng bản *Pending Config* đang chờ xử lý và sau đó nhấp vào *Deploy*. Kết quả sẽ thấy tiến trình chạy của các lệnh. Khi quá trình triển khai hoàn tất, trạng thái là **SUCCESS**.

c) Cấu hình giao diện truy cập tới các server

Thực hiện gán *policy int_access_host* cho các cổng Ethernet1/2 trên tất cả các thiết bị Leaf để định nghĩa kiểu kết nối truy cập. Trong đó, cấu hình Access VLAN 230 được áp dụng cho Leaf-1 và Leaf-2, còn Leaf-3 được gán với VLAN 231, tương ứng với các mạng nội bộ đã tạo. Sau khi hoàn tất cấu hình cho từng giao diện, tiến hành triển khai bằng tính năng Deploy Config để áp dụng cấu hình lên thiết bị.

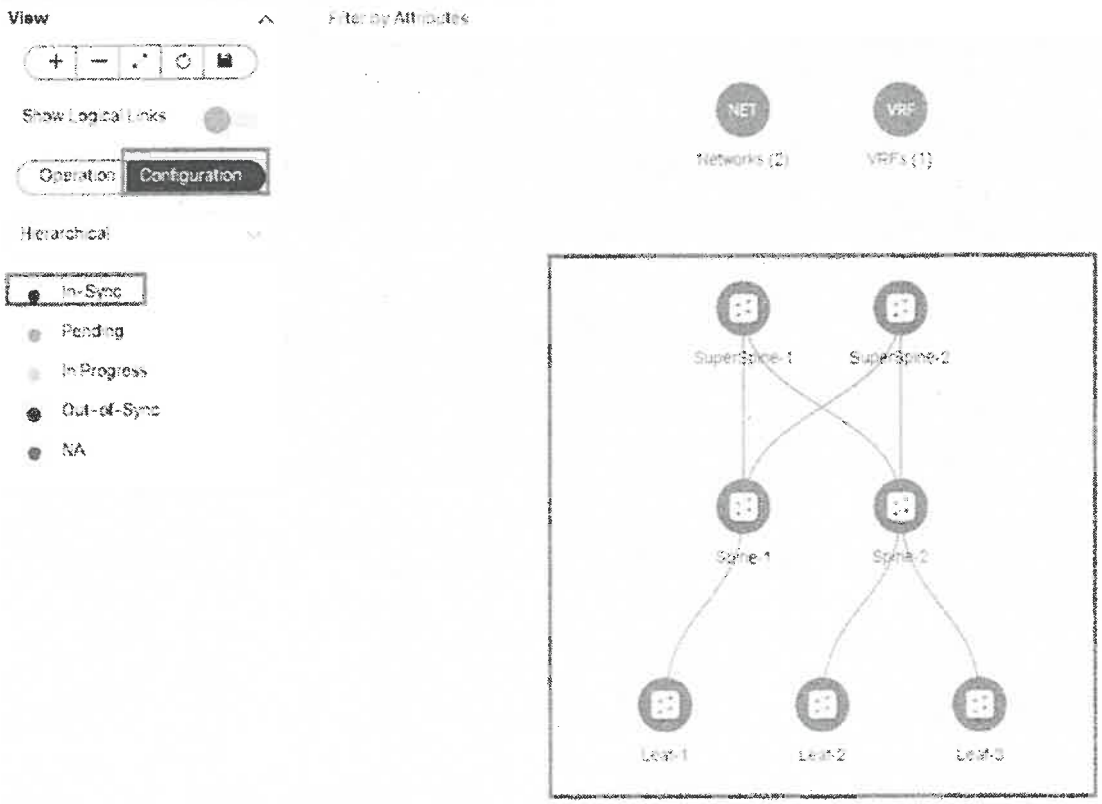
d) Tạo VRF và mạng

Tiến hành tạo một VRF mới có tên dcloud với VLAN ID 2000 để làm không gian định tuyến riêng cho hệ thống. Trong VRF này, tạo hai mạng *dcloud_net_10* và *dcloud_net_11*, tương ứng với các subnet 10.10.10.0/24 và 10.10.11.0/24, sử dụng VLAN ID lần lượt là 2300 và 2301.

Các mạng này sau đó được đính kèm vào các thiết bị Leaf tương ứng: dcloud_net_10 gắn với Leaf-1 và Leaf-2 trên VLAN 230, còn dcloud_net_11 gắn với Leaf-3 trên VLAN 231. Cuối cùng, sử dụng chức năng Deploy Config để triển khai toàn bộ cấu hình lên hệ thống hạ tầng mạng.

