Tecnologies Avançades d'Internet
Pràctica 2: *Infraestructura iproute2*

curs 2019-2020

1 Objectius

El principal objectiu d'aquesta pràctica és entendre el funcionament de la infraestructura IP del kernel de GNU/Linux a través de les comandes que ens proporciona iproute2 (**ip** i **tc**).
Tot i que algunes de les funcions que es demanaran a la pràctica es podrien realitzar amb les antigues comandes de configuració (*ifconfig*, *route*, *ifdown*, *ifup*), ja obsoletes, **en aquesta pràctica utilitzeu les `iptools`**. Totes les configuracions s'hauran de realitzar a partir de les comandes que ens proporciona iproute2 amb l'excepció de l'emmascarament d'adreces IP, que s'haurà de realitzar amb l'aplicació *iptables* (vista a la pràctica 1).

2 Introducció al CORE

Per realitzar la pràctica utilitzarem un emulador de xarxa anomenat CORE (Common Open Research Emulator). Aquest emulador permetrà tenir múltiples contenidors que es comportaran com a sistemes Linux aïllats entre si. Aquest contenidors, donat que són sistemes Linux, podran ser clients, hosts, routers o hubs, entre d'altres.

El CORE[1] permet realitzar proves i tasques d'administració en els diferents `containers` (o sistemes) d'una manera còmoda i eficaç. A més, els contenidors tot i que són semblants a les màquines virtuals que podríem utilitzar en VMWare[5] són molt més lleugeres reduint així el consum de recursos.

L'emulador es proporciona totalment configurat dintre d'una màquina virtual. D'aquesta manera es podran realitzar les pràctiques tant al laboratori com a qualsevol altre màquina que suporti VMWare.

Els contenidors de CORE es podran trobar, sempre que s'hagi iniciat prèviament l'emulació, en el directori de la màquina virtual: */tmp/pycore.xxxx* on *xxxx* serà l'identificació de l'emulació actual (pot variar entre emulacions).
Dins d'aquest directori trobarem un directori **.conf* per a cada un dels contenidors. Aques directori **.conf* és el directori *home* del contenidor.
Per exemple el directori */tmp/pycore.xxxx/gateway.conf/* contindrà l'estructura de directoris del `gateway` per a l'emulació `xxxx`.

En el cas de la pràctica només heu de llençar una emulació alhora.

2.1 Aspectes a considerar

Per a la correcta consecució de la pràctica heu de tenir en compte el següent:

- Tots els passos a realitzar s'han de fer des de la línia de comandes, no serà vàlid modificar la configuració del CORE per aconseguir la funcionalitat que us demanem, excepte si s'especifica el contrari.
- Assegureu-vos d'utilitzar sempre l'escenari de CORE adient per cada pràctica.

  Per aquesta segona pràctica el fitxer de l'escenari és el **`p2.imn`**.

  Un cop heu arrencat la màquina virtual que conté el CORE instal.lat cal que us descarregueu aquest escenari. Per a fer-ho, en la màquina virtual, descarregueu-vos el fitxer p2.imn del campus virtual i deixeu-ho a l'escriptori.
- Una vegada parem l'emulació, tota la configuració així com el contingut dels `containers` s'esborrarà. Així doncs serà important guardar els scripts de manera regular.
- La màquina virtual serà utilitzada per altres grups, enrecordeu-vos de no deixar scripts en aquesta.

3 Enunciat

3.1 Esquema principal

Tal i com veiem a la figura 1 el muntatge realitzat a l'escenari **p2.imn** consisteix en un *gateway*, el node `n2`, amb dues interfícies de sortida (`eth0` i `eth1`) connectades a dues xarxes independents. Ens referirem a les IP's assignades a cadascuna d'aquestes dues interfícies com a `IP_router_eth0` per la de la interfície `eth0` i `IP_router_eth1` per la `eth1`.

Cadascuna de les dues xarxes té el seu propi gateway a partir del qual podem accedir a Internet. El gateway de la xarxa `192.168.0.0/24` és el que té com IP `192.168.0.1`. El gateway de la xarxa `192.168.1.0/24` és el que té com IP `192.168.1.1`.

