



CE Workgroup

Status of Embedded Linux September 2013

Tim Bird

Architecture Group Chair

LF CE Workgroup



CE Workgroup

Outline

Kernel Versions
Technology Areas
CE Workgroup Projects
Other Stuff
Resources



CE Workgroup

Outline

Kernel Versions

Technology Areas

CE Workgroup Projects

Other Stuff

Resources



CE Workgroup

Kernel Versions

- Linux v3.6 – 30 Sep 2012 – 71 days
- Linux v3.7 – 10 Dec 2012 – 71 days
- Linux v3.8 – 18 Feb 2013 – 70 days
- Linux v3.9 – 28 Apr 2013 – 69 days
- Linux v3.10 – 30 June 2013 – 63 days
 - I predicted July 7, 2013 – (7 days off)
- Linux v3.11 – 2 Sep 2013 – 64 days
- Linux v3.11 (no –rc yet)
 - I predict 3.12 on ...



CE Workgroup

Kernel Versions

- Linux v3.6 – 30 Sep 2012 – 71 days
- Linux v3.7 – 10 Dec 2012 – 71 days
- Linux v3.8 – 18 Feb 2013 – 70 days
- Linux v3.9 – 28 Apr 2013 – 69 days
- Linux v3.10 – 30 June 2013 – 63 days
 - I predicted July 7, 2013 – (7 days off)
- Linux v3.11 – 2 Sep 2013 – 64 days
- Linux v3.11 (no –rc yet)
 - I predict 3.12 on ... 8 Nov 2013 – 68 days



CE Workgroup

Linux v3.6

- Android RAM console functionality integrated into pstore
- CANFD support for CAN protocol
 - CAN with flexible data rate
- LED oneshot mode
 - Sysfs interface for certain one-time LED/gpio manipulations
- "Suspend to Both"
 - Create resume image both in RAM and on disk
 - If power dies during suspend, disk image can be used to resume



CE Workgroup

Linux v3.7

- ARM multi-platform support
 - See <http://lwn.net/Articles/496400/>
- ARM 64-bit support (Aarch64)
- Cryptographically signed kernel modules
 - See <https://lwn.net/Articles/470906/>
- Perf trace (alternative to strace)
 - Allows intermingling kernel trace events with `syscall` events
- Runtime power management for audio
- Kernaldoc system can output in HTML5 format



CE Workgroup

Linux v3.8

- F2FS – flash-friendly file system
 - See <https://lwn.net/Articles/518988/>
- New thermal governor subsystem
- Memory control group support for accounting for kernel memory usage
 - Stack and slab accounting and limits
- Cpuidle support for big.LITTLE



CE Workgroup

Linux v3.9

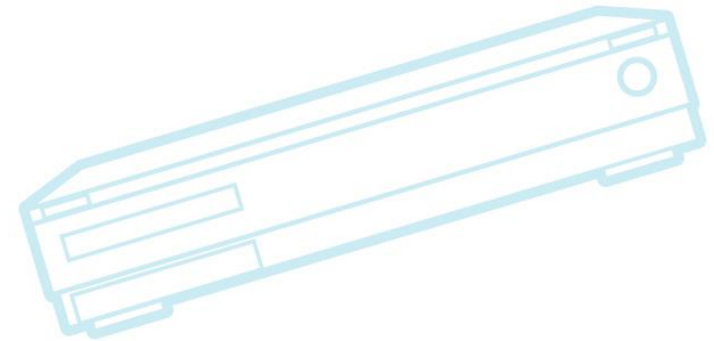
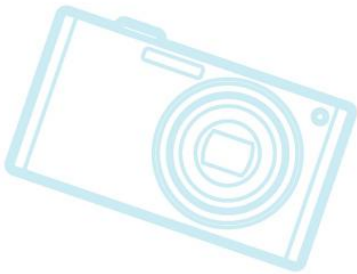
- Ftrace snapshots
 - Grab a snapshot of a running trace without stopping
- KVM virtualization for Cortex A15 processors
- PowerPC support for transactional memory
- CONFIG_EXPERIMENTAL=y
 - And should be gone soon
- 'make menuconfig' now has "save" and "load" buttons



CE Workgroup

Linux v3.9 (cont.)

- Descriptor-based GPIO
 - Access GPIOs by descriptor (e.g. by name in addition to by number)
 - Allows for grouping GPIOs - for “atomic” operations
 - Possibly useful for handling realtime issues
 - See <http://lwn.net/Articles/533632/>





CE Workgroup

Linux v3.10

- Full tickless (more later)
- Single zImage for ARM
 - Lots more platforms support multi-platform kernels
 - Arnd Bergmann shooting for almost-complete coverage by v3.12
- Multi-cluster power management
 - Partial support for big.LITTLE PM
 - <https://lwn.net/Articles/539082/>



CE Workgroup

Linux v3.10 (cont.)

