An empirical survey of performance and energy efficiency variation on Intel processors

Aniruddha Marathe Lawrence Livermore National Laboratory marathe1@llnl.gov

Nirmal Kumbhare University of Arizona nirmalk@email.arizona.edu Yijia Zhang Boston University zhangyj@bu.edu

Ghaleb Abdulla Lawrence Livermore National Laboratory abdulla1@llnl.gov Grayson Blanks Lawrence Livermore National Laboratory blanks1@llnl.gov

Barry Rountree Lawrence Livermore National Laboratory rountree4@llnl.gov

ABSTRACT

Traditional HPC performance and energy characterization approaches assume homogeneity and predictability in the performance of the target processor platform. Consequently, processor performance variation has been considered to be a secondary issue in the broader problem of performance characterization. In this work, we present an empirical survey of the variation in processor performance and energy efficiency on several generations of HPC-grade Intel processors. Our study shows that, compared to the previous generation of Intel processors, the problem of performance variation has become worse on more recent generation of Intel processors. Specifically, the performance variation across processors on a large-scale production HPC cluster at LLNL has increased to 20% and the runto-run variation in the performance of individual processors has increased to 15%. We show that this variation is further magnified under a hardware-enforced power constraint, potentially due to the increase in number of cores, inconsistencies in the chip manufacturing process and their combined impact on processor's energy management functionality. Our experimentation with a hardwareenforced processor power constraint shows that the variation in processor performance and energy efficiency has increased by up to 4x on the latest Intel processors.

CCS CONCEPTS

• General and reference → Empirical studies; • Hardware → Power estimation and optimization; *Platform power issues*;

KEYWORDS

Empirical studies, Performance analysis, Energy distribution

ACM Reference Format:

Aniruddha Marathe, Yijia Zhang, Grayson Blanks, Nirmal Kumbhare, Ghaleb Abdulla, and Barry Rountree. 2017. An empirical survey of performance and energy efficiency variation on Intel processors. In *Proceedings of E2SC'17: Energy Efficient Supercomputing (E2SC'17)*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3149412.3149421

- 2017, 1007, 12, 17, 2017, Denver, CO, COM

© 2017 Association for Computing Machinery. ACM ISBN 978-1-4503-5132-4/17/11...\$15.00

https://doi.org/10.1145/3149412.3149421

Figure 1: Comparison of sequential processor performance on three Intel processors for a computation-heavy workload

1 INTRODUCTION

HPC performance optimization efforts have traditionally focused solely on application performance characterization with the assumption that the performance variation on the underlying platform is predictable within a small, known bound. However, with the increase in the complexity of both the processor power management features and the system software, performance variation has become an increasingly challenging problem towards improving overall system efficiency[22]. Run-to-run variation is typically attributed to system noise which is primarily caused by system processes[13], on-node and off-node resource contention[5], and platform bugs[22]. Inter-processor performance variation is typically caused by the process inaccuracies introduced during the chip manufacturing process which affect processor's dynamic frequency throttling operation and energy efficiency[2]. In this paper, we study both the run-to-run and the inter-processor performance variations with an emphasis on inter-processor variation on several generations of Intel processors.

Inter-processor variation occurs in identical processors in the same stock keeping unit (SKU) that operate at different effective frequencies. Modern Intel processors increasingly rely on dynamic overclocking of the Turbo Boost Technology to achieve maximum performance possible for a given type of workload and operating conditions[2]. For a processor, the effective frequency attained by Turbo Boost depends on the number of active cores, the type



ACM acknowledges that this contribution was authored or co-authored by an employee, contractor, or affiliate of the United States government. As such, the United States government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for government purposes only. *E2SC'17. Nov. 12–17. 2017. Denver. CO. USA*

of workload and processor's power and thermal headroom. This variation in frequency on seemingly homogeneous processors is typically a collective effect of transistor-level variation introduced by the CMOS manufacturing process, variations in other node level components, and thermal conditions[3, 20]. We call this type of variation in the processor performance, power and thermal characteristics due to process variation *manufacturing variability*.

Figure 1 compares the performance variation on three Intel processors *viz.*, Sandy Bridge, Ivy Bridge and Broadwell (from left to right) for a computation-heavy, embarrassingly parallel benchmark called *Firestarter*[10]. The boxplot shows the performance of Firestarter on two set of processors over several runs on the three clusters normalized to the best-performing run. The plot shows that Ivy Bridge and Broadwell show progressively worse performance variation on both sockets compared to Sandy Bridge. This observation motivates us to study the performance variation on the three clusters for several HPC benchmarks with different degrees of compute-boundedness, cache and memory access patterns, and types of instructions.

