HAproxy v praxi

zkušenosti z integrace a provozu

Michal Rychlík

InstallFest
6.3.2016

Obsah

- Základy
- Jak to bylo, je a bude v LMC
- Migrace
- Monitoring
- Ukázka konfigurace
- Novinky ve verzi 1.6

Co je HAPROXY a kdo to používá

- Transparentní http / tcp proxy
- Load-balancer

Virgin america

stackoverflow

serverfault

GitHub

YOUPORN

Základní funkce

- tcp / http load-balancer
- SSL terminace
- healt check backend serverů
- nerovnoměrné rozdělení provozu
- socket api

Základní funkce

- modifikace / vkládání hlaviček
- balancování podle dotazu
  - url, hlavičky, zdrojová IP
- session affinity
  - zdrojová IP, hash url, cookie
- logování
- statistiky

Plus / minus

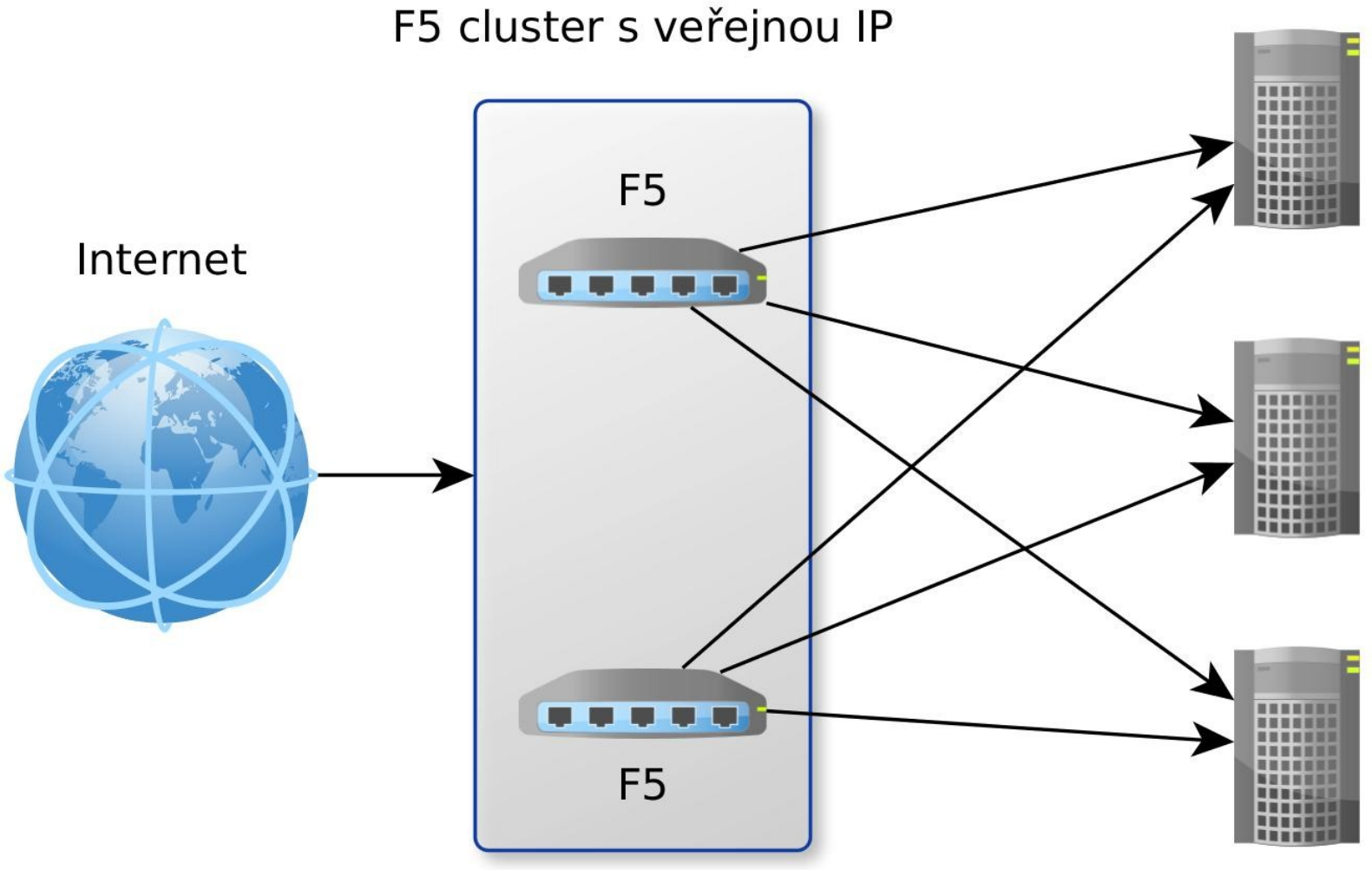
- ✔ openSource
- ✔ SW
- ✔ škálování
- ✔ napojení na puppet, consul
- ✔ Hapee
- ✔ Aloha

