Equilibrage de charges

Jean-Marc Pouchoulon

Mai 2019

1 Objectifs du TP et organisation.

1.1 Les compétences à acquérir à la fin de cette séance sont les suivantes :

— Construire deux architectures d'équilibrage de charge sous Linux Virtual Server("Direct Routing" et Dnat) au niveau 4.

— Equilibrer et manipuler des flux http avec un équilibreur de Niveau 7 : Haproxy.

1.2 Organisation, recommandations et notation du TP.

L'équilibrage de charge n'est pas une "Rocket Science" mais reste délicat à configurer. Il n'y a pas de place pour l'approximation, chaque configuration a son importance.

Il vous explicitement demandé de faire valider votre maquette par l'enseignant avant de la démonter. Ces "checks" permettront de vous noter.

Un compte rendu succinct (fichiers de configuration , copie d'écran montrant la réussite de la construction ...) est demandé et à rendre sur Moodle Didex.

Une partie "Tips and Tricks" est là pour vous aider , lisez-la avant de démarrer le TP.

Pour les maquettes GNS3 vous travaillerez par groupe de deux (Un maquettant le mode "Direct Routing" et l'autre le mode "NAT"). Pour les maquettes physiques vous travaillerez par groupe de 4.

Vous réaliserez au choix une des deux maquettes.

2 Linux Virtual Server.

Vous devez donc réaliser un travail exploratoire réalisé sous GNS3 et ensuite réaliser une deux maquettes "Linux Virtual Server" avec vos postes de travail et le matériel de la licence :

— Une utilisant Linux Virtual Server en mode "NAT".

— L'autre utilisant Linux Virtual Server en mode "Direct Routing".

Dans les deux cas une VIP ("Virtual IP Address") est utilisée. C'est celle qui recois la requête du client http. Les requêtes reçues sont transformées par LVS et redirigées vers les RIPs ("Real IP Address") ;

Vous aurez besoin de vous isoler du réseau de la salle via un routeur Cisco 1800 et vous pourrez utiliser un switch pour plus de confort.

Le NAT est nécessaire pour vous permettre d'installer des paquets sur vos machines si nécessaires.

Si votre machine ne dispose pas de deux cartes réseaux vous utiliserez des convertisseurs USB ethernet.

Le fait d'avoir un Apache et un Nginx vous permet de différencier facilement un "Real IP Server" d'un autre en affichant les headers de la requête http : chaque serveur génère des headers différents.

2.1 Réalisation d'une maquette Linux Virtual Server NAT

En mode NAT LVS travaille au niveau IP en se basant sur le mécanisme de NAT. Le schéma de la maquette GNS3 est le suivant :

