Nftables (II): NAT

Per a què els ordinadors puguin establir comunicacions entre ells han d'estar identificats i això es fa amb l'adreçament: a cada ordinador s'hi assigna una adreça IP que l'identifica de manera única en la xarxa, No obstant, a causa de la limitació de l'espai d'adreces del protocol IPv4 no és possible que a Internet tots els ordinadors tinguin una única adreça. Per solucionar aquest problema, el que es fa és disposar a voluntat d'una quantitat (generalment gran) d'adreces privades per als diferents ordinadors de la nostra xarxa local (assignades per nosaltres ja sigui estàtica o dinàmicament) però utilitzar només una sola adreça pública (assignada per l'ISP que tinguem contractat) per comunicar tots aquests ordinadors amb Internet. Per a què això funcioni cal tenir a la nostra xarxa local un sistema que, quan un ordinador de la nostra xarxa vulgui comunicar-se amb Internet, tradueixi la seva adreça privada concreta a la única adreça pública existent. Aquest sistema es diu NAT ("Network Adress Translation") i l'implementen els encaminadors, que són els dispositius que fan de "frontera" entre la xarxa local i la pública.

> NOTA:Cal tenir en compte que NAT és una sol.lució de compromís perquè presenta una sèrie de problemes. Per exemple, el retard, ja que en efectuar la traducció de cada adreça en els encapçalaments dels paquets es produeix un alentiment. A més per cada paquet s'ha de decidir si s'efectua la traducció o no, que també alenteix el procés. Tampoc es pot fer un seguiment dels paquets, ja que poden passar per diversos encaminadors executant el NAT i això fa que siguin més difícils de seguir, múltiples salts de NAT i pot provocar que algunes aplicacions no funcionin perquè s'amaguen les adreces de destinació i d'origen. Tots aquests problemes se sol.lucionen amb l'IPv6.

Recordem els rangs d'IPs privades que es poden fer servir (tota IP que no estigui dins d'aquests rangs, és pública i, per tant, cau fora de la nostra administració): de classe A tenim la xarxa sencera 10.0.0.0/8 ; de classe B tenim el rang de la IP 172.16.0.0/16 fins a la 172.31.0.0/16 i de classe C tenim el rang de la IP 192.168.0.0/24 fins a la 192.168.254.0/24

Teoria: SNAT

El cas descrit anteriorment, és a dir, el canvi en l'adreça d'origen dels paquets de l'adreça privada d'un node de la LAN per la pública del sistema enrutador-tallafocs quan el trànsit és originat des de la xarxa local cap a l'exterior és el tipus de NAT més habitual amb diferència i és el que se sol entendre quan es parla de NAT genèricament però més específicament hauríem de parlar de **SNAT**, de "Source NAT"). SNAT és un mecanisme que permet compartir una direcció IP pública per molts equips que tenen cadascú una IP privada diferent. D'aquesta manera, es permet l'accés a "Internet" a tots aquests equips minimitzant el problema de l'escassetat de direccions IPv4 (públiques).

En el SNAT el destí del paquet "veurà" que l'origen del paquet és el "router" i, per tant, serà a ell a qui li reenviarà la resposta (més en concret, a un port determinat del "router"). Quan aquesta resposta arribi al "router", aquest consultarà una taula (la "taula NAT") que tindrà en memòria on s'associa la IP d'origen real (i també el port d'origen) amb la seva IP (i el port que ha utilitzat el router per connectar amb l'exterior) i reenviarà convenientment aquesta resposta a l'origen real.

A la taula NAT hi ha quatre dades: IP origen, port origen, IP router, port router. D'aquesta manera, el router és capaç de reenviar una resposta rebuda per un determinat port seu a un determinat port de l'origen. Per tant, és perfectament possible que cada origen pugui tenir diferents connexions en paral.lel (cadascuna creades en un port diferent) perquè el router saps quina és quina en tot moment. Normalment, el port que fa servir el router per accedir a l'exterior és el mateix que el port utilitzat per l'origen, però si hi hagués més d'un origen utilitzant el mateix port, el router llavors automàticament en farà servir un altre (i,per tant, també canviarà al paquet el port d'origen).

