追蹤Linux封包流程之排列組合

簡世昕 林盈達

國立交通大學資訊科學系

300 新竹市大學路1001號

Tel: 03-5712121 ext. 56667

E-MAIL：{shchien,ydlin}@cis.nctu.edu.tw

摘要

在單一閘道器中整合各種功能，來妥善的處理流入封包，是目前最受注意的議題。在本文中我們將討論封包在防火牆(Firewall)、網路位址轉換(NAT)、路由(Routing)、虛擬私有網路(VPN)、入侵偵測系統(IDS)、內容過濾(Content Filter)、頻寬管理(Bandwidth Management)等元件之間的流程追蹤和運作關係，並就元件的先後排列順序來做分析，其中考量的依據集中在安全性、相容性、管理難易等方面。當IPsec封包進入VPN解封裝之前，需由Firewall提供屏障，降低VPN元件被外部攻擊的機會。若封包遇到NAT和VPN的相容問題，解決的最好辦法是先通過NAT再作VPN，避免IPSec封包的表頭被修改。利用Firewall和IDS之間的互動關係，我們可以減少更多安全上的漏洞。最後，我們將總結出一張元件先後順序關係表，可以看出兩兩元件之間有那些相互的關係。

關鍵字：Integrated Gateway, Firewall, NAT, Routing, VPN, IDS, Content Filter, Bandwidth Management, Compatibility

1. 簡介

封包在不同廠商所研發的閘道器下，其流程並不盡然相同，主要是為了使用者操作方便和一些安全性的考慮，減少系統管理者撰寫防火牆規則出錯的機會，

以達到用最簡單的管理方式作最大的保護功能。由於在封包的流程中，前一個元件有時會提供處理後的封包資訊給下一個元件作判斷，因此，我們有必要去觀察封包在這些元件之間的流程關係，來了解兩兩元件之間是如何工作和互動。另外，也方便將來有新元件的加入時，可以正確的規畫應該放置在那個掛鉤點上，來避免相容性問題和達到最佳化的效能。

究竟封包是如何從網路上傳遞到本地主機上呢?首先，我們先了解在主機上是用怎麼樣的資料結構來儲存封包的，當網路卡驅動程式從網路卡上抓取到封包後，它會配置一塊sk_buff記憶體用以儲存封包內容，以後封包在系統中就是以sk_buff的形式表現，其資料結構定義在<linux/skbuff.h>中。sk_buff主要目的是為了提供通訊協定層以標準的函式或方法對應用程式的封包進行處理，其中各通訊協定主要利用如表1所示的四個指標和二個長度資訊來操作這塊封包緩衝區，使得我們只需要依靠指標的轉換，就可以很有效率且安全地對封包資料的標頭和結尾資訊作增加和移除，而不用為了在各層通訊協定間傳遞資訊，而去不斷的複製資料，浪費不必要的空間和時間。

欄位	意義
head	指向sk_buff的起點
data	指向真正資料的起點
tail	指向真正資料的終點
end	指向sk_buff的終點
len	真正資料的長度
truesize	sk_buff的總長度

表1:sk_buff的重要欄位

在下一節中，我們將介紹封包接收的基本處理，來觀察封包是如何從網路上傳送到應用程式去。第三節則介紹封包在這些元件中的流程分析。第四節則深入討論在上一節的元件相容問題。最後是我們的結論。