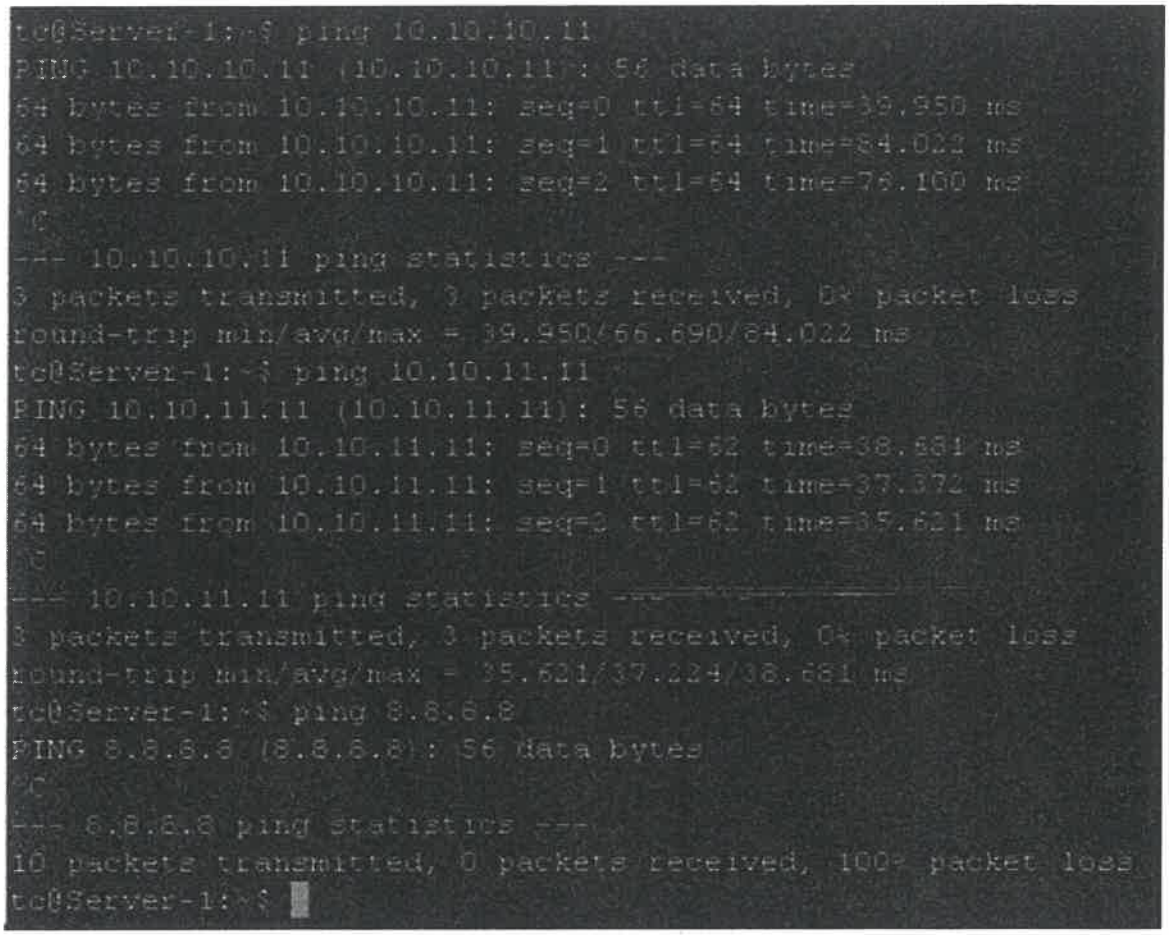
e) Xác minh kết nối topology fabric

Quay lại trang NDFC chính và xem cấu hình topology được ghi nhận.



Hình 3. 5 Topology

Để xác minh kết nối, lúc này quay lại MTPuTTY cho server-1 và ping 10.10.10.11 (server-2), 10.10.11.11 (server-3) và 8.8.8.8 (đại diện cho một tài nguyên bên ngoài) từ server-1. Kết quả là, kết nối giữa các server hiện đang hoạt động và kết nối đến tài nguyên bên ngoài, 8.8.8.8, vẫn thất bại.



Hình 3. 6 Xác minh kết nối giữa các server và kết nối tới tài nguyên bên ngoài

3.2.3.3 Tạo và Cấu hình External Fabric

External fabric đại diện cho kết nối từ trung tâm dữ liệu ra thế giới bên ngoài (WAN hoặc Internet).

a) Tạo External Fabric

Đặt các trường theo các giá trị trong bảng dưới

Trường hoặc Điều khiển	Giá trị
Fabric Name	External
Pick Template	External_Fabric
General Parameters	

BGP AS #	65002
Fabric Monitor Mode	unchecked

Bảng 3. 1 Các thông tin để tạo external fabric

b) Thêm thiết bị và gán vai trò

Truy cập *Fabric External* và thực hiện thêm thiết bị *External Router* thông qua tính năng *Discover*. Sau khi thiết bị được phát hiện, tiến hành thêm vào fabric và gán vai trò là edge router. Tiếp theo, sử dụng tùy chọn *Recalculate Config* để tính toán lại cấu hình, kiểm tra trước khi triển khai, và cuối cùng thực hiện Deploy để áp dụng thay đổi. Khi trạng thái chuyển sang **SUCCESS**, việc cấu hình đã hoàn tất.

3.2.3.4 Cấu hình kết nối giữa các fabric (IFC)

a) Cấu hình Liên kết giữa các Fabric

Đặt các trường theo các giá trị trong bảng dưới:

Trường	Thông tin của link thứ nhất	Thông tin của link thứ hai
Link Type	Inter-Fabric	Inter-Fabric
Link Sub-Type	VRF-LITE	VRF-LITE
Source device	SuperSpine-1	SuperSpine-2
Source interface	Ethernet1/1	Ethernet1/1
Destination Fabric	External	External
Destination device	External-Router	External-Router
Destination interface	Ethernet1/3	Ethernet1/4
Source BGP ASN	65001	65001
Source IP Address/Mask	192.168.1.1/30	192.168.1.5/30
Destination IP	192.168.1.2	192.168.1.6
Destination BGP ASN	65002	65002
Link MTU	9216	9216

Bảng 3. 2 Thông tin cấu hình liên kết

b) Cấu hình đính kèm VRF

Trong Fabric, tiến hành gắn *VRF dcloud* vào hai thiết bị SuperSpine-1 và SuperSpine-2. Kích hoạt chế độ Attach, giữ nguyên VLAN ID là 2000, sau đó cấu hình VRF Lite Extension với DOT1Q_ID lần lượt là 20 và 21 để phân biệt. Sau cùng, triển khai cấu hình bằng *Deploy Config*, và xác nhận khi trạng thái hiển thị **SUCCESS**.

