Joyent

Fork Yeah!
The Rise & Development of illumos

Bryan Cantrill
VP, Engineering

bryan@joyent.com
@bcantrill

WTF is illumos?

- An open source descendant of OpenSolaris

- ...which itself was a branch of Solaris Nevada

- ...which was the name of the release after Solaris 10

- ...and was open but is now closed

- ...and is itself a descendant of Solaris 2.x

- ...but it can all be called "SunOS 5.x"

- ...but not "SunOS 4.x" — that's different

- Let's start at (or rather, near) the beginning...

SunOS: A people's history

- In the early 1990s, after a painful transition to Solaris, much of the SunOS 4.x engineering talent had left

- Problems compounded by the adoption of an immature SCM, the Network Software Environment (NSE)

- The engineers revolted: Larry McVoy developed a much simpler variant of NSE called NSElite (ancestor to git)

- Using NSElite (and later, TeamWare), Roger Faulkner, Tim Marsland, Joe Kowalski and Jeff Bonwick led a sufficiently parallelized development effort to produce Solaris 2.3, "the first version that worked"

- ...but with Solaris 2.4, management took over day-to-day operations of the release, and quality slipped again

Solaris 2.5: Do or die

- Solaris 2.5 absolutely had to get it right — Sun had new hardware, the UltraSPARC-I, that depended on it

- To assure quality, the engineers "took over," with Bonwick installed as the gatekeeper

- Bonwick granted authority to "rip it out if it's broken" — an early BDFL model, and a template for later generations of engineering leadership

- Solaris 2.5 shipped on schedule and at quality — and engineers would never again relinquish control of the risk management of the operating system

The rise of the Bonwick Youth

- In the mid-1990s, it became a foregone conclusion that UNIX would die at the hands of Windows NT (!)

- Hard to believe, but Sun was the *only* computer vendor whose Windows NT strategy was to beat Windows NT

- Sun's dedication to this vision — the operating system as a nexus of innovation — attracted an entire new generation of engineers to the company

- New generation became known as the "Bonwick Youth"

- As talent attracted talent, there emerged an opportunity to become to OS innovation in the 2000s what Xerox PARC was to computing innovation in the 1970s

Revolutionary ideas

- By mid-2001, the trajectories set out by the move from SunOS 4.x to Solaris a decade prior had reached their logical conclusion...

- Development started on more radical ideas, each of which would become revolutionary in its own right:

  - ZFS (née Pacific)
  - DTrace
  - Zones (née Project Kevlar)
  - Fault Management Architecture (FMA)
  - Service Management Facility (SMF, née Greenline)
  - FireEngine/Crossbow
  - Least Privilege

- These were technologies invented and initiated by *engineers*, not managers or marketers

- These projects reflected the *people* behind them, and their *personal* histories with the operating system

- Each of these projects represented an *opinion* — borne of both experience and frustration — about what the operating system should be

- **Organizations don't innovate — people do**

- As the rise of Linux (and particularly x86) forced the market price of OS acquisition to zero, it became clear that *open sourcing the operating system* was the right business decision

- The OS had always been a *loss-leader* for Sun; the company made money from complementary goods (hardware, support, etc.)

- Moreover, open sourcing the system would assure that our innovations would transcend products and become *advances* in the state of the art

- This was entirely in character for the company that had pioneered "open systems" some fifteen years prior!

Open source!

- Starting in the late 1990s, there was serious talk of open sourcing the operating system...

- ...but in a system with deeply proprietary roots and a troubled legal past, it takes a lot more than just talk

- In 2004, Jonathan Schwartz made it a priority to open source the operating system

- In January of 2005, OpenSolaris was born when DTrace became the first of the system to be open sourced

- The rest of the OS was open sourced in June 2005

- The OS was developed in the open from that point — changes to the OS were open source as they integrated

Leaping the proprietary chasm

- Very, very few established software projects have leapt the chasm from proprietary to open

- Unfortunately, not all elements of the OS could be open sourced; some contracts prevented some small (but important) bits from being open sourced

- To allow for such proprietary drivers, Sun developed a file-based copy-left license, the Common Development and Distribution License (CDDL)

- Contrary to public claims by some ex-Sun employees, this was *not* done to be deliberately GPL incompatible!

