You think you're not a target?
A tale of 3 developers…

Chris Lamb
Debian Project Leader
@lolamby

New York Linux User Group
Two Sigma, NYC
20th March 2018

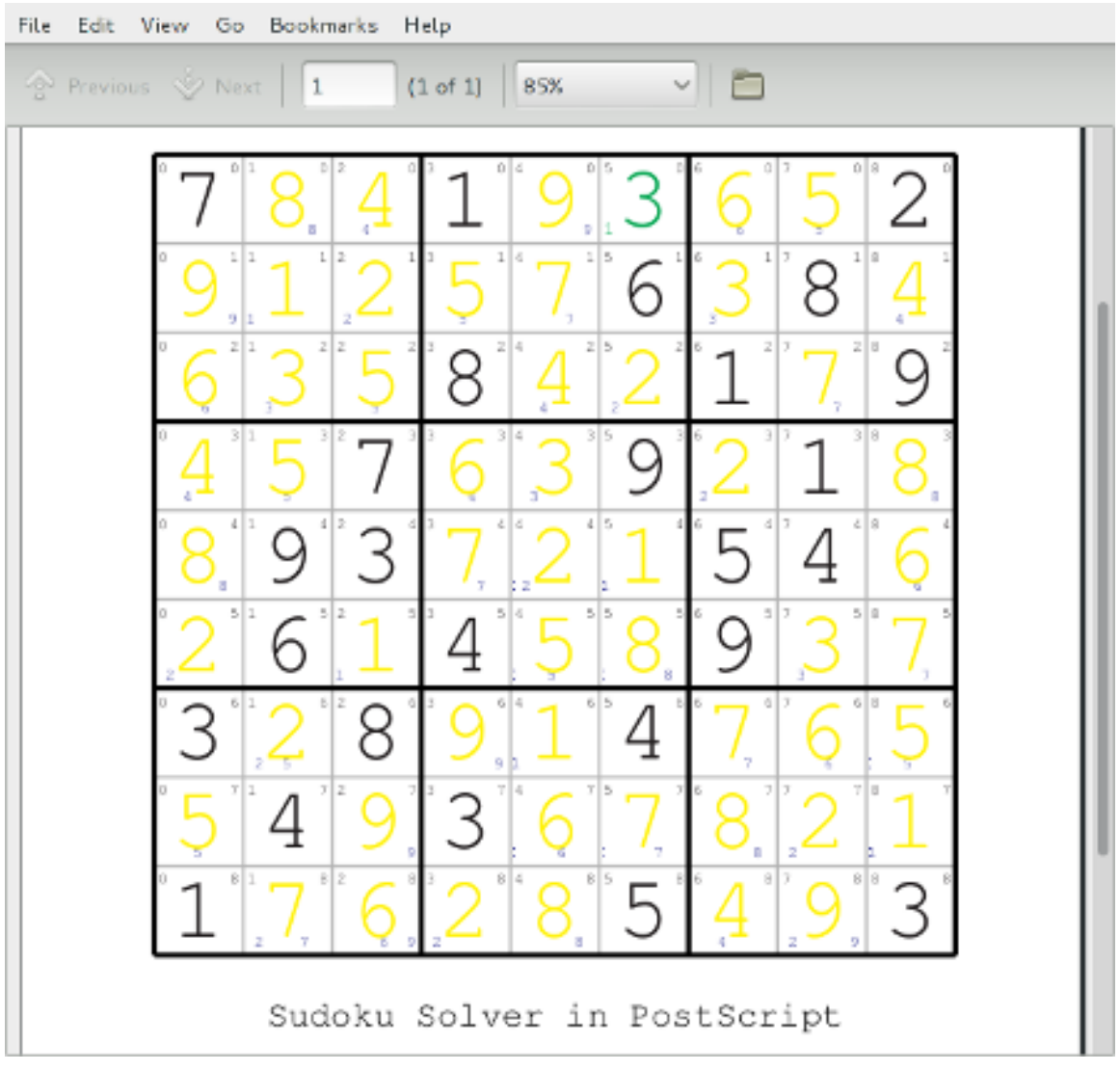
My name is…

Debian Project Leader

OpenSource.org Board Member

Free software developer for 10+ years

Freelance software developer


File Edit View Go Bookmarks Help
Previous Next
1
(1 of 1)
85%
Sudoku Solver in PostScript

```
< zed0> can you get cp to give a progress bar like wget?
```

Damn right you can.

```
#!/bin/sh
cp_p()
{
   strace -q -ewrite cp -- "${1}" "${2}" 2>&1 \
      | awk '{
        count += $NF
            if (count % 10 == 0) {
               percent = count / total_size * 100
               printf "%3d%% [", percent
               for (i=0;i<=percent;i++)
                  printf "="
               printf ">"
               for (i=percent;i<100;i++)
                  printf " "
               printf "]\r"
            }
         }
         END { print "" }' total_size=$(stat -c '%s' "${1}") count=0
}
```

