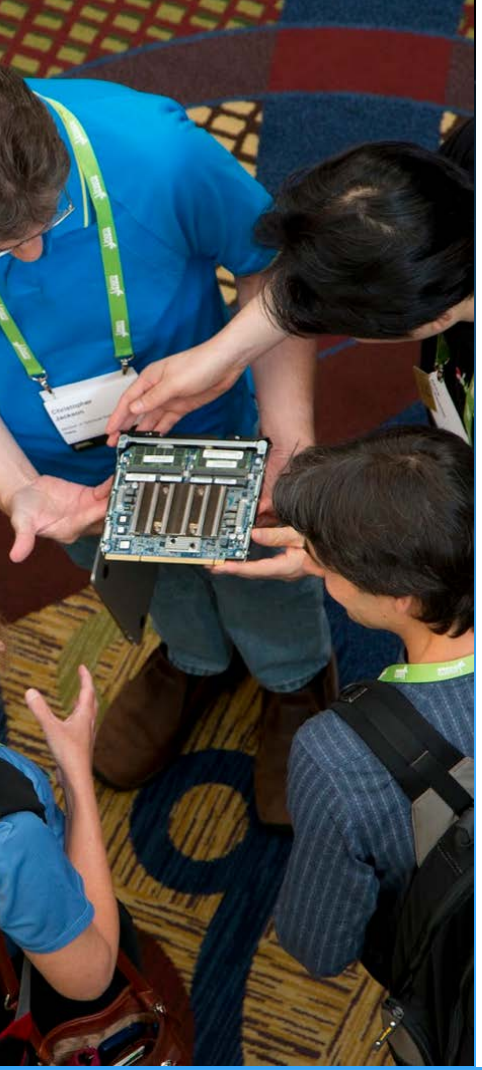
LAS16-205: The State of AOSP Common android-4.4 Kernel

Amit Pundir



Session Layout

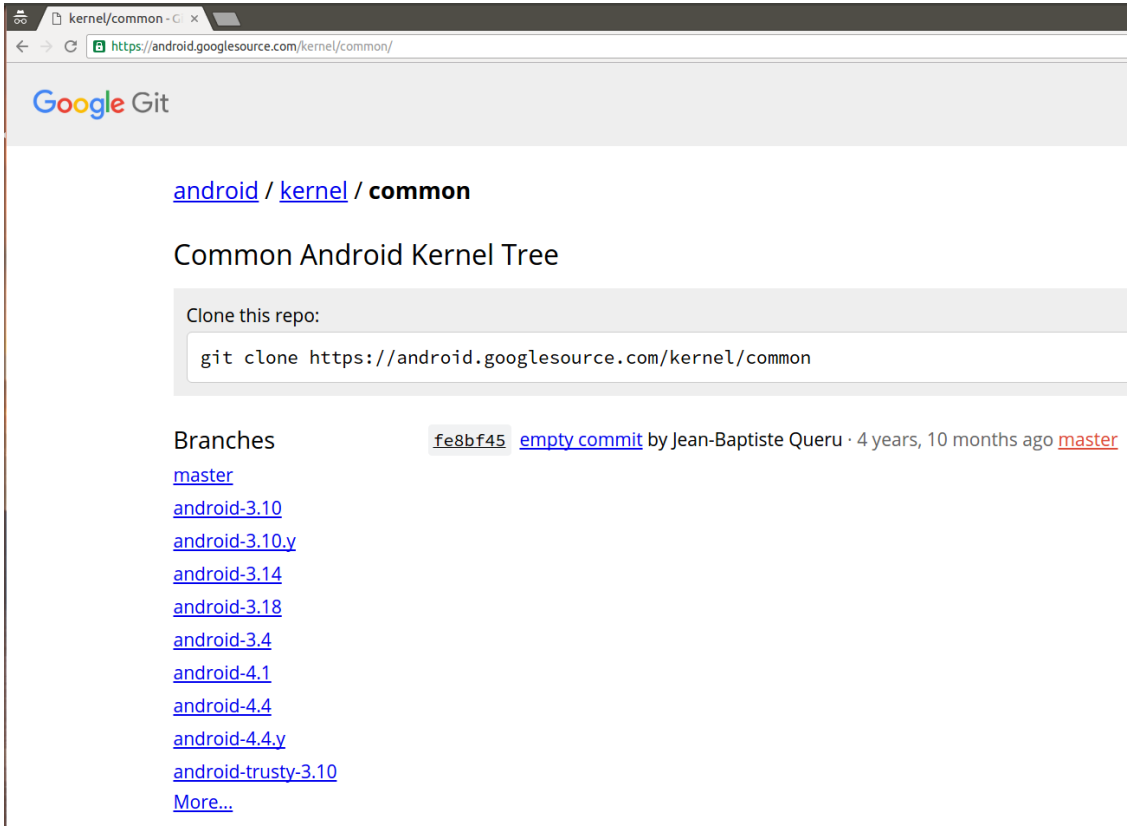
- AOSP Common Kernel
  - Rationale / Brief introduction and Evolution
- Linux v4.4 vs android-4.4
  - Git diff stats - v4.4..android-4.4
  - Brief introduction of functionality and upstream status
- linaro-android-llct
  - Git diff stats - v4.8-rc5..linaro-android-4.8
- Android kernel patches in Staging
  - Brief update on ongoing staging activities

