Prepared by Minh Nguyen – M365 Consultant

Phone / Zalo: 0983 003533

+++

CẨM NANG TOÀN DIỆN: MICROSOFT ENTRA PRIVATE ACCESS (ZTNA) - TỪ KIẾN TRÚC ĐẾN THỰC CHIẾN

PHẦN 1: TỔNG QUAN & TRIẾT LÝ KIẾN TRÚC BẢO MẬT HIỆN ĐẠI

1. Sự lụi tàn của VPN truyền thống và Sự trỗi dậy của ZTNA

Trong nhiều thập kỷ, VPN (Client-to-Site) là xương sống của truy cập từ xa. Tuy nhiên, VPN có một điểm yếu chí mạng: Nó cấp quyền truy cập ở cấp độ Mạng (Network-level). Khi một thiết bị vượt qua cổng VPN, nó được đặt vào bên trong mạng LAN, có thể ping và rà quét toàn bộ hệ thống. Đây là con đường lây lan ưa thích của Ransomware.

Microsoft Entra Private Access là giải pháp Zero Trust Network Access (ZTNA) ra đời để thay thế VPN. Triết lý của nó là **Micro-segmentation (Vi phân đoạn)**: Không cấp quyền truy cập vào mạng lưới, chỉ cấp quyền truy cập vào đích danh một Ứng dụng (App-level access). Người dùng chỉ nhìn thấy những gì họ được phép thấy, mọi hạ tầng cốt lõi khác đều bị ẩn giấu.

2. Kiến trúc cốt lõi: Hoạt động như thế nào?

Kiến trúc của Entra Private Access là mô hình **Cloud-routed (Định tuyến qua đám mây)** kết hợp với **Identity-Centric (Lấy danh tính làm trung tâm)**:

- **Không cần mở Port:** Không cần NAT hay mở bất kỳ Inbound Port nào trên Firewall doanh nghiệp. Mọi kết nối đều là Outbound từ một máy ảo nhỏ gọi là **Connector** đặt tại On-premise kết nối lên Microsoft Entra.
- **GSA Client & WFP Filter:** Trên máy tính người dùng cài phần mềm Global Secure Access (GSA) Client. Nó sử dụng màng lọc Windows Filtering Platform (WFP) can thiệp sâu vào nhân hệ điều hành để phân luồng giao thông.

- **Cơ chế Synthetic IP và NRPT:** Khi người dùng truy cập một tên miền nội bộ (FQDN), GSA Client sử dụng bảng Name Resolution Policy Table (NRPT) để đánh chặn yêu cầu DNS cục bộ. Thay vì trả về IP thật của Server nội bộ, nó trả về một **Synthetic IP (IP ảo, thường là dải 6.6.x.x)**. Máy Client sẽ gửi dữ liệu vào cái IP ảo này, GSA tóm lấy, bọc vào Tunnel (đường hầm) và gửi lên Cloud, sau đó Cloud đẩy về Connector để giải mã và đẩy vào Server thật.

3. Nghịch lý ZTNA: "Ping mù nhưng App vẫn chạy"

Một trong những cú sốc lớn nhất với các IT Admin là: Tại sao Ping FQDN thì báo `Request timed out` nhưng mở Web hoặc File Share thì vẫn chạy ầm ầm?

- **VPN hoạt động ở Layer 3 (Network):** Thông là thông tất cả, bao gồm giao thức ICMP (Ping).
- **ZTNA hoạt động ở Layer 4-7 (Transport & Application):** GSA Client chỉ mã hóa và chuyển tiếp TCP/UDP. Lệnh Ping dùng ICMP không có Port, GSA Client sẽ thẳng tay loại bỏ (Drop) ngay tại máy Client.
- **Giá trị bảo mật:** Ngăn chặn tuyệt đối các kỹ thuật dò quét mạng nội bộ (Network Scanning). Hacker không thể dùng công cụ để quét tìm Server. Để test mạng trong ZTNA, phải dùng PowerShell test TCP Port (ví dụ: `Test-NetConnection -Port 80`) hoặc truy cập thẳng giao thức đó.

