

Load-Balancing für den schmalen Geldbeutel

Thomas Heil <heil@terminal-consulting.de>

01.12.2011



Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion
von Load-Balancing

Notwendigkeit von
Load-Balancing

Notwendigkeit von
Load-Balancing II

Notwendigkeit von
Load-Balancing III

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Überblick

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion von Load-Balancing

Notwendigkeit von Load-Balancing

Notwendigkeit von Load-Balancing II

Notwendigkeit von Load-Balancing III

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

- Wer bin ich
- Warum dieses Thema?
- Aufgaben und Funktion von Load-Balancing
- Notwendigkeit von Load-Balancing
- Load-Balancing Techniken
- Wie wähle ich den richtigen LB aus?
 - Hardware vs. Software
 - oder beides? kommerziell oder frei
- Die grosse Herausforderung, Die Anwendung hinter dem LB
- Der schmale Geldbeutel oder die Open Source Philosophie
- Beispiele / Anwendungsszenarien
- Praxisteil, Fazit, Fragen?

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion
von Load-Balancing

Notwendigkeit von
Load-Balancing

Notwendigkeit von
Load-Balancing II

Notwendigkeit von
Load-Balancing III

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

■ Administrator bei der bookandsmile GmbH

- <http://www.billigflieger.de>
- Load Balancing
- Monitoring
- Virtualisierung

■ OpenSouce Consultant

- OpenSource Consultant
- VPN, Proxies, Firewall
- Virtualisierung, Verteilung, Überwachung
- Load Balancing, HA

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion
von Load-Balancing

Notwendigkeit von
Load-Balancing

Notwendigkeit von
Load-Balancing II

Notwendigkeit von
Load-Balancing III

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

■ OpenWrt Entwickler

- Package Maintainer für HAProxy, Openswan, memcached, pound, varnish

■ OpenSource Enthusiast

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion
von Load-Balancing

Notwendigkeit von
Load-Balancing

Notwendigkeit von
Load-Balancing II

Notwendigkeit von
Load-Balancing III

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

- erste Versuche mit Debian
 - HAProxy, keepalived, iptables
- OpenBSD
 - HAProxy, CARP, PF
- Installation, Betrieb und Wartung äusserte sich zeitintensiv und schwer nachvollziehbar
- Einarbeitung der Kollegen oft schwierig
- Neuinstallation nach Hardwarschaden benötigte etwa einen Tag
- Suche nach dem absoluten Minimum (bzgl. Installation, Sicherung und Wiederherstellung)
- nachvollziehbaren Anpassungen
- zentralem Accounting, Verkehrsfluss

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion
von Load-Balancing

Notwendigkeit von
Load-Balancing

Notwendigkeit von
Load-Balancing II

Notwendigkeit von
Load-Balancing III

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

- ein LB gehört in jede Landschaft, die Webserver bzw Webservices anbietet, selbst bei sehr kleinen Setups
- auch intern sind LB durchaus hilfreich
 - Balancing von Mysql-Slaves unter Berücksichtigung von Slave Delays
- vollständige Realisierung mit freien Bausteinen (OpenSource)
- damit jeder Zugang zu Technologie und Verständnis erlangen kann
- Geld sollte hauptsächlich in menschliche Realisierung, anstatt Hardware und Lizenzen fließen
- Aufwand für Anpassungen / Konfigurationen fallen auch bei Anschaffung von Hardware und Lizenzen an
- es sollte immer ein entsprechendes Test-System in der Planung berücksichtigt werden

Aufgaben und Funktion von Load-Balancing

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion
von Load-Balancing

Notwendigkeit von
Load-Balancing

Notwendigkeit von
Load-Balancing II

Notwendigkeit von
Load-Balancing III

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

■ Verteilung hoher Anfragelasten auf nachgelagerte Server

1 `das bedeutet in erste Linie Scaling nicht Hochverfügbarkeit`

■ Realisierung dieser Aufgabe über bestimmte Verfahren / Layer

- Layer 2 (Bonding)
- Layer 3 und Layer 4 (Direct Server Return, NAT)
- Layer 7 (Reverse Proxy, Application Proxy)

■ Realisierung von Firewall-Funktionalitäten nach gegebenen Ansprüchen

■ Umsetzung von Hochverfügbarkeitskonzepten

■ verlässliche Erkennung, ob Anfragen an nachgelagerte Server abgesetzt werden können

■ Realisierung von Server Updates im laufenden Betrieb, ohne Downtime

- ein Server kann die Last von Anfragen nicht abarbeiten

- 1 das betrifft vor allem Spitzen, also Zeiten in denen besonders viele Anfragen
- 2 an den Server gestellt werden

- Web-Inhalte sind mehr und mehr dynamisch

- Nutzer verweilen wesentlich länger auf der Webseite bzw. dem Service

- damit belegen sie für einen längeren Zeitraum Ressourcen, die bei statischen Inhalten anderen sofort wieder zur Verfügung gestanden hätten

- kostengünstigerer Ausbau von Ressourcen (kaufe lieber preiswerter)

- Pizzabox vs. Bolide

- von jeder Sorte sollten mind. zwei vorhanden sein, um Ausfälle kompensieren zu können

- mehrere Server erhöhen leider auch die Ausfallwahrscheinlichkeit

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion von Load-Balancing

Notwendigkeit von Load-Balancing

Notwendigkeit von Load-Balancing II

Notwendigkeit von Load-Balancing III

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

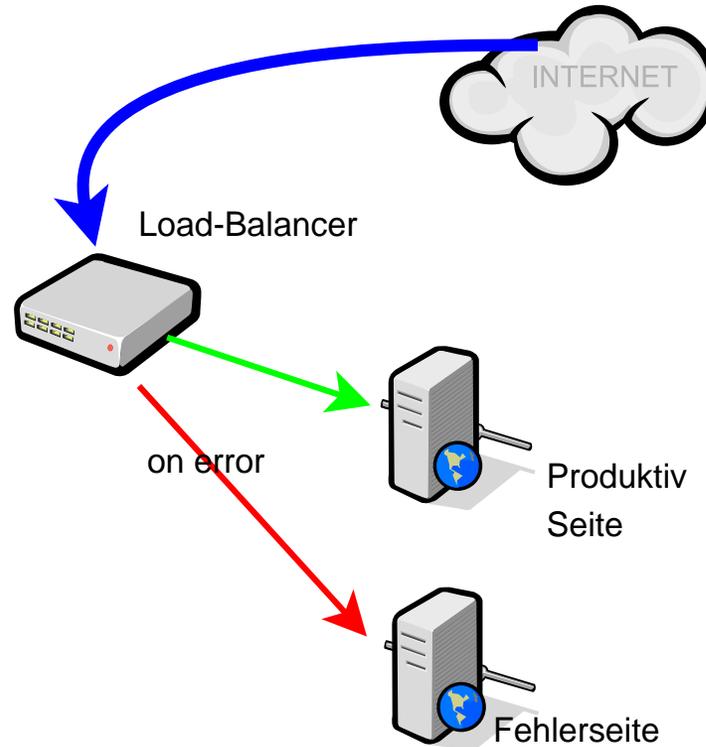
Reverse Proxy

SNAT TProxy

SSL Termination z.B. mit Wildcard Zertifikat

Das TIME_WAIT und CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

■ ein einfaches Fehlerseiten Szenario



- Fehlerseite bsp. über kleine VM
- bei Überlast oder Wartung springt so die Fehlerseite automatisch ein
- immerhin besser als ein Timeout oder eine weisse Seite