- ✗ nastavení Kernelu
- ✗ ve VM může zatížení HV brzdit provoz
- ✗ neudrží nastavení z api při reload

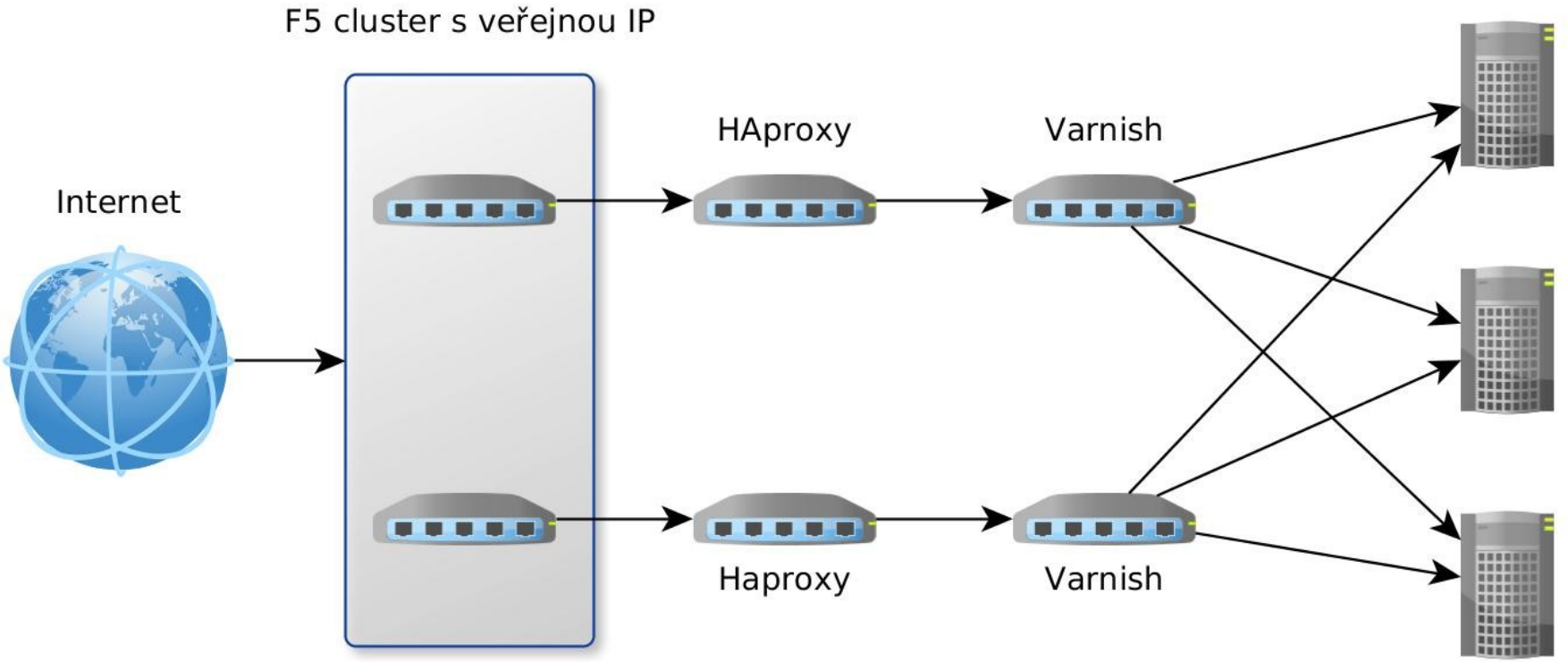
Specifika prostředí LMC

- produkce - 88 serverů v 39 poolech
- přes LB pouze aplikační servery
- provoz přes front balancer cca 100Mbit/s
- 10 testovacích prostředí