c) Cấu hình External router

Đặt các trường theo các giá trị trong bảng dưới:

Trường (Field)	**Giá trị (Value)**	
Layer-3 interface	Ethernet1/3	Ethernet1/4
Encapsulation dot1q VLAN ID	20	21
VRF Name	dcloud	dcloud
Subinterface IPv4 Address/ Netmask	192.168.1.2/30	192.168.1.6/30
Subinterface MTU	9216	9216
Neighbor IPv4 Address	192.168.1.1	192.168.1.5
Neighbor ASN	65001	65001

Bảng 3. 3 Thông số cấu hình các kết nối của external router

Bây giờ sẽ thực hiện cấu hình giao diện loopback. Giao diện loopback này đại diện cho tài nguyên bên ngoài với địa chỉ IP 8.8.8.8.

Trong cửa sổ *External Fabric Overview*, nhấp vào *Interfaces* và sau đó, nhấp vào *Actions* > *Create interface*. Trong giao diện Create interface, nhập các giá trị, nhấp vào *Save* và nhấp vào *Deploy*.

Field or Parameter	**Value**
Type	Loopback
Select a device	External-Router
Loopback ID	8

Field or Parameter	Value
Interface VRF	dcloud
Loopback IP	8.8.8.8
Enable interface	checked

Bảng 3. 4 Thông số cấu hình interface loopback của external router

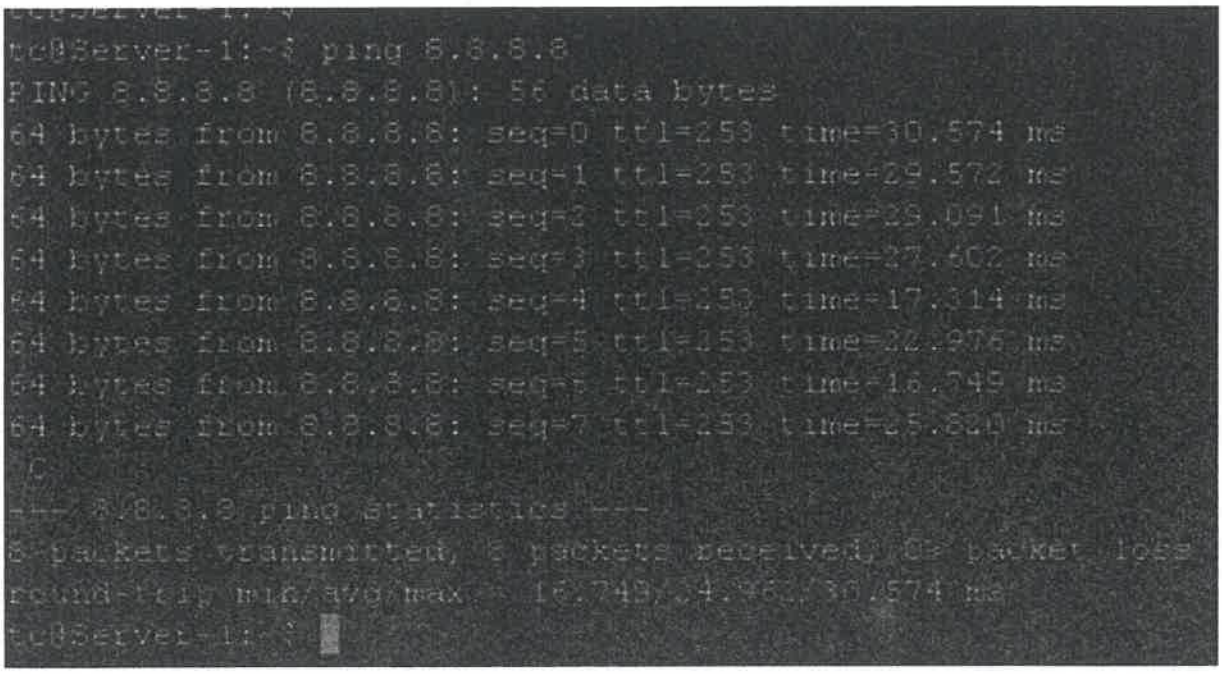
Sau khi nhập xong thông số cấu hình, thực hiện kiểm tra cấu hình và deploy.

Kế đến, thực hiện cấu hình quảng bá BGP của mạng 8.8.8.8/32 từ bộ định tuyến bên ngoài đến các Super Spine.

Trong trang *Create Policy*, cấu hình kết nối với các giá trị được thể hiện trong bảng

