

Apache Performance Tuning mittels Caching

René Pfeiffer <pfeiffer@luchs.at>

CaT

12. Mai 2007



Apache Performance Tuning mittels Caching

Ein paar Worte über Caches

Ein paar Worte über Caches

- Caches zeigen immer nur alte Informationen.

Ein paar Worte über Caches

- Caches zeigen immer nur alte Informationen.
- Caches sind gefährlich, weil sie Daten fressen.

Ein paar Worte über Caches

- Caches zeigen immer nur alte Informationen.
- Caches sind gefährlich, weil sie Daten fressen.
- Caches verzögern alles enorm.

Ein paar Worte über Caches

- Caches zeigen immer nur alte Informationen.
- Caches sind gefährlich, weil sie Daten fressen.
- Caches verzögern alles enorm.
- Caches tun seltsame und undurchschaubare Dinge.

Ein paar Worte über Caches

- Caches zeigen immer nur alte Informationen.
- Caches sind gefährlich, weil sie Daten fressen.
- Caches verzögern alles enorm.
- Caches tun seltsame und undurchschaubare Dinge.
- Caches machen Web Development unmöglich.

Ein paar Worte über Caches

- Caches zeigen immer nur alte Informationen.
- Caches sind gefährlich, weil sie Daten fressen.
- Caches verzögern alles enorm.
- Caches tun seltsame und undurchschaubare Dinge.
- Caches machen Web Development unmöglich.

Diese Aussagen sind alle **falsch!**

Subversive Cache Infiltration (SIC!)

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.
- Festplatten haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.
- Festplatten haben Caches.
- Netzwerkkarten haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.
- Festplatten haben Caches.
- Netzwerkkarten haben Caches.
- Router haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.
- Festplatten haben Caches.
- Netzwerkkarten haben Caches.
- Router haben Caches.
- Applikationen haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.
- Festplatten haben Caches.
- Netzwerkkarten haben Caches.
- Router haben Caches.
- Applikationen haben Caches.
- Betriebssysteme haben Caches.

Subversive Cache Infiltration (SIC!)

- CPUs haben Caches.
- Grafikkarten haben Caches.
- Festplattencontroller haben Caches.
- Festplatten haben Caches.
- Netzwerkkarten haben Caches.
- Router haben Caches.
- Applikationen haben Caches.
- Betriebssysteme haben Caches.

Seltsamerweise funktionieren all diese Systeme trotzdem (ja, tun sie).

Wege von Webseiten

Wege von Webseiten

- Webbrowser stellen Fragen (HTTP Requests).

Wege von Webseiten

- Webbrowser stellen Fragen (HTTP Requests).
- Webserver antworten (HTTP Responses).

Wege von Webseiten

- Webbrowser stellen Fragen (HTTP Requests).
- Webserver antworten (HTTP Responses).
- Zwischen Frage und Antwort liegt viel Netzwerk.