En una configuració estàndard només utilitzaríem un dels dos gateways, `192.168.0.1`o bé `192.168.0.1`per poder sortir a Internet. En el nostre esquema, però, configurem el nostre gateway (`n2`) per a que pugui utilitzar els dos gateways als que està connectat amb la possibilitat de fer balanceig de càrrega entre ells. Podrem indicar-li al gateway quin percentatge d'ús volem que s'utilitzi per cada interfície.

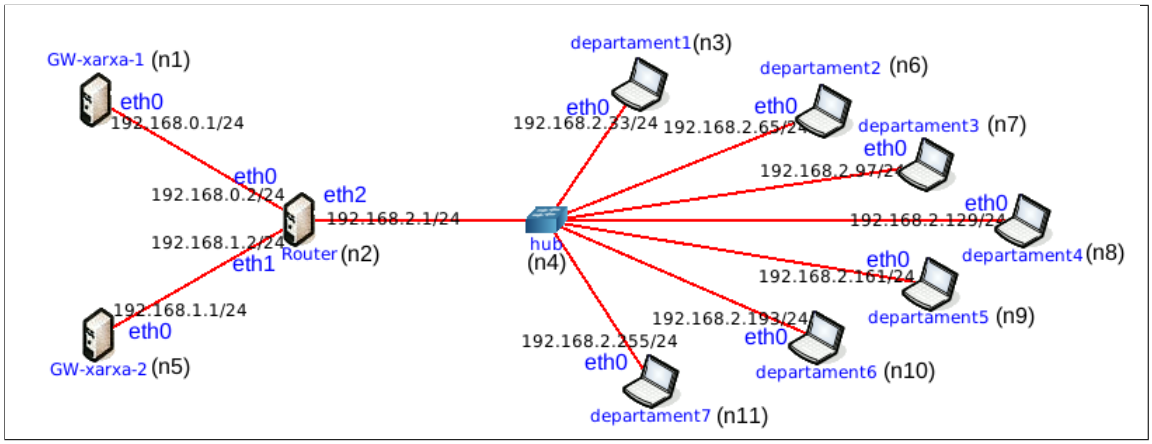

Figura 1: *Escenari de la part Obligatòria.*

La tercera interfície del gateway és la `eth2`, que tindrà assignada la IP `192.168.2.1`. Aquesta interfície l'anomenarem interfície d'entrada ja que serà la que estarà connectada a les xarxes d'on penjaran els nostres ordinadors clients. Hem de veure aquests ordinadors clients com si estiguessin repartits pels diferents laboratoris, aules o departaments de la nostra institució.

D'aquesta manera, haurem d'organitzar els clients en diferents subxarxes. Això ens permetrà donar o tallar l'accés a Internet a les subxarxes, des del *gateway*.

Concretament dividirem el rang `192.168.2.0/24`en set subxarxes diferents tal i com mostra la taula 1.

Taula 1: *Exemple dels rangs d'IPs per les subxarxes d'entrada.*

	Subxarxa	Rang vàlid d'IPs
departament 1	192.168.2.32/27	192.168.2.33∼62
departament 2	192.168.2.64/27	192.168.2.65∼94
departament 3	192.168.2.96/27	192.168.2.97∼126
departament 4	192.168.2.128/27	192.168.2.129∼158
departament 5	192.168.2.160/27	192.168.2.161∼190
departament 6	192.168.2.192/27	192.168.2.193∼222
departament 7	192.168.2.224/27	192.168.2.225∼254

La IP del gateway associada a la seva interfície `eth2` es pot configurar de dues maneres diferents:

Solució 1: La IP associada a la `eth2` del gateway podria ser `192.168.2.1` amb una màscara de 24 bits (255.255.255.0). Els diferents rangs d'IPs vàlids pels ordinadors dels diferents departaments haurien de ser els que veiem a la taula 1 però la màscara dels ordinadors clients també hauria de ser de 24 bits en lloc de 27. D'aquesta forma, **tant el gateway com els clients estaran a la mateixa xarxa**. **Tot i això, després utilitzarem les subxarxes definides (amb màscara de 27 bits) per identificar el grup de workstations d'on provenen els paquets** i aplicar les configuracions que toqui a nivell de subxarxa.