Resumint,

1.-Un equip d'una LAN amb IP privada (suposarem 192.168.3.14) vol accedir a una pàgina web (port 80 TCP) com "www.lerele.net". Realitza la consulta DNS i obté que l'equip que allotja la pàgina té la direcció 76.74.254.126

2.-Consulta la seva taula d'encaminament i com l'equip destí no està a la seva xarxa, envia la petició de la pàgina a la seva porta d'enllaç definida (que suposarem que té la IP 192.168.3.254)

3.-El gateway, en comprovar que la direcció IP de destí no és la seva, l'enviarà a la seva pròpia porta d'enllaç, que ja serà una direcció IP pública. Abans de què el paquet surti per la seva tarja de xarxa externa, però, el gateway canvia la direcció IP d'origen (192.168.3.14) per la seva direcció IP pública (que suposarem que és 80.58.1.14) i es guarda la petició en una taula en memòria (anotant també el port d'origen, que suposarem que és el 50158 TCP).

4.-El paquet viatja per Internet saltant de router a router (eventualment fent també els seus propis processos NAT si fossin necessaris, però en aquest exemple, per simplicitat, això ho obviarem) fins que arriba al seu destí

5.-L'equip 76.74.254.126 rep la petició des de la direcció 80.58.1.14 i la respon. Per tant, el paquet de resposta tindrà com direcció IP d'origen 76.74.254.126, direcció IP de destí 80.58.1.14, port d'origen 80 TCP i port de destí 50158 TCP

6.-Aquesta resposta arriba a la tarja externa del gateway de la nostra LAN, que consulta la seva taula en memòria i comprova (gràcies al port origen allà registrat) que correspon amb una petició realitzada des de l'equip 192.168.3.14, així que modifica la direcció IP de destí del paquet per aquesta IP i l'envia directament.

Hi ha un tipus de "Source NAT" específic anomenat "**IP masquerading**", el qual és el que passa quan la direcció IP pública que sustitueix a la IP d'origen és dinàmica (el cas més habitual en connexions a Internet domèstiques). En el "Source NAT" clàssic la IP pública és estàtica (uns circumstància molt rara avui dia)

Teoria: DNAT

Hi ha un altre tipus de NAT que s'implementa quan el trànsit és originat des de l'exterior cap a un servidor que tinguem funcionant a la xarxa local; aquest tipus de NAT -no tan habitual- se sol anomenar **DNAT** (de "Destination NAT" o també "NAT invers"), i consisteix en canviar la IP de destí del paquet (que sol ser la IP pública de la tarja externa del nostre gateway...si no fos aquesta el propi "router" ja descartaria el paquet en qüestió) per una IP privada d'alguna màquina de la nostra LAN.

Cal tenir en compte que aquest tipus de NAT es realitza només quan el gateway detecta que aquest paquet no es correspon amb cap connexió iniciada des de la nostra LAN (és a dir, ha de ser un equip extern qui inicii la connexió) i té prèviament definida la regla DNAT adequada (si no la té, el paquet es descarta).

Aquest tipus de NAT, el qual se sol anomenar també "Port forwarding", s'utilitza quan volem que algun servidor funcionant en una màquina de la nostra LAN sigui accessible des de l'exterior. Els passos concrets són aquests:

1.-Un equip qualsevol d'Internet amb direcció IP pública 150.212.23.6 vol connectar-se per SSH (port 22 TCP) a l'equip "miservidor.midominio.com". Realitza una consulta DNS i obté com a resposta que aquest equip té la direcció IP 85.136.14.7

2.-Estableix la connexió (suposem que el port d'origen que utilitza és 23014 TCP) amb l'equip 85.136.14.7, que resulta ser un dispositiu NAT que no té cap servei SSH escoltant al port 22 TCP però que té una regla DNAT que fa que tot el que li arribi a aquest port ho reenviarà a un equip de la seva xarxa local (suposarem que és el 10.0.0.2). Per tant, canvia la direcció IP de destí (85.136.14.7) per la 10.0.0.2 i registra aquesta associació a una taula en memòria

3.-A l'equip 10.0.0.2 li arriba una sol.licitud al port 22 TCP. La resposta consegüent del servidor SSH tindrà llavors les següents característiques: IP d'origen 10.0.0.2, port d'origen 22 TCP, IP de destí 150.212.23.6 i port de destí 23014 TCP

4.-El dispositiu NAT canvia la direcció IP d'origen per la seva direcció IP pública (85.136.14.7) i el paquet de resposta arriba al seu destí.