PHẦN 2: LỘ TRÌNH TRIỂN KHAI VÀ QUY HOẠCH CẤU HÌNH

1. Giai đoạn 1: Quick Access (Lift and Shift)

Quick Access là ứng dụng mặc định giúp doanh nghiệp bê nguyên xi mô hình VPN cũ lên Cloud. Bạn khai báo một dải mạng lớn (ví dụ: `192.168.11.0/24`).

- **Ưu điểm:** Nhanh, dễ cấu hình, gom toàn bộ Traffic vào để theo dõi (Discovery) xem người dùng đang thực sự truy cập vào những Server nào.
- **Nhược điểm:** Tiềm ẩn rủi ro về vòng lặp định tuyến (Routing Loop) nếu không thiết kế kỹ, đi ngược lại triết lý Zero Trust.

2. Giai đoạn 2: Micro-segmentation qua Enterprise Applications (Best Practice)

Sau một thời gian dùng Quick Access, Best Practice là xóa bỏ nó và chuyển sang cấp quyền theo từng Enterprise App độc lập.

- Tạo từng App riêng lẻ (Ví dụ: `Ketoan-FileServer`, `HR-WebApp`).
- Khai báo đích danh IP tĩnh (`/32`) hoặc FQDN của đúng Server đó.
- Gán quyền (Assign) cho đúng nhóm User cần thiết.
- **Kết quả:** Bảo mật tuyệt đối. Nếu dính mã độc, nó chỉ bị giới hạn ở đúng Port và IP của App đó, không thể lây lan sang các Server hạ tầng (DNS/AD) khác.

3. Sức mạnh của Private DNS

Không bao giờ dùng IP để truy cập, hãy ép người dùng sử dụng **FQDN (Tên miền)**. Việc khai báo Private DNS (Ví dụ: `*.internal.company.com`) trên Entra Portal sẽ giúp GSA Client tự động cấy luật NRPT xuống máy khách, đảm bảo trải nghiệm người dùng liền mạch bất kể họ đang ở nhà hay ở công ty.

PHẦN 3: GIẢI PHẪU CÁC "BÃI MÌN" KIẾN TRÚC VÀ CÁCH KHẮC PHỤC (TROUBLESHOOTING)

Trong quá trình triển khai, GSA Client có thể gặp những lỗi kiến trúc rất sâu ở tầng Kernel. Dưới đây là các "bệnh kinh điển" và giải pháp xử lý.

1. Lỗi Vòng lặp định tuyến DNS (Ouroboros DNS Loop)

Hiện tượng: GSA Client liên tục báo Disconnected, mạng LAN chập chờn, phân giải tên miền chết cứng dù ping IP công cộng (8.8.8.8) vẫn sống. **Nguyên nhân:** Khi bạn cấu hình Quick Access gom nguyên dải mạng (VD: `192.168.11.0/24`), dải này vô tình **bao trùm luôn cả IP của máy chủ DNS nội bộ kiêm Connector** (VD: `192.168.11.1`). Khi máy Client hỏi DNS đến con `.1`, GSA tóm lấy gói tin Port 53, ném lên Cloud, Cloud ném về Connector `.1`, Connector lại đi hỏi chính nó -> Vòng lặp vô tận gây sập hệ thống.

Giải pháp (Kỹ thuật đục lỗ Port 53):

- Trong cấu hình Network Segments của Quick Access, thay vì để Port trống (bắt tất cả), hãy tạo 2 Rule để ngoại trừ Port 53:
    - Rule 1: Destination `192.168.11.0/24` | Ports: **`1-52`** | TCP, UDP

    - Rule 2: Destination `192.168.11.0/24` | Ports: **`54-65535`** | TCP, UDP
- Hệ quả: GSA sẽ bỏ qua luồng DNS. Gói tin DNS đi thẳng qua Card mạng vật lý, giải quyết triệt để vòng lặp.

