



# The OpenBSD hypervisor in the wild, a short story

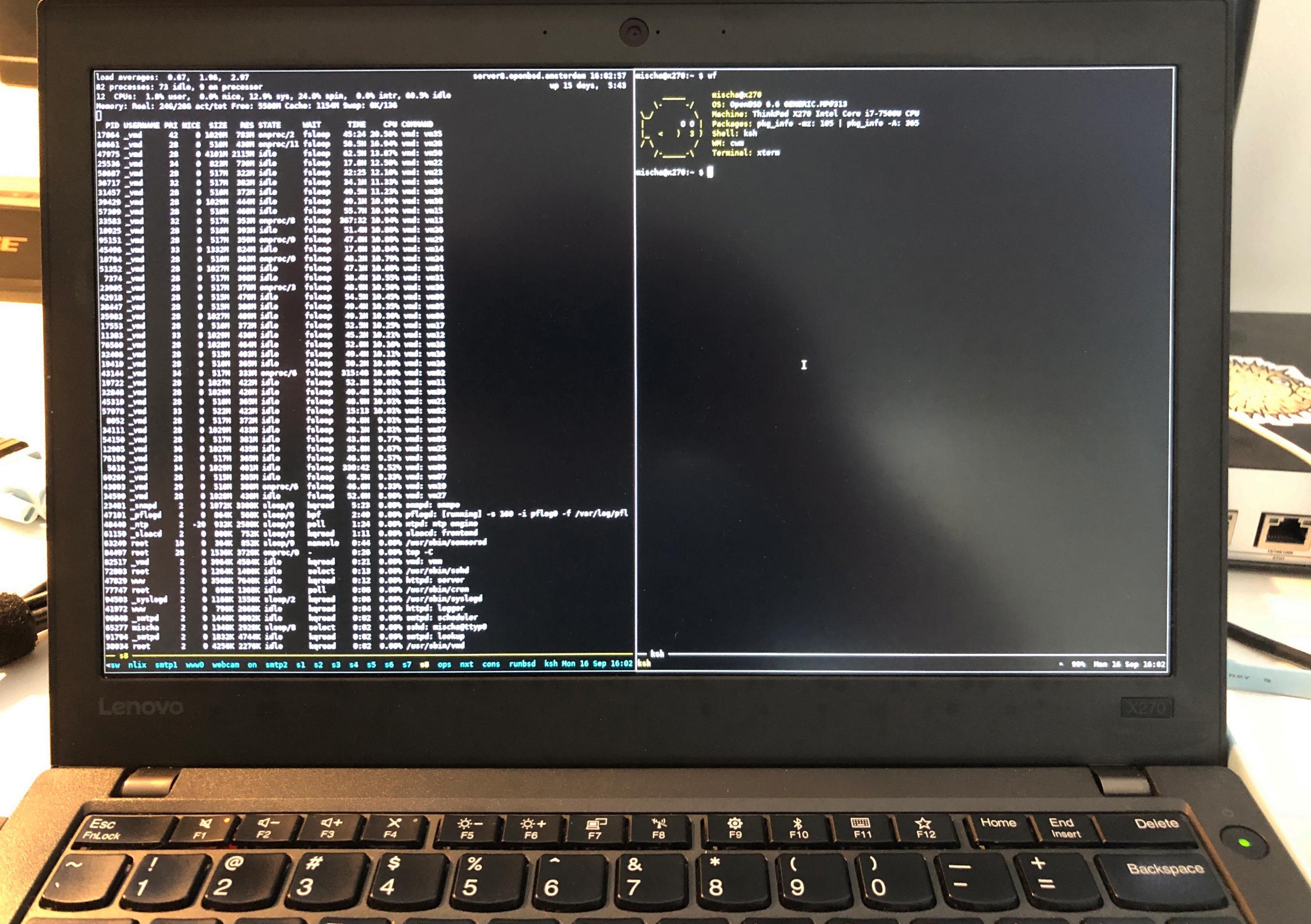
# Sorry...