2. 封包接收的基本處理

如圖1所示，當網路卡收到封包後，它會發出中斷來通知核心，於是核心會

去呼叫適當的中斷處理常式(Interrupt Service Routine)來處理。此時，這個常式會去配置一塊sk_buff，並將封包放入此塊記憶體中，通常也會使用DMA(Direct Memory Access,直接存取記憶體)不經過CPU直接傳送到記憶體中。接著netif_rx()函式所作的像是單一收集點，將從各個不同的網路卡驅動程式中所接收到的封包，統一放入佇列中給更上層的通訊協定處理。接著系統會檢查是否有軟中斷(softirq)發生，若有就呼叫net_rx_action()，將封包從佇列中取出來。接著必須通過ptype_all和ptype_base這二個串列，來比對packet_type的欄位，若有多個通訊協定與之匹配，則sk_buff將被複製多份給相對應的通訊協定常式分別處理。當封包傳遞到IP協定層，會對IP封包作完整性的檢查，包括IP檢驗值、IP表頭欄位等，接著會處理IP路由、資料碎片的問題。當封包傳遞到第四層時，若為TCP封包則會對TCP表頭作完整檢驗，查看是屬於那個socket後，封包會依據不同的socket狀態而有不同的處理方法，若是UDP封包則處理方式較簡單，先對標頭作檢驗，選取接收的socket，再將UDP封包放入接收socket的緩衝區內。接著會在Socket層呼叫data_ready()代表完成第四層的動作，這時呼叫接收封包的系統呼叫(System call)的應用程式會被喚醒，完成封包的接收動作。

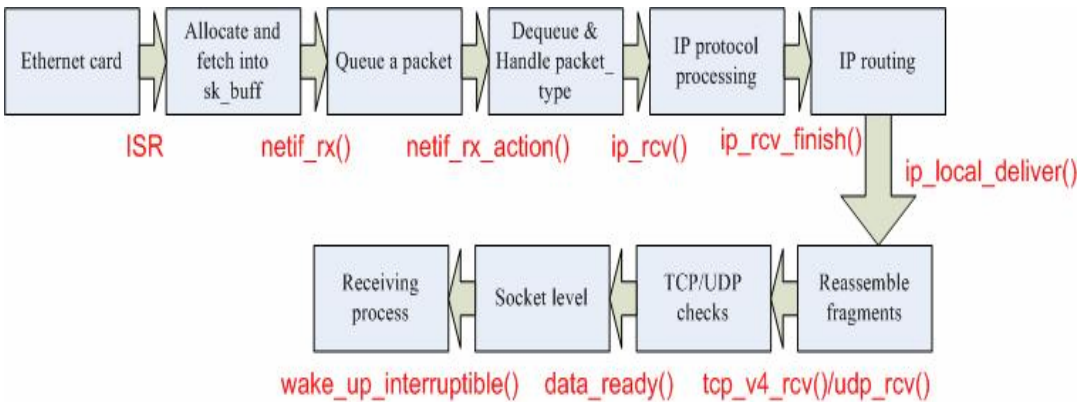

圖1:內部封包接收的基本處理

3．封包在元件間的流程分析

前一節從系統內部接收封包的角度來追蹤流程，了解到封包的基本接收處理

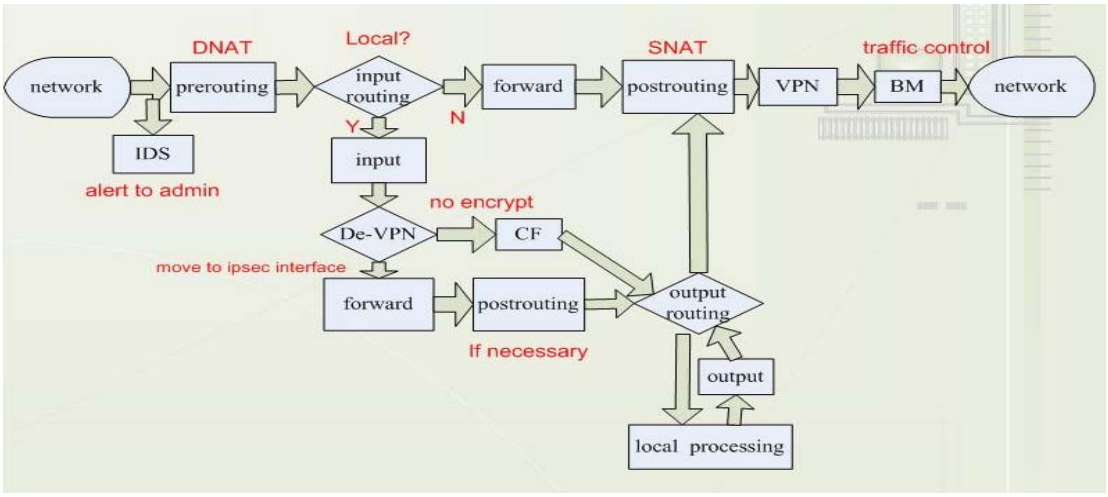

圖2:封包在元件之間的運作流程圖

方式，本節將繼續深入介紹封包在不同元件間的運作流程和處理方法。如圖2所示，這是一個典型封包處理流程的例子，我們可以分成以下步驟:

