FLASH

Fast Linux Advanced Scheduler Hardware

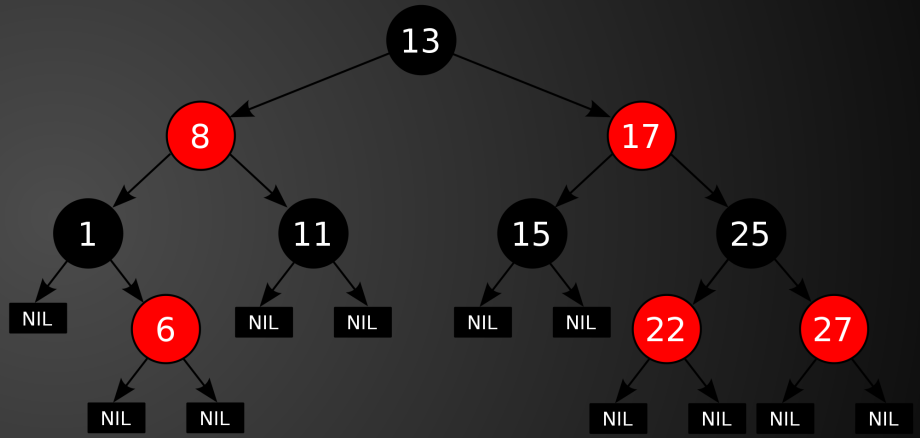
Mark Aligbe, Chae Jubb

What's the Problem?

- Context switching isn't always free
- Tasks suffer degraded performance from optimal (batch)

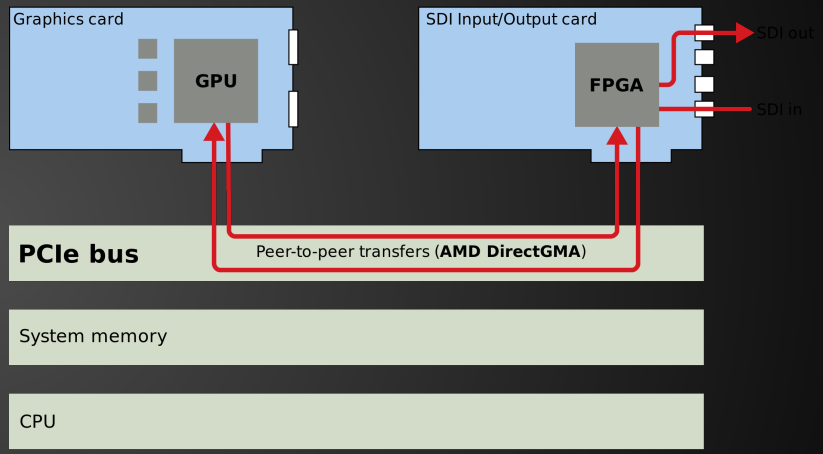
Loss in Performance

- CFS is a (too) smart algorithm
- Data and instruction cache invalidation
- CFS also hinders TLB performance



Solution

- Scheduler hardware unit
- Simple lookup from the operating system
- Energy and area efficient



Previous Hardware Schedulers

- Real-time embedded applications
  - Fixed number of tasks
  - Tightly coupled hardware unit
- Simple algorithms
  - Usually few tasks