load averages: 0.67, 1.96, 2.97  
 82 processes: 73 idle, 9 on processor  
 12 CPUs: 1.0% user, 0.0% nice, 12.9% sys, 24.0% spin, 0.0% intr, 68.3% idle  
 Memory: Real: 240/266 act/tot Free: 5500M Cache: 1125M Swap: 0K/12G  
 PID USERNAME PRI NICE SIZE RSS STATE WAIT TIME CPU COMMAND  
 17004 \_vnd 42 0 1820M 763M enproc/2 fleep 45:24 26.50% vnd: ve25  
 60061 \_vnd 28 0 516M 430M enproc/11 fleep 58.3M 16.94% vnd: ve26  
 47975 \_vnd 28 0 410M 2115M idle fleep 62.3M 13.67% vnd: ve20  
 25530 \_vnd 24 0 82M 730M idle fleep 17.0M 12.59% vnd: ve22  
 50087 \_vnd 28 0 517M 322M idle fleep 32:25 12.16% vnd: ve23  
 36717 \_vnd 22 0 517M 362M idle fleep 24.2M 11.33% vnd: ve24  
 31457 \_vnd 28 0 516M 372M idle fleep 40.3M 11.23% vnd: ve20  
 30420 \_vnd 28 0 162M 444M idle fleep 40.3M 10.99% vnd: ve26  
 57269 \_vnd 28 0 516M 460M idle fleep 55.7M 10.04% vnd: ve25  
 23583 \_vnd 22 0 517M 250M enproc/8 fleep 267:32 10.04% vnd: ve13  
 10025 \_vnd 28 0 516M 302M idle fleep 51.4M 10.00% vnd: ve26  
 95151 \_vnd 28 0 517M 350M enproc/9 fleep 47.0M 10.00% vnd: ve20  
 45496 \_vnd 23 0 1332M 824M idle fleep 17.0M 10.00% vnd: ve14  
 10784 \_vnd 28 0 516M 363M enproc/9 fleep 40.2M 10.79% vnd: ve24  
 51252 \_vnd 28 0 1627M 460M idle fleep 47.2M 10.00% vnd: ve21  
 7374 \_vnd 28 0 517M 300M idle fleep 30.4M 10.55% vnd: ve21  
 22965 \_vnd 28 0 517M 370M enproc/3 fleep 40.3M 10.55% vnd: ve20  
 42918 \_vnd 28 0 515M 470M idle fleep 34.3M 10.45% vnd: ve20  
 28447 \_vnd 28 0 515M 360M idle fleep 40.4M 10.35% vnd: ve25  
 29503 \_vnd 28 0 1627M 460M idle fleep 40.3M 10.30% vnd: ve26  
 17553 \_vnd 28 0 516M 372M idle fleep 52.3M 10.25% vnd: ve17  
 11383 \_vnd 33 0 1629M 430M idle fleep 45.2M 10.21% vnd: ve12  
 76560 \_vnd 28 0 1628M 464M idle fleep 52.0M 10.10% vnd: ve18  
 32406 \_vnd 28 0 515M 460M idle fleep 40.4M 10.11% vnd: ve20  
 18410 \_vnd 28 0 516M 260M idle fleep 30.2M 10.00% vnd: ve16  
 43144 \_vnd 34 0 517M 332M enproc/6 fleep 315:40 10.00% vnd: ve22  
 19722 \_vnd 28 0 1627M 422M idle fleep 52.3M 10.01% vnd: ve11  
 22848 \_vnd 28 0 1629M 420M idle fleep 40.4M 10.01% vnd: ve20  
 45310 \_vnd 28 0 516M 340M idle fleep 40.0M 10.01% vnd: ve21  
 57070 \_vnd 33 0 520M 422M idle fleep 15:13 10.01% vnd: ve22  
 8052 \_vnd 28 0 517M 372M idle fleep 50.0M 9.93% vnd: ve24  
 34111 \_vnd 28 0 1629M 430M idle fleep 40.3M 9.93% vnd: ve27  
 54150 \_vnd 28 0 517M 362M idle fleep 45.0M 9.77% vnd: ve26  
 12005 \_vnd 26 0 1629M 430M idle fleep 35.0M 9.67% vnd: ve25  
 76100 \_vnd 28 0 517M 360M idle fleep 40.0M 9.57% vnd: ve26  
 5616 \_vnd 34 0 1629M 461M idle fleep 330:42 9.52% vnd: ve26  
 69260 \_vnd 28 0 515M 360M idle fleep 40.0M 9.30% vnd: ve27  
 43903 \_vnd 29 0 516M 360M enproc/8 fleep 46.0M 9.30% vnd: ve20  
 34590 \_vnd 28 0 1629M 430M idle fleep 52.0M 9.00% vnd: ve27  
 23481 \_mpmd 2 0 2072K 2000K sleep/0 hread 5:23 0.00% mpmd: sleep  
 47301 \_pflogd 4 0 0.0K 500K sleep/0 kpf 2:00 0.00% pflogd: [running] -> 300 -i pflogd -f /var/log/pfl  
 48440 \_ntp 2 -20 932K 2500K sleep/0 poll 1:24 0.00% ntpd: ntp engine  
 61339 \_slacd 2 0 0.0K 752K sleep/0 hread 1:11 0.00% slacd: frontend  
 63240 root 10 0 304K 852K sleep/0 monole 0:44 0.00% /usr/sbin/monole  
 84497 root 28 0 1536K 3720K enproc/0 - 0:26 0.00% top -c  
 62517 \_vnd 2 0 364K 4504K idle hread 0:21 0.00% vnd: vnm  
 72603 root 2 0 1280K 1400K idle select 0:13 0.00% /usr/sbin/child  
 47820 www 2 0 2500K 7640K idle hread 0:12 0.00% httpd: server  
 77747 root 2 0 0.0K 1300K idle poll 0:06 0.00% /usr/sbin/cron  
 94553 \_syslogd 2 0 1108K 1550K sleep/2 hread 0:06 0.00% /usr/sbin/syslogd  
 41972 www 2 0 700K 2000K idle hread 0:04 0.00% httpd: logger  
 95646 \_netpd 2 0 1440K 2002K idle hread 0:02 0.00% netpd: scheduler  
 65277 mischa 2 0 1340K 2028K sleep/8 select 0:02 0.00% schd: mischagttypes  
 91794 \_netpd 2 0 1832K 4744K idle hread 0:02 0.00% netpd: lookup  
 30034 root 2 0 4250K 2270K idle hread 0:02 0.00% /usr/sbin/vmd  
 - 58