Minulost - F5



Minulost: F5 → HAproxy → Varnish



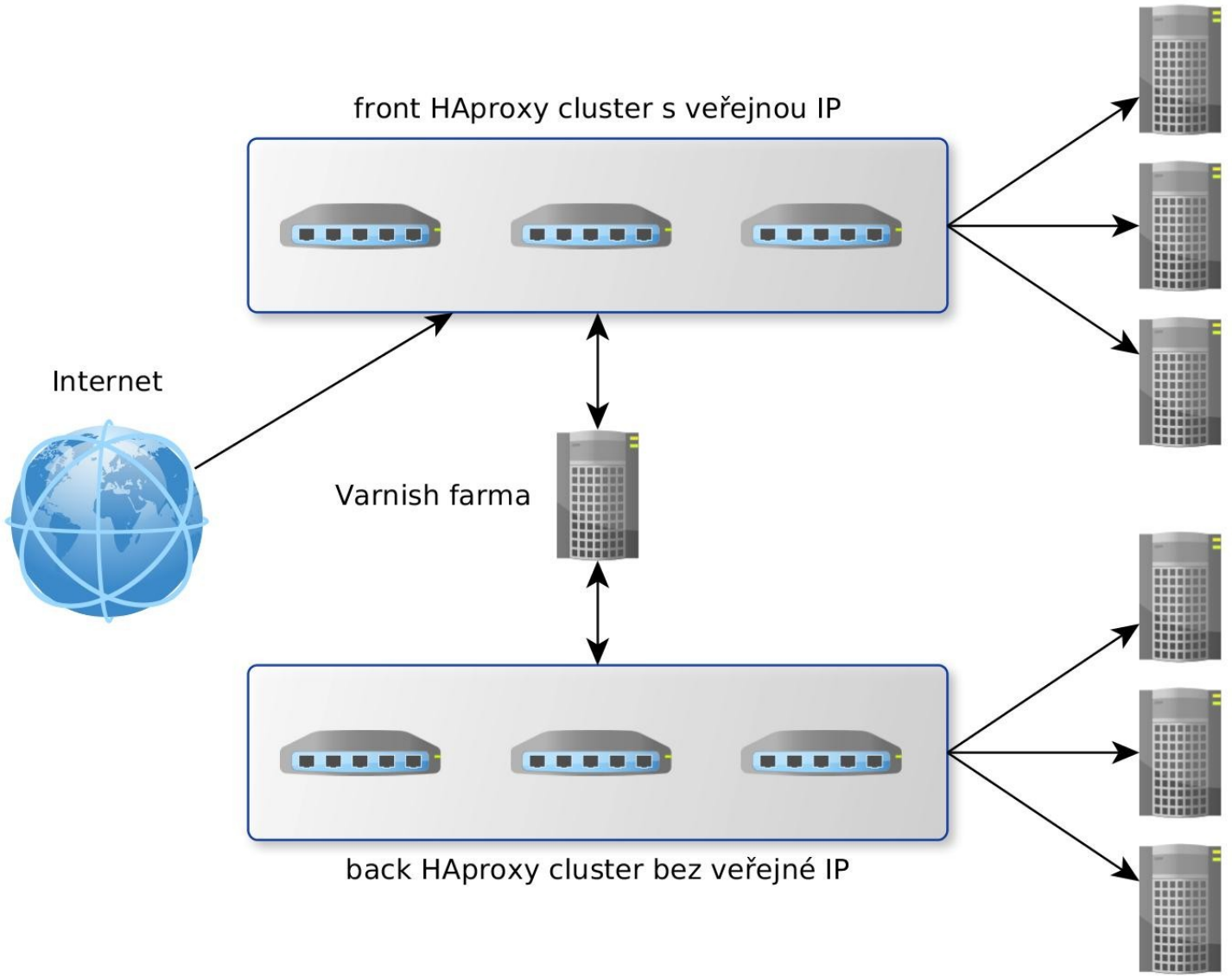
Proč migrace na HAproxy?

- Migrace do cloudu (openstack)
- spousta vývojových prostředí
- dynamické vytváření a konfigurace
- Více možností oproti varnish
  - online statistiky, api, SSL, cookie session affinity
- Corosync HAproxy = možnost vypnutí F5
- Nástroje pro monitoring a grafy

Vývoj konfigurace a testování

- přepis pravidel z varnishe
  - směrujeme pomocí hlaviček
- na integračním prostředí
- testovány jednotlivé služby
- nejčastější problémy
  - dotaz padá na jiný pool
  - krátký timeout
  - healt checky plní logy (default metoda options)
  - testy nezasílají cookie (session affinity)
- 3 měsíce

Současnost: HAproxy



Migrace

- backend → frontend
- po blokách 4 poolů
- po skupinách ve veřejných DNS
- 6 týdnů

Výsledek

- Snížena odezva našich webů
- Výrazně lepší přehled co se děje
- Méně prvků na cestě
- Lepší session affinity
- Snadnější konfigurace
  - například snadné zablokování url které mají být dostupné jen z vnitřní sítě

Problémy

- bug v init skriptu → neukončující se procesy
- rozpad clusteru kvůli DNS

- 502 bad gateway
- php opcache
- tcp_tw_reuse, tcp_tw_recycle
- SNI a default certifikát
- corosync – dva nody se stejnou IP po rozpadu sítě