Constraining processor power has been shown to magnify the inter-processor variation in performance and energy efficiency. Previous studies have shown Intel Sandy Bridge and Intel Ivy Bridge clusters to exhibit up to 30% and 60% inter-processor performance variation, respectively, upon severely constraining processor power[3, 7, 9, 12, 17, 20]. These studies suggest that future generations of processors may show higher impact of manufacturing variation on performance and power efficiency as they scale up in number of cores. Our study confirms, for the first time on Broadwell, that this variation is worse than anticipated, which further complicates achieving better system efficiency.

In this work, we analyze several types of performance and energy efficiency variations on three generations of Intel processors. Specifically, we present the following observations on the Intel Broadwell cluster relative to the performance on our Intel Sandy Bridge and Intel Ivy Bridge clusters at LLNL:

- For a computation-heavy workload, the variation in sequential processor performance has increased from 4.7% to 13.5% in the median case and from 7% to 17% in the worst-case.
- The inter-core performance variation for the computationheavy workload has increased from 2.5% to 5%.
- Under a hardware-enforced power limit, the worst-case variation in processor performance for several benchmarks has increased significantly from 30% on Sandy Bridge to 1.4x on Ivy Bridge and 4x on Broadwell for severe power limits. Our analysis shows that existing methods to model the performance variations are inadequate to capture the non-linear relationship between the power limit and the observed metrics of performance and power usage.
- The variation in power usage across processors has also increased at higher power limits from 10% to 20%.

2 EXPERIMENTAL SETUP

This section describes in detail our experimental setup in terms of the hardware, platform configuration and applications.

2.1 Cluster specification

Table 1 lists the specifications of the three HPC clusters at LLNL on which we performed our experiments: Cab, Catalyst and Quartz. Nodes in Cab and Catalyst are connected using InfiniBand interconnects whereas, nodes in Quartz are connected using Intel Omni-Path interconnects. Each node in clusters comprises of two Intel processors. Memory specified in Table 1 is equally divided among both the processors on each node. Both the processors are connected via Intel QPI. All three clusters run the Tri-Lab Operating System Software (TOSS) which is based on Red Hat Enterprise Linux Server 7. Hyper-threading is enabled by default on Catalyst and Quartz and requires root privileges to disable. Therefore we leave one hyperthread idle on each core (unless otherwise specified) to minimize the effects of system noise[13]. Hyper-threading on Cab is disabled.

2.2 Software tools

We used the Intel compiler tool chain and MVAPICH2 to build all benchmarks. We used -O2 option to enable compiler-level optimizations and -qopenmp to enable OpenMP threads. For powerlimiting using Intel RAPL and reading performance counters, we used a lightweight monitoring library called *libPowerMon*[14] with *msr-safe*[19]. Intel Turbo was enabled so that the applications could extract maximum performance under the power limit[3].

2.3 Design of Experiments

For benchmarking, we used EP, MG, CG, and FT from the NAS Parallel Benchmark Suite[4], STREAM[15], Firestarter[10], Prime95 [1] and DGEMM[8]. These benchmarks were selected because of the following reasons (1) they have different average and peak power consumption, (2) the chosen input problem sizes for STREAM, CG and MG keep DRAM power consumption high, (3) Firestarter and Prime95 are compute-intensive benchmarks designed to keep the CPU power consumption close to its Thermal Design Power (TDP).

For our power-uncapped experiments, we ran Firestarter for 60 seconds over 75 times and collected end-to-end measurements on all three clusters. For our power-capped runs, we configured the benchmarks to run for at least 120 seconds to capture potential effects of steady-state temperatures on processor power consumption. We ran each benchmark 20 times and reported median measurements. We chose an input problem size of 2³⁸ for EP. We used Class D input size for MG (MG.D) with an iteration count of 80. We selected benchmarks CG (CG.C) and FT (FT.C) to operate on a problem size of class C for an iteration count of 1000 and 330, respectively. We chose to run Prime95 and Firestarter for 120 seconds before terminating them externally. For STREAM, we selected an array size of 100 million elements with an iteration count of 1700. For DGEMM, we used a 2-D matrix size of 2700 x 2700. We ran a single instance of each benchmark on each processor so as to eliminate inter-node and inter-processor communication.