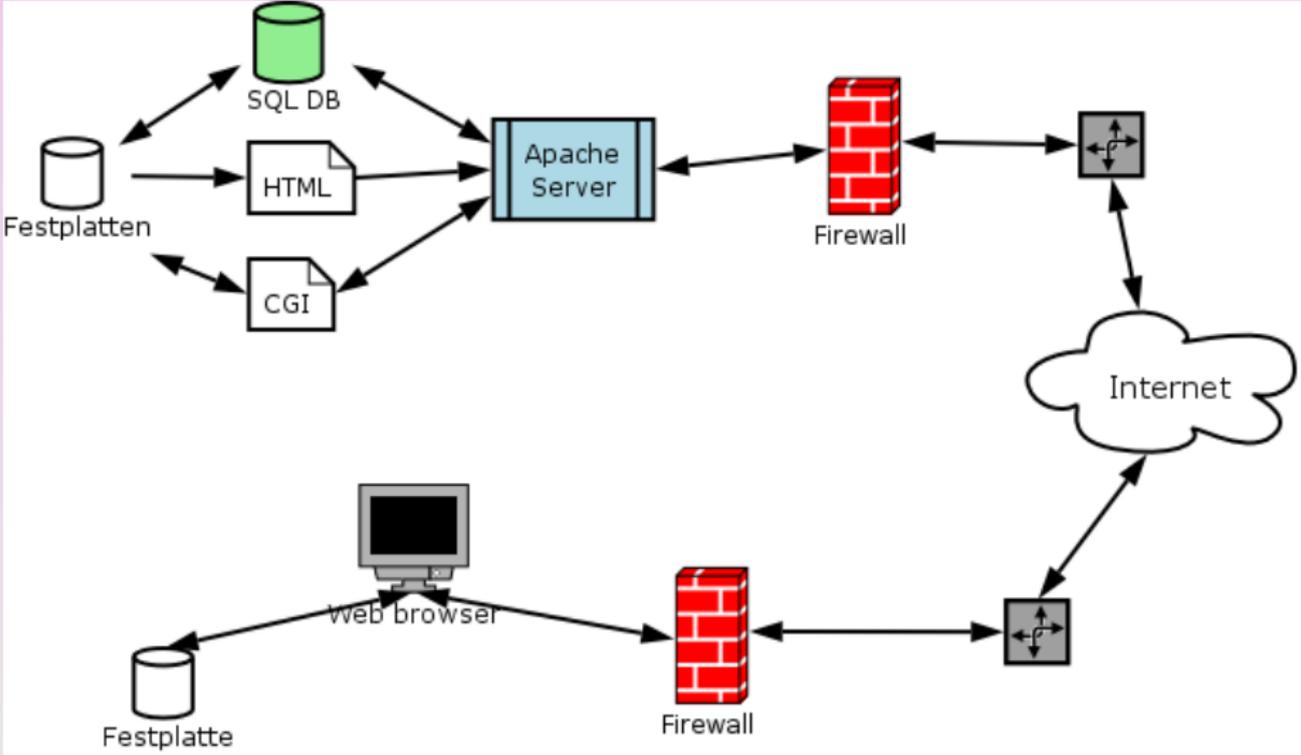
Wege von Webseiten

- Webbrowser stellen Fragen (HTTP Requests).
- Webserver antworten (HTTP Responses).
- Zwischen Frage und Antwort liegt viel Netzwerk.
- Manche Antworten werden lokal vom Webbrowser gespeichert.