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion von Load-Balancing

Notwendigkeit von Load-Balancing

Notwendigkeit von Load-Balancing II

Notwendigkeit von Load-Balancing III

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

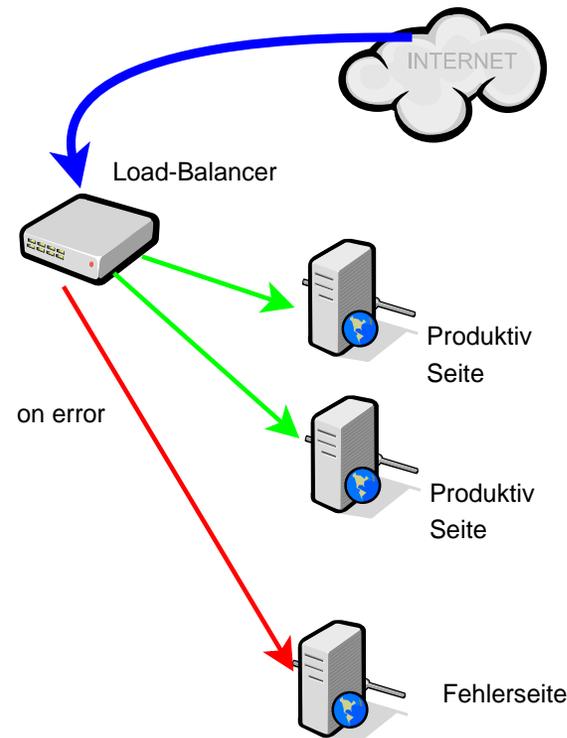
Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

■ besseres Fehlerseiten Szenario



- Fehlerseite wieder über kleine VM
- bei Überlast oder Wartung springt so die Fehlerseite automatisch ein
- Ausfälle eines Produktiv-Systems führen nicht länger zu einer Fehlerseite

Überblick

Einleitung

Wer bin ich

Wer bin ich

Warum dieses Thema

Warum dieses Thema

Aufgaben und Funktion von Load-Balancing

Notwendigkeit von Load-Balancing

Notwendigkeit von Load-Balancing II

Notwendigkeit von Load-Balancing III

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Überblick

LB-Techniken

DNS

Layer 2 / 3 / 4

Layer 7

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

LB-Techniken

Überblick

LB-Techniken

DNS

Layer 2 / 3 / 4

Layer 7

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Verteilung über Round-Robin DNS

```

1  dig -t google.com
2  .root-servers.net. 90288 IN A 192.58.128.30
3  j.root-servers.net. 90288 IN AAAA 2001:503:c27::2:30
4  l.root-servers.net. 90288 IN A 199.7.83.42
5  k.root-servers.net. 90288 IN A 193.0.14.129
6  ...

```

- Ausfall eines Servers schliesst einen nicht unerheblichen Teil der Nutzer aus
- Scale not HA
- geographisches LB, aka akamai
- auch hier: Anzahl der Server steigert die Ausfallwahrscheinlichkeit

Überblick

LB-Techniken

DNS

Layer 2 / 3 / 4

Layer 7

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Datalink Ethernet

- Bonding (z.B. 802.3ad) Link Aggregation
- Mac based LB (active active CARP¹)

■ Layer 3 und 4 (Netzwerk und Transport)

- Direct Routing (z.B. LVS)
- Tunneling (z.B. LVS)
- DNAT oder Full-Nat (z.B. LVS)

■ Aufwand sehr gering, da die Abarbeitung im Kernel abläuft

■ Kombinationen mit Hochverfügbarkeit möglich

- VRRP²)

¹**CARP** - Common Address Redundancy Protocol

²**VRRP** - Virtual Router Redundancy Protocol

Überblick

LB-Techniken

DNS

Layer 2 / 3 / 4

Layer 7

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Application

- Reverse Proxy basierende Lösungen
- im Gegensatz zu NLBs³ brauchen diese wesentlich mehr Ressourcen
- sie sind langsamer und können wesentlich weniger Traffic verarbeiten
- können jedoch auf Protokoll Ebene (http, rdp, smtp) arbeiten
- damit sind bsp. URL Switching u.v.a. mehr möglich
- Vertreter sind:
 - varnish, haproxy, pound, pen, stunnel
 - apache, lighttpd, nginx
- Wieviel Traffic haben Sie?

³**NLB** - Network Load Balancer

Überblick

LB-Techniken

Was sollte der LB noch können

Was noch I

Was noch II

Was noch III

Was noch, Persistence

Was noch,
Verteilungsverfahren

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

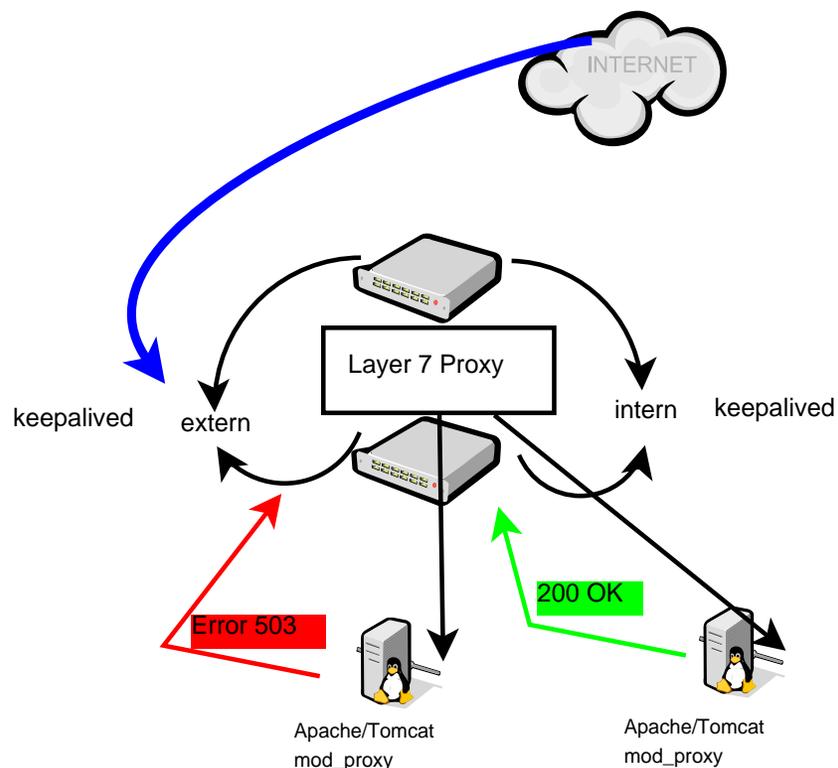
Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Was sollte der LB noch können