3 PERFORMANCE VARIATION WITHOUT POWER CAPPING

In this section, we present our detailed analysis of the variation in performance and energy efficiency of Intel processors observed in Figure 1. Figure 2 compares core-level performance variation

Cluster	Node	Intel	Architecture	Clock	Cores per	Processors	Memory per	Processor
	Count	Processor ID		Speed (GHz)	Processor	per node	Node (GB)	TDP (W)
Cab	1296	Xeon E5-2670	Sandy Bridge	2.6	8	2	32	115
Catalyst	324	Xeon E5-2695-v2	Ivy Bridge	2.4	12	2	128	115
Quartz	2688	Xeon E5-2695-v4	Broadwell	2.1	18	2	128	120

Table 1: Cluster Configuration



Figure 2: Comparison of core-level performance variation on the best-, median- and worst-performing nodes of Sandy Bridge and Broadwell clusters for Firestarter.

of Firestarter on two processors on our Sandy Bridge and Broadwell clusters in terms of number of iterations completed within one minute. We present several observations that show that processor and core performance have become worse on Broadwell compared to Sandy Bridge. First, the median core performance on the worst node on our Sandy Bridge cluster is up to 5% lower than the median core performance of the best node, whereas, the same for our Broadwell cluster has increased to 13.5% (note that we have removed the outliers in the plot). This shows the degree of performance non-homogeneity on Broadwell. Second, the worst-case core performance on our Broadwell cluster is 17%, which is significantly worse than the worst-case performance on our Sandy Bridge cluster which is 7%. Third, the median processor performance (shown by the green, blue and red lines) between processor 0 and processor 1 on Sandy Bridge show up to 1% difference. However, that difference in median processor performance increases to up to 5% on Broadwell (3% even on the median nodes). This result has a direct impact on how performance metrics must be treated on Broadwell. For example, to show that a new performance optimizing method actually yields the expected improvement, the evaluation must conduct sufficient number of runs to show a median improvement adjusted to the median core-level variation. Fourth, on both clusters, core 0 on each processor typically shows lower performance compared to

other cores on the processor. Core 0 also shows more variation in sequential performance than other cores. Intel Ivy Bridge cluster showed worse median performance than Sandy Bridge but better median performance than Broadwell. Due to limited space, we do not show our results on Ivy Bridge in the rest of this section.

Variation in performance of Processor 0 Core 0

Figure 3 shows the distribution of core 0 performance with respect to its operating frequency on processor 0 for several runs on all nodes of our Sandy Bridge and Broadwell clusters. Each dot represents one run of Firestarter. The core frequencies on Broadwell are more uniformly distributed than on Sandy Bridge, which shows that the best and worst-performing nodes for the Broadwell cluster in Figure 2 are not outliers in contrast to the Sandy Bridge cluster. The difference between the best-performing runs (upper end of the red and green dots) on Broadwell is 15% compared to 7% on Sandy Bridge. We observe a strong correlation between the operating frequency and best-case performance of the cores which explains the difference in the best-case performance for red and green dots for both Sandy Bridge and Broadwell. We observe that the core operating frequency is weakly correlated to core temperatures on Broadwell (data not shown in the paper). This suggests that the variation in operating frequency occurs due to manufacturing variation



Figure 3: Correlation between core-level performance variation with frequency for Processor 0, core 0 on Sandy Bridge and Broadwell clusters for Firestarter.

in the processors. Moreover, the difference between the absolute best and worst performing cores on best and worst nodes is up to 20.1% on Broadwell compared to 11.9% on Sandy Bridge. For the median node, the run-to-run variation on Broadwell has increased to 10% from 5.2% on Sandy Bridge. The operating frequency appears to limit the best-case performance on individual cores which suggests that system noise may have caused the run-to-run variation within a core. Although cores other than core 0 show slightly lower variation, they follow similar trend on Broadwell relative to Sandy Bridge. Thus, Figure 3 shows that both manufacturing variation and system noise potentially induce higher variation in performance and energy efficiency on Broadwell than Sandy Bridge.

Performance variation with hyper-threads on Broadwell Figure 4 shows the normalized performance of Firestarter on both physical and logical threads of the best, median and worst processors (including processors 0 and 1) on the Broadwell cluster. Figure 4 shows that the performance variation becomes worse when running on both threads compared to a single thread per core per processor (Figure 2 (b)) in the following ways. First, the inter-thread performance variation on a processor gets magnified to up to 8% for the best processor and up to 10% on the worst processor compared to the inter-core performance variation in Figure 2 (b). Second, we observe that the run-to-run performance variation in the individual threads is consistently higher than that in individual cores in Figure 2 (b). Within each core, hyper-thread 1 shows consistently better performance than hyper-thread 0 across all cores and processors which is surprising. Lastly, at the processor level, the relative worstcase performance (as shown by the boxplots in red) is as low as 78% compared to the best-case performance. Our analysis extends previously reported performance variation for hyper-threads[13].



Figure 4: Variation in thread-level performance of Firestarter on Broadwell. From top to bottom, the three horizontal bands of boxplots show thread-level performance on the best, median and worst processors. Alternate white and grey background bands show groups of threads that belong to the same core. Red boxplots show the spread of the minimum thread performance on each processor.

4 PROCESSOR VARIATION UNDER A HARDWARE-ENFORCED POWER LIMIT

This section describes our evaluation of evolution of variation in processor performance and energy efficiency over three generations of Intel processors under several processor power limits.

An empirical survey of performance