CE Workgroup

Embedded Linux Community Update
June 2020

Tim Bird

Principal Software Engineer, Sony Electronics

Nature of this talk…

- Quick overview of lots of embedded topics
- A springboard for further research
  - If you see something interesting, you have a link or something to search for
- Some overlap with material given previously
  - I may go quickly over some older slides
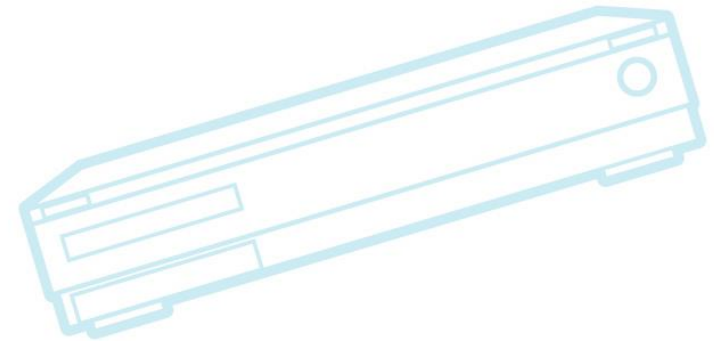- Not comprehensive!
  - Just stuff that I saw

Outline

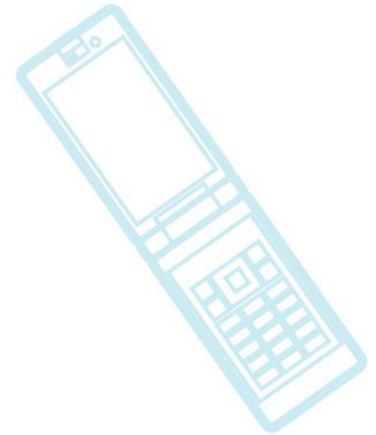Linux Kernel
Technology Areas
Conferences
Industry News
Resources

Outline

Linux Kernel

Technology Areas

Conferences

Industry News

Resources

Kernel Versions

- Linux v5.3 – 15 Sep 2019 – 70 days
- Linux v5.4 – 24 Nov 2019 – 70 days
- Linux v5.5 – 26 Jan 2020 – 63 days
- Linux v5.6 – 29 Mar 2020 – 63 days
- Linux v5.7 – 31 May 2020 – 63 days
- Current kernel = v5.8-rc2
  - Merge window is closed – no new features for 5.8
  - Expect 5.8 kernel on August 2

Kernel Versions

- Linux v5.3 – 15 Sep 2019 – 70 days
- Linux v5.4 – 24 Nov 2019 – 70 days
- Linux v5.5 – 26 Jan 2020 – 63 days
- Linux v5.6 – 29 Mar 2020 – 63 days
- Linux v5.7 – 31 May 2020 – 63 days
- Current kernel = v5.8-rc2
  - Merge window is closed – no new features for 5.8
  - Expect 5.8 kernel on August 2
- *Note: There is no discernable impact from COVID-19*

Linux v5.3

- New pidfd feature – to handle pid reuse
- Scheduler utilization clamping
- 0.0.0/8 IPv4 address support
  - Allows 16 million new IPv4 addresses
- Added CONFIG_PREEMPT_RT config option
  - But not the final code yet
- init_on_alloc and init_on_free boot options
  - pre/post-initialize memory from heap allocations
- See https://kernelnewbies.org/Linux_5.3

Linux v5.4

- EROFS graduated from staging
- exFAT added to staging
- fs-verity feature added
- boot-time entropy fix
  - Fix for commit that was reverted in late 5.3
  - Prevents get_random() from blocking on boot
  - Implementation based on clock jitter, by Linus himself
  - See https://lwn.net/Articles/802360/

Linux v5.5

- ARM64 has full support for ftrace
- MIPS supports kcov – coverage analysis
- KUnit testing framework added
- CPU scheduler's load-balancing algorithm was replaced
  - Follow-on to PELT (Per Entity Load Tracking) work
    - See https://lwn.net/Articles/732021/ for PELT info
- sysctl() system call was removed
  - Use /proc/sys/... instead

Linux v5.6

- WireGuard VPN feature added to kernel
- Work on 2038 issues for ALSA
  - New 64-bit structure for some operations
- Mechanism to disable SELinux at module load time is deprecated (system runtime)
  - Plan is to add a painful delay (increasing with each kernel release) in order to discourage future use
- Bootconfig tool to add super-long command-lines arguments to kernel
- F2FS gained support for compression

Linux v5.7

- Pointer authentication and return-address signing added for ARM64 (explained later)
- Thermal events can effect scheduling
  - scheduler takes into account thermal status and tries to reduce load on hot CPUs
  - https://lwn.net/Articles/788380/
- exFAT fs module in staging was removed
  - Replaced with a new version better integrated into existing filesystem subsystem
    - New version contributed by Samsung
- Kunit results can now be output on debugfs

Linux v5.7 (cont.)

