ULL | Universidad de La Laguna

S(TIC)

Balanceador basado en Software Libre (HAProxy)

Abilio Domingo Hernández
adomher@ull.edu.es

Motivaciones

- Problemas rendimiento balanceador Cisco en pruebas de carga automatrícula
- Mayor control de la solución
- Es una solución fácilmente escalable apoyado en Chef y máquinas virtuales
- Uso de hardware genérico
- Fácil implementación de entornos de preproducción
- Ahorro en costes de mantenimiento
- El conocimiento de la infraestructura se queda en la ULL
- Mayor número de funcionalidades
- Rendimiento HAProxy en entorno virtualizado
  - : 1000 clientes, 200 peticiones, 200mil muestras -› 9400 peticiones/seg

Situación previa

- 1 pareja de balanceadores Cisco ACE 4700 en modo activo-pasivo
- 1 pareja de balanceadores Cisco CSS 11500 (End Of Life en 2011)
- Política de balanceo roundrobin
- Sin terminación ssl en el balanceador
- Sin source nat
- Sin soporte del fabricante (por ahorro de costes)

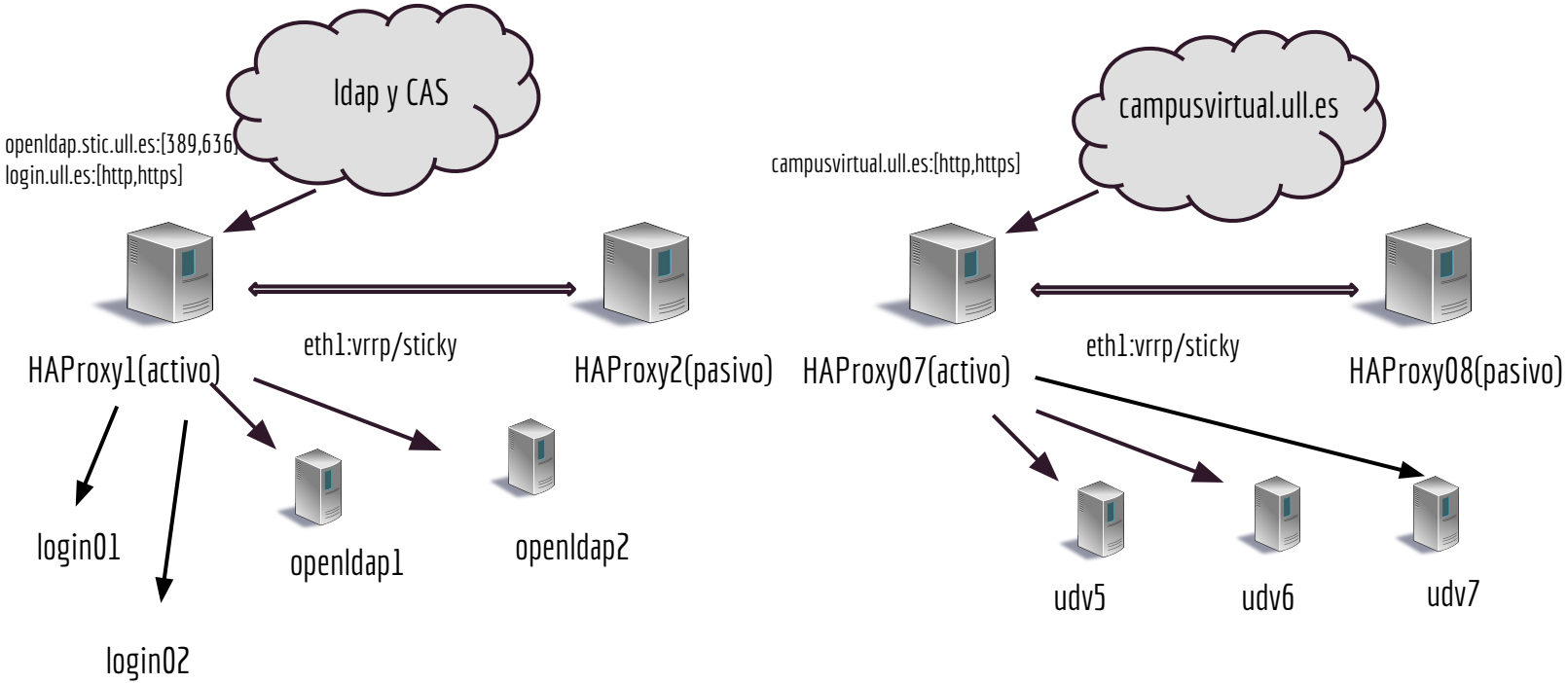
Situación actual

- 3 parejas de nodos virtualizados VMware
  - 2 cores 2.2GHz
  - 4 GB RAM
  - 40GB de disco duro
- HAProxy 1.5.14
- HA mediante Keepalived 1.2.13 y replicación de tablas sticky de HAProxy en red aislada
- Cada pareja de nodos HAProxy en modo activo/pasivo
- Política de balanceo roundrobin
- Sin source nat
- Configuración gestionada a través de Chef
- Monitorización mediante Nagios y grafana
- +13M sesiones balanceadas

Servicios balanceados actualmente HAProxy

- Campusvirtual ULL
- Servicio de autenticación centralizada (CAS)
- Openldap
- Páginas webs de los usuarios
- Próximamente:
  - Correo electrónico
  - Web institucional

Servicios balanceados actualmente HAProxy



Universidad de La Laguna

Balanceador basado en Software Libre (HAProxy)

S(TIC)

Interfaz web

> General process information

pid = 1861 (process #1, nbproc = 1)
uptime = 12d 16h43m42s
system limits: memmax = unlimited; ulimit-n = 10291
maxsock = 10291; **maxconn** = 4096; **maxpipes** = 1024
current conns = 149; current pipes = 7/11; conn rate = 12/sec
Running tasks: 1/166; idle = 98 %

