

Depenguinating Your Infrastructure



Allan Jude -- ScaleEngine Inc.

Introductions

Allan Jude

- 12 Years as FreeBSD Server Admin
- Architect of the ScaleEngine CDN (HTTP and Video)
- Host of BSDNow.tv & TechSNAP.tv Podcasts
- Former Professor @ Mohawk College (2008-2011) teaching Network Engineering and Security Analysis
- Extensive work with Puppet to manage our 80+ servers in 31 data centers in 11 countries
- Lots of work with ZFS to manage large collections of videos as well as extremely large website caches (15+ million objects in ff*ff directories)

Bhyve

- One of the first users of bhyve
- Exposed many hard set limitations, like max vCPUs
- Exposed bugs with ZFS
- Requested ZFS enabled loader
- Loaned hardware (dual E5 2620v2) to Peter and Neel to develop new features
 - r260410 - Posted Interrupt Processing
 - r260532 - Virtual Interrupt Delivery

Background

ScaleEngine Inc.

- Video Streaming and HTTP CDN
- 150 TiB/month HTTP (Varnish & NGINX)
- 200 TiB/month Video (Wowza)
- Managed by Puppet
- Video streaming uses a commercial java application (Wowza Media Server)
- Runs great on FreeBSD with OpenJDK7
- However, Video transcoding uses native Linux libraries
- FreeBSD Linuxulator is only 32bit

The Problem

- Managing CentOS server is not my specialty
- Our server management system is predicated on FreeBSD and Jails
- Our 'Transcoding' Servers (Linux) did not have jails, so could not also serve HTTP traffic, or otherwise be part of our infrastructure, they were aberrations

Bhyve to the Rescue

- With FreeBSD 10.0 bhyve became an option
- Thanks to grub2-bhyve, can boot CentOS
- Few bugs, means we run -HEAD for now
- Can treat a bhyve similar to how we treat a jail in our existing management system
- Allows the host system to be a normal part of our infrastructure, serve HTTP etc.

Puppet Recipe

- A modified version of our Jail recipe (see BSDCan 2013 slides)
- Uses Michael Dexter's `vm.sh`
- Few modifications will be necessary to make it more like `ezjail`
- We really like `ezjail`

```
class bhyve($bhyve_name) {
  exec { "create_bhyve_${bhyve_name}":
    provider      => "shell",
    command       => "/usr/local/bhyve/vm create ${bhyve_name}",
    creates       => "/usr/local/bhyve/vm/${bhyve_name}",
    require       => [
      Package["sysutils/bhyve-vm"],
    ],
    notify        => Service["bhyve_${bhyve_name}"],
  }
  service { "bhyve_${bhyve_name}":
    ensure      => running,
    name        => "bhyve",
    enable      => true,
    hasrestart  => true,
    hasstatus   => true,
    start       => "/usr/local/bhyve/vm start ${bhyve_name}",
    stop        => "/usr/local/bhyve/vm stop ${bhyve_name}",
    restart     => "/usr/local/bhyve/vm restart ${bhyve_name}",
    status      => "/usr/local/bhyve/vm status ${bhyve_name}",
  }
}
```


Performance

Performance has been very good, better than expected considering Peter keeps saying they never test performance

Network is fine to gigabit, CPU utilization is a little high, will be interesting to see what e1000 driver and netmap can do for this

openssl speed sha256

Lenovo T530: i5-3320M (2.60 GHz), 16 GB RAM

Host:

The 'numbers' are in 1000s of bytes per second processed.

type	16 bytes	64 bytes	256 bytes	1024 bytes	8192 bytes
sha256	43492.59k	104753.70k	181075.46k	223218.66k	239949.14k

Guest:

The 'numbers' are in 1000s of bytes per second processed.

type	16 bytes	64 bytes	256 bytes	1024 bytes	8192 bytes
sha256	43314.18k	103957.81k	180355.07k	223440.63k	240056.17k

Conclusions

- Deploying Linux on top of FreeBSD gives you the tools you are used to
- Means the machine can still do FreeBSD things in addition to Linux
- Great for 'Appliances', we're considering moving our phone server to bhyve
- Performance seems to be 100%

Podcasts

BSDNow.tv is a weekly video podcast featuring News, Interviews and Tutorials about the BSD family of Operating Systems. Hosted by Kris Moore (founder of PC-BSD) and Myself.

TechSNAP.tv is a weekly sysadmin video podcast covering an OS agnostic range of security and production issues of interest to those working, studying or interested in the field.

Twitter: @allanjude

Email: freebsd@allanjude.com