- BPF and PREEMP_RT can now coexist
- LLVM support integrated into kernel build system
  - Can use LLVM=1 on make command line
    - May also need LLVM_IAS=1 as well to use LLVM assembler

Linux v5.8

- Inline encryption for filesystems (more later)
- kgdb can now work with the boot console
  - Allows debugging earlier in the boot process
- A new generic kernel event notification system was added (more on this later)
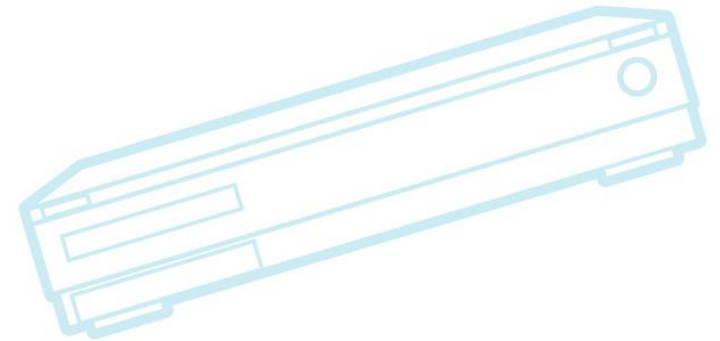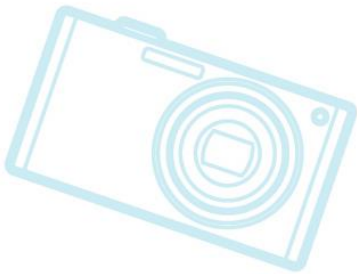
Interesting stats

- What companies are employing maintainers?
  - The top 5 employers supporting maintainers
    - That is, developers with non-author commit signoffs

Employer	Non-author sign-offs	Percent
Red Hat	2560	19.3%
Linaro	1377	10.4%
Intel	986	7.4%
Linux Foundation	878	6.6%
Google	787	5.9%

- Almost half of all patches go through gatekeepers at just 5 companies

- Table data: https://lwn.net/Articles/821813/

Signed tags

- Linus is asking people to use "signed tags"
  - Recipient of pull request can be sure it really came from the person it appears to be from
- Out of 214 subsystem trees, 167 are now using signed tags (78%)
- Is a nice sign of new process being adopted
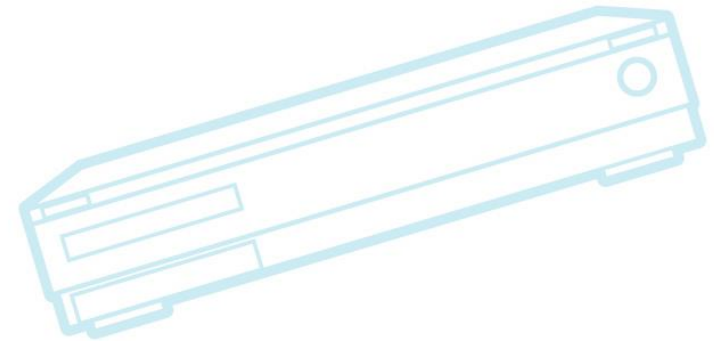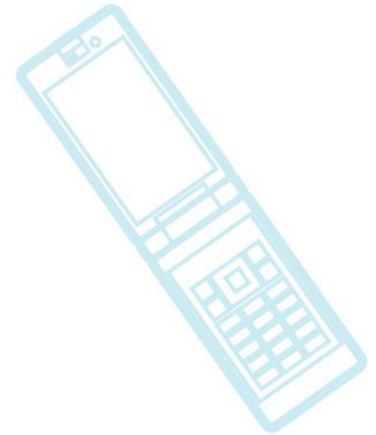- See https://lwn.net/Articles/821813/

Paid development

- 90% of kernel developers are paid by their employer to work on the kernel
  - But there are areas that no one is paid to work on (dedicated)
    - E.g. There is no paid documentation person or team
  - There's still a lot of anxiety about unsupported or under-resourced areas of the kernel
  - Will be a discussion at kernel summit about documentation

Outline

Linux Kernel

Technology Areas

Conferences

Industry News

Resources

Technology Areas

- Audio
- BPF
- Core kernel
- Development
- Documentation
- Filesystems
- Graphics
- Languages

- Memory
- Security
- Testing
- Tracing
- Toolchains
- Tools
- Build Systems

Audio

- Work in ALSA drivers to support new 64-bit structures
  - One benefit is cleanup of fields to avoid year-2038 problems
- New patch proposed for Qualcomm protection domain restart helpers
  - Feature specific to qualcomm SoCs
  - Allows AVS Audio to run in a separate address space
    - Can crash&recover without disrupting other domains

BPF

- Replacing kernel operations structures
  - Ability to replace a structure of function pointers
  - Can load new functions as BPF modules
  - Can create a structure in user-space to override the current in-kernel structure
  - Use a BPF program to replace the in-kernel structure
- Caveats:
  - Must be a structure pre-designated to support this kind of replacement
  - Is only used for TCP congestion-control algorithms (for now)
  - See https://lwn.net/Articles/811631/

BPF (cont.)