1. 當封包從某一網路進入到閘道器後，入侵偵測系統會擷取所有進入的封包，以特徵模式(signatures)來辨別是否為蓄意的入侵或有攻擊的意圖。在比對出符合的攻擊模式後，反應模組會立即通知系統管理員或切斷任何可疑連線。
2. 進入Prerouting後，NAT處理函式會去辨別目的位址是否需要改寫。
3. 接著會作輸入路由處理，檢查IP是否需要轉送出去，如果需要轉送封包，則進行Forward過濾，否則，送往Input過濾。在送往本機的封包中，可分成三種情況:

   情況1:如果發現為VPN封包，則將之解封裝後送往IPsec的介面，進行Forward和Postrouting的過濾和轉換。

   情況2:若非VPN封包，但卻是網站回應封包，則執行網頁內容過濾。

   情況3:若只是單純的本機封包，則跳過上述動作。
4. 由於VPN封包已被解開，得到封包真正的目的位址後，必須再作一次輸出路由，此時主機才會收到真正要傳送到閘道器的封包。
5. 在傳往Postrouting時，若有需要，則會進行來源位址的改寫，接著才會

作VPN的封裝。

6. 最後，封包要從閘道器出去時，會作頻寬管理，接著送往下一個網路去。

4．元件相容性

現在，我們看另外一個例子，如圖 3 所示，若將解封裝 VPN 的動作放置在 Input 過濾封包之前，並且將 VPN 封裝的動作放在改寫目的位址 Postrouting 之前的話，我們可發現一些議題是需要討論的，分別為 Firewall 和 VPN 的安全問題、VPN 和 NAT 的相容問題、IDS 和 Firewall 的管理難易問題等。以下將深入的介紹這三種不同的先後模式，在封包處理過程中所會造成的影響，並與原先的模式做比較，觀察兩者各具有什麼優缺點。

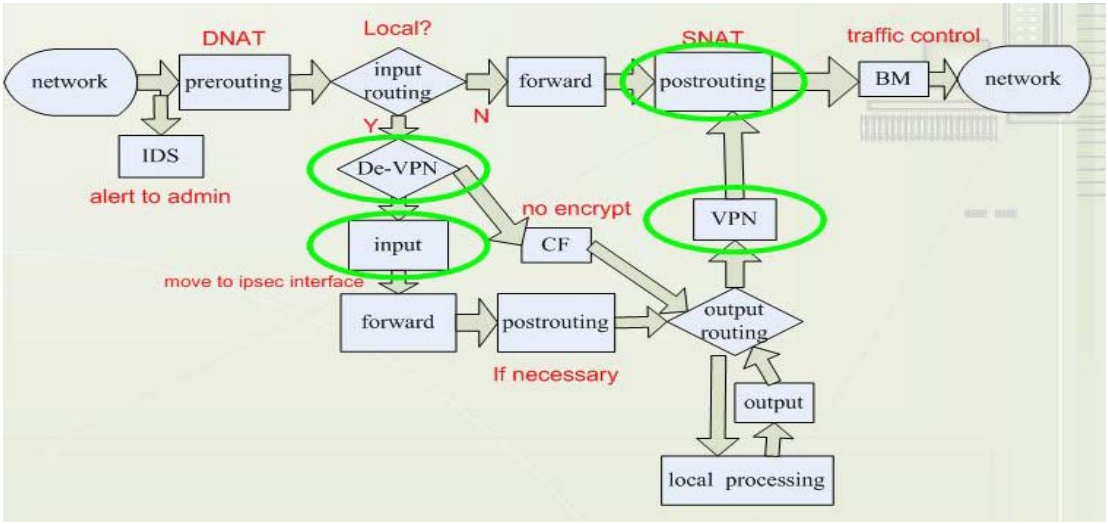

圖3: 另一個合法封包在元件之間的運作流程圖

4.1 防火牆和IPsec的安全問題

防火牆(firewall)主要是作封包控管、保護內部網路及阻止未經許可的連線進入。而 VPN 主要是利用 IPsec(IP Security)作加密技術，IPsec 結合了加密(Encryption)、認証(Authentication)、密鑰管理(Key Management)、數位檢定(Digital Certification)等安全標準，具有高度的保護能力，然而，VPN 和防火牆的前後關係該如何設置，才能確保提供安全可靠的 VPN 服務，卻又不傷害到防火牆的原

先規則?