■ Health Checks

- ❑ Server antwortet zwar noch auf ICMP und HTTP, liefert jedoch WSOD⁴ oder 503
- ❑ Beispiel: HAProxy, Apache, Tomcat



⁴WSOD - White Screen of Death

- Überblick
- LB-Techniken
- Was sollte der LB noch können
- Was noch I
- Was noch II
- Was noch III
- Was noch, Persistence
- Was noch, Verteilungsverfahren
- Direct Routing
- NAT
- IP Tunneling
- Reverse Proxy
- SNAT TProxy
- SSL Termination z.B. mit Wildcard Zertifikat
- Das TIME_WAIT und CONNTRACK Problem Grenzen von Layer 4 - 7 Proxies
- Auswahl des richtigen Load Balancers, Verfahrens oder entsprechender Kombinationen
- Die grosse Herausforderung

■ Health Checks, Direkt versus Indirekt

- Direkt: check port 80 with 200 OK

```
1 option httpcheck
2 server 01 1.2.3.4:80 cookie 01 maxconn 20 check
```

- der Server selbst liefert den Status, anderer Port ist möglich
- INDIRECT: check port xyz with 200OK über einen eigenen TCP Service

```
1 server 01 1.2.3.4:52080 cookie f01 track server-status/f-01 maxconn 100
2 #
3 backend server-status
4 server f-01 1.2.3.5:9201 check inter 20s rise 2 fall 2
5 #
```

- ein Service mit anderer IP-Adresse / Port liefert den Status eines Servers
- in HAProxy Worten: benutze den Status eines anderen Backends

Überblick

LB-Techniken

Was sollte der LB noch können

Was noch I

Was noch II

Was noch III

Was noch, Persistence

Was noch, Verteilungsverfahren

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Überblick

LB-Techniken

Was sollte der LB noch können

Was noch I

Was noch II

Was noch III

Was noch, Persistence

Was noch, Verteilungsverfahren

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

■ Zusammenfassung Checks

- oft reicht ein einfacher TCP Check nicht aus (http vs. pure tcp)
- es muss darauf geachtet werden, dass Checks nicht zu häufig ausgeführt werden
- nicht das die Server nur noch mit Checks beschäftigt sind
- Wie verhält sich der Check unter Last Bedingungen?
- Flapping muss berücksichtigt werden
- achten Sie darauf, dass die angestrebte Lösung diese Möglichkeiten besitzt

■ Persistence

- der LB muss eine Anfrage von einem Client immer wieder an den gleichen Server schicken
- z.B. wäre ohne globales Session Management kein Login möglich (PHPSESSION)

■ Persistence

- der LB muss also wissen oder sich merken, welche Anfrage an welchen Server ausgeliefert wurde
- was passiert, wenn ein Server unpässlich ist
- was passiert, wenn einige Pakete verloren gehen und die Anfrage erneut abgesetzt werden muss
- Je nach Verfahren entsteht ein nicht unerheblicher Aufwand, der vor allem CPU und Speicher kostet
- Mit wieviel Speicher werden HW-Appliances ausgeliefert? (2-4GB RAM)
- Ausweg Cookie Insertion (Proxy injiziert Cookie und sendet ihn zurück an den Client)
- IP-Hashing-basierend, wenn bsp. TCP verwendet wird sind keine Cookies möglich
- Wie ist die IP Verteilung der Clients (NAT, Proxy etc.)

Überblick

LB-Techniken

Was sollte der LB noch können

Was noch I

Was noch II

Was noch III

Was noch, Persistence

Was noch, Verteilungsverfahren

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Überblick

LB-Techniken

Was sollte der LB noch können

Was noch I

Was noch II

Was noch III

Was noch, Persistence

Was noch, Verteilungsverfahren

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

- der erste der antwortet
 - muss keine gute Auslastung / Verteilung bedeuten, da viele Faktoren die Antwortzeiten bestimmen
- der die wenigstens Verbindungen hat
 - sinnvoll, wenn lange Sessionzeiten gebraucht werden
- round robin
 - nicht deterministisch, aber sorgt dennoch für eine gute Verteilung auf alle beteiligten Server
 - sehr einfach umzusetzen
- alle Verfahren lassen sich mit Wichtungen verbinden um bsp. vorhandene Hardwareklassen berücksichtigen zu können (schwache vs. starke Server)
- und natürlich IP-Hashing

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

Direct Server Return

Direct Server Return II

Direct Server Return

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Direct Routing

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

Direct Server Return

Direct Server Return II

Direct Server Return

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Funktionsweise

- Umschreiben der Ziel-Mac-Adresse auf einen oder mehrere Server mittels `ipvsadm` und entsprechender Kernel Module
- Server sendet die Antwort direkt
- sehr performant

■ Randbedingungen

- Backend darf auf ARP auf loopback nicht antworten
- kann damit nicht auf allen Systemen betrieben werden

```
1 ipvsadm -A -t <VIP>:80 -s wrr
2 ipvsadm -a -t <VIP>:80 -r 192.168.1.11:80 -g
3 ipvsadm -a -t <VIP>:80 -r 192.168.1.12:80 -g
4 sysctl -w net.ipv4.ip_forward=1
5 sysctl -w net.ipv4.conf.lo.arp_ignore=1
6 sysctl -w net.ipv4.conf.lo.arp_announce=2
7 #
```

- LB und Server müssen sich innerhalb des gleichen Netzwerkes befinden
- Anpassungen auf den Servern notwendig

■ Erweiterung mittels keepalived

```
1  #on the balancer
2  virtual_server 192.168.1.10 80 {
3      delay_loop 30
4      lb_algo wrr
5      lb_kind DR
6      persistence_timeout 500
7      protocol TCP
8      real_server 192.168.1.11 80 {
9          weight 1
10         TCP_CHECK {
11             connect_port 80
12             connect_timeout 3
13         }
14     }
15     real_server 192.168.1.12 80 {
16         weight 2
17         TCP_CHECK {
18             connect_port 80
19             connect_timeout 3
20         }
21     }
22 }
```

```
1  #auf den Webservern
2  #iptables-mod-nat-extra
3  iptables -t nat -A PREROUTING -p tcp -d 192.168.1.10 --dport 80 -j REDIRECT
```

■ Service checks dank keepalived (Framework um ipvsadm)