Monitoring

graphite

collectd

Grafana

logstash

kibana

elastic

Statistiky

Statistics Report for pid 19192

> General process information

pid = 19192 (process #1, nbproc = 1)
uptime = 0d 21h11m58s
system limits: memmax = unlimited; ulimit-n = 64089
maxsock = 64089; **maxconn =** 32000; **maxpipes =** 0
current conns = 6721; current pipes = 0/0; conn rate = 240/sec
Running tasks: 1/6776; idle = 81 %

- 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

all_in_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				1	646	-	1	60	16 000	107 115			29 878 730	2 412 284 947	0	0	0					OPEN								

apache

	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
ap01:80	0	0	-	6	120		0	60	-	117 598	103 888	1s	26 030 684	2 194 356 333		0		0	0	0	0	17h47m UP	L7OK/200 in 1ms	1	Y	-	0	0	0s	-
ap02:80	0	0	-	0	0		0	0	-	0	0	?	0	0		0		0	0	0	0	17h47m DOWN	L4TOUT in 2001ms	1	Y	-	1	1	17h47m	-
Backend	0	0		6	120		0	60	3 200	117 598	103 888	1s	26 030 684	2 194 356 333	0	0		0	0	0	0	17h47m UP		1	1	0		0	0s	

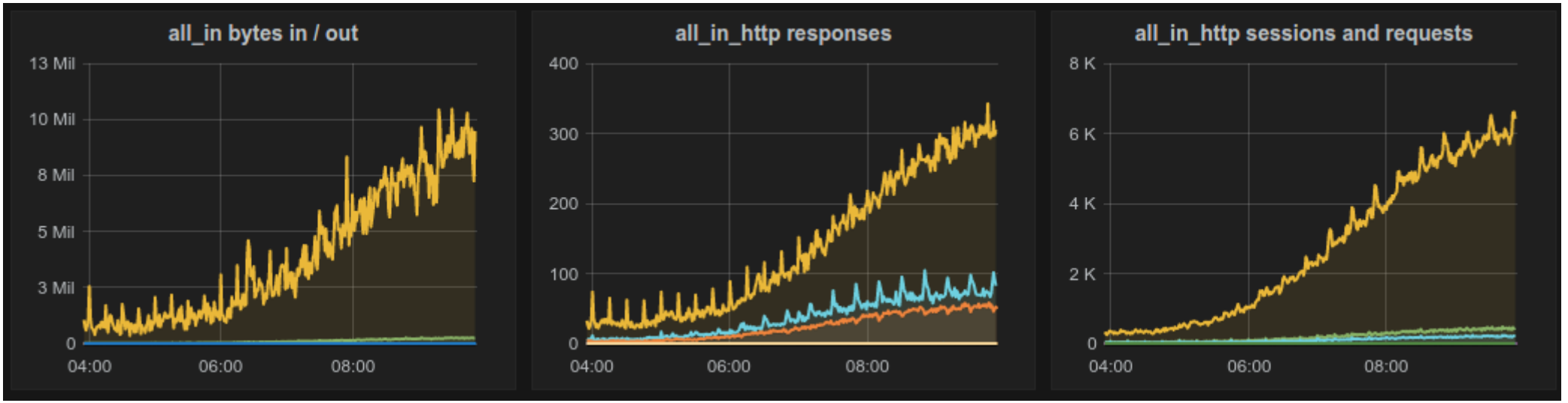