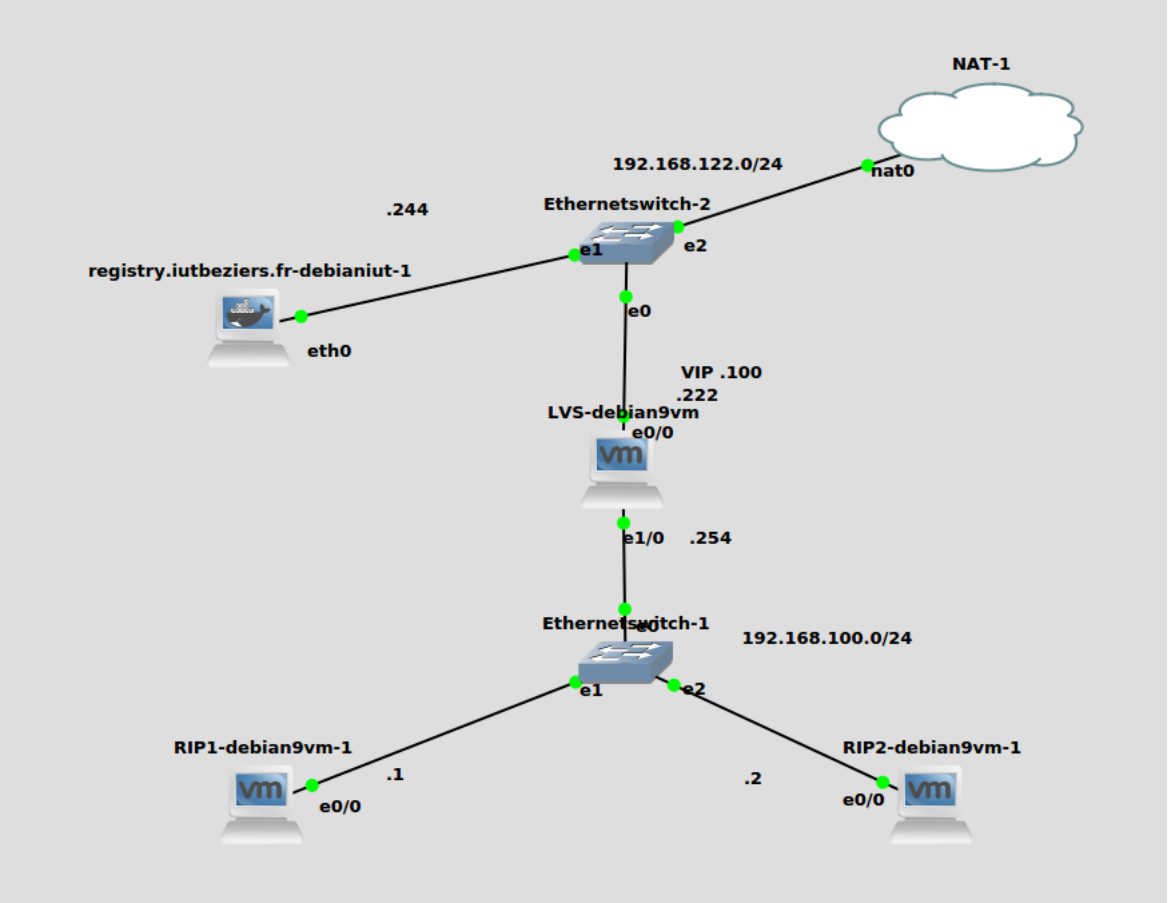

FIGURE 1 – LVS DNAT

Clefs pour réussir la configuration LVS Nat :

— Les "Real IP Servers" ont comme passerelle par défaut le LVS.

— Le routage entre les cartes réseaux est activé dans le LVS.

2.2 Caractérisation de LVS NAT.

1. Faites deux captures de trames illustrant le mécanisme utilisé par LVS NAT pour TCP et UDP.
2. Faites en sorte que nat1 supporte deux fois plus de connexions que nat2 sur le flux http.Testez-le via la commande ab (apache bench) issue du paquet Debian apache2-utils. Voir `http://www.linuxvirtualserver.org/docs/scheduling.html` et `http://www.keepalived.org/pdf/asimon-jres-paper.pdf` Afin de suivre l'évolution du nombre de connexions utilisez la commande :

   ```
   ipvsadm -L -n
   ```

3. Rendez la connexion permanente pour 3600 secondes entre votre client et un RIP. Quel est l'intérêt de réaliser cette configuration ? Via la commande précédente visualisez que vous êtes bien toujours connecté au même RIP depuis votre poste client.

2.3 Réalisation d'une maquette Linux Virtual Server DR.

C'est le mode le plus performant de LVS car il travaille au niveau 2. Le schéma de la maquette est le suivant :