AOSP common kernel

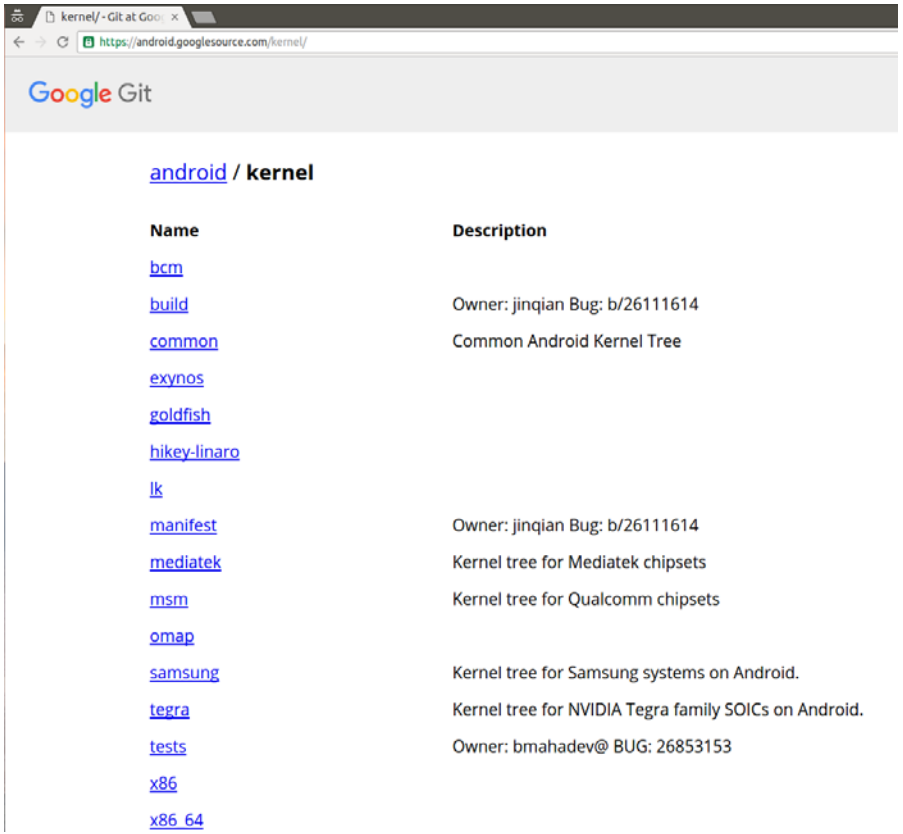
AOSP Common Kernel

- Features tailored for Android needs e.g. Interactive Gov, UID routing
- Features rejected by upstream owing to implementation concerns e.g. MTP/PTP, Paranoid networking
- Testbed for features to be pushed upstream overtime e.g. EAS
- Features which are available mainline but Android still using earlier in-house implementations e.g. PPPoPNS, PPPoLAC
- Vendor/OEM features which can be useful for others as well e.g. sdcard_fs

