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

Holger Levsen
h01ger
based on Chris Lamb's slides

MiniDebConf Brasil
Curitiba
13th April 2018

My name is…

Holger Levsen

...e eu nao falo Português, desculpa!!

My nick is…

h01ger

went to CCCongress before Linux / IPv4

Debian user since 1995

Free software contributor for 15+ years

Freelancer

more about Debian and me

Debian contributor since 2001, DD since 2007

ex-debconf organizer 2004-2014 and dc-commitee member

video.debian.net

Debian Edu

piuparts.debian.org

jenkins.debian.net

qubes-os.org user

coreboot user

Ask me anything!

feel free to interrupt

or ask me later

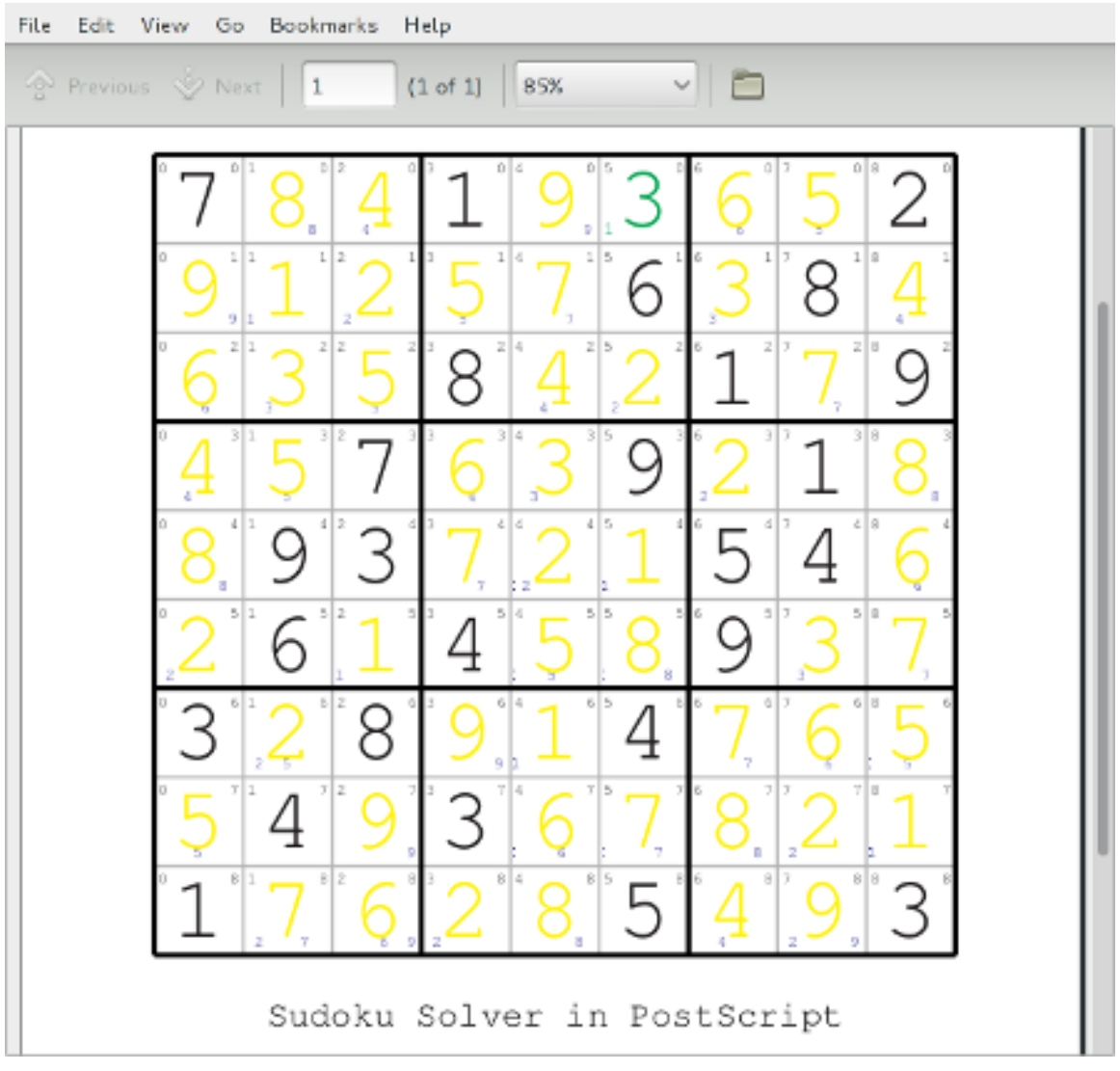
This talk is about

Reproducible Builds

in Debian started at DebConf13 by Lunar

I've been working on it since 2014

Slides taken from Chris 'lamby' Lamb


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

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


← Caro


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?