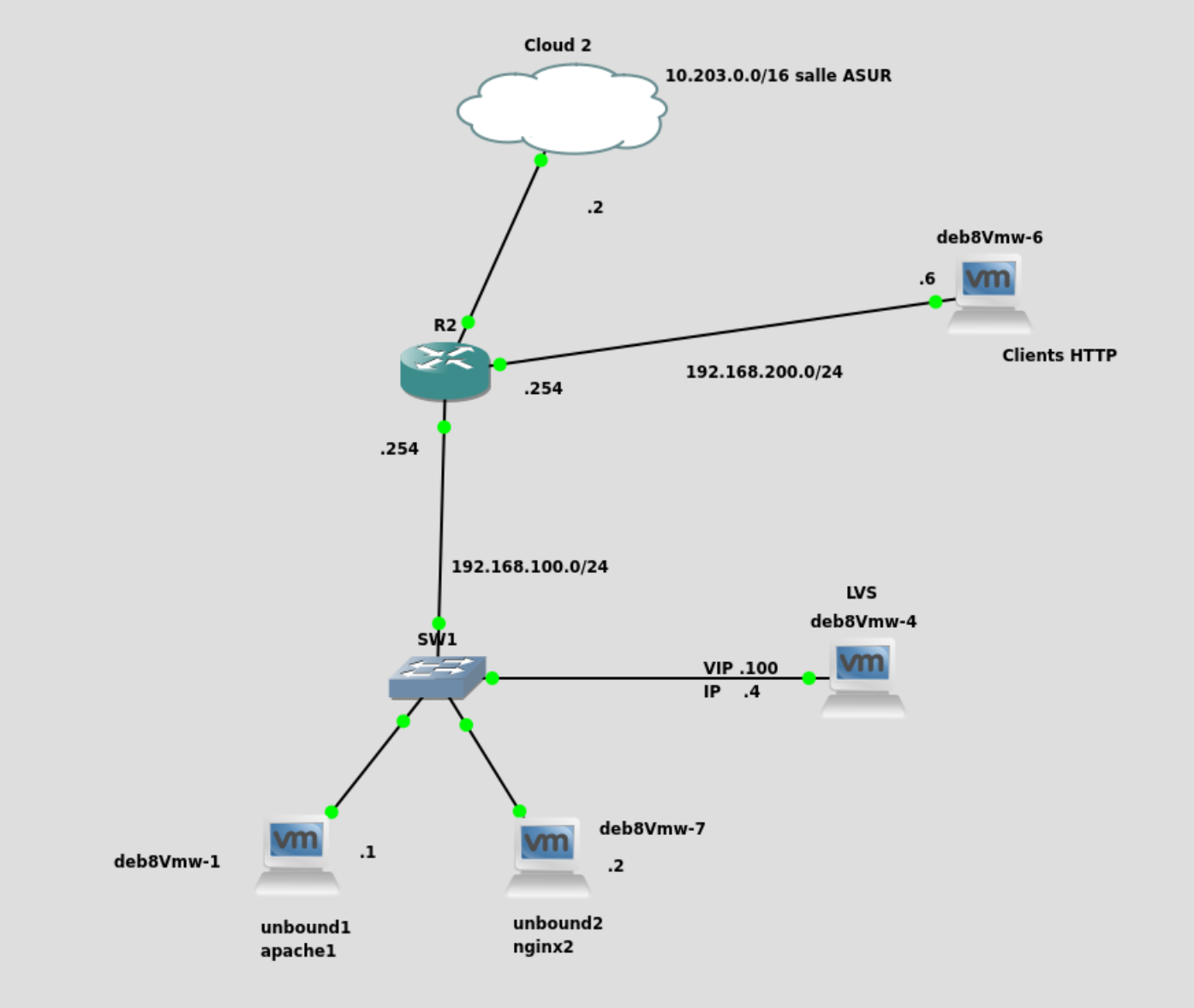

FIGURE 6 – LVS DR

Clefs pour réussir la configuration LVS DR :

— Le flux rentre par le "Director", le director change l'adresse ethernet de destination du paquet avec celle d'un RIP (Real IP Server). Le paquet est transmis au RIP, le RIP qui a aussi la VIP répond directement au client.

— Les RIP ont donc aussi la VIP du director configurée sur la loopback (masque /32) mais ne doivent pas répondre aux requêtes ARP sur leur VIP.Seul le directeur LVS doit le faire. Il faut donc désactiver la résolution arp.(voir "tips and tricks" pour désactiver la réponse aux requêtes arp sur le RIP).

On peut ne pas configurer de VIP sur les RIPs si on utilise IPTABLES. C'est une méthode moins performante mais plus simple à mettre en oeuvre. (voir "tips and tricks" pour utiliser iptables).

— Le routage est activé dans le LVS.

— Le client se trouve forcément sur un autre réseau que le LVS.

— Il faut une VIP distincte de l'adresse réseau du LVS.

2.4 Caractérisation de LVS DR.

1. Faites deux captures de trames illustrant le mécanisme utilisé par LVS DR pour équilibrer le service web.

3 Tips and tricks

3.1 Configuration du NAT sur un routeur Cisco.

```
interface FastEthernet0/0
ip address dhcp
ip nat outside
ip virtual-reassembly
duplex auto
speed auto
!
interface FastEthernet0/1
ip address 192.168.0.1 255.255.255.0
ip nat inside
ip virtual-reassembly
duplex auto
speed auto
!
ip nat inside source list 100 interface FastEthernet0/0 overload
!
access-list 100 permit ip 192.168.0.0 0.0.0.255 any
```