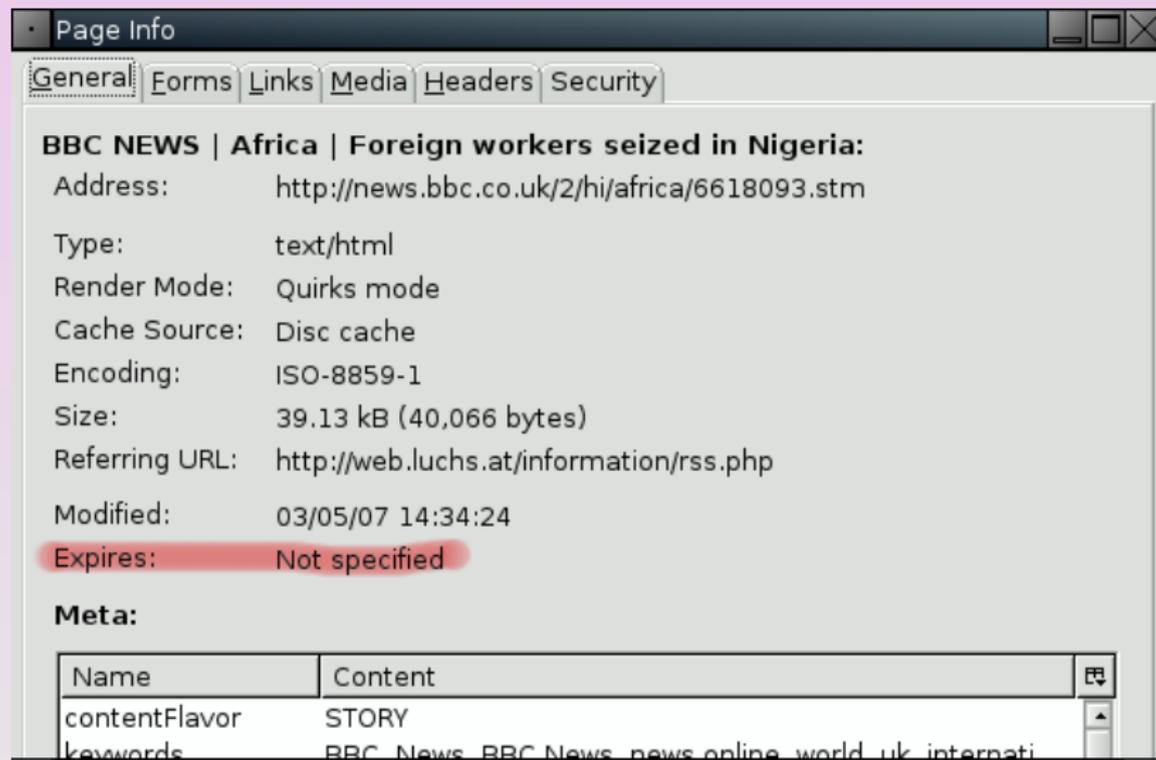
Wege von Webseiten

- Webbrowser stellen Fragen (HTTP Requests).
- Webserver antworten (HTTP Responses).
- Zwischen Frage und Antwort liegt viel Netzwerk.
- Manche Antworten werden lokal vom Webbrowser gespeichert.
- Lokale Kopien schlagen *jede* Breitbandanbindung.

Von Platte zu Platte



Ablaufdaten - Wie lange darf der Cache?



The screenshot shows a 'Page Info' dialog box with several tabs: 'General', 'Forms', 'Links', 'Media', 'Headers', and 'Security'. The 'General' tab is selected, displaying the following information:

- BBC NEWS | Africa | Foreign workers seized in Nigeria:**
- Address: <http://news.bbc.co.uk/2/hi/africa/6618093.stm>
- Type: text/html
- Render Mode: Quirks mode
- Cache Source: Disc cache
- Encoding: ISO-8859-1
- Size: 39.13 kB (40,066 bytes)
- Referring URL: <http://web.luchs.at/information/rss.php>
- Modified: 03/05/07 14:34:24
- Expires: Not specified** (highlighted in red)

Below the main information is a 'Meta:' section with a table:

Name	Content	
contentFlavor	STORY	
keywords	BBC News BBC News news online world uk internati	

Ablaufdaten setzen mit Apache

Ablaufdaten setzen mit Apache

- Apache Modul `mod_expires` kann Ablaufdaten setzen.

Ablaufdaten setzen mit Apache

- Apache Modul `mod_expires` kann Ablaufdaten setzen.
- `mod_expires` kontrolliert bzw. setzt:
 - ▶ `Expires`: HTTP Header
 - ▶ `max-age` Direktive im `Cache-Control` HTTP Header

Ablaufdaten setzen mit Apache

- Apache Modul `mod_expires` kann Ablaufdaten setzen.
- `mod_expires` kontrolliert bzw. setzt:
 - ▶ `Expires`: HTTP Header
 - ▶ `max-age` Direktive im `Cache-Control` HTTP Header
- Beide Informationen sind wichtig für Caches.
 - ▶ Inhalte werden sinnvoll zwischengespeichert.
 - ▶ Inhalte werden rechtzeitig aktualisiert.

Ablaufdaten setzen mit Apache

- Apache Modul `mod_expires` kann Ablaufdaten setzen.
- `mod_expires` kontrolliert bzw. setzt:
 - ▶ `Expires`: HTTP Header
 - ▶ `max-age` Direktive im `Cache-Control` HTTP Header
- Beide Informationen sind wichtig für Caches.
 - ▶ Inhalte werden sinnvoll zwischengespeichert.
 - ▶ Inhalte werden rechtzeitig aktualisiert.
- `mod_expires` ist meist per Default inaktiv.

Ablaufdaten setzen mit Apache (2)

Ablaufdaten setzen mit Apache (2)

- `mod_expires` **einschalten**: `ExpiresActive On`

Ablaufdaten setzen mit Apache (2)

- `mod_expires` **einschalten**: `ExpiresActive On`
- `mod_expires` **kann Ablaufdaten global festlegen**.
 - ▶ `ExpiresDefault ``access plus 1 month```

Ablaufdaten setzen mit Apache (2)

- `mod_expires` **einschalten**: `ExpiresActive On`
- `mod_expires` **kann Ablaufdaten global festlegen.**
 - ▶ `ExpiresDefault ``access plus 1 month```
- `mod_expires` **kann gezielt Ablaufdaten setzen.**
 - ▶ `ExpiresByType text/html ``access plus 3 hours```
 - ▶ `ExpiresByType application/xml ``modification plus 2 weeks```