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

Direct Server Return

Direct Server Return II

Direct Server Return

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

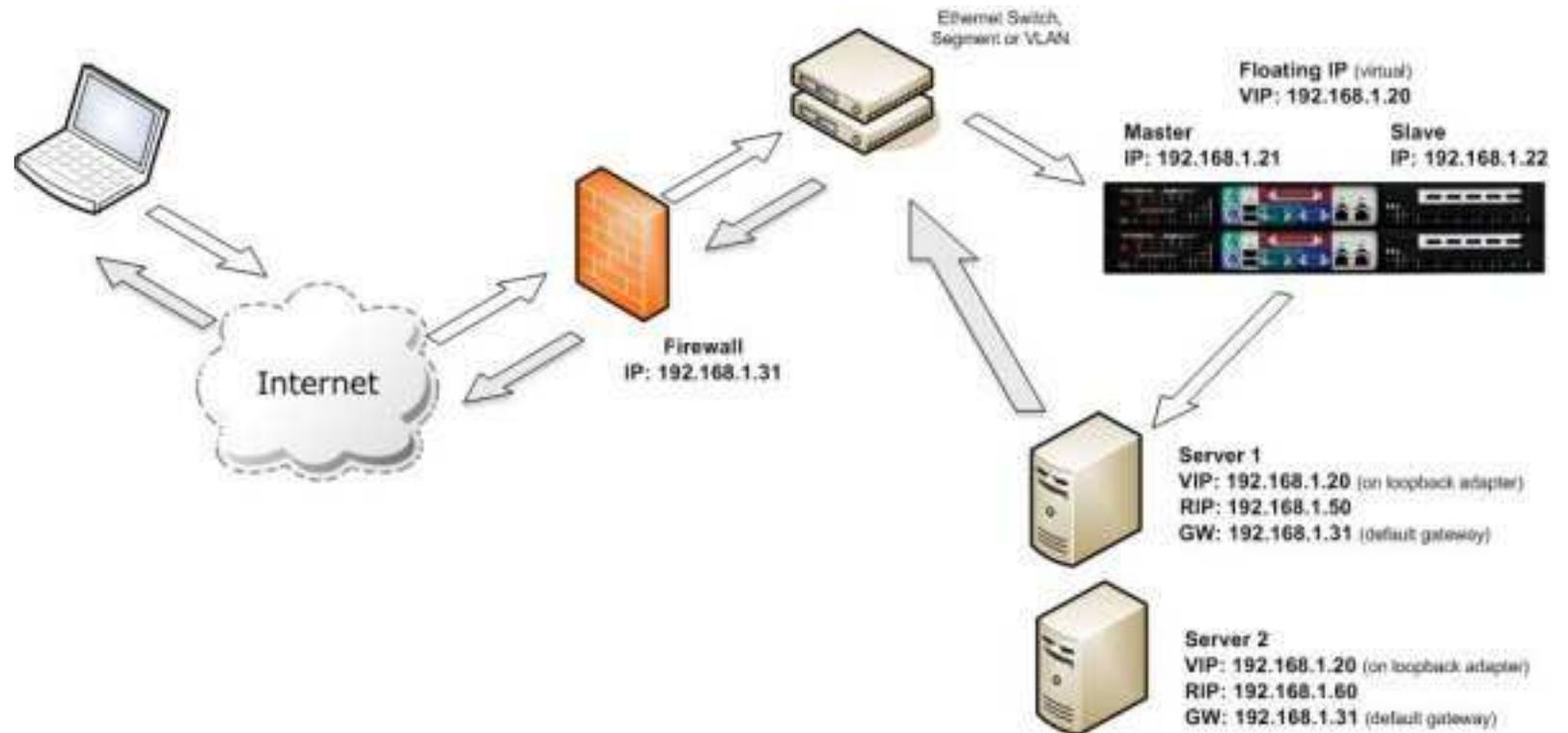
Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Schema von loadbalancer.org



Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

Direct Server Return

Direct Server Return II

Direct Server Return

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

Einfaches DNAT

LVS/NAT

LVS/Nat

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

NAT

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

Einfaches DNAT

LVS/NAT

LVS/Nat

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Beispiel:

- Umschreiben der Zieladresse auf einen Netzbereich

```
1 iptables -t nat -A PREROUTING -p tcp --dport 80 -j  
   DNAT --to 172.16.0.2-172.16.0.3
```

■ PRO

- einfache Realisierung

■ CON

- keine Einflussmöglichkeiten
- Hosts müssen aufeinander folgen
- Hosts müssen immer verfügbar sein

```
1 nicht zu gebrauchen
```

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

Einfaches DNAT

LVS/NAT

LVS/Nat

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard ZertifikatDas TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
ProxiesAuswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
KombinationenDie grosse
HerausforderungDer schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Beispiel:

- Umschreiben der Zieladresse auf einen oder mehrere Server mittels `ipvsadm` und entsprechender Kernel Module

```

1 ipvsadm -A -t <VIP>:80 -s wrr
2 ipvsadm -a -t <VIP>:80 -r 192.168.1.11:80 -m
3 ipvsadm -a -t <VIP>:80 -r 192.168.1.12:80 -m