Així podrem tallar, o no, l'accés a Internet a les subxarxes directament mitjançant **les regles d'emmascarament** amb iptables. La idea consisteix en només fer emmascarament a les subxarxes que poden tenir accés a internet.

Solució 2: Una altra possibilitat seria assignar les IPs i màscares reals de cada subxarxa (192.168.2.X/27) als ordinadors clients i assignar 7 IPs diferents a la interfície d'entrada (`eth2`) del gateway. De totes maneres, us recomanem que utilizeu la proposta anteriorment plantejada per no complicar excessivament la configuració del gateway.

3.2 Funcionalitat de la pràctica

La pràctica consisteix en aplicar una configuració mitjançant `iptables` i les eines de `iproute2` que permeti el següent:

1 Permetre o tallar l'accés a Internet als clients de les diferents subxarxes definides.

2 En el gateway s'ha de poder fer balanceig de càrrega entre les seves interfícies de sortida: `eth0` i `eth1`.

3 El gateway ha de poder limitar l'ample de banda màxim permès per cadascuna d'aquestes subxarxes.

Si vulguéssim ser totalment estrictes amb la configuració d'aquest bandwidth hauríem de tenir en compte el factor de balanceig de càrrega de les diferents interfícies de sortida i la velocitat de les xarxes on estan connectades. Per no complicar excessivament la pràctica no tindrem en compte aquests **paràmetre**s.

3.3 Passos a realitzar

És important tenir clar que **no** hem de modificar els fitxers de configuració del sistema ni cap configuració gràfica del CORE; en reiniciar, el sistema perdrà totes les modificacions que haguem realitzat[1].

Els passos que s'haurien d'anar seguint per provar totes les configuracions són:

1. PERMETRE EL FORWARDING

Per poder redirigir els paquets de les xarxes d'entrada cap a les de sortida, necessitem que el sistema permeti el forwarding de paquets. Per defecte, la majoria de distribucions basades en GNU/Linux ho desactiven per qüestions de seguretat. En el nostre cas, però, el gateway ja estarà configurat amb aquesta opció, de manera que no farà falta activar el flag "/proc/sys/net/ipv4/ip_forward".

2. PERMETRE L'EMMASCARAMENT O NO SEGONS LA SUBXARXA D'ORIGEN

Per poder realitzar aquesta opció s'ha d'utilitzar l'aplicació "iptables" (més informació d'iptables a [3]).

Cal que creeu dos scripts:

activaEmmascarament.sh: Aquest escript ha de fer la traducció d'adreces per a **totes les subxarxes** de manera que tinguin connexió a internet.

modificaAccesPerSubxarxa.sh: Aquest script ha de **modificar** la configuració establerta per l'execució de l'anterior script, de manera que es pugui especificar si es vol permetre, o bé tallar l'accés a internet, a una determinada subxarxa. A aquest escript li haureu de passar, doncs, un parell de **paràmetres**:

- opció d'activar/desactivar
- subxarxa

3. CONFIGURACIÓ DE LES TAULES DE ROUTING

És recomanable que proveu, primer un i després l'altre, la connexió a Internet de cada xarxa de sortida (`192.168.0.1` i `192.168.1.1`), configurant l'entrada per defecte a la taula de routing amb la comanda **ip route**.
Seguidament haureu d'eliminar aquesta entrada per defecte i investigar com poder organitzar el balanceig de càrrega entre les dues interfícies de sortida. Podeu trobar informació al capítol 4 de [4].

[1] Si la màquina virtual es queda desconfigurada o porta problemes, reinicieu la màquina **real**: Ctrl Alt F2 i Ctrl Alt Supr

Cal que creeu un script:

canviaBalanceig.sh: Aquest script ha de modificar la configuració per defecte i establir el balanceig de càrrega entre les interficies del router `eth0` i `eth1`. Caldrà doncs, que a aquest script li passeu dos **paràmetres**:

- el percentatge de càrrega associat a la interfície `eth0`.
- el percentatge de càrrega associat a la interfície `eth1`.

