Limits and the Practical Usability of BSDs, a Big Data Prospective

Predrag Punoševac
predragp@cs.cmu.edu

The Auton Lab
Carnegie Mellon University

June 11, 2016
Thanks

Thanks to organizers for this great meeting and for giving me the opportunity to speak.
Who am I?
Intro

- Who am I?
- What is the Auton Lab?
Intro

- Who am I?
- What is the Auton Lab?
- Why don’t we just use SCS computing facilities?
Intro

- Who am I?
- What is the Auton Lab?
- Why don’t we just use SCS computing facilities?
- How did I get myself into this mess?
Intro

- Who am I?
- What is the Auton Lab?
- Why don’t we just use SCS computing facilities?
- How did I get myself into this mess?
- Origins of BSDs in the Auton Lab?
Intro

- Who am I?
- What is the Auton Lab?
- Why don’t we just use SCS computing facilities?
- How did I get myself into this mess?
- Origins of BSDs in the Auton Lab?
The Auton Lab is a statistical data mining and machine learning group started by Andrew Moore and Jeff Schneider (Uber) currently directed by Artur Dubrawski. Please say something about number of users, CPUs, GPUs, storage space. Started because SCS computing facilities didn’t support 64-bit computing at that time. We operate autonomously from SCS and CMU computing facilities and treat their network as a hostile. Based upon some forensic evidence FreeBSD was used in the past. June 24 of 2013 I put the System Administrator’s hat. I found the computing infrastructure in disarray in part due to accretion as a main asset growth/acquisition method.
Chronology

1969 · · · · · · Ken Thompson starts work on UNIX.
1971 · · · · · · UNIX first edition.
1975 · · · · · · Ken Thompson sabbatical at the UC Berkeley.
1976 · · · · · · Bill Joy starts work on BSD.
1977 · · · · · · 1.0 BSD.
1992 · · · · · · 4.4 BSD.
1994 · · · · · · NetBSD 1.0.
1995 · · · · · · Theo De Raadt forks OpenBSD from NetBSD on the 18th of October.
1996 · · · · · · OpenBSD 1.2 released in July.
Genealogy Tree

- Intro
- Chronology
- Chronology II
- Genealogy Tree
- General
- Limitations
- Scientific
- Computing
- Continuation
- misc issues
- NetBSD
- OpenBSD
- pf.conf and pfctl
- OpenBSD cons
- FreeBSD
- TrueOS
- TurnKey Appliance
- FreeNAS
- pfSense
- DragonFly BSD
- HAMMER
- Dark Clouds
- References
Say something about other talks which are happening at the same time which might be more interesting and FreeBSD centric:

1. Open/LibreSSL in FreeBSD
2. Through the Wire Measurement and Improvement of a software based IPsec implementation
3. FreeBSD and GDB

Give the outline of my talk and why it is not really a lecture but more a round table discussion.
General Limitations

- Lack of hardware and software vendor support!
- No NVIDIA CUDA drivers means GPU is impossible.
- Lack of proprietary compilers (Portland Group Fortran, C and C++ compilers and tools, Intel (possibly available for FreeBSD))
- No MATLAB®. FreeMAT, GNU Octave, and SciLab are not real alternatives neither as numerical computing environment nor as programming language. Power of MATLAB is in toolboxes! For most students that is the first programming language.
- No Wolfram Mathematica® which is de facto standard general computer algebra program.
The ports collection is a volunteer project (both OpenBSD and FreeBSD) which means that state of important ports is hit-and-miss.

There is just too much open source software which is required in scientific computing to make in-house effort worthwhile.

ATLAS, BLAS, LAPACK, Boost, GCC 5.0 and higher due to threading requirement (scientists are poor code writers!), OpenCV, GDB, Valgrind ...

R is typically first serious domain specific language that students use after hitting limits with MATLAB (generally in great condition on both Open and Free), Python (numpy, scipy, matplotlib, pandas, scikit-learn, ipython, pip). Until recently OpenBSD didn’t have 3.5 flavour.
Continuation

- **No Rocks Cluster Distribution**
- **Hadoop and Apache Spark** are finally available on FreeBSD.

- **No Caffe** (deep learning framework) which is even pain on RHEL due to Ubuntu activism.

- **General problem** is that lot of software is developed on Ubuntu without any consideration for other Linux distros let alone other OSs.

