

# big.LITTLE mini-summit

Amit Kucheria, Power Management Tech Lead

LCE 2012, Copenhagen



## **Asymmetric cores**

- Capacity
- Energy
- Latencies
- Operating points
- Cache types



# In-kernel Switcher (IKS)

- Pros
  - Minimal kernel changes
  - Available now through Linaro
- Cons
  - Half the cores used



# Heterogenous MP (HMP)

- Pros
  - All cores can be used
- Cons
  - Large changes to Linux kernel
  - Production-ready only next year
    - Basic feature-set for partners 1Q 2013
    - Upstreaming several months
    - Optimisations



# Being a catalyst...

- Solving long standing problems
  - Better CPU qiesce
  - Better scheduling
- Useful for SMP (A9, A15)



### Mini-summit agenda

- Plenary Robin Randhawa
  - Whirlwind tour of experimental results on TC2
- Session 1 (09:00 09:55)
  - Status overview
  - Making Linux work with asymmetric systems
- Session 2 (10:00 10:45)
  - The Bluesky session: What would the ideal power-aware kernel do? (45 mins)
- Session 3 (11:00 11:55)
  - Back to reality: What do we have today and the sequence of steps to get to where we want to be (55 mins)
- Session 4 (12:00 13:00)
  - Workloads and Test Automation (30 mins)
  - General Discussions on further work and Wrap-Up (30 mins)



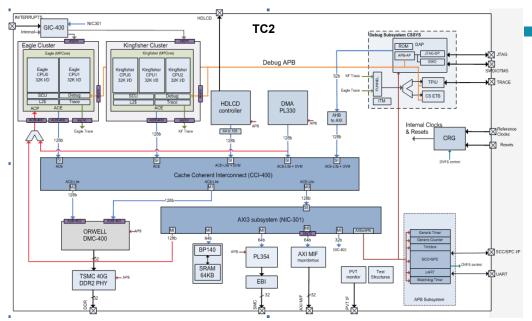
# big.LITTLE on TC2

## Robin Randhawa





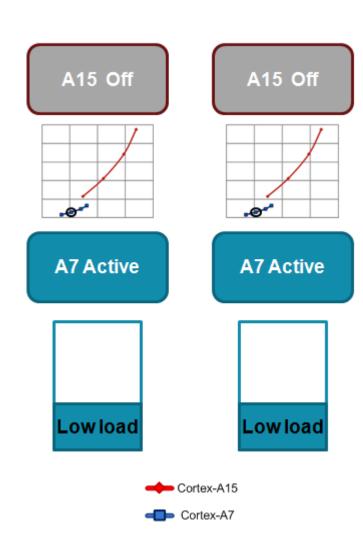
#### ARM's Test Chip 2 (TC#2): An Overview





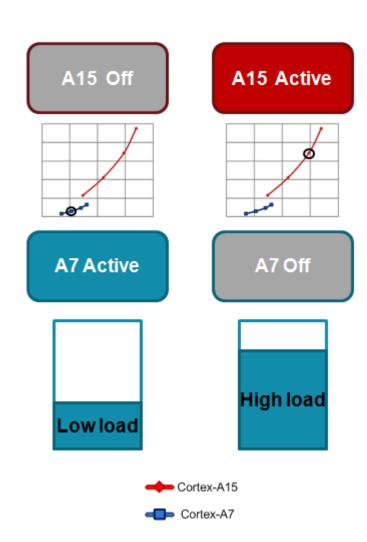
- A Versatile Express core tile publically available:
  - Capabilities
    - 2 x A l 5 (r2p l) @ up to 1.2 Ghz
    - 3 x A7 (r0pl) @ up to IGhz
    - CCI/DMC/GIC/ADB (r0p0)
    - DMA (PL330)
      - 2GB external DDR2 memory@ 400Mhz
    - 64k internal SRAM
  - Coresight debug (including JTAG and ITM trace but no STM)
  - No GPU
  - cpufreq support: Independent for each cluster with limited voltage scaling
  - cpuidle support: Cluster power gating

