Status of Embedded Linux

Tim Bird
CELF Architecture Group Chair
Outline

- Kernel Versions
- Technology Areas
- CELF Contract Work
- Embedded Distributions
- Conferences
- What to do?
- Resources
Linux Kernel Versions
Kernel Versions

- Linux v2.6.29 – 23 Mar 2009
- Linux v2.6.30 – 10 June 2009
- Linux v2.6.31 – 9 Sep 2009
- Linux v2.6.32 – 2 Dec 2009
- Linux v2.6.33 – 24 Feb 2010  
  - 4 kernels in 11 months
- Linux v2.6.34-rc4 – 13 Apr 2010  
  - Expect 2.6.34 some time in May??
Linux v2.6.29

- SquashFS
  - Andrew Morton – “Merge it! We’ve long needed a filesystem named after a vegetable.”
  - Linus Torvalds – “You make a very compelling case.”

- BTRFS

- Kernel mode setting
  - Ability to set graphics mode in kernel

- Asynchronous Function Calls
  - To allow concurrency in initcalls at boot
Linux v2.6.30

- TOMOYO security module
  - Path-based mandatory access control
- Integrity Measurement
  - Kernel interaction with TPM hardware
- Threaded interrupts
Linux v2.6.31

- **Ftrace features**
  - Generic filters, function profiler, new tracepoints, new documentation
- **SMACK security module logging**
- **Performance counters**
  - [http://lwn.net/Articles/311850/](http://lwn.net/Articles/311850/)
- **kmemleak - detect kernel memory leaks**
  - [http://lwn.net/Articles/187979/](http://lwn.net/Articles/187979/)
Linux v2.6.32

- devtmpfs
  - Dynamic, fast population of /dev
  - http://lwn.net/Articles/331818/
- Timechart tool
  - New tool to create SVG chart of kernel events
  - See http://blog.fenrus.org/?p=5
- Runtime Power Management (core)
  - http://lwn.net/Articles/347573/
Linux v2.6.33

- LZO kernel compression
- Perf improvements
  - Probe, bench, diff, perl scripts, filters
  - ARM support
- Ramzswap/Compcache
- Android patches removed from -staging
Linux v2.6.34-rcx

- LogFS
  - Log-structured flash filesystem
Patches to watch

- kbuild: kconfig CROSS_COMPILE option
  - Puts ARCH and CROSS_COMPILE into files in build directory
  - Makes it easier to build for multiple architectures
  - Mainline effort stalled, was recently re-submitted again
Patches to watch (cont.)

- asm-generic
  - Simplifications and cleanup of architecture support
  - Should make it easier to add architectures in future
- See “Porting Linux” presentation by Jon Masters
- See http://lwn.net/Articles/357803/ and http://lwn.net/Articles/307713/
Technology Areas
Technology Areas

- Audio, Video, Graphics
- Bootup Time
- File Systems
- Power Management
- Real-time
- Security
- System Size
- Tracing/Tools
Audio/Video/Graphics

- GStreamer
  - Is de-facto standard for video management for handhelds
- OpenCore/OMX
  - Android media layer
- DirectFB
  - Continued uptake in TV domain
- OpenGL ES
  - De-facto standard for 3D API (/Gaming in embedded?)
- Clutter (GTK UI Framework)
Bootup Time

- Bootup time is a hot topic (due to netbooks)
- Intel 5-second boot on Moblin
  - See http://lwn.net/Articles/299483
- Asynchronous function calls
  - Mainlined in 2.6.29
  - See http://lwn.net/Articles/314808
- `scripts/bootgraph.pl` for visualization of initcalls
- New Readahead techniques
- Aggressive application optimizations (especially for X)
Bootup Time (cont.)