varnish

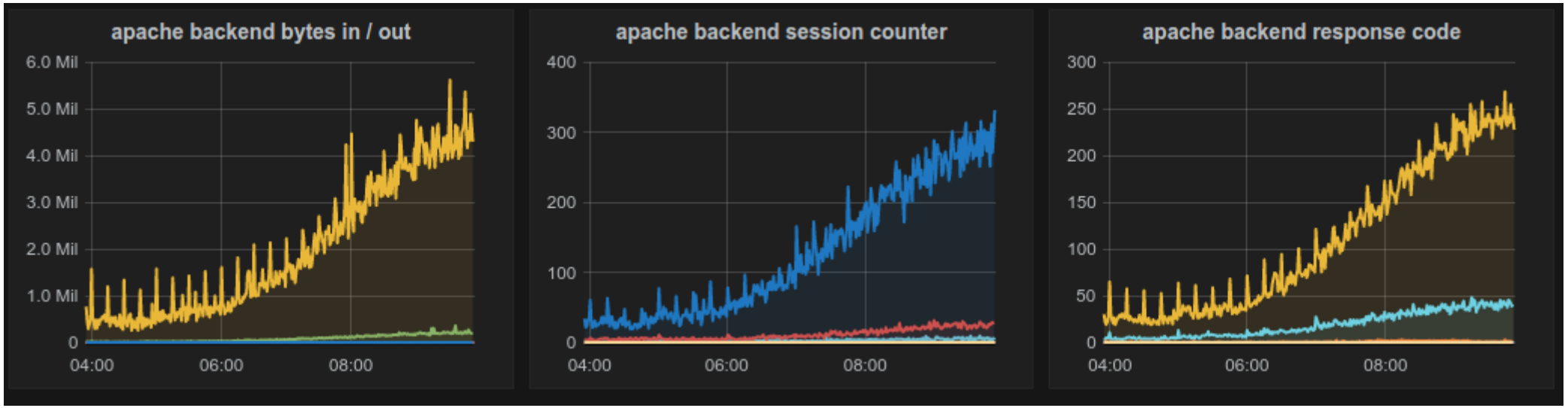
	Queue			Session rate			Sessions						Bytes		Denied		Errors			Warnings		Server								
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle
localhost:80	0	0	-	0	0		0	0	-	0	0	?	0	0		0		0	0	0	0			1	-	Y				-
varnish1:80	0	0	-	0	612		0	26	-	23 200	23 200	6s	9 643 160	627 969 251		0		0	0	0	0	17h47m UP	L7OK/200 in 1ms	1	Y	-	0	0	0s	-
Backend	0	0		0	612		0	27	3 200	23 200	23 200	6s	9 643 160	627 969 251	0	0		0	0	0	0	17h47m UP		1	1	1		0	0s	

Logy

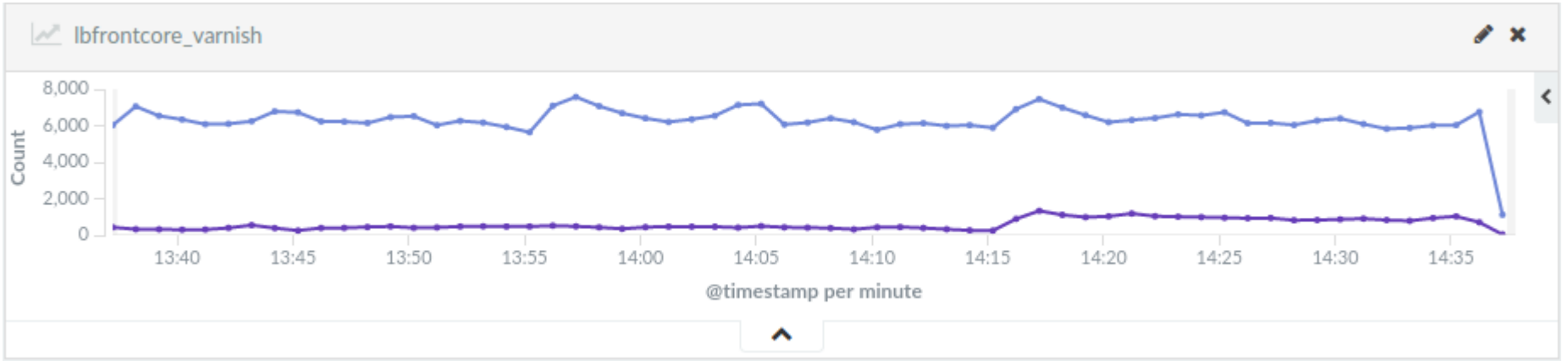
```
46.165.195.139:34890 [31/Dec/2015:00:43:00.157] all_in_http apache/ap03.prod.lmc.cz:80
0/0/0/-1/269 502 5423 - - SHNN 1043/970/4/1/0 0/0 "GET / HTTP/1.1"

77.48.31.6:63241 [06/Mar/2016:10:24:01.183] all_in_http apache/ap04.prod.lmc.cz:80
0/0/0/2/2 200 1696 - - --VN 4018/3605/14/4/0 0/0 {Mozilla/5.0 (Windows NT 10.0; Win64;
x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/46.0.2486.0 Safari/537.36
Edge/13.10586|exporter.lmc.cz} "GET /blesk-new.htm HTTP/1.1"
```