- Book Review: BPF Performance Tools
  - Looks like a nice book, and lots of neat tools
  - A sign that BPF should probably be taken seriously
  - See https://lwn.net/Articles/813114/
- Recent discussions about BPF integration with LSM for kernel runtime security instrumentation (KRSI) patch set
  - KRSI needs high performance (wants a special mechanism to replace crypto calls with static jumps)
  - LSM wants general mechanisms (no special cases for BPF)
 - See https://lwn.net/Articles/813261

BPF (even more)

- Some signs that BPF "fever" is waning
  - Discussion by Linux that "user mode helper" feature of BPF should be removed
  - Feature was intended to allow BPF to replace netfliter
  - But:
    - After two years in kernel, there are no users
    - It's very complicated
    - Linus and others question if there is a better way to handle the compatibility

Core kernel

- Year 2038 status
- New generic event notification system
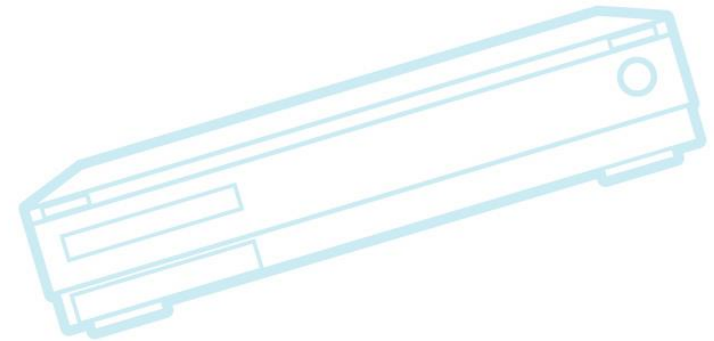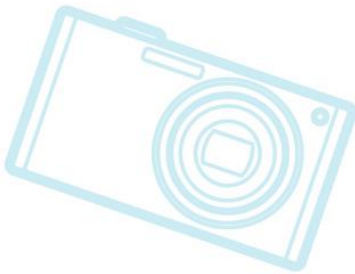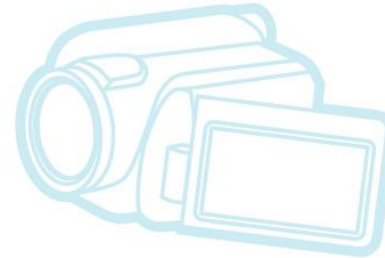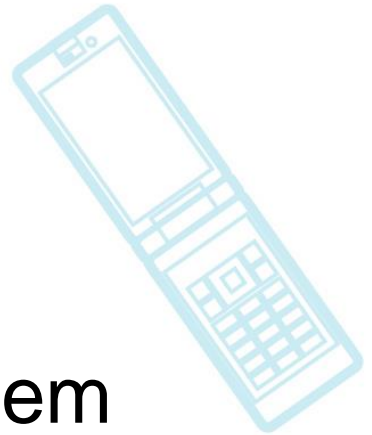
Year-2038 status

- Very good blog entry by Arnd Bergmann
  - https://www.linaro.org/blog/the-end-of-an-era/
  - Describes work from 2014 to 2019 to fix 2038 issues
  - Over 1000 files changed
  - Interns from Outreachy program were used to help with huge effort
  - Some "simple" fixes required other changes that were big
    - leap second handling, time overflow handling, cleaning out unmaintained architectures, etc.
  - Library work (user-space) is in-progress
  - See https://lwn.net/Kernel/Index/#Year_2038_problem

New Event Notification System

- New event notification system (watch_queue?)
  - Is fairly simple
    - Uses a regular Unix Pipe
    - Described by one developer as "a breath of fresh air"
  - Avoids polling by clients
  - Avoids dependency on networking code
    - Avoids netlink ("Friends don't let friends use netlink")
  - May replace other notification systems in future
    - For now, only used for keyring notifications
  - See https://lwn.net/Articles/760714
  - See Documentation/watch_queue.rst (in 5.8 kernel) for latest API details

Development

- Longer lines are OK!!
  - Relaxed warning against lines longer than 80 columns
  - Checkpatch.pl modified so warning occurs at 100 character line length
  - Lines within 80-chars are still preferred
  - Most people have terminal windows that can handle 100 columns now
  - See https://www.phoronix.com/scan.php?page=news_item&px=Linux-Kernel-Deprecates-80-Col

Documentation

- Good article on how to contribute to kernel documentation:
  - https://lwn.net/Articles/810404/
- Specific Tasks:
  - Remove all warnings
    - Specific tips on how to address changes required in kerneldoc messages
  - Add unreferenced kerneldoc info
    - Use scripts/find-unused-docs.sh
  - Fix typos
    - This is a good place to start to learn process
    - Leave some typos for other beginners

Documentation Tasks (cont.)

