Energy Awareness :: The Next Step

● Who we are

● Where do we come from?

● Where are we at?

● Where do we want to be?
Energy Awareness :: The Past

On October 2007 the Completely Fair Scheduler by Ingo Molnár was released (2.6.23).

The main goals of CFS were:

- To maximize overall CPU utilization and interactive performance.
- To balance runnable tasks across available resources (cores) and minimize the idling of cores.

At the time, CFS had no notion of energy cost.
Energy Awareness :: The Past

On October 2008 the HTC Dream with a 528 MHz Qualcomm MSM7201A (ARM11) processor running Android 1.6* was released.
- The Scheduler had only one core/queue to manage.

On February 2011 the Samsung Galaxy S2 with a 1.2 GHz dual core Exynos (ARM Cortex A9) processor running Android 2.3 was released.
- The scheduler had to balance runnable tasks across two identical cores.
Energy Awareness :: The Past

On April 2013 the Samsung Galaxy S4 with an Exynos 5 Octa 5410 CPU – Quad-core 1.6 GHz Cortex-A15 & quad-core 1.2 GHz Cortex A7 – running Android 4.2.2 was introduced.

- The scheduler could only see one (big.LITTLE) cluster at a time depending on the load.

IKS was a further improvement of Cluster Migration where each pair of a 'Big' & 'Little' core (virtual core) was visible to the scheduler.
Energy Awareness :: The Past

Heterogeneous multi-processing (HMP) was the latest improvement to big.LITTLE where all cores are available to the scheduler at the same time.

The asymmetric SMP design posed unique problems to the scheduler:

- Still tried to balance runnable tasks across available cores, without knowing the big.LITTLE performance difference.
- Still didn’t know the energetic cost of it's decision.

This led to some performance issue and wasted energy.
Energy Awareness :: The Past

To further complicate matters, the scheduler, cpufreq, cpuidle & hotplug decision centres were:
- Periodically running to gauge the (past) cpu utilization.
- Independently deciding the best course of action.
- Not taking into account the big picture (perf + power).
- Fighting over control of the platform.
- Not communicating among them.
- Not sharing a common strategy.
- Undoing each other's work.
- Working in total isolation.
- Reactive in nature.
Energy Awareness :: The Present

In 2013 with the advent of aSMP SoCs there was a reckoning to further improve the scheduler’s decision making by:

- Fully integrating CPUFREQ & CPUIDLE into the scheduler itself.
- Creating a new scale invariance load tracking (0…1024).
- Creating a new Power Model for the most energy efficient decisions.
Energy Awareness :: The Present

The goals of EAS is to allow sophisticated load balancing heuristics by:

- Tuning the scheduler to optimise for energy savings without affecting performance.
- Correctly scheduling on big.LITTLE (aSMP) systems.
- Improving power management on SMP systems.
- Providing proper mechanisms for thermal management.
- Tightly integrating all important decision making centres (CPUFREQ, CPUIDLE) into the scheduler.
- Making the scheduler aware of the idle state of the CPU.
- Revamp how the CPU capacity is handled by the scheduler with a scale invariance load tracking.
Energy Awareness :: The Present

To come up with an EAS Power Model the engineer has to select a random SoC and go through the painstaking process of getting the data by hand in the lab.

This kind of handcrafted data poses the following problems:
- Creates a very simplistic & error prone Power Model based on static measurement of a random SoC.
- Power Model does not adapt to the changing SoC circumstances like frequency & temperature.
- Doesn’t take into account static power, idle discharge, L2 Cache among others.
Energy Awareness :: The Future

“Accurate Real Time Information is POWER”

We need to start building a new generation of **self aware systems** that automagically adapt to the unique characteristics of the platform they're running on without any kind of maintenance or “magic knob” tuning from engineering or users.

Furthermore we need to take away the complexity of creating a proper Power Model so it can be useful to the EAS subsystem.
Energy Awareness :: The Future

The AURORA Chip (Phase #1)

The AURORA chip is a hardware tailored component that monitors & reports the power readings of all key SoC systems to it’s kernel part that in turn will feed EAS with the most accurate Power Model in real time.

This uniquely adapted & accurate Power Model will take away the complexity and enable EAS to make the most adequate energy wise decisions without compromising performance.
Energy Awareness :: The Future

Key benefits of the AURORA Chip (Phase #1):

- The most accurate Power Model precisely adapted to the specifics of the platform it is running on.
- Factors in key Power Model data such as L2 Cache (freq/temp), per Core Frequency/Temperature/Voltage combination and Static Power (C-states & Idle).
- Adapts to the unique characteristics of every single SoC that comes out of the FAB.
- Takes away all the manual & intensive engineering lab labour of coming up with a (generic) Power Model.
Energy Awareness :: The Future

The AURORA Chip (Phase #2):

The next iteration of the AURORA Chip will take on a more prominent role by offloading some of the calculations that the scheduler has to periodically do among other things.

This in turn would usher in a new way of computing where all the strenuous demands will be met in the most advanced independent and self optimizing way possible.
Thank You

Carlos Lloret
[carlos.lloret@iondroid.com | @CeKMTL]
#BUD17
For further information: www.linaro.org
BUD17 keynotes and videos on: connect.linaro.org