Common Android Kernel Tree

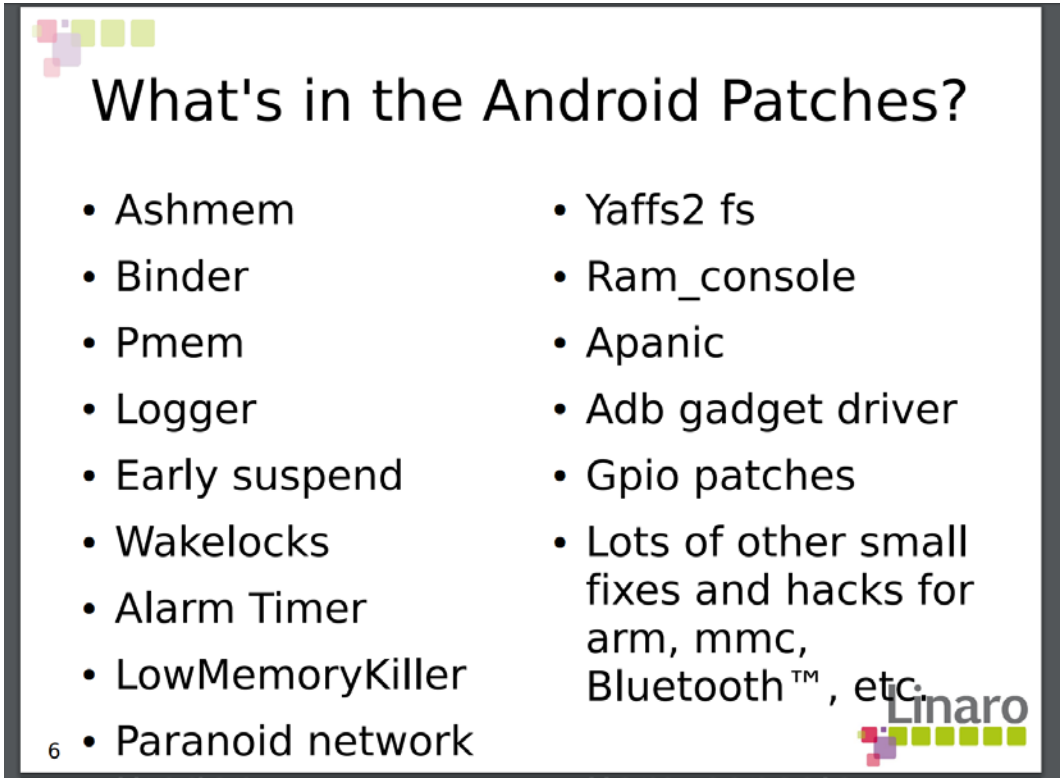


Kernel Tree for Vendor Chipsets



AOSP Patchset Evolution

- Android patchset snapshot taken from one of John's talk at ELC 2011.



Linux v4.4 vs android-4.4

v4.4..android-4.4 git diff stats

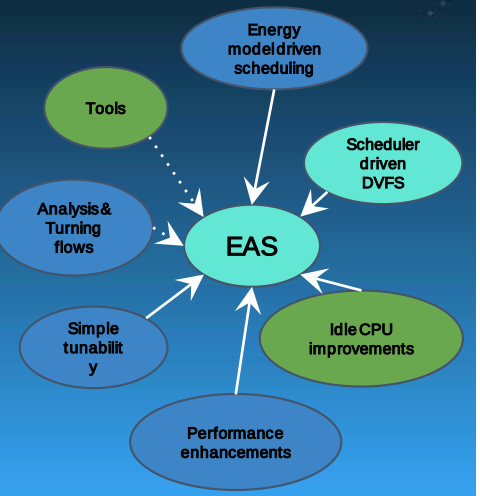
- 523 files changed, 46634 insertions(+), 1634 deletions(-)
  - Including UPSTREAM/ BACKPORT / FROMLIST fixes
- ~14% Networking
- ~09% Energy Aware Scheduling
- ~09% USB Gadgets
- ~09% Atomic Display Framework
- ~08% Verity Boot
- ~08% Sdcard FS
- ~04% FIQ debugger
- ~04% Input
- ~03% Cpufreq
- Rest: Documentation, include, kernel, arch, mm..

