

Rails Cluster

Mehrserverlösung zum Betrieb von RoR
Applikationen

Julian Fischer
fischer@enterprise-rails.de
<http://www.enterprise-rails.de>

Einleitung

Einleitung

Über mich

Über mich



Julian Fischer

- ▶ Geschäftsführender Gesellschafter der Avarteq GmbH
- ▶ Twitter: <http://www.twitter.com/railshoster>
- ▶ E-Mail: fischer@enterprise-rails.de

- ▶ Lehrbeauftragter im Fach „Ruby on Rails“ an der HTWdS
- ▶ Ruby und Ruby on Rails Programmierer
- ▶ Enterprise-Rails.de - Head of Hosting

Einleitung

Über die Avarteq GmbH

Über die Avarteq GmbH



- ▶ **Gegründet im Nov. 2008**
aus zwei Einzelunternehmungen
- ▶ **Beteiligung der Key-Systems GmbH**
verwaltet derzeit 2,5 Mio. Domains für Kunden aus über 200 Ländern
- ▶ **Teamgröße: 14 Personen**
7 Vollzeit, 7 Teilzeit/Freelancer

Einleitung

Portfolio

Über die Avarteq GmbH



- ▶ Umfasst alle Phasen eines Web-Projekts
- ▶ Beratung
- ▶ Konzeption und Screendesign
- ▶ Ruby on Rails Entwicklung. Inhouse & vorort.
- ▶ Ruby on Rails Hosting
RailsHoster.de - Enterprise-Rails.de

Wie erstellt man einen hochwertigen Rails Cluster?

Vorarbeit

Vorarbeit

Wahl des Rechenzentrums

Wahl des RZ



	Hetzner	Thomas Krenn AG
Load Balancer	Software, Apache or Nginx	2 LB Appliances, autom. IP-Failover
Storage	NFS	NetApp Storage Appliance, redundante HW, separate 1Gbit Verbindung zu jedem App-Server
App-Servers	EQ4-EQ9	HW nach individuellem Bedarf, z.B. SuperMicro Basis, 2xIntel Xeon Quad Core CPUs, redundantes Netzweil, Hardware-Raid mit BBU
Privates Netzwerk	Anzahl der Server muss im Voraus bekannt sein	Neue Server können jederzeit hinzugenommen werden.
Support im RZ	HDD defekt? Ticket an RZ Support. Downtime. Tausch der HDD.	Rechner piept? Anruf durch RZ: HDD könnte defekt sein. Diagnose. HDD hotswap. Keine Downtime!

Vorarbeit

Dimensionierung

Dimensionierung



- ▶ Wieviele Server werden benötigt?
- ▶ Welche Server-Rollen wird es geben?
App, DB, LB, BG-Jobs, ...
- ▶ Welche HW ist für die jeweilige Rolle am besten geeignet?
- ▶ ...

Vorarbeit System-Design

- ▶ Welche System-Komponente kommt auf welchen Server?
- ▶ APP und DB auf die gleichen Server?
- ▶ Dedizierte DB Server?
- ▶ Wohin mit dem Suchserver?
- ▶ Wo werden Job-Queues laufen?
- ▶ ...

Hardware

Hardware Load Balancer

Load Balancer



- ▶ Selten der Flaschenhals
- ▶ Oft Single Point of Failure (SPOF)
- ▶ Zwei sind besser als einer!
- ▶ Was tun wenn einer kaputt geht?
- ▶ Redundanter LB mit automatischem IP Failover!

Hardware Storage

Load Balancer



- ▶ **Wohin mit den Uploads?**
Profilbilder, etc.
- ▶ **NFS?**
- ▶ **SPOF**
- ▶ **Vergleichsweise langsam**

Load Balancer



- ▶ Redundante iSCSI Storage Appliance
- ▶ Jeder App Server wird mit einer separaten GBit-Leitung angeschlossen
- ▶ iSCSI = SCSI über TCP
= Remote Blockdevice
- ▶ Weniger Overhead als NFS
- ▶ Clusterdateisystem benötigt.

OCFS2, GFS, ...

Hardware

Netzwerk

- ▶ Wie sind die Rechner des Clusters miteinander verbunden?
- ▶ Privates (!) physikalisches Netzwerk
- ▶ ≥ 1 GBit/s
 - ▶ Minimale Latenz durch die Verteilung der Komponenten

Hardware

Scale out

- ▶ Neue Server sollen ins private Netzwerk aufgenommen werden.
- ▶ Rechner sollten an einem gemeinsamen Switch hängen.
- ▶ RZ mit Flexibilität ist gefragt.
- ▶ Massenhoster sind meist wenig flexibel.