- Multiple ftrace buffers
- Memory pressure control group support
 - Allows for notification if memory gets low
 - <http://lwn.net/Articles/531077/>



CE Workgroup

Linux v3.10 (cont.)

- Full tickless (full dynamic tick)
 - Under some circumstance, some processors may run with no periodic ticks at all
 - Previous CONFIG_NO_HZ used dynamic tick, but only when CPU was idle
 - New option is tri-state: periodic, idle, full
 - Boot CPU cannot be 'full' tickless
 - CPU cannot be full tickless with more than one process
 - <https://lwn.net/Articles/549580/>



CE Workgroup

Linux v3.11

- Power-efficient workqueues
 - Allow work to be done on any CPU, to avoid waking sleeping CPUs
- LZ4 kernel image compression
- Checkpatch –fix
 - Attempt to fix some simple errors
- F2fs continues to mature
 - Lots of patches from Samsung



CE Workgroup

Linux v3.11 (cont.)

- Zswap
 - "Zswap is a lightweight, write-behind compressed cache for swap pages. It takes pages that are in the process of being swapped out and attempts to compress them into a dynamically allocated RAM-based memory pool. ... This results in a significant I/O reduction and performance gains for systems that are swapping"
- See <https://lwn.net/Articles/551401/>



CE Workgroup

Linux 3.12 (probable)

- Full-system idle detection
 - Tricky rcu-based implementation to allow for fast indication of individual CPU idleness (using per-cpu variable), AND fast detection of global CPU idleness (single global variable)
- New cpu-idle driver that build on multi-cluster power management
 - I.e. Getting closer to support for “big.LITTLE” CPU scheduling
- Lots of device drivers converting over to device tree
 - More on this later



CE Workgroup

Things to watch

- Android features
 - Volatile ranges
 - ION memory allocator
- big.LITTLE MP scheduling
 - See <https://lwn.net/Articles/501501/>
 - See the In-Kernel-Switcher work
 - <https://lwn.net/Articles/549473/>
- Single zImage support on ARM
- Support for transactional memory instructions
 - Could be as big a deal (eventually) as locking primitives



Things to watch (longer-term)

- Non-volatile mass memory
 - Interesting remarks by Linus in LinuxCon 2012 keynote question and answer
 - Skeptical it will happen “this year” – it’s always within a few years of happening.
 - Won’t change a lot of kernel algorithms
 - Will mostly change filesystems
 - Byte-addressable storage has big implications for long-term storage
 - Applications will still segregate data between persistent and non-persistent groups
 - Things take longer to change than people think



CE Workgroup

Outline

Kernel Versions

Technology Areas

CE Workgroup Projects

Other Stuff

Resources



CE Workgroup

Bootup Time

- Not much new stuff this time
 - (Note: Watch out for device tree overhead...)
 - Sending device tree into random number pool



Graphics

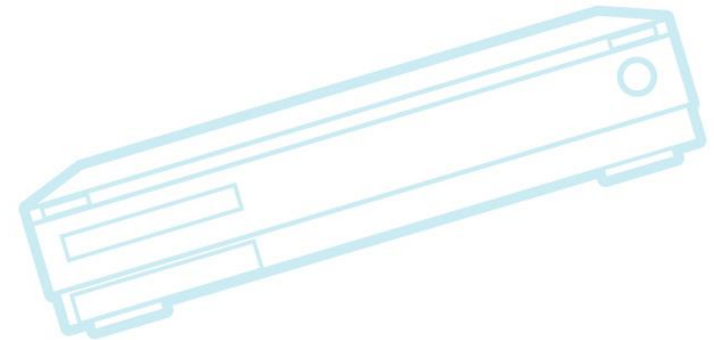
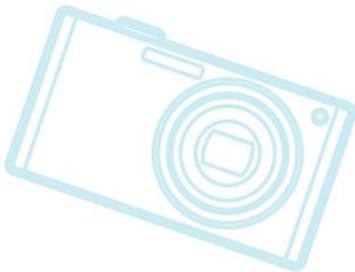
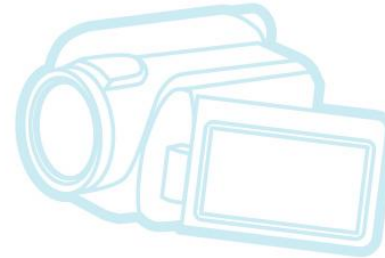
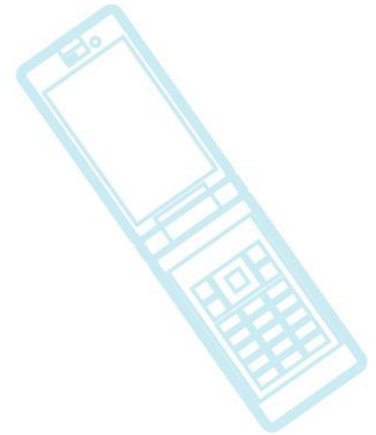
- “Mir vs. Wayland” battles in Desktop realm
- Android using skia, OpenGL ES, surfaceflinger, etc.
- Other embedded still using X, fbdev, Qt, Gtk/Cairo
 - Qt now at Digia (sold by Nokia)
- Kernel Mode Setting (KMS) support
 - Control graphics modes in kernel
 - Prevent race conditions in user-space
 - Maybe switch HWComposer from fbdev to KMS