Networking

- Paranoid networking
  - Restrict network access to certain group of users
  - Largely perceived as Android hacks with hardcoded AIDs mapped to userspace groupids
    - Expect this filtering to be based on network namespaces
    - It will require a fair bit of userland changes and that is unlikely to happen.
  - “Network filtering for control groups” may be the answer?
- UID based routing
  - Route packets differently based on the user ID that owns the socket
    - Allow userspace to configure routing rules based on UID ranges.
  - Upstream options include netfilters / iptables or even namespaces.
    - Too expensive to that using iptables.
  - Android devs seem to be working on a revised patchset for upstream submission.

Networking

- Netfilter: qtaguid, quota2, idletimer
  - Data usage tracking &limiting
    - qtaguid and quota2 modules to do per uid tracking and accounting
    - May be replaced with NFQUEUE?
  - IDLETIMER
    - Help ConnectivityService deal with quiet interfaces
      - Track and send uevents when interface becomes active again
    - Last upstream submission got mixed reviews
- PPP: PPPoLAC and PPPoPNS
  - For legacy VPN support
  - May be switch to mainline PPTP and PPPoL2TP interfaces
    - Userspace need to be changed
    - GPL pppd plugins available but that may not work out for Android



Energy Aware Scheduling

- To make Linux fully aware of the power/performance capabilities of the CPUs and optimize energy consumption.
  - EAS core
  - SchedFreq (cpufreq gov)
    - Scheduler-aware CPU frequency management.
  - SchedTune (boosting mechanism)
    - Scheduler centric power-performance tunable
  - WALT (PELT load-tracking replacement)
    - Window Assisted Load Tracking to track CPU utilization
- Under active development and testing phase.
  - Patches/Discussion can be tracked on linux-arm.git

USB Gadgets

- ConfigFS USB gadget patches
  - Android Functions
    - MTP/PTP
      - In-kernel drivers are no go, advised to use functionFS instead
    - Audio Source and Accessory drivers
  - Android device class (/sys/class/android_usb/android0) interface
    - Should read usb state changes from /sys/class/udc/*/state instead
- RNDIS fixes
  - Data aggregation (multi packet) support
  - Few misc tethering fixes

Atomic Display Framework

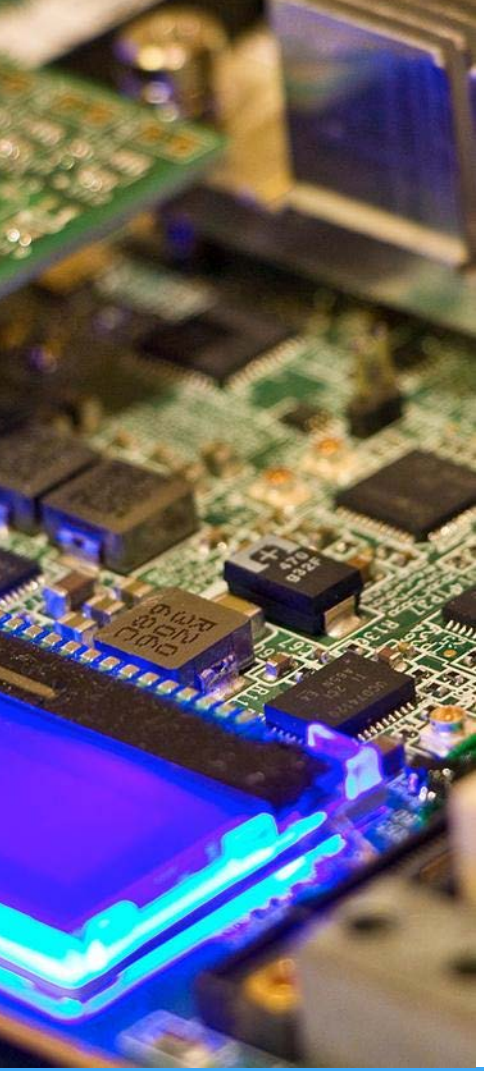
- Set of display buffers and configurations to be atomically updated
- Sits between Android's hwcomposer HAL and kernel driver
- On its way out in favor of DRM/KMS enhanced with Atomic mode setting and Explicit fence support.
  - Will not be used in next android-lts kernel.

