

The Debian Archive

A look behind the CABAL

Team-Motto: Fuck it! What could possibly go wrong?

Joerg Jaspert

joerg@debian.org

Taiwan MiniConf 2009

Taipei, Taiwan, September 2009



debian

Outline

- 1 Team
- 2 History
- 3 Layout
- 4 Queues
- 5 Dinstall
- 6 Package life



debian

The Team

Me fail English? That's impossible.

FTP Masters

- Joerg Jaspert - Ganneff
- Mark Hymers - mhy



debian

The Team

Me fail English? That's impossible.

FTP Masters

- Joerg Jaspert - Ganneff
- Mark Hymers - mhy

FTP Assistants

- Frank Lichtenheld - djpig
- Mike O'Connor - stew
- Barry deFreese - bdefreese
- Torsten Werner - twerner
- Alexander Reichle-Schmehl - tolimar
- Chris Lamb - lamby



debian

Outline

- 1 Team
- 2 History
- 3 Layout
- 4 Queues
- 5 Dinstall
- 6 Package life



debian

History - Scripts

Motto: Fuck it! What could possibly go wrong?

- Ian Murdock/Jackson makeshift scripts
- Monolithic perl dinstall written by Guy Maor
 - Not exactly secure: upload, shipit
 - Not even using PGP either.
- Katie: a rewrite in python. Known as DAK.



debian

History - Archive

That guy impressed me and I am not easily impressed. Wow. BLUE Hair.

- Architectures / Releases:

Bo 1: i386

Hamm 2: + m68k

Slink 4: + sparc, alpha

Potato 6: + powerpc, arm

Woody 11: + mips, mipsel, hppa, ia64, s390

Sarge 11 (unofficial amd64 release)

Etch 11: + amd64, - m68k (there is a semi-official etch-m68k release)

Lenny 12: + armel

Squeeze ?: + kfreebsd-i386, kfreebsd-amd64, - arm, ...

- Proposed Architectures: avr32, sh{3,4}



debian

History - Archive size

Ganneff - if it goes wrong, we make it “Fuck it what do we care?”, but dont tell anyone.

- Releases:

Buzz	229Mb
Rex	306Mb
Bo	1.2Gb
Hamm	2.1Gb
Slink	4.1Gb
Potato	6.7Gb
Woody	27Gb
Sarge	57Gb
Etch	82Gb
Lenny	125Gb
squeeze+X	1Tb?

- Daily pushes:

2005 200Mb up to 2Gb, average 1Gb. Once a day.

2009 1Gb up to 6Gb, average 2Gb. 4 times a day.



debian

Outline

- 1 Team
- 2 History
- 3 Layout
- 4 Queues
- 5 Dinstall
- 6 Package life



debian

Suites

You can't keep blaming yourself. Just blame yourself once, and move on.

- A suite is an (upcoming) Debian release...
 - etch
 - lenny
 - squeeze
 - sid (will never be released)



debian

Suites

You can't keep blaming yourself. Just blame yourself once, and move on.

- A suite is an (upcoming) Debian release...
 - etch
 - lenny
 - squeeze
 - sid (will never be released)
- ... or an extension to one
 - etch-proposed-updates
 - lenny-proposed-updates
 - experimental (takes from sid)



debian

Suites

You can't keep blaming yourself. Just blame yourself once, and move on.

- A suite is an (upcoming) Debian release...
 - etch
 - lenny
 - squeeze
 - sid (will never be released)
- ... or an extension to one
 - etch-proposed-updates
 - lenny-proposed-updates
 - experimental (takes from sid)
- Suites are also accessible via their functional name
 - oldstable (etch)
 - stable (lenny)
 - stable-proposed-updates
 - testing (squeeze)
 - unstable (sid)



debian

Suites

You can't keep blaming yourself. Just blame yourself once, and move on.

- A suite is an (upcoming) Debian release...
 - etch
 - lenny
 - squeeze
 - sid (will never be released)
- ... or an extension to one
 - etch-proposed-updates
 - lenny-proposed-updates
 - experimental (takes from sid)
- Suites are also accessible via their functional name
 - oldstable (etch)
 - stable (lenny)
 - stable-proposed-updates
 - testing (squeeze)
 - unstable (sid)
 - rc-buggy (experimental)



debian

Some Release Team Member didn't like rc-buggy all that much...



debian

Archive Layout

why - I was upset to discover (via grepping my logs) that I actually said that quote in one of the subtitles verbatim