Ablaufdaten setzen mit Apache (2)

- `mod_expires` **einschalten**: `ExpiresActive On`
 - `mod_expires` **kann Ablaufdaten global festlegen.**
 - ▶ `ExpiresDefault ``access plus 1 month```
 - `mod_expires` **kann gezielt Ablaufdaten setzen.**
 - ▶ `ExpiresByType text/html ``access plus 3 hours```
 - ▶ `ExpiresByType application/xml ``modification plus 2 weeks```
 - **Unterschied `access` und `modification`**
 - ▶ `access` nimmt Clientzugriffszeit
 - ▶ `modification` Modification Time der Datei
- als Basis.

mod_expires **in der** httpd.conf

mod_expires in der httpd.conf

```
<IfModule mod_expires.c>
    ExpiresActive On
    ExpiresByType text/html "access plus 3 hours"
    ExpiresByType text/xml "access plus 3 hours"
    ExpiresByType image/gif "access plus 4 weeks"
    ExpiresByType image/jpg "access plus 4 weeks"
    ExpiresByType image/png "access plus 4 weeks"
    ExpiresByType video/quicktime "access plus 2 months"
    ExpiresByType audio/mpeg "access plus 2 months"
    ExpiresByType application/pdf "access plus 2 weeks"
    ExpiresByType application/ps "access plus 2 weeks"
    ExpiresByType application/xml "modification plus 2 weeks"
</IfModule>
```

Obige Konfiguration kann in die globale Sektion, in Virtual Hosts oder Directory Deklarationen. Angegebene Zeitspannen sind Beispiele.

Der Squid Proxy

Der Squid Proxy

- Squid ist ein HTTP/FTP High Performance Proxy.

Der Squid Proxy

- Squid ist ein HTTP/FTP High Performance Proxy.
- Squid Proxies können
 - ▶ vernetzt in Hierarchien
 - ▶ vor HTTP/FTP Clients und
 - ▶ vor HTTP/FTP Servern agieren.

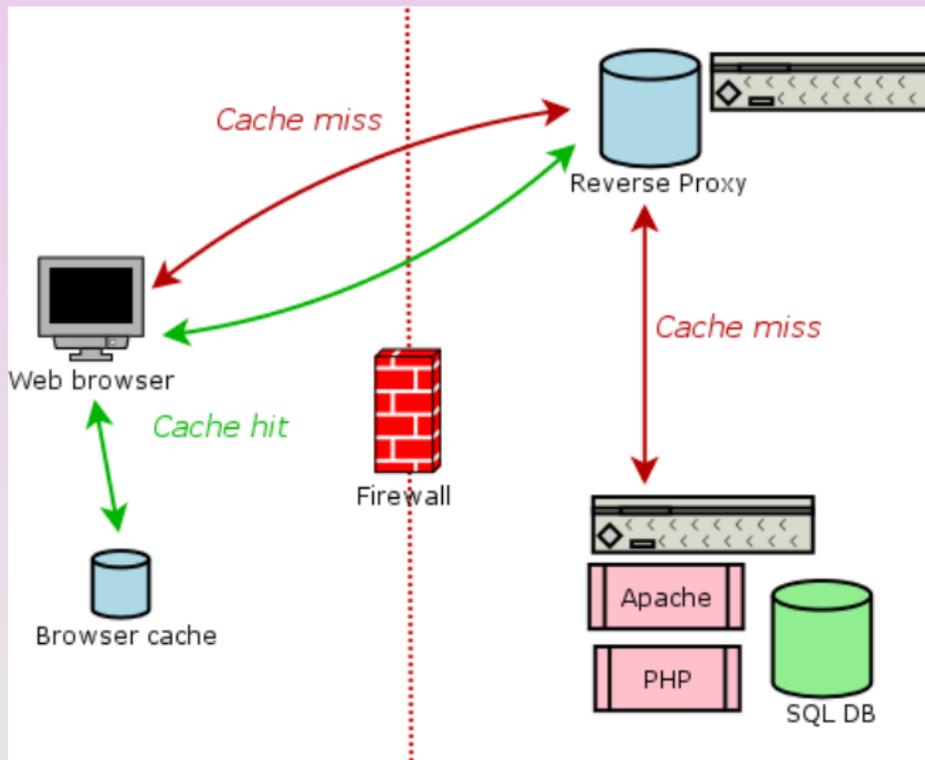
Der Squid Proxy

- Squid ist ein HTTP/FTP High Performance Proxy.
- Squid Proxies können
 - ▶ vernetzt in Hierarchien
 - ▶ vor HTTP/FTP Clients und
 - ▶ vor HTTP/FTP Servern agieren.
- Squids on-disk Format skaliert bis zu mehreren 100 GB.