- Specific Tasks (cont.):
  - Remove, fix or tag outdated documentation
    - Lots of old stuff
  - Organize content into better groups
  - Improve HTML look
    - Improve the style-sheet for HTML output
  - Make rst2pdf tool work with kernel docs
  - Write more documentation
    - Still lots of undocumented areas
- Would be nice to have automated testing to indicated "health status" of kernel docs
- See this video, from kernel recipes 2019:
  - https://www.youtube.com/watch?v=1LuAIUKqKDk

Filesystems

- F2FS gets compression (already mentioned)
- New io_uring system for asynchronous I/O
  - Already have AIO system – this one is better
  - A ring buffer is shared between kernel and user-space
  - User-space can stuff opcodes (commands) into the buffer, and the kernel can execute them, without any syscalls
  - More complex operations are envisioned using BPF
    - That's under heavy discussion
  - See https://lwn.net/Articles/810414/
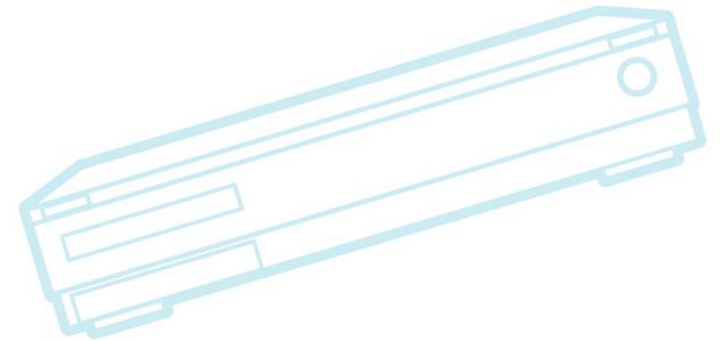
Graphics

- Panfrost driver for ARM Mali GPUs
  - Open Source driver
  - Some support by ARM for development work
  - Support for Mali T720, T820 and T860
  - Support for normal desktops (including Wayland)
  - Currently only support for OpenGL ES <= 2.0
  - Does not support Vulkan yet

- Source: LinuxConfAU talk by Robert Foss
  - https://linux.conf.au/schedule/presentation/68/

Graphics (cont.)

- "Everything Awesome about GPU Drivers"
  - LinuxConf AU talk by Daniel Vetter
  - Good overview of state of Linux kernel graphics stack
  - Source: https://linux.conf.au/schedule/presentation/86/
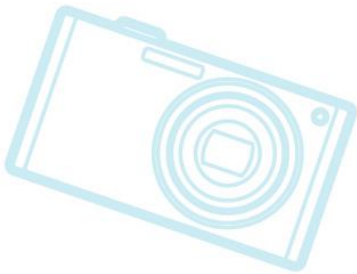
Memory

- Facebook contributing OOMD to systemd
  - Facebook's OOMD is a replacement for Linux OOM handler
    - Faster and more configurable
  - Plan to contribute to systemd so it's automatically integrated into most Linux distributions
  - Probably take about 1 year to show up
  - See https://www.phoronix.com/scan.php?page=news_item&px=Systemd-Facebook-OOMD

Real-Time

- PREMPT_RT patch is closer to being fully upstream
  - Latest patch set for v5.6 was released June 16
    - About 14000 lines of changes
    - Don't know if out-of-tree stuff is required or is just in-progress development or polish
  - Based on CONFIG_PREEMPT_RT being added in v5.3, I was expecting less out-of-tree code
    - But that's not too bad
- BSP and PREEMPT_RT can now coexist:
  - https://lore.kernel.org/bpf/20200214133917.304937432@linutronix.de/

Security

- Control-flow integrity
    - Kernel return address signing
- Kernel Runtime Security Instrumentation
- Wireguard VPN
- Inline encryption for filesystems

Security

- Control-flow integrity for the kernel
  - New patch set for kernel hardening
    - Not mainlined yet – it's in development
  - Reduce ability to redirect code execution
  - Validate function pointer or return address on heap or stack
  - Compiler can collect function pointers into tables that are verified before being used
  - Arm has support for "shadow stacks"
    - Only stores return addresses
    - Shadow stack location is kept secret to prevent tampering
  - Requires Clang (LLVM 10) support for latest features
- See https://lwn.net/Articles/810077/

Kernel return address signing

- New patch set from ARM developer to implement kernel return address signing
  - Added for ARM64 in Linux v5.7
- Uses gcc 7's –msign-return-address feature
  and ARM pointer authentication
  - Cryptographically signs return values (pointers) on the stack
    - In top 24-bits of pointer
  - Validates them before returning

- See https://lwn.net/Articles/804982/

Security (cont.)