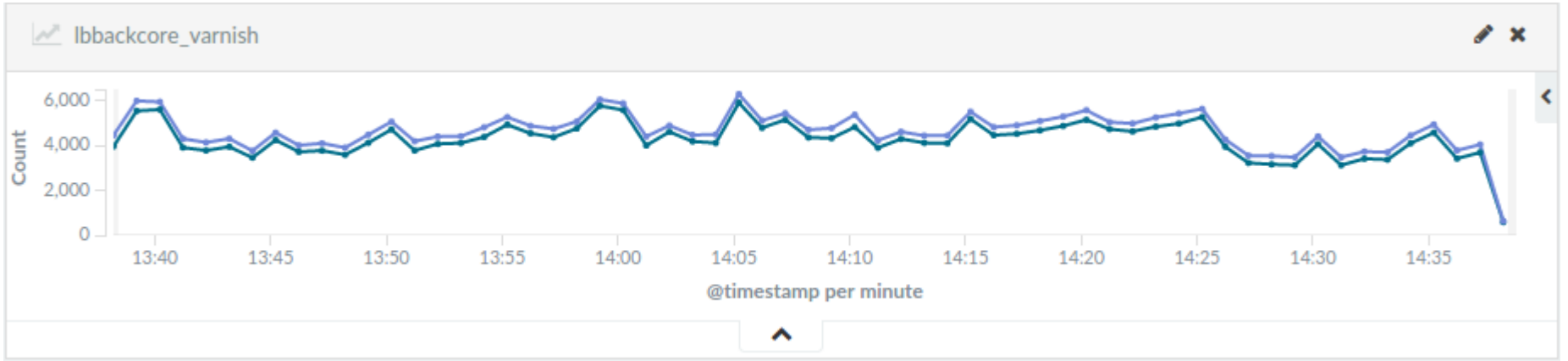
Grafy ze statistik



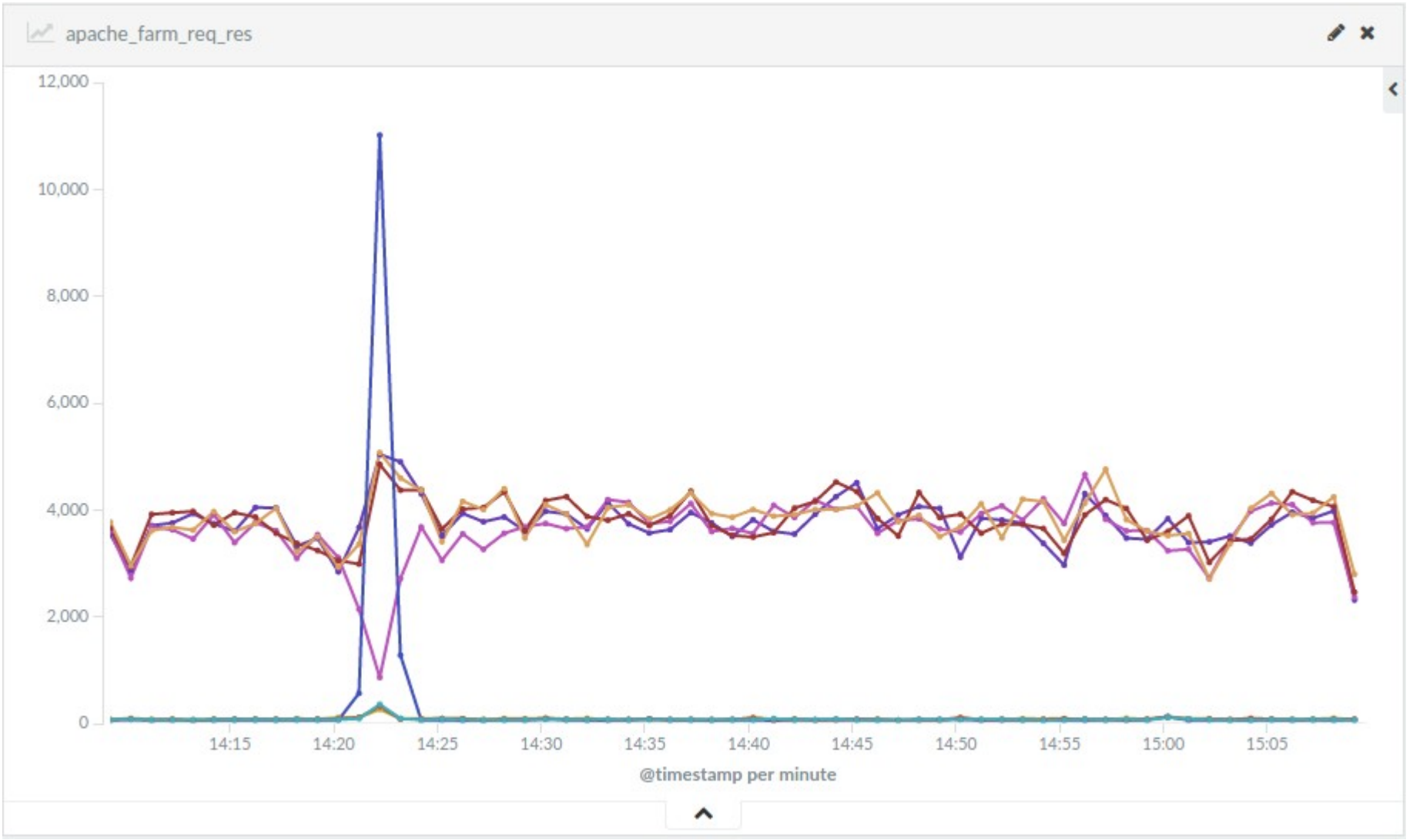


Kibana

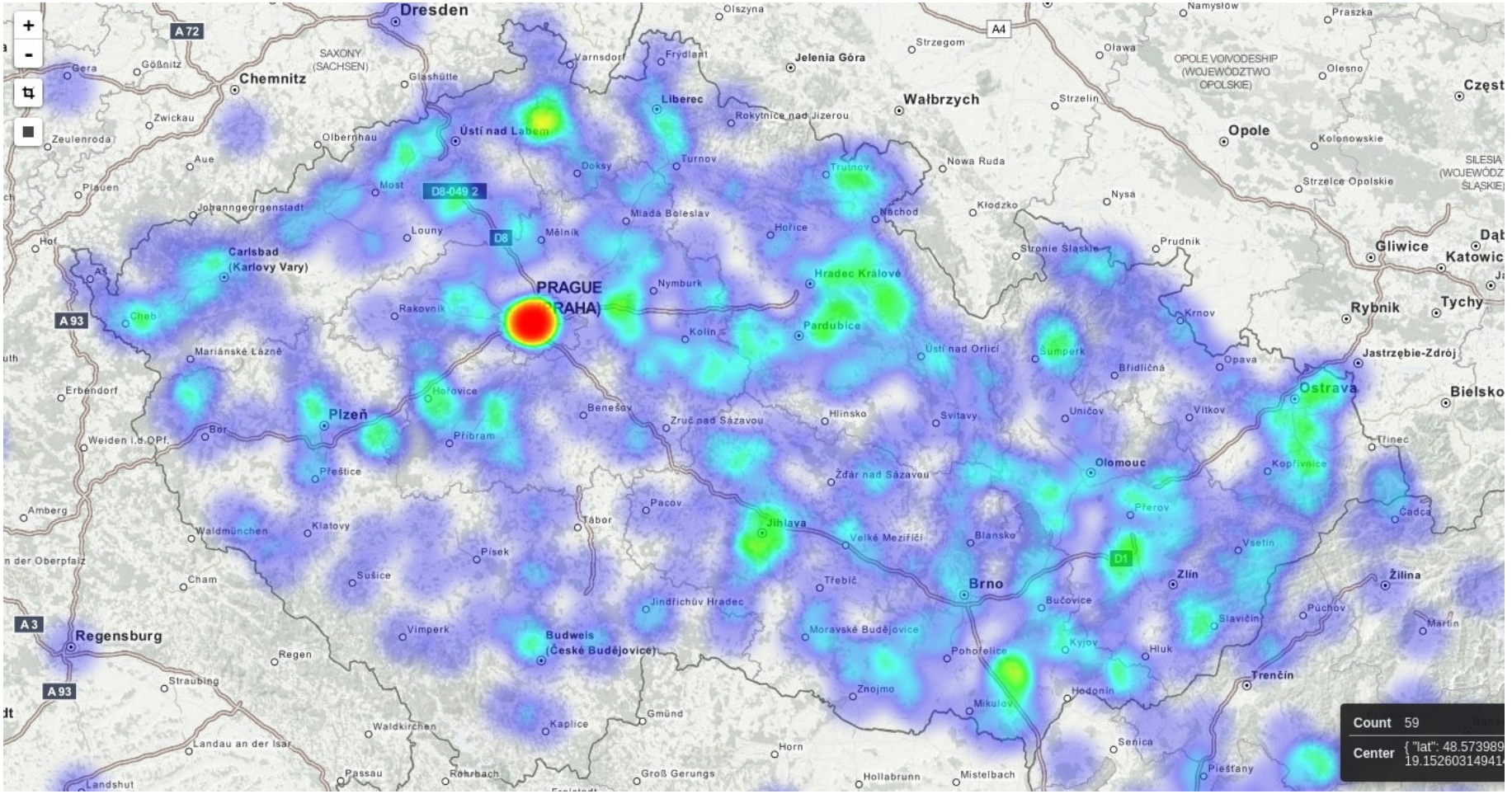

lbfrontcore_varnish
8,000
6,000
4,000
2,000
0
Count
13:40
13:45
13:50
13:55
14:00
14:05
14:10
14:15
14:20
14:25
14:30
14:35
@timestamp per minute


lbbackcore_varnish
6,000
4,000
2,000
0
Count
13:40
13:45
13:50
13:55
14:00
14:05
14:10
14:15
14:20
14:25
14:30
14:35
@timestamp per minute

