

Open Source Virtualisation

Felix Krohn

<felix@kro.hn>

gpg: 0x1C246E3B

17. Januar 2010



Übersicht

- 1 Basics
- 2 Software
- 3 'Lösungen'



- 1 Basics
 - Isolierung
 - Paravirtualisierung
 - Virtualisierung
- 2 Software
- 3 'Lösungen'



Virtualisierungsarten – für und wider

Isolierung – Paravirtualisierung – “full virtualisation”

- unterschiedliche Einsatzzwecke
- unterschiedliche Methoden zur Trennung einzelner Instanzen
- unterschiedliche Performanbewert



Isolierung

- auch: Operating System-Level Virtualization
- Einzelne Instanzen werden durch Erstellen eigener Systemumgebung voneinander isoliert
- also eher eine Art besseres chroot, d.h.
 - Alle Instanzen haben das gleiche Betriebssystem
 - Alle Instanzen teilen sich einen Kernel
 - Kein/Weniger Overhead
 - Benötigte Hardware für Instanz wird virtualisiert



Beispiele

- linux-vserver
- Virtuozzo/OpenVZ
- FreeBSD jails
- Solaris zones (containers)



Paravirtualisierung

- Enge Zusammenarbeit zwischen Virtualisierungs- und Gastsystem
- Gast muß auf Virtualisierungssystem portiert werden (können)
- Gast "weiß", daß es in einer virtualisierten Umgebung ausgeführt wird
- Dadurch relativ gute Performance möglich
- Benchmarks gab's schon heute Mittag :)



Beispiele

- Xen PV
- ...
- im prinzip alles mit PV Treibern : -)



Beispiele

- Xen PV
- ...
- im prinzip alles mit PV Treibern : -)



“full virtualisation”

- Unterstützt im Prinzip alle Gastsysteme, die für \$ARCH ausgelegt sind
- Hardware übernimmt handling von Events, Exceptions und Ressourcen
- Umkehrschluss: bedingt unterstützung durch Hardware (VMX/SVM)



Beispiele

- Linux KVM
- Xen HVM
- VMware ESXi
- diverse andere (MSFT, Parallels, ...)



- 1 Basics
 - Isolierung
 - Paravirtualisierung
 - Virtualisierung
- 2 Software
- 3 'Lösungen'



linux-vserver

- patchset gegen Vanilla (experimental)
- verschiedene patchsets gegen Linux 2.4 und 2.6 mit oder ohne GRS (stable)
- linux-vserver.org/
- hat(te) einige Einschränkungen



- Open-Source Virtual Machine Monitor der ersten Stunde
- Paravirtualisierung und Virtualisierung
- XenLinux basiert ursprünglich auf Linux 2.6.18
- in-kernel DomU support seit Anfang 2008
- neuere linux-Kernel bieten pv_ops
- ports für diverse BSD, Chromium OS, ...



Xen - pv_ops

- wilde forward- und backporterei in alle Richtungen
- Seit 2.6.23 stückweise Integration in den mainline kernel
- ...
- oder auch nicht



Xen - pv_ops

- wilde forward- und backporterei in alle Richtungen
- Seit 2.6.23 stückweise Integration in den mainline kernel
- ...
- oder auch nicht



Xen - pv_ops

- wilde forward- und backporterei in alle Richtungen
- Seit 2.6.23 stückweise Integration in den mainline kernel
- ...
- oder auch nicht



Xen - pv_ops

- wilde forward- und backporterei in alle Richtungen
- Seit 2.6.23 stückweise Integration in den mainline kernel
- ...
- oder auch nicht



Xen - pv_ops

- Dom0 noch 10-20% langsamer als 2.6.18-version
- default für xen 4.x (unstable)
- dom0 patch qualität nicht ausreichend für molnar & Co
- aus Xen-Sicht: "maintainer relations issues"
- "At current progress, I would guess 2.6.33 merge window is a good target"
- we'll see that ...



Xen - pv_ops

- Dom0 noch 10-20% langsamer als 2.6.18-version
- default für xen 4.x (unstable)
- dom0 patch qualität nicht ausreichend für molnar & Co
- aus Xen-Sicht: "maintainer relations issues"
- "At current progress, I would guess 2.6.33 merge window is a good target"
- we'll see that ...



UnFUCK²⁰¹⁰

Linux KVM

- im Vanilla Kernel enthalten
- einfacher einzurichten als xen
- einfacher zu benutzen als xen
- läuft als userland prozess auf dem Host
- KSM: Kernel SamePage Merging a.k.a.
- work in progress, performance variiert je nach verwendeter Version und Prozessor
- RHEV, Proxmox, diverse Desktop Tools



- 1 Basics
 - Isolierung
 - Paravirtualisierung
 - Virtualisierung
- 2 Software
- 3 'Lösungen'