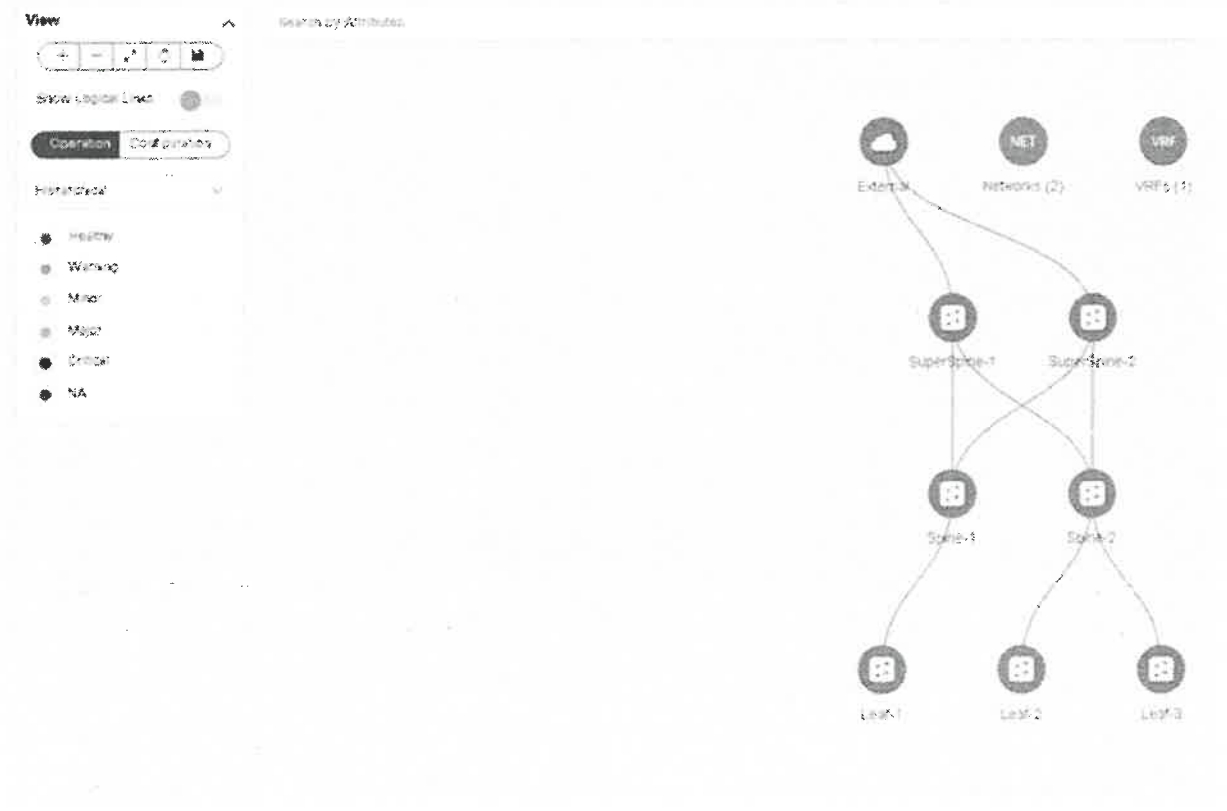
Field	Value
BGP AS #	65002
VRF Name	dcloud
IP prefix to advertise	8.8.8.8/32

Bảng 3. 5 Thông tin cấu hình để thực hiện quảng bá BGP



Hình 3. 7 Các server đã có thể kết nối tới tài nguyên bên ngoài

Sau khi kiểm tra và deploy cấu hình, các server đã có thể kết nối tới tài nguyên bên ngoài. Để xác minh điều này, thực hiện ping 8.8.8.8 (đại diện cho tài nguyên bên ngoài) từ server 1. Kết quả là, server 1 đã có thể kết nối tới tài nguyên bên ngoài, không bị drop gói tin nào.



Hình 3. 8 Topology trong NDFC sau khi thực hiện cấu hình

3.2.4 Nhận xét

Việc triển khai công nghệ VXLAN BGP EVPN trong mô hình mô phỏng không chỉ mang lại lợi ích về mặt kiến trúc và hiệu năng, mà còn được tối ưu hóa đáng kể nhờ vào công cụ quản lý Cisco Nexus Dashboard Fabric Controller (NDFC). Thông qua quá trình mô phỏng, có thể nhận thấy rõ ràng rằng NDFC đóng vai trò then chốt trong việc đơn giản hóa các tác vụ phức tạp liên quan đến triển khai và quản trị mạng VXLAN.

Trước hết, NDFC giúp đơn giản hóa cấu hình nhờ giao diện điều khiển tập trung và hỗ trợ tự động hóa nhiều thao tác cấu hình lặp đi lặp lại. Người quản trị không còn cần thiết phải sử dụng các dòng lệnh thủ công để triển khai hàng loạt tham số mạng như VRF, VNI, VLAN hay gán mạng cho các interface — tất cả đều có thể được thực hiện qua giao diện đồ họa một cách nhanh chóng và chính xác. Điều này giảm thiểu đáng kể sai sót cấu hình, đặc biệt hữu ích trong các môi trường mô phỏng, thử nghiệm hoặc triển khai thực tế.

Bên cạnh đó, NDFC cung cấp khả năng hiển thị và giám sát mạng toàn diện, giúp người vận hành có cái nhìn trực quan về toàn bộ topology mạng VXLAN, bao gồm trạng thái kết nối, luồng traffic, bảng điều khiển cấu hình và cảnh báo. Điều này giúp việc kiểm tra, xác minh và theo dõi hoạt động mạng trở nên đơn giản và hiệu quả hơn, đặc biệt trong giai đoạn kiểm chứng mô hình.

Không kém phần quan trọng, khả năng khắc phục sự cố cũng được nâng cao đáng kể. Nhờ vào các công cụ phân tích tích hợp trong NDFC, người quản trị có thể dễ dàng truy vết luồng EVPN/VXLAN, xác định nhanh các lỗi về định tuyến, ánh xạ VNI hoặc phân vùng VRF — từ đó xử lý sự cố một cách chính xác và nhanh chóng.

