The OpenBSD hypervisor in the wild, a short story
Sorry...
Who dis

- Began at XS4ALL (ISP) in 1995
- Working for $vendor since 1998
- Started with FreeBSD in 1998
- Hosting / Co-Location since 1999
What about you?

- who is using OpenBSD?
- who is using \texttt{vmm(4)/vmd(8)}?
- who is on OpenBSD Amsterdam? ;)

How it all began

Always on the lookout for easy segmentation and virtualisation

- Started with and still using jails(8)
- Used bhyve(8)
- Using vmm(4)/vmd(8)
How it all began

• Spare rackspace
• Spare hardware
• Spare IP space
• Domain with something BSD
• Contributing back to the community
• How far can we take this
• Let's go!
Something BSD

mischa@x270:~ $ whois openbsd.amsterdam
Domain Name: openbsd.amsterdam
Registry Domain ID: DNML139608-SIDN
Registrar WHOIS Server: whois.rrproxy.net
Registrar URL: http://www.key-systems.net/tld/amsterdam
Updated Date: 2019-05-30T00:15:03Z
Creation Date: 2010-05-30T17:49:38Z
Registry Expiry Date: 2020-05-30T17:40:38Z
Registrar: Key-Systems LLC
Registrar IANA ID: 1345
Registrar Abuse Contact Email: abuse@key-systems.net
Registrar Abuse Contact Phone: +49.694939850
Domain Status: clientTransferProhibited https://iann.org/opp#clientTransferProhibited
Registry Registrant ID: Redacted for privacy
Registrant Name: Redacted for privacy
Registrant Organization: M Peters
Registrant Street: Redacted for privacy
Registrant City: Redacted for privacy
Registrant State/Province:
Registrant Postal Code: Redacted for privacy
Registrant Country: NL
Registrant Phone: Redacted for privacy
Registrant Phone Ext:
Registrant Fax: Redacted for privacy
Registrant Fax Ext:
Where is it?

- Amsterdam!
- XS4ALL (KPN) Datacenter
- Dell R610 -> Foundry FLS448 -> Foundry MLX-4
Started on Twitter

OpenBSD Amsterdam
@OpenBSDAms

#OpenBSD VPS #6 on #OpenBSD #VMM #RUNBSD

2:13 PM - 10 Jun 2018

Xeon(R) CPU E3-1220 V2 @ 3.10GHz w/ 8G RAM
What are people willing to pay

How much are you willing to pay for a OpenBSD VM in Amsterdam with 512MB/50G/v4+v6

- 57% < 5
- 39% => 5 < 10
- 4% => 10 < 15

148 votes • Final results

4:39 PM - 11 Jun 2018

9 Retweets 4 Likes
Proper machine online

For all the people wh voted, server #2 is ready!! #announcement #OpenBSD #RUNBSD

cert.security.amsterdam/server2.html

Start contributing to #OpenBSD while running a VPS!

OpenBSD Amsterdam @OpenBSDAms
How much are you willing to pay for a OpenBSD VM in Amsterdam with 512MB/50G/v4+v6
Show this thread

7:59 AM - 1 Jul 2018

3 Retweets 4 Likes

Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40GHz w/ 32G RAM
First donation

Donated to the OpenBSD Foundation! Thank you all for making this possible!

#OpenBSD #RUNBSD
Statistics

- Latest donation €370
- 2018 €1850 (6 months)
- 2019 €1700 (YTD)
- Total: €3550
- Active Hosts: 8
- Active VMs: 280
What do you get?

- Opinionated VM
What do you get?

• Opinionated VM
• 512M RAM
• 50G Disk
• IPv4 assigned via DHCP
• IPv6 statically assigned (/56 is assigned to a host)
  • Host is gateway for each VM
Setup
BASE

Everything we use is in base
BASE

Everything we use is in base
- perl(1)
- vmm(4)/vmd(8)
- dhcppd(8)
- autoinstall(8)
- siteXX.tgz
- httpd(8)
- sensorsd(8)
- vi(1)
perl(1)

- /etc/vm.conf
- /etc/dhcpd.conf
- /var/www/htdocs/install/<MAC>-install.conf
- /etc/doas.conf
- user creation
- vm image creation
vm.conf(5)

socket owner : _vmdusers

switch "uplink_vlan931" {
  interface bridge931
}

vm "vm13" {
  disable
  owner alice
  disk "/var/vmm/vm13.qcow2"
  interface tap {
    switch "uplink_vlan931"
    lladdr fe:e1:bb:f1:c8:01
  }
}

dhcpd.conf(5)

option domain-name "openbsd.amsterdam";
option domain-name-servers 46.23.80.26;

subnet 46.23.93.0 netmask 255.255.255.0 {
    option routers 46.23.93.1;
    server-name "server8.openbsd.amsterdam";

    host vm13 {
        hardware ethernet fe:e1:bb:f1:c8:13;
        fixed-address 46.23.93.13;
        filename "auto_install";
        option host-name "puffy.openbsd.amsterdam";
    }
}