- Kernel Runtime Security Instrumentation (KRSI)
  - Allows to mitigate a security attack while it's in progress
  - Provides flexible hook for monitoring and mitigation
  - Implemented as LSM that can run eBPF programs
  - See https://lwn.net/Articles/798157/
  - Recent news:
    - Disagreements between LSM and BPF developers over ways to enhance performance

WireGuard VPN tunnel

- Faster and simpler than Ipsec and OpenVPN
  - Added in kernel v5.6
- Aim is to be as easy to use as SSH
  - Simple generation of public/private key pairs
  - Similar mode of distribution for public keys
- Uses Linux ip commands to set up tunnel
- Allows roaming by both sides of tunnel
- Uses state-of-the-art cryptography
  - High-speed cryptography, suitable for embedded
- Amenable to security audits
  - Due to much simpler code base

Inline encryption for filesystems

- Allows kernel to offload encryption and decryption to the storage device
- Normally, encrypted filesystem has lots of overhead
  - On main CPU or in crypto hardware and busses as data is copied multiple times
- This is faster
  - Kernel manages setup, then storage device handles cryptography
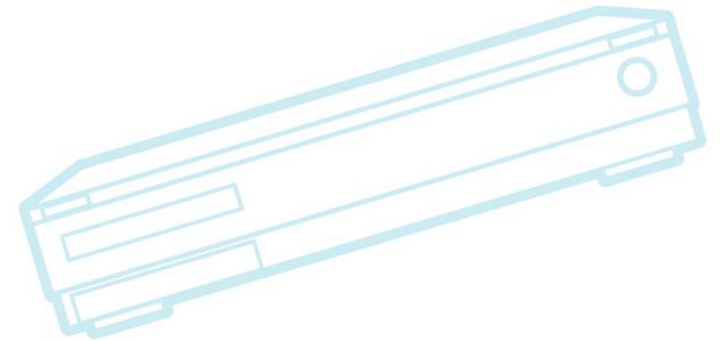- Kernel can also validate that encryption worked as expected
- See https://lwn.net/Articles/797309/

Tracing

- New Bootconfig system
  - Extra boot configuration
  - Allows passing a large set of options to the kernel during boot
    - Was not a good fit for device tree
  - Passes a tree-structured key-value list
  - Data is loaded with initrd
  - Used primarily to pass kernel command line items for ftrace and early tracing
  - Mainlined in 5.6
  - See https://lwn.net/Articles/806002/

Tracing (cont.)

- Babeltrace 2.0 released
  - New version of the babeltrace trace manipulation toolkit
  - For viewing, converting, transforming, and analyzing traces
  - See https://lwn.net/Articles/810395/
  - And https://babeltrace.org
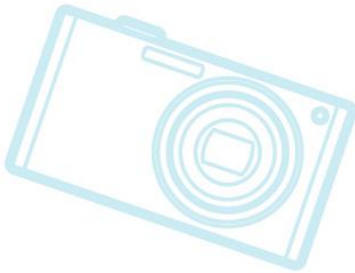
Toolchains

- Static analysis framework for GCC
  - Analyses intermediate representation of code
  - Uses a plugin architecture, for now
  - 2 facility categories for now:
    - memory allocation errors
    - file handling errors
  - Adds meta-data to the diagnostic message
    - Lots of information about the problem
    - e.g. Can indicate the Common Weakness Enumeration (CWE) entry for a problem
  - See https://lwn.net/Articles/806099/

Tools

- Continuing efforts to create tools for bridging gap between git and e-mail
  - New tool: get-lore-mbox
    - Can download a thread related to message id into a local .mbox format (mail archive)
    - Can download just the patch in the thread
    - Can automatically add tags (from different e-mails in the thread) to the patch
      - E.g. reported-by, acked-by, tested-by, etc.
    - Very handy for retrieving patches that were mangled by your email system
    - See https://lwn.net/Articles/811528/
- More work is going on to make additional tools
  - https://github.com/gitgitgadget/gitgitgadget
    - Converts from github pull request to e-mail patches
- Use kernel "workflow" mailing list to watch discussions