In action:

```
% cp_p /mnt/raid/pub/iso/debian/debian-2.2r4potato-i386-netinst.iso /dev/null
76% [=============================================>                    ]
```

Three developers…


Alice


My Awesome Software
Download Source
or
Download .exe / .deb / .rpm


My Awesome Software
Download Source
or
Download .exe / .deb / .rpm


Bob
AR-15
RAIDMAX


← Eve


Eve

The four essential freedoms

A program is free software if the program's users have the four essential freedoms:

- The freedom to run the program as you wish, for any purpose (freedom 0).
- The freedom to study how the program works, and change it so it does your computing as you wish (freedom 1). Access to the source code is a precondition for this.
- The freedom to redistribute copies so you can help your neighbor (freedom 2).
- The freedom to distribute copies of your modified versions to others (freedom 3). By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this.

General problem

Can view source code for malicious flaws

But Users install pre-compiled packages

Can we trust the compilation process?

```
Reading package lists... Done
Building dependency tree... Done
The following extra packages will be installed:
  apache2-bin apache2-data apache2-utils file libalgorithm-c3-perl libapr1 libaprutil1 libaprutil1-dbd-sqlite3 libaprutil1-ldap
  libarchive-extract-perl libb-hooks-endofscope-perl libb-hooks-op-check-perl libbareword-filehandles-perl libcgi-fast-perl libcgi-pm-perl
  libclass-c3-perl libclass-c3-xs-perl libclass-method-modifiers-perl libclass-xsaccessor-perl libcpan-changes-perl libcpan-meta-perl
  libdata-optlist-perl libdata-perl-perl libdata-section-perl libdevel-caller-perl libdevel-global-destruction-perl libdevel-lexalias-perl
  libencode-locale-perl libexpat1 libexporter-tiny-perl libfcgi-perl libffi6 libfile-slurp-perl libgdbm3 libgetopt-long-descriptive-perl
  libgmp10 libgnutls-deb0-28 libhogweed4 libhtml-parser-perl libhtml-tagset-perl libhttp-date-perl libhttp-message-perl libicu55
  libimport-into-perl libindirect-perl libio-html-perl libio-stringy-perl libldap-2.4-2 liblexical-sealrequirehints-perl
  liblist-moreutils-perl liblog-message-perl liblog-message-simple-perl liblua5.1-0 liblwp-mediatypes-perl libmagic1 libmodule-build-perl
  libmodule-implementation-perl libmodule-pluggable-perl libmodule-runtime-perl libmodule-signature-perl libmoo-perl libmoox-handlesvia-perl
  libmro-compat-perl libmultidimensional-perl libnamespace-autoclean-perl libnamespace-clean-perl libnettle6 libnghttp2-14 libp11-kit0
  libpackage-constants-perl libpackage-stash-perl libpackage-stash-xs-perl libpadwalker-perl libparams-classify-perl libparams-util-perl
  libparams-validate-perl libpath-tiny-perl libpod-latex-perl libpod-markdown-perl libpod-readme-perl librole-tiny-perl libsasl2-2
  libsasl2-modules libsasl2-modules-db libsoftware-license-perl libsqlite3-0 libssl1.0.2 libstrictures-perl libsub-exporter-perl
  libsub-exporter-progressive-perl libsub-identify-perl libsub-install-perl libsub-name-perl libtasn1-6 libterm-ui-perl libtext-soundex-perl
  libtext-template-perl libtimedate-perl libtry-tiny-perl libtype-tiny-perl libtype-tiny-xs-perl libunicode-utf8-perl liburi-perl
  libvariable-magic-perl libxml2 mime-support openssl perl perl-modules rename sgml-base ssl-cert xml-core
Suggested packages:
  www-browser apache2-doc apache2-suexec-pristine | apache2-suexec-custom libarchive-cpio-perl
  libsasl2-modules-otp libsasl2-modules-ldap libsasl2-modules-gssapi-mit libsasl2-modules-gssapi-heimdal
  libdevel-stacktrace-perl libwww-perl ca-certificates perl-doc libterm-readline-gnu-perl libterm-readline-perl-perl make libb-lint-perl
  libcpanplus-dist-build-perl libcpanplus-perl libfile-checktree-perl libobject-accessor-perl sgml-base-doc openssl-blacklist debhelper
The following NEW packages will be installed:
  apache2 apache2-bin apache2-data apache2-utils file libalgorithm-c3-perl libapr1 libaprutil1 libaprutil1-dbd-sqlite3 libaprutil1-ldap
  libarchive-extract-perl libb-hooks-endofscope-perl libb-hooks-op-check-perl libbareword-filehandles-perl libcgi-fast-perl libcgi-pm-perl
  libclass-c3-perl libclass-c3-xs-perl libclass-method-modifiers-perl libclass-xsaccessor-perl libcpan-changes-perl libcpan-meta-perl
  libdata-optlist-perl libdata-perl-perl libdata-section-perl libdevel-caller-perl libdevel-global-destruction-perl libdevel-lexalias-perl
  libencode-locale-perl libexpat1 libexporter-tiny-perl libfcgi-perl libffi6 libfile-slurp-perl libgdbm3 libgetopt-long-descriptive-perl
  libgmp10 libgnutls-deb0-28 libhogweed4 libhtml-parser-perl libhtml-tagset-perl libhttp-date-perl libhttp-message-perl libicu55
  libimport-into-perl libindirect-perl libio-html-perl libio-stringy-perl libldap-2.4-2 liblexical-sealrequirehints-perl
  liblist-moreutils-perl liblog-message-perl liblog-message-simple-perl liblua5.1-0 liblwp-mediatypes-perl libmagic1 libmodule-build-perl
  libmodule-implementation-perl libmodule-pluggable-perl libmodule-runtime-perl libmodule-signature-perl libmoo-perl libmoox-handlesvia-perl
  libmro-compat-perl libmultidimensional-perl libnamespace-autoclean-perl libnamespace-clean-perl libnettle6 libnghttp2-14 libp11-kit0
  libpackage-constants-perl libpackage-stash-perl libpackage-stash-xs-perl libpadwalker-perl libparams-classify-perl libparams-util-perl
  libparams-validate-perl libpath-tiny-perl libpod-latex-perl libpod-markdown-perl libpod-readme-perl librole-tiny-perl libsasl2-2
  libsasl2-modules libsasl2-modules-db libsoftware-license-perl libsqlite3-0 libssl1.0.2 libstrictures-perl libsub-exporter-perl
  libsub-exporter-progressive-perl libsub-identify-perl libsub-install-perl libsub-name-perl libtasn1-6 libterm-ui-perl libtext-soundex-perl
  libtext-template-perl libtimedate-perl libtry-tiny-perl libtype-tiny-perl libtype-tiny-xs-perl libunicode-utf8-perl liburi-perl
  libvariable-magic-perl libxml2 mime-support openssl perl perl-modules rename sgml-base ssl-cert xml-core
0 upgraded, 114 newly installed, 0 to remove and 1 not upgraded.
Need to get 23.8 MB of archives.
After this operation, 97.9 MB of additional disk space will be used.
Do you want to continue? [Y/n]
```