- Allen Unkenrufen zum Trotz: Projekt lebt
- 2007: Citrix kauft XenSource
- 2009: Citrix XenServer 5 wird gratis
- basiert auf xen 3.3.1 und Linux 2.6.18
- freies xen wird die von XenServer genutzte API übernehmen



in action

```

fx@gargleblaster: ~
Checking root file system...fsck 1.40-WIP (14-Nov-2006)
/dev/xvda1: clean, 8643/262144 files, 69187/524288 blocks
done.
EXT3 FS on xvda1, internal journal
Setting the system clock..
Cleaning up ifupdown....
Loading device-mapper support.
Checking file systems...fsck 1.40-WIP (14-Nov-2006)
done.
Setting kernel variables...done.
Mounting local filesystems...done.
Activating swapfile swap...done.
Setting up networking....
Configuring network interfaces...done.
INIT: Entering runlevel: 2
Starting system log daemon: syslogd.
Starting kernel log daemon: klogd.
* Not starting internet superserver: no services enabled.
Starting periodic command scheduler: crond.
Starting web server (apache2)....

Debian GNU/Linux 4.0 xentest1 hvc0
xentest1 login: []

fx@gargleblaster: ~
done.
EXT3 FS on xvda1, internal journal
Setting the system clock..
Cannot access the Hardware Clock via any known method.
Use the --debug option to see the details of our search for an access method.
Cleaning up ifupdown....
Loading device-mapper support.
Checking file systems...fsck 1.40-WIP (14-Nov-2006)
done.
Setting kernel variables...done.
Mounting local filesystems...done.
Activating swapfile swap...done.
Setting up networking....
Configuring network interfaces...done.
INIT: Entering runlevel: 2
Starting system log daemon: syslogd.
Starting kernel log daemon: klogd.
* Not starting internet superserver: no services enabled.
Starting OpenBSD Secure Shell server: sshd.
Starting periodic command scheduler: crond

fx@gargleblaster: ~
Activating swap...done.
Checking root file system...fsck 1.40-WIP (14-Nov-2006)
/dev/xvda1: clean, 8680/262144 files, 62826/524288 blocks
done.
EXT3 FS on xvda1, internal journal
Setting the system clock..
Cleaning up ifupdown....
Loading device-mapper support.
Checking file systems...fsck 1.40-WIP (14-Nov-2006)
done.
Setting kernel variables...done.
Mounting local filesystems...done.
Activating swapfile swap...done.
Setting up networking....
Configuring network interfaces...done.
INIT: Entering runlevel: 2
Starting system log daemon: syslogd.
Starting kernel log daemon: klogd.
* Not starting internet superserver: no services enabled.
Starting periodic command scheduler: crond.

Debian GNU/Linux 4.0 test2 hvc0
test2 login: []

fx@gargleblaster: ~
xentop - 22:31:05 Xen 3.2.1
4 domains: 1 running, 3 blocked, 0 paused, 0 crashed, 0 dying, 0 shutdown
Mem: 2084568k total, 1560216k used, 524352k free CPUs: 4 @ 2400MHz

```

NAME	STATE	CPU(sec)	CPU(%)	MEM(k)	MEM(%)	MAXMEM(k)	MAXMEM(%)	VCPUS
dom0	-----r	3407	0.1	205824	9.9	no limit	n/a	4
4	0	0	0	0	0	0	2149621440	
node0001	--b---	3	0.1	262144	12.6	524288	25.2	1
1	0	0	0	0	0	0	2149621440	
test2	--b---	3	0.1	524288	25.2	1048576	50.3	1
1	0	0	0	0	0	0	2149621440	
test3	--b---	1	0.0	524288	25.2	1048576	50.3	1
1	0	0	0	0	0	0	2149621440	

```

NETS NETTX(k) NETRX(k) VBDs VBD_00 VBD_RD VBD_WR SSD

```

in action

(kurzdemo xenserver gui, virtuozzo und proxmox)



Proxmox

- Distribution mit Webinterface zur Administration virtueller Server
- unterstützt sowohl OpenVZ als auch Linux KVM



VM Manager

- Virtual Machines
- Appliance Templates

Configuration

- System
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

List Create Migrate

Running Maintenance Tasks

No active Tasks

Cluster Node 'proxmox-104'

Online

VMID	Status	Name	Uptime	Disk	Memory	CPU
101	running	mailgateway-21	35 minutes	4.30%	2.67%	0.00%
106	running	zimbra.proxmox.org	2 hours	8.05%	67.58%	6.00%
107	running	webproxy	35 minutes	8.91%	48.97%	0.00%
108	running	winxp	19 hours	32.00 GB	86.91%	3.00%
109	running	win2008-server	35 minutes	32.00 GB	89.94%	1.00%
116	running	daniwiki	35 minutes	4.98%	27.38%	0.00%

Cluster Node 'proxmox-105'

Online

VMID	Status	Name	Uptime	Disk	Memory	CPU
102	running	cyan	35 minutes	5.39%	1.88%	0.00%
105	running	win2003	2 hours	32.00 GB	86.91%	0.00%
110	running	debian-etch	34 minutes	2.90%	1.38%	0.00%

Cluster Node 'proxmox-106'

Online

VMID	Status	Name	Uptime	Disk	Memory	CPU
103	running	mediawiki-intranet	35 minutes	4.86%	27.26%	0.00%
104	running	centos-5-1	36 minutes	3.91%	0.63%	0.00%
111	running	ubuntu-804-64bit	33 minutes	32.00 GB	86.91%	0.00%
112	running	exch_2007	22 minutes	100.00 GB	96.00%	0.00%



Proxmox

- linux kernel 2.6.24 (aus Ubuntu 8.04 LTS)
- Debian Lenny
- kvm v86
- jede Menge Patches für jede menge toller sachen (DRBD, Treiber-Backports, ...)
- free as in "free speech"



Proxmox - Ausblick

- im Test: 3 verschiedene Kernelversionen
- bisheriger Kernel: 2.6.24 mit KVM, OpenVZ und vielen eigenen Patches
- 2.6.18 mit RH's KVM-backport und offizieller OpenVZ-Unterstützung
- 2.6.32 mit KVM und KSM, aber ohne OpenVZ



Fragen?

„Reality is a thing of the past.“

Matrix, 1999



UnFUCK²⁰¹⁰