Tổng thể, sự kết hợp giữa VXLAN EVPN và NDFC không chỉ hợp lý về mặt kỹ thuật mà còn mang lại tính khả thi cao trong vận hành, giúp giảm tải đáng kể công việc cấu hình thủ công, đồng thời tăng cường tính ổn định và kiểm soát mạng trong cả môi trường mô phỏng lẫn triển khai thực tế

3.3. Kết luận chương 3

Chương 3 đã trình bày chi tiết về quá trình triển khai mô phỏng VXLAN trong môi trường mạng trung tâm dữ liệu sử dụng Cisco Nexus Dashboard Fabric Controller (NDFC). Chương này bắt đầu bằng việc cung cấp một cái nhìn tổng quan về dòng sản phẩm Cisco Nexus và nền tảng Cisco Nexus Dashboard, làm rõ hơn về vai trò và khả năng của các thành phần cốt lõi trong giải pháp.

Phần trọng tâm của chương là việc triển khai mô phỏng VXLAN, bao gồm giới thiệu mô hình mô phỏng cụ thể được sử dụng, kịch bản triển khai. Quan trọng nhất, chương này đã hướng dẫn chi tiết qua các bước thực hiện cụ thể, từ cấu hình ban đầu đến triển khai hoàn chỉnh một fabric VXLAN, chứng minh cách NDFC đơn giản hóa đáng kể quá trình phức tạp này.

Qua các bước triển khai đã thể hiện khả năng của Cisco NDFC trong việc đơn giản hóa cấu hình một mạng VXLAN EVPN. Việc sử dụng NDFC không chỉ giúp giảm thiểu thời gian và công sức cần thiết cho việc cấu hình mạng, mà còn tăng cường độ chính xác và giảm thiểu rủi ro lỗi do cấu hình thủ công. Điều này khẳng định NDFC là một công cụ mạnh mẽ và hiệu quả để triển khai các kiến trúc mạng hiện đại trong trung tâm dữ liệu.

KẾT LUẬN

Bài đề án này đã đi sâu vào nghiên cứu và trình bày về công nghệ VXLAN và ứng dụng của nó trong các trung tâm dữ liệu hiện đại, đặc biệt tập trung vào việc triển khai thực tế với Cisco Nexus Dashboard Fabric Controller (NDFC).

Chương 1 đã cung cấp một cái nhìn tổng quan toàn diện về công nghệ VXLAN, từ khái niệm cơ bản, nguyên nhân ra đời cho đến cấu trúc gói tin, các thành phần cốt lõi như VTEP, VNI, và mặt phẳng điều khiển. Chương này cũng làm rõ cách thức chuyển tiếp lưu lượng Unicast và BUM, cũng như cơ chế khám phá VTEP từ xa và học địa chỉ. Cuối cùng, đã phân tích chi tiết ưu điểm vượt trội của VXLAN so với VLAN truyền thống trong việc xây dựng các mạng quy mô lớn và linh hoạt, đồng thời chỉ ra một số hạn chế cần lưu ý.

Chương 2 tập trung vào ứng dụng của VXLAN trong bối cảnh các trung tâm dữ liệu. Chương này đã xem xét sự phát triển của kiến trúc mạng trung tâm dữ liệu, từ mô hình ba lớp truyền thống đến mô hình Clos hiện đại và vai trò của các giao thức như Spanning-tree và vPC. Trọng tâm của chương này là việc làm rõ mối quan hệ giữa ảo hóa mạng, khái niệm overlay và underlay, và đặc biệt là công nghệ EVPN-VXLAN. EVPN-VXLAN đã được chứng minh là một giải pháp mạnh mẽ, kết hợp những ưu điểm của VXLAN với mặt phẳng điều khiển BGP-EVPN, mang lại khả năng mở rộng, ổn định và hiệu quả cao cho các trung tâm dữ liệu.

Cuối cùng, chương 3 đã trình bày một kịch bản triển khai mô phỏng VXLAN trong mạng trung tâm dữ liệu sử dụng Cisco Nexus Dashboard Fabric Controller (NDFC). Chương này đã cung cấp tổng quan về dòng thiết bị Cisco Nexus và nền tảng NDFC, sau đó mô tả chi tiết mô hình, kịch bản và các bước thực hiện cụ thể. Qua quá trình triển khai, những nhận xét đã được đưa ra, nhấn mạnh rằng NDFC đã đơn giản hóa đáng kể quá trình cấu hình và quản lý các fabric VXLAN phức tạp, giảm thiểu lỗi thủ công.

Thông qua đề án này, VXLAN, đặc biệt khi kết hợp với EVPN và được quản lý bởi các công cụ như Cisco NDFC, là giải pháp tối ưu cho việc xây dựng các mạng trung tâm dữ liệu thế hệ mới. Các mạng này đáp ứng được yêu cầu về khả năng mở rộng, tính linh hoạt, tính sẵn sàng cao và hỗ trợ hiệu quả cho môi trường ảo hóa và điện toán đám mây.

Sau khi đã hoàn thành việc thiết kế và cấu hình VXLAN EVPN trong một site đơn lẻ một cách hiệu quả thông qua NDFC, hướng nghiên cứu tiếp theo là mở rộng sang kiến trúc EVPN VXLAN đa site (multi-site), khai thác triệt để sức mạnh của tự động hóa.

Kiến trúc EVPN VXLAN đa site cho phép kết nối liền mạch và thống nhất các trung tâm dữ liệu hoặc các khu vực địa lý khác nhau, biến chúng thành một mạng lưới lớn hơn, linh hoạt và dễ quản lý. Điều này không chỉ cung cấp khả năng mở rộng vượt trội cho hạ tầng đám mây và ứng dụng phân tán, mà còn là yếu tố then chốt cho các doanh nghiệp có hoạt động phân tán, yêu cầu khả năng khôi phục sau thảm họa (Disaster Recovery) mạnh mẽ hoặc cần tối ưu hóa hiệu suất ứng dụng trên phạm vi rộng.

