Trends and Challenges in Performance Analysis and How Tools and PerfMon are Changing

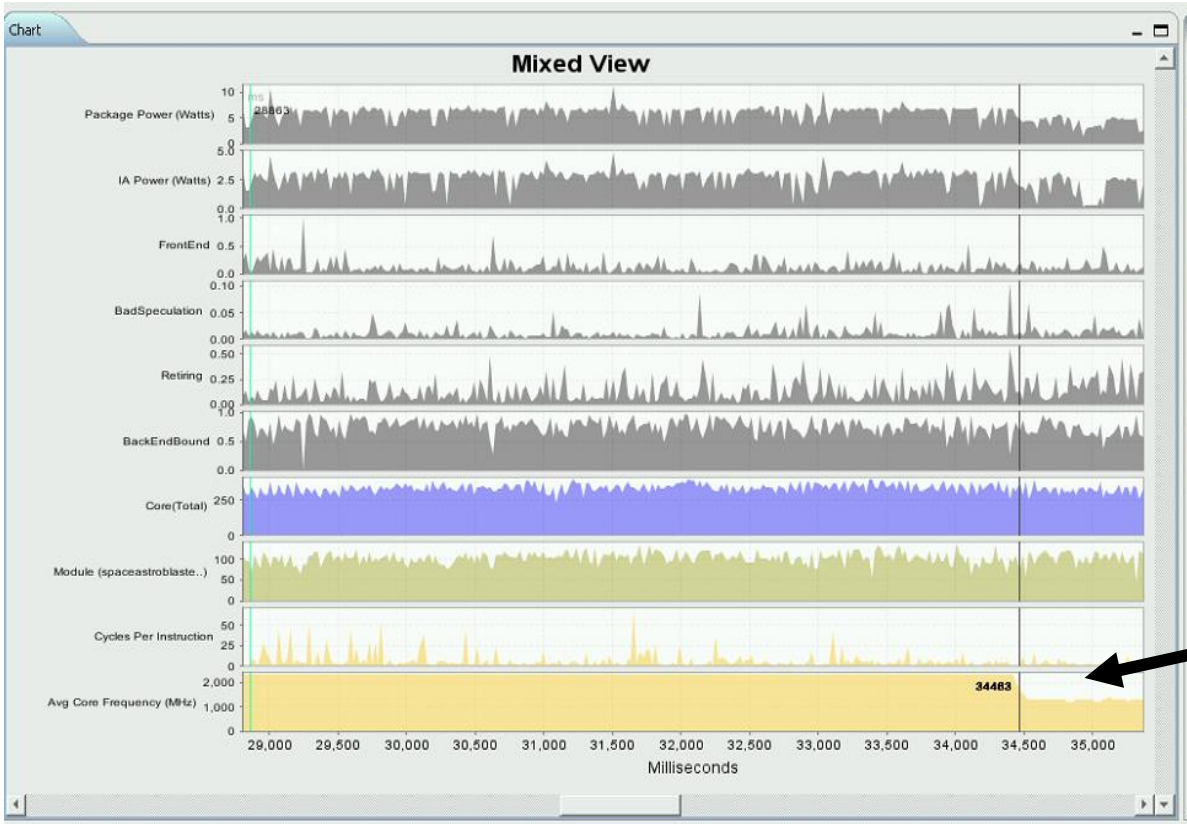
Michael Chynoweth - Sr. Principal Engineer Intel Corporation

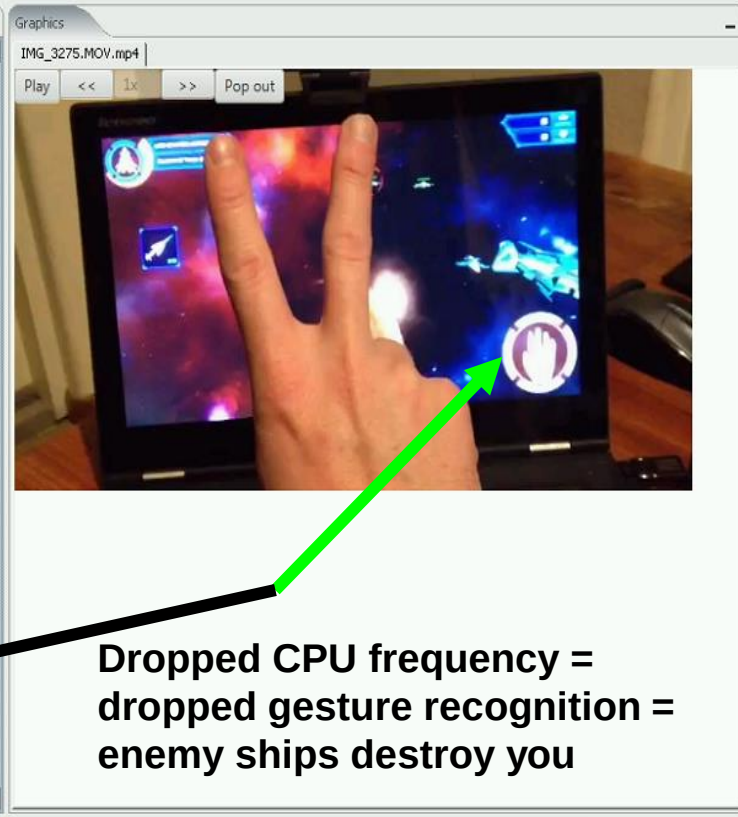
Contributors: Joe Olivas, Ryan Mclaughlin, Patrick Konsor, Rajshree Chabukswar, Sneha Gohad, Ahmad Yasin, Charlie Hewett