CE Workgroup

File Systems

- F2FS
- Flash Filesystem Tuning guide
- exFAT





CE Workgroup

F2FS

- Flash-friendly filesystem
- Mainlined in Linux version 3.8
- Log-structured, with lots of tweaks
 - E.g. hot vs. cold data separation
- Written by Samsung
 - They continue to enhance it
 - e.g. Support for security attributes in 3.12
- Not sure if it's shipped in anything yet



Flash Filesystem tuning

- CE WG project to analyze filesystem performance on eMMC
- Goal: to test different block-based filesystems on flash media
 - Specifically, want to measure the effect of different kernel tuning options (IO scheduler, flash geometry vs. flash part attributes and workload characteristics)
- Result document is NOW available at:
 - http://elinux.org/File_Systems#Comparison_of_flash_filesystems



CE Workgroup

exFAT

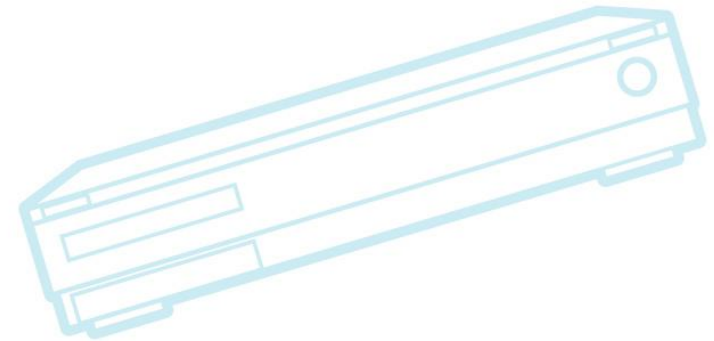
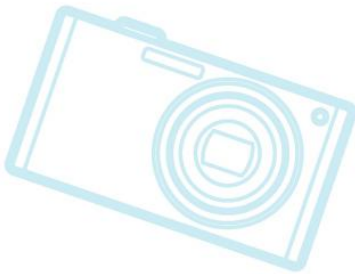
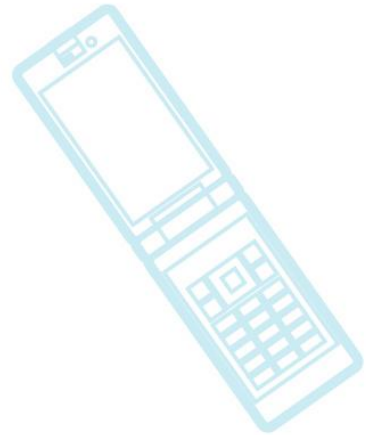
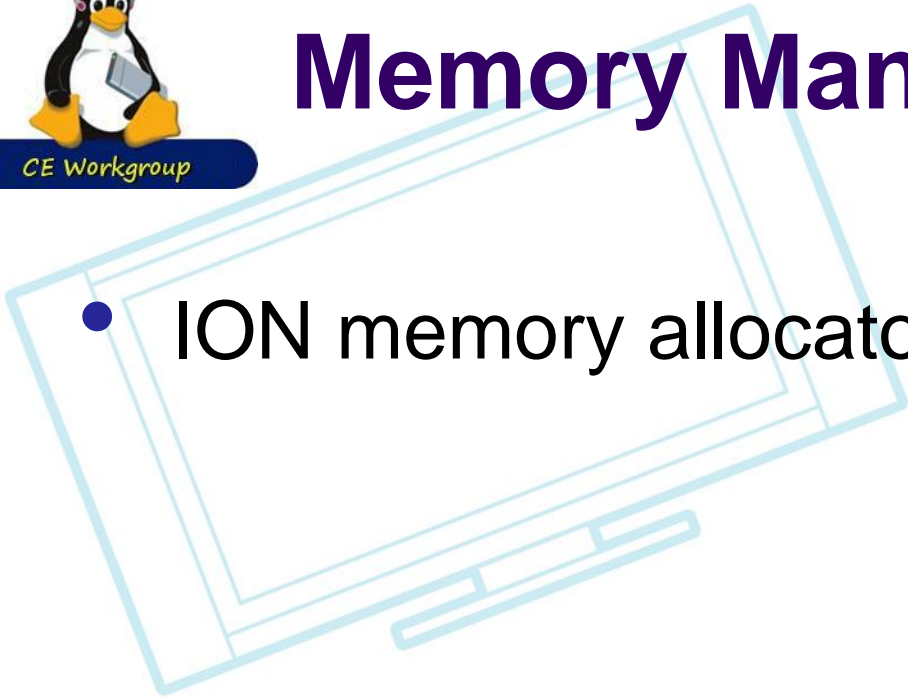
- Weird sequence of events
- Background: exFAT filesystem is covered by Microsoft patents
 - Used for sd cards – almost a requirement to support it
- exFAT code released by independent Russian developer
 - “Liberated” from Samsung
 - Not sure about license
 - But some code may have been derived from kernel
- Samsung released code a few weeks later
- I wouldn't use this code