```

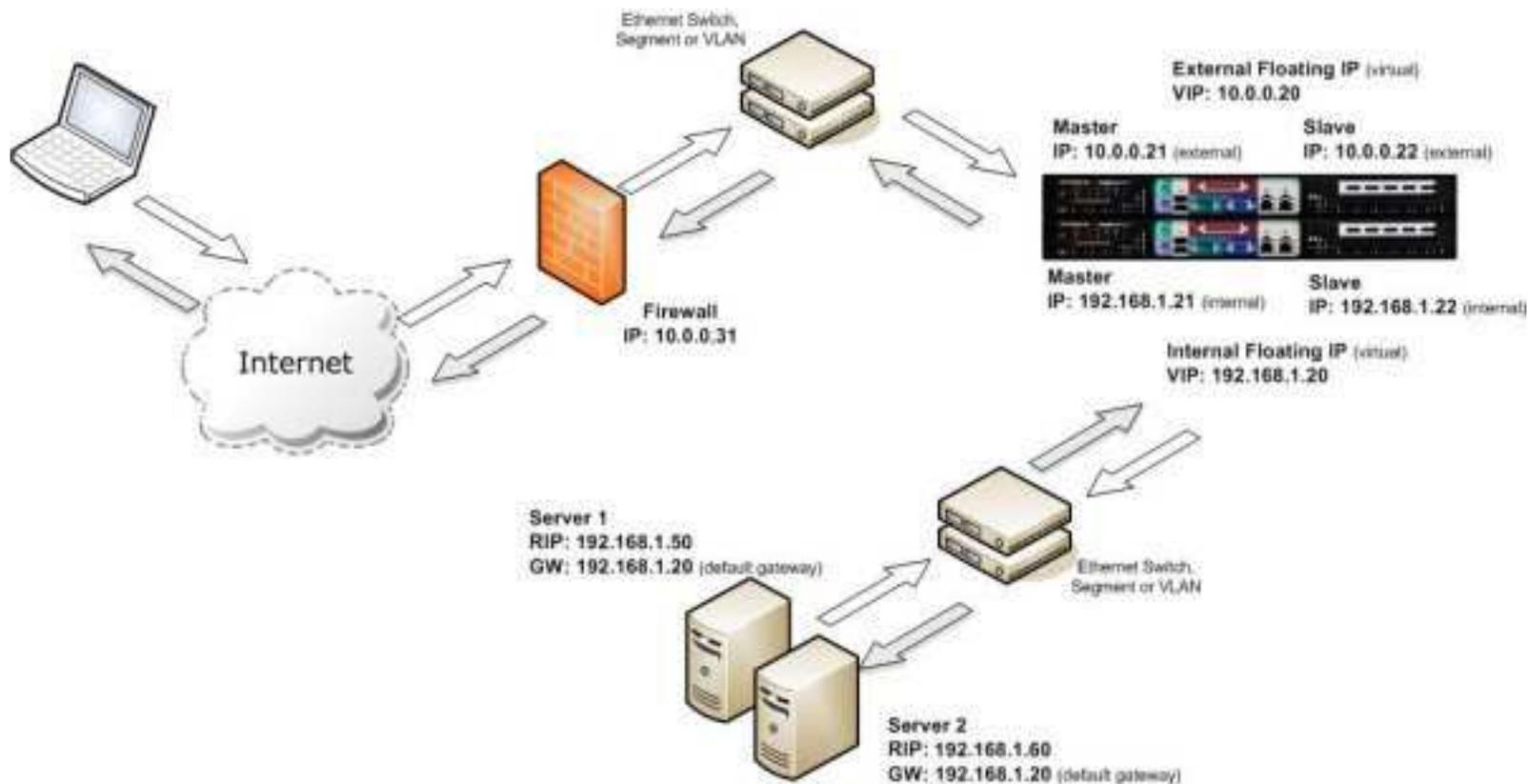
■ PRO

- einfache Realisierung
- keine Anpassung der Zielrechner notwendig

■ CON

- keine Service Checks
- LB muss Default Gateway für alle beteiligten Server sein
- oder entsprechend konfiguriertes Policy Routing

■ Schema von loadbalancer.org



Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

Einfaches DNAT

LVS/NAT

LVS/Nat

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

LVS/IP-Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Beispiele - / -
Anwendungsszenarien

IP Tunneling

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

LVS/IP-Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Beispiele - / -
Anwendungsszenarien

■ Beispiel:

- arbeitet wie Direct Routing
- die VIP wird auf jedem Backend konfiguriert (nur tun (ip tunnel) Interface)
- LB sendet Anfragen über entsprechenden IP Tunnel (ipip), jedoch mittels DNAT

```
1 ipvsadm -A -t <VIP>:80 -s wrr
2 ipvsadm -a -t <VIP>:80 -r 192.168.1.11:80 -i
3 ipvsadm -a -t <VIP>:80 -r 192.168.1.12:80 -i
```

■ PRO

- Server können durch Tunnel in fast beliebigen Netzen stehen

■ CON

- keine Service Checks
- Anpassung auf Servern notwendig (IP Tunnel)

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

HAProxy

SNAT / HAProxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Reverse Proxy

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

HAProxy

SNAT / HAProxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Beispiel:

- Anfrage wird vom Proxy entgegengenommen und dann nach speziellen Regeln an einen oder mehrere Server weitergeleitet

```
1 listen http:80
2 mode http
3 option xforwardfor
4 server 01 192.168.1.1:80
5 server 02 192.168.1.2:80
```

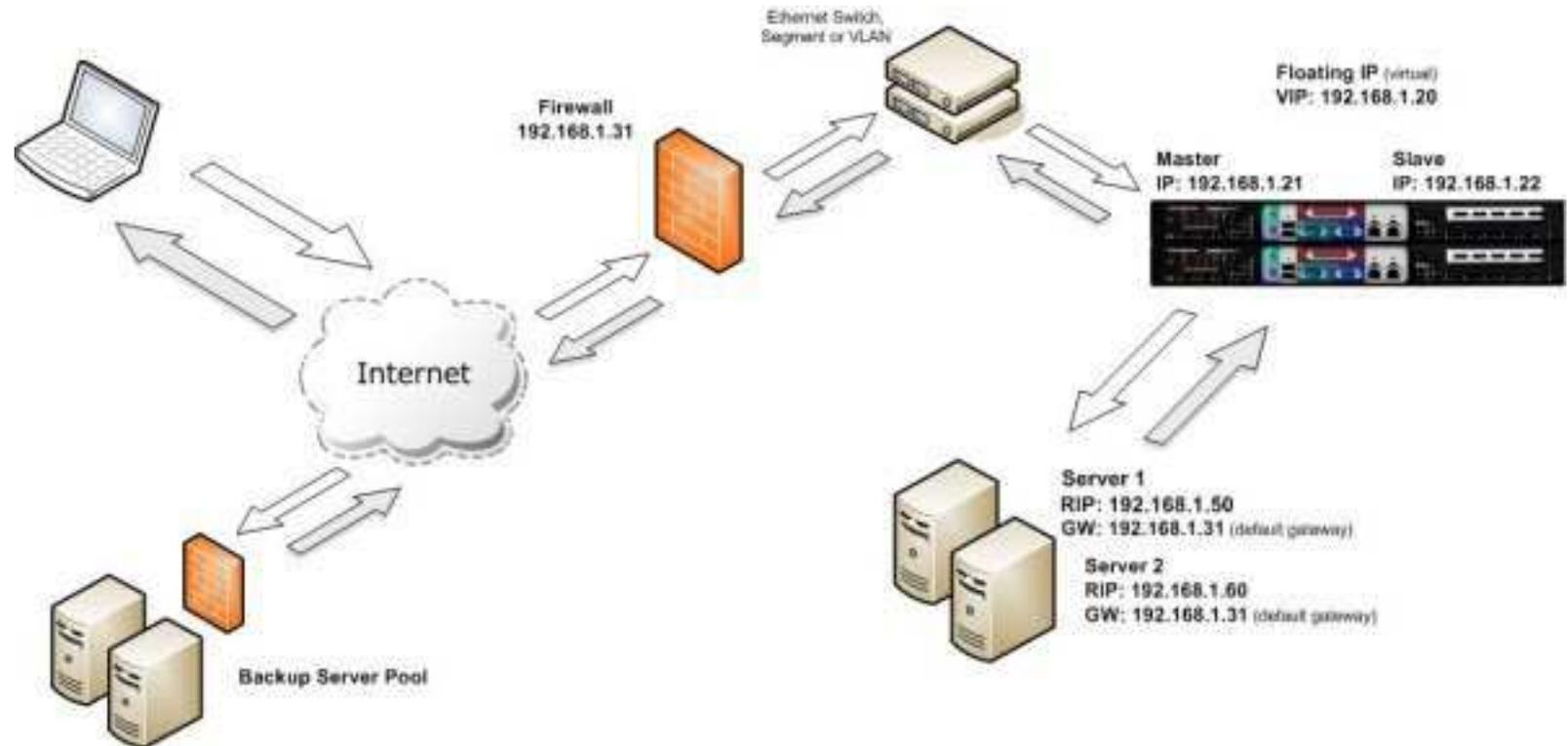
■ PRO

- einfach zu realisieren, Service Checks eingebaut

■ CON

- Server sieht nicht die Absender IP des Clients
- Anpassung auf Servern für Logging und Access-Control mit X-Forwarded-For