3.2 Configuration d'une seconde adresse IP sur une interface (DR et NAT).

```
# The primary network interface
allow-hotplug eth0
iface eth0 inet static
address 192.168.100.1
netmask 255.255.255.0
gateway 192.168.100.254

# Secondary address
iface eth0 inet static
address 192.168.100.100
netmask 255.255.255.0
```

3.3 Configuration d'une seconde adresse IP sur une interface de loopback d'un RIP (DR).

```
# The loopback network interface
auto lo
iface lo inet loopback
up ip add add 192.168.100.100/32 dev lo
down ip add del 192.168.100.100/32 dev lo

# THE primary network interface
allow-hotplug eth0
iface eth0 inet static
address 192.168.100.1
netmask 255.255.255.0
gateway 192.168.100.254
```

3.4 Désactivation de l'arp pour les interfaces RIP des serveurs.

```
echo "0" >/proc/sys/net/ipv4/ip_forward
echo "1" >/proc/sys/net/ipv4/conf/all/arp_ignore
echo "2" >/proc/sys/net/ipv4/conf/all/arp_announce
echo "1" >/proc/sys/net/ipv4/conf/default/arp_ignore
echo "2" >/proc/sys/net/ipv4/conf/default/arp_announce
echo "1" >/proc/sys/net/ipv4/conf/lo/arp_ignore
echo "2" >/proc/sys/net/ipv4/conf/lo/arp_announce
echo "1" >/proc/sys/net/ipv4/conf/eth0/arp_ignore
echo "2" >/proc/sys/net/ipv4/conf/eth0/arp_announce
```

```
dans /etc/sysctl.conf
net.ipv4.conf.lo.arp_ignore = 1
net.ipv4.conf.lo.arp_announce = 2
```

Il peut être nécessaire de supprimer le cache arp sur le routeur :

```
router#clear ip arp 192.168.122.100
router#sh arp
Protocol  Address          Age (min)  Hardware Addr   Type   Interface
Internet  192.168.0.1            -   c201.2582.0001  ARPA   FastEthernet0/1
Internet  192.168.0.2            8   8ac2.6215.c067  ARPA   FastEthernet0/1
Internet  192.168.122.1           3   5254.00f3.4471  ARPA   FastEthernet0/0
Internet  192.168.122.2           6   000c.2950.1984  ARPA   FastEthernet0/0
Internet  192.168.122.3           3   000c.2942.841e  ARPA   FastEthernet0/0
Internet  192.168.122.236         -   c201.2582.0000  ARPA   FastEthernet0/0
```

3.5 Utilisation de iptables pour ne pas configurer de VIP sur les RIPs

```
# si 10.0.0.40 est la VIP
 iptables -t nat -A PREROUTING -d 10.0.0.40 -j REDIRECT
```

4 Briques logicielles.

4.1 Utilisez un client en ligne de commandes :httpie.

httpie vous permet de faire du web en mode cli et en particulier de visualiser les headers HTTP. Curl peut aussi être utilisé à cet effet.

```
apt get install httpie
# si 10.0.0.40 est la VIP
http -h 10.0.0.40
```