EuroBSDCon 2019 - Lillehammer
System hostname = puffy.openbsd.amsterdam
Password for root = [password]
Which speed should com0 = 115200
Network interfaces = vio0
IPv4 address for vio0 = dhcp
IPv6 address for vio0 = 2a03:6000:6f64:613::13
IPv6 default router = 2a03:6000:6f64:613::1
Setup a user = alice
Password for user = [password]
Public ssh key for user = ssh-ed25519 AAAAC3N...U7KKt alice@domain.tld [password]
Which disk is the root disk = sd0
What timezone are you in = Europe/Amsterdam
Location of sets = http
Server = server8.openbsd.amsterdam
Set name(s) = -x* +xb* +xf* +site*
Continue anyway = yes
Continue without verification = yes
siteXX.tgz

- `installurl(5)`:
  
  https://cdn.openbsd.org/pub/OpenBSD

- `sysctl.conf(5)`:
  
  kern.timecounter.hardware=tsc

- `rc.conf.local(8)`:
  
  ntpd_flags="-s"
sndiod_flags=NO
httpd(8)

/etc/httpd.conf

server "default" {
    listen on * port 80
    root "/htdocs/autoinstall"
    location "/pub/OpenBSD/6.5/amd64/*" {
        root "/htdocs/6.5"
        request strip 4
        directory { auto index }
    }
}

sensorsd(8)

/etc/sensorsd.conf

drive:command=/etc/sensorsd/drive %t %n %2 %s

#!/bin/sh
#
#   %t     The type of sensor.
#   %n     The sensor number.
#   %2    The sensor's current value.
#   %s    The sensor status.
#
#drive:command=/etc/sensorsd/drive %t %n %2 %s
#Subject: Sensor drive0 changed
#Raid state: drive0 online OK
echo "Current raid state: ${1}${2} ${3} ${4}" | mail -s "$(hostname) ${1}${2} ${4}" -r noreply@domain.tld mischa@domain.tld
Deploying!
server8:~ # cat _deploy.conf
# Server config for <MAC>-install.conf
SERVER="server8"
DOMAIN="openbsd.amsterdam"
# IP / MAC config
IP_PREFIX="46.23.93"
IP_START=100
IPV6_PREFIX="2a03:6000:6f64"
IPV6_START=600
MAC_PREFIX="fe:e1:bb:f1:c8"
# .conf locations
VMS="/home/mischa/vms"
ETC="/etc"
IMAGES="/var/vmm"
HTDOCS="/var/www/htdocs/default"
# vm.conf
MEMORY="512M"
DISKSIZE="50G"
FORMAT="qcow2"
VMDUSERS="_vmdusers"
SWITCH="uplink_vlan931"
INTERFACE="bridge931"
# dhcpcd.conf
ROUTER="46.23.93.1"
DNS="46.23.80.26"
SUBNET="46.23.93.0"
NETMASK="255.255.255.0"
Deploy-flow

- form > email > file
- run deploy.pl on the host
- restart dhcpd
- reload vmd
- start vm
- run installer - Hit (A)
Form

Type-in your name *
Alice

RAM
Standard 512M

email *
alice@example.com

HDD
Standard 50G

and your SSH public key *
ssh-ed25519 FRhkxldn1...sDZUdP

disk format
qcow2

referral code
OBSD-XXXX-XXXX

hostname *
example

note?
I like VMs

username *
alice

Book it
Email > ~/vms/vm13.txt

date="2019/09/21"
payment=""
donated=""
name="Alice"
email="alice@domain.tld"
sshkey="ssh-ed25519 AAAAC3N...U7KKt alice@domain.tld"
hostname="puffy"
username="alice"
note=""
memory="512M"
disk2=""
format="qcow2"
referral=""
server8:~ # deploy.pl
autoinstall(8) files:
useradd(8) creation:
   alice
vmm(4)/vmd(8) files:
   vm13 /var/vmm/vm13.qcow2 created (size 50G)
server8:~ # vmctl reload
server8:~ # rcctl restart dhcpd
dhcpd(ok)
dhcpd(ok)
server8:~ # vmctl start -c vm13
Connected to /dev/ttyph (speed 115200)
The Regents of the University of California. All rights reserved.
Copyright (c) 1995-2019 OpenBSD. All rights reserved. https://www.OpenBSD.org
OpenBSD 6.5 (RAMDISK_CD) #3: Sat Apr 13 14:55:38 MDT 2019
deraadt@amd64.openbsd.org:/usr/src/sys/arch/amd64/compile/RAMDISK_CD
real mem = 520093696 (496MB)
avail mem = 500412416 (477MB)
mainbus0 at root
bios0 at mainbus0
acpi at bios0 not configured
cpu0 at mainbus0: (uniprocessor)
cpu0: Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40GHz, 2401.04 MHz, 06-3f-02
...
 sd0 at scsibus0 targ 0 lun 0: <VirtIO, Block Device, > SCSI3 0/direct fixed
sd0: 51200MB, 512 bytes/sector, 104857600 sectors
...
root on rd0a swap on rd0b dump on rd0b
erase ^?, werase ^W, kill ^U, intr ^C, status ^T