ischaq@270: ~ \$ lf

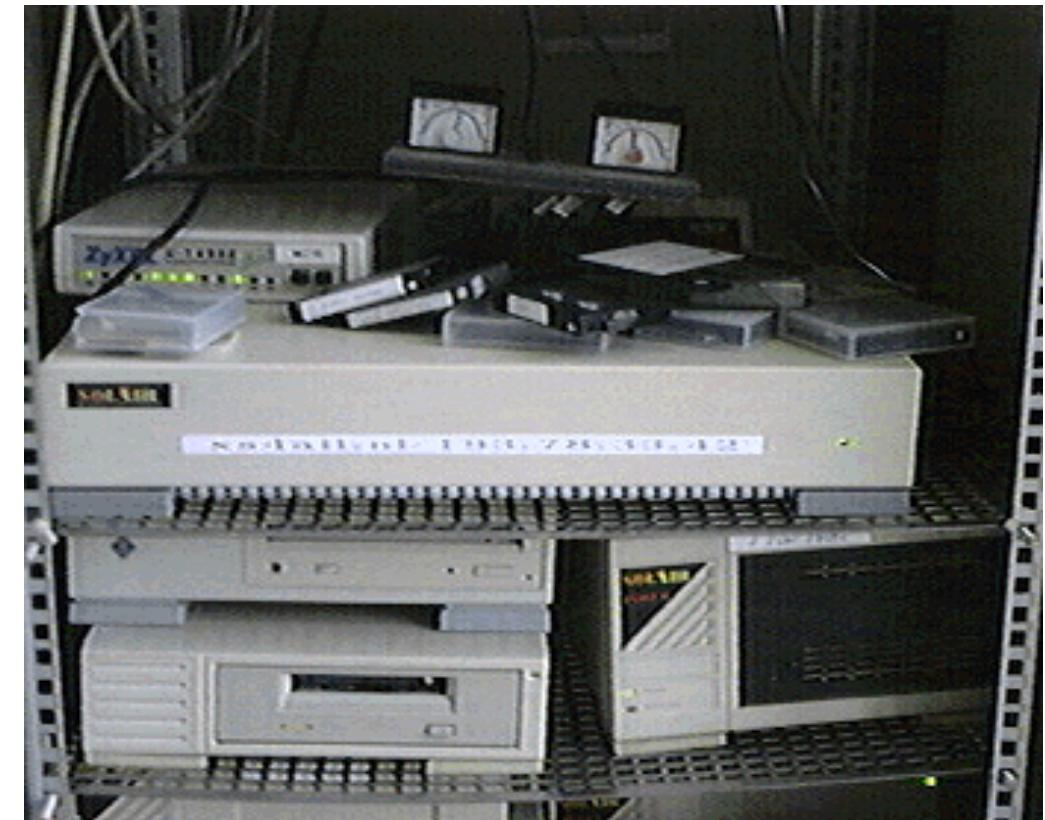
```
mische@x270: ~$ OS: OpenSUSE 6.0 GENERIC.HW#313  
Machine: ThinkPad X270 Intel Core i7-7560U CPU  
Packages: pkg_info -rxz: 105 | pkg_info -A: 365  
Shell: ksh  
WM: CDE  
Terminal: xterm
```