CE Workgroup

Memory Management

- ION memory allocator





Ion memory allocator

- Allows sharing of memory areas between kernel subsystems (and devices)
 - Which reduces copies
- Different devices have different memory constraints (cached, contiguous, etc.)
 - ION can select memory areas matching the least-common-denominator of the constraints
- ION can manage cache relationship to memory
- But, ION uses arm-specific page accessors, and allows hardware-specific optimizations, so it may have difficulty getting mainlined



Power Management

- Power-aware scheduling:
 - Small-task packing
 - Try to migrate tasks to allow more CPUs to go idle
 - Task placement on mixed cpu_power systems
 - Move large tasks to faster CPUs
- Resources:
 - <http://lwn.net/Articles/546664> - overview
 - <http://lwn.net/Articles/552885> - some resistance
 - Ingo Molnar wants to consolidate this power stuff in the scheduler – rather than spread out into power/cpufreq/cpuidle/scheduler systems



big.LITTLE scheduling

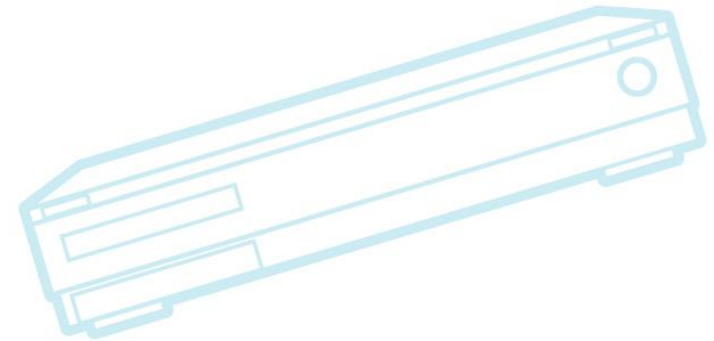
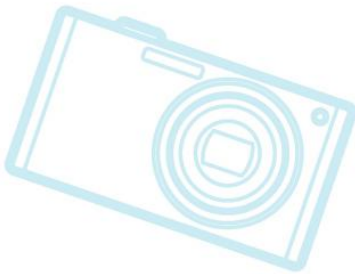
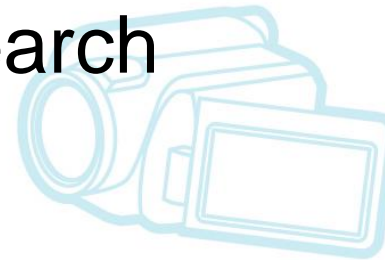
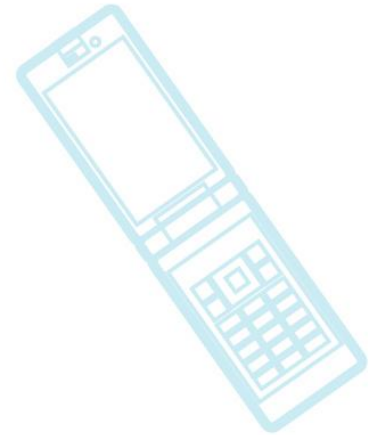
- Lots of work recently on big.LITTLE scheduling
 - Multi-cluster power scheduling
 - In-kernel-switcher work
 - See talk at LCJ by Nakagawa-san of Renesas
 - One User Space Approach to big.LITTLE MP System on Real Silicon
- Still waiting for real-product results ??