OpenSolaris challenges

- That certain critical bits had to remain proprietary made forking the operating system technically difficult...

- And that virtually all Solaris implementation knowledge lived within Sun's walls made it a practical impossibility

- The community had the *right* to fork, but not the *power*

- This led down the primrose path to open source governance: governing boards, elections, constitutions

- And because all development on the system realistically required copyright assignment to Sun, OpenSolaris (sadly) remained a Sun puppet

- Worse, some among Sun's middle management fancied themselves puppeteers...

The "OpenSolaris missile crisis"

- In fall of 2007, Sun decided to create a new OpenSolaris-based distribution called… OpenSolaris

- Rightfully, the OpenSolaris Governing Board (OGB) was upset about this — it made the community that much more of a charade to have its very name hijacked

- Humanity-wasting governance orgy ensued...

- A compromise was reached in the end that allowed for de-escalation, but it became clear that — despite stated intentions — Sun did not want an independent OGB

- The community was deflated; OGB composition became more and more Sun dominated, and OpenSolaris was in the doldrums for the next three years

- Ailing Sun was bought by Oracle in 2009, with the acquisition closing in February 2010

- Sun was concisely eulogized by Scott McNealy:

  Kicked butt, had fun, didn't cheat, loved our customers, changed computing forever.

- After the acquisition closed, it became clear that Scott's eulogy was (ironically?) the antithesis of Oracle

- Over 2010, it became clear that Oracle had absolutely *no* interest in OpenSolaris

- Despite that it would be obviously stupid, there was even a move within Oracle to actually *close* the system!

Meanwhile, the birth of illumos

- Starting in the summer of 2010, Garrett D'Amore at Nexenta — with help from Rich Lowe, Jason King and others — began the process of either writing the closed bits from scratch or porting them from BSD

- By early August, an entirely open system was booting

- Dubbed "illumos" (from *illuminare*, Latin for illuminate) and made available on August 3, 2010

- Notably, the announcement included both code and a working demo — leading with technology, not rhetoric

- illumos was *not* designed to be a fork, but rather an entirely open downstream repository of OpenSolaris

The undignified death of OpenSolaris

- On Friday, August 13th, 2010, an internal memo was circulated by the putative Solaris leadership:

  > *We will distribute updates to approved CDDL or other open source-licensed code following full releases of our enterprise Solaris operating system. In this manner, new technology innovations will show up in our releases before anywhere else.* ***We will no longer distribute source code for the entirety of the Solaris operating system in real-time*** *while it is developed, on a nightly basis.*

- This was *never* publicly announced; updates just silently stopped on August 18th, 2010

- And it was a lie anyway: Solaris 11 was released on November 9, 2011 — and there was no source release

The Solaris diaspora

- The Solaris diaspora (which was already underway) was greatly accelerated by Oracle's depraved act

- Within 90 days, the entire DTrace team had left Oracle, all primary inventors of ZFS had left Oracle and primary engineers for both zones and networking had left Oracle

- Fortunately, Oracle's loss was illumos's gain: nearly all of these engineers went to companies betting on illumos

- It is no surprise that these engineers have continued to innovate — and that a year later, their innovations (big and small!) have landed in illumos

The illumos innovation irony

- There is no copyright assignment in illumos; code is under the CDDL, but copyright remains with the holder

- But with the closing of Solaris, Oracle has effectively abandoned the CDDL — which it can only do because it was assigned copyright on all community contributions

- However, Oracle does *not* own copyright on the illumos innovations; it cannot take this code back until it abides by its own license!

- In illumos, we have seen critical innovations and bug fixes in ZFS, DTrace, Zones and other core technologies

- *These innovations will never be in Oracle Solaris*

- illumos is the repo of record for critical OS technologies!