Page 270 - 8



# 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 [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 raxspace
- Spare hardware
- Spare IP space
- Domain with something BSD
- Contributing back to the community
- How far can we take this
- Let's go!

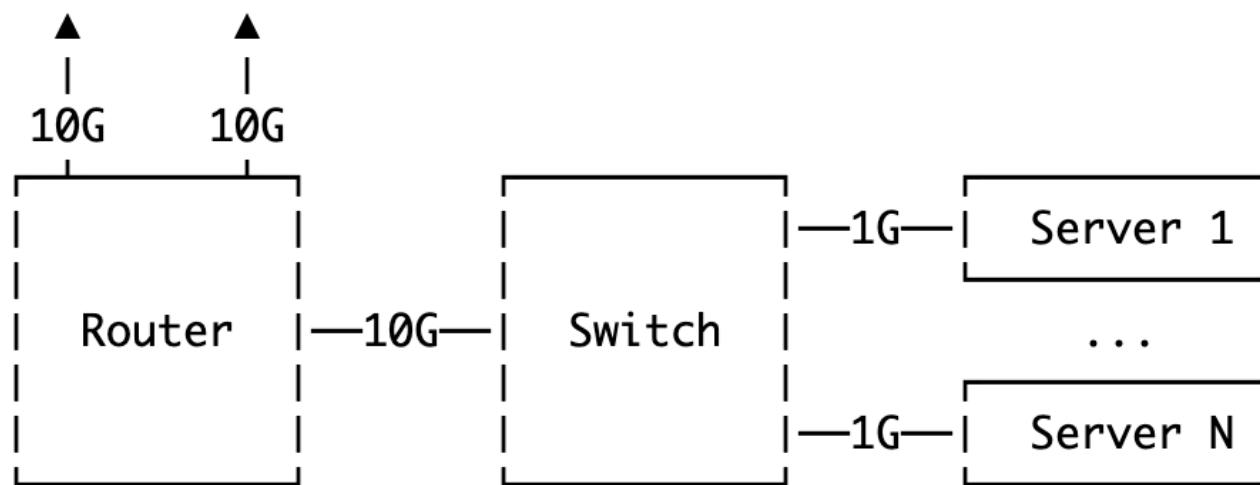


# Something BSD

```
mischa@x270:~ $ whois openbsd.amsterdam
Domain Name: openbsd.amsterdam
Registry Domain ID: DNM_139668-SIDN
Registrar WHOIS Server: whois.rrpproxy.net
Registrar URL: http://www.key-systems.net/tld/amsterdam
Updated Date: 2019-05-30T00:15:03Z
Creation Date: 2018-05-30T17:40:38Z
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: +49.68949396850
Domain Status: clientTransferProhibited https://icann.org/epp#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



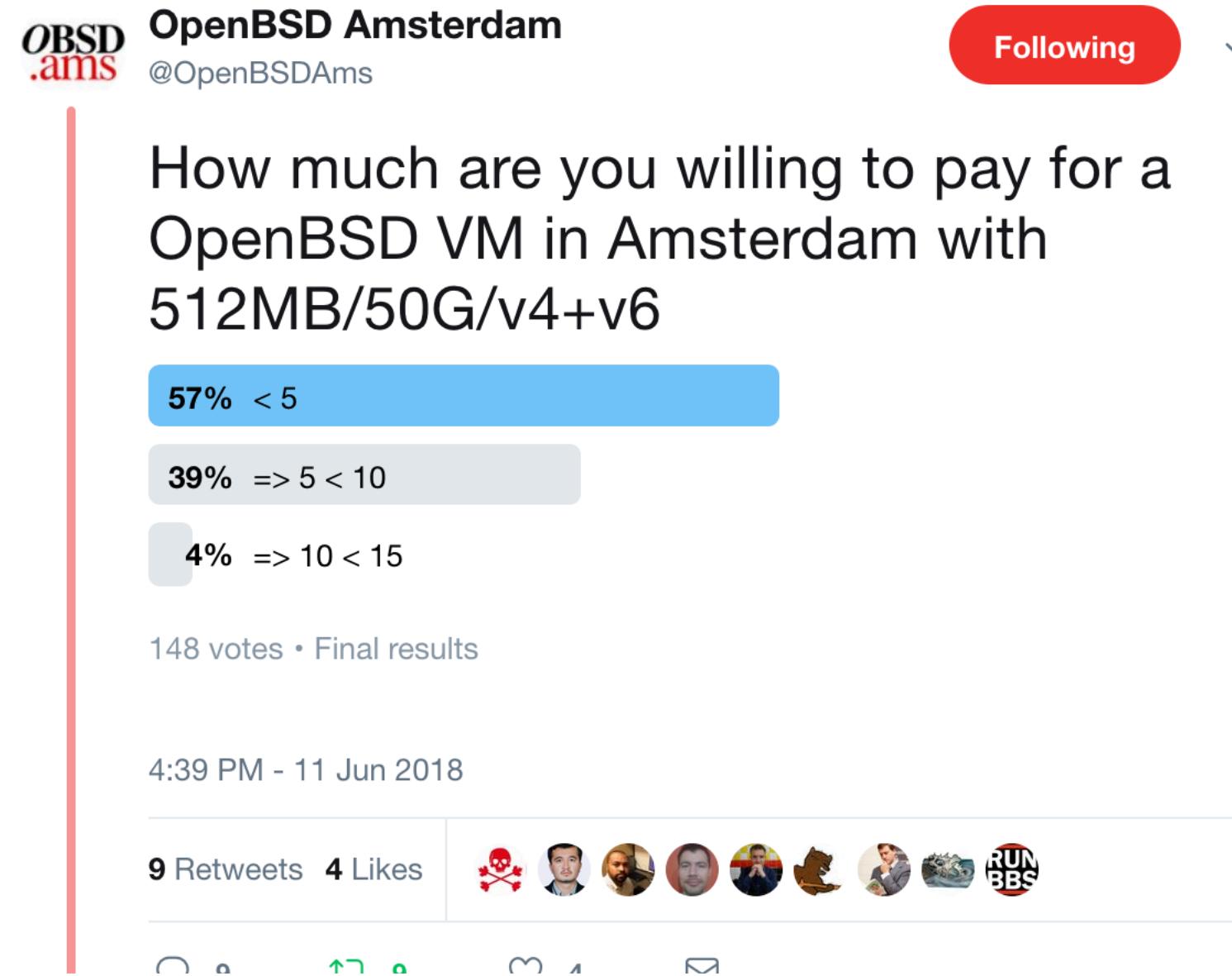
# Started on Twitter



---

<sup>s<sup>1</sup></sup> Xeon(R) CPU E3-1220 V2 @ 3.10GHz w/ 8G RAM