- **情況 1:IPsec 在防火牆之前先行服務**

這樣的架構最主要的好處在於能將解封裝後的 IPsec 封包，在進入信任網域之前，就先行被防火牆所檢查。不過由於未受到防火牆保護的原故，IPsec 易遭受到阻斷式攻擊(Deny of Service,DoS)，不斷被要求 VPN 服務，使得所有 VPN 連線受到干擾，甚至斷線結束。若解封裝的封包未通過 IPsec 的認証，且 IPsec 無封包過濾功能的話，因為防火牆收到的都是已解除 IPsec 的封包，若 IPsec 無法提供防火牆的認証資訊,得必須再多作一次認証,造成一些不必要的額外負擔。

- **情況 2:防火牆在 IPsec 之前先行服務**

將防火牆放置在 IPsec 前隔絕和外部網路的連繫，來保護 IPsec 避免遭到阻斷式攻擊，同時減少防火牆額外的負擔。雖然減少了外在的威脅，卻有一些內部安全的隱憂。因為 IPsec 封包被封裝加密,防火牆必須允許通訊協定為 50、51(AH 和 ESP)的 IP 封包通過。此外,為了自動金鑰交換機制(Internet Key Exchange, IKE)也要通過監聽埠為 500 的 UDP 封包。使得防火牆無法得知 VPN 內部封包的真實來源和目的位址，造成封包控管產生漏洞，讓 VPN 使用者擁有進入內部網域存取特定資源的存取權限(Access Control)。主要的原因是因為 VPN 只提供資料在傳送時的安全保護，卻無法控管內部網路資源的存取。

4 .2 NAT和IPsec的相容問題

NAT 和 IPsec 本質上是不相容的兩種協議。IPsec 主要是用 AH(Authentication Header)和 ESP(Encapsulating Security Payload)組成來保護封包完整性。而 NAT 卻是更改封包中的目的或來源的位址和埠口，同時修正 TCP/UDP 表頭的檢驗值內容。所以當系統必須同時使用 NAT 與 IPsec 時，必須要考量一些問題。

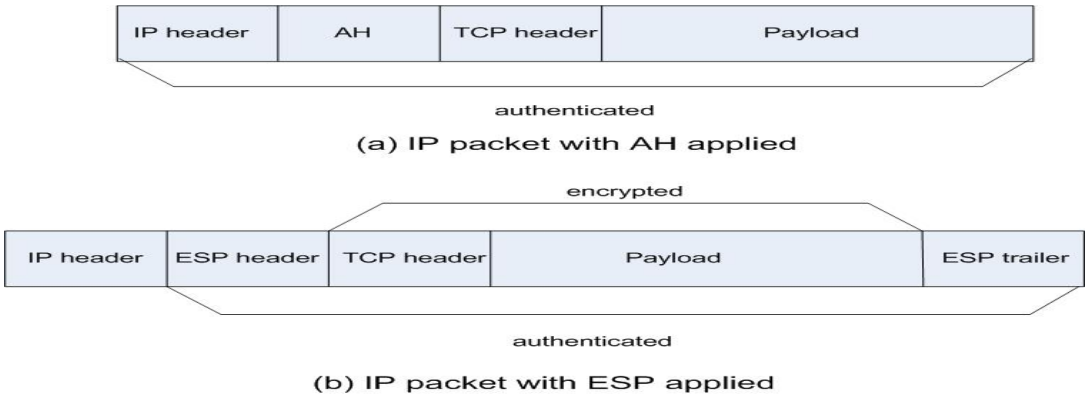

圖 4:AH 和 ESP 的檢驗方法

■ **情況 1:IPsec 穿越 NAT 的情況**

以下分別就 IPsec 的 AH 和 ESP 兩種安全協定在不同傳輸模式下(Transport mode /Tunnel mode)，穿越 NAT 所可能遇到的問題。