Hướng phát triển này sẽ tập trung vào việc tăng cường tự động hóa trong triển khai và quản lý đa site. NDFC sẽ đóng vai trò trung tâm trong việc tự động hóa quá trình kết nối các fabric VXLAN EVPN riêng lẻ, thiết lập chính sách đồng bộ và quản lý lưu lượng giữa các site. Điều này giảm thiểu đáng kể công sức cấu hình thủ công, loại bỏ lỗi tiềm ẩn và tăng tốc độ triển khai các dịch vụ mới trên toàn bộ hạ tầng mạng phân tán, đảm bảo tính nhất quán và hiệu quả vận hành tối đa.

TÀI LIỆU THAM CHIẾU

[1] "Data Center Virtualization Market to Reach USD 28.9 Billion by 2032, Driven by the Growing Need for Scalable, Cost-Effective Infrastructure Solutions | Research by SNS Insider," SNS Insider pvt ltd, 06 December 2024. [Online].

[2] Firas Ahmed, Somit Maloo, CCNP and CCIE Data Center Core DCCOR 350-601 Official Cert Guide, Cisco Press, 2020.

[3] George, "VXLAN: the Future for Data Center Networks," FS, 10 July 2024. [Online].

[4] Zhang Yuting, Chen Li, "What Is VXLAN?," Huawei, 25 November 2024. [Online].

[5] Rich Castagna, "Techtarget," 10 benefits of server virtualization for businesses, 26 February 2021. [Online].

[6] H. Farag, "VXLAN," 31 Oct 2018. [Online].

[7] I. Mirza, "What is VXLAN?," CBT Nugets powered by Adept, 27 Feb 2024. [Online].

[8] Kinza Yasar, Peter Loshin, Ben Lutkevich, How to design and build a data center, TechTarget, 2024.

[9] Deepti Chandra, Data Center Deployment with EVPN/VXLAN, Juniper Networks.

[10] Somit Maloo, Iskren Nikolov, Cisco Data Center Fundamentals, Cisco Press, 2022.

[11] Venkata Josyula, Malcolm Orr, Greg Page, Cloud Computing: Automating the Virtualized Data Center, Cisco Press, 2012.

[12] David Jansen, Lukas Krattiger, Shyam Kapadia, Building Data Centers with VXLAN BGP EVPN: A Cisco NX-OS Perspective, Cisco Press, 2017.

[13] "VXLAN Packet Encapsulation Format," Huawei. [Online].

[14] V. Deshpande, "VXLAN Series – Multicast usage in VXLAN," VMware by Broadcom. [Online].

KiểmTraTàiLiệu

BÁO CÁO KIỂM TRA TRÙNG LẶP

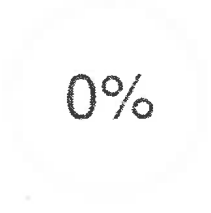
Thông tin tài liệu

Tên tài liệu:	De an tot nghiep_Vu Thi Thanh Tu
Tác giả:	Vũ Thị Thanh Tú
Điểm trùng lặp:	5
Thời gian tải lên:	15:14 01/08/2025
Thời gian sinh báo cáo:	15:17 01/08/2025
Các trang kiểm tra:	82/82 trang

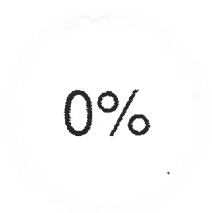
Kết quả kiểm tra trùng lặp

Có 5% nội dung trùng lặp

Có 95% nội dung không trùng lặp



Có 0% nội dung người dùng loại trừ



Có 0% nội dung hệ thống bỏ qua

Nguồn trùng lặp tiêu biểu

aws.amazon.com 123docz.net wikimaytinh.com

Vũ Thị Thanh Tú

BỘ KHOA HỌC VÀ CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BÁO CÁO GIẢI TRÌNH
SỬA CHỮA, HOÀN THIỆN ĐỀ ÁN TỐT NGHIỆP

Họ và tên học viên: Vũ Thị Thanh Tú

Chuyên ngành: Kỹ thuật Viễn thông

Khóa: 2023 đợt 2

Tên đề tài: Nghiên cứu giải pháp VXLAN trong các trung tâm dữ liệu

Người hướng dẫn khoa học: TS. Vũ Tuấn Lâm

Ngày bảo vệ: 19/07/2025

Các nội dung học viên đã sửa chữa, bổ sung trong đề án tốt nghiệp theo ý kiến đóng góp của Hội đồng chấm đề án tốt nghiệp:

TT	Ý kiến hội đồng	Sửa chữa của học viên
1	Rà soát hiệu chỉnh theo ý kiến đóng góp của phản biện và các thành viên hội đồng	Tiếp thu góp ý của Hội đồng, học viên đã rà soát hiệu chỉnh theo ý kiến đóng góp của phản biện và các thành viên hội đồng tại mục 2.1.1 của chương 2
2	Bổ sung, phân tích để làm rõ sự lựa chọn VXLAN, phần mô hình mô phỏng trong chương 3	Tiếp thu góp ý của Hội đồng, học viên đã bổ sung, phân tích sự lựa chọn VXLAN, phần mô hình mô phỏng tại mục 3.2 trong chương 3
3	Đánh số lại đề mục	Học viên đã rà soát và chỉnh sửa lại đề mục