- active UP
- active UP, going down
- active DOWN, going up
- active or backup DOWN
- active or backup DOWN for maintenance (MAINT)
- active or backup SOFT STOPPED for maintenance
- backup UP
- backup UP, going down
- backup DOWN, going up
- not checked

Note: "NOLB"/"DRAIN" = UP with load-balancing disabled.

Display option:
- Scope :
- Hide 'DOWN' servers
- Refresh now
- CSV export

External resources:
- Primary site
- Updates (v1.5)
- Online manual

campusvirtual_ull_es_http

		Queue			Session rate			Sessions						Bytes		Denied		Errors			Warnings		Server								
		Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
	Frontend				0	36	-	3	36	2 000	114 314			54 443 424	54 622 221	0	0	12					OPEN								
	udv_web5	0	0	-	0	21		1	19	-	36 622	21 832	18s	17 884 537	17 304 006		0		0	0	2	1	12d4h UP	L7OK/200 in 1ms	1	Y	-	3	1	31m	-
	udv_web6	0	0	-	0	23		0	20	-	40 202	22 038	19s	18 523 154	19 130 229		0		0	0	0	0	12d16h UP	L7OK/200 in 1ms	1	Y	-	0	0	0s	-
	udv_web7	0	0	-	0	21		2	20	-	37 424	21 834	3s	18 035 733	18 187 986		0		1	0	0	0	12d4h UP	L7OK/200 in 1ms	1	Y	-	3	1	30m	-
	Backend	0	0		0	36		3	36	200	114 314	65 704	3s	54 443 424	54 622 221	0	0		1	0	2	1	12d16h UP		3	3	0		0	0s	

Choose the action to perform on the checked servers : Apply

campusvirtual_ull_es_https

		Queue			Session rate			Sessions						Bytes		Denied		Errors			Warnings		Server								
		Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
	Frontend				12	67	-	143	506	2 000	7 037 163			60 403 485 465	554 245 997 763	0	0	138					OPEN								
	udv_web5	0	0	-	10	47		40	218	-	2 337 598	182 723	0s	19 922 100 315	185 489 011 171		0		28	2 066	74	42	12d5h UP	L7OK/200 in 2ms	1	Y	-	3	1	30m45s	-
	udv_web6	0	0	-	0	51		53	327	-	2 308 365	184 344	2s	21 373 725 154	177 003 532 346		0		16	2 229	88	49	12d16h UP	L7OK/200 in 4ms	1	Y	-	0	0	0s	-
	udv_web7	0	0	-	2	52		50	208	-	2 390 246	182 774	1s	19 107 659 996	191 753 454 246		0		22	1 858	102	56	12d4h UP	L7OK/200 in 3ms	1	Y	-	3	1	30m	-
	Backend	0	0		12	67		143	506	200	7 037 163	549 841	0s	60 403 485 465	554 245 997 763	0	0		66	6 153	264	147	12d16h UP		3	3	0		0	0s	