1.NAT 對 AH Transport/Tunnel Mode 的影響

如圖 4(a)所示，AH 會對整個 IP 封包產生一個雜湊值，包括 IP 表頭和資料區塊，接收方會利用此雜湊值來驗証封包傳送的過程中，是否有無修改變化，若認証失敗，則接收方會丟棄此封包。因此，若 NAT 改變來源位址或埠口，從而導致 AH 完整性認証失敗，顯然 AH 協議和 NAT 之間無法正常運作。

2.NAT 對 ESP Transport Mode 的影響

如圖 4(b)所示，ESP 類似 AH 的功能，也會對 IP 封包作資料完整性的檢驗，同時還結合了加密演算法。但與 AH 不同的是，ESP 並未把 IP 表頭包含在內，因此，不會有 AH 類似的問題產生。但是 ESP 會對資料區塊作加密保護，NAT 在轉換位址和埠口後，同步修改 TCP/UDP 的檢驗值，造成接收端無法通過 ESP 的認証後，如果 NAT 不進行檢驗值的更新，則導致 TCP/UDP 校驗錯誤。

3.NAT 對 ESP Tunnel Mode 的影響

因為 IPsec Tunnel Mode 是將原封包封裝到另一封包的資料區塊，再新加上一個 IP 表頭。因此 NAT 只會改到最外頭的 IP 位址，原來的 IP 表頭並未更動，所以可通過 TCP/UDP 的檢驗值檢查。

■ **情況 2:NAT 穿越 IPsec 的情況**

由於 NAT 會先作完轉換 IP 位址或埠口後，IPsec 之後的加密和資料完整性檢查動作皆不受影響，在此種模式下，可以正常傳送封包。缺點是 IPsec 無法得知被 NAT 轉換前的真正封包來源位址，因為位址已被轉換為公有的 IP 位址了。

■ **解決相容性方法**

上述情況中只有 NAT 穿越 IPsec 時，封包才能暢通無阻地通過。若是 IPsec 的封包穿越 NAT 的話，必須使用其它合用的方法來修正不相容的問題。

1． IPsec 封包不受 NAT 修改

利用修改 iptables 的規則，允許 IPsec 封包可以通過 Postrouting 表格而不作 NAT 的轉換，如此 IPsec 封包可不受影響。此種方法最大的好處是簡單容易，但該子網路卻無法繼續使用 NAT 的功能，除非是重新修改其規則。其 iptables 的規則如下: 除了要作 IPsec 的子網路位址之外，其它封包將會偽裝成閘道器位址。

```
Iptables –t nat –A POSTROUTING –o eth0 –s gw_addr –d \ !sub_addr –j MASQUERADE
```

2．封裝成 UDP 封包

此方法為在發送 IPsec 封包前，先封裝在 UDP 封包內，經過 NAT 轉換後，接收方收到封包後，會除去 IP 表頭並解封裝 UDP 封包，因此，可忽略 NAT 所做的任何改變對 IPsec 的影響。此種方法不需要對 IPsec 協定作更動，所以可適用在 IPsec 的所有協定(AH/ESP)和模式(Transport/Tunnel)。不過，由於需要對主機作修改，增加額外的位元組來處理 UDP 封裝。另外，整個 IPsec 的安全關聯(Security Association,SA)的協商時間也會明顯增加。

3．更改 IPsec 協定

由於AH和NAT兩者無法相容，因此我們可以利用ESP null[1]來替代AH作封包完整性的檢查，而不作IP表頭的檢查。另外，因為IPsec認証會對封包的TCP/UDP欄位作檢查，為了消除NAT的更改IP位址或埠口時，同時也會修改到檢驗值，在傳輸封包前，我們可以將傳輸層的檢驗值關閉。因此作封裝時不再需要額外的負