Hà Nội, ngày 4 tháng 8 năm 2025

Ký xác nhận của

CHỦ TỊCH HỘI ĐỒNG CHẤM ĐỀ ÁN	THƯ KÝ HỘI ĐỒNG	NGƯỜI HƯỚNG DẪN KHOA HỌC	HỌC VIÊN
PGS.TS Lê Nhật Thăng	TS. Nguyễn Thị Thu Hiên	TS. Vũ Tuấn Lâm	Vũ Thị Thanh Tú

BỘ KHOA HỌC VÀ CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BIÊN BẢN
HỌP HỘI ĐỒNG CHẤM ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

Căn cứ quyết định số Quyết định số 1098/QĐ-HV ngày 26 tháng 06 năm 2025 của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ. Hội đồng đã họp vào hồi...9....giờ...55.phút, ngày 19 tháng 07 năm 2025 tại Học viện Công nghệ Bưu chính Viễn thông để chấm đề án tốt nghiệp thạc sĩ cho:

Học viên: **Vũ Thị Thanh Tú**

Tên đề án tốt nghiệp: **Nghiên cứu giải pháp VXLAN trong các trung tâm dữ liệu**

Chuyên ngành: **Kỹ thuật viễn thông** Mã số: **8520208**

Các thành viên của Hội đồng chấm đề án tốt nghiệp có mặt: ..05.../ 05

TT	HỌ VÀ TÊN	TRÁCH NHIỆM TRONG HĐ	GHI CHÚ
1	**PGS. TS. Lê Nhật Thăng**	Chủ tịch	
2	**TS. Nguyễn Thị Thu Hiên**	Thư ký	
3	**TS. Nguyễn Hồng Thủy**	Phản biện 1	
4	**PGS.TS. Nguyễn Chiến Trinh**	Phản biện 2	
5	**TS. Lê Anh Ngọc**	Uỷ viên	

Các nội dung thực hiện:

1. Chủ tịch Hội đồng điều khiển buổi họp. Công bố quyết định của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ.
2. Người hướng dẫn khoa học hoặc thư ký đọc lý lịch khoa học và các điều kiện bảo vệ đề án tốt nghiệp của học viên. (có bản lý lịch khoa học và kết quả các môn học cao học của học viên kèm theo).
3. Học viên trình bày tóm tắt đề án tốt nghiệp.
4. Phản biện 1 đọc nhận xét (có văn bản kèm theo)
5. Phản biện 2 đọc nhận xét (có văn bản kèm theo)
6. Các câu hỏi của thành viên Hội đồng:

1) Làm rõ thêm về Data center
2) Tại sao chọn mô phỏng VXLAN bằng Netsim?
3) Hãy đề xuất việc áp dụng VXLAN vào thực tế
4) Phân tích ưu nhược điểm của VXLAN?

7. Trả lời của học viên:

1) Đã trình bày rõ đề cập đến đề tài và khả năng áp dụng thực tế

2) Việc triển khai dự kiến thực hiện theo quy mô ... Cisco là hãng hỗ trợ công nghệ mạng, trong quá trình triển khai là ...

8. Thư ký đọc nhận xét về quá trình thực hiện đề án tốt nghiệp của học viên *(có văn bản kèm theo)*.

9. Hội đồng họp riêng:

- Bầu Ban kiểm phiếu:
  1. Trưởng Ban kiểm phiếu: TS. Lê Anh Ngọc
  2. Uỷ viên Ban kiểm phiếu: TS. Nguyễn Hải Thanh
  3. Uỷ viên Ban kiểm phiếu: Nguyễn Thị Thu Hiên
- Hội đồng chấm đề án tốt nghiệp bằng bỏ phiếu kín.
- Ban kiểm phiếu làm việc:
- Trưởng Ban kiểm phiếu báo cáo kết quả kiểm phiếu *(có Biên bản họp Ban kiểm phiếu kèm theo)*
- **Điểm trung bình của đề án tốt nghiệp:** .....8,0.....

Kết luận:

1. Các nội dung cần chỉnh sửa, hoàn thiện sau bảo vệ đề án tốt nghiệp:
   Rà soát kiểm chỉnh theo ý kiến đã góp phần hiệu và thành viên hội đồng.
   Bổ sung phần đề cập và lý lựa chọn VXLAN, phân mô hình mô phỏng chuẩn 3
   Danh mục tài liệu để mục
2. Đề nghị Học viện công nhận (hoặc không) và cấp bằng (hoặc không) thạc sĩ cho học viên:
   Đề nghị công nhận và cấp bằng thạc sĩ cho học viên
3. Đề án tốt nghiệp có thể phát triển thành đề tài nghiên cứu cho NCS.....không...(1/5).....

Buổi làm việc kết thúc vào..11h00.... cùng ngày.

Chủ tịch	**Thư ký**
PGS. TS. Lê Nhật Thăng	TS. Nguyễn Thị Thu Hiên

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc

----------oOo----------

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

(Dùng cho người phản biện)

Tên đề tài luận văn: Nghiên cứu giải pháp VXLAN trong các trung tâm dữ liệu.

Chuyên ngành: Kỹ thuật viễn thông.

Mã chuyên ngành: 60.52.02.08

Họ và tên học viên: Vũ Thi Thanh Tú

Họ và tên người nhận xét: Nguyễn Chiến Trinh

Học hàm, học vị: Phó Giáo sư - Tiến sĩ.

Chuyên ngành: Điện tử - Viễn thông

Cơ quan công tác: Học viện Công nghệ Bưu chính Viễn thông.

Số điện thoại: 0915400946.....E-mail: trinhnc@ptit.edu.vn.