Software

Software Application Server

Application Server



- ▶ = Phusion Passenger
- ▶ Speicherverbrauch abschätzen,
App-Instanzen einstellen
- ▶ Auf Upload-Endungen achten und ggf.
PHP deaktivieren
- ▶ ...

Application Server



- ▶ Einen geeigneten Session-Store wählen
 - ▶ **Cookie-St.** ist ein guter Anfang
 - ▶ **Memcached-St.**, etw. vergesslich, memcached benötigt
 - ▶ **ActiveRecord-St.**, nicht der Schnellste, benötigt keine Extra-Wurst

Application Server



- ▶ Storage einbinden
 - ▶ In App-Ordner verlinken
 - ▶ Storage-Mounts überwachen

Software

Database Server

Database Server



- ▶ MySQL, Postgres, ... CouchDB wird interessanter
- ▶ Scale-Out der DB ist immer hässlich
- ▶ Dennoch nicht den MySQL NDB Cluster verwenden, es sei denn ihr habt einen Fulltime-DB-Admin.
- ▶ DB-Server gut ausrüsten bzw. aufrüstbar gestalten
- ▶ Viele kleine, schnelle Festplatten verwenden

Software Database Replikation?

Database Server



- ▶ **Ja** aber nicht unbedingt zum Last-Scale-out.
- ▶ „Echtzeit“-Backup der Datenbank.

Database Server



- ▶ Welche Art der Replikation?
- ▶ Finger weg von Multimaster-Gefrickel, es sei denn eure Anwendung löst Datensatzkonflikte auf.
- ▶ Master/Slave ist ein guter Anfang.
- ▶ Kein automatisches Failover
- ▶ Manuelles Failover schneller als DB-Server neu aufzusetzen. Aber nehmt euch nach dem Recovery Zeit für die Reintegration des Masters.

Software

Suchen und Finden

Suchen und Finden



- ▶ Sphinx ist gut, Solr auch!
- ▶ Suchserver laufen lange neben der Anwendung
- ▶ Wenn Sie groß werden, wollen Sie ein eigenes Zimmer! Munin hilft den richtigen Zeitpunkt zu erkennen.

Software

Hintergrundverarbeitung

Hintergrundverarbeitung



- ▶ BackgroundRB, Starling&Workling, ...
- ▶ Arbeiter wollen auch irgendwann ein eigenes Zimmer.
- ▶ Bis dahin tut es auch ein beliebiger Rechner mit etwas Freiraum. CPU und RAM.

Software Deployment

Deployment



- ▶ Capistrano
- ▶ Anpassung des Deployment-Rezepts an die vorhandenen Server/Server-Rollen
- ▶ Initiales Deployment schmerzt am meisten.

Backups

Backups

OMG!

Backups



- ▶ Automatisierte Backups erhalten die Freundschaft.
- ▶ Dateisystem und Datenbank.
- ▶ Backups packen, verschlüsseln und verstauen. Duplicity sei dank!
- ▶ Ab und zu ein Backup entpacken und kontrollieren. Kaputte Backups = keine Backups.

Backups



- ▶ Wieviele Backups passen auf den Backupserver?
- ▶ Backups sollten rotiert werden.
z.B. tägliches Backup, 7 Backups werden vorhalten.
- ▶ $\text{Größtes Backup} * 7 \leq \text{Backupspace}$
- ▶ Backupspace überwachen

Monitoring

„Nur weil ich paranoid bin, heißt es nicht, dass
sie nicht hinter mir her sind.“

(Werauchimmer)

Woher weiß der Admin,
dass etwas kaputt ist?

Monitoring

Nagios

Nagios



- ▶ Nagios überwacht Schlüsselmerkmale eines jeden Rechners. So sollte er eingerichtet werden.
- ▶ Stimmt etwas nicht, benachrichtigt er der/die verantwortliche Person(en). Dito.

- ▶ **Was wird überprüft?**
 - ▶ **Swap.** Viel swap = wenig RAM = Laaaangssaaahmm.
 - ▶ **LOAD.** Mehr sysload als Cores? Prozesse müssen warten.
 - ▶ **HTTP.** Ist da noch ein Webserver?
 - ▶ **APP.** Gibt meine Anwendung noch den String XY aus?
 - ▶ **und vieles mehr!**