[1] ESP協定對於防篡改或造假的能力並不如AH，所以雖然ESP有類似驗證功能，卻不可完全取代 AH。

擔，但此方法只適用在NAT傳輸IPsec的ESP協定，而不支援AH。

4．使用 RSIP(Realm-Specific IP)協定

由 IETF 正在研究的 RSIP 協定，可用來替代 NAT 和 IPsec 相配合工作。RSIP 用來將公有 IP 位址租借給私有網域中的 RSIP 主機，RSIP 要求更動主機，使得可以將 RSIP 閘道器所給定的資訊綁在每一個封包傳送，用來和外面的公有網域作端對端連線，卻不會破壞其通訊協定的特徵，保存封包的完整性。不過由於該協定尚未成熟，而且除了要求每一部客戶主機支援 RSIP 外，還得有對應的 RSIP 閘道器，降低該協定的可用性。

4.3 防火牆和IDS的管理問題

入侵偵測系統(IDS)主要是監視網路上所有的封包，並以特徵模式(signatures)來辨別是否為蓄意的入侵或有攻擊的意圖。IDS 可因不同的策略而有不同的放置條件，若放在防火牆之外，可用來偵測外部的攻擊；若放在防火牆之內，則可用來保護內部網域，避免病毒的入侵。由於 IDS 大多採用 pass-by 方式來監聽網路上的封包流，限制了其本身的反制功能，而且防火牆的策略都是事先設置好的，無法動態設置策略，缺少對攻擊應變的靈活性，不能更有方法的保護網路安全，所以 IDS 和防火牆聯合互動的目的就是更有效地反制所發生的攻擊事件，進而保護整個網路。如圖 5 所示，當封包進入閘道器時，會複製一份封包給 IDS 作策略檢查，如果在比對特徵資料庫中發現是攻擊事件，那麼 IDS 會發給防火牆一個相對應的動態反制策略，防火牆就會依據該動態策略來進行過濾。然而，這種聯合互動會有潛在的危機，因為 IDS 面臨有漏報率和誤報率的問題，如果有心者利用偽裝的封包來讓 IDS 判斷錯誤，進而使得防火牆策略擋住合法位址封包進入。因此市面上也出現一種新趨勢，將兩者功能結合為一，可以檢測入侵，還可以即時防禦的 IDP(Intrusion Detection & Prevention)。

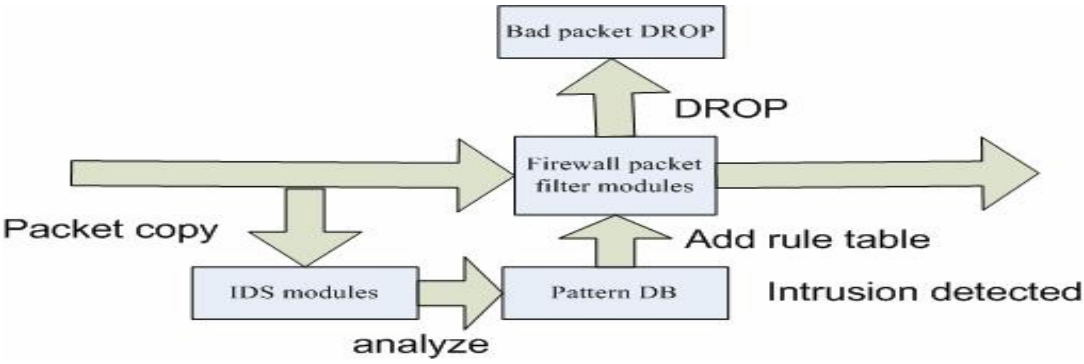

圖 5:IDS 和防火牆整合工作原理

4.4 元件間位置關係

如表 2 所示，斜線左邊為區域網路往廣域網路的情況，相當於由我們的電腦發送封包到網際網路的情況，右邊反之，在表格左邊縱排元件為前置元件，上面橫排元件為後置時的關係。在研究過各種不同的組合，我們歸納出一些元件的先後順序關係。