## **IKS: CPU Migration**



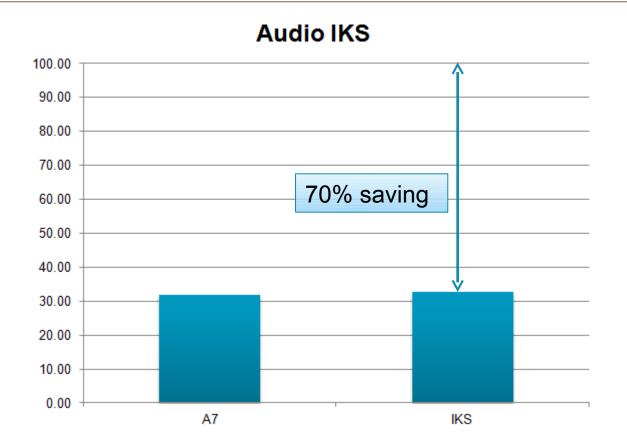
- big.LITTLE extends DVFS
- DVFS algorithm monitors load on each CPU
- When load is low it can be handled on a LITTLE processor
- When load is high the context is transferred to a big processor
- The unused processor can be powered down
- When all processors in a cluster are inactive the cluster and its L2 cache can be powered down

## **IKS: CPU Migration**



- big.LITTLE extends DVFS
- DVFS algorithm monitors load on each CPU
- When load is low it can be handled on a LITTLE processor
- When load is high the context is transferred to a big processor
- The unused processor can be powered down
- When all processors in a cluster are inactive the cluster and its L2 cache can be powered down

#### IKS: Results for Audio on TC2



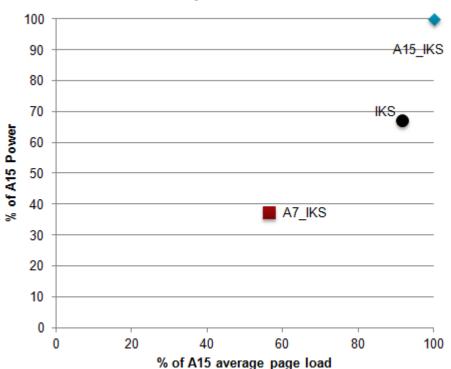
TC2:
A15 up to 1.2 GHz
A7 up to 1 GHz
Better results expected on representative silicon.

- Power compared to executing the use case on A15
- IKS does not use A15s during Audio run



#### IKS: Results for BBench + Audio on TC2



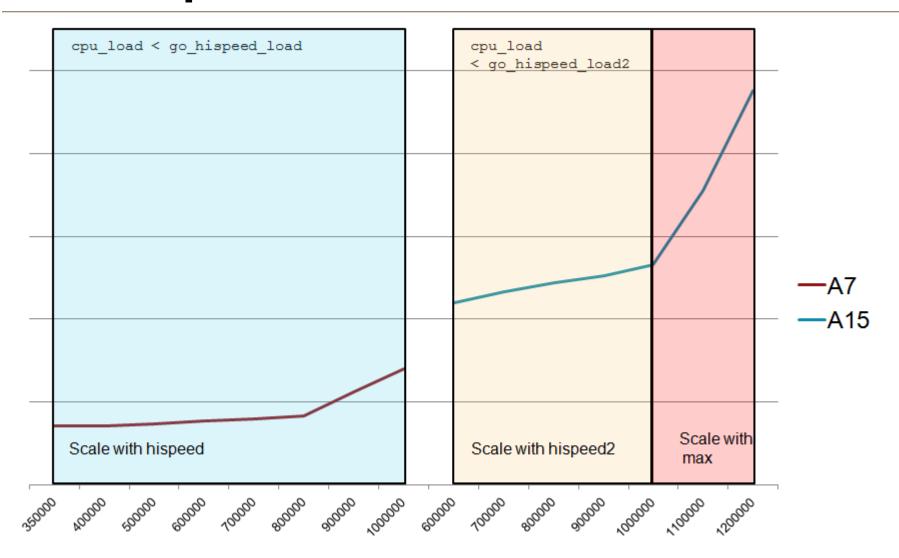


TC2:
A15 up to 1.2 GHz
A7 up to 1 GHz
Better results expected on representative silicon.

