

KS 2014

Reviewing new APIs/ABIs

Michael Kerrisk, man7.org

August 2014

Outline

1	Reviewing new APIs/ABIs	1-1
1.1	Reviewing new APIs/ABIs	1-1

Outline

1	Reviewing new APIs/ABIs	1-1
1.1	Reviewing new APIs/ABIs	1-1

Summary

- Insufficient testing
- Missing unit tests/poor test coverage
- Missing/poor specifications
- Insufficient review(er)s
- CC linux-api@

Insufficient testing

- I find bugs in about 50% of *released* syscalls
- Typically: poor coverage testing of new APIs
- Apparently: sometimes no testing *at all* of new/specific features



- Silent regressions...
- Example: inotify `IN_ONESHOT`
 - Initially: `IN_ONESHOT` (by design) didn't trigger `IN_IGNORED`
 - 2.6.3x: fsnotify reworking
 - Now: `IN_ONESHOT` does trigger `IN_IGNORED`
 - No-one noticed... (I spotted in mid-2014)
 - Presumably unfixable

Missing/poor specifications

- Often no / thin spec
- How can we review?
- How can we test? (implementation versus intention)
- Subtle details can cause pain for user space
 - Example: inotify rename events
 - `IN_MOVED_FROM` and `IN_MOVED_TO` events not necessarily contiguous
 - `IN_MOVED_FROM` and `IN_MOVED_TO` event pair is **not** atomically inserted
- Missing specs result in buggy implementations
 - Example: `recvmsg()` *timeout* argument
 - <http://article.gmane.org/gmane.linux.network/324996>



The 'timeout' argument implements three cases:

- 1) 'timeout' is NULL: the call blocks until 'vlen' datagrams are received.
- 2) 'timeout' points to {0, 0}: the call (immediately) returns up to 'vlen' datagrams if they are available. If no datagrams are available, the call returns immediately, with the error EAGAIN.
- 3) 'timeout' points to a structure in which at least one of the fields is nonzero. The call blocks until either:
 - a) the specified timeout expires
 - b) 'vlen' messages are received

In case (a), if one or more messages has been received, the call returns the number of messages received; otherwise, if no messages were received, the call fails with the error EAGAIN.

If, while blocking, the call is interrupted by a signal handler, then:

- * if 1 or more datagrams have been received, then those datagrams are returned (and interruption by a signal handler is not (directly) reported by this or any subsequent call to `recvmsg()`).
- * if no datagrams have so far been received, then the call fails with the error EINTR.



Missing documentation

- APIs and feature additions still regularly appear without a man-pages patch

- Not enough reviewers
- Poor designs slip through too easily
 - Example: good case that `O_TMPFILE` should have been a new syscall
- I try to do a lot of this, but my time is unpaid and very bursty
 - (I'd be interested in conversations about changing that situation...)



- `linux-api@vger.kernel.org`
- <https://www.kernel.org/doc/man-pages/linux-api-ml.html>
- Provides heads-up for
 - man-pages
 - glibc
 - LTP
 - Tools (strace, trinity, ...)
 - Others