# What are people willing to pay



# Proper machine online

 **OpenBSD Amsterdam**  
@OpenBSDAms Following ▾

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

[openbsd.amsterdam/server2.html](http://openbsd.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 

---

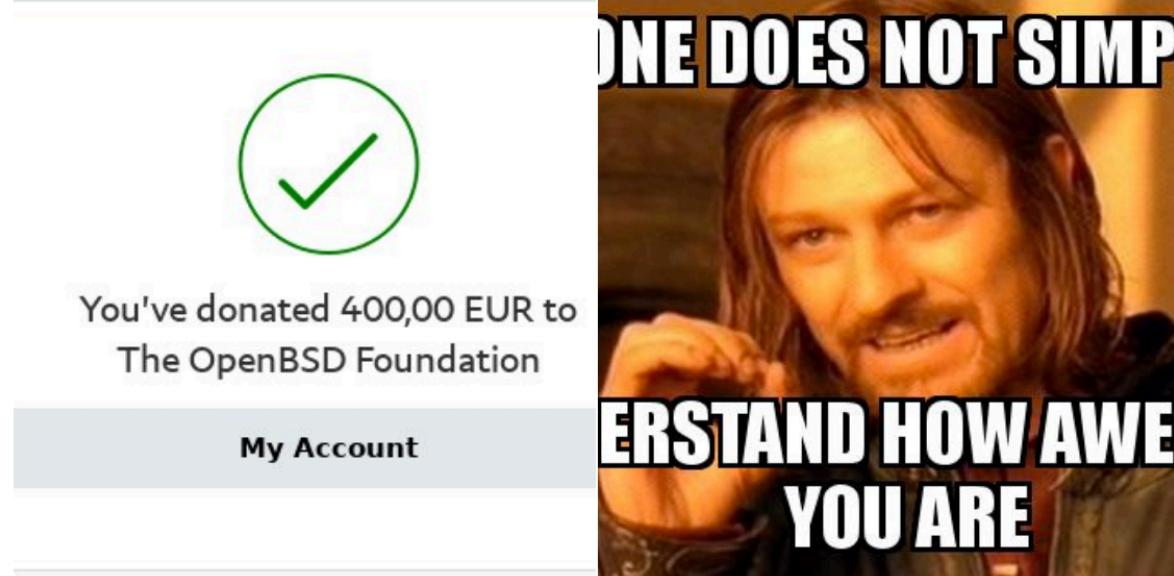
s<sup>2</sup> Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40GHz w/ 32G RAM

# First donation

 OpenBSD Amsterdam  
@OpenBSDAms

Following ▾

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



You've donated 400,00 EUR to The OpenBSD Foundation

My Account

8:46 PM - 27 Jul 2018

5 Retweets 31 Likes

The image shows a tweet from the account @OpenBSDAms. The tweet text reads: "Donated to the OpenBSD Foundation! Thank you all for making this possible! #OpenBSD #RUNBSD". Below the text is a screenshot of a donation confirmation page showing a green checkmark and the message "You've donated 400,00 EUR