Miscellaneous

- Long term support upgraded to 6 years for two LTS kernels!!
  - 4.19 and 5.4 will be supported for six years instead of two years

kernel.org/releases.html#longterm

Longterm release kernels

Version	Maintainer	Released	Projected EOL
5.4	Greg Kroah-Hartman & Sasha Levin	2019-11-24	Dec, 2025
4.19	Greg Kroah-Hartman & Sasha Levin	2018-10-22	Dec, 2024
4.14	Greg Kroah-Hartman & Sasha Levin	2017-11-12	Jan, 2024
4.9	Greg Kroah-Hartman & Sasha Levin	2016-12-11	Jan, 2023
4.4	Greg Kroah-Hartman & Sasha Levin	2016-01-10	Feb, 2022
3.16	Ben Hutchings	2014-08-03	Jun, 2020

- See https://kernel.org/releases.html#longterm

Outline

Linux Kernel
Technology Areas
Conferences
Industry News
Resources
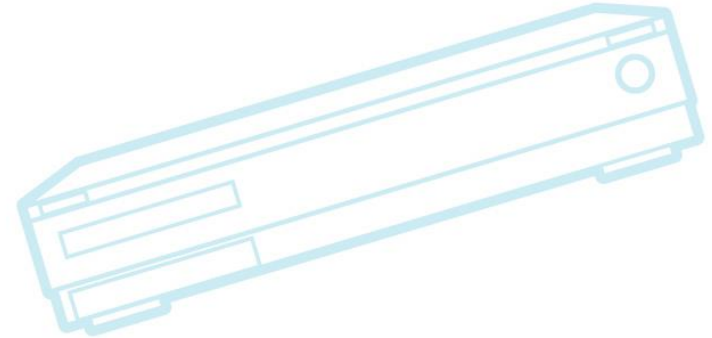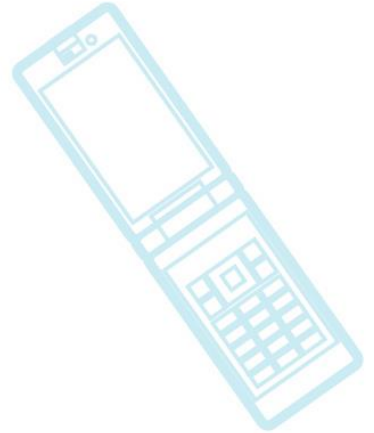
Conferences – 2020 (plan)

- Embedded Linux Conference 2020
  - June 22-24, Austin, Texas, USA
- Linux Plumbers
  - August 24-28, Halifax, Canada
- Open Source Summit Japan
  - September 15-16, Tokyo, Japan (??)
  - Web site not updated with any change yet
- ELC Europe 2020
  - October 26-28, Dublin, Ireland (??)

Conferences - 2020

- Embedded Linux Conference 2020
  - ~~June 22-24, Austin, Texas, USA~~
  - June 29-July 1, online virtual event
- Linux Plumbers
  - ~~August 24-28, Halifax, Canada~~
  - August 24-28, online virtual event
- Open Source Summit Japan
  - September 15-16, Tokyo, Japan (??)
  - Web site not updated with any change yet
- ELC Europe 2020
  - October 26-28, Dublin, Ireland (??)

COVID-19 issues?

- ~~Will conferences get canceled?~~
  - When will conferences no longer be cancelled?
  - Unknown at this point
- ELCE has not been canceled yet
  - But things are not looking good
- Some positives:
  - New virtual events allow many more people to join from around the world
  - There is a Slack channel for OSSNA/ELC attendees
    - Many attendees from many different locations who say this is their first conference

Embedded Linux Conference 2020

- Only $50 to join virtually
- Not at a great time for attendees from Asia
- But sessions are recorded
  - Attendees can watch them any time
  - Can contact speakers on Slack during event
- Interesting issues with labor practices in Japan
  - Sony has a policy to avoid requiring people to work in middle of night on consecutive days
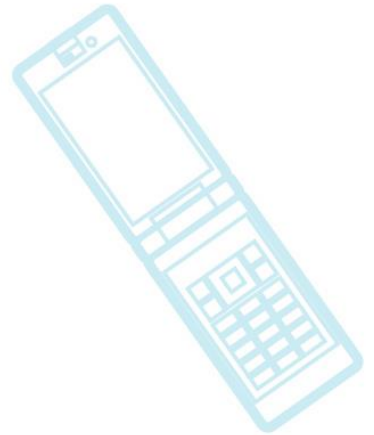   - Is OK if they physically travel, but not if they attend remote online event