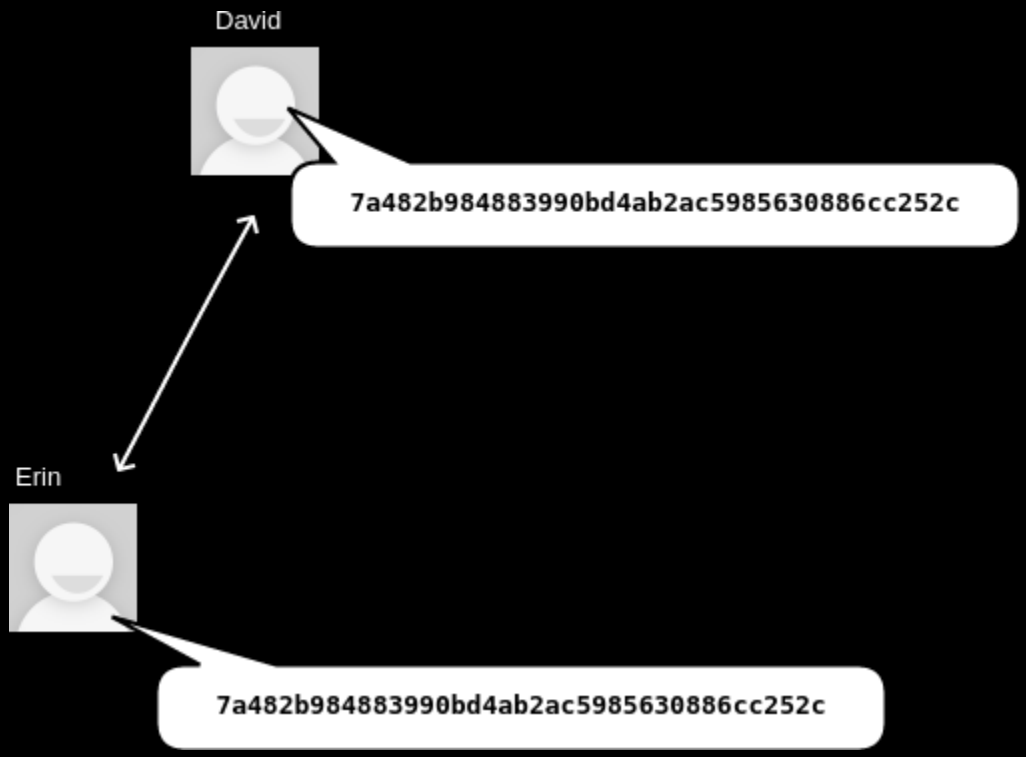
cker explains how he put "backdoor" in hundreds of Linux Mi wnloads