■ Schema von loadbalancer.org



Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

HAProxy

SNAT / HAProxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SNAT TProxy mit
HAProxy

TProxy SNAT / HAProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

SNAT TProxy

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SNAT TProxy mit HAProxy

TProxy SNAT / HAProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ Beispiel:

- Anfrage wird vom Proxy entgegengenommen und im Transparent Modus an einen oder mehrere Server weitergeleitet

```
1 listen http:80
2 mode http
3 source 0.0.0.0 usesrc clientip
4 server 01 192.168.1.1:80
5 server 02 192.168.1.2:80
```

■ PRO

- keine Modifikation an den Servern notwendig (Proxy wird quasi unsichtbar)

■ CON

- LB muss Default Gateway sein
- LB muss TProxy und Policy Routing unterstützen

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SNAT TProxy mit HAProxy

TProxy SNAT / HAProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load Balancers, Verfahrens oder entsprechender Kombinationen

Die grosse Herausforderung

Der schmale Geldbeutel oder
Die OpenSource Philosophie

■ Building Blocks

```
1 iptables -t mangle -N DIVERT
2 iptables -t mangle -A PREROUTING -p tcp -m socket -j DIVERT
3 iptables -t mangle -A DIVERT -j MARK --set-mark 1
4 iptables -t mangle -A DIVERT -j ACCEPT
5 ip rule add fwmark 1 lookup 100
6 ip route add local 0.0.0.0/0 dev lo table 100
7
8 echo 1 > /proc/sys/net/ipv4/conf/all/forwarding
9 echo 1 > /proc/sys/net/ipv4/conf/all/send_redirects
10 echo 1 > /proc/sys/net/ipv4/conf/eth0/send_redirects
11
12 opkg install iptables-mod-tproxy iptables-mod-ipopt
    iptables-mod-contrack iptables-mod-nat ip6tables
```

■ Prerouting schiebt TCP nach Divert

- DIVERT bekommt FWMARK 1 und sorgt bei der Beantwortung einer Anfrage dafür, das HAProxy die Anfrage wieder bekommt

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

SSL Termination mit
Pound / Stunnel

SSL Termination mit
Pound II

SSL Termination mit
Pound und TProxy

SSL Termination mit
Pound und TProxy

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

SSL Termination

z.B. mit Wildcard Zertifikat

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

SSL Termination mit
Pound / Stunnel

SSL Termination mit
Pound II

SSL Termination mit
Pound und TProxy

SSL Termination mit
Pound und TProxy

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

■ Beispiel:

- Anfrage wird von Pound per https entgegengenommen und dann an einen Reverse Proxy weitergeleitet

```
1 ListenHTTPS
2     Port 443
3 End
4 Service
5     BackEnd
6         Port 10443
7     End
8 End
```

■ PRO

- auf internen Server wird http gesprochen
- Inspektion des Traffics möglich

■ CON

- SSL kann den LB schnell auslasten
- Die Anwendung muss mitspielen z.B.
AddHeader "X-SSL-Request: 1"

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

SSL Termination mit
Pound / Stunnel

SSL Termination mit
Pound II

SSL Termination mit
Pound und TProxy

SSL Termination mit
Pound und TProxy

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

■ Beispiel:

```
1  frontend pound1_443_auf10443
2      bind :10443
3      mode http
4      option forwardfor except 127.0.0.1
5      acl acl_web1 -f /etc/haproxy-regex-web1.lst
6      acl acl_web2 -f /etc/haproxy-regex-web2.lst
7      use_backend web1 if acl_web1
8      use_backend web2 if acl_web2
9      default_backend devnull1
10 backend web1
11     .. server ...
12 backend web2
13     .. server ...
```

- je nach Regex auf Hostname wird ein anderes Backend angesprochen
- devnull liefert timeout, damit gibt es per se keine nicht konfigurierten Hosts
- viele *.wildcarddomain.de möglich
- vor allem für kleinere Shops oder Test-Instanzen sinnvoll
- Kombinationen mit Direct Routing sinnvoll

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

SSL Termination mit
Pound / Stunnel

SSL Termination mit
Pound II

SSL Termination mit
Pound und TProxy

SSL Termination mit
Pound und TProxy

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

■ Änderungen:

- es ist auch möglich pound als transparenten Proxy zu konfigurieren
- der wiederum transparent an HAProxy weiterleiten kann

1

```
http://blog.loadbalancer.org/transparent-proxy-of-ssl-traffic-using-pound-to-haproxy-backend-patch-and-howto/
```

■ PRO:

- keine Anpassungen an den Servern notwendig
- End to End transparent

■ CON:

- komplexe Konfiguration
- LB muss Default Gateway sein
- limitierte Sessions, da SSL und iptables Mehraufwand

SSL Termination mit Pound und TProxy

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

SSL Termination mit
Pound / Stunnel

SSL Termination mit
Pound II

SSL Termination mit
Pound und TProxy

SSL Termination mit
Pound und TProxy

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

```
1 ListenHTTPS
2     Address 0.0.0.0
3     Port 443
4     MaxRequest 16384
5 End
6 Service
7     BackEnd
8         Address 127.0.0.1
9         Port 10443
10        TProxy 1
11    End
12 End
13
14 iptables -t mangle -N DIVERT
15 iptables -t mangle -A PREROUTING -p tcp -m socket -j DIVERT
16 iptables -t mangle -A DIVERT -j MARK --set-mark 1
17 iptables -t mangle -A DIVERT -j ACCEPT
18
19 iptables -t mangle -A OUTPUT -s 127.0.0.1 -p tcp -sport 10443 -j DIVERT
20 iptables -t mangle -A OUTPUT -d 127.0.0.1 -p tcp -dport 10443 -j DIVERT
21
22 ip rule add fwmark 1 lookup 100
23 ip route add local 0.0.0.0/0 dev lo table 100
```