2. Intelligent Local Access (ILA) và Hiện tượng Tromboning

Hiện tượng: Người dùng mang máy tính lên văn phòng cắm cáp LAN, truy cập Server nội bộ, nhưng thấy chậm. Mở log Traffic lên thấy gói tin vẫn bị gán nhãn `Action: Tunnel`. Đáng lẽ ở LAN thì gói tin phải đi thẳng qua Switch (Bypass). Hiện tượng này gọi là Tromboning (Mạng chạy vòng lên mây rồi vòng về). **Nguyên nhân:** Tính năng ILA (Tự động nhận diện mạng nội bộ) bị thất bại trong bài test ngầm. Mỗi vài phút, ILA gửi truy vấn DNS hỏi tên miền gốc (VD: `company.local`) đến DNS Server nội bộ. Nếu DNS Server trả về lỗi hoặc trả về IP không khớp dải mạng đã cấu hình trên Portal, ILA kết luận: "Đang ở Internet" và ép vào Tunnel.

Giải pháp:

- Truy cập máy chủ DNS nội bộ.
- Tạo một **Record A** cho khoảng trống (Tên miền gốc / Root Domain), trỏ về IP của máy chủ DNS.
- Khởi động lại GSA Client, kiểm tra log `Advanced Diagnostics -> Forwarding Profile -> Private network definitions`. Nếu thấy cập nhật đủ thông số, ILA sẽ hoạt động lại, luồng mạng sẽ hiển thị `Action: Local Access` (Bypass Tunnel hoàn toàn).

3. Bí ẩn tàng hình của File Share (SMB - PID 4)

Hiện tượng: Người dùng truy cập File Share (Port 445) thành công, nhưng mở Traffic Log của GSA Client không hề thấy bất kỳ bản ghi nào. **Nguyên nhân:** Giao thức SMB trong Windows không chạy bởi ứng dụng bình thường (User-Mode), mà chạy bởi Trình điều khiển Kernel (`mrxsmb.sys`) dưới quyền tiến trình hệ thống tối cao: **System (Process ID 4)**. Để tránh gây màn hình xanh (BSOD) hoặc phá vỡ Kerberos, Microsoft thiết kế GSA **tuyệt đối bypass mọi luồng giao thông từ PID 4 khi ở mạng LAN**. Do đó, gói tin đi thẳng mạng vật lý và không bị ghi log. (Tuy nhiên, nếu ở ngoài Internet, cơ chế Fallback sẽ bắt gói tin này và đẩy vào hầm, lúc đó log sẽ xuất hiện tên Process là `<unknown>`).

4. Xung đột xác thực NLA của RDP trong mạng LAN

Hiện tượng: Ở ngoài Internet thì RDP vào Server nội bộ bình thường. Về văn phòng (trong LAN), RDP bằng FQDN lại bị báo lỗi không thể kết nối. **Nguyên nhân:** Khi ở LAN, ILA kích hoạt, GSA đóng vai trò làm **Local Proxy**. Khi RDP (yêu cầu Network Level Authentication -

NLA cao ngặt nghèo bằng Kerberos) đi qua Local Proxy bằng IP ảo, NLA phát hiện sự bất thường (SPN Mismatch) và tự động chặt đứt kết nối để chống tấn công Man-in-the-Middle. **Best Practice:**

- **Ngoại mạng (Internet):** Luôn dùng **FQDN** để GSA bắt vào Tunnel.
- **Nội mạng (LAN):** Yêu cầu người dùng RDP bằng **IP thật** hoặc **Tên ngắn (Shortname)**. GSA sẽ không nhận diện được định dạng này, bỏ qua đánh chặn, giúp RDP và Kerberos hoạt động mượt mà.

PHẦN 4: BẢO MẬT NÂNG CAO VÀ CONDITIONAL ACCESS

ZTNA không chỉ là kết nối, nó là Định danh (Identity). Mọi kết nối đều phải đi qua bộ lọc Conditional Access (CA) của Entra ID.

1. Vé VIP (PRT) và Trải nghiệm Xác thực Ngầm (SSO)

Nhiều Admin ngạc nhiên khi cấu hình CA Policy yêu cầu MFA (Xác thực đa yếu tố) khi mở RDP, nhưng mở RDP lại vào thẳng mà không hiện popup MFA.

- **Bản chất:** GSA Client giao tiếp ngầm với Cloud. Nó lấy **Primary Refresh Token (PRT)** - chiếc vé VIP được cấp khi người dùng đăng nhập Windows/M365 - gửi lên Entra ID.
- Nếu trong PRT đã có sẵn Claim `MFA = Satisfied` (đã từng MFA trong ngày), Entra ID sẽ đánh giá là Pass và cấp quyền Silent Authentication (Xác thực ngầm). Điều này mang lại trải nghiệm SSO tuyệt vời.

