[CASE STUDY] Thiết Kế Site Trung Tâm SD-WAN Có Bảo Mật: Cái Giá Của "An Toàn" Khi Dữ Liệu Phải Qua Tường Lửa

Ở các doanh nghiệp lớn, đặc biệt là khi triển khai SD-WAN theo mô hình **Hub & Spoke**, việc bảo vệ trung tâm dữ liệu (Data Center) bằng **Firewall** là chuyện hiển nhiên. Nhưng khi bắt tay vào thiết kế thực tế, **"gài" thêm FW** giữa các thành phần SD-WAN – đặc biệt giữa **cEdge và Controllers** – lại mở ra một loạt vấn đề khiến nhiều anh em phải vò đầu. Cùng mổ xẻ case sau – một thiết kế phổ biến mà nhiều tổ chức đang áp dụng:

Kiến Trúc:

- Các **Controllers** (vBond, vSmart, vManage) được đặt tại **On-Prem DC**.
- CEdge trung tâm (DC cEdge) kết nối WAN đa vùng: MPLS1, MPLS2, Metro-E, Internet.
- CEdge này phải đi qua một cụm **Firewall DC (DC FW Cluster)** trước khi tới các Controllers.
- Cạnh đó là hàng loạt remote cEdges ở các chi nhánh, kết nối về trung tâm theo mô hình Hub & Spoke.

🔐 Yêu Cầu Bảo Mật:

1. **Toàn bộ traffic từ WAN về Controllers phải qua Firewall**.
2. **DC cEdge phải thiết lập kết nối TLS với Controllers trên mỗi TLOC** – đây là yêu cầu **bắt buộc** để **tunnel dataplane SD-WAN** có thể hình thành.

⚠️ Hai Vấn Đề Kỹ Thuật Cực Kỳ Hay Gặp:

❌ Vấn đề 1: Không Có Route VPN0 đến Controllers

- Mỗi TLOC của DC cEdge (VPN0) không có tuyến đi đến bộ điều khiển nếu firewall không cho phép.
- Kết quả: DC cEdge không thể thiết lập **TLS** đến vSmart/vBond/vManage qua các đường này.

- Hệ quả: **Remote cEdges** sẽ không thể hình thành tunnel dữ liệu SD-WAN được.
- Giải pháp tạm: Cho phép remote cEdge thiết lập TLS tới Controllers thông qua Hub – nhưng đây là **workaround**, không bền vững.

🧱 **Vấn đề 2: TLS/DTLS bị chặn bởi ACL ngầm**

- Ngay cả khi các tunnel từ xa gửi TLS/DTLS về, gói tin sẽ **bị drop ngầm** bởi ACL nội tại trên interface TLOC của DC cEdge.
- Lý do? **Cisco SD-WAN có một implicit ACL trên tất cả TLOC interface**, thường chỉ cho phép traffic "expected" (control plane từ DC) đi vào.
- Nếu không cấu hình rõ ràng (VD: policy-based access), traffic TLS/DTLS từ remote bị lọc sạch, không báo lỗi rõ ràng.

☑️ Bài học rút ra:

- Khi thiết kế **Hub Site có Firewall**, chúng ta **phải tính đến Control Plane Traffic** đi vòng qua firewall: từ cEdge đến Controllers.
- ACL trên interface TLOC của cEdge cần được **tùy chỉnh cho phép lưu lượng TLS/DTLS** từ remote nếu mô hình dùng Hub Relay cho control plane.
- Đừng quên kiểm tra kỹ **route trong VPN0**, đây là một phần dễ bỏ sót khi firewall filter theo vùng mạng.

💬 **Bạn đã từng gặp lỗi "mọi thứ đều đúng mà vẫn không hình thành BFD/OMP giữa remote và hub" chưa? Rất có thể đây chính là nguyên nhân.**

Giải thích hiện tượng: Vì sao các Controllers lại "tự kết nối nhau qua IP public"?

Bối cảnh:

- Bạn NAT các IP nội bộ của vBond, vSmart, vManage thành **IP Public** trên FortiGate để các remote cEdge có thể kết nối từ Internet.
- FortiGate là thiết bị **NAT ngoài**, nằm giữa DC và Internet.