CE Workgroup

System Size

- Kernel size
- Library size
- Automated reduction research





CE Workgroup

Kernel size

- Cooperative memory relinquishment
 - Volatile Ranges
 - Lexmark work (membroker and ANR malloc)
 - See talk at ELC 2013 – "SystemWide Memory Management without Swap"



Library reduction

- olibc – bionic libc
 - Has good features from Android, and is smaller and more configurable than glibc

glibc 2.11 : /lib/libc.so	→ 1,208,224 bytes
uClibc 0.9.30 : /lib/libuClibc.so	→ 424,235 bytes
bionic 2.1 : /system/lib/libc.so	→ 243,948 bytes

- See ELC 2013 talk by Jim Huang
- Kconfig for egllibc
 - Ability to configure parts of libc to use

libc-2.17.so reduced from	1.2M -> 830K
ld-2.17.so reduced from	128K -> 120K
libm-2.17.so reduced from	610K -> 580K

- See ELC 2013 talk by Khem Raj



Advanced Size Optimization of the Linux Kernel

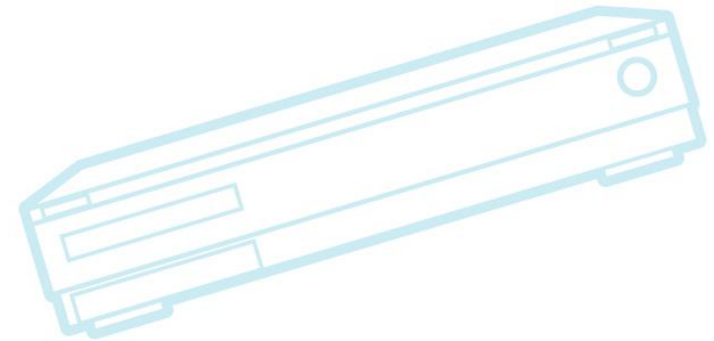
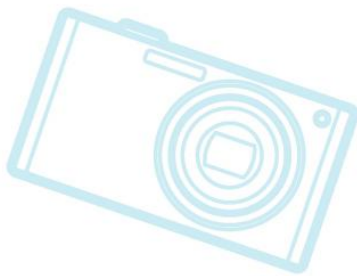
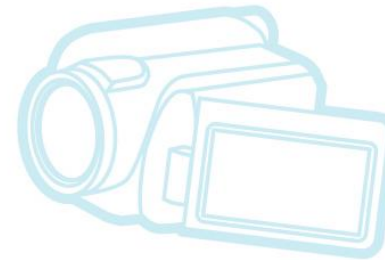
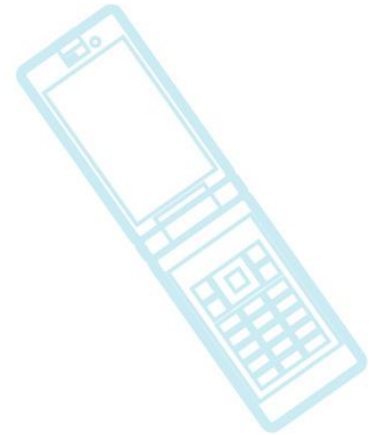
- “Auto-reduce” project
- Find automated ways to reduce the kernel
 - Link-time optimization
 - System call elimination
 - Kernel command-line argument elimination
 - Kernel constraint system
- Additional research
 - Link-time re-writing
 - Cold-code compression
- See Tim Bird’s presentation on advanced size optimization of the kernel
 - Notes and slides available at:
http://elinux.org/System_Size_Auto-Reduction



CE Workgroup

Security

- SMACK
- SE-Linux





CE Workgroup

SMACK

- SMACK for Tizen
 - Simplified rule set (3 tiers, 40,000 rules)
 - See <http://lwn.net/Articles/55278>



CE Workgroup

SE-Linux

- SE-Android
 - Implementation of SE-Linux for Android systems
- SE-Linux was previously too big for embedded
 - Early embedded SE-Linux required 2M
 - Desktop SE ruleset is 900,000 rules
- However, SE-Android only has 1658 rules and 263 types (71K policy size)
- <http://selinuxproject.org/page/SEAndroid>
 - Especially:
http://www.internetsociety.org/sites/default/files/Presentation02_4.pdf



CE Workgroup

Tracing

- Ktap
 - Dynamic tracing, without the overhead of compiling into a module
 - Adds an interpreter to the kernel
 - Single module, that leverages ftrace, kprobes, etc.
 - Prints results in ASCII
 - Good session in LinuxCon Japan by Jovi Zhang



Device Tree

- New requirements for implementing ARM board support and drivers
 - Separation of hardware description from code
 - Can use instead of “platform data”
- Has some problems
 - I have found it very complicated to use
 - Not mature yet
 - E.g. dma, pinctrl still being developed
 - Everyone defining their own bindings
 - Not enough documentation and examples
 - No type-checking or compile-time optimization



Device tree (cont.)