Health

Solution?

NHS cyber-attack: Amber Rudd says lessons must be learnt

35 minutes ago | Health

Share

Top S

NHS 'm
The NHS
systems
attack, th
35 min

Watson
defeat'
8 hour

Tory pl
past
30 min

How does this help?

Alice → Blackmail will be discovered

Bob → Compromise will be detected

Carol → Tampered laptop will be discovered

Removes incentive to attack in the first place

“Reproducible” builds allows verification that no flaws have been introduced during the compilation process

"Builds with the same dependencies"... ✖

"Reliable" builds... ✖

Identical build results

```
# sha1sum ./my-binary
```

Alice	7a482b984883990bd4ab2ac5985630886cc252c
Bob	7a482b984883990bd4ab2ac5985630886cc252c
Carol	d0f65b7de7a49e818b8095538d3a0f783cc9c27

Wait, software isn't reproducible already?

Timestamps

Timezones & locales

Dictionary/hash/database ordering

Build paths

Non-deterministic file ordering

Build parallelism

Users, groups, `umask`, environment variables

Other advantages

Easier to test changes — minimal diffs

Cache ratio — save time, money & CO_2

Detect corrupted build environments

Finds bugs!

Predictable OpenID secret

```
# Build.PL
$build->config_data(OpenIDConsumerSecret=>int(1e15*rand()));


# /usr/share/perl5/GBrowse/ConfigData.pm
{
 'OpenIDConsumerSecret' => '639098210478536',
 'cgibin' => '/usr/lib/cgi-bin/gbrowse',
 'conf' => '/etc/gbrowse',
 [..]
},
```

Every installation shares the same secret!

Random characters in manpages

```
-This manual page documents the usageoof WikipediaFS.
+This manual page documents the usage of WikipediaFS.
```

```
memcpy(&buf[1], &buf[2], strlen(buf)-1);
```

```
memcpy(3): The memory areas must not overlap
```

" n\\011" → "\111" → maps to capital "I"

```
- memcpy(&buf[1], &buf[2], strlen(buf)-1);
+ memmove(&buf[1], &buf[2], strlen(buf)-1);
```

Fails to build 0.46% of the time

```
x = f(u('abc'), 16)
y = f(u('abc'), 16)
self.assertEqual(sorted(set(x)), [u('a'), u('b'), u('c')])


AssertionError: Lists differ: [u'a', u'b'] != [u'a', u'b', u'c']
```

$$({}_3C_2)*(2/3)^{16} - ({}_3C_1)*(1/3)^{16} =\sim 0.46\%$$

Debian & reproducible builds

"Torture test"