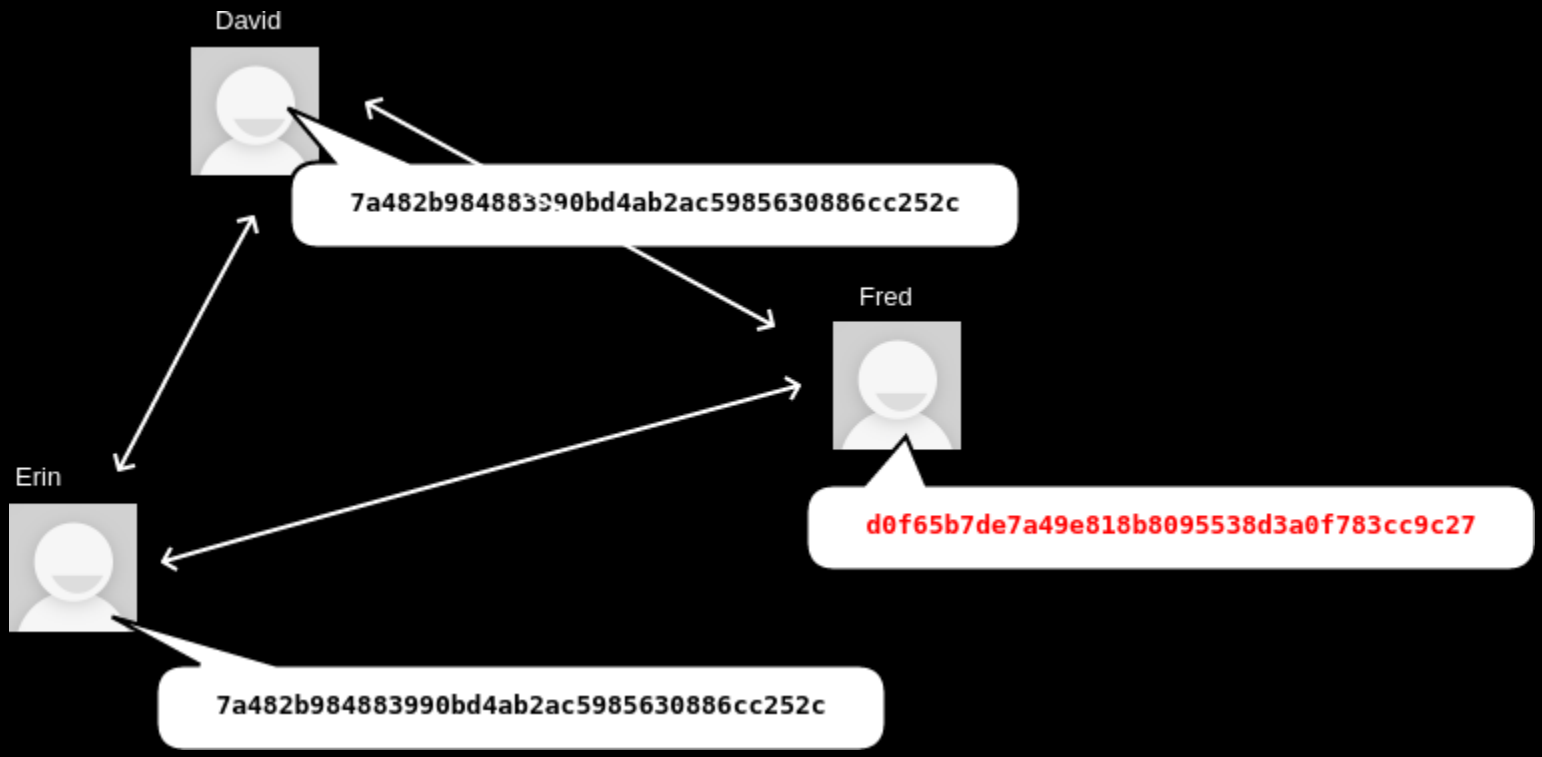
hacker said their prime motivation for the backdoor was to build a botnet.

Solution?

By Zack Whittaker for Zero Day | February 22, 2016 -- 01:28 GMT (01:28 GMT) | Topic: Security


David
7a482b984883990bd4ab2ac5985630886cc252c


David
7a482b984883990bd4ab2ac5985630886cc252c
Erin
7a482b984883990bd4ab2ac5985630886cc252c


David
7a482b984883990bd4ab2ac5985630886cc252c
Fred
Erin
d0f65b7de7a49e818b8095538d3a0f783cc9c27
7a482b984883990bd4ab2ac5985630886cc252c

How does this help?

Alice → Blackmail will be uncovered

Bob → Compromise detected

Carol → Tampered laptop will be discovered

Reduces incentive to attack in the first place

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

"Reliable" builds... ✖

"Builds with the same dependencies"... ✖

Identical build results

Wait…

Dictionary/hash/database ordering

Parallelism in builds

Timestamps

Build paths

Non-deterministic file ordering

Users, groups, `umask`, environment variables, etc.

