

2014 April 18th

VM Armageddon

Running a load of VM's on one box in a
“secure” manner.

“So why do you need to do this anyway...?”

One of my side projects (a game!) requires each player to at one point need a VM that they can play in.

Due to the nature of the game, they might need it for a long time. Or, in fact, forever. I might need to run a lot of them.

Ways to do this

- VMware
- KVM
- Abuses of QEMU (ARMv5 anyone?)
- Really strange versions of BSD 2.1
- LXC
- Docker
- OpenVZ
- OpenBSD Jail's

VMWare

- I use it already for all of my personal servers / systems
- Holy crap look at the overhead per VM!

CPU:	1 vCPU
Memory:	16 MB
Memory Overhead:	81.16 MB

KVM

- Full OS / CPU emulation
- Pretty well air gapped in security
- Overhead is large
- About 200MB per host [1]
- Plus on top the raw HDD image that can only be deduped by a fancy filesystem (see XFS / Btrfs)

[1] <http://rwmj.wordpress.com/2013/02/13/what-is-the-overhead-of-qemukvm/>

Abuses of QEMU (ARMv5)

- Another CPU arch (the x86 qemu is also quite large in overheads in the same way that KVM is)
 - can cause issues if players want to move their own binaries on to the VM
- With 8MB of RAM to the guest:
 - ARMv5: 24mb
 - i386: 63mb (2.625x larger)
- “Death by timers”
 - With 100 running idle, 1.5 cores where at 100%
 - (on a i7-2600 CPU @ 3.40GHz) x 4

Old school BSD

Some people have ported BSD 2.11 (Older than me) to more interesting and easy to emulate platforms.

The architecture it was ported to is insane (Stacks grow up etc etc)

I value my sanity more than 128kb per VM.

LXC (Linux Containers)

- Not a VM, Everyone uses the same Kernel
 - I hope that your kernel is secure `└┐`
- File system is backed by the host's file system
 - You can “oversell” the whole thing to AWS t.1 Micro levels.
- Not really designed for random strangers to run code in.
- “`echo c > /proc/sysrq-trigger`” will crash the host

Docker

- Uses LXC but in a more secure manner [1]
- Gaining momentum fast but not really production ready
- Not audited enough to know where the security aspects are (better than LXC but thats almost irrelevant)

[1] <http://blog.docker.io/2013/08/containers-docker-how-secure-are-they/>

OpenVZ

- Same ground as LXC and Docker but powered by a custom kernel.
- Trusted by a metric load of VPS Hosts and other hosting providers to be secure enough.
- 500 VM's running busybox's init consumes:
 - 768MB RAM
 - 5% of CPU

[Net,Open,Free] BSD Jails

- Have also been trusted by the hosting world and people who are paranoid for years
- Work in the same way as OpenVZ and LXC
- Pretty much the same results
- *Slightly* less RAM usage per process.

I had "Containers" before it was cool



tl;dr

- If you care about security
 - qemu / Xen / KVM
- If you care about RAM usage
 - OpenVZ / BSD Jails
 - Docker / LXC
- If you care about sanity
 - Docker
- If you are mad
 - Emulate a strange PIC system and run BSD 2.11 on it

ty. Questions?

Ben Cox

@benjojo12

ben@benjojo.co.uk

Also many ty's to @gamingrobot (Morgan Creekmore) who helped a massive deal in some of this research