- Most core ZFS engineers now in illumos community: e.g. Matt Ahrens, Eric Schrock, George Wilson, Adam Leventhal, Bill Pijewski and Brendan Gregg

- ZFS engineers from several different ZFS-based systems (illumos-based, FreeBSD-based, and others) formed a joint working group

- First concrete technology artifact was a proposal for SPA versioning that allows disjoint features from different vendors without requiring conflicting versions

- That this was a most pressing problem speaks for itself: many different companies are betting big on illumos ZFS!

- Matt Ahrens added the "refratio" property to understand the compression win for referenced data

- Matt Ahrens added the ability to get estimated progress for zfs send and receive (!)

- Bill Pijewski and Jerry Jelinek added Zones-based I/O throttling, allowing for optimal disk operation in light of multi-tenant demands

- Dan McDonald and Sumit Gupta implemented UNMAP for STMF, allowing for better ZFS-backed iSCSI LUNs

- Many more features in the pipe: feature flags, background destroy, resumable send

DTrace

- Virtually all DTrace engineers are now in the illumos community: Adam Leventhal, Brendan Gregg, Eric Schrock, Dave Pacheco

- Added a log-linear quantization action (llquantize())

- Added KVM support with vmregs[] variable

- A bunch of little stuff: allowed tracemem() to take a variable-size; added toupper()/tolower(); allowed lltostr() to take an optional base

- Added USDT provider reaping (a long-time issue)

- Eric Schrock added the CTF-aware print() action!

Zones

- Primary zones engineer (and JumpStart tech lead!) Jerry Jelinek is very active in the illumos community

- Particular focus on high-tenancy systems; it is highly likely that the machines that run the most number of zones in production are doing so today on illumos

- Added many per-zones kstats (e.g. for CPU usage, VM activity, I/O activity, etc.)

- Reimplemented rcapd to be much more reliable, especially under heavy load

- Fixed many bugs related to many-zone configurations

- Added -z and -Z options to svcs (and -L!)

KVM

- In addition to work on the traditional core technologies of ZFS, DTrace and Zones, we have added a new one: Joyent team ported KVM from Linux to illumos

- KVM makes use of hardware (specifically, Intel) virtualization support to allow one to run arbitrary x86 guests at nearly bare-metal speed

- Work was presented by Joyent at KVM Forum this year; currently running in production in the Joyent cloud

- illumos-based KVM has three distinct advantages over Linux-based KVM: ZFS, DTrace and Zones

- illumos distributions SmartOS and OpenIndiana have KVM support by default

illumos Distributions

- There are already several illumos-based distributions, with more on the way

- OpenIndiana includes IPS packaging and has both server and desktop variants

- SmartOS is a Joyent distribution features NetBSD packaging and is designed for cloud computing

- illumian is a new Nexenta distribution that features Debian packaging

- These distributions *complement* each other — they are growing the illumos pie more than dividing it!

illumos community values

- We value the freedom to fork over governance; code over discussion; innovation over democracy

- David Clark of IETF fame put it best (h/t @kebesays):

  We reject: kings, presidents and voting.
  We believe in: rough consensus and running code

- Instead of a BDFL, we have a benevolent oligarchy — and so far we have achieved consensus with little effort

- We value *utility*; we don't do things for their own sake!

- These values were embodied in the recent illumos hackathon (h/t @ahl), which was a resounding success!

- Come to the #lisa11 illumos BOF tonight!

- Start playing around with a distro that suits you

- Start playing around with the code:

  - http://src.illumos.org or

  - https://github.com/illumos/illumos-gate

- If you're looking for something to do or have questions, don't hesitate to ask:

  - Mailing list: developer@lists.illumos.org

  - IRC: #illumos on irc.freenode.net

- If you don't know where to go, you can always tweet or message Deirdré Straughan (@deirdres), Joyent's SmartOS community manager

Thank you!

- @gedamore for conceiving of illumos, birthing it, and leading the community

- @richlowe for his tremendous work from the beginning on all aspects of illumos

- @openindiana and @aszeszo for OpenIndiana

- @joshwilsdon, @johnnysunshine and @rmustacc for their work on the SmartOS distribution

- @ahl for organizing the illumos hackathon — and all of the hammer-swinging butt-kickers who attended!