- devtmpfs
  - A form of devfs again (oh no!)
  - 0.6 second faster boot-time
    - On desktops
    - Need reports of savings on embedded systems
Bootup patches to watch

- Speed up symbol resolution during module loading
  - Binary search in symbol table avoids linear lookup cost per symbol when loading a module
  - Developer, Alan Jenkins, reports saving 20% of time on coldboot
    - Depends on modules being loaded
- Also, Carmelo Amorosso has patches for hash table lookup (even faster, but requires symbol table storage change)
  - See presentation at ELC 2009
File Systems - SquashFS

- Compressed, read-only FS
- Mainlined in 2.6.29
  - Was out-of-tree for years
  - CELF contracted author to mainline it
- Recent news:
  - Patches for LZMA support were submitted to kernel mailing list Dec 7.
  - See V2 patches submitted December 10
    - See http://lkml.org/lkml/2009/12/10/456
  - It looks like it missed the 2.6.34 merge window
File systems - UBIFS

• General purpose flash filesystem, built on wear-leveling layer above flash
  • See http://lwn.net/Articles/276025
• Mainlined in 2.6.27
• Recent news:
  • Toshiba has good report on UBIFS performance and issues:
    • JLS presentation does not appear to be available, but Jamboree presentation is at http://elinux.org/images/f/f8/CELFJamboree30-UBIFS_update.pdf
  • CELF has a candidate project to fund UBIFS mount-time speedups, but is having trouble finding a contractor
More File Systems

- **LogFS**
  - Log-structured file system which keeps meta-data on flash
  - Mounts really fast (no need to build in-memory tables to start accessing)
  - Mainlined (after much work) in 2.6.34
  - CELF supported this a few years ago
    - Introduction
      - [http://lwn.net/Articles/234441/](http://lwn.net/Articles/234441/)
    - Announcement of mainlining
      - [http://lwn.net/Articles/377741](http://lwn.net/Articles/377741)

- **AXFS**
  - Advanced XIP file system
  - Where is it??
File Systems Issues

- Patches of interest:
  - VFAT patent workaround
    - 2 attempts by Andrew Tridgell to work around Microsoft VFAT long-name patent
    - First attempt was controversial, because functionality was lost
      - New approach preserves functionality
  - http://lwn.net/Articles/339641
  - VFS-based union mounts
    - See http://lkml.org/lkml/2009/5/18/289
Power Management

- Wakelocks submitted by Google
  - Kernel PM developers rejected the approach
  - But this opened a discussion of the requirements
  - See http://lwn.net/Articles/318611

- Runtime Power Management
  - Adds ability to suspend and resume individual system components
  - See http://lwn.net/Articles/347573/
  - See Kevin Hillman’s slides at:
    - http://elinux.org/ELC_2010_Presentations

- In progress – asynchronous suspend and resume
Real-time – RT-preempt

• Interrupt threads - Mainlined 2.6.30
  • Tutorial by Mike Anderson about interrupt threads and how to use them

• Sleeping Spinlocks
  • Spinlock name cleanup merged in 2.6.33
    • See http://lkml.org/2009/12/6/162
  • Macros for doing switch from spinlocks to semaphores remains to be merged

• Some issues remain with “Big Kernel Lock”
  • CELF may fund work here in 2010
Real-time – dual kernel

- Xenomai
  - Dual-kernel Real-time for Linux
  - Successor to Adeos/I-Pipe
  - See http://www.xenomai.org/
  - See Phillipe Gerum’s keynote at ELC Europe 2009:
    - State of real-time linux…
System Size / Memory

- smem
- Kernel image compression
- Ramzswap
- mem_notify patches
- -ffunction_sections
- XIP
  - Not much new development (AXFS?)
smem

- Tool for analyzing system memory
  - Better numbers than ps or top
    - Reports *Proportional Set Size* (PSS), which accounts better for shared pages, and doesn’t lie, like RSS
  - Home page: http://www.selenic.com/smem
  - See http://lwn.net/Articles/329458
- 0.9 release was made in November
  - See http://lwn.net/Articles/361497
  - Release includes a man page and smemcap
    - smemcap = lightweight snapshotting tool for using smem with an embedded system
Kernel Image Compression

- LZMA support
  - Support for LZMA kernel image compression (up to 30% better than gzip)
  - Added in 2.6.29

- LZO kernel image compression
  - See http://lwn.net/Articles/350985
  - Compression not so great, but faster than others
  - Merged in 2.6.33
RamzSwap (compcache)