- ausgehender Verkehr muss ebenfalls nach Divert, so das Pound die Anfrage auch zurückbekommt

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

TIME_WAIT

TIME_WAIT II

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Das TIME_WAIT und CONNTRACK Problem Grenzen von Layer 4 - 7 Proxies

■ Erläuterung

- da der LB auf TCP aufsetzt, wird eine Session die erledigt ist in den `time_wait` Zustand übergehen
- dort verbleibt sie solange wie es der Kernel (`sysctl`) vorsieht; (15-60 Sekunden) sind das absolute Minimum
- Systeme, die sehr viele Session zu bearbeiten haben, brauchen entsprechend grosse Session Tabellen

```

1  1000 session / s, Session dauer 0,1s = 0,1 / (1 /
2  1000) = 100 aktive Sessions
   60s / ( 1 /1000) = 60000 TIME_WAITS

```

- TIME_WAIT ist sehr preiswert, allerdings müssen entsprechende Session Tabellen im Arbeitsspeicher vorgehalten werden.

```

1  # tcp-time-wait buckets pool size
2  # net.ipv4.tcp_max_tw_buckets = 180000

```

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard ZertifikatDas TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

TIME_WAIT

TIME_WAIT II

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
KombinationenDie grosse
HerausforderungDer schmale Geldbeutel
oder
Die OpenSource
Philosophie

[Überblick](#)

[LB-Techniken](#)

[Was sollte der LB noch können](#)

[Direct Routing](#)

[NAT](#)

[IP Tunneling](#)

[Reverse Proxy](#)

[SNAT TProxy](#)

[SSL Termination
z.B. mit Wildcard Zertifikat](#)

[Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies](#)

[TIME_WAIT](#)

[TIME_WAIT II](#)

[Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen](#)

[Die grosse
Herausforderung](#)

[Der schmale Geldbeutel
oder
Die OpenSource
Philosophie](#)

■ weitere Parameter

```

1   net.ipv4.tcp_fin_timeout = 30
2   net.ipv4.tcp_max_orphans = 8192
3   net.ipv4.ip_local_port_range = 16384    64000
4   net.ipv4.tcp_sack=0
5   net.ipv4.tcp_window_scaling=0
6
7   net.ipv4.tcp_tw_reuse = 1
8   net.ipv4.tcp_tw_recycle = 1
9
10  net.ipv4.tcp_max_tw_buckets = 16777216
11  net.ipv4.tcp_max_syn_backlog = 262144
12  net.core.netdev_max_backlog = 15000
13  net.core.somaxconn = 262144
14
15  net.core.rmem_max = 8738000
16  net.core.wmem_max = 6553600
17  net.ipv4.tcp_rmem = 8192 873800 8738000
18  net.ipv4.tcp_wmem = 4096 655360 6553600
19  vm.swappiness=0
20
21  fs.file-max=400000
22  fs.nr_open=400000
23
24  #+ nf_conntrack part

```

■ Erlaube sehr viele Waits, sonstige Verbindungen

■ dimensioniert auf 16GB RAM oder mehr

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Kriterien

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Beispiele - / -
Anwendungsszenarien

Auswahl des richtigen Load Balancers, Verfahrens oder entsprechender Kombinationen

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Kriterien

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Beispiele - / -
Anwendungsszenarien

■ Durchsatz

- Frage in die Runde: Wieviel Durchsatz brauchen Sie?
- Wieviele Session (maximal) werden erwartet?
- Nicht alle sind Slashdot oder Google

■ Bedienbarkeit

- Wie einfach bzw. wie schwer ist es Änderungen vorzunehmen?
- verstehen die Entwickler / Programmierer wie das LB funktioniert?
Bsp: Reverse Proxy: Download ohne Content-Length

■ Pros und Cons

- muss es Hardware sein, oder geht auch Software
- muss es ein kommerzielles Produkt sein?
Hilft das den Admins / Entwicklern
- Wie hoch liegt die TCO wirklich

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

**Die grosse
Herausforderung**

Die Anwendung

Die Anwendung II

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

Die grosse Herausforderung

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Die Anwendung

Die Anwendung II

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

- Trennung von statischen und dynamischen Inhalten
- `/static/img` vs. `/*.*`
- statische Inhalte lassen sich beispielsweise viel schneller mittels `lighttpd` ausliefern
- Realisierung mittels URL Switching
 - Tuning von Apache und dynamischen Inhalten
 - deaktivieren von `keepalive` (wirkt sich auch sehr negativ beim Einsatz von URL Switching aus)
- Lasttests der Anwendung
 - Die Anwendung sollte vor dem Start ein Profiling erfahren
 - oder: Auslastung zu Anfang sollte $\leq 40\%$ betragen

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Die Anwendung

Die Anwendung II

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

■ beliebte Fehler

- dynamische Downloads ohne Content-Length
- zu große Post oder Get Requests
- Verwendung von absoluten URLs oder Protokollen

■ Fehlerszenarien

- Eruiieren von Do's and Dont's
- Dokumentation (z.B. wie wird ein Server aus dem Grid entfernt für Downtime oder Updates
- Automatisierung und zentrale Verwaltung der Webserver Konfigurationen (z.B. puppet)

■ Sonstiges

- Überwachung der Anwendung, trotz LB
- Implementierung von Langzeitstatistiken und vor allem Logging

Überblick

LB-Techniken

Was sollte der LB noch
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

schmalere Geldbeutel

OpenWrt

OpenWrt II

Der schmale Geldbeutel oder Die OpenSource Philosophie

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

schmaler Geldbeutel

OpenWrt

OpenWrt II

■ absolutes Minimum

- Realisierung von Load-Balancing Mechanismen
- Firewall Funktionen
- VPN Verbindungen
 - OpenVPN, OpenSwan, tinc, Strongswan ...
- Überwachung / Starten von Diensten, Statistiken
 - monit, xinetd, curl, collectd, luci ...

