Intel Running Average Power Limit Technology

Stephanie Labasan[†○], Jeffrey Shafer[†]

Robin Goldstone*, Barry Rountree[❖]

University of the Pacific[†], Lawrence Livermore National Laboratory[○❖], Integrated Computing & Communications*, Center for Applied Scientific Computing[❖]*

Team Venus took advantage of Intel's power management features, specifically Running Average Power Limit (RAPL) interfaces, in order to build the best performing system given a defined power budget. By enforcing power consumption limits on our Intel EZ-2670 2.6GHz processors, the team was able to simultaneously decrease system power consumption to within 26A and maintain comparable performance to the system with the default limits.

Motivation

Given a defined power limit of 26A, it was necessary to obtain the highest performance possible by capping the power consumption limit. Intel's Running Average Power Limit (RAPL) Technology allows users to manipulate processors to maximize a system's power budget.

RAPL Interfaces

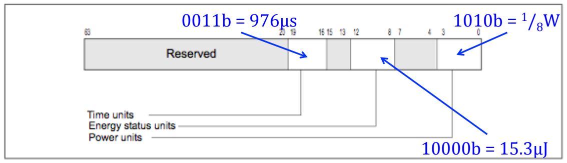
- *Power Limit*: Specifies power limit, time window, lock bit, clamp bit, and enable bit
- *Energy Status*: Specifies total energy consumed
- *Perf Status*: Provides information on the performance effects due to power limits
- *Power Info*: Provides information on the range of parameters for a given domain, for example, min power, and max power
- *Policy*: four-bit priority information to divide hardware among domains

Conclusion

Linpack: The results were inconsistent with what we expected to occur. We expected both power and performance to decrease as we reduced the power cap, but found that performance actually increased at our lowest power cap setting. We hypothesize that a lower performance occurred at a higher power cap because more energy is dedicated to cooling the system, taking away from the computation.

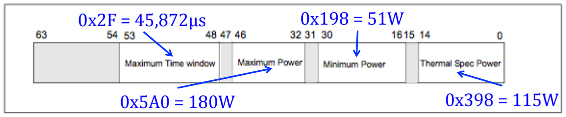
CAM: Studying CAM's performance at varying power cap limits gave us more reasonable results. As we gradually lowered the power cap value, the amount of current drawn similarly decreased, while the runtime increased.

Model Specific Registers



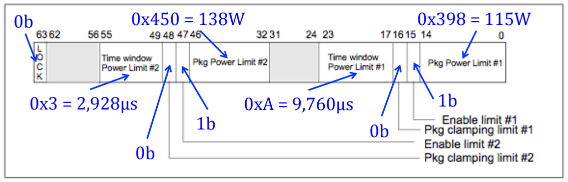
MSR_RAPL_POWER_UNIT: rdmsr 0x606 = 0xA1003

- Scaling factors supplied to allow for precision in a finite number of bits



MSR_PKG_POWER_INFO: rdmsr 0x614 = 0x2F05A001980398

- Inferred minimum time window = 0x1 = 976µs



MSR_RAPL_POWER_LIMIT: rdmsr 0x610 = 0x6845000148398

- Lock: If set, all write attempts ignored until next RESET
- Clamping Limit: Allow going below OS-requested performance and throttle states in the CPU
- Package Power Limit #1: Sets the lower "average" limit
- Package Power Limit #2: Sets the higher "peak" limit

Diagrams courtesy of Intel's® 64 & IA-32 Architectures SW Developer's Manual

Case Study 1: Linpack

Challenge: Because Linpack heavily stresses the processors, memory subsystem, and interconnect, it was our primary target for applying the power capping technology. We tested several cap values on both seven and eight node cluster configurations. Our optimal configuration (in green below) resulted in a power savings of 9% with only a 3% reduction in performance.

8 Nodes

Power Cap #1/#2	**Amps**	**Performance**	**Runtime**
115W/138W*	**26.7A**	**2391GF/s**	**3589s**
100W/115W	25.4A	2290GF/s	3687s
85W/100W	**24.2A**	**2333GF/s**	**3681s**

7 Nodes

Power Cap #1/#2	**Amps**	**Performance**	**Runtime**
115W/138W*	24A	2006GF/s	3502s
138W/150W	24.1A	2074GF/s	3386s

* indicates default power cap

Case Study 2: CAM

Challenge: When running the high detail "f05_g16" data set at the default power cap across eight nodes, we found that, as with Linpack, we were exceeding the power limit. Power capping proved to be extremely beneficial in optimizing our overall efficiency.

8 Nodes

Power Cap #1/#2	**Amps**	**Runtime**
115W/138W*	**26.4A**	**2319s**
110W/115W	26A	2362s
95W/100W	**24.6A**	**2382s**
85W/95W	22.3A	2494s

* indicates default power cap

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

LLNL-POST-599036