4. LIMITACIÓ DE L'AMPLE DE BANDA UTILITZAT PER CADA SUBXARXA

Abans de configurar el gateway per limitar l'ample de banda segons cada subxarxa d'origen, és recomanable que es provi de limitar l'ample de banda total de cadascuna de les interfícies de sortida. Al capítol 9 de [4] trobareu informació sobre com fer-ho.

Veureu que al capítol 9.2.2 de [4] se'ns recomana utilitzar un **tipus de política de cua** anomenat **`tbf`** (*token bucket filter*) per limitar aquest ample de banda. Mitjançant aquesta politica de cues (*token bucket filter*) es pot fixar el bandwidth en kbits associat a una cua d'una interfíe.

Per provar que la configuració funciona correctament, des de dues consoles diferents podeu executar una de les següents comandes en cada consola:

```
watch -n 1 tc -s qdisc show dev eth0
watch -n 1 tc -s qdisc show dev eth1
```

Haureu de veure la cua que heu creat. Des d'una altra consola genereu tràfic. En la consola del `watch` veureu com els datagrames s'encuen en una de les dues cues.

Un cop provat que la limitació de l'ample de banda funciona correctament, heu de modificar aquesta configuració per associar un bandwidth diferent par a cada una de les subxarxes de les workstations. Trobareu informació sobre com fer-ho als capítols 9 i 12 de [4].

És important que proveu que la vostra configuració funciona des de les diferents subxarxes de les workstations.

Per portar a terme aquesta configuració haureu d'utilitzar les comandes

`tc qdisc`, i **`tc filter`**.

Cal que creeu dos scripts:

activaAmpleDeBanda.sh: Aquest script ha d'activar una limitació de l'ample de banda per defecte, per a cada una de les subxarxes.

modificaAmpleDeBandaPerSubxarxa.sh: Aquest script ha de **modificar** la configuració establerta amb l'execució de l'anterior script, de manera que es pugui especificar un nou ample de banda associat a una determinada subxarxa. Li haureu de passar, doncs, un parell de **paràmetres**:

- subxarxa
- nou ample de banda

5. TREURE LA LIMITACIÓ DE L'AMPLE DE BANDA DEL GATEWAY

Un cop totes les configuracions anteriors funcionin correctament us adonareu que també haureu limitat l'ample de banda del mateix gateway.

Si mireu les cues i genereu tràfic des del gateway veureu que el tràfic del gateway s'encúa en una de les cues que heu generat.

Això no interessa. Cal que el gateway aprofiti al màxim l'ample de banda disponible. Per aquest motiu **s'ha de treure aquesta limitació.**

Cal que creeu un script:

resetAmpleDeBandaGateway.sh: Aquest script ha de modificar la configuració per defecte per tal de fer passar el tràfic generat des del gateway per una cúa que tingui configurat el màxim ample de banda.

Recordeu que per provar i testejar totes aquestes configuracions el sistema ens proporciona aplicacions com `tcpdump` o `wireshark`, que ens permet *sniffar* què és el que realment està passant per les diferents interfícies de xarxa del sistema.

3.4 Part No obligatòria

3.4.1 Escenari de la Part Opcional

Per realitzar la part opcional d'aquesta pràctica es proporciona un nou escenari que s'ha de carregar amb el CORE. Aquest escenari s'anomena **p2-opcional.imn**. En la màquina virtual, Descarregueu-vos el fitxer p2-opcional.imn del campus virtual i deixeu-ho a l'escriptori.
L'escenari opcional es pot veure a la figura 2. **Recordeu que no heu de modificar res de l'escenari**. S'ha de fer tot per línia de comandes.

