Quality Assurance in Lenny+1

Lucas Nussbaum
lucas@debian.org
<pusling> QA is just about maintaining packages no one else wants
Quality Assurance

- **Goal:** Improve the quality of Debian as a whole
- Not really a team, more a central place to discuss and work on QA stuff
- #debian-qa, debian-qa@lists.d.o

What do "we" do?

- Maintain the PTS, DDPO (Debian Developers’ Packages Overview), and many other tools
  - Don’t write your own tools for your team!
  - Write them for Debian, get them used by everybody!
- Develop tools to run archive-wide checks (and do the checks and mass bug filings)
- Take care of orphaned/neglected packages
- Track inactive maintainers (Missing In Action team)
(NEW) Bapase : BAd PAckages SEarch

- Use various sources of info about packages
- And a scoring system
- To get a list of suspicious packages, and orphan/remove them

Examples: Packages which are very buggy, neglected/useless, orphaned, etc.

<table>
<thead>
<tr>
<th>Package</th>
<th>Score</th>
<th>Action</th>
<th>O age</th>
<th>Testing</th>
<th>Migrate</th>
<th>Popcon</th>
<th>Wrapping</th>
<th>RFS</th>
<th>Bugs</th>
<th>Last upload</th>
<th>NMUs</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>metamail</td>
<td>114643</td>
<td>Should be removed (O pkg) (since 2008-01-28)</td>
<td>1404 (2004-10-01)</td>
<td>0</td>
<td>0</td>
<td>3263</td>
<td>O</td>
<td>0</td>
<td>12</td>
<td>151 (2008-03-08)</td>
<td></td>
<td>still has rdevps, prop_rm filed on rdep.</td>
</tr>
<tr>
<td>electric</td>
<td>112643</td>
<td>Should be removed (O pkg) (since 2008-01-30)</td>
<td>1238 (2005-03-17)</td>
<td>0</td>
<td>0</td>
<td>415</td>
<td>O</td>
<td>0</td>
<td>2</td>
<td>202 (2008-01-17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>libgstreamer-perl</td>
<td>103699</td>
<td>Should be orphaned (since 2008-02-06)</td>
<td>105 (2008-04-23)</td>
<td>0</td>
<td>0</td>
<td>47</td>
<td>ita</td>
<td>0</td>
<td>3</td>
<td>92 (2008-05-06)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(NEW) Ultimate Debian Database

GSOC project by Christian von Essen
Co-mentored with zack & HE

Gather a lot of existing data on Debian in the same SQL DB

=> SQL queries combining different data sources:
  - RC bugs sorted by popcon of the package they affect
  - (co-)maintainers with the most bugs
  - ...

Lucas Nussbaum  Quality Assurance in lenny+1
(NEW) Ultimate Debian Database (2)

Current status:

- data sources imported on a regular basis (cron or ssh-triggered).
- Packages/Sources (for Debian and Ubuntu), popularity contest (for Debian and Ubuntu), testing migrations, history of uploads, bugs, orphaned packages, lintian, carnivore, ...
- Runs on udd.debian.net, accessible from master

```bash
psql -U guest -h udd.debian.net udd
```

http://wiki.debian.org/UDD
Popular packages not in testing

```
SELECT DISTINCT unstable.source, insts
FROM (SELECT DISTINCT source FROM sources
WHERE distribution = 'debian' and release = 'sid')
AS unstable, popcon_src
WHERE unstable.source NOT IN
  (SELECT source FROM sources
   WHERE distribution = 'debian'
   AND release = 'lenny')
AND popcon_src.source = unstable.source
ORDER BY insts DESC;
```
Popular packages not in testing

lzo 64910
gnome-cups-manager 32241
db4.6 20780
ffmpeg-debian 13045
freetype1 10544
flashplugin-nonfree 7110
perlftlib 6752
nvidia-graphics-drivers 3855
wxwindows2.4 3643
dvi2tty 2252
kdebase-runtime 1760
easytag 1714
g-wrap 1586
yaird 1503
slocate 1501
youtube-dl 1401
Most popular orphaned packages

```
select pkgs.source, type, insts
from orphaned_packages pkgs, popcon_src pc
where pkgs.source = pc.source
and pkgs.type in ('O', 'ITA')
ORDER BY pc.insts DESC
```
## Most popular orphaned packages