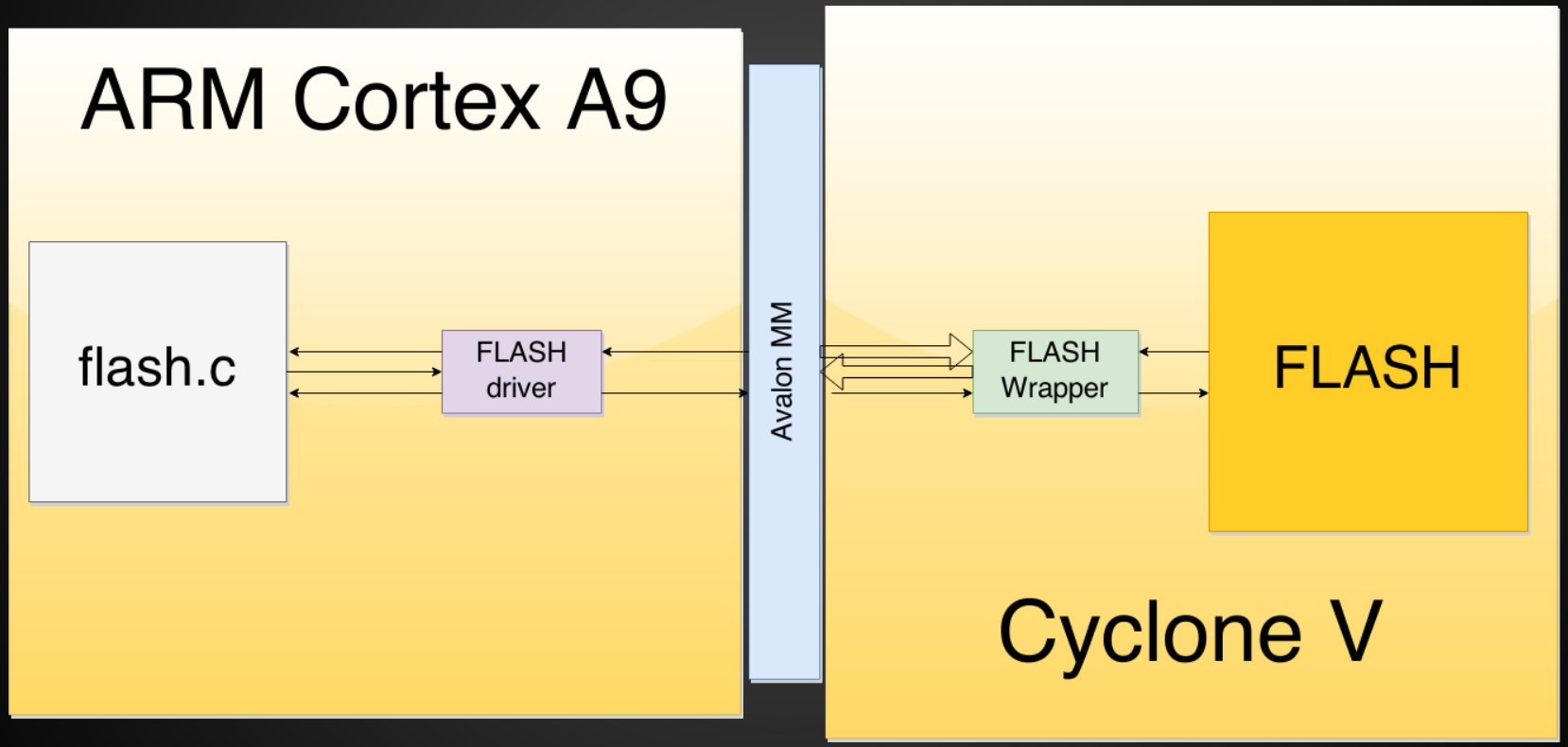
Challenges in Applying to Linux

- Number of tasks not fixed
- Higher number of tasks
- Complicated scheduling algorithm
  - CFS

Our Scheduling Algorithm

- Simple CFS implementation
- Virtual runtime
  - Weighted by priority
  - Same relative weights as usual kernel CFS implementation

Architecture



Modifying the Kernel

- Mirror task state changes to device
- Create light scheduler shell around required functions
- Uses lookups

Hardware - Kernel Interaction

- Device sends timer tick interrupt
- Kernel can request next task

/dev/

Testing Environment and Beyond

- Prototype in CARGO -> Proof of concept with SoC -> Accelerator units in modern CPUs
- Perf counters: cache and TLB performance
- Hackbench + oprofile: scheduler residency
- Interactivity test: kernel make + imgur (user experience and make completion time)

Results

:(

Engineering Experiences

Platform Challenges

- DE2 is ancient
- CARGO simulator uses an emulated kernel, and the actual kernel lacks correct performance counting
- Going from SystemC to Quartus was an endeavor in versions
- CtoS

Platform Realizations

- SoCKIT is so much simpler
- CtoS

Obstacles to Adoption

- Need a new piece of hardware
  - Unless integrated into processor like MMU
- Kernel modifications
  - Module is not sufficient

Questions?

Image Sources

http://en.wikipedia.org/wiki/Direct_memory_access
http://en.wikipedia.org/wiki/Red%E2%80%93black_tree
http://blog.sqlauthority.com/2013/10/09/big-data-buzz-words-what-is-mapreduce-day-7-of-21/
http://www.cs.montana.edu/~chandrima.sarkar/AdvancedOS/CSCI560_Proj_main/fig1.jpg
http://www.skomplekt.com/newsimg//soc.jpg