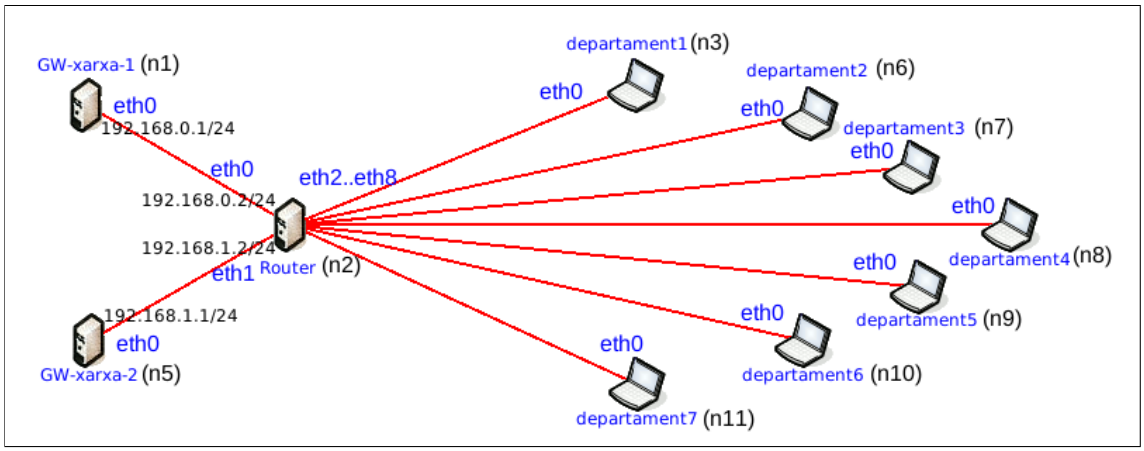

Figura 2: *Escenari de la part opcional.*

3.4.2 Descripció

1 **Diferenciar entre les set subxarxes**

En aquest escenari el gateway, té definides una interfície diferent per a cada una de les subxarxes, `eth2`, `eth3`, ... `eth8`. Caldrà que utilitzeu la comanda `ip a` per assignar la IP que toqui a cada una de les interfícies. Recordeu que heu d'indicar la màscara de xarxa \27 per a cada una de les IPs que assigneu.

Cal que configureu les taules d'encaminament tant del gateway com dels clients, per a diferenciar entre les set diferents subxarxes a on estan els clients.

Per a fer-ho creeu els escripts:

configuraSubxarxes.sh: Aquest escript s'executarà en el `gateway`. Associarà una IP a cada una de les interfícies que connecten amb les subxarxes: `eth2..eth8`.

configuraSubxarxa.sh: Aquest escript s'executarà en el `client`. Rebrà com a **paràmetre**:

- una subxarxa.

Caldrà configurar la taula d'encaminament del client en funció de la subxarxa a on està situat.

Aquesta part opcional està valorada amb 0.75 punts.

2 **Marcatge de paquets utilitzant iptables**

Per a marcar els paquets en funció de la subxarxa d'origen, en lloc d'utilitzar el `realm` en la taula d'encaminament del `gateway`, caldrà que utilitzeu `iptables`, concretament el target `MARK`.

Per dirigir els paquets a la cua que toqui segons la seva marca, caldrà que utilitzeu unaltre filtre diferent del `route`. El nou filtre que utilitzareu serà el **fwmark** (fw). Al capítol 9 i 12 de [4] trobareu informació sobre com fer-lo servir.

Creeu un nou escript amb la nova configuració:

canviaFiltrePerFW.sh: Aquest script ha de fer el nou marcatge mitjançant `iptables` i afegir els nous filtres `fw`.

Aquesta part opcional està valorada amb 1'25 punts.

4 Altres aspectes, recomanacions, ...

- La imatge de la màquina virtual esta al directori:

  `/opt/vmware/Debian-7.x_32-bit`

  Per a iniciar-la, cal que obriu el fitxer `Debian-7.x 32-bit.vmx` des de la línia de comandes:

  ```
  vmplayer Debian-7.x\ 32-bit.vmx &
  ```

- S'ha d'iniciar sessió a la màquina virtual com a `root` on la password és "root".

- El servidor HTTP que hi ha instal.lat en el gateway serveix els fitxers que hi hagin en el directori `/tmp/pycore.xxxx/gateway.conf/var.www` (on `xxxx` serà l'identificador del CORE per a l'emulació actual).

- Tots utilitzareu la mateixa màquina virtual. No us deixeu pràctiques, scripts, ... a les mateixes.

- Recordeu també que quan una emulació es para, tots els fitxers desats en els linux containers dels nodes s'esborren. Us pot servir per a començar la pràctica de nou. **Recordeu-vos però de passar els scripts que vulgueu conservar abans de parar la simulació al vostre compte**.

- La comanda `iptables-save` mostra per pantalla la configuració de totes les taules de `iptables`. Us pot ser d'utilitat.