Welcome to the OpenBSD/amd64 6.5 installation program.
(I)nstall, (U)pgrade, (A)utoinstall or (S)hell? a

EuroBSDCon 2019 - Lillehammer
What did we find
socket owner

Set the control socket owner to the specified group.

https://marc.info/?l=openbsd-cvs&m=153003284400760&w=2
tap(4) interfaces

$ cd /dev
$ ls -al tap*
crw-------  1 root  wheel   93,   0 Apr 25 09:28 tap0
crw-------  1 root  wheel   93,   1 Apr 25 09:28 tap1
crw-------  1 root  wheel   93,   2 Apr 25 09:28 tap2
crw-------  1 root  wheel   93,   3 Apr 25 09:28 tap3

$ for i in $(jot 50 4 50); do doas sh MAKEDEV tap$i; done
share password??

jot -rcs '' 20 33 126

-\texttt{r} \quad \text{Generate random data. By default, jot generates sequential data}

-\texttt{c} \quad \text{This is an abbreviation for -w \%c.}

-\texttt{w word} \quad \text{Print word with the generated data appended to it. Octal, hexadecimal, exponential, ASCII, zero-padded, and right-adjusted representations are possible by using the appropriate printf(3) conversion specification inside word, in which case the data is inserted rather than appended.}

-\texttt{s string} \quad \text{Print data separated by string. Normally, newlines separate data.}

added to ~/.ssh/authorized_keys

ascii http://www.ascii-table.com/
stopping VMs

Used to do this with:

$ vmctl show | for i in $(awk '!/ID| - / {print $1}'); do doas vmctl stop $i; sleep 30; done

Now there is:

$ doas vmctl stop -aw

---

https://marc.info/?l=openbsd-cvs&m=153806854327569&w=2
starting VMs

$ vmctl show | for i in $(awk '!/ID/ {print $1}'); do doas vmctl start $i; sleep 30; done

Or

$ vmctl show | for i in $(awk '!/ID/ {print $1}'); do doas vmctl start $i; sleep 90; done
arpq

$ sysctl net.inet.ip.arpq.drops
net.inet.ip.arpq.drops=524

$ sysctl net.inet.ip.arpq.maxlen
net.inet.ip.arpq.maxlen=50

$ doas sysctl net.inet.ip.arpq.maxlen=1024
What users experience
Clock drift

Clock drifts, sometimes more severe.

# Sync clock every 15 minutes
*/15 * * * * /usr/sbin/rdate -s pool.ntp.org

https://openbsd.amsterdam/clock.html
High CPU interrupts

VMs have a constant high intr CPU state:

CPU states: 0.0% user, 0.0% nice, 0.1% sys, 0.0% spin, 98.0% intr, 1.9% idle
Connectivity drops

Cron

`*/5 * * * * /sbin/ping -c3 <gateway> > /dev/null`

Cron + tmux

`@reboot /usr/bin/tmux new -d 'while true; do ping -i5 <gateway>; done' \;`
Unresponsive VM

When `vmctl stop -f <vm-name>` doesn't work.\(^6.5\)

/etc/doas.conf

permit nopass <vm-owner> as root cmd pkill args -9 -f <vm-name>

User runs:

$ doas pkill -9 -f vm13

\(^6.5\) `vmctl stop <vm-name> -f` ([https://marc.info/?l=openbsd-cvs&m=155916557307145&w=2](https://marc.info/?l=openbsd-cvs&m=155916557307145&w=2))
Wishlist / Future

- iPXE
- no/less clock drift
- using `switch(4) / L3`?
- automate more
- deploy 300 VMs!
Couldn't be possible without!

Mike Larkin (@mlarkin2012)
Reyk Flöter (@reykfloeter)
Carlos Cardenas (@cobracmdr)
Stefan Kempf
Claudio Jeker
Jasper Lievisse Adriaanse (@jasper_la)
Ori Bernstein (@oribernstein)
Roman Zolotarev (@romanzzolotarev)
Thank you!

More information https://openbsd.amsterdam
Deploy script https://git.high5.nl/deploy.pl
Twitter https://twitter.com/OpenBSDAms
Mastodon https://bsd.network/@OpenBSDAms
Just URLs

More information https://openbsd.amsterdam
Deploy script https://git.high5.nl/deploy.pl
Twitter https://twitter.com/OpenBSDAms
Mastodon https://bsd.network/@OpenBSDAms