Outline

Linux Kernel

Technology Areas

Conferences

Industry News

Resources

Industry News

- Trade associations
- Technology to watch
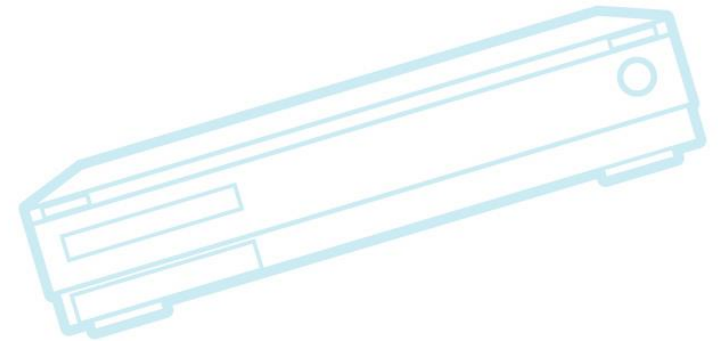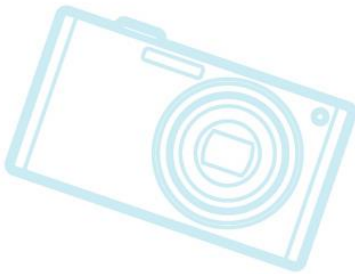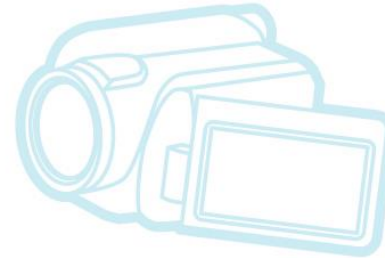
Trade associations

- Linux Foundation
  - Joint Development Foundation
    - Helps with migration from OSS project or spec, to international standard
    - OpenChain is first project
  - MLflow – machine learning platform
    - Used for tracking ML assets, packaging, and model sharing and collaboration
  - Watch for lots of announcements next week at Open Source Summit North America
    - I'm on the LF board, and I can't remember what else has already been announced and what hasn't

Linaro

- Celebrated it's 10th anniversary on June 18th
  - https://www.linaro.org/blog/linaro-a-decade-of-development/
- Has done a LOT of great work over the years
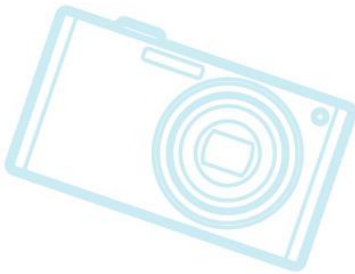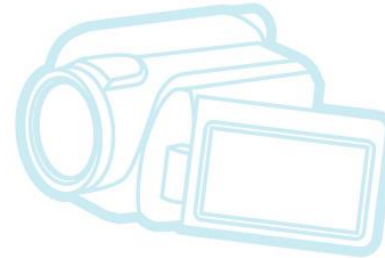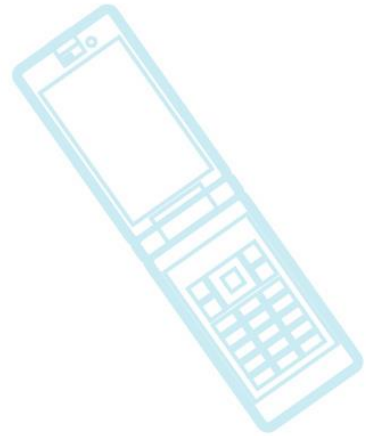 - Can see from contribution stats and maintainer stats that they have had a real impact

Technology to watch

- Microsoft IPE patchset
- Systemd homed
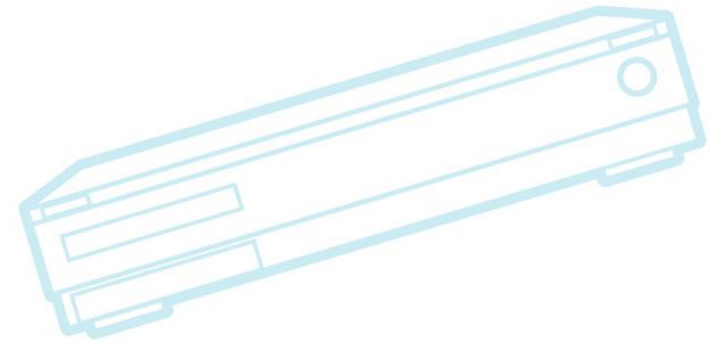
Microsoft IPE patchset