Kibana


apache_farm_req_res
12,000
10,000
8,000
6,000
4,000
2,000
0
14:15
14:20
14:25
14:30
14:35
14:40
14:45
14:50
14:55
15:00
15:05
@timestamp per minute

Kibana


Dresden
Chemnitz
Liberec
Jelenia Góra
Wałbrzych
Opole
PRAGUE
(PRAHA)
Plzeň
Ostrava
Brno
Olomouc
Zlín
Regensburg
Budweis
(České Budějovice)
Carlsbad
(Karlovy Vary)
Ústí nad Labem
Hradec Králové
Pardubice
Jihlava
Gliwice
Rybnik
Tychy
Bielsko
Katowic
Częst
Žilina
Trenčín
Count 59
Center { "lat": 48.573989
19.15260314941

Základní konfigurace

- global
- defaults
- frontend
- backend
- listener

Základní konfigurace - global

```
global
  chroot  /var/lib/haproxy
  crt-base  /etc/haproxy/ssl
  daemon
  group  haproxy
  user  haproxy
  log  10.0.36.64 local0
  maxconn  32000
  pidfile  /var/run/haproxy.pid
  ssl-default-bind-ciphers
ECDH+AESGCM:DH+AESGCM:ECDH+AES256:DH+AES256:ECDH+AES128:DH+AES:ECDH+3DES:DH+3DES:RSA+AESGC
M:RSA+AES:RSA+3DES:!aNULL:!MD5:!DSS:!AES256
  stats  socket /var/lib/haproxy/stats mode 660 level admin
```

Základní konfigurace - defaults

```
defaults
  compression  algo gzip
  compression  type text/* application/xml application/x-javascript application/json
  errorfile  400 /etc/haproxy/errors/400.http
  errorfile  403 /etc/haproxy/errors/403.http
  errorfile  408 /dev/null
  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
  log  global
  maxconn  16000
  mode  http
  option  httplog
  option  dontlognull
  # option  dontlog-normal
  retries  3
  stats  enable
  timeout  connect 10000ms
  timeout  client 50000ms
  timeout  server 50000ms
  timeout  check 10000ms
```

Základní konfigurace - frontend

```
frontend all_in_https
  bind 0.0.0.0:443,:10443 ssl crt lblmccz.pem crt . no-sslv3
  capture  request header User-Agent len 200
  capture  request header Host len 40
  default_backend  apache
  acl  https ssl_fc
  acl  secured_cookie res.hdr(Set-Cookie),lower -m sub secure
  rsprep  ^(Set-Cookie:\ JSESSIONID.*|Set-Cookie:\ G2_SID.*) \1;\ Secure if https !
secured_cookie
  acl  pdxml path_beg /pdxml
  acl  g2 hdr_beg(host) g2.lmc.cz
  acl  private_net src 10.0.0.0/8
  http-request  deny if !private_net pdxml g2
  acl  poradna_url path_beg /poradna
  acl  poradna_host hdr_beg(host) www.jobs.cz www.topjobs.sk www.prace.cz
  http-request  add-header X-Rule poradna if poradna_url poradna_host
  use_backend  apache_poradna if { hdr(X-Rule) poradna }
  http-request  set-header X-Forwarded-Proto https
  reqidel  ^X-Forwarded-Proto:.*
```

Základní konfigurace - backend

```
backend apache
  balance  roundrobin
  cookie  APID insert indirect
  mode  http
  option  httplog
  option  http-server-close
  option  forwardfor
  option  httpchk  HEAD /isEnabled
  timeout check  5s
  server ap01.deploy.lmc.cz:80 ap01.deploy.lmc.cz:80 cookie ap01.deploy.lmc.cz:80 check
  server ap02.deploy.lmc.cz:80 ap02.deploy.lmc.cz:80 cookie ap02.deploy.lmc.cz:80 check
```

Co je nového ve verzi 1.6

- uvozovky v konfiguračních souborech
- Lua (možnost Let's Encrypt)
- log tag
- dns překlad za běhu
- možnost používat proměnné v konfiguraci
- email notifikace
- zpracování těla http stránek
- ochrana proti slow-post útoku
- ukládání stavu serverů při reloadu
- externí healt checky

Děkuji za pozornost

michal@rychlik.it