Điều thú vị nằm ở cách mà Cisco SD-WAN controllers tự phát hiện và giao tiếp:

- Trong quá trình khởi tạo **control plane**, mỗi Controller sẽ **quảng bá TLOC hoặc địa chỉ mình đang dùng** vào hệ thống (vBond/vSmart thông báo IP cho vManage, và ngược lại).
- Nếu lúc này các controller **thấy nhau qua địa chỉ public**, thì thông tin đó sẽ được ghi nhận và dùng để thiết lập kết nối nội bộ giữa chúng.
- Vì bạn NAT ra Internet, nên khi controller 1 kết nối controller 2, nó không "biết" IP private của controller kia – chỉ biết IP public được quảng bá.

Hệ quả: Controller nằm cùng site lại **cố kết nối lẫn nhau qua IP Public**, mà route đó thì không tồn tại (vì nằm sau FortiGate). Đây là hành vi mặc định do **DNS, discovery info, và TLOC advertisement** không được ép buộc rõ IP nội bộ.

🛠️ Cách 1: Giải pháp Hairpin NAT trên Firewall

- **Hairpin NAT (loopback NAT)** cho phép traffic từ bên trong FortiGate (VD: vSmart) đi ra public IP của vBond (cũng nằm trong nội bộ), rồi FortiGate sẽ **loopback NAT ngược về IP private**.
- Kiểu NAT này được gọi là "**NAT reflection**" hoặc "**U-turn NAT**" trong nhiều tài liệu.

Cấu hình thường gặp trên FortiGate:

- Tạo một policy cho phép traffic từ zone LAN đi đến IP public đã NAT của controller.
- Áp dụng **NAT policy ngược chiều** để phản hồi đúng source IP.
- Bật tính năng **NAT hairpin hoặc enable nat-loopback**.

Gợi ý cải tiến dài hạn

Bạn có thể:

1. **Tách riêng NAT cho remote cEdge** và không dùng IP public cho các controller khi các thiết bị nằm cùng DC.
2. Cấu hình **DNS nội bộ** cho controller trỏ về IP private thay vì IP public.

3. Dùng **policy-based routing (PBR)** hoặc SD-WAN policies để ép controller nội bộ chỉ đi theo đường LAN (không qua NAT).
4. Dùng collapse Core (L2,L3 Switching/Routing) trong Data Center.

Tóm lại:

Controllers "thấy nhau qua IP public" là vì bạn đã NAT các controller và quá trình control plane discovery của Cisco SD-WAN không phân biệt ai là "internal". Cách xử lý của bạn dùng **Hairpin NAT** là hợp lý, nhưng nên đánh giá kỹ xem có cần **chuyển sang kiến trúc nội bộ thuần private IP** để tránh complexity trong troubleshooting và vận hành.

🔥 Cách 4: Thiết kế Backbone Thu Gọn – Tối Ưu Hóa Kiến Trúc Trung Tâm Dữ Liệu cho SD-WAN 🔥

Trong các triển khai SD-WAN ở cấp độ doanh nghiệp, đặc biệt tại **Hub Site trung tâm**, một thách thức lớn là làm sao để vừa đơn giản hóa thiết kế mạng, vừa đảm bảo tính linh hoạt cao cho chuỗi dịch vụ bảo mật và điều khiển (service chaining). Và giải pháp "backbone thu gọn" (collapsed backbone) ra đời như một phương pháp đầy hiệu quả.

Thiết Kế Backbone Thu Gọn là gì?

Thay vì phân chia riêng biệt tầng L2 (access) và L3 (distribution/core), mô hình **Collapsed Backbone Design** gộp hai lớp lại thành một cặp **switch tổng hợp đa năng** (L2/L3 aggregation switch), đóng vai trò vừa định tuyến vừa chuyển mạch. Đây là backbone vật lý và logic duy nhất tại hub, nơi kết nối toàn bộ thiết bị và kết nối WAN.

Tại hub:

- Tất cả các thiết bị như Firewall, Edge Router, Controller, Server,... đều **kết nối vào cặp switch L3 HA này**.
- Switch hỗ trợ cả cổng L2 và L3 để **hỗ trợ đa dạng dịch vụ và phân vùng VLAN**.
- Backbone được gọi là "thu gọn" vì không còn sự phân lớp truyền thống mà vẫn đảm bảo khả năng scale và resilience thông qua cặp switch HA.