- Legacy layout until Potato

- Everything below dists/\$SUITE

```
dists/$SUITE/  
  binary-all/   binary-alpha/ [...]   
  disks-alpha/  disks-arm/   [...]   
  source/  
  upgrade-alpha/ upgrade-i386/ [...]
```

- Each of those then had

```
admin/ base/ comm/ devel/ doc/ editors/  
[...] web/ x11/
```



debian

Archive Layout

why - I was upset to discover (via grepping my logs) that I actually said that quote in one of the subtitles verbatim

- Legacy layout until Potato

- Everything below dists/\$SUITE

```
dists/$SUITE/  
  binary-all/   binary-alpha/ [...]   
  disks-alpha/  disks-arm/   [...]   
  source/  
  upgrade-alpha/ upgrade-i386/ [...]
```

- Each of those then had

```
admin/ base/ comm/ devel/ doc/ editors/  
[...] web/ x11/  
Packages Packages.gz Release [Sources]
```



debian

Archive Layout

why - I was upset to discover (via grepping my logs) that I actually said that quote in one of the subtitles verbatim

- Legacy layout until Potato

- Everything below dists/\$SUITE

```
dists/$SUITE/  
  binary-all/   binary-alpha/ [...]   
  disks-alpha/  disks-arm/   [...]   
  source/  
  upgrade-alpha/ upgrade-i386/ [...]
```

- Each of those then had

```
admin/ base/ comm/ devel/ doc/ editors/  
[...] web/ x11/  
Packages Packages.gz Release [Sources]
```

- Huge traffic hit for mirrors at release time.
- Lots of duplicated files between releases.



debian

It makes you wanna die



Archive Layout

why - Ganneff airlines: departing from a window near you

- Pool layout since woody
 - Initially “just” a side effect
 - All package related files (binaries, source) now below `pool/$COMPONENT/$HASH/$SOURCEPACKAGENAME`

```
pool/main/m/mailfilter/  
mailfilter_0.3.3-1_alpha.deb  
mailfilter_0.6.2-1_mipsel.deb  
mailfilter_0.6.2.orig.tar.gz  
mailfilter_0.3.3-1.dsc  
mailfilter_0.6.2-1.dsc
```
 - Hash special case for libraries, there its `lib$HASH`



debian

Archive Layout

why - Ganneff airlines: departing from a window near you

- Pool layout since woody
 - Initially “just” a side effect
 - All package related files (binaries, source) now below `pool/$COMPONENT/$HASH/$SOURCEPACKAGENAME`

```
pool/main/m/mailfilter/  
mailfilter_0.3.3-1_alpha.deb  
mailfilter_0.6.2-1_mipsel.deb  
mailfilter_0.6.2.orig.tar.gz  
mailfilter_0.3.3-1.dsc  
mailfilter_0.6.2-1.dsc
```
 - Hash special case for libraries, there its `lib$HASH`
 - Indices and control information (Packages/Sources files) still below `dists/$SUITE`
 - Lots of disk space saved, lots of traffic at release time saved



debian

Archive Layout

why - Ganneff airlines: departing from a window near you

- Pool layout since woody

- Initially “just” a side effect
- All package related files (binaries, source) now below
`pool/$COMPONENT/$HASH/$SOURCEPACKAGENAME`

```
pool/main/m/mailfilter/  
mailfilter_0.3.3-1_alpha.deb  
mailfilter_0.6.2-1_mipsel.deb  
mailfilter_0.6.2.orig.tar.gz  
mailfilter_0.3.3-1.dsc  
mailfilter_0.6.2-1.dsc
```

- Hash special case for libraries, there its lib\$HASH
- Indices and control information (Packages/Sources files) still below
`dists/$SUITE`
- Lots of disk space saved, lots of traffic at release time saved
- With legacy layout a mirror would be at least 780Gb today
that would waste more than 300Gb



debian

Pool layout makes kittens happy



an

Outline

- 1 Team
- 2 History
- 3 Layout
- 4 Queues
- 5 Dinstall
- 6 Package life



debian

Queues

Oh, so they have Internet on computers now!

- UploadQueue
 - Processed by debianqueued
 - On multiple debian.org hosts, final “master” on ftp-master.debian.org
 - Maintainer can delete files using signed .commands files
 - Processed roughly every 5 minutes
 - debian queued on ftp-master moves complete uploads to **unchecked**



debian

Queues

Oh, so they have Internet on computers now!