- Change in maintainership
 - Grant Likely transferred maintainership to others
 - Not enough review of bindings
- Discussion about having device tree be long-lived ABI to kernel
 - Should be usable by other operating systems
 - Maybe move out of kernel repository
- Lots of discussions planned at ARM mini-summit/Kernel Summit
 - Lots of presentations at ELC Europe
- See http://elinux.org/Device_Tree



CE Workgroup

Outline

Kernel Versions

Technology Areas

CE Workgroup Projects

Other Stuff

Resources



CE Workgroup

CEWG Contract Work 2013

- eMMC tuning guide completed
- Open Project Proposal period



CE Workgroup

eMMC tuning guide

- Description:
 - This project analysed EXT4, BTRFS and F2FS on a variety of block-based flash parts on a few different development boards
 - Output is a document describing best practices for tuning Linux block-based filesystems for block-based flash filesystems
 - Also, methods and scripts for filesystem testing
- Contractor: Cogent Embedded
- Status: Just completed
- Document at:
http://elinux.org/File_Systems#Comparison_of_flash_filesystems



CE Workgroup

Open Project Proposals

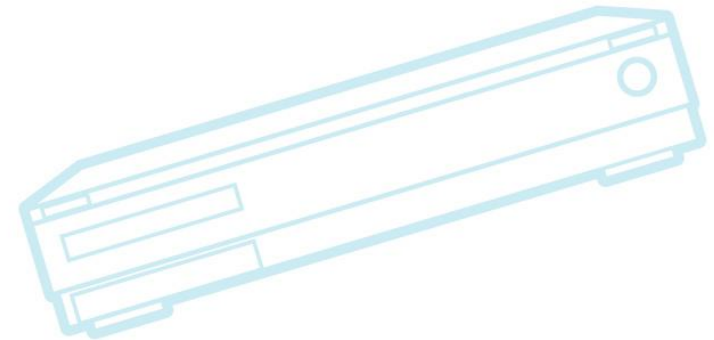
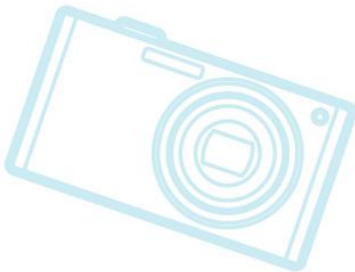
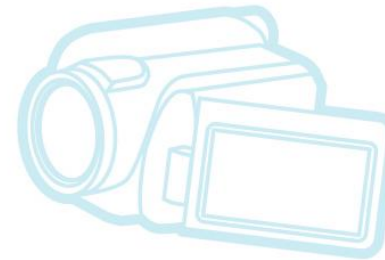
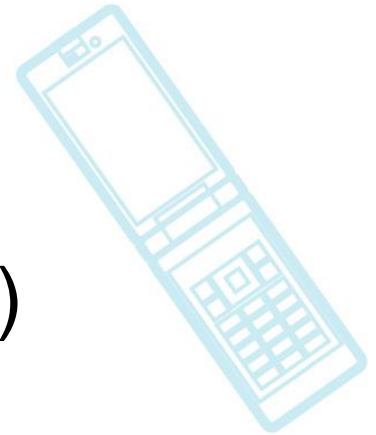
- Proposal period is now open
- See http://elinux.org/CEWG_Open_Project_Proposal_2013
- Looking for ideas for projects to fund.
- Deadline is October 2



CE Workgroup

Other Projects

- Long Term Support Initiative (LTSI)





CE Workgroup

Long Term Support Kernel for Industry

- LTSI 3.4 is available now
- Held a workshop at LinuxCon Japan
 - Discussed testing phase of project
 - Discussed promotion of project
- Program for free hardware for LTSI kernel testing
- Linux 3.10 is next community Long Term Stable kernel