Agenda

- Power
  - Problem#1: Power, QoS and Routing Power to the Right Component
  - Problem#2: Spinning to Avoid Sleep States
- Profiling without perturbing the system

Power, QoS and Routing Power to the Right Component





Dropped CPU frequency = dropped gesture recognition = enemy ships destroy you

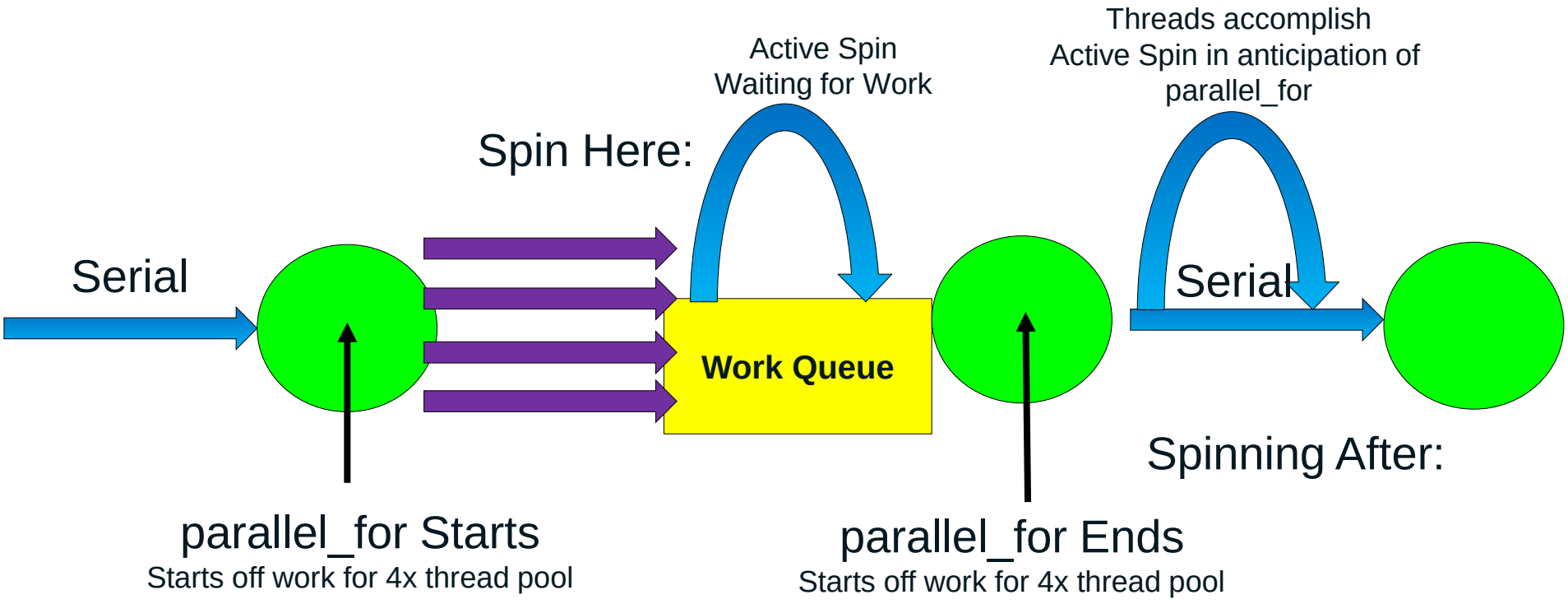
Sync Phone Video with Overtime Profile

Popular Benchmark: Cstates Enabled/Disabled Impact on uArch States (TLB States)

Event Names	C-states Enabled	C-states Disabled	% of C0 Cycles Difference
CPU_CLK_UNHALTED.THREAD	7.59E+10	6.88E+10	
DTLB_LOAD_MISSES.STLB_HIT	4.94E+08	6.26E+08	13.01%
DTLB_LOAD_MISSES.WALK_ACTIVE	6.93E+09	4.54E+09	33.66%
DTLB_STORE_MISSES.STLB_HIT	9.18E+07	1.21E+08	2.88%
DTLB_STORE_MISSES.WALK_ACTIVE	9.57E+08	6.26E+08	4.66%
		Total	54.22%