Verified Boot

- Block-level integrity protection and forward error correction for system and vendor/oem read-only partitions
- dm-verity
  - Mostly upstream cherry-picks / backports
- dm-verity-fec
  - Upstream backport
  - limit error correction recursion and add sysfs attribute for stats
- dm-android-verity
  - Setup verity root A/B or seamless update support

Sdcard FS

- Sdcard file system wrapper derived from wrapfs
  - Implement FAT32 emulation layer inside kernel
- Fuse alternative for emulated storage in AOSP
  - Emulated storage access through FUSE add lot of performance overhead
  - Sdcardfs is found to perform better. Used in Samsung phones for a while now
- Didn't find any upstream RFC for this version
  - Found one discussion thread on wrapfs Wrapfs: a stackable file system dated 2001.
  - AID_*(Android static uids) usage in sdcardfs won't help in upstreaming in current state

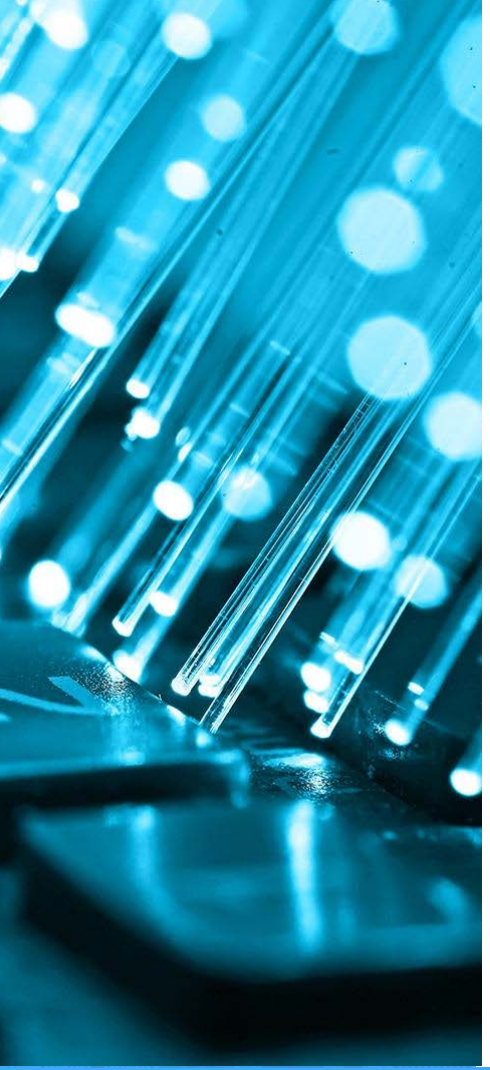


FIQ Debugger

- Low level kernel debugger for ARM
- Uses Fast IRQ (FIQ) interface for debugging
- Parts of it already integrated with upstream KDB implementation
- In Progress:
  - Fiq debugger for ARMv8
  - Extend NMI watchdog
  - IPI FIQ for ARM

Input

- Keychord and Keyreset(Keycombo) driver
  - Driver to handle different key press combinations
  - Chunks of it already upstreamed as part of SYSRQ driver
- Generic GPIO input support
  - Supports keyboard matrices
  - Direct inputs/outputs
  - Axes connected to gpios



Interactive Governor

- Select operating frequency of the processor depending on the user interaction
- Aggressive on-demand governor
- NACKed in favor of scheduler-driven cpufreq selection
- Revised implementation using schedfreq is submitted for RFC recently