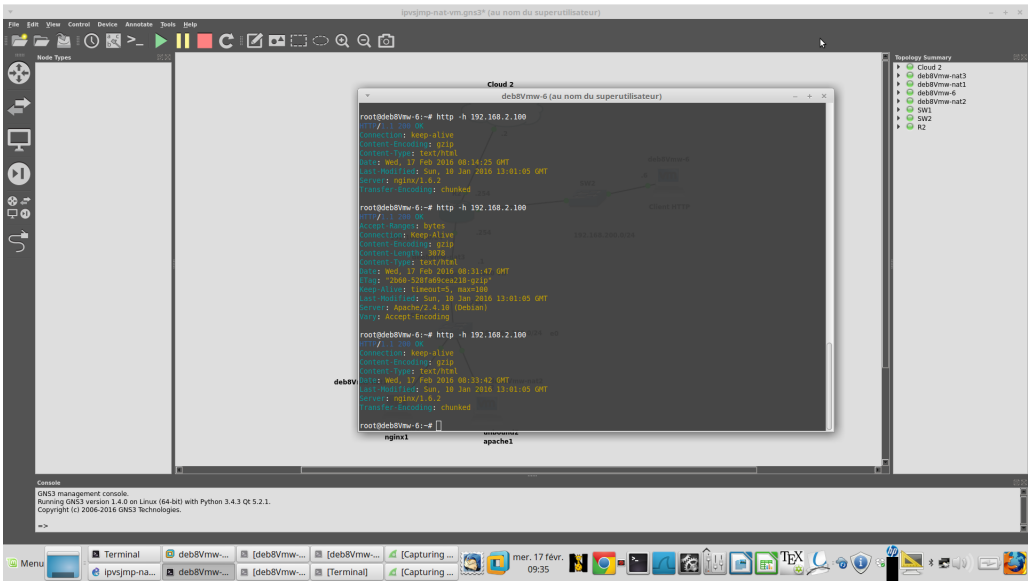

FIGURE 9 – Le web en mode CLI avec httpie

4.2 Manipulation du LVS avec la ligne de commande.

ipvsadm est l'utilitaire qui permet de manipuler LVS en ligne de commande. Les commandes ipvsadm-save et ipvsadm-restore permettent de sauvegarder et de restaurer les configurations LVS. La commande suivante permet de "flusher" la configuration :

```
ipvsadm-restore < /dev/null
```

4.3 Configuration d'un resolver/cache unbound

Après installation de unbound , créez le fichier /etc/unbound/unbound.d/simple.conf et décommentez dans /etc/default/unbound la ligne relative au fichier de configuration.

```
server:
        interface: 0.0.0.0
        access-control: 10.0.0.0/16 allow
        access-control: 127.0.0.0/8 allow
        access-control: 192.168.0.0/16 allow
        verbosity: 1

forward-zone:
      name: "."
      forward-addr: 10.6.0.1
```

4.4 Pour information : désactivation de l'ICMP redirect au cas en mode "direct routing"

```
echo 0 > /proc/sys/net/ipv4/conf/all/send_redirects
echo 0 > /proc/sys/net/ipv4/conf/default/send_redirects
echo 0 > /proc/sys/net/ipv4/conf/eth0/send_redirects
```

5 Toujours plus haut... avec Haproxy un loadbalancer de niveau sept.

Vous utiliserez la syntaxe backend et frontend pour haproxy.cfg.

1. A quoi correspondent les différentes options dans haproxy.cfg ?
2. Configurez HAProxy comme reverse proxy de www.iutbeziers.fr .
3. Configurez HAProxy avec deux acls pour qu'un /ent vous renvoie vers `https://www.didex.fr` et qu'un /univ vous renvoie vers `https://moodle.umontpellier.fr` .
4. Créez deux containers ou deux serveurs web (NGINX et APACHE), utilisez haproxy pour load balancer le traffic sur les deux serveurs .
5. Utilisez haproxy comme terminaison SSL pour les deux load balancers .
6. Mettez en place des stats et expliquez leurs contenus . voir `https://www.datadoghq.com/blog/monitoring-haproxy-performance-metrics/`

5.1 Tips and tricks

— Exemple de configuration HAPROXY.

— Lancement en mode debug de Haproxy.

— Commande de vérification d'un Certif SSL.

