

illumos

State of the Community

POSIX Update

Garrett D'Amore

Advanced Micro Devices, Inc.

illumos is...

- Open Source System V (Solaris) Derived
- Development home for DTrace and ZFS
- Kernel + core for OmniOS, SmartOS, NexentaStor
- Four years on its own, but really ~25 years old

Recent Developments

- Network Virtualization
- Virtualization (LX brand, Bhyve, etc.)
- Persistent L2ARC
- SMB2 / SMB3
- Hardware drivers
- POSIX & Compatibility

Pace

- Good stuff happening
- But,
 - without the frenetic churn
 - this is a good thing
 - but,
 - its also a concern

State of POSIX in Tree

- Mostly complete support for POSIX.1-2001, SUSv3
- Failings are in user land mostly (e.g. more, pax, etc.)
- Minimal support for POSIX 2008 updates. Largest of which is `locale_t` (`newlocale(3c)`& friends)
- Some 2008 interfaces present, but not exposed, and many removed interfaces exposed

Driving Real \$\$

- Standbys: Nexenta, Joyent, Delphix, OmniTI
- New comers: RackTop, Pluribus Networks
- “Stealthy” players: e.g. Tegile
- Total “market” value: > \$100M (perhaps >>)

CHALLENGES & Opportunities

No “Formal”
Organization

things we lack

- Bounties
- Governance (?!?)
- Ownership
- Marketing
- Legal presence

Opportunities for non-Gurus

- Write & improve documentation
- Localization
- Test development
- Marketing

~150 Contributors

~0 ♀ Contributors

Where are the
women?

Hire a Female Kernel
Engineer. Today.

Groupthink.

— *people are sheeple*

Linux

– “easy” choice

FUD.

(Not Elmer.)

Drift & **Divergence**

Lots of private forks

less is more

less(1) is more(1)

less **BLOAT** is more
(better)

less (compatible) is more (compatible)

less *talking* is more coding

illumos-core

illumos-core vision

- focus on code
- standards compliant - *by default*
- less is more
- self-hosting
- self-testing
- cross-platform
- illumos coding standards (RTI ready by default)

Standards

- SunOS 4.x (and 3.x) is truly **ancient (32-bit SPARC only!)**
- XPG3 - this is really, really old
- POSIX is really, really old (1988)
- XPG4 - is also really old
- SUS - aka XPG4v2 - also old
- SUSv2 - UNIX98
- SUSv3 - POSIX.1 2001
- SUSv4 - POSIX.1 2008 - work in progress

Ancient & Useless Stuff

- `htable`, `gettable` - last possible use was in 1990 (DNS anyone?!)
- Wireless USB. DOA.
- 16-bit PCMCIA. Seriously?!?
- SunOS 4.x `a.out` support
- `kssl` - no TLS support
- NCA/NL7C - Global Zone only, actually slows all other traffic down
- LMS (AMT/HECI) - Can you say security risk/pointless?

readdir_r(3c) case study

- Standardized by POSIX in 1995 (IEEE Std. 1003.1c-1995)
- Earlier Draft version (Draft 6) supported in Solaris 2.6 - 11
- By default
- Solaris is the only common operating system to still do this

readdir_r(3c) case study

- `struct dirent *readdir_r(DIR *, struct dirent ***);`
- unless...
 - `_POSIX_C_SOURCE - 0 >= 199506L` *or...*
 - `_POSIX_PTHREAD_SEMANTICS` *or...*
 - `_LP64` *or...*
 - `_FILE_OFFSET_BITS != 32`
- in which case...
 - `int readdir_r(DIR *, struct dirent *, struct dirent **);`
 - note that return value semantics are **different!**

readdir_r(3c) resolution

- `struct dirent *readdir_r(DIR *, struct dirent ***);`
- unless...
 - `_POSIX_C_SOURCE - 0 >= 199506L` *or...*
 - `_POSIX_PTHREAD_SEMANTICS` *or...*
 - `_LP64` *or...*
 - `_FILE_OFFSET_BITS == 32` *or...*
 - `!_SUNOS_SOURCE`
- in which case...
 - `int readdir_r(DIR *, struct dirent *, struct dirent **);`
 - (note that you now get POSIX compliant by DEFAULT!)

accomplishments so far

- less(1) is more(1)
- uname & puname
- fexecve(2), O_TTY_INIT, O_DIRECTORY
- removed a lot of bloat (~200KLOC)
- suppression and updates for POSIX 2008 symbols we already had (-D_XOPEN_SOURCE=700)
- s/awk/nawk/
- no more dbus or hal dependencies
- rich self tests for POSIX (and non-POSIX) interfaces (>3000 tests)
- 85 separate bugs/features integrated (started Aug 14, approx 2 per day)
- vi(1) is real vi, not vim(1)

uname(2)

- `uname -s` == “illumos”
- `uname -r` == 0.9.x (x == months since Aug 2010)
- `puname(1)` utility
 - `puname -S uname -s -r` == “SunOS 5.11”
 - `puname -S uname -v` == “alternate-uname”
 - `puname -U` to restore
 - implemented as new flag in process *uarea*

uname breakage

- 64-bit Studio (checks for SunOS > 5.7)
- hald - fixed (in separate git repo), breaks X
- config.guess (auto*) - can override in configure command line
- pkgdepend - (fixed?) via terminatorlib/cwd.py
- DTrace printf test (test checks uname == SunOS)

Remaining Compliance Work for C APIs

- System Interfaces (POSIX.1-2008)
 - memory streams
 - dprintf (universe explodes...)
 - pthreads changes
- C11 support
 - interfaces to improve security

Shell & Utilities Work

- pax (not star, thank you very much)
- modernize (slightly) awk
- lots of minor updates (mv, cp, ls, etc.)
- replace ksh93 with dash + libedit
- Test suite development, Standard(s) trolling

Can I Help?

YES! YOU CAN!

- Code & Documentation reviews
- Testing
- Help upstreaming
- Code contributions (pull requests!)
- <http://bitbucket.org/gdamore/illumos-gate>
 - “core” branch

Thank you!!