Miscellaneous

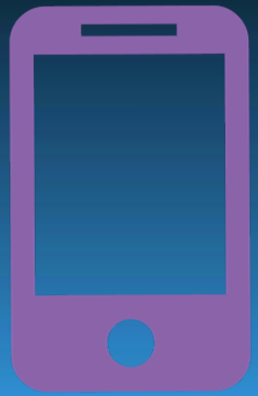
- MM
  - Private Anonymous memory
    - Anon memory tagged/named by userspace to track and debug physical memory usage
- Cgroup
  - Android hooks/checks to move a task across control groups
- Timerslack_PID
  - set/change the scheduling of the background threads by changing their timerslack value using PR_SET_TIMERSLACK_PID
  - Already upstreamed as /proc/<pid>/timerslack_ns interface
- Power
  - wakeup reason logging, report wakeup source
- UID_CPUTIME
  - Per UID based cpu time statistics exported to /proc
  - Used by BatteryStats service

Miscellaneous

- ARCH
  - Image.gz/dtb, Image-dtb support
  - Low level printk, CONFIG_CMDLINE_EXTEND support
- Tracing
  - MMC, GPU, Min/Max cpufreq, lowmemory kill events tracing
- MMC
  - Sysfs interface for IO latency histogram
  - Additional retries on SD detection
  - Embedded sdio support and other sdio fixes
- Dual Role USB Phys sysfs interface
  - Sysfs interface to track and change the state of dualrole usb phys

Miscellaneous

- Memory State Time driver
  - New memory_state_time driver tracks time spent in different DDR frequency and bandwidth states
- Android Pipe &goldfish emulator etc etc
  - Android emulator support
  - Already upstreamed
- SELinux hooks
  - For compatibility with Android Muserspace and whitelisting tracefs filesystem
- Android config fragments
  - Mainline kernel config fragments already upstreamed

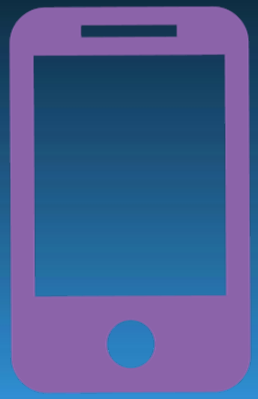


LMG
Mobile

kernel/common/android-4.4 cleanup

- Patches dropped from android-4.4
  - Switch Class &GPIO dropped in favor of ExtCon
  - UID_STAT &Network activity stats driver dropped in favor of qtaguid and quota2
  - Armv6 DCC tty driver in favor of upstream DCC driver
  - Removed few duplicate HID, SElinux, Power, Debug fixes etc
  - Removed few obsolete Android composite gadget, MMC fixes etc