2. Ép buộc MFA hiển thị (Force Re-authentication)

Nếu quy định công ty bắt buộc hễ mở Server Kế toán là phải rút điện thoại ra bấm MFA:

- Vào Conditional Access Policy > mục **Session**.
- Bật **Sign-in frequency** và cấu hình **Every time** (hoặc 1 Hours).
- Lúc này, PRT cũ sẽ bị từ chối, GSA Client sẽ gọi OS Broker của Windows bật Popup ép người dùng nhập MFA.

3. Sự bất đồng trình duyệt: Edge vs. Chrome

Khi set CA Policy *Sign-in frequency: Every time*:

- **Microsoft Edge:** Tuân thủ 100%, hiện popup MFA. Lý do: Edge tích hợp sâu với Windows Web Account Manager (WAM).
- **Google Chrome:** Vượt rào vào thẳng không cần MFA. Lý do: Chrome chạy Sandbox, có Cookie Jar riêng, đẩy Session Cookie cũ lên Cloud và Cloud chấp nhận vì Chrome không giao tiếp được với WAM để gửi cờ bắt buộc xác thực lại.
- **Giải pháp (Bắt buộc cho Doanh nghiệp):** Yêu cầu cài đặt Extension **"Microsoft Single Sign On" (Windows 10 Accounts)** trên tất cả trình duyệt Chrome của công ty. Extension này làm cầu nối giữa Chrome và Windows WAM, ép Chrome tuân thủ tuyệt đối lệnh hiển thị MFA của Conditional Access.

PHẦN 5: LỜI KHUYÊN & BEST PRACTICES THỰC CHIẾN

1. **Dọn sạch VPN cũ (WFP Conflict):** Không bao giờ cài đặt GSA Client trên máy tính đang chạy các phần mềm VPN truyền thống (OpenVPN, WireGuard, FortiClient...). Hai bên sẽ tranh giành màng lọc WFP ở nhân hệ điều hành, dẫn tới xung đột định tuyến và màn hình xanh (BSOD). Đã dùng ZTNA thì phải gỡ sạch VPN cũ.
2. **Điểm yếu của giao thức UDP:** Entra Private Access hiện tại là bậc thầy về TCP, nhưng hỗ trợ UDP còn hạn chế (chủ yếu là Port 53 DNS và 443 QUIC). Nếu doanh nghiệp có các ứng dụng Legacy dùng UDP đặc thù (Voice IP chuyên biệt, truyền phát Video), cần cân nhắc kỹ vì GSA có thể Drop các luồng UDP này.
3. **Bật Source IP Restoration:** Mặc định IIS/Web Server nội bộ sẽ thấy mọi truy cập đều đến từ IP của máy Connector (VD: `192.168.11.1`), gây khó khăn cho việc Audit. Hãy vào *Global Secure Access -> Global settings -> Session management*, bật **Source IP restoration**. GSA sẽ bọc IP WAN thật của người dùng vào biến `X-Forwarded-For`. Sau đó, thêm trường này vào log của IIS, Server sẽ đọc được chính xác IP gốc từ xa của người dùng.
4. **Tận dụng Web Content Filtering:** Nếu đã mua bộ License, hãy bật luôn Internet Access Profile. Sử dụng tính năng Secure Web Gateway để chặn các trang Web đen, mạng xã hội, hoặc mã độc. Màng lọc này sẽ chặn ở mức Endpoint bất kể nhân viên đang ngồi ở văn phòng hay quán cafe, thay thế hoàn toàn nhu cầu mua các bộ Firewall vật lý đắt tiền để lọc Web.

Tài liệu được đúc kết từ quá trình nghiên cứu, triển khai thực chiến và chẩn đoán sâu ở cấp độ hệ thống. Mục tiêu cốt lõi của Microsoft Entra Private Access không chỉ là thông mạng, mà là thay đổi hoàn toàn tư duy thiết kế chu vi bảo mật: "Xác thực mọi thứ, Không tin tưởng bất kỳ ai" (Never Trust, Always Verify).