Der Squid Proxy

- Squid ist ein HTTP/FTP High Performance Proxy.
- Squid Proxies können
 - ▶ vernetzt in Hierarchien
 - ▶ vor HTTP/FTP Clients und
 - ▶ vor HTTP/FTP Servern agieren.
- Squids on-disk Format skaliert bis zu mehreren 100 GB.
- Squid freut sich auch über viel RAM. 😊

Squid als Reverse Proxy



Squid als Reverse Proxy - Konfiguration

Squid als Reverse Proxy - Konfiguration

- Ausgangspunkt
 - ▶ Webserver 172.16.23.43 mit Virtual Host `www.luchs.at`
 - ▶ Squidproxy 172.16.23.42

Squid als Reverse Proxy - Konfiguration

- Ausgangspunkt
 - ▶ Webserver 172.16.23.43 mit Virtual Host `www.luchs.at`
 - ▶ Squidproxy 172.16.23.42
- Vorbereiten in der `squid.conf`:

```
http_port      172.16.23.42:80 vhost vport
http_port      127.0.0.1:80
icp_port       0
cache_peer     172.16.23.43 parent 80 0 originserver default
```

Squid als Reverse Proxy - Konfiguration

- Ausgangspunkt

- ▶ Webserver 172.16.23.43 mit Virtual Host `www.luchs.at`
- ▶ Squidproxy 172.16.23.42

- Vorbereiten in der `squid.conf`:

```
http_port      172.16.23.42:80 vhost vport
http_port      127.0.0.1:80
icp_port       0
cache_peer     172.16.23.43 parent 80 0 originserver default
```

- Squid weiß nun, daß nicht zwischengespeicherter Requests an 172.16.23.43 auf Port 80 weitergeleitet werden sollen.

Squid als Reverse Proxy - Zugriffskontrolle

Squid als Reverse Proxy - Zugriffskontrolle

- Squid versteht Zugriffskontrollen (Access Lists)

Squid als Reverse Proxy - Zugriffskontrolle

- Squid versteht Zugriffskontrollen (Access Lists)
- Explizites Erlauben jeglichen Zugriffs auf unseren Webservice:

```
acl          all src 0.0.0.0/0.0.0.0
acl          manager proto cache_object
acl          localhost src 127.0.0.1/255.255.255.255
acl          accel_hosts dst 172.16.23.42 172.16.23.43
http_access  allow accel_hosts
http_access  allow manager localhost
http_access  deny manager
http_access  deny all
deny_info   http://www.luchs.at/ all
```

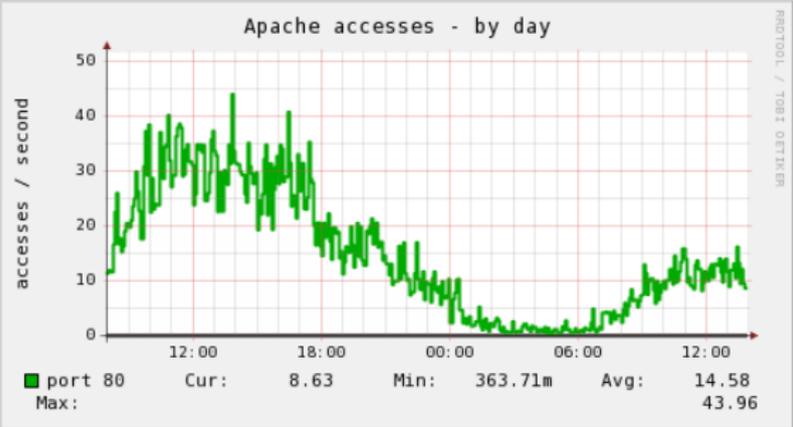
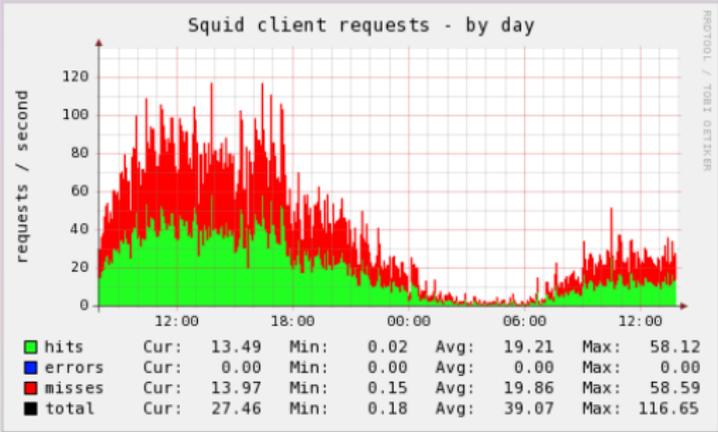
Squid als Reverse Proxy - Zugriffskontrolle