Time & date

Hostname & domain name

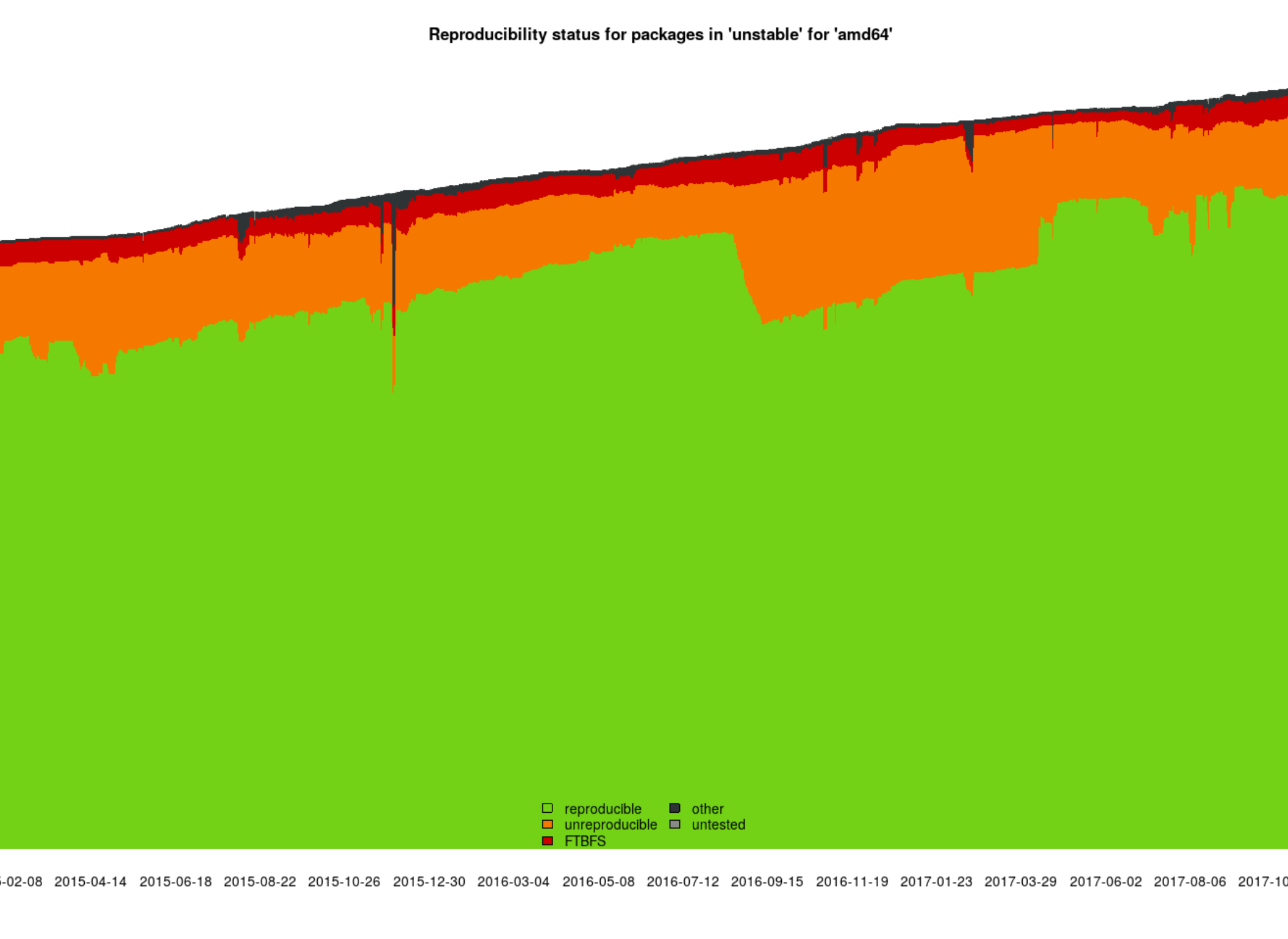
Filesystem (`disorderfs`)

Timezone & locale

`uid` & `gid`

Kernel & CPU type

First rebuild in 2013	24% packages reproducible
March 2018	93% packages reproducible