Choose the action to perform on the checked servers : Apply

admin

	Queue			Session rate			Sessions						Bytes		Denied		Errors			Warnings		Server								
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
Frontend				0	3	-	3	3	2 000	17 643			2 350 498	47 375 713	0	0	2					OPEN								
Backend	0	0		0	0		0	0	200	0	0	0s	2 350 498	47 375 713	0	0		0	0	0	0	12d16h UP		0	0	0		0		

Universidad de La Laguna

Balanceador basado en Software Libre (HAProxy)

S(TIC)

Interfaz web

openldap_389

	Queue			Session rate			Sessions					Bytes		Denied		Errors			Warnings		Server								
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
Frontend				0	70	-	174	259	2 000	839 936			2 946 762 660	8 222 432 130	0	0	0					OPEN								
openldap1	0	0	-	0	31		91	157	-	418 894	418 894	5s	1 473 220 492	4 100 908 032		0		1	444	0	1	12d50m UP	L7OK/0 in 1ms	1	Y	-	0	1	53m13s	-
openldap2	0	0	-	0	70		83	248	-	421 045	421 045	10s	1 473 542 168	4 121 524 098		0		0	485	0	2	12d2m UP	L7OK/0 in 1ms	1	Y	-	0	1	42m49s	-
Backend	0	0		0	70		174	259	200	839 936	839 939	5s	2 946 762 660	8 222 432 130	0	0		1	929	0	3	12d16h UP		2	2	0		0	0s	

Choose the action to perform on the checked servers : Apply

openldap_636

	Queue			Session rate			Sessions					Bytes		Denied		Errors			Warnings		Server								
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
Frontend				0	43	-	80	215	2 000	263 855			1 024 169 417	23 479 773 915	0	0	0					OPEN								
openldap1	0	0	-	0	25		51	148	-	131 582	131 582	2s	616 390 029	12 085 077 301		0		0	9	0	0	12d50m UP	L7OK/0 in 1ms	1	Y	-	0	1	53m12s	-
openldap2	0	0	-	0	25		29	159	-	132 271	132 271	8s	407 779 388	11 394 696 614		0		0	53	0	0	12d2m UP	L7OK/0 in 2ms	1	Y	-	0	1	42m49s	-
Backend	0	0		0	50		80	211	200	263 855	263 853	2s	1 024 169 417	23 479 773 915	0	0		0	62	0	0	12d16h UP		2	2	0		0	0s	

Choose the action to perform on the checked servers : Apply

login_http

	Queue			Session rate			Sessions					Bytes		Denied		Errors			Warnings		Server								
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
Frontend				0	8	-	1	9	2 000	26 767			5 791 435	10 096 491	0	0	0					OPEN								
login01	0	0	-	0	5		0	6	-	3 025	701	1m	635 358	868 381		0		0	0	0	0	4d2h UP	L7OK/301 in 1ms	1	Y	-	3	5	7d9h	-
login02	0	0	-	0	3		1	9	-	8 031	619	10s	903 958	1 921 953		0		0	0	0	0	4d2h UP	L7OK/301 in 0ms	1	Y	-	6	5	7d12h	-
login3	0	0	-	0	7		0	6	-	4 097	602	5d6h	2 170 158	2 826 583		0		0	0	0	0	5d6h MAINT		1	Y	-	3	8	5d8h	-
login4	0	0	-	0	6		0	6	-	11 604	586	5d6h	2 081 961	4 479 574		0		0	0	0	0	5d6h MAINT		1	Y	-	0	8	5d8h	-
Backend	0	0		0	8		1	9	200	26 767	2 508	10s	5 791 435	10 096 491	0	0		0	0	0	0	12d16h UP		2	2	0		0	0s	

