

Load Balancing mit Freier Software

Linux Virtual Server (LVS)

Volker Dormeyer
<volker@ixolution.de>

GNU/LinuxTag 2005

Inhalt

- 1 **Einleitung**
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 **Techniken**
 - Linux Virtual Server (LVS)
 - Zielservers Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 **Zusammenfassung**
 - Nebeneffekt
 - Produkte
 - Literatur

Inhalt

- 1 **Einleitung**
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 **Techniken**
 - Linux Virtual Server (LVS)
 - Zielservers Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 **Zusammenfassung**
 - Nebeneffekt
 - Produkte
 - Literatur

Terminologie

VIP Virtuelle IP Adresse

DIP Director IP (Load Balancer selbst)

RS Real Server (Zielsysteme hinter Load Balancer)

RIP Real Server IP Adresse

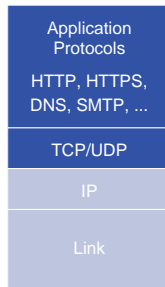
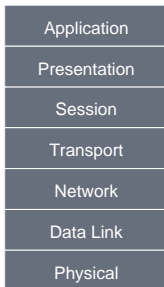
CIP Client IP Adresse

Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 Techniken
 - Linux Virtual Server (LVS)
 - Zielservers Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 Zusammenfassung
 - Nebeneffekt
 - Produkte
 - Literatur

Auf welchen Schichten arbeitet ein Load Balancer?

- OSI Schichten 4-7



Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 **Techniken**
 - **Linux Virtual Server (LVS)**
 - Zielserver Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 Zusammenfassung
 - Nebeneffekt
 - Produkte
 - Literatur

Was ist LVS?

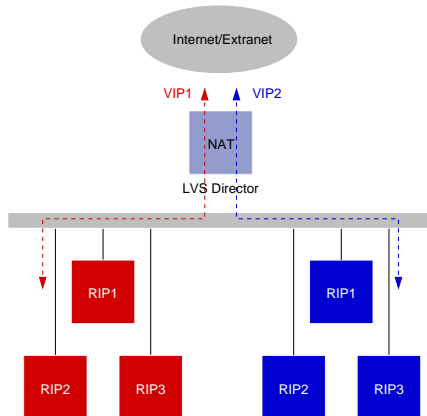
- Schicht 4 Switch
- Verteilt eingehende TCP/IP Verbindungen und UDP Pakete über im Hintergrund stehende Zielsysteme (Real Server)
- Linux Kernel Komponente (Netfilter Modul)

Forwarding Methoden

- LVS NAT
- LVS DR (Direct Routing)
- LVS TUN (Tunneling)

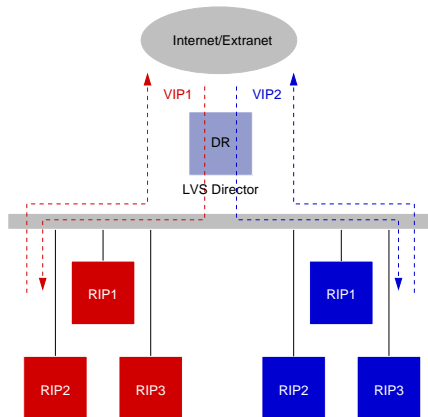
LVS NAT

- Hin- und Rückweg führen durch den Load Balancer



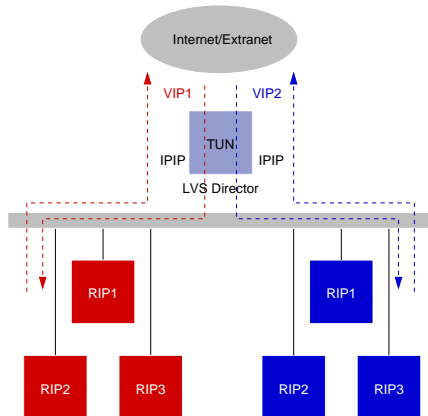
LVS DR (Direct Routing)

- Nur der Hinweg führt durch den Load Balancer



LVS TUN (Tunneling)

- IP IP Kapselung



LVS NAT

Pro

- Unabhängig vom Zielsystem
- Einfach zu verstehen und umzusetzen

Contra

- Protokollunterstützung für NAT nicht immer gegeben
- Geringerer Durchsatz in großen Umgebungen

LVS DR

Pro

- Applikationsunabhängig
- Sehr hoher Durchsatz

Contra

- Zusätzliche Konfiguration auf Zielsystemen nötig
- Load Balancer muß mit den Zielsystemen im gleichen Netzsegment stehen

LVS TUN

Pro

- Applikationsunabhängig
- Hoher Durchsatz
- Zielsysteme könne in Fremdnetzen stehen

Contra

- Zusätzliche Konfiguration auf Zielsystemen nötig

LVS Scheduler

- Bestimmung des Zielsystems für eine Verbindung

Scheduler Typen (gewichtet/ungewichtet)

- Round Robin
- Least Connection
- Locality Based Least Connection (replication)
- Hashing (source/destination)
- Shortest Expected Delay
- Never Queue
- Weighted Least Incoming Bps