- Squid versteht Zugriffskontrollen (Access Lists)
- Explizites Erlauben jeglichen Zugriffs auf unseren Webservice:

```
acl          all src 0.0.0.0/0.0.0.0
acl          manager proto cache_object
acl          localhost src 127.0.0.1/255.255.255.255
acl          accel_hosts dst 172.16.23.42 172.16.23.43
http_access  allow accel_hosts
http_access  allow manager localhost
http_access  deny manager
http_access  deny all
deny_info    http://www.luchs.at/ all
```

- Niemals derart Forward und Reverse Proxybetrieb mischen!
 - ▶ Das führt zu offenen Proxies und mehr Spam!

Erfolgsstatistiken



Was man noch tun kann...

Was man noch tun kann...

- MySQL Query Cache verwenden.
- Apache 2.2 hat Caches für das Dateisystem.
- Linux Kernel Tuning (Dateisystem, Memory Management).
- PHP Caching (Accelerators).
- Kompression in allen Formen und Farben.
- Kritische Komponenten mit GCC und Optionen entschlacken.
- ...

Danke! Noch Fragen?

```
Stopping sshd: [ OK ]
Stopping xinetd: [ OK ]
Stopping crond: [ OK ]
Saving random seed: [ OK ]
Stopping NFS statd: [ OK ]
Stopping portmapper: [ OK ]
Shutting down kernel logger: [ OK ]
Shutting down system logger: [ OK ]
Shutting down interface eth0: [ OK ]
Shutting down interface eth1: [ OK ]
Shutting down loopback interface: [ OK ]
Disabling IPv4 packet forwarding: [ OK ]
Starting killall: [ OK ]
Sending all processes the TERM signal... [ OK ]
Sending all processes the KILL signal... [ OK ]
Syncing hardware clock to system time [ OK ]
Turning off swap: [ OK ]
Turning off quotas: [ OK ]
Unmounting file systems: VFS: Busy inodes after unmount. Self-destruct in 5 seconds. Have a nice day... [ OK ]

Halting system...
flushing ide devices: hda hdb
System halted.
```

Die Linuxgazette

Dieses Thema wurde auch in einem Artikel der [Linuxgazette](#) in der Ausgabe [132](#) behandelt. Die Linuxgazette publiziert seit mehr als 10 Jahren Interessantes, Wissenswertes, Tips und Nützliches rund um Linux®. Jede(r) ist eingeladen einen Artikel zu schreiben. Man wird von der Redaktion betreut, wenn man nett ist, und obendrein ist man dann in einer Zeitschrift mit ISSN (1934-371X) verewigt! Noch dazu wird man weltweit gespiegelt und gelesen! Wann schreibst Du Deinen Artikel? 😊

Über dieses Dokument

- Autor: René Pfeiffer
- Erstellt mit \LaTeX und \LaTeX Beamer Class
- Dokumentensammlung unter
<http://web.luchs.at/information/docs.php>

Copyright (C) 2007 by René Pfeiffer <lynx@luchs.at>. This material may be distributed only subject to the terms and conditions set forth in the Open Publication License, v1.0 or later (the latest version is presently available at <http://www.opencontent.org/openpub/>).