<table>
<thead>
<tr>
<th>source</th>
<th>type</th>
<th>insts</th>
</tr>
</thead>
<tbody>
<tr>
<td>net-tools</td>
<td>ITA</td>
<td>73981</td>
</tr>
<tr>
<td>ispell</td>
<td>ITA</td>
<td>65844</td>
</tr>
<tr>
<td>vbetool</td>
<td>ITA</td>
<td>33978</td>
</tr>
<tr>
<td>mdbtools</td>
<td>ITA</td>
<td>30376</td>
</tr>
<tr>
<td>libdc1394</td>
<td>ITA</td>
<td>21908</td>
</tr>
<tr>
<td>imlib</td>
<td>O</td>
<td>21308</td>
</tr>
<tr>
<td>toshset</td>
<td>O</td>
<td>9271</td>
</tr>
<tr>
<td>gnome-libs</td>
<td>O</td>
<td>8458</td>
</tr>
<tr>
<td>reiserfsprogs</td>
<td>ITA</td>
<td>8270</td>
</tr>
<tr>
<td>htdig</td>
<td>O</td>
<td>7637</td>
</tr>
<tr>
<td>libnet</td>
<td>ITA</td>
<td>7185</td>
</tr>
<tr>
<td>xli</td>
<td>O</td>
<td>6914</td>
</tr>
<tr>
<td>pyogg</td>
<td>O</td>
<td>4595</td>
</tr>
<tr>
<td>pyvorbis</td>
<td>O</td>
<td>4502</td>
</tr>
<tr>
<td>procinfo</td>
<td>ITA</td>
<td>3843</td>
</tr>
<tr>
<td>gnuift</td>
<td>O</td>
<td>3698</td>
</tr>
</tbody>
</table>
Maintainers with the most lintian errors

(using carnivore!)

```sql
SELECT (SELECT cn.name || ' <' || ce.email || '>
FROM carnivore_names cn, carnivore_emails ce
WHERE ce.id = cn.id AND ce.id = (SELECT id
FROM carnivore_emails WHERE (temp.package_type IS NULL
AND email = (SELECT substring(p.maintainer from '.+ <(.+)>')
FROM packages p WHERE p.package = temp.package
AND p.distribution = 'debian' AND p.release = 'sid'
LIMIT 1)) OR (temp.package_type = 'source'
AND email = (SELECT substring(s.maintainer from '.+ <(.+)>')
FROM sources s WHERE s.source = temp.package
AND s.distribution = 'debian' AND s.release = 'sid'
LIMIT 1)) LIMIT 1) AS person,
SUM(temp.tag_count) AS tag_count
FROM (SELECT l.package AS package, l.package_type AS package_type,
COUNT(l.tag) AS tag_count
FROM lintian l WHERE (l.tag_type = 'error' OR l.tag_type = 'warning')
AND (l.package_type IS NULL OR l.package_type = 'source')
GROUP BY l.package, l.package_type) AS temp
GROUP BY person ORDER BY tag_count DESC;
```
Maintainers with the most lintian errors

4515 Debian Qt/KDE Maintainers <debian-qt-kde@lists.debian.org>
2782 Debian OpenOffice Team <debian-openoffice@lists.debian.org>
1556 Debian QA Group <packages@qa.debian.org>
1104 Stefan Hornburg <racke@debian.org>
1098 Isaac Clerencia <isaac@debian.org>
870 Debian GCC Maintainers <debian-gcc@lists.debian.org>
835 Debian GIS Project <pkg-grass-devel@lists.alioth.debian.org>
788 GNU Libc Maintainers <debian-glibc@lists.debian.org>
643 Michael Stone <mstone@cs.loyola.edu>
630 Debian X Strike Force <debian-x@lists.debian.org>
630 Jose Luis Rivas <ghostbar38@gmail.com>
621 Debian Java Maintainers <pkg-java-maintainers@lists.alioth.debian.org>
597 Mutsumi <ishikawa@isoternet.org>
581 Gurkan Sengun <gurkan@phys.ethz.ch>
571 Muammar El Khatib <muammarelkhatib@gmail.com>
541 NIIBE Yutaka <gniibeb@debian.org>
532 Turbo Fredriksson <turbo@papadoc.bayour.com>
529 Ari Pollak <compwiz@bigfoot.com>
524 LI Daobing <lidaobing@gmail.com>
520 Javier Fernandez-Sanguino Pen~a <jfs@ieeeesb.etsit.upm.es>
So, what can you do?

- Talk to us!
  Join #debian-qa, debian-qa@lists.d.o
- Adopt orphaned packages you care about.
  wnpp-alert (in devscripts) is your friend.
- Improve our tools (see bugs filed against qa.debian.org)
- Think about mass bug filings about real issues, and talk to us!
But this is supposed to be a BOF!

(Ideas of) things to discuss:

- >700 orphaned packages! Is that OK?
  - Should we aggressively remove them?
  - Ignore them?
  - Split Debian into a main (supported) and an universe (unsupported) section?

- Status of MIA (Missing In Action, tracking of inactive maintainers)

- How could we make the QA team more attractive?

- Ideas of new checks/tests/MBF/stuff to do for lenny+1?
  - Anyone interested in working on piuparts?