Other advantages

Minimal diffs on "deliberate" changes

Cache ratio — save time, money & CO_2

Detect corrupted build environments

Remove build-dependencies

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 of this build 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

- 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

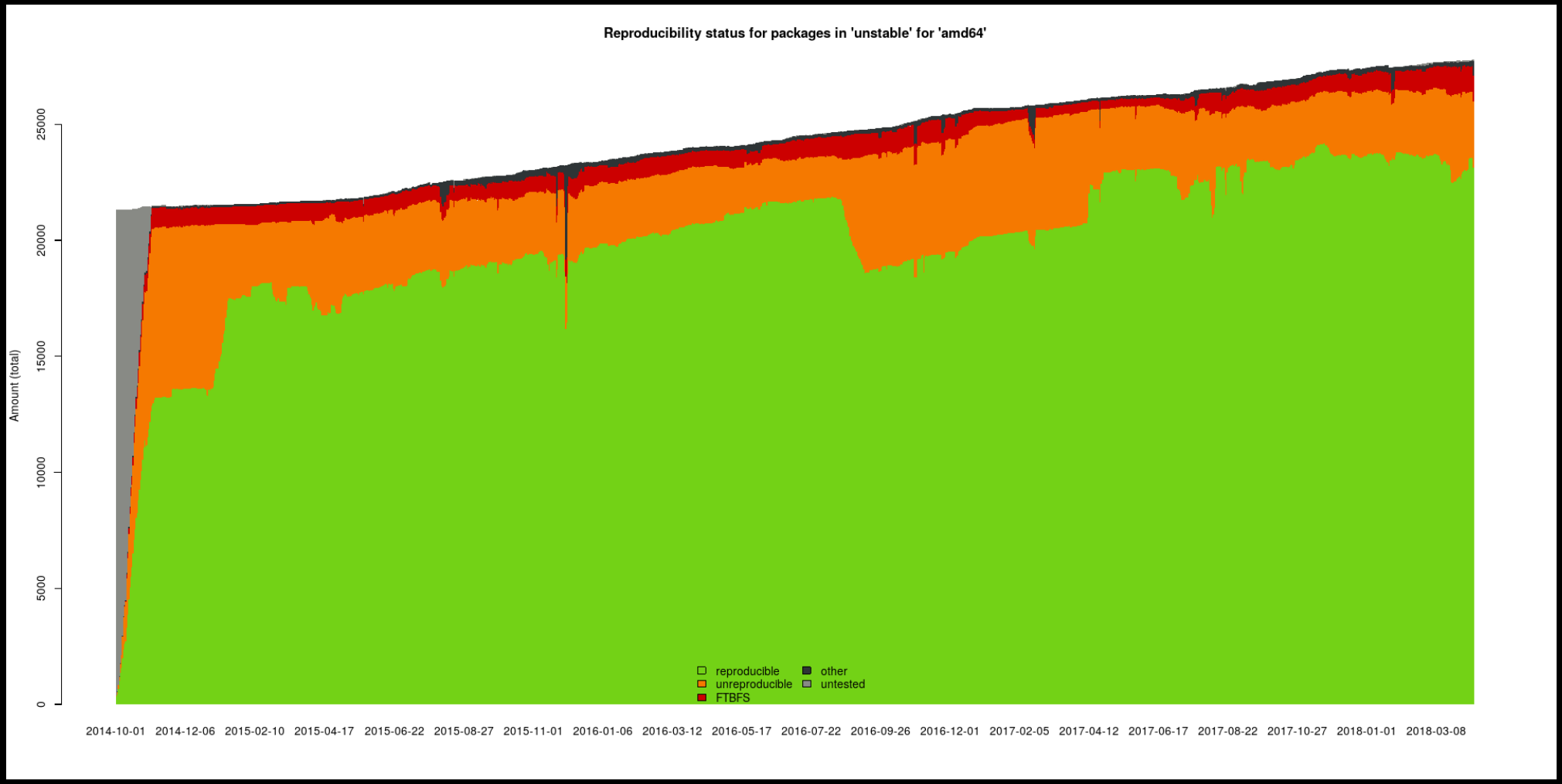
Filesystem (`disorderfs`)