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Các trung tâm dữ liệu hiện nay đang trở thành nền tảng để phát triển và cung cấp dịch vụ, cũng như cho các công nghệ khác nhau như điện toán đám mây, dữ liệu lớn, AI, …. Đề tài "Nghiên cứu giải pháp VXLAN trong các trung tâm dữ liệu" nghiên cứu triển khai giải pháp công nghệ cho TT dữ liệu do vậy có ý nghĩa khoa học và thực tiễn, hỗ trợ các doanh nghiệp trong quá trình xây dựng, vận hành, khai thác TT dữ liệu.

II/ Nội dung của luận văn, các kết quả đã đạt được:

Luận văn bao gồm 3 chương: giới thiệu tổng quan công nghệ VXLAN, giải pháp VXLAN cho TT dữ liệu, triển khai mô phỏng VXLAN trong TT dữ liệu trên thiết bị Cisco. Bố cục của đề án hợp lí, tuy nhiên nội dung các chương trình bày trong đề án chưa bám sát đề cương được duyệt. Cụ thể một số vấn đề của đề án cần hoàn thiện:

- Nên có phân tích giải pháp VXLAN và khả năng ứng dụng cho TT dữ liệu so với các giải pháp khác.
- Cần có mô tả mô hình và kịch bản mô phỏng chi tiết.
- Chương 3 cần có các nhận xét về kết quả thử nghiệm đánh giá cho giải pháp VXLAN để khẳng định lựa chọn giải pháp và các ưu điểm đem lại cho TT dữ liệu.

- Hình thức trình bày: đánh số lại các mục, các hình vẽ còn mờ, trích dẫn tài liệu tham khảo.

III/ Những vấn đề cần giải trình thêm:

1) Phân tích sự cần thiết và nêu ưu điểm khi triển khai VXLAN cho TT dữ liệu.
2) Chỉ rõ trong thử nghiệm hoạt động và kết quả của VXLAN trong phần mô phỏng chương 3?

IV/ Kết luận:

Đề án đạt yêu cầu của đề án hướng ứng dụng tốt nghiệp Thạc sĩ.

Đồng ý cho phép học viên bảo vệ trước hội đồng.

Hà Nội, ngày tháng năm 2025

NGƯỜI NHẬN XÉT

PGS.TS Nguyễn Chiến Trinh

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc

BẢN NHẬN XÉT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

(*Dùng cho người phản biện*)

Tên đề tài đề án tốt nghiệp: "**Nghiên cứu giải pháp VXLAN trong các trung tâm dữ liệu**"

Chuyên ngành: Kỹ thuật viễn thông.....

Mã chuyên ngành: 8520208 .....

Họ và tên học viên: Vũ Thị Thanh Tú.....

Họ và tên người nhận xét: Nguyễn Hồng Thủy .....

Học hàm, học vị: Tiến sĩ.....

Chuyên ngành: Kỹ thuật viễn thông.....

Cơ quan công tác: Bộ Công an.....

Số điện thoại:.098115026C.....E-mail: Enzonguyen775@gmail.com

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Đề tài có cơ sở khoa học và thực tiễn cao và cấp thiết vì:

- Nền tảng kỹ thuật Trung tâm dữ liệu, kỹ thuật VXLAN (mạng LAN ảo mở rộng) đã được nghiên cứu, đánh giá, có thể ứng dụng cho nội dung nghiên cứu của đề tài.

- Nhu cầu cấp thiết và sự phát triển mạnh mẽ của các mô hình mới của Trung tâm dữ liệu tại Việt Nam đòi hỏi nghiên cứu các ứng dụng như VXLAN.

II/ Nội dung của đề án tốt nghiệp, các kết quả đã đạt được:

Nội dung cơ bản của đề án tốt nghiệp đã bám sát mục tiêu, yêu cầu và nội dung của đề cương đã được duyệt. Những kết quả nghiên cứu lý thuyết, cơ sở khoa học có đủ căn cứ để học viên đề xuất giải pháp ứng dụng VXLAN trong trung tâm dữ liệu.

Việc triển khai mô phỏng VZXLAN trong trung tâm dữ liệu bởi một dòng thiết bị chuyển mạch VXLAN đã có kết quả, minh chứng một phần cho giải pháp đề xuất của đề tài. Cisco Systems là một trong những nhà cung cấp chính

trên thế giới về các sản phẩm cho mạng trung tâm dữ liệu, do vậy kết quả nghiên cứu và mô phỏng có thể ứng dụng thực tiễn.

III/ Những vấn đề cần giải thích thêm:

Tuy nhiên, học viên cần giải thích thêm và làm rõ một số vấn đề sau:

- Chỉnh sửa, bổ sung định nghĩa Trung tâm dữ liệu mang tính khoa học, khái quát hơn.
- Bổ sung luận giải rõ về việc chọn mô phỏng VXLAN bằng thiết bị Cisco? So sánh kết quả mô phỏng bằng Cisco với một số loại thiết bị khác.
- Học viên hãy đề xuất việc ứng dụng VXLAN vào xây dựng trung tâm dữ liệu lớn, phạm vi rộng như các trung tâm dữ liệu nhà nước và doanh nghiệp lớn của Việt Nam hiện nay sẽ phát sinh những vấn đề chính gì về an ninh mạng và biện pháp giải quyết?

IV/ Kết luận:

Đồng ý cho phép học viên bảo vệ đề án tốt nghiệp sau khi làm rõ và chỉnh sửa các vấn đề nêu tại phần III trên.

Ngày 15 tháng 7 năm 2025

NGƯỜI NHẬN XÉT

Ts. Nguyễn Hồng Thủy

NHẬN XÉT, GÓP Ý