CE Workgroup

Outline

Kernel Versions

Technology Areas

CE Workgroup Projects

Other Stuff

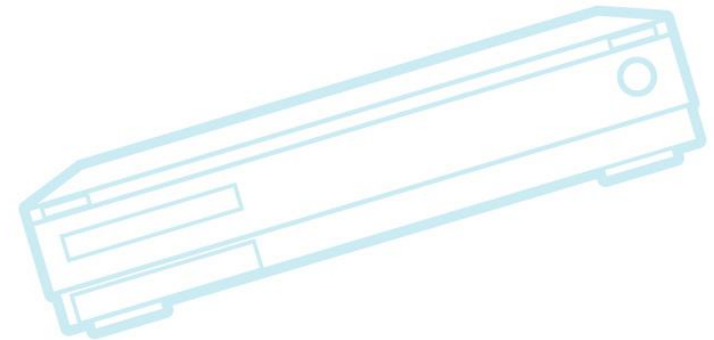
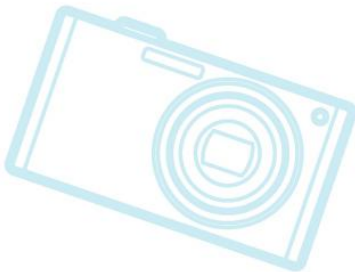
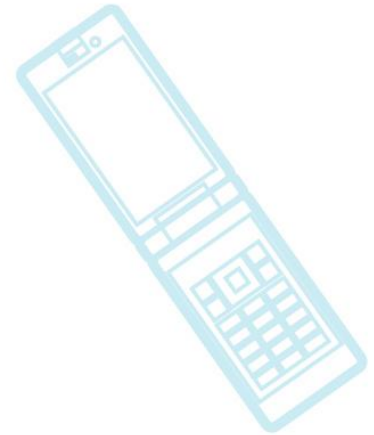
Resources



CE Workgroup

Other Stuff

- Tools
- Build Systems
- Events
- Miscellaneous





Tools

- Cortex
 - Coredump filter
 - Generates sparse coredump
 - See ELC 2013 presentation by Tristan Lelong
 - "Debugging for production systems"
- Debugging techniques
 - Good overview by Kevin Dankwardt at ELC 2013
 - "Survey of Linux Kernel Debugging Techniques"



Testing frameworks

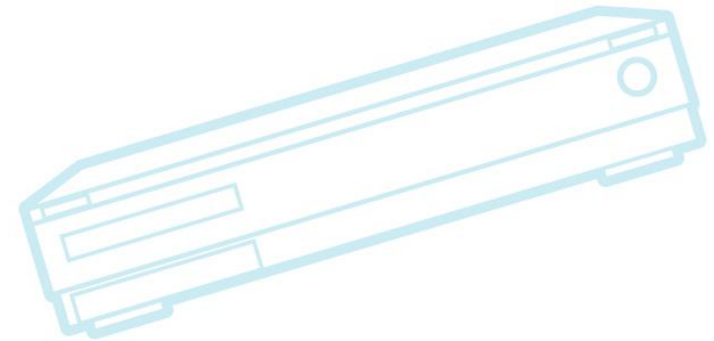
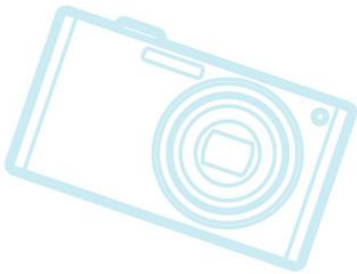
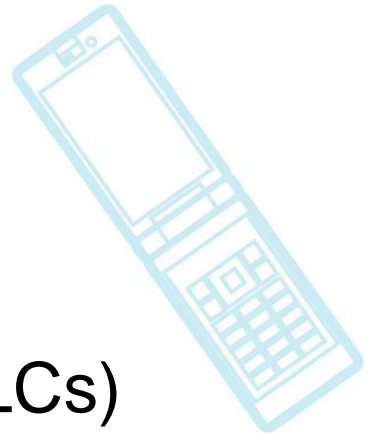
- Autotest
 - Simple framework
 - Not cross-compiler aware?
- LAVA
 - Linaro test framework
- "Kernel Testing Tools and Techniques" BOF by Matt Porter at ELC 2013
- CE workgroup considering reviving test activity for 2013
 - Need input...



CE Workgroup

Build Systems

- Yocto project
 - Lots of talks at ELC (and previous ELCs)
 - Sean Hudson – good introduction tutorial
 - Saul Wold – current status
 - Tutorials now online
- Buildroot still hanging in there





Distributions

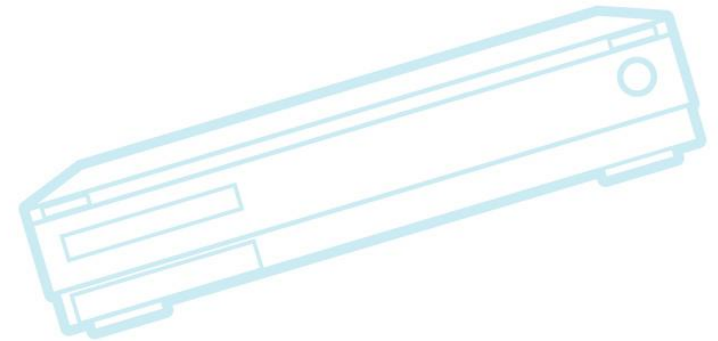
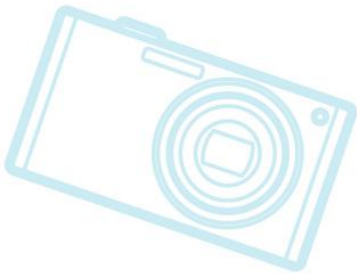
- Tizen – may be a serious competitor in embedded distros
 - Needs to open up a bit more (but it looks like it's happening)
 - Replacing Bada at Samsung
 - Shipping in phones??
- Android use in non-CE embedded
 - Headless android
- Yocto Project = the new in-house distro
- Angstrom
 - Very common on development boards