Timezone & locale

`uid` & `gid`

Kernel & CPU type

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


Reproducibility status for packages in 'unstable' for 'amd64'
Amount (total)
0
5000
10000
15000
20000
25000
2014-10-01
2014-12-06
2015-02-10
2015-04-17
2015-06-22
2015-08-27
2015-11-01
2016-01-06
2016-03-12
2016-05-17
2016-07-22
2016-09-26
2016-12-01
2017-02-05
2017-04-12
2017-06-17
2017-08-22
2017-10-27
2018-01-01
2018-03-08
reproducible
unreproducible
FTBFS
other
untested

isdebianreproducibleyet.com

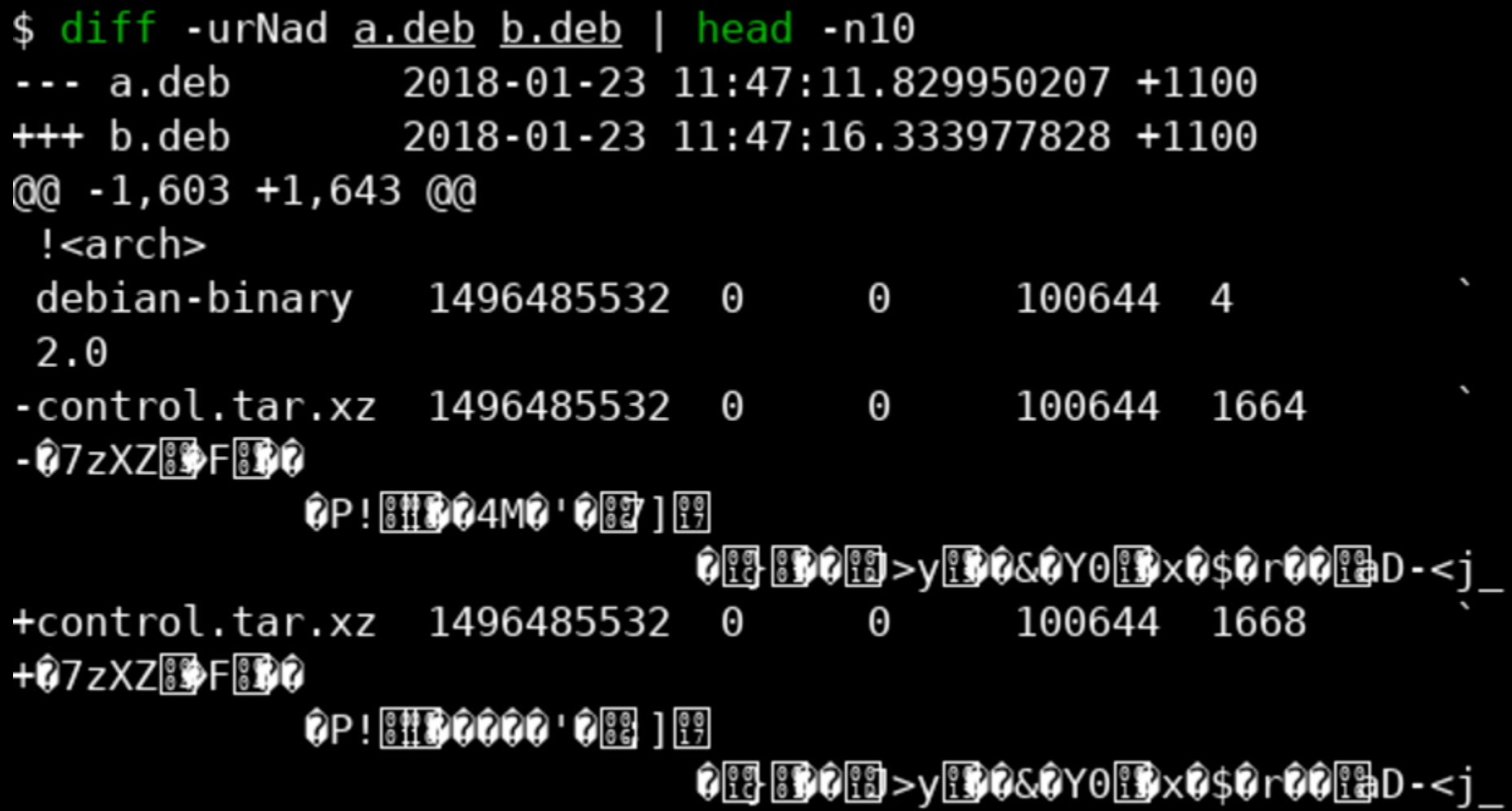
Beyond Debian…

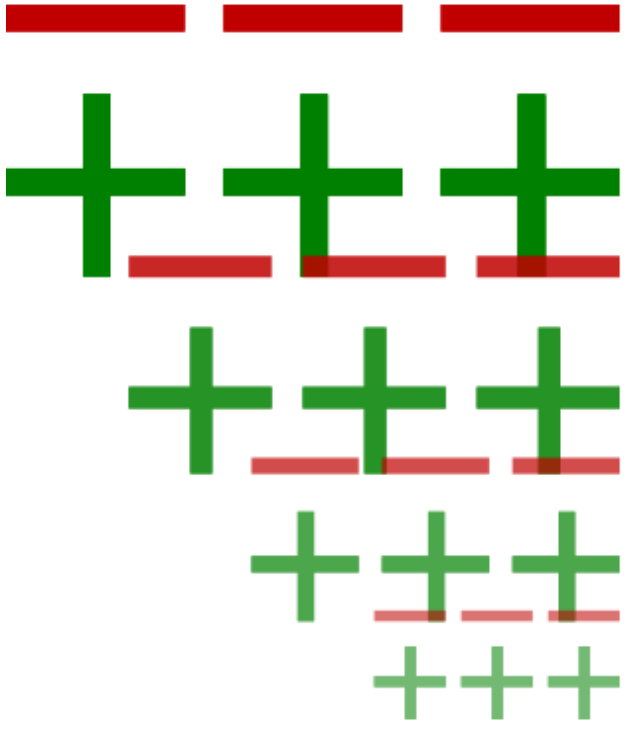
coreboot, OpenSUSE, OpenWRT, NetBSD, FreeBSD, Archlinux, Tails, Qubes, F-Droid, NixOS, Guix, Basel, Meson, etc.