- UploadQueue

- Processed by debianqueued
- On multiple debian.org hosts, final “master” on ftp-master.debian.org
- Maintainer can delete files using signed .commands files
- Processed roughly every 5 minutes
- debian queued on ftp-master moves complete uploads to **unchecked**

- unchecked

- Processed every 15 minutes except during “dinstall”
- Automated package checks
 - Checksums
 - Signature
 - Version
 - ...
- If existing package -> go to **accepted**
- If byhand in upload .changes -> go to **Byhand**
- If new package -> go to **NEW**



debian

What? A NEW package? Are you sure?



debian

Queues

why - oh god. Ganneff has been on TV? I'm surprised the sets didn't all explode

- NEW/Byhand

- Packages with new overrides or byhand parts need human processing
- Byhand package usually needs unpacking of files somewhere
- NEW needs lots of checks for the package, when all is OK overrides are added.
- If package is accepted (overrides added) -> go to **accepted**
- If package is rejected it ends in queue/reject



debian

And if people want to discuss a rejection...



debian

...and really insist on their point...



...we can still win.



debian

Queues

I am so smart! I am so smart! S-M-R-T! I mean S-M-A-R-T...

- accepted
 - “Final” queue
 - Files get installed into pool/ during “dinstall”
 - Visible as `http://incoming.debian.org/`



debian

Queues

I am so smart! I am so smart! S-M-R-T! I mean S-M-A-R-T...

- accepted
 - “Final” queue
 - Files get installed into pool/ during “dinstall”
 - Visible as `http://incoming.debian.org/`
- Holding, done, reject - not queues per se, but in there.



debian

Outline

- 1 Team
- 2 History
- 3 Layout
- 4 Queues
- 5 Dinstall
- 6 Package life



debian

dinstall

Hi, my name's Moe. Or as the ladies like to refer to me, 'hey you in the bushes'

- “Magic” Shell script
- Current runtime between 3 and 4 hours (4 times a day)



debian

Tiresome long job



루니

Debian

dinstall

I'm normally not a praying man, but if you're up there, please save me Superman.

- “Magic” Shell script
- Current runtime between 3 and 4 hours (4 times a day)
- Installs packages,
- updates indices,
- does “everything we need to make our users happy”
- Currently 47 functions to run
- Most timeconsuming:
 - Cleanup of apt-ftparchive databases: 100 minutes
 - Indices (Packages/Sources/Contents) with apt-ftparchive: 60 - 70 minutes
 - Wanna-Build trigger: 20 minutes
 - Package installs (accepted queue): 5 - 50 minutes
- From those only cleanup runs backgrounded



debian

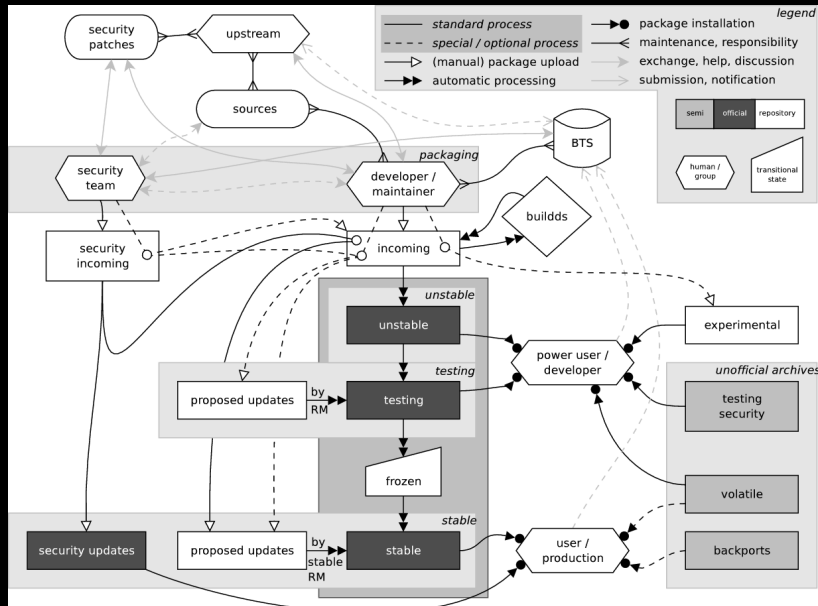
Outline

- 1 Team
- 2 History
- 3 Layout
- 4 Queues
- 5 Dinstall
- 6 Package life



debian

The life of a package in Debian



Thank you for sleeping through this talk

Ganneff - (trust full, when had i been that drunk?)

Joerg Jaspert

ftpmaster@debian.org

<http://ftp-master.debian.org/>

<https://ftp-master.debian.org/git/dak.git>



debian