Hi ha un tipus de "Destination NAT" específic anoemant "PAT" ("Port Address Translation") o simplement "**redirect**", el qual és un mecanisme que només canvia els ports (normalment el de destí) però no les IPs. El PAT és útil quan un servidor ubicat al mateix sistema que Nftables no escolta al seu port estàndard però no volem fer que els client hagin d'especificar manualment el port no estàndard al que han d'anar quan connectin a aquest servidor.

Teoria: la taula NAT

Ja s'implementi SNAT o DNAT, en l'encaminador es manté una taula amb les traduccions que va fent de manera que pugui portar el control de quins paquets corresponen a cada connexió individual encara que totes elles comparteixin la mateixa IP pública.

Si a més de produir canvis en les adreces IP el router també canvia els números de ports (en el cas del SNAT, el port d'origen i en el cas del DNAT el port de destí -en el que popularment s'anomena "port-forwarding" o "redirecció de ports"-) estarem parlant d'un NAT de fa servir la tècnica **PAT** ("Port Address Translation") per realitar la seva funcionalitat. En aquest tipus de S/DNAT (el més habitual amb diferència) l'encaminador fa servir de manera dinàmica números elevats de ports per identificar de manera única la connexió d'un ordinador local a l'exterior. La idea és que s'ha de discriminar a quin ordinador de la xarxa local corresponen les dades que arriben o provenen d'una única adreça pública, ja que si només es tradueixen adreces no es pot diferenciar (perquè tenint una sola adreça pública totes les adreces privades s'acaben traduint a aquesta).

El mecanisme PAT concret implementat en un NAT de tipus SNAT és el següent: l'encaminador registra a la seva taula NAT la IP privada de l'ordinador de la nostra LAN que vol "sortir a Internet" juntament amb el número de port client que hagi obert per iniciar la connexió i vincula aquesta parella de valors, respectivament, amb la seva IP pública i un número de port propi, que pot ser el mateix que l'utilitzat per l'ordinador origen (o no, si ja està "agafat"). D'aquesta manera, aquest port propi permet a l'encaminador saber a quina connexió en concret de les que pugui tenir establertes correspon el trànsit de la xarxa pública. La figura següent mostra un exemple de taula SNAT d'un encaminador fent servir PAT:

RED PRIVADA		INTERNET		Prot.
IP interna	Puerto interno	IP externa	Puerto externo	
10.0.0.13	1050	83.77.100.34	1050	TCP
10.0.0.15	1050	83.77.100.34	12311	TCP

NOTA: Cada entrada de la taula NAT té associats dades addicionals que no es mostren al diagrama anterior, com ara el seu "timeout", més enllà del qual l'entrada és eliminada, entre altres.

En el cas dels NATs de tipus DNAT, el funcionament del PAT és similar però en aquest cas l'associació no es realitza dinàmicament cada cop (i mentre) que s'estableix la connexió sinò que s'estableix de forma estàtica a la configuració del router, de tal manera que aquest sempre sàpiga que les peticions provinents de l'exterior dirigides a la IP pública del router (no pot ser una altra) i a un determinat port públic del router les haurà de reenviar a una determinada IP privada i, en aquesta, a un determinat port. La majoria de routers comercials ofereixen un apartat específic dins del seu panell de control web per indicar aquesta associació permanent, normalment anomenat "Port Forwarding" o similar.

Pas previ: activar l'"Ip-Forwarding"

El més habitual és utilitzar la taula NAT en un sistema que tingui dues tarjes, per tal de funcionar com a "router", on es rep el tràfic per una tarja i es redirigeix a l'altra, que és per on surt. No obstant, per a què aquesta funcionalitat es pugui fer servir, primer hem d'activar el "IP forwarding" entre ambdues tarjes per tal de què el tràfic arribat per una d'elles pugui "passar" internament a l'altra i viceversa.