Reproducibility status for packages in 'unstable' for 'amd64'
reproducible
unreproducible
FTBFS
other
untested
-02-08
2015-04-14
2015-06-18
2015-08-22
2015-10-26
2015-12-30
2016-03-04
2016-05-08
2016-07-12
2016-09-15
2016-11-19
2017-01-23
2017-03-29
2017-06-02
2017-08-06
2017-10

isdebianreproducibleyet.com

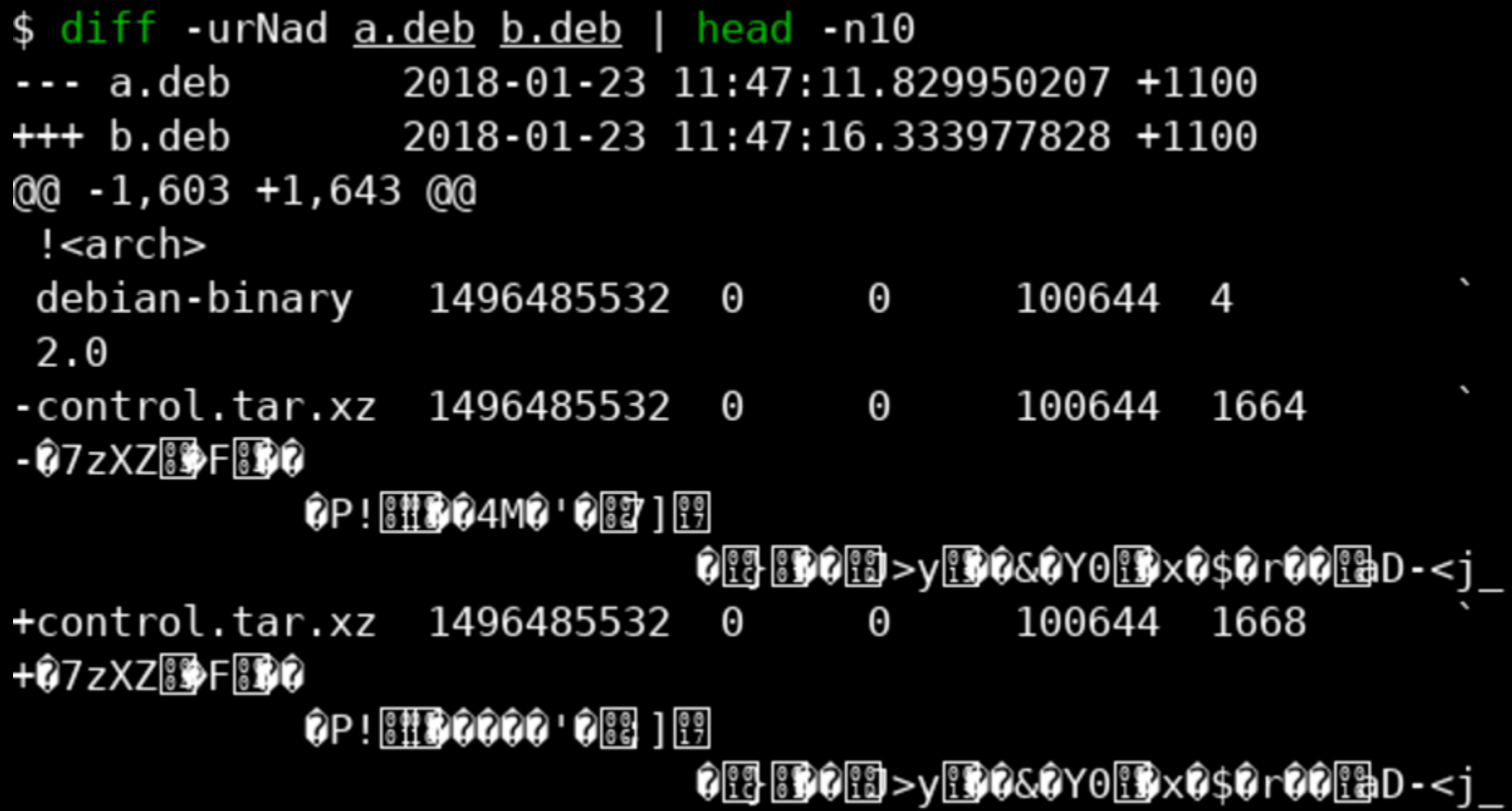
Beyond Debian…

coreboot, Fedora, LEDE, OpenWRT, NetBSD, FreeBSD, Archlinux, Qubes, F-Droid, NixOS, Guix, etc.

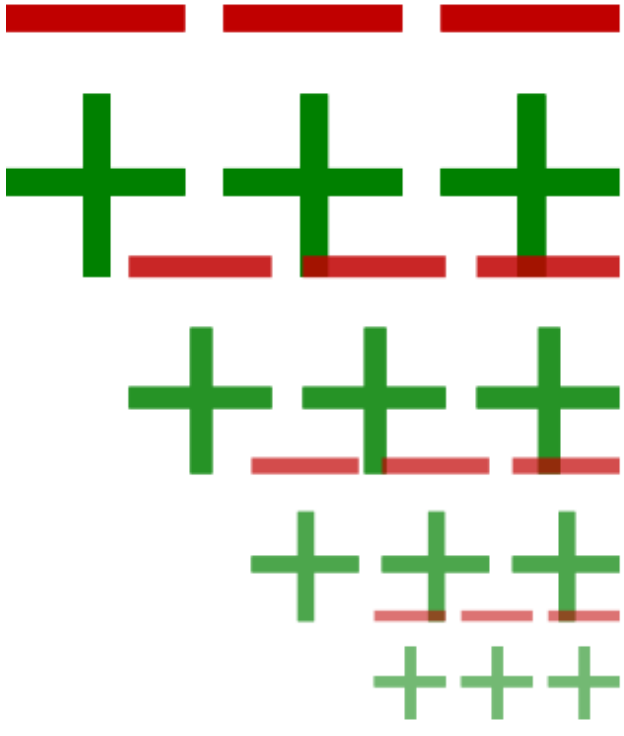
Other projects now using Debian's testing framework