to The OpenBSD Foundation". To the right of the text is a meme image featuring a man with a mustache and long hair, with the text "ONE DOES NOT SIMPLY UNDERSTAND HOW AWESOME YOU ARE" overlaid.

# Statistics

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

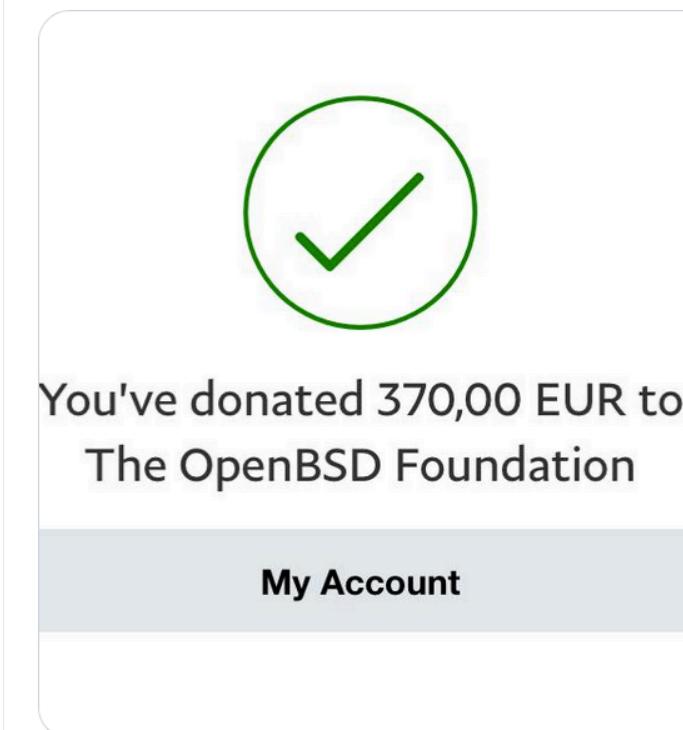
 OpenBSD Amsterdam  
@OpenBSDAms

We just donated €370 to the #OpenBSD Foundation, totaling €3550 so far.

In August 16 new VMs were added and 14 VMs were renewed.

Thank you all!

#OpenBSD #RUNBSD



3:16 PM · Sep 5, 2019 · TweetDeck

View Tweet activity

15 Retweets 46 Likes

# 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



CATS : ALL YOUR BASE ARE BELONG  
TO US.

# BASE

Everything we use is in base

- perl(1)
- vmm(4)/vmd(8)
- dhcpcd(8)
- autoinstall(8)
- siteXX.tgz
- httpd(8)
- sensorsd(8)
- vi(1)



CATS : ALL YOUR BASE ARE BELONG  
TO US.

## 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";
    }
}
```