- Provides compressed swap, in memory
  - Uses LZO-compressed RAM based block device as swap disk
- Mainlined in 2.6.33
- See http://lwn.net/Articles/334649
- Home page: http://compcache.googlecode.com/
- http://code.google.com/p/compcache/wiki/Performance
  - Author writes: “Embedded devices: TODO”
  - Someone please test and post their results…
Compcache results

RAM usage with/without compcache

Memory (kB)

0 10000 20000 30000 40000 50000 60000 70000 80000 90000 100000

Web pages footprint

0 NULL A B C D E F
Mem_notify patch

- Long history of memory notification mechanisms
  - See http://lwn.net/Articles/266586/
- Latest memory notify patch was modification to cgroup system
  - http://elinux.org/Memory_Management#OOM_notification_in_cgroups
  - Discussions got tangled up in cgroup issues and generic event mechanisms
    - See http://lkml.org/lkml/2009/7/7/409
- Work was funded by CELF
S Size Patches to watch

- CONFIG_PRINTK_VERBOSITY
  - Allows compiling out printk messages below a certain printk level
Tracing

- Ftrace
- LTTng
- SystemTap
- Perf Counters
FTrace

- Is a new system to provide kernel tracing
- Generic framework for adding tracing to the kernel
  - Provides multiple tracers, selectable at runtime
  - Infrastructure for tracepoint definition, data capture (ring buffer), tracer control and trace output
- Core mainlined in 2.6.27
  - More bits coming (generic filtering in 2.6.31)
- See
  http://people.redhat.com/srostedt/ftrace-world.odp
Ftrace

- New stuff in 2.6.33:
  - Regular expression support in tracing filters
  - Tracing of accesses and modifications to arbitrary kernel variables
  - Dynamic probes for ftrace
    - Perf tool can place and use dynamic tracepoints

Perf

- Perf tool now has “diff” mode which shows change in performance between two runs
- Scripting support in perf - http://lwn.net/Articles/373842
Security

- Security Modules:
  - Tomoyo
  - SMACK
  - Embedded SELinux
- Integrity Measurement
Tomoyo Linux

- Path-based mandatory access control security module
- See http://lwn.net/Articles/277833/
- See http://elinux.org/TomoyoLinux
- Mainlined in 2.6.30
  - Was a big deal to get a path-based LSM into the kernel
    - Presentations on using Tomoyo with Android at Japan Technical Jamborees 27 and 28
Smack

- Simple Mandatory Access Control Kernel
  - Simple name-based MAC
- Mainlined in 2.6.25 kernel
- Seems good for embedded (low overhead)
  - CELF started project to evaluate TV use case
  - See http://lwn.net/Articles/292291
Embedded SE Linux

- I haven’t heard much lately
- See Mike Anderson’s presentation – Using SELinux on a router
Integrity Measurement

- Allows kernel to interact with TPM and ensure trusted operation
- See [http://lwn.net/Articles/302043/](http://lwn.net/Articles/302043/)
- Mainlined in 2.6.30
- See also [http://lwn.net/Articles/137306](http://lwn.net/Articles/137306)
- Search: “Integrity Measurement Architecture IBM”
Miscellaneous Stuff

• Device Trees for ARM
  • Device Trees is a mechanism to pass info from bootloader to kernel
    • Supported in X86 and PPC
    • If all platforms supported it, it could allow more uniform device drivers
    • Also, allows for a single binary to run on multiple platforms
  • Is a new hot topic

• DLNA
  • Open Source DLNA summit 3 – Planned for October in Cambridge, UK
  • CELF provided DLNA hardware and specs to OSS developers
Embedded Distributions
Embedded Distributions

- Moblin
- Android
- Maemo
- Ubuntu netbook
- Chrome
• Android 2.0 SDK (Eclair) released (28 Oct 2009)
• Number of devices **still** increasing
  • Non-phones starting to appear
  • “People of Lava” has announced an Android TV
• Android patches removed from mainline kernel (2.6.33)
  • Patches lived in drivers/staging for a while, but they’re gone now
  • Greg Kroah-Hartman talked about this yesterday
Meego!!