  Podeu volcar aquesta sortida cap a un fitxer i a posteriori restaurar-ho amb la comanda `iptables-restore`:

  ```
  iptables-save > iptables.conf
  iptables-restore iptables.conf
  ```

  Podeu utilitzar aquestes comandes i el fitxer `iptables.conf` en els scripts de configuració `gateway.sh` i `client.sh`.

- A continuació us indiquem com ho podeu fer per copiar els fitxers del vostre compte de pràctiques cap al **gateway|workstation** i a la inversa. Cal que hagueu iniciat la simulació amb el CORE.

  Noteu que un cop l'emulació està iniciada, en la màquina virtual s'han creat els directoris:

Per al gateway `/tmp/pycore.XXXX/gateway.conf`

Per a les workstations `/tmp/pycore.XXXX/workstation-nX.conf`

- **Per copiar els fitxers que tenim en el nostre compte de pràctiques cap al `gateway|workstation` un cop l'emulació està iniciada**:
  1. Obriu un terminal en la màquina virtual `Debian`.
  2. Executeu la comanda sense les cometes dobles:
     `"scp tai-a1@deic-dc1.uab.cat:iptables/*.sh ."`
  3. Per a transferir aquest fitxer cap al gateway o workstation ho podeu fer de forma visual o per comandes:

     De forma visual (a) Obriu en la màquina virtual el gestor de fitxers. Cliqueu en en 'Systema de Fitxers'. Navegueu fins el directory `/root/`. En aquest directori és a on tindreu els fitxers que heu copiat del vostre comtpe de pràctiques.

     (b) Des de l'emulador, obriu una consola en el `gateway/workstation` amb doble click sobre la icona del node.

     (c) Des de l'aplicació de gestió de fitxers arrossegueu el fitxer que voleu copiar a la consola del `gateway/workstation`. Us apareixerà en la consola la ruta d'origen del fitxer.

     (d) Des de la consola del `gateway/workstation` afegiu a la ruta la comanda per copiar el fitxer en el directori actual:
     `cp /root/XXX.sh .`
     No us oblideu del punt final!

     Per línia de comandes Des del terminal de la màquina virtual a on us heu copiat els fitxers del vostre compte de pràctiques executeu:
     `cp *.sh /tmp/pycore.40397/gateway.conf/`
     i ja està.

- **Per desar la feina que tenim en el `gateway/workstation` cap al vostre compte de pratiques**, directori iptables, fer:
  1. Obriu un terminal en el `gateway/workstation` del CORE.
  2. Executeu la comanda sense les cometes dobles:
     `"scp *.sh tai-a1@deic-dc1.uab.cat:iptables"`

  Si voleu transferir els fitxers cap a la màquina virtual, ho podeu fer visualment amb el gestor de fitxers o bé executant la comanda:

  `cp *.sh /root`

- Si teniu algun problema amb alguna de les màquines virtuals reinicieu la màquina del laboratori.
- Per aplicar les configuracions que us demanem, recordeu que no s'ha de necessitar reinicialitzar la màquina virtual ni cap dels `containers` del CORE en cap moment.
- Quan hagueu acabat de treballar amb la màquina virtual, amb la `Debian`, penseu que cal aturar degudament el seu sistema operatiu. Cal que executeu la seqüència: `Ctr+Alt+Supr`. Quan s'hagi acabat el procés d'aturada del sistema operatiu, abans de que torni a carregar-se, tanqueu l'aplicació *vmplayer*.

Referències

[1] **CORE Homepage.** URL: *http://www.nrl.navy.mil/itd/ncs/products/core*. Pàgina web oficial del CORE..

[2] **IP command reference.** URL: *http://linux-ip.net/gl/ip-cref/ip-cref.html*. Manual de la comanda ip.

[3] **Manual d'iptables.** Manual del sistema de iptables accessible a partir de la comanda "man 8 iptables".

[4] **Linux Advanced Routing & Traffic Control.** URL: *http://www.lartc.org*. HOWTO de Linux Advanced Routing & Traffic Control utilitzant l'infraestructura de iproute2.

[5] **VMWare homepage.** URL: *http://www.vmware.com*. Pàgina web oficial de l'empresa vmware.