# autoinstall(8)

/var/www/htdocs/autoinstall/fe:e1:bb:f1:c8:13-install.conf

```
# vm13-install.conf
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 dhcpcd
- reload vmd
- start vm
- run installer - Hit (A)

# Form

Type-in your name \*

Alice

email \*

alice@example.com

and your SSH public key \*

ssh-ed25519 FRhkxldn1...sDZUdP

hostname \*

example

username \*

alice

RAM

Standard 512M

HDD

disk format

Standard 50G

qcow2

referral code

OBSD-XXXX-XXXX

note?

I like VMs

**Book it**

## Email > ~/vms/vm13.txt

```
date="2019/09/21"
payment=""
donated=""
name="Alice"
email="alice@domain.tld"
sshkey="ssh-ed25519 AAAAC3N...U7Kt alice@domain.tld"
hostname="puffy"
username="alice"
note=""
memory="512M"
disk2=""
format="qcow2"
referral=""
```

# deploy.pl

```
server8:~ # deploy.pl
autoinstall(8) files:
    vm13 /var/www/htdocs/default/fe:e1:bb:f1:c8:13-install.conf created
useradd(8) creation:
    alice
vmm(4)/vmd(8) files:
    vm13 /var/vmm/vm13.qcow2 created (size 50G)
```

```
server8:~ # vmctl reload
server8:~ # rcctl restart dhcpcd
dhcpcd(ok)
dhcpcd(ok)
server8:~ # vmctl start -c vm13
Connected to /dev/ttyp0 (speed 115200)
Copyright (c) 1982, 1986, 1989, 1991, 1993
        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
```

# What did we find

# socket owner

The screenshot shows a tweet from the account **OpenBSD src Changes** (@OpenBSD\_src). The tweet, posted at 12:25 PM - 26 Jun 2018, contains the following text:

reyk@ modified usr.sbin/vmd: Add "socket owner" to allow changing the owner of the vmd control socket. This allows to open vmctl control or console access to other users that are not in group wheel. Access for non-root users still defaults to read-only actions unless you chang...

The tweet has received 5 Retweets and 9 Likes. The interface shows a red "Following" button and a dropdown arrow.

socket owner :group

Set the control socket owner to the specified group.

---

socket <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
```

- r Generate random data. By default, jot generates sequential data
- c This is an abbreviation for -w %c.
- w word 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.
- s string Print data separated by string. Normally, newlines separate data.

added to `~/.ssh/authorized_keys`



Reyk Flöter  
@reykfloeter

Replying to @blakkheim @NicoSchottelius and @datacenterlight

Better? Login is puffy@, cloud-agent can now generate a random password and write it as a comment into `.ssh/authorized_keys` - I shamelessly stole the idea from [@OpenBSDAms](#).

```
$ ssh puffy@2a0a:e5c0:2:2:0:c8ff:fe68:bf16
Last login: Wed Jun  5 22:05:52 2019 from 2001:8e0:2002:8913:2eaa:c7ca:253c:d589
OpenBSD 6.5 (GENERIC.MP) #0: Wed Apr 24 23:38:54 CEST 2019
```

Welcome to OpenBSD: The proactively secure Unix-like operating system.

Please use the sendbug(1) utility to report bugs in the system. Before reporting a bug, please try to reproduce it with the latest version of the code. With bug reports, please try to ensure that enough information to reproduce the problem is enclosed, and if a known fix for it exists, include that as well.

```
vm0200c868bf16$ head -1 .ssh/authorized_keys
# XXXXXXXXXXXXXXXXXX
vm0200c868bf16$ doas -s
doas (puffy@vm0200c868bf16) password: XXXXXXXXXXXXXXXXXX
vm0200c868bf16#
```

# 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
```

---

<sup>a</sup> <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
```

---

clock <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

---

<sup>intr</sup> <https://marc.info/?l=openbsd-misc&m=154834783313341&w=2>

# 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.<sup>6.5</sup>

`/etc/doas.conf`

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

User runs:

```
$ doas pkill -9 -f vm13
```

---

<sup>6.5</sup> `vmctl stop <vm-name> -f` (<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 ([@cobracmder](#))

Stefan Kempf

Claudio Jeker

Jasper Lievisse Adriaanse ([@jasper\\_la](#))

Ori Bernstein ([@oribernstein](#))

Roman Zolotarev ([@romanzolotarev](#))

47

# 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>