Reproducible Builds summits (Athens, Berlin)

```
# diff -urNad file1 file2
--- file1  2017-06-18 12:37:03.179186661 +0800
+++ file2  2017-06-18 12:37:04.811193648 +0800
@@ -1 +1 @@
-This is the first file.
+This is the second file.
```


$ diff -urNad a.deb b.deb | head -n10
--- a.deb 2018-01-23 11:47:11.829950207 +1100
+++ b.deb 2018-01-23 11:47:16.333977828 +1100
@@ -1,603 +1,643 @@
!<arch>
debian-binary 1496485532 0 0 100644 4 `
2.0
-control.tar.xz 1496485532 0 0 100644 1664 `
-7zXZ F
P! 4M']
>y &Y0 x $ r aD-<j_
+control.tar.xz 1496485532 0 0 100644 1668 `
+7zXZ F
P! ']
>y &Y0 x $ r aD-<j_

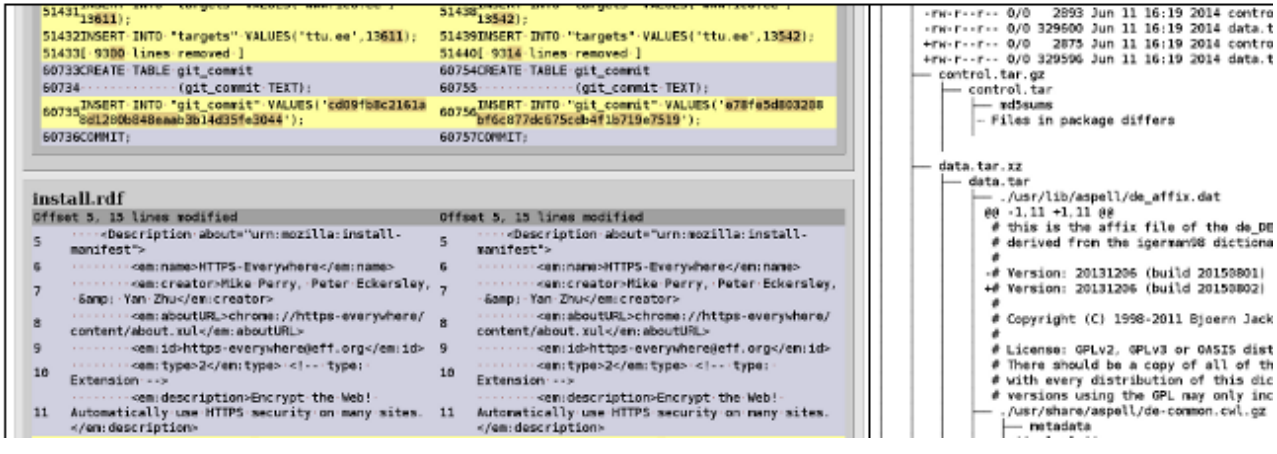

I SHOULD BUILD A BETTER DIFF



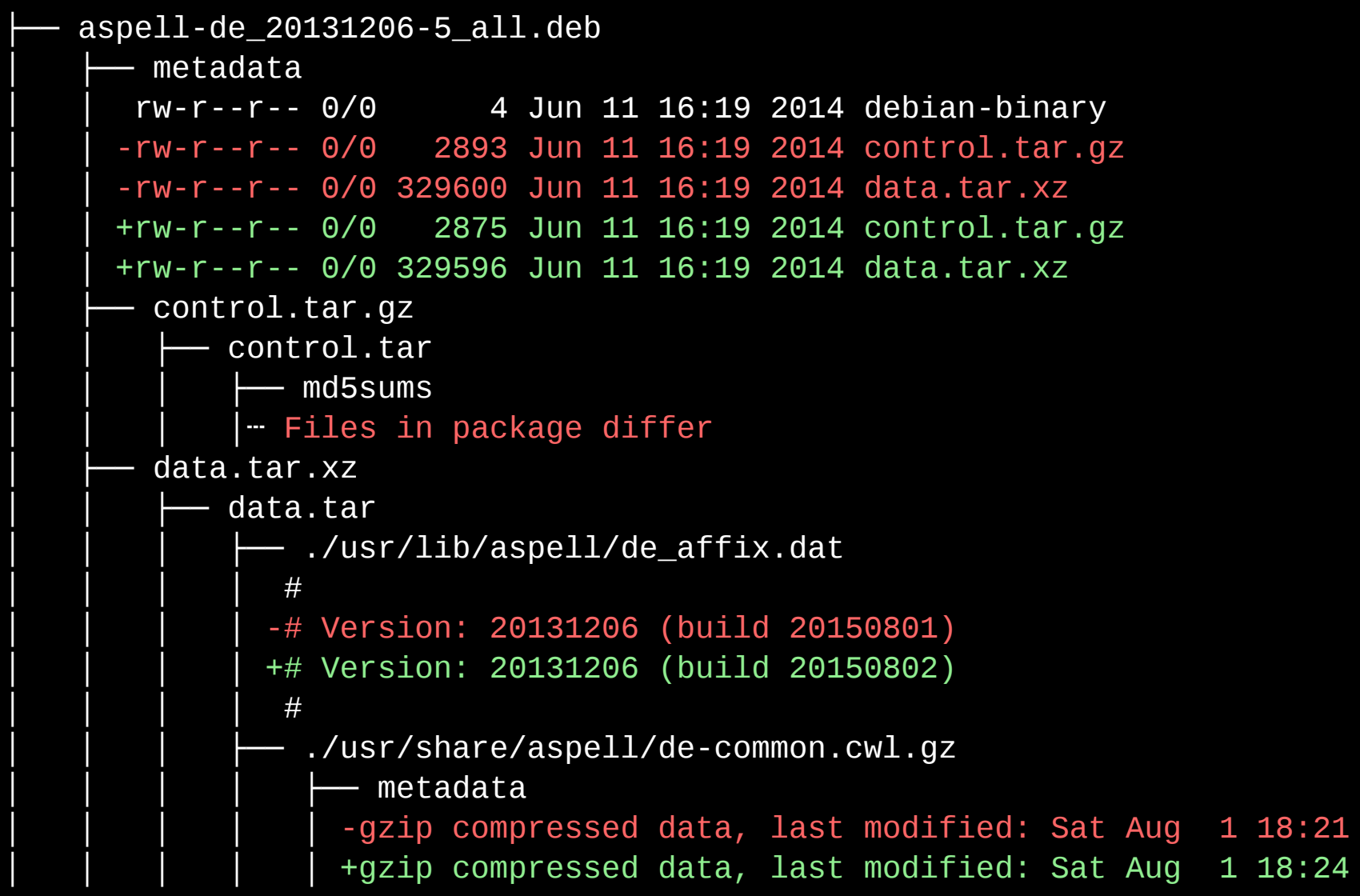
diffoscope

in-depth comparison of files, archives, and directories

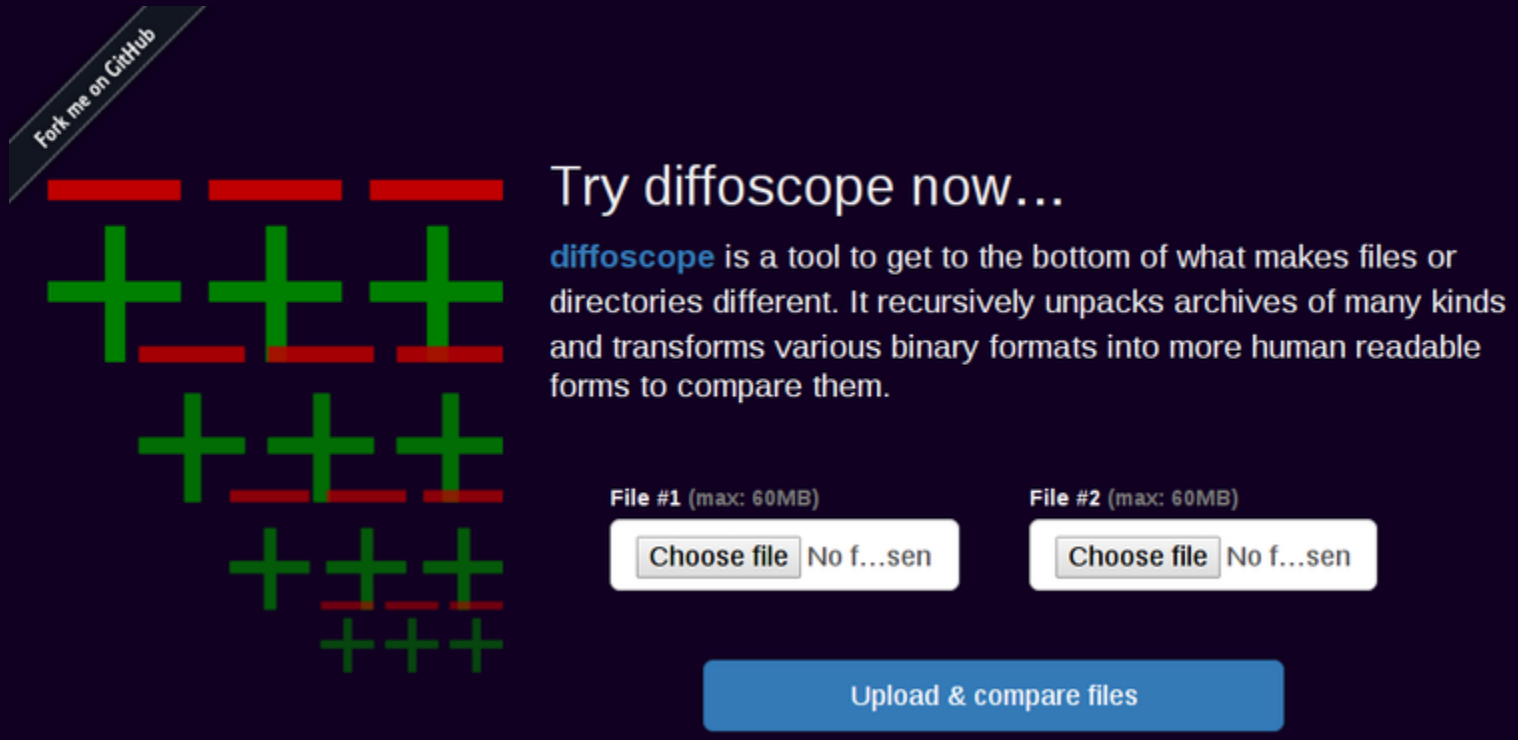
> *diffoscope* will try to get to the bottom of what makes files or directories different. It will recursively unpack archives of many kinds and transform various binary formats into more human readable form to compare them. It can compare two tarballs, ISO images, or PDF just as easily.