linaro-android-llct



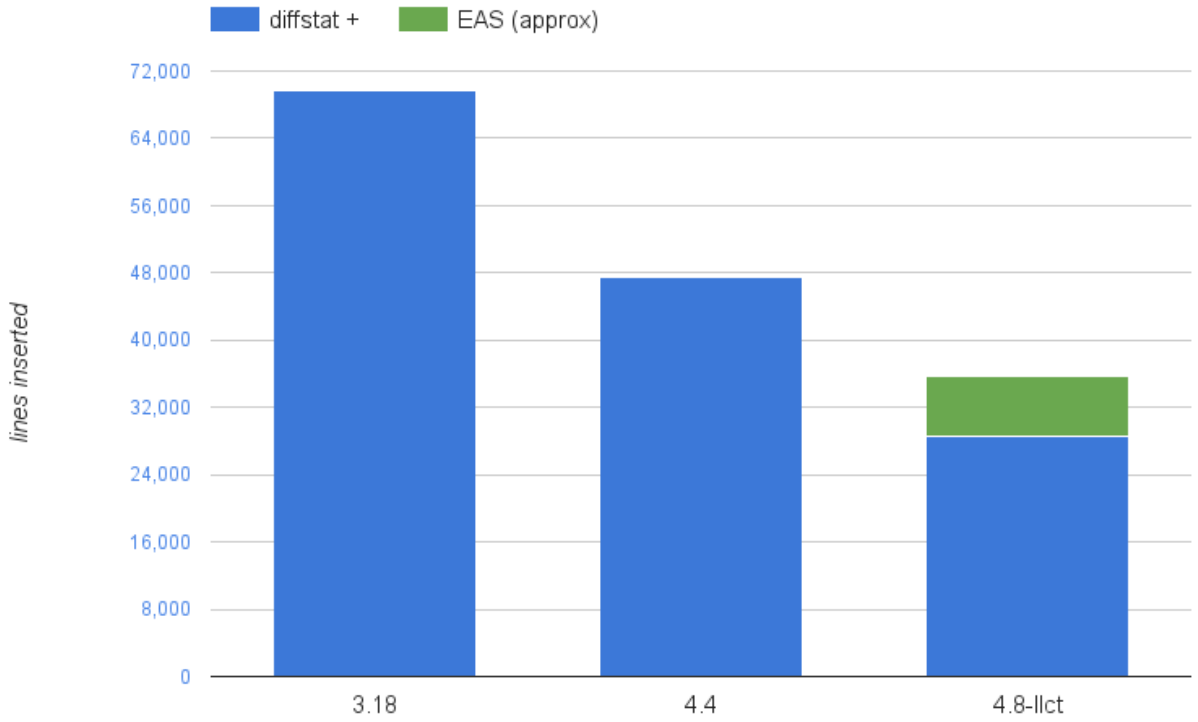
LMG
Mobile

linaro-android-llct

- Android patchset / features rebased to latest Linux tag.
- Find / Report / Fix Android regressions or ABI breakages in upstream kernel.
- Testbed for patches which we are trying to submit upstream for RFC.
- Can be used as a reference experimental preview to what we think the next AOSP common branch might look like.
  - Cherry-picked and submitted fixes from linaro-android-3.18 to AOSP/experimental/android-3.18
  - Shared linaro-android-4.1 branch with AOSP for reference
  - AOSP picked linaro-android-4.4 as it is for experimental/android-4.4

v4.8-rc5..linaro-android-llct git diff stats

- 337 files changed, 28515 insertions(+), 478 deletions(-)
- ~22% Networking
- ~15% USB Gadgets
- ~07% FIQ Debugger
- ~13% Sdcard FS
- ~07% Input
- ~05% Cpufreq
- ~04% Verity Boot


diffstat +
EAS (approx)
lines inserted
72,000
64,000
56,000
48,000
40,000
32,000
24,000
16,000
8,000
0
3.18
4.4
4.8-llct

android-4.4..llct git diff stats

- Not easy to compare the absolute diff stat comparison in previous slide
  - More UPSTREAM/ BACKPORT/ FROMLIST fixes in older versions
  - No EAS patches in llct yet
- Few patches / features dropped from AOSP in favor of upstream features.
  - Atomic Display Framework is going out in favor of DRM/KMS.
    - So now is the time to redo your display drivers if you are still using ADF
- Few patches / features pushed upstream.
  - Timerslack_ns, MMC, Config fragments and Tethering fixes.

AOSP patches in Staging

Android Patches in Staging

- AShMem
  - Anonymous SHared MEMory
    - Unpin/Discard memory units under memory pressure
  - “Memory unpinning” using “volatile-ranges” was attempted
    - Turned out to be a bigger ordeal than planned just to support a file-desc based memory sharing model
  - Memfd is a more reasonable replacement for Ashmem
    - Adding “memory unpinning” in Memfd is in TODO

- LowMemoryKiller
  - Kill applications under memory constraints
  - More aggressive version of out-of-memory killer
  - “mempressure notifier” as LMK replacement?
    - Send memory pressure notifications to userspace
    - Probably still being worked on / used internally by Android developers