- Moblin transferred by Intel to Linux Foundation
- Moblin + Maemo = Meego
- Intel and Nokia joining forces to create single embedded distribution
- Ari Jaaski of Nokia, will talk about this at the Collaboration Summit
  - ELC had sessions on work done for Moblin, how to program Maemo, etc.
  - Collaboration Summit will have a whole track on Meego
Contract Work
CELF Contract Work

- smem
- OOM notifications in cgroups
- SquashFS enhancements
- SMACK on TV analysis
- Device-trees for ARM
- -ffunction-sections
- Open Project Proposal
- Matt Mackall – maintainer work
SMACK on TV Analysis

- Smack
  - Simple Mandatory Access Control Kernel (SMACK)
  - Mainlined in 2.6.25 kernel
  - See http://lwn.net/Articles/292291
  - Seems good for embedded (low overhead)
    - CELF started project to evaluate TV use case

- Project status:
  - White paper is on TV use case is on elinux wiki security page
Function-sections

- `-ffunction-sections` puts each function in its own linker section
- Allows eliminating more code by the linker
  - Which results in space savings
- Also, could support code re-ordering for bootup time reduction or improved cache utilization:
  - See [http://www.celinux.org/elc08_presentations/DDLink%20FunctionReorder%202008%202004.pdf](http://www.celinux.org/elc08_presentations/DDLink%20FunctionReorder%202008%202004.pdf)
- Watch Denys Vlasenko's presentation
Open Project Proposal

- Project ideas were collected in Dec. and Jan.
- See http://elinux.org/CELF_Open_Project_Proposal_2010
- Bids were solicited Feb. and Mar.
- It took longer than anticipated to solicit bids and arrange the vote
  - CELF members are now voting on the projects
  - Should have results for first round of 2010 contractor projects by the end of this week
- Most selected projects should get kicked in next two weeks
Matt Mackall

- Is now in MAINTAINERS file as an embedded Linux maintainer
  - Joined David Woodhouse
- CELF sponsors Matt to do miscellaneous development and advocacy for embedded Linux kernel work
- Got a problem interacting with linux-embedded?
  - E-mail Matt or David!
Hardware Donations

- CELF donates hardware for some open source developers
  - Recently: Beagle-boards, touchpad, PS3
- E-mail me if you need hardware for a project that you think CELF would be interested in
eLinux Wiki

- http://elinux.org/
- Wikipedia embedded Linux developers
- Very good for some technical areas
  - Stagnant in some areas
- Promotion
  - Contest – best editor in next 6 weeks wins a Lenovo internet tablet
  - Look for information on site (later this week)
- Please use it and post stuff
Conferences

• Future
  • LinuxCon Japan
  • ELC Europe 2010
  • ELC 2011

• See http://www.embeddedlinuxconference.com/
LinuxCon Japan

- September 27-29 in Tokyo, Japan
  - In Roppongi Hills
- CELF is a major sponsor
- Should be a good embedded track
ELC Europe 2010

- Cambridge, UK
- October 27, 28
- Tentative joint events (Oct 26)
  - Open Source DLNA summit
  - GStreamer conference
What can you do to help?

• Work at top of tree
  • Version gap is still biggest problem with embedded people working in community
  • I know it’s difficult
  • If your hardware isn’t supported: bug your semiconductor vendor and/or use a different board for top-of-tree work

• Don’t wait for someone else to test new features!!
  • Post results to wiki
  • Come to next conference and tell us what happened)
What to Test…

- New flash filesystems
  - LogFS, UBIFS, Yaffs2
- Boot-time:
  - Use devfs and measure improvement in boot times
- Size features:
  - Ramzswap on embedded platforms
  - -ffunction-sections
  - smem
- New tracing features:
  - trace-cmd, LTTng
- Runtime power management
Resources

- LWN.net
  - http://lwn.net/
  - If you are not a subscriber, please do so
  - This is an invaluable community resource that needs your support

- http://kernelnewbies.org/Linux_2_6_??

- eLinux wiki
  - http://elinux.org/

- Linux-embedded mailing list
  - http://vger.kernel.org/vger-lists.html#linux-embedded
Thanks!

Keep up the good work!