Choose the action to perform on the checked servers : Apply

login_https

	Queue			Session rate			Sessions					Bytes		Denied		Errors			Warnings		Server								
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
Frontend				3	39	-	98	178	2 000	1 955 263			3 915 021 429	11 351 569 491	0	0	143					OPEN								
login01	0	0	-	1	25		47	166	-	350 849	91 469	1s	936 148 689	2 369 637 083		0		109	72	7	341	4d2h UP	L7OK/301 in 2ms	1	Y	-	3	5	7d9h	-
login02	0	0	-	1	28		51	136	-	320 311	83 365	1s	860 213 282	2 272 408 899		0		112	75	10	410	4d2h UP	L7OK/301 in 2ms	1	Y	-	6	4	7d12h	-
login3	0	0	-	0	39		0	76	-	643 008	124 285	5d6h	1 003 790 183	3 187 661 712		0		0	358	6	3	5d6h MAINT		1	Y	-	3	8	5d8h	-
login4	0	0	-	0	27		0	77	-	641 469	123 669	5d6h	1 114 869 275	3 521 861 797		0		0	349	4	2	5d6h MAINT		1	Y	-	0	8	5d8h	-
Backend	0	0		3	39		98	178	200	1 955 263	422 788	1s	3 915 021 429	11 351 569 491	0	0		221	854	27	756	12d16h UP		2	2	0		0	0s	

Universidad de La Laguna

Balanceador basado en Software Libre (HAProxy)

S(TIC)

Estadística login.ull.es:https 24horas

Universidad de La Laguna

Balanceador basado en Software Libre (HAProxy)

S(TIC)

Estadística campusvirtual.ull.es:https 24horas

¿Cómo gestionamos un nodo?

- Asignar parámetros dns
- Instalación desatendida de la máquina
- Mediante Chef
  - Bootstrap del nodo con el rol stic_base
  - Cookbook stic_HAProxy
    - Atributo node_role
    - La cookbook instala los paquetes necesarios de HAProxy y keepalived
    - Busca sus peers de keepalived
    - Configura reglas shorewall y routing avanzado
    - Aplica templates de servicios balanceados
    - Aplica parámetros de sysctl
- Tiempo de despliegue del nodo: 30 minutos
- Gestión de los real servers mediante interfaz web de HAProxy

Ejemplo template

```
global
  log 10.107.26.47 local1 info
  log 10.107.26.24 local1 info
  maxconn 4096
  user root
  group root
  tune.ssl.default-dh-param 2048
  ca-base /etc/ssl/certs
  crt-base /etc/ssl/certs
  max-spread-checks 1
  stats socket /run/admin.sock mode 600 level admin

defaults
  log     global
  mode    http
  retries 3
  timeout client 300s
  timeout connect 30s
  timeout server 300s
  #option dontlog-normal
  option httplog
  option redispatch
  option splice-auto
  option splice-request
  option splice-response
  balance  roundrobin


peers mypeers
  peer haproxy07 10.107.88.42:1026
  peer haproxy08 10.107.88.43:1026
  <% if node.chef_environment == 'kitchen' %>
  peer <%= node['hostname'] %> 127.0.0.1:1026
  <% end %>

# Set up application listeners here.

listen campusvirtual_ull_es_http
  bind 193.145.118.61:80 interface eth2 transparent
  #redirect scheme https code 301 if !{ ssl_fc }
  option httpchk OPTIONS * HTTP/1.1\r\nHost:\ www.campusvirtual.ull.es
  mode tcp
  option tcplog
  server udv_web5 193.145.118.44:10080 check inter 15s
  server udv_web6 193.145.118.45:10080 check inter 15s
  server udv_web7 193.145.118.46:10080 check inter 15s
  stick-table type ip size 10k expire 600s store gpc0 peers mypeers
  stick on src ipmask(32)      #Hacemos sticky por ip de origen
  source 0.0.0.0 usesrc client #Hacemos proxy transparente, no se natea source ip

listen campusvirtual_ull_es_https
  bind 193.145.118.61:443 interface eth2 transparent
  #redirect scheme https code 301 if !{ ssl_fc }
  option httpchk OPTIONS * HTTP/1.1\r\nHost:\ www.campusvirtual.ull.es
  mode tcp
  option tcplog
  stick-table type ip size 10k expire 600s store gpc0 peers mypeers
  stick on src ipmask(32)      #Hacemos sticky por ip de origen
  server udv_web5 193.145.118.44:10443 check check-ssl verify none inter 15s
  server udv_web6 193.145.118.45:10443 check check-ssl verify none inter 15s
  server udv_web7 193.145.118.46:10443 check check-ssl verify none inter 15s
  source 0.0.0.0 usesrc client #Hacemos proxy transparente, no se natea source ip