1. Firewall 和 NAT：當封包從區域網路送出時，要先經過防火牆過濾，避免遭到 NAT 作位址轉換，成為閘道器的 IP 位址；若封包從廣域網路送來，得先經過 NAT，得到真正的目的地位址後，才能對 IP 封包作過濾動作。
2. NAT 和路由：當封包從廣域網路接收，得先經過 NAT 轉換，得到真正的目的地位址後，再作路由轉送到下一站去。
3. VPN 和路由：同樣的，當封包從廣域網路接收下來，經過解封裝後，才能得到真正的原來封包內容，此時路由才有意義。
4. IDS：入侵偵測系統在整個封包流程中，扮演的是一個監視者的角色，可以因不同的策略而放置在流程中不同的位置。封包抵達時，再複製一份封包過去，比對後若發現為攻擊的情況，再警告給系統管理員知道。
5. CF：由於網址過濾和網頁內容過濾必須看到封包內容，得到達應用層才能作處理判斷。因此，當封包從區域網路送出去時，會先比對禁止網址的串列，再往下處理 TCP 封包，像是 NAT、路由、VPN 等等。若封包從廣域網路收進來，路徑則是剛好相對，封包會處理完位址轉換、解封裝、路由等等，到

達應用層後，才處理網站的回應封包，檢查檔案內容格式，若能通過比對的，才往後傳遞給使用者。

rear

front	Firewall	NAT	Routing	VPN	IDS	CF	BM
Firewall		L/R	D/D	L/L	R/R	M/M	D/D
NAT	R/L		D/M	L/L	D/D	I/M	D/L
Routing	D/D	D/I		D/I	D/D	I/M	D/D
VPN	R/R	R/R	D/M		D/D	I/M	D/D
IDS	L/L	D/D	D/D	D/D		D/D	D/D
CF	I/I	M/I	M/I	M/I	D/D		D/D
BM	D/D	D/R	D/D	D/D	D/D	D/D	

表格內容：Lan to Wan/Wan to Lan

文字說明：M:Must I:Impossible L:Likely R:Rarely D:Don't care

表2:元件間關係圖

5. 結論

由以上的介紹，我們得知封包內部的流程和在閘道器中各元件間的流程比較，及相互使用時的一些替換上的可能性。其中也發現了一些固定的規則，如當VPN處理流入封包時，必須由防火牆架設於較外層以提供保護，避免遭受到服務阻斷攻擊，造成無法使用VPN連線。解封裝後的封包，可得到原來的封包內容，但仍需再經過一次封包過濾，將禁止進入的來源位址阻擋在外，保護區域網路安全。在使用IPsec和NAT的時後，建議先使用NAT作位址轉換後，再通過IPsec的方法，避免IPsec的封包因為修改IP位址欄位，造成驗証失敗，而必須作額外的處理動作。而IDS和防火牆的互動，可以提供更效率、更安全的防護措施，來動態加入過濾規則，使網路隱患降到最低傷害。

雖然目前各界都在研究和解決元件在整合時所遇到的難題，卻沒有一個確切的方法出現，正因為最佳設計的方法皆是因地置宜，主要考慮在於是否有架設需求、執行效率、管理機制、封包控管、重複使用性等去評估，而有不同的設計和

考量，若能修改整合套件，也是不錯的方法。另外，在安全性問題上，沒有任何技術是保証絕對安全的，唯有網路管理者根據自身的經驗和使用者的習慣來做調整，才能真正的保障網路安全，避免真正合法的使用者不得其門而入。

6. 參考資料

[1] 蔡孟甫、曹世強、林盈達，「NetBSD核心網路安全模組: IPFilter及IPSec」，http://speed.cis.nctu.edu.tw/~ydlin/miscpub/hands-on_security_kernel_modules.pdf

[2] RFC3715 ”IPsec-Network Address Translation Compatibility Requirements” B. Aboba, W. Dixon, Mar. 2004

[3] IPtables, http://www.iptables.org/

[4] Dansguardian,http://dansguardian.org/

[5] 「IPsec VPN 的難題：Firewall 與 NAT 的配置」，http://www.iii.org.tw/ncl/document/IPSecVPN.htm

[6] Linux IP Masquerade HOWTO, http://en.tldp.org/HOWTO/IP-Masquerade-HOWTO/