LVS Connection Persistence

- Auch bekannt als “Stickyness” oder “Server Affinität”
- Schicht 4 Persistence
 - Quell-IP Adresse basiert
 - Quell-IP Adresse und Quell-/Zielport basiert
 - Quell-Netzwerk basiert
- “Connection Template” verfällt nach Ablauf der benutzerdefinierten Zeitangabe

Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 **Techniken**
 - Linux Virtual Server (LVS)
 - **Zielserver Überwachung**
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 Zusammenfassung
 - Nebeneffekt
 - Produkte
 - Literatur

Zielserver Überwachung (1)

- Methoden in Schichten 3-7
 - ICMP Pings
 - Port Verfügbarkeit
 - Applikationsverfügbarkeit (HTTP GET, etc.)
 - Applikationsergebnis Analyse
- In Band / Out of Band Überwachung
- GNU/Linux Zielserver
 - Feedback Agent
 - Dynamische Anpassung der Gewichte nach Zielserverauslastung

Zielserver Überwachung (2)

- Verbindungslimits
 - Limitierung der Verbindungsanzahl zu ausgewählten Zielsystemen
 - Mit automatischer Korrektur der Gewichte
- Mögliche Anpassung mit benutzerdefinierten Skripten
- Benutzerdefinierte dynamische Anpassung der Gewichte auf Basis ausgewählter Parameter der Zielsysteme

Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 **Techniken**
 - Linux Virtual Server (LVS)
 - Zielserver Überwachung
 - **OSI Schicht 7**
 - Hochverfügbarkeit
- 3 Zusammenfassung
 - Nebeneffekt
 - Produkte
 - Literatur

Schicht 4 Switch-Probleme

- Megaproxies

Schicht 7 Switching

- Parsen des Applikationsprotokolls

Methoden

- URL-Switching (Hash + Mustervergleiche)
- Parsen der Sitzungs-ID (HTTPS)
- Parsen von Cookies

Schicht 7 Persistency

- Anhand Ziel URL
- Sitzungs-ID (HTTPS)
- Cookie Daten

Inhalt

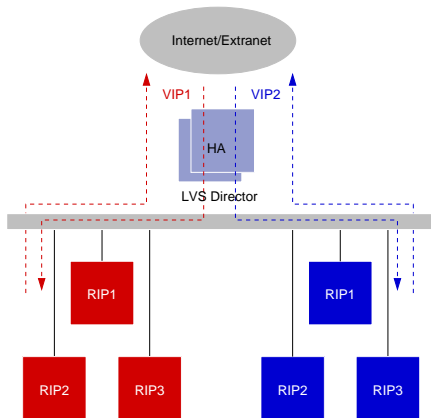
- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 **Techniken**
 - Linux Virtual Server (LVS)
 - Zielserver Überwachung
 - OSI Schicht 7
 - **Hochverfügbarkeit**
- 3 Zusammenfassung
 - Nebeneffekt
 - Produkte
 - Literatur

Hochverfügbarkeit

- Aktiv/Passiv
- Aktiv/Aktiv
- Synchronisation der Verbindungen (saru)

Heartbeat

- STONITH
- ipfail



Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 Techniken
 - Linux Virtual Server (LVS)
 - Zielservers Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 **Zusammenfassung**
 - **Nebeneffekt**
 - Produkte
 - Literatur

Nebeneffekt des Load Balancing

- Vereinfachte Wartungsmöglichkeiten ohne Stopp der Applikation
- Automatischer “Fallback”

Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 Techniken
 - Linux Virtual Server (LVS)
 - Zielsver Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 **Zusammenfassung**
 - Nebeneffekt
 - **Produkte**
 - Literatur

Produkte

Load Balancing

- CLUSTERIP (Schicht 4)
- LVS (Schicht 4, mit KTCPVS auch Schicht 7)
- Pen (Schicht 4, nur TCP)
- Pound (Proxy, Schicht 7)

Zielserver Überwachung

- Keepalived
- Ldirectord
- Mon

Hochverfügbarkeit

- Heartbeat
- Keepalived

Inhalt

- 1 Einleitung
 - Begriffserläuterung
 - OSI und IP Schichten
- 2 Techniken
 - Linux Virtual Server (LVS)
 - Zielservers Überwachung
 - OSI Schicht 7
 - Hochverfügbarkeit
- 3 Zusammenfassung
 - Nebeneffekt
 - Produkte
 - Literatur

Literatur



Chandra Kopparapu
Load Balancing Servers, Firewalls and Caches
Wiley Computer Publishing, 2002.



Matthew Syme, Philip Goldie
Optimizing Web Performance With Content Switching
Prentice Hall, 2003.



Karl Kopper
The Linux Enterprise Cluster
Build a HA Cluster with Commodity Hardware and Free Software
New Starch Press, 2005.



Evan Marcus, Hal Stern
Blueprints for High Availability
Wiley Publishing, 2003.

Fragen

Fragen? / Diskussion