- **Why do we need such software diversity?** We really don’t! However when research is going nowhere the easiest thing to do is to blame system administrator who is depriving you of that magical peace of software which will write Ph.D. dissertation for you.
DoD compliance issues and similar.

Lack of full/paravirtualization like Xen Dom0 (never really had a gut to try NetBSD version) forced us to opt for Linux KVM. We are heavy users of OS level virtualization (FreeBSD Jails). Functional application virtualization would be nice (running Microsoft Office or WebEx would be really nice). Is it too late for Bhyve?

Developers and students familiarity with anything besides Ubuntu. I have had hard time pushing Springdale Linux (RHEL) as a standard computing platform. Often OS X users are as illiterate as Windows users.

Diversity of hardware (even a tiny difference in RAID card) and OS has adverse effect on sysadmin productivity.
**NetBSD**

- [http://predrag.freeshell.org/](http://predrag.freeshell.org/)
- Walter Neto porting WABPL (Wasabi Systems) so that OpenBSD can suck less.
- Have interesting regression tools and mail for example.
- Cross-compiling does not add up to portability!
- Everything is dead except amd64, arm, and mips64 (network) anyway.
- OpenBSD vs NetBSD developers Polish interview reveals that not a single NetBSD developer runs NetBSD at work, some not even at home so why should I?
- Never really understood lack of Xen Dom0 advertising. Does it really work?
People might have noticed that my personal web page is hosted by SDF Public Access UNIX System which runs NetBSD (used to be Alpha not sure now).
Comparing to FreeBSD and RHEL in particular, OpenBSD feels simpler, more straight-forward, more stable, better documented, sane defaults, and finished.

Secure by default! Really! Privilege separation, mitigation, pledge, random stack gap, W^X, ASLR, stack smashing protection, arc4random, encryption, no PAM, doas ...

We use for PF, DNS (unbound), LDAP, OpenVPN gateway, remote Monitoring/Telemetry, centralized logging, web proxy, sftp server with chrooted accounts.

Easy backup with altroot.

Easy binary upgrade and update with openup.

Lot of in-house tools: PF, OpenSSH, OpenNTPD, OpenSMTPD, mandoc, LibreSSL.

Native ldapd, snmpd, syslogd (TCP and TLS), sensorsd, L2PT/IPsec, softraid, dhcpd, authpf, relayd.
pf.conf and pfctl

```bash
# $OpenBSD: pf.conf,v 1.54 2014/08/23 05:49:42 deraadt
#
# See pf.conf(5) and /etc/examples/pf.conf

set skip on lo

block return    # block stateless traffic
pass            # establish keep-state

# By default, do not permit remote connections to X11
block return    in on ! lo0 proto tcp to port 6000:6010

Atomic rule set load

# pfctl -nvf /etc/pf.conf
# pfctl -s all
# systat pf

Don’t flush it!

# pfctl -F all
```
OpenBSD cons

- According to Henning Brauer OpenBSD sucks due to the lack of modern file system, poorly SMP aware network stack (and PF), insufficient optimization.

- It looks like the only valid concern a year after this talk is lack of modern file system.

- Personally I would really like to see Walter Neto finishing porting of WAPBL to OpenBSD. HAMMER would be truly nice but I am not sure I want to see OpenBSD getting more cumbersome.

- The only thing I never got to work for my needs is relayd. We use Nginx.
"The reason why ZFS requires so much memory is because it includes its own separate cache system based on the ARC algorithm. It’s a fantastic algorithm, but it certainly violates IBM patents, which is why it was removed from PostgreSQL and omitted from Linux. Perhaps Oracle has a license or a sufficient patent portfolio to protect Solaris and Unbreakable Linux users, but the CDDL does not confer any patent protection for ZoL or FreeBSD users.”

- I actually prefer hardware over software RAID.
- ZFS has copy-on-write, check-sums, and consistency check which beats the pants out of XFS but what about HAMMER?
- Built in backup (ZFS snapshots and remote replication) which beats the pants out of XFS but what about hammer mirror, hammer stream, hammer snapshots, and most importantly hammer history?
- Well integrated with Jail infrastructure via iocage.
- FreeBSD - a lesson in poor defaults (sendmail vs dma)
- Start but not finished shelf: bsnmpd, bsdhwmon, ...
“Long time FreeBSD user just starting out with TrueOS. Is there a cheat-sheet/document that lists differences and potential programs that exist in TrueOS, but not in the main FreeBSD?”