Chuỗi Dịch Vụ (Service Chain) – VLAN & Bảo Mật

Thiết kế backbone thu gọn cho phép tạo ra các chuỗi dịch vụ nội tuyến (inline) bằng VLAN. Dưới đây là 2 use-case phổ biến:

🔐 1. TLS/DTLS Service Chain – Hub Edge → Firewall → Controllers

- Dành cho kết nối **control plane** giữa Edge và các thiết bị điều khiển (controller).

- Giao tiếp được mã hóa qua TLS hoặc DTLS.
- Chuỗi dịch vụ định tuyến VLAN gồm:
    - VLAN20: kết nối Edge đến Firewall (Gig1)
    - VLAN20: tiếp tục từ DMZ của Firewall
    - VLAN10: từ Inside của Firewall đến Controller (Gig0)

👉 VLAN được dùng để xác định luồng dịch vụ rõ ràng và cách ly bảo mật.

2. IPsec Service Chain – Hub Edge → Firewall → Internet

- Dành cho **data plane** đi từ Remote Edge qua Firewall trước khi ra Internet.
- Các VLAN được định nghĩa như sau:
    - VLAN20: kết nối từ Edge (Gig1) đến Firewall DMZ
    - VLAN11: kết nối từ Outside của Firewall ra Internet thông qua ISP
- **SVI 11** (Switch Virtual Interface) đại diện cho gateway định tuyến đến ISP.

👉 Điều này giúp thực thi chính sách NAT, IDS/IPS, DLP,... trước khi IPsec tunnel được thiết lập qua Internet.

Lợi Ích Thiết Kế

1. **Giảm chi phí thiết bị** – Không cần phân lớp access/distribution riêng biệt.
2. **Đơn giản hóa quản lý** – Tập trung cấu hình tại 2 switch backbone chính.
3. **Tối ưu bảo mật theo chuỗi VLAN** – Hỗ trợ dễ dàng cho các service chain bảo mật nội tuyến.
4. **Hỗ trợ đa dịch vụ WAN** – Internet, MPLS1, MPLS2, Metro-E,... đều kết nối được trực tiếp qua L3 interface.
5. **Khả năng chịu lỗi cao** – Với mô hình HA (high availability) trên cặp switch backbone.

Kết luận

Collapsed backbone không chỉ là một lựa chọn "cắt giảm chi phí", mà là **xu hướng thiết kế thông minh** cho các tổ chức đang chuyển dịch sang kiến trúc SD-WAN hiện đại, đặc biệt là khi cần tích hợp sâu các chuỗi dịch vụ bảo mật, kiểm soát truy cập, và giám sát hiệu năng.

Nếu bạn đang thiết kế Hub Site cho SD-WAN, hãy cân nhắc mô hình này. Và đừng quên — VLAN và các interface ảo (SVI) là “vũ khí chiến lược” để xây chuỗi bảo mật trong SD-WAN fabric!

Cách dùng SDWAN Policy

SD-WAN (SD-WAN Policies) để **ép traffic giữa các controller nội bộ chỉ đi qua đường LAN, không qua IP Public hay NAT**. Đây là một **giải pháp "chuẩn chỉnh" hơn** so với Hairpin NAT, vì ta kiểm soát đường đi của lưu lượng ngay từ thiết bị SD-WAN, **tránh lệ thuộc vào firewall**. Sau đây là hướng dẫn chi tiết và các lưu ý khi triển khai.

Mục tiêu

Các controller như **vBond, vSmart, vManage** đang nằm trong cùng một DC (hoặc cùng site), nhưng do bị NAT ra Public IP, nên chúng cố gắng **kết nối với nhau bằng IP public qua Internet**.

Giờ ta muốn:

- Các controller **giao tiếp nội bộ qua IP private**
- SD-WAN cEdge tại DC **route đúng traffic nội bộ controller qua LAN interface**, **bỏ qua tunnel TLOC và NAT**

Hướng triển khai

1. Tạo Localized Data Policy hoặc Centralized Control Policy

Chính sách này dùng để ép traffic đi theo đường mong muốn (direct LAN route thay vì qua TLOC).