- Performance is measured as from page loading times of BBench
- Results normalised to power and performance consumed on same use case run on A15 only

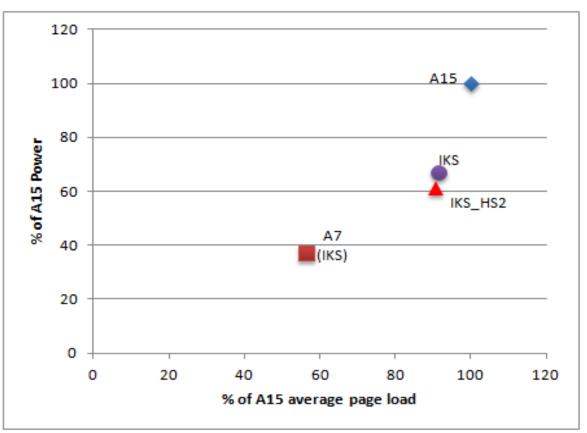


## IKS: Hispeed2



### IKS: Results: Bbench + Audio

#### BBench page + Audio



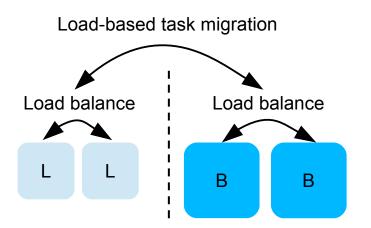
TC2:
A15 up to 1.2 GHz
A7 up to 1 GHz
Better results expected on representative silicon.

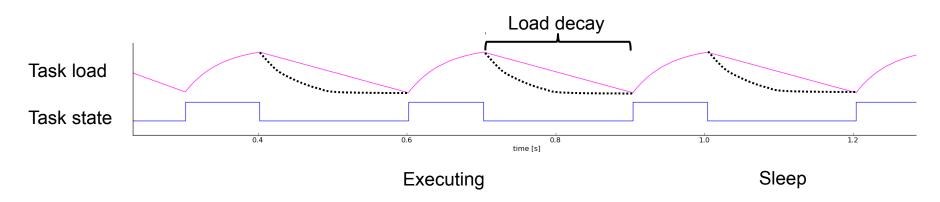
Power improves with no performance cost



#### MP solution – more details

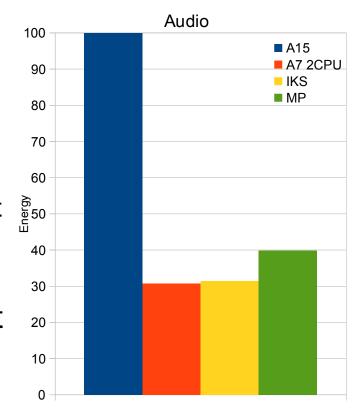
- Scheduler modifications:
  - Treat big and LITTLE cpus as separate scheduling domains.
  - Use PJT's load-tracking patches to track individual task load.
  - Migrate tasks between the big and the LITTLE domains based on task load.





#### MP: ARM TC2: Audio

- Workload: Audio (mp3 playback)
- Performance/Energy target:
  - A7 energy
- Status:
  - Audio related task do not use A15s, but the power consumption is still significantly more than A7 alone.
- MP not as power efficient as IKS yet
- Todo:
  - Target spurious wake-ups on A15. All the extra power comes from the A15's which shouldn't be used at all.

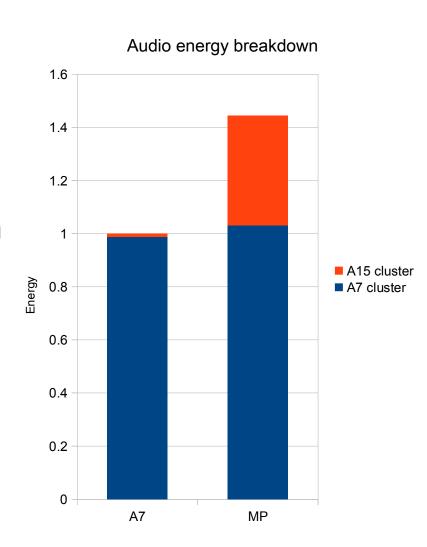


	Energy
A7	30.79%
MP	39.86%



## MP: Audio workload analysis

- Where is the extra energy spent with MP?
  - Need a look at why A15's consume power when they are not necessary
- We see unwarranted wake ups on A15
  - No user threads running on A15
  - Tend to favour CPU0
  - Examples:
    - tick\_sched\_timer (99.7% on CPU0)
    - Hrtimers
    - Workqueue





### MP – Top Issues

- Spurious wakeups
  - A15s are woken up by scheduler ticks (mainly)
  - Workqueues
  - Timers
  - RCU
  - Scheduler ticks
- cpu wakeup prioritisation
  - Pick the cheapest target cpu
- Balancing
  - Scale invariance
  - Load accumulation rate
  - Spread load to A7s when A15s are overloaded
  - Pack vs. spread
- Cluster aware cpufreq governors