```

HA

- Implementado mediante keepalived en configuración máster-backup
- Interfaz dedicada en red aislada para el intercambio de mensajes vrrp unicast
- El proceso de keepalived hace tracking de las interfaces de la máquina y de que el servicio HAProxy esté arrancado
- En caso de fallo la conmutación es rápida ya que el nodo que está en modo backup también tiene el servicio HAProxy arrancado
- Cuando se produce una conmutación los nodos mandan una alerta por email

Shorewall

- Diferentes zonas de seguridad
  - Gestión
  - HA
  - Balanceo
- Sólo políticas específicas para la interfaz de administración de la máquina y ha
- Para las interfaces que proporcionan servicios balanceados, zona de balanceo de shorewall, la política por defecto es ACCEPT
- Manipulación de las tablas de rutas para advanced routing y solventar el problema del routing asimétrico

¿ Advanced routing ?

- Problema de routing asimétrico:
  - Los real servers tienen como puerta de enlace el firewall institucional
  - El tráfico de los servicios balanceados llega por la interfaz de HAProxy y el real server lo envía a través de su default gateway
- Solución:
  - Tablas de routing alternativas
  - Marcado de paquetes
  - Uso de fichero providers de shorewall

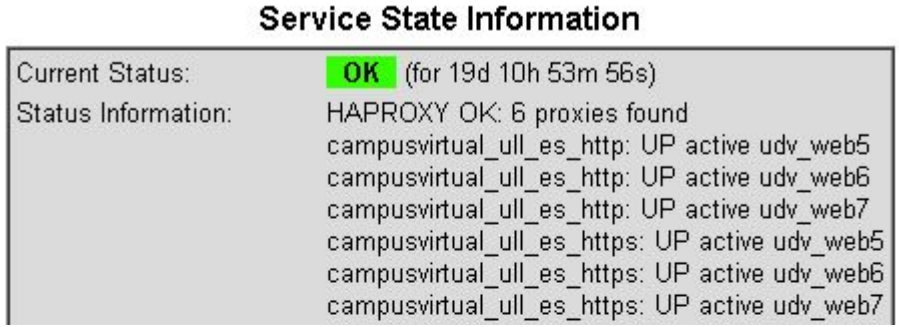
¿ Advanced routing ?

/etc/shorewall/providers

```
1  # Generated by Chef for haproxy2.stic.ull.es
2  # Local modifications will be overwritten.
3  # Bibliografia: http://shorewall.net/manpages/shorewall-providers.html
4
5
6  #
7  #NAME  NUMBER MARK DUPLICATE INTERFACE GATEWAY OPTIONS COPY
8  HAproxy 100      -       -       lo        -     tproxy -
9
10 eth2    101     101     main    eth2    detect  track   eth2
11
12 eth3    102     102     main    eth3    detect  track   eth3
```

Monitorización

- Chequeos en nagios
  - CPU
  - Disco
  - Memoria
  - Estado del servicio
  - Estado de los real servers
- Grafana



Service State Information

Current Status:	OK (for 19d 10h 53m 56s)
Status Information:	HAPROXY OK: 6 proxies found campusvirtual_ull_es_http: UP active udv_web5 campusvirtual_ull_es_http: UP active udv_web6 campusvirtual_ull_es_http: UP active udv_web7 campusvirtual_ull_es_https: UP active udv_web5 campusvirtual_ull_es_https: UP active udv_web6 campusvirtual_ull_es_https: UP active udv_web7

Dificultades

- La curva de aprendizaje de chef
- Implantar una infraestructura nueva desde cero sin el conocimiento previo
- Al principio complicamos demasiado la gestión de la configuración con chef
- No hay réplica de la tabla de conexiones en caso de failover del nodo activo, se perdió mucho tiempo investigando
- Al principio complicamos demasiado las reglas de shorewall
- Implementación de advanced routing a través de shorewall
- Trabajar en una herramienta web para la gestión unificada de las interfaces de HAProxy
- Las propias de la migración de los servicios del esquema con balanceador Cisco a HAProxy

ULL | Universidad de La Laguna

S(TIC)

Balanceador basado en Software Libre (HAProxy)

Abilio Domingo Hernández
adomher@ull.edu.es