Với mô hình Hub Site có vEdge/cEdge tại DC, bạn nên dùng Centralized Control Policy để đảm bảo tính đồng bộ.

2. Tạo SLA Class hoặc TLOC Preference (tuỳ yêu cầu)

Tuy nhiên, trong trường hợp này, **ta không cần SLA Tracking**, mà chỉ cần **Match + Action Forward** cụ thể theo interface LAN.

3. Cấu hình Match cho controller IP nội bộ

Giả sử IP private các controller:

- vManage: 10.1.1.1
- vSmart: 10.1.1.2
- vBond: 10.1.1.3

Bạn có thể match như sau:

```
control-policy LAN-Controller-Only
 sequence 10
  match
   source-data-prefix 10.1.1.0/24
   destination-data-prefix 10.1.1.0/24
  action accept
 default-action reject
```

Mục tiêu: chỉ cho phép control plane traffic giữa các controller nội bộ đi qua đường private IP.

4. Gắn Policy vào Site/Device phù hợp

```
apply-policy
 control-policy LAN-Controller-Only direction inbound
```

💡 Chỉ cần gán tại **site chứa các controller**, không cần propagate toàn bộ remote edge.

📌 **Kỹ thuật nâng cao: Dùng data-policy để điều khiển flow đi ra ngoài**

Nếu bạn đang dùng các controller chạy trong container (VD: vManage docker trong UCS), bạn có thể kết hợp:

- **Data-Policy:** để xử lý cả traffic data đi từ controller ra ngoài
- **TLOC route preference:** để ép traffic *không bao giờ chọn TLOC ra Internet/public*

Lưu ý:

1. **Bản chất Discovery vBond vẫn thấy IP Public** do nó đang NAT – nên bạn có thể cần ép lại IP quảng bá bằng lệnh:

system-ip 10.1.1.1

vbond 10.1.1.3 local

2. **Nếu đã publish DNS record trỏ về IP public**, cần cân nhắc sửa lại bản ghi DNS nội bộ để các controller resolve đúng IP private.

☑ Lợi ích khi làm theo cách này:

- Không cần **Hairpin NAT** – tránh tạo dependency vào Fortigate.
- Controller luôn sử dụng **đường LAN nội bộ ổn định, tốc độ cao**, không bị delay do tunnel qua WAN.
- Dễ quản lý, scale-out khi bạn dùng nhiều controller (HA/Cluster).

🎯 Mục tiêu

- vManage: 10.1.1.1
- vSmart: 10.1.1.2
- vBond: 10.1.1.3

Traffic giữa các controller này **phải đi nội bộ qua LAN** tại site DC.

🧱 Cấu hình Control Policy ép route nội bộ

Bước 1: Tạo prefix-list các IP private controller

vEdge(config)# policy

vEdge(config-policy)# prefix-list CONTROLLER-LOCAL

vEdge(config-prefix-list)# ip-prefix 10.1.1.0/24

vEdge(config-prefix-list)# exit

Bước 2: Tạo Control Policy cho control-plane

vEdge(config-policy)# control-policy FORCE-LAN-CTRL

vEdge(config-control-policy)# sequence 10

vEdge(config-sequence)# match

vEdge(config-match)# source-data-prefix CONTROLLER-LOCAL

vEdge(config-match)# destination-data-prefix CONTROLLER-LOCAL

vEdge(config-match)# exit

vEdge(config-sequence)# action accept

vEdge(config-sequence)# exit

vEdge(config-control-policy)# default-action reject

vEdge(config-control-policy)# exit

Bước 3: Áp dụng policy tại site DC (site chứa controllers)

Giả sử site của bạn là site ID 100 (xem với show sdwan system status)

vEdge(config-policy)# apply-policy

vEdge(config-apply-policy)# site-list DC-SITE

vEdge(config-apply-policy-site-list)# site-id 100

vEdge(config-apply-policy-site-list)# exit

vEdge(config-apply-policy)# control-policy FORCE-LAN-CTRL direction inbound

vEdge(config-apply-policy)# exit

vEdge(config-policy)# exit

☑ Kiểm tra lại

Sau khi commit cấu hình (commit and-quit), bạn có thể kiểm tra:

show policy from-vsmart control local-applications

Hoặc kiểm tra routing thực tế:

show ip route vpn 0 10.1.1.0/24

Đảm bảo rằng route tới các IP controller là **qua interface LAN** chứ không phải TLOC tunnel hoặc IP public.