1. Installer (ZFS on the root even as z-mirror). Vanilla FreeBSD can do this now.
2. `pc-sysinstall` script and customizable configuration files are superior for automatic or customized installation to vanilla FreeBSD installer.
4. Update manager. FreeBSD can do this.
5. Life Preserver (management tool for ZFS snapshots and replication). Due to the bugs I switched to `zfsnap` and `zxfer`.
6. Warden (Jail management). Quickly abandoned in favor of `iocage` which is now abandonware (1.7.4) and being rewritten in Go.
7. Sane[r] defaults.
Rad Hat wannabe

TrueOS feels and looks like a Red Hat wannabe without massive corporate structure behind it. LDAP is broken for me. What is `/usr/local/etc/pcbsd.conf` and `pc-ldap.conf`?
TurnKey Appliance

- Only FreeBSD has them and very few comparing to Linux.
- Making learning curve gentler.
- Hiding details to accomplish previous is never a good idea.
- Typically has too many features but never the one I really need.
- Don’t do it if you are in for a long haul!
FreeNAS

- Gateway to ZFS and first encounter with FreeBSD after 7 years.
- My first exposure to embedded installations (NanoBSD).
- It has defaults picked by people who know infinitely more than I about ZFS.
- I still have fun reading FreeNAS configuration files.
- Learned a lot about proper storage monitoring including SMART daemon, telemetry data collection, and backups (snapshots and replication).
- We still have two main file servers running 9.2.1.9.
- It has way too many features which I will never use.
9.10 dilemma

Is trying to upgrade from 9.2.1.9 to 9.10 worth it? Root on ZFS is ok but improved NFS performance would be really nice. How about more radical move trying to import pool to TrueOS 10.3 and ditch FreeNAS altogether. Diversity is the enemy of productivity. LDAP authentication and authorization is broken for me on TrueOS. I would definitely have to learn more about zil and l2arc.
pfSense

- pfSense is a semi-interesting turnkey network appliance based now of FreeBSD 10.3 after running 8.xxx many years under the hood.
- PF on FreeBSD is a looong story...
- ZFS for root not possible unlike FreeNAS but embedded installation based on NanoBSD is interesting (OpenBSD has **flashrd**)
- Using web server to configure a firewall. Seriously?
- some people consider OpenBSD just an OS, not a firewall... They even came up with the name for ”real firewall”. It is called pfSense.
- Captive portal is an interesting feature which I have never implemented with OpenBSD (not sure that **authpf** is really that practical).
- I run in the VirtualBox to learn about PF/network monitoring/diagnostic tools. I was also curios about backup scheme.
Serious about security should forget about intrusion detection, firewalls, MAC and maybe start with zero-sum game theory. It looks like my government is not there yet. The Tao of Network Security Monitoring: Beyond Intrusion Detection by Richard Bejtlich leaves mathematician to wonder.
Originally considered in the role of a file server OS but lack of LDAP client was a deal breaker. LDAP client became fully functional with 4.0 release. I was probably the first one to test it based on FreeBSD documentation and obsolete ports security/pam-ldap and net/nss-ldap. I never tested with net/nss-pam-ldapd and net/nss-pam-ldapd-sasl.

Monitoring troubles (No Monit client, SNMP only v1 with LibreNMS, Collectd doesn’t compile).

Lack of drivers. My message to users@dragonflybsd about missing HDD answered by Matt that driver is missing as not ported from FreeBSD. Are hardware RAID cards really tested?

Jail infrastructure not updated. It would be nice to integrate with HAMMER.

HAMMER still have bugs according to Tomohiro Kusumi.

Installer is very crude and automatic, unattended installation is for practical purposes not possible.

Small user base and miniscule developer base without critical mass.

Very nice but even at home tricky to use at best.
Time permitting give a full demo of how HAMMER history over NFS, HAMMER mirror-stream, and backup with `rsync --inplace` work in my home network environment.
What lies ahead?

I was repeatedly asked by bosses to evaluate "cloud solutions".

AWS. Not as bad as originally sounded since OpenBSD now runs as DomU. Thanks Mike Belopuhov for PVHVM drivers and Antoine Jacoutot for creating the image!

Docker

OpenStack God forbid!
References


The Auton Lab System Administrator’s Guide. Is it worth it to write?