Kernel development:
How things go wrong

(And why you should participate anyway)

Jonathan Corbet
LWN.net
corbet@lwn.net
Kernel development is a success

~5 releases/year
> 10,000 changes/release
> 1000 developers/release

Linux is showing up in everything

...it works!
Kernel development is a success

~5 releases/year
> 10,000 changes/release
> 1000 developers/release

Linux is showing up in everything

...it works!
So why talk about failure?
High profile failures give the kernel a bad name.
“A key Linux contributor has admitted the developer community can be intimidating and hard to break into.”
“A key Linux contributor has admitted the developer community can be intimidating and hard to break into.”

(Seen on slashdot - must be true)
Failure can teach us things
“It's fine to celebrate success, but it is more important to heed the lessons of failure.”
“It's fine to celebrate success, but it is more important to heed the lessons of failure.”
--Bill Gates
“A bridge, under its usual conditions of service, behaves simply as a relatively smooth level surface on which vehicles can move. Only when it has been overloaded do we learn the physical properties of the materials from which it is built.”

-- Herbert Simon
One note

The kernel community does not lack for clowns.
One note

The kernel community does not lack for clowns.

I am not talking about them.
This talk will be naming names

Every developer I name has my respect!
“Hey, all my other theories made sense too. They just didn't work. But as Edison said: I didn't fail, I just found three other ways not to fix your bug.”

-- Linus Torvalds
Tux3

A next-generation filesystem by Daniel Phillips

2008-07-23  Initial announcement
2008-11-25  Booting as root filesystem
2009-08-16  Last commit
“Do NOT fall into the trap of adding more and more stuff to an out-of-tree project. It just makes it harder and harder to get it merged. There are many examples of this.”

-- Andrew Morton
Daniel kept adding features
...then lost interest
“Anyway, Andrew Morton was right, we should have merged into mainline as soon as Tux3 was booting as root.”
-- Daniel Phillips
Lessons

Out-of-tree code is nearly invisible
  Few users
  Few contributors
  Little momentum
Lessons

Get it into the mainline early!
em28xx

...a video4linux driver

2005-11-08   Initial driver merge
...
2008-01-05   Markus Rechberger's final em28xx patch
2008-11-02   Replacement patch rejected
2009-08-09   Markus's final kernel patch
“Companies should be aware that if they try to submit any code to you they will loose the authority over _their_ work.”
-- Markus Rechberger
Another example

May, 2004

Hans Reiser tries to block the addition of new functionality to reiserfs.
Lessons

Contributing means losing control

Others *will* improve your code
“The fact is, maintainership does _not_ mean ownership. It means that you should be _responsible_ for the code, and you get credit for it, but if problems happen you do NOT “own” it. Not at all.”

-- Linus Torvalds
2.5.x IDE

2002-02-15  Martin Dalecki's first “IDE cleanup” patch
2002-03-08  IDE18, subsystem takeover
2002-08-09  IDE115 merged
2002-08-16  Martin quits, all IDE work reverted
“Breakage is the price you have to pay for advancements”
-- Martin Dalecki
Don't break things!

Listen when people complain
Deadline scheduler

Con Kolivas's scheduler rewrite

2007-03-04 First post
2007-03-05 Linus amenable to merging
2007-03-19 Linus gets irritated
2007-04-13 Molnar posts CFS
2007-07-10 CFS merged for 2.6.23
2007-07-25 Con leaves the kernel community
“So, I've had enough. I'm out of here forever. I want to leave before I get so disgruntled that I end up using windows.”

-- Con Kolivas
Lessons

Improve the kernel for everybody
...or at least don't make it worse
Lessons

Some parts of the kernel are hard to change.
Lessons

Participate in the wider discussion -ck list did not help
Lessons

Aim for a solution to the problem
...rather than inclusion of specific code
reiser4

2002-10-29  First code post
2003-07-24  2.6.0-test merge request
2004-08-19  Added to 2.6.8.1-mm2
2005-09-11  Push for 2.6.14
2006-07-20  Push for 2.6.19
2006-10-11  Hans Reiser arrested
What were the problems?

Non-POSIX filesystem behavior
Numerous technical difficulties
Hard-to-reproduce benchmarks
Antagonistic approach to others
Memories of reiser3
Lessons

Linux is not a research system
Lessons

Visionary brilliance will not excuse a poor implementation
Lessons

It's better not to accuse others of conspiring against you

Photo: Rob!
Lessons

The community remembers past actions
Developers also think far into the future

Photo: krupp
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-11</td>
<td>DTrace debuts</td>
</tr>
<tr>
<td>2005-10</td>
<td>RHEL4 introduces SystemTap</td>
</tr>
<tr>
<td>2008-07</td>
<td>FTrace merged</td>
</tr>
<tr>
<td>2009-06</td>
<td>Perf Events merged</td>
</tr>
<tr>
<td>2009-09-22</td>
<td>SystemTap 1.0 released</td>
</tr>
<tr>
<td>????</td>
<td>SystemTap merged</td>
</tr>
</tbody>
</table>
2008 Kernel Summit

50% had tried to use SystemTap
20% succeeded
“I thought everyone learned the lesson behind SystemTap's failure: when it comes to tooling/instrumentation we don't want to concentrate on the fancy complex setups and abstract requirements drawn up by CIOs as development isn't being done there. Concentrate on our developers today, and provide no-compromises usability to those who contribute stuff.”

-- Ingo Molnar
In other words...

If kernel developers don't see the value
...it won't go in.
TALPA

Posted in August 2008
Never merged as such

The goal:
Provide hooks for virus scanners
Problems with TALPA

Kernel developers disliked it
  Why bother with broken security models?

Badly-expressed requirements
  No threat model
  Solutions not needs
fanotify

Merged in August, 2010 (2.6.36)

Provides hooks for virus scanners
What changed?

Featured a cleanup of file event notification
   Replaced inotify and dnotify

Rephrased requirement:
   “Enable virus scanners to hook into file operations without using rootkit techniques.”
Lesson

Patches must be sold to developers
Not managers or customers
Other examples

Android wakelocks
Distributed storage
TuxOnIce
Wireless extensions
CML2
msleep()
Xen
utrace
...

...
Why bother?
It's not as hard as it seems
A slightly elite club

“Well, you don't get to be a kernel hacker simply by looking good in speedos”
-- Rusty Russell
Jobs

If you show that you can get code into the kernel, you will get job offers.
Influence

It's how you get the kernel to meet your needs.
“If we don't succeed we run the risk of failure”
-- Dan Quayle
Questions?