Cách dùng DNS nội bộ để ép các Controller (vBond, vSmart, vManage) giao tiếp với nhau **qua IP private,**

tránh việc bị "bắt tay nhau qua IP Public" như bạn đang gặp phải.

☑ Vì sao dùng DNS nội bộ lại hiệu quả?

Trong Cisco SD-WAN, **controller discovery và communication** có thể dựa vào **DNS hostname**, đặc biệt là với **vBond** khi khởi tạo kết nối.

➡ Nếu DNS trả về **IP public** (do ghi ở ngoài Internet), thì các controller và thiết bị remote sẽ kết nối qua IP đó — kể cả khi bạn đang ở cùng site.

➡ Nếu bạn **kiểm soát DNS**, bạn có thể làm cho các thiết bị trong nội bộ **resolve hostname controller về IP private**, từ đó **tránh việc đi qua NAT/public**.

🧭 Mô hình lý tưởng:

Giả sử:

- vbond.sdwan.local → IP public: 203.0.113.10 (bản ghi ngoài Internet)
- IP private thật sự của vBond tại DC: 10.1.1.3

🛠️ Cách triển khai DNS nội bộ

1. Triển khai một DNS nội bộ (Internal DNS Server)

- Có thể dùng Windows Server, dịch vụ DNS trên Firewall, hoặc máy Linux với bind9, dnsmasq,...
- Tạo các bản ghi DNS cho các hostname controller:

vbond.sdwan.local A 10.1.1.3

vsmart.sdwan.local A 10.1.1.2

vmanage.sdwan.local A 10.1.1.1

🎯 Trên DNS nội bộ – các bản ghi hostname controller sẽ trả về IP nội bộ.

2. Cấu hình thiết bị trong nội bộ DC dùng DNS nội bộ

Đối với Controller (vBond, vSmart, vManage):

- Truy cập Linux shell (vshell) nếu dùng bản dạng ISO.
- Sửa file /etc/resolv.conf hoặc cấu hình DHCP nội bộ cấp DNS server nội bộ.
- Kiểm tra:

ping vbond.sdwan.local

Trả về IP 10.1.1.3 là thành công.

3. Trên các cEdge hoặc vEdge tại DC:

Bạn cũng cần thiết bị SD-WAN **trong nội bộ** sử dụng DNS nội bộ.

```
vpn 0
 dns 10.1.1.10   ! IP của DNS nội bộ
```

Hoặc trong cấu hình hệ thống:

```
system
 host-name dc-cedge
 site-id 100
 system-ip 10.0.0.1
 organization-name "vnpro"
 vbond vbond.sdwan.local local
```

→ Khi resolve vbond.sdwan.local, thiết bị sẽ dùng IP 10.1.1.3 → đi nội bộ, không cần NAT.

☑ Ưu điểm

- **Không phụ thuộc vào Hairpin NAT** hoặc điều chỉnh ACL phức tạp.
- Giữ đúng mô hình SD-WAN với khả năng mở rộng tốt.
- Tách biệt **remote site (xài DNS public)** và **DC nội bộ (xài DNS nội)** một cách sạch sẽ.

🔁 Nếu dùng chung DNS nội và DNS public?

Bạn có thể dùng:

- **Split Horizon DNS**: Cùng 1 hostname, DNS nội trả về IP private, DNS ngoài trả về IP public.
- Trên FortiGate hoặc Router có thể tạo **DNS override hoặc DNS forwarding** theo zone.
- Nếu dùng vManage cluster, hãy đảm bảo các node đều resolve hostname cluster IP về IP nội bộ.

- Với các hệ thống hybrid cloud, bạn có thể dùng DNS conditional forwarding: Nếu truy vấn *.sdwan.local → forward về DNS nội.