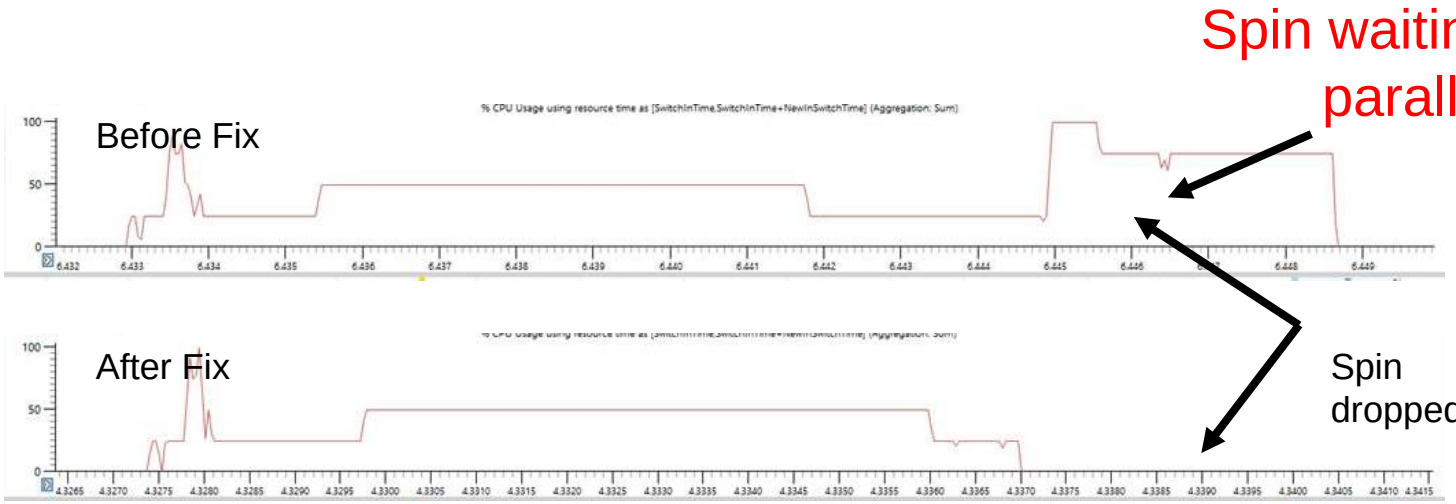
DTLB Cycles >50% of C0 cycle increase

Spinning to Avoid C-States



Spinning to Avoid C-States

	OpenCV With Active Spin	OpenCV Without Active Spin	Total Power Savings
	8.32 W	6.15 W	2.2 W



22% drop single core utilization and 2.2 W of power

Active Spin Costs 2.2W

Spins Are Not Always Easy to Catch: Kernel Spin for Milliseconds

Disasm	Latency(SKYLAKE)	Static ASM Notes
inc ebx		LOOP_START
test dword ptr [rip+4fba4], ebx	1	REG_PLUS_2K_LOAD
jnz lc0001d2a		
mov rax, qword ptr [rip+4fb9f]		REG_PLUS_2K_LOAD
test rax, rax	1	
jnz lc001bbe6		
pause		
mov rcx, rsi	1	
call lc0001e00		
test dword ptr [rcx+e0], 10000		
jnz lc001bc08		
mov rax, qword ptr [rcx+48]		
ret		
mov rcx, rax		
mov rax, qword ptr [rsi+70]		
cmp rax, r13	1	
jnz lc001bbf4		
rdtsc	24	SKYLAKE_HIGH_LATENCY_INS...
shl rdx, 20		
or rax, rdx	1	
cmp rax, rdi	1	
jb lc001bc00		
lea rcx, ptr [r14+rax*1]	1	
mov rdi, rax	1	
cmp rcx, rbp	1	
jb lc0001d10		LOOP_END

Spin on pause and timing

Resolving Issues with Power

- Power Decisions Impact to Performance
  - Visibility in tools helps resolve issues quickly
  - Improving our C-State and P-State decisions
- Detecting Active-Spin-On-Pause
  - KBLR/CFL and beyond have added event for Pause instruction
    - ROB_MISC_EVENTS.PAUSE_INST
    - Can also estimate the percentage of time in pause
      - PAUSE_COST = ROB_MISC_EVENTS.PAUSE_INST*140/CPU_CLK_UNHALTED.THREAD

Profiling Without Performance Monitoring Interrupts (or Ring Transitions)

- Extended Processor Event Based Sampling (PEBS) Definition
  - Supports use of PEBS-like triggering on all programmable and fixed counters
- Advantages of Extended PEBS
  - More precise event attribution
  - Avoids need for an expensive Performance Monitoring Interrupt (~10k cycles)
  - Avoids “blind spots” when interrupts masked
  - Information required is captured on PEBS assist rather than in the profiling driver

Conclusions

- Power, QoS and Routing Power to the Right Component
  - Visibility into Power Control Unit Decisions
- Active spins to avoid sleep states
- Profiling without perturbing the system
  - Extended PEBS will reduce PMIs, overhead and perturb the system less

Backup