Other projects 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.
```

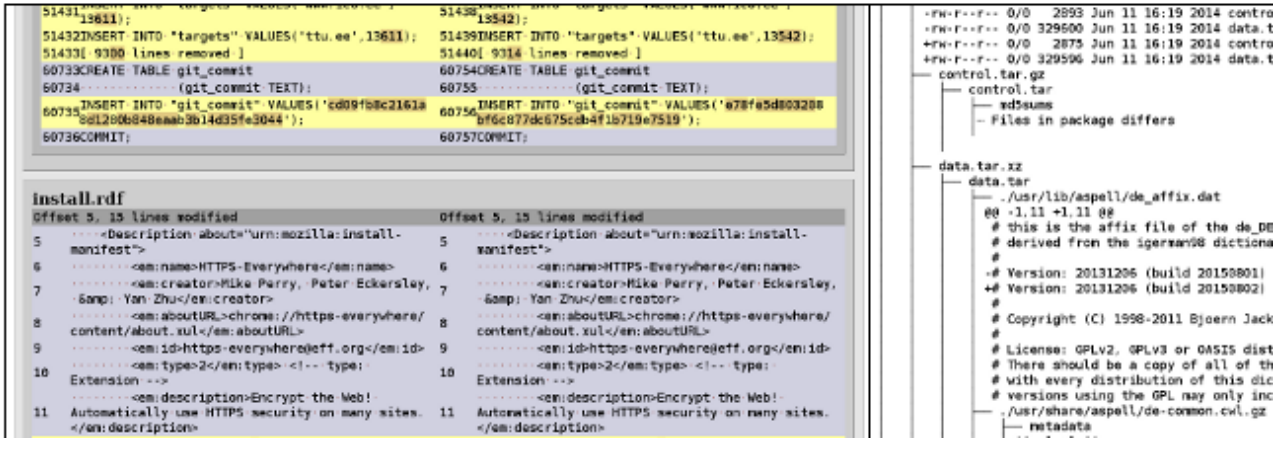

$ diff -urNad a.deb b.deb | head -n10
--- a.deb 2018-01-23 11:47:11.829950207 +1100
+++ b.deb 2018-01-23 11:47:16.333977828 +1100
@@ -1,603 +1,643 @@
!<arch>
debian-binary 1496485532 0 0 100644 4 `
2.0
-control.tar.xz 1496485532 0 0 100644 1664 `
-7zXZ F
P! 4M']
>y &Y0 x $ r aD-<j_
+control.tar.xz 1496485532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



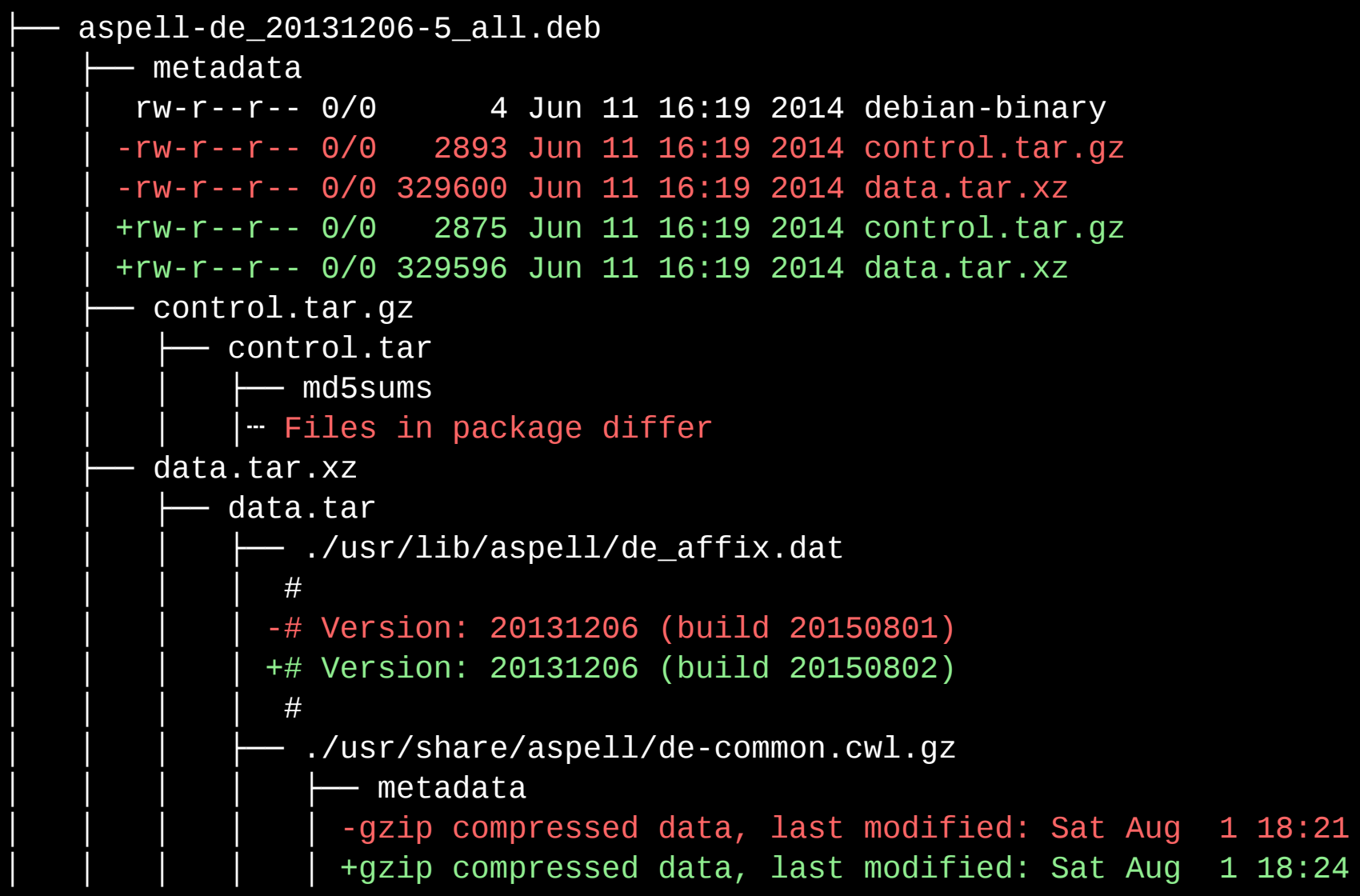
diffoscope

in-depth comparison of files, archives, and directories

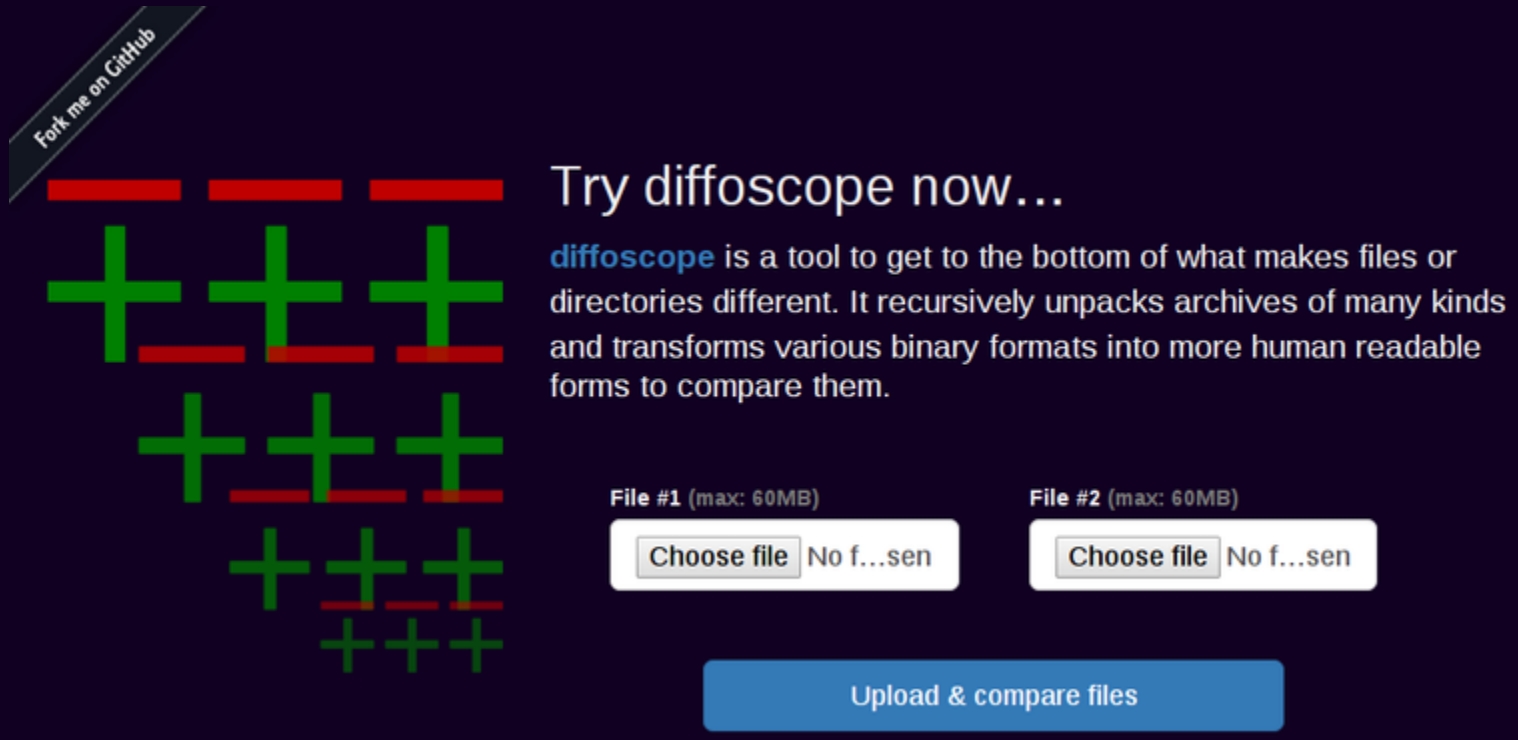
> *diffoscope* will try to get to the bottom of what makes files or directories different. It will recursively unpack archives of many kinds and transform various binary formats into more human readable form to compare them. It can compare two tarballs, ISO images, or PDF just as easily.



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

Show differences in security uploads

diffoscope != definition of reproducible!

Binary blobs (eg. router images)

What's left to do?

Source code

Programming errors

Backdoors / obfusticated code

Weak algorithms

Code with "testing" modes

```
$ apt install python-pywt-doc
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following NEW packages will be installed:
  python-pywt-doc
0 upgraded, 1 newly installed, 0 to remove and 4 not upgrad
Need to get 102 kB of archives.
After this operation, 978 kB of additional disk space will
WARNING: The following packages are not reproducible!
  python-pywt-doc
Install these packages anyway? [y/N]
```

Toolchain fixes (GCC!)

Improving developer tools

Mandating Debian packages be reproducible?

Defeating *Trusting Trust*…?

Get involved!

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

Obrigado!

h01ger
holger@debian.org

reproducible-builds.org
tests.reproducible-builds.org