— Créer un certificat SSL

```
global
        log /dev/log    local0
        log /dev/log    local1 notice
        chroot /var/lib/haproxy
        stats socket /run/haproxy/admin.sock mode 660 level admin
        stats timeout 30s
        user haproxy
        group haproxy
        daemon

        # Default SSL material locations
        ca-base /etc/ssl/certs
        crt-base /etc/ssl/private
        tune.ssl.default-dh-param 2048

        # Default ciphers to use on SSL-enabled listening sockets.
        # For more information, see ciphers(1SSL). This list is from:
```

```
        #  https://hynek.me/articles/hardening-your-web-servers-ssl-ciphers/
        ssl-default-bind-ciphers ECDH+AESGCM:DH+AESGCM:ECDH+
        AES256:DH+AES256:ECDH+AES128:DH+AES:ECDH+3DES:DH+3DES:RSA+AESGCM:RSA+AES:RSA+3DES:!aNULL:!MD5:!DSS
        ssl-default-bind-options no-sslv3

defaults
        log     global
        mode    http
        option  httplog
        option  dontlognull
        timeout connect 5000
        timeout client  50000
        timeout server  50000
        errorfile 400 /etc/haproxy/errors/400.http
        errorfile 403 /etc/haproxy/errors/403.http
        errorfile 408 /etc/haproxy/errors/408.http
        errorfile 500 /etc/haproxy/errors/500.http
        errorfile 502 /etc/haproxy/errors/502.http
        errorfile 503 /etc/haproxy/errors/503.http
        errorfile 504 /etc/haproxy/errors/504.http

frontend f_yahoo
    bind *:8001
    default_backend b_yahoo

frontend f_publinet
    bind *:8002
    default_backend b_publinet

frontend f_wwwiutbeziers
    bind *:8003
    default_backend b_wwwiutbeziers

frontend publinet-https
        bind 192.168.20.134:443 ssl crt /etc/ssl/certs/cert.pem
        reqadd X-Forwarded-Proto:\ https
        rspadd Strict-Transport-Security:\ max-age=31536000

frontend f_publinet
    bind *:8002
    default_backend b_publinet

frontend f_wwwiutbeziers
    bind *:8003
    default_backend b_wwwiutbeziers

frontend publinet-https
        bind 192.168.20.134:443 ssl crt /etc/ssl/certs/cert.pem
        reqadd X-Forwarded-Proto:\ https
        rspadd Strict-Transport-Security:\ max-age=31536000
        default_backend b_publinet

backend b_yahoo
    balance roundrobin
    # Poor-man's sticky
    # balance source
    # JSP SessionID Sticky
    # appsession JSESSIONID len 52 timeout 3h
    option httpchk HEAD / HTTP/1.0
    option forwardfor
    option http-server-close
    server y1 77.238.184.150:80 maxconn 32 check
    server y2 188.125.73.108:80 maxconn 32 check


backend b_publinet
    balance roundrobin
    # Poor-man's sticky
    # balance source
    # appsession JSESSIONID len 52 timeout 3h
    http-request set-header Host publinet.ac-montpellier.fr
    reqirep  ^(GET|POST|HEAD)\ (.*)     \1\ /publinet/resultats\2
```

```
    acl response-is-redirect res.hdr(Location) -m found
    rspirep ^Location:\ (http|https)://\/(.*)    Location:\ \1://192.168.1.34/publinet/resultats\2
              if response-is-redirect
    option httpchk HEAD / HTTP/1.0
    option http-server-close
    server p1 195.83.225.163:80 maxconn 32 check
    server p2 195.83.225.164:80 maxconn 32 check


backend b_wwwiutbeziers
    option forwardfor
    option httpchk HEAD / HTTP/1.0
    option http-server-close
    server w1 194.199.227.80:80 maxconn 32 check


listen stats :1936
    mode http
    stats enable
    stats hide-version
    stats realm Haproxy\ Statistics
    stats uri /
    stats auth pouchou:pouchou

listen admin
    bind *:8080
    stats enable

/usr/sbin/haproxy -f /etc/haproxy/haproxy.cfg -d
[ALERT] 020/113657 (2647) : Error(s) found in configuration file : /etc/haproxy/haproxy.cfg
[ALERT] 020/113707 (2647) : Proxy 'f_wwwiutbeziers': unable to find required default_backend: 'b_wwwiutbeziers'.
[WARNING] 020/113707 (2647) : parsing [/etc/haproxy/haproxy.cfg:84] : backend 'b_publinet'.....

openssl req -x509 -newkey rsa:2048 -keyout key.pem -out ca.pem -days 1080 -nodes -subj '/CN=*/O=IUTBEZIERS RT licpro./C=FR'
cp key.pem cert.pem
cat ca.pem >> cert.pem


awk 1 ORS='\\n' cert.pem
-----BEGIN PRIVATE KEY-----\nMIIEvgIBADANBgkqhkiG9w0BAQEFAASCBKgwggSkAgEA
....
-----END CERTIFICATE-----\n

# testez les certificats
openssl s_client -connect example.com:443 -ssl3

http --verify=no https://192.168.20.134

HTTP/1.1 302 Found
Cache-Control: no-cache
Content-length: 0
Location: http://publinet.ac-montpellier.fr/publinet/resultats
Strict-Transport-Security: max-age=31536000
```

6 URLs intéressantes pour LVS

- https://www.server-world.info/en/note?os=Ubuntu_16.04&p=lvs
- https://www.server-world.info/en/note?os=Ubuntu_16.04&p=lvs&f=2
- https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Load_Balancer_Administration/s1-lvs-direct-VSA.html#s2-lvs-direct-arptables-VSA
- https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Load_Balancer_Administration/s2-lvs-direct-iptables-VSA.html