Per activar l'"IP forwarding" només cal escriure el valor "1" dins de l'arxiu "/proc/sys/net/ipv4/ip_forward" (així, per exemple: *echo 1 | sudo tee /proc/sys/net/ipv4/ip_forward*). Una altra opció equivalent, en comptes d'editar manualment aquest arxiu seria utilitzar la comanda *sysctl*, així: *sudo sysctl -w net.ipv4.ip_forward=1*

NOTA: Recordem que sempre podrem mirar el valor actual de l'opció indicada executant *sysctl net.ipv4.ip_forward*

No obstant, l'activació feta de qualsevol de les maneres anteriors és temporal (és a dir, el valor "1" només es manté mentre la màquina estigui encesa). Si volem que l'IP forwarding" sigui permanent, haurem de modificar l'arxiu anomenat "99-sysctl.conf" que hi ha dins la carpeta "**/etc/sysctl.d**" de manera que hi aparegui la següent línia: *net.ipv4.ip_forward=1* (i, en el cas de fer servir Ipv6, la línia *net.ipv6.conf.all.forwarding=1*) i llavors reiniciar per començar a aplicar-lo (o bé executar la comanda *sudo sysctl -p* per aplicar-lo ipso-facte, sense haver de reiniciar).

NOTA: L'"IP forward" es pot implementar, alternativament, simplement afegint la línia *IPForward=yes* a l'arxiu "*.network" corresponent a qualsevol de les tarjes de xarxa gestionades per systemd-networkd (amb una sola que s'indiqui ja funciona per totes) En realitat, el que fa aquesta línia és executar la comanda *echo 1 > /proc/sys/net/ipv4/ip_forward* automàticament cada cop que la tarja en qüestió s'activi

NAT a la pràctica

Per activar la funcionalitat NAT a Nftables primer hem de crear una taula del tipus específic "nat" i seguidament crear al seu interior alguna cadena o bé de tipus "postrouting" (si es vol implementar SNAT) o bé de tipus "prerouting" (si es vol implementar DNAT) o totes dues, (si es volen implementar els dos tipus de NAT). És a dir (la prioritat 100 i -100 és la per defecte en les cadenes SNAT i DNAT, respectivament):

sudo nft add table ip nat
sudo nft add chain ip nat postrouting { type nat hook postrouting priority 100 \; }
sudo nft add chain ip nat prerouting { type nat hook prerouting priority -100 \; }

A partir d'aquí:

a) Per fer "IP masquerading" en un sistema que faci de "router", la regla que hem d'escriure és (suposant que la seva tarja interna estigui connectada a la LAN 192.168.1.0/24 i enp0s8 estigui connectada a "Internet"):

sudo nft add rule ip nat postrouting ip saddr 192.168.1.0/24 oifname enp0s8 masquerade

NOTA: L'"IP masquerading" es pot implementar, alternativament, simplement afegint la línia *IPMasquerade=yes* a l'arxiu .network corresponent a la tarja enp0s8 (aquesta línia ja implica *IPForward=yes*). En realitat, el que fa aquesta línia es executar internament la comanda *nftables* anterior

b) Per fer SNAT en un sistema que faci de "router" dirigit a un determinat destí (amb IP 1.2.3.4, per exemple), la regla que hem d'escriure és (suposant que la seva tarja interna estigui connectada a la LAN 192.168.1.0/24 i enp0s8 estigui connectada -o tingui definida una ruta- a aquest destí):

sudo nft add rule ip nat postrouting ip saddr 192.168.1.0/24 oifname enp0s8 snat 1.2.3.4

c) D'altra banda, si la màquina on tenim el Nftables funcionant té també un servidor HTTP (per exemple) en marxa escoltant al port 1234 en comptes del port per defecte (80), podem fer que totes les peticions provinents de l'exterior al nostre port 80 es redireccionin automàticament al port 1234 sense que els clients (en aquest cas, els navegadors) tinguin constància. És a dir, fer un redireccionament PAT amb la següent regla:

sudo nft add rule ip nat prerouting iifname enp0s8 tcp dport 80 redirect to 1234

d) Seguint amb el mateix exemple anterior, si la màquina on tenim el "router-tallafocs" funcionant (que és la que rep les peticions dels navegadors externs d'Internet) no fos la mateixa que la que corre el servidor HTTP (és a dir, si el tallafocs ofereix una protecció al davant del servidor web, el qual funciona al "darrera" -en una màquina amb IP privada 192.168.1.4, per exemple-), es podria fer igualment una redirecció de màquina (+port). És a dir, fer un DNAT, així:

sudo nft add rule ip nat prerouting iifname enp0s8 tcp dport 80 dnat 192.168.1.4:80