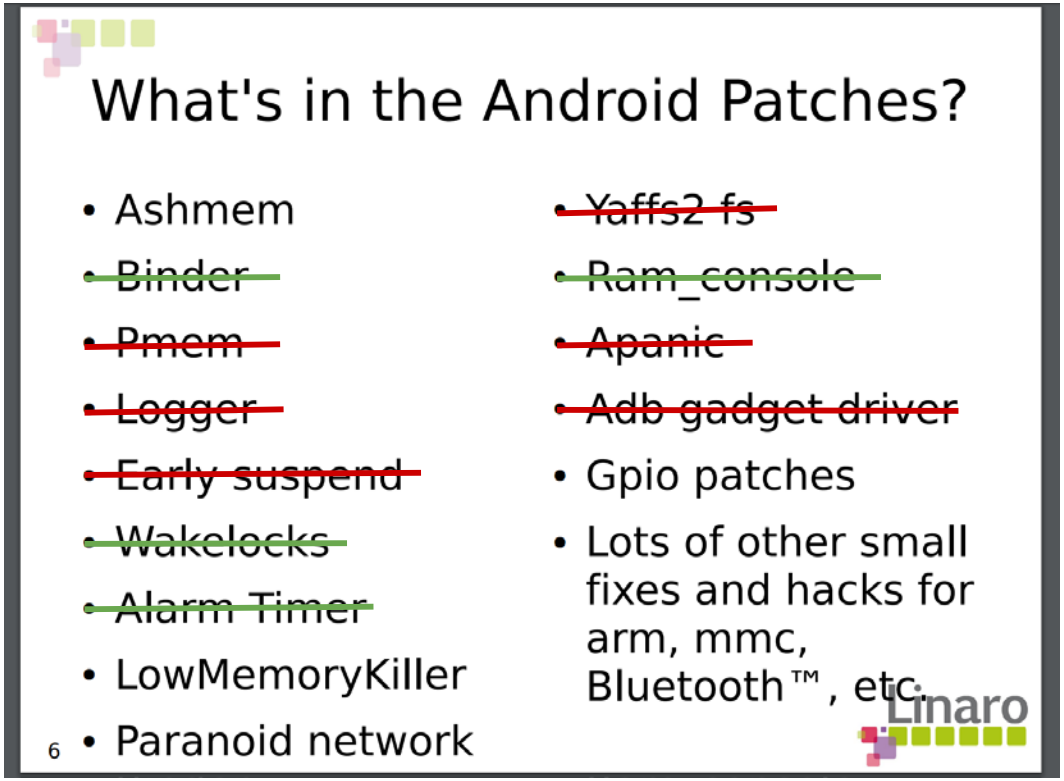
Android Patches in Staging

- Sync / Sw_Sync
  - APIs to synchronise buffers in complicated DMA pipelines
  - Sync In process of being de-staged from Staging to dma_buf
    - Sync points deprecated to use dma_buf fences.
  - Sw_sync to be moved as debugfs interface

- ION
  - Managing different pools of memory with different constraints and sharing these between devices.
  - In Progress
    - Device tree bindings to expose cma regions as ion heaps
    - New IOCTL interface to support caching using dma_buf APIs
    - Better discovery of available heap IDs

AOSP Patchset Evolution

- A quick look at where does Android patchset from 5 years ago stand today.



References

- LWN: The LPC Android microconference 2015
- LWN: The volatile volatile ranges patch set
- LWN: The mempressure control group proposal
- [YouTube] how the Linux networking stack is made to work on Android devices
- [LPC 2013] Android Netfilter Changes
- LWN: Network filtering for control groups
- [android-paltform] Will be FUSE removed?
- [developer.arm.com] Energy Aware Scheduling
- [linux-usb] About Data Aggregation with RNDIS and u_ether Driver
- Linaro Wiki: Make Android use upstream PPP VPN code
- OpenIoT &ELC 2016: Atomic Display Support in Upstream - Daniel Vetter, Intel
- Linaro Connect: SFO15-311: ConfigFS Gadget - An Introduction
- Android Verified Boot
- Strictly Enforced Verified Boot with Error Correction

Thank You

#LAS16

For further information: www.linaro.org

LAS16 keynotes and videos on: connect.linaro.org