CE Workgroup

Events

- LinuxCon US
 - September 2012 – New Orleans
- Embedded Linux Conference Europe 2013
 - October 21-23, 2013 – Edinburgh, Scotland
- Embedded Linux Conference 2014
 - April, 2013 – San Jose





CE Workgroup

eLinux wiki

- <http://elinux.org>
 - Web site dedicated to information for embedded Linux developers
 - The wikipedia of embedded linux!
- Hundreds of page covering numerous topic areas: bootup time, realtime, security, power management, flash filesystem, toolchain, editors
- Working on wiki projects:
 - Video transcription project



Miscellaneous

- Kernel Community civility
 - Recent discussion about being nicer to people on LKML
 - Sarah Sharp complained about abusive language and attitude on LKML
 - Some say harshness is needed to maintain quality
 - Others say system works OK as is
 - Will be discussed at kernel summit
- Status of industry = Healthy
 - Over 1.5 billion devices shipped with embedded Linux
 - Still going strong



CE Workgroup

Kernel Versions
Technology Areas
CE Workgroup Projects
Other Stuff
Resources



Resources

- LWN.net
 - <http://lwn.net/>
 - If you are not subscribed, please do so
- Kernel Newbies
 - http://kernelnewbies.org/Linux_3.0
- eLinux wiki - <http://elinux.org/>
 - Especially <http://elinux.org/Events> for slides
- Celinux-dev mailing list
- LinuxCon Japan slides
 - <http://events.linuxfoundation.org/events/linuxcon-japan/program/presentations>



CE Workgroup

Survey for ELCE game

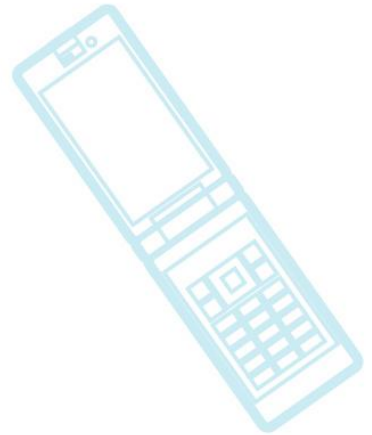
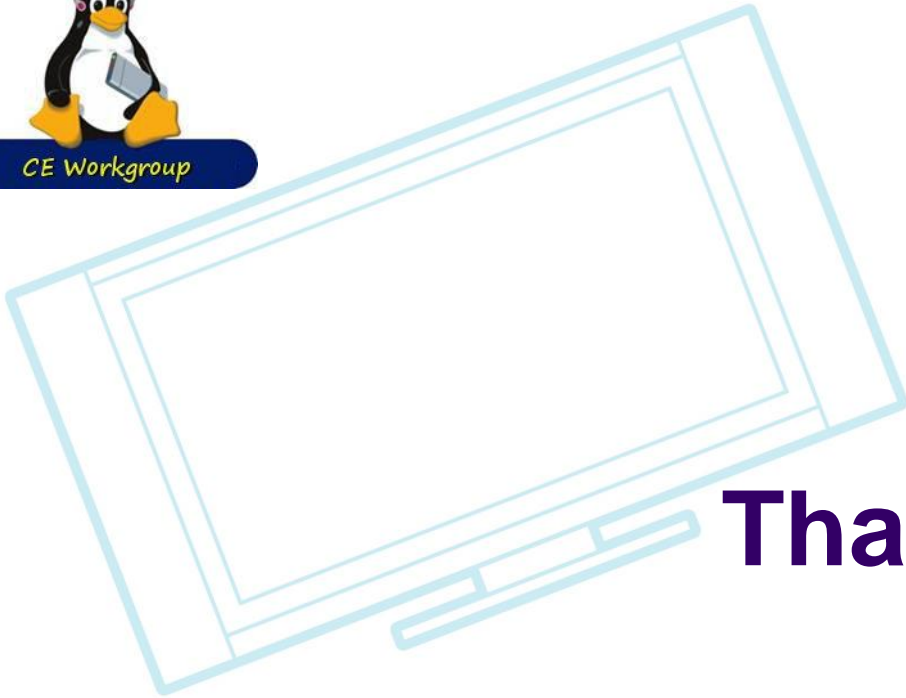
- Please go to:

- <http://embeddedlinuxconference.com/>

and take the survey!



CE Workgroup



Thanks!