NOTA: Una altra funcionalitat interessant del DNAT, combinada amb l'acció de duplicar paquets del Nftables, és redirigir una còpia de determinats paquets a un destí alternatiu, així: *sudo nft add rule ip nat prerouting dup to 172.20.0.2*

A continuació es presenta un script d'exemple que converteix un sistema Linux en un "router" gràcies a l'aplicació de les següents regles (sempre i quan s'hagi implementat prèviament l'"IP masquerading"!), on la tarja eth0 representa que està conectada a la LAN interna i la eth1 a "Internet" (òbviament, els ordinadors de la LAN hauran de tenir com a porta d'enllaç per defecte la direcció IP de eth0):

*Regla SNAT que permet l'accés a Internet de la xarxa interna a través de la IP pública del "router"
*Regla DNAT que redirigeix el tràfic provinent d'Internet a un servidor web intern (192.168.1.4)
*Regles "Filter" que només permeten cert tràfic que travessa el router (no tot) i respondre a "pings".

```
#!/usr/sbin/nft -f
flush ruleset
table inet filter {
        #No permeto cap tràfic amb destí al "router" específicament, excepte...
        chain input { type filter hook input priority 0; policy drop;
                      #...el tràfic "loopback", que és necessari...
                       iifname lo accept
                       #...i els "pings" provinents de l'exterior
                       iifname eth1 ip protocol icmp accept
        }
        #No permeto cap tràfic originat des del "router" específicament...
        chain output { type filter hook output priority 0; policy drop; }
        #... ni el que el travessi indiscriminadament...
        chain forward { type filter hook forward priority 0; policy drop;
                        #...a no ser que vagi de dins a fora...
                        iifname eth0 oifname eth1 accept
                        #...o que vagi de fora a dins però només si està relacionat amb tràfic anterior
                        iifname eth1 oifname eth0 ct state established,related accept
        }
}
table ip nat {
        chain prerouting   { type nat hook prerouting priority -100;
                             iifname eth1 tcp dport 80 dnat 192.168.1.4:80
        }
        chain postrouting { type nat hook postrouting priority 100;
                            ip saddr 192.168.1.0/24 oifname eth1 masquerade
        }
}
```

EXERCICIS:

Per realitzar aquests exercicis utilitzarem dues màquines VirtualBox ("MaquinaA" i "MaquinaB") amb un sistema Server (Ubuntu o Fedora) instal.lat. Assegura't de què "MaquinaA" tingui dues tarjes: una en mode "adaptador pont" i una altra en mode "xarxa interna" (suposarem que aquesta és enp0s8) i "MaquinaB" en té només una, en mode "xarxa interna".

1.-a) Assigna a la tarja en mode "xarxa interna" de "MaquinaA" una direcció IP fixa (amb la comanda *ip address add...* per anar més ràpid). D'altra banda, assigna a la única tarja de "MaquinaB" també una direcció IP fixa (igualment amb *ip address add...*) de la mateixa xarxa que la de la IP de "MaquinaA".

b) Comprova que entre les MVs es facin ping i que "MaquinaA" té connexió a qualsevol ordinador de la xarxa de l'institut -o Internet- (gràcies a la seva tarja en mode "adaptador pont") però que "MaquinaB" no

c) Configura (amb la comanda *ip route add default via ...*) la porta d'enllaç de "MaquinaB" per a què aquesta sigui "MaquinaA". Comprova que igualment, des de "MaquinaB" es continua sense poder accedir a la LAN de l'institut -o Internet- fent un ping des de "MaquinaB" a -per exemple- 8.8.8.8.

d) Instal.la el paquet "tshark" (si és Ubuntu) o "wireshark-cli" (si és Fedora) a la "MaquinaA" mentre vas executant els "pings" de l'apartat anterior i executa a "MaquinaA" la comanda *tshark -ni enp0s8*. El que hauries de veure a la pantalla és que els pings de "MaquinaB" sí que arriben a "MaquinaA", però aquesta no fa de "trampolí" a Internet.

e) El que has de fer per "activar el trampolí" és activar l'"IP forwarding" a "MaquinaA"; és a dir, fer que el que rebi una de les seves tarjes de xarxa es transmeti a l'altra i viceversa. Fes-ho. Tot i així, "MaquinaB" encara no rebrà cap resposta als pings enviats a 8.8.8.8. Comprova-ho.