https://diffoscope.org/


├── aspell-de_20131206-5_all.deb
│ ├── metadata
│ │ rw-r--r-- 0/0 4 Jun 11 16:19 2014 debian-binary
│ │ -rw-r--r-- 0/0 2893 Jun 11 16:19 2014 control.tar.gz
│ │ -rw-r--r-- 0/0 329600 Jun 11 16:19 2014 data.tar.xz
│ │ +rw-r--r-- 0/0 2875 Jun 11 16:19 2014 control.tar.gz
│ │ +rw-r--r-- 0/0 329596 Jun 11 16:19 2014 data.tar.xz
│ ├── control.tar.gz
│ │ ├── control.tar
│ │ │ ├── md5sums
│ │ │ │┄ Files in package differ
│ ├── data.tar.xz
│ │ ├── data.tar
│ │ │ ├── ./usr/lib/aspell/de_affix.dat
│ │ │ │ #
│ │ │ │ -# Version: 20131206 (build 20150801)
│ │ │ │ +# Version: 20131206 (build 20150802)
│ │ │ │ #
│ │ │ ├── ./usr/share/aspell/de-common.cwl.gz
│ │ │ │ ├── metadata
│ │ │ │ │ -gzip compressed data, last modified: Sat Aug 1 18:21
│ │ │ │ │ +gzip compressed data, last modified: Sat Aug 1 18:24

Android APK files, Android boot images, Ar(1) archives, Berkeley DB database files, Bzip2 archives, Character/block devices, ColorSync colour profiles (.icc), Coreboot CBFS filesystem images, Cpio archives, Dalvik .dex files, Debian .buildinfo files, Debian .changes files, Debian source packages (.dsc), Device Tree Compiler blob files, Directories, ELF binaries, Ext2/ext3/ext4/btrfs filesystems, FreeDesktop Fontconfig cache files, FreePascal files (.ppu), Gettext message catalogues, GHC Haskell .hi files, GIF image files, Git repositories, GNU R database files (.rdb), GNU R Rscript files (.rds), Gnumeric spreadsheets, Gzipped files, ISO 9660 CD images, Java .class files, JavaScript files, JPEG images, JSON files, LLVM IR bitcode files, MacOS binaries, Microsoft Windows icon files, Microsoft Word .docx files, Mono 'Portable Executable' files, Ogg Vorbis audio files, OpenOffice .odt files, OpenSSH public keys, OpenWRT package archives (.ipk), PDF documents, PGP signed/encrypted messages, PNG images, PostScript documents, RPM archives, Rust object files (.deflate), SQLite databases, SquashFS filesystems, Statically-linked binaries, Symlinks, Tape archives (.tar), Tcpdump capture files (.pcap), Text files, TrueType font files, XML binary schemas (.xsb), XML files, XZ compressed files, etc.



try.diffoscope.org

Future work

Communicating concept to end-users?

Toolchain fixes

Improving developer tools

Mandating Debian packages be reproducible?

Source code remains vulnerable

Get involved!

Visit:	`reproducible-builds.org`
Follow:	**@ReproBuilds** on Twitter
Join:	`#reproducible-builds` on `irc.oftc.net`
Fix:	Bugs and toolchain issues!

Thank you!

@lolamby
lamby@debian.org

chris-lamb.co.uk
reproducible-builds.org