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

schmaler Geldbeutel

OpenWrt

OpenWrt II

■ absolutes Minimum

- Realisierung von Load-Balancing Mechanismen
- Firewall Funktionen
- VPN Verbindungen
 - OpenVPN, OpenSwan, tinc, Strongswan ...
- Überwachung / Starten von Diensten, Statistiken
 - monit, xinetd, curl, collectd, luci ...
- alles auf Basis von OpenWrt

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

schmaler Geldbeutel

OpenWrt

OpenWrt II

■ absolutes Minimum

- Realisierung von Load-Balancing Mechanismen
- Firewall Funktionen
- VPN Verbindungen
 - OpenVPN, OpenSwan, tinc, Strongswan ...
- Überwachung / Starten von Diensten, Statistiken
 - monit, xinetd, curl, collectd, luci ...
- alles auf Basis von OpenWrt
 - X86 bare metal, KVM, XEN, Hyper-V, VMWare, Virtualbox

■ Building Blocks

- OpenWrt Buildchain, ein paar Patches, ein paar Modifikationen, ein paar neue Tools
- im Augenblick auf 32Bit begrenzt

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination
z.B. mit Wildcard Zertifikat

Das TIME_WAIT und
CONNTRACK Problem
Grenzen von Layer 4 - 7
Proxies

Auswahl des richtigen Load
Balancers, Verfahrens
oder entsprechender
Kombinationen

Die grosse
Herausforderung

Der schmale Geldbeutel
oder
Die OpenSource
Philosophie

schmaler Geldbeutel

OpenWrt

OpenWrt II

- SMP und PAE Unterstützung
- threaded OpenSSL, neue Kernel Module
- gcc, varnish, pound, acpid, keepalived u.v.a.
- rudimentäre initrd Unterstützung (z.B. booten vom USB Stick)
 - uclibc basierend (Basis Image 45MB)
 - busybox als bash Ersatz
 - hotplug statt udev (udev zu gross)
- klein oder gross
 - gleiches Framework egal ob Seagate Dockstar oder X86 Server
 - Test versus Produktiv
 - mit wenig Hardwareaufwand lassen sich komplexe Netzwerk Szenarien realisieren

■ Warum OpenWrt

- einfach zu installieren und dabei sehr klein
 - damit wird jedem der Zugang zu dieser Technologie ermöglicht
 - Test-Setups kosten fast nichts mehr
 - aufgrund der Größe kann auch einfach die Image Datei gesichert werden
 - oder eben nur die Konfigurationen und die Paketlist
- ```
1 daraus lässt sich mittels imagebuilder ein neues Image erzeugen
2 Neuinstallation == neues Image
```
- Kosten für die Realisierung und Anpassung eines Setups fließen vollständig in Hardware und Dienstleistungen ein
  - Was wir verkaufen ist Zeit und Know-How
  - Skalierung jederzeit möglich
  - kein Cluster Framework notwendig

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

schmaler Geldbeutel

OpenWrt

OpenWrt II

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

schmaler Geldbeutel

OpenWrt

OpenWrt II

## ■ Wir arbeiten an einer NDA<sup>5</sup>

- Mixture aus LVS, Pound, stunnel, HAProxy ...
- Web-GUI's
- Installation, Backup, Upgrade und Recovery Prozeduren
- am sysctl Tuning
- an der Farbe des Servers

## ■ Und die Kosten

- entstehen nur für Beratungsleistungen und spezielle Features
- kommen zu einem nicht unwesentlichen Teil dem OpenWrt Projekt zugute

## ■ Timeframe

- Mitte bis Ende 2012

---

<sup>5</sup>**NDA** - Network Distribution Appliance

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

Beispiele

## Beispiele - / - Anwendungsszenarien

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

Beispiele

## ■ MySQL Slave LB

- mehrere Slave werden unter Berücksichtigung von Seconds behind Master angesprochen

## ■ redundanter LB für Apache

- zwei LB via VRRP intern und extern und vielen Webservern

## ■ redundanter LB für IIS mit TProxy Snat

- mehrere IIS mit Service Checks hinter einem HAProxy
- LB == Gateway

## ■ redundanter LB für Streaming Server

- CRTMPSEVER auf 1935 (TCP, TProxy SNAT)
- LB == Gateway

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

Beispiele

## ■ SSL Termination mit Pound

- aufbrechen des SSL Traffics (transparent) oder per SNAT

## ■ Direct Server Return und SSL-Termination

- Wenn der Terminator nicht mehr ausreicht (SSL ist sehr CPU hungrig)

## ■ Beispiel Session TPROXY

```
1 opkg install iptables-mod-tproxy iptables-mod-ipopt iptables-mod-contrack iptables
 -mod-nat
2 opkg install haproxy
3 #!/etc/firewall.user
4 iptables -t mangle -N DIVERT
5 iptables -t mangle -A PREROUTING -p tcp -m socket -j DIVERT
6 iptables -t mangle -A DIVERT -j MARK --set-mark 1
7 iptables -t mangle -A DIVERT -j ACCEPT
8 #!/etc/rc.local
9 ip rule add fwmark 1 lookup 100
10 ip route add local 0.0.0.0/0 dev lo table 100
```

Überblick

LB-Techniken

Was sollte der LB noch  
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

Beispiele

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

Beispiele

```
1 #machine one
2 config 'interface' 'wan'
3 option 'ifname' 'eth0'
4 option 'proto' 'static'
5 option 'type' 'none'
6 option 'ipaddr' '192.168.1.10'
7 option 'gateway' '192.168.1.14'
8 option 'dns' '141.1.1.1'
9 option 'netmask' '255.255.255.0'
10
11 config 'interface' 'lan'
12 option 'ifname' 'eth1'
13 option 'proto' 'static'
14 option 'type' 'bridge'
15 option 'ipaddr' '192.168.0.10'
16 option 'netmask' '255.255.255.0'
17
18 #machine two
19 config 'interface' 'lan'
20 option 'ifname' 'eth0'
21 option 'proto' 'static'
22 option 'type' 'none'
23 option 'ipaddr' '192.168.0.11'
24 option 'gateway' '192.168.0.10'
25 option 'netmask' '255.255.255.0'
```

Überblick

LB-Techniken

Was sollte der LB noch können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

Beispiele

```
1 global
2 uid 0
3 gid 0
4 daemon
5
6 defaults
7 mode http
8 retries 5
9 maxconn 120000
10 contimeout 130000
11 clitimeout 130000
12 srvtimeout 130000
13 option redispatch
14
15 listen app1
16 bind :8080
17 mode http
18 #transparent do not use transparent until your 100 percent sure
19 maxconn 200
20 stats enable
21 stats uri /
22
23 frontend http-port-80
24 bind :80
25 mode http
26 maxconn 5000
27 option dontlognull
28 default_backend www
29
30 backend www
31 mode http
32 balance roundrobin
33 source 0.0.0.0 usesrc clientip
34 server server01 192.168.0.11:80 maxconn 200 check
35 server server01 192.168.0.12:80 maxconn 200 check
```

Überblick

LB-Techniken

Was sollte der LB noch  
können

Direct Routing

NAT

IP Tunneling

Reverse Proxy

SNAT TProxy

SSL Termination  
z.B. mit Wildcard Zertifikat

Das TIME\_WAIT und  
CONNTRACK Problem  
Grenzen von Layer 4 - 7  
Proxies

Auswahl des richtigen Load  
Balancers, Verfahrens  
oder entsprechender  
Kombinationen

Die grosse  
Herausforderung

Der schmale Geldbeutel  
oder  
Die OpenSource  
Philosophie

Beispiele - / -  
Anwendungsszenarien

**Fragen, Hinweise. Danke für Ihre  
Aufmerksamkeit**