> NOTA: A més a més d'activar l'"IP-Forwarding", cal que la cadena FORWARD tingui com acció per defecte ACCEPT; això és important però, en general, ja és així per defecte, així que en principi no ens preocuparem per això ara.

¿Per què no funciona encara el "ping" tot i haver activat l'"IP-Forwarding"? Perquè no hi ha configurat cap tipus d'enrutament. Fixeu-vos: el paquet "ping" sortint a l'exterior a través de la porta d'enllaç de "MaquinaA" conté com a IP d'origen la de "MaquinaB"; per tant, la resposta "ping" obtinguda de l'exterior tindrà com a IP de destí aquesta IP. No obstant, en rebre aquesta resposta el "router" de l'institut, no sabrà on localitzar aquest destí perquè no reconeix la xarxa de "MaquinaB" com a pròpia (recordem que és una "xarxa interna" i, per tant, invisible). Així doncs, el "router" de l'institut descarta el paquet. Hi ha dues possibles solucions: o bé indicar un enrutament (estàtic) al "router" de l'institut per dir que tots els paquets amb IP de destí "MaquinaB" vagin a parar a la tarja en mode adaptador pont de "MaquinaA" (aquesta tarja sí és accesible per ell perquè pertany a la seva mateixa xarxa LAN) o bé configurar un SNAT en "MaquinaA" per a què abans d'enviar el paquet a la seva porta d'enllaç la seva IP d'origen original sigui substituïda per la de la tarja en mode adaptador pont de "MaquinaA" i així evitar, gràcies a les vinculacions establertes dins la seva taula NAT, aquest problema. Optarem per aquesta opció en no haver de configurar el "router" del centre.

> NOTA:El problema seria el mateix (i la solució és, per tant, la mateixa) si volguéssim fer ping no a Internet sinó a qualsevol ordinador de la xarxa de l'institut: en voler contestar a la IP d'origen original, un ordinador del centre veuria que no pertany a la seva xarxa i redireccionaria la resposta a la seva porta d'enllaç (el "router" del centre), no arribant mai, doncs, la resposta al seu destí, "MaquinaB".

f) Executa a "MaquinaA" la comanda *nft* adient per activar el "masquerading". Un cop fet això, hauràs de veure que la comanda ping 8.8.8.8 que estava executant-se a "MaquinaB" comença a rebre les respostes.

2.-a) Posa en marxa a "MaquinaB" un servidor Netcat escoltant al port 4444 i un altre al port 5555. Executa a "MaquinaA" les comandes *nft* adients per activar el DNAT necessari per a què es pugui accedir a aquests serveis des de la màquina real (és a dir, per a què tot el tràfic entrant per la tarjeta en mode "adaptador pont" i que vagi dirigit el port 4444 de "MaquinaA" passi a redireccionar-se a la IP de "MaquinaB" i al seu port 4444/TCP i tot el que vagi dirigit al port 5555 de "MaquinaA" passi a la IP de "MaquinaB" i el seu port 5555.

b) Independentment de la configuració actual de "MaquinaA" i "MaquinaB", ¿què faria aquesta hipotètica comanda (suposant que la taula "nat" i la cadena "prerouting" estan definides al sistema) : *sudo nft add rule ip nat prerouting udp dport {53,5353} ip saddr 192.168.1.0/24 dnat 208.67.222.222:53* ? ¿Per a què creus que un "hacker" dolent podria servir fer això?

3.- Instal.la ara a "MaquinaA" un servidor SSH, canvia el seu port d'escolta (això es fa editant la directiva *Port* de l'arxiu "/etc/ssh/sshd_config") i posa'l en marxa. Comprova des d'un màquina remota (com pot ser la màquina real o "MaquinaB", és igual) que ara cal que afegeixis al client ssh el paràmetre *-p nº* per poder accedir-hi. Afegeix ara una regla Nftables de tipus DNAT "redirect" per tal de què no calgui afegir aquest paràmetre als clients. Un cop afegida, comprova que, efectivament, pots tornar a executar el client *ssh* sense haver d'especificar cap paràmetre *-p*