- IPE (Integrity Policy Enforcement) system
  - Targeted at embedded systems
  - Implemented as a Linux security module (LSM)
  - Developer provides list of checks that a binary must pass, in order for system to execute it
    - Ex: The binary has a valid dm-verity signature
  - Not in mainline yet
  - See https://www.zdnet.com/article/microsoft-announces-ipe-a-new-code-integrity-feature-for-linux/
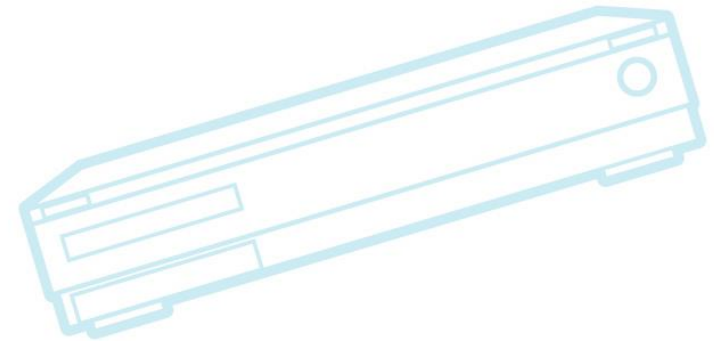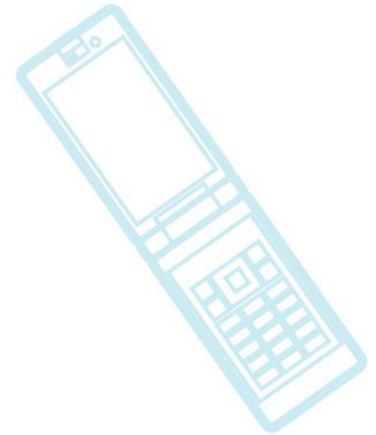
system homed

- New system to manage user home directories
  - Puts user identity in new JSON file, that is co-located with home directory
    - independent of /etc/passwd and /etc/shadow
  - Makes home directory portable between systems
  - Home dir can be encrypted
  - Adds new 'homectl' command for managing users and home directories
- https://www.techrepublic.com/article/linux-home-directory-management-is-about-to-undergo-major-change/

Outline

Linux Kernel

Technology Areas

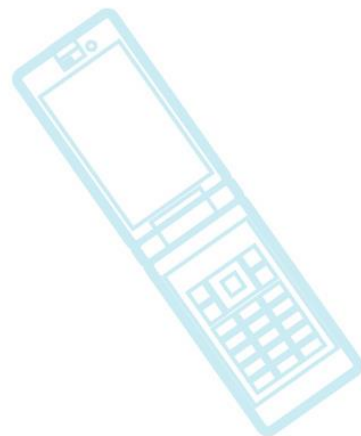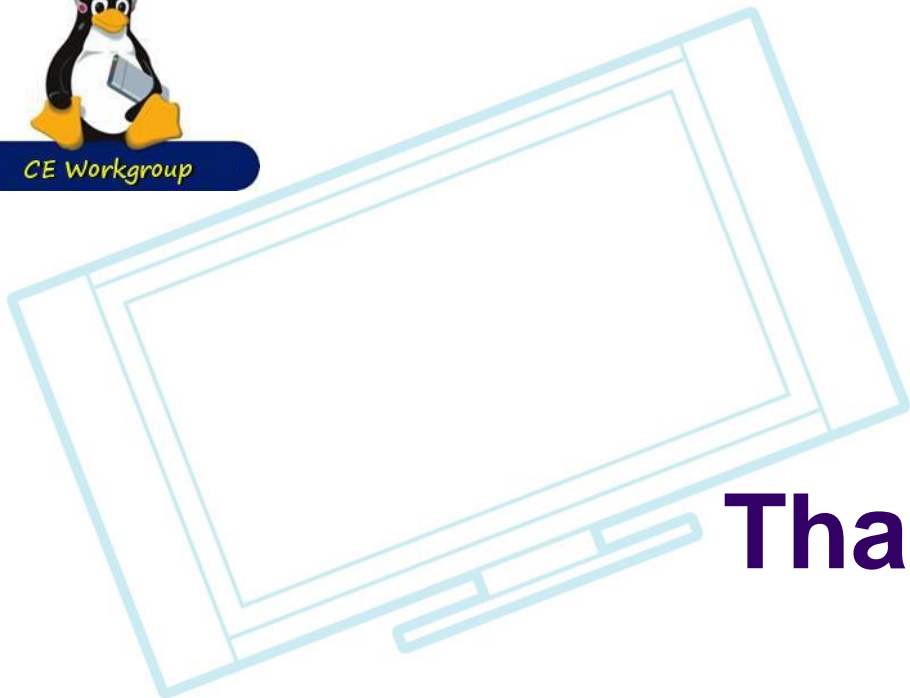Conferences

Industry News

Resources

Resources

- LWN.net – https://lwn.net
  - If you are not subscribed, please do so
  - Some content is delayed by 2 weeks for non-subscribers (some links in this presentation)
- Linux Gizmos – https://linuxgizmos.com
- Phoronix - https://www.phoronix.com/
- Google

Thanks!