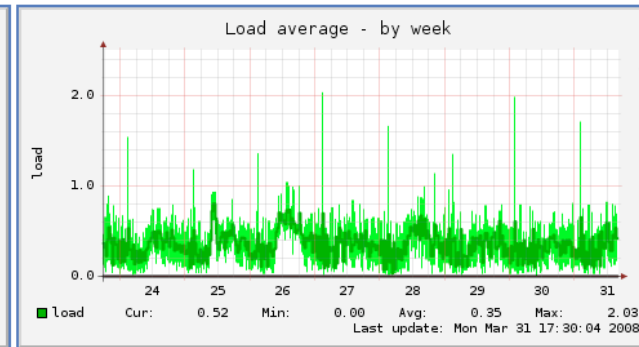
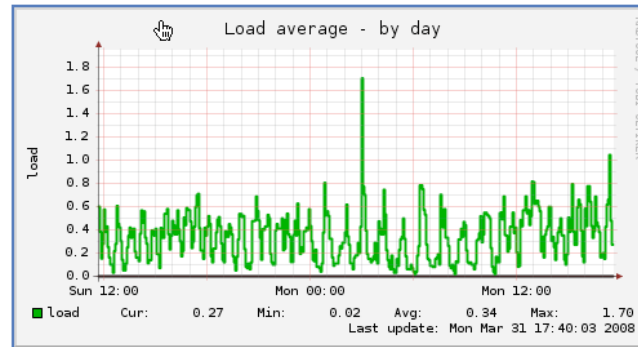
Monitoring

Ressourcen-Historie

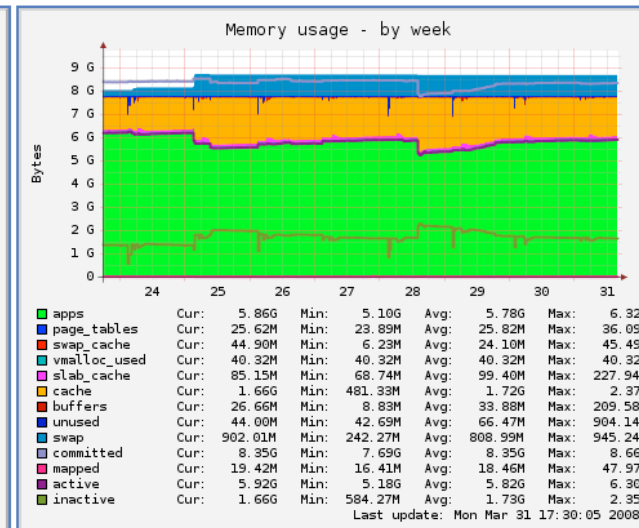
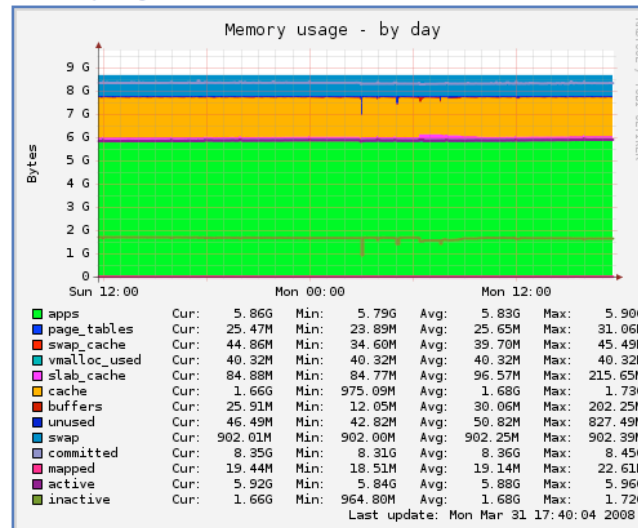
Ressourcen-Historie

► Munin

:: [Load average](#)



:: [Memory usage](#)



Ressourcen-Historie



- ▶ Zeichnet die Auslastung bestimmter Ressourcen auf.
- ▶ Trend-Analyse
- ▶ Was war nach meiner PR-Aktion auf den Servern los?
- ▶ Wann muss ich zusätzliche Server kaufen?
- ▶ Hat meine APP ein Speicherleck?

Wartung

▶ OS-Aktualisierungen.

Minor-, Major-, Distributions-Updates.

- ▶ Für größere Updates werden einzelne Server aus dem Verbund gelöst, aktualisiert und danach wieder eingefügt.

▶ Ruby Updates. Ruby, Rails, Gems, ...

▶ Rotation und Stichproben der Logs.

▶ Überprüfung der Backups.

▶ ...

Fazit



Ist ja alles gut zu wissen, aber...

Fazit



**... wer kommt neben all diesen
Dingen noch zum programmieren?**

Enterprise
 **Rails**

Fragen?

Vielen Dank für Ihre Aufmerksamkeit!

Headquarter:
<http://www.avarteq.de>

Blog:
[http://
www.treibstoff.de](http://www.treibstoff.de)

Rails Enterprise Hosting:
[http://www.enterprise-
rails.de](http://www.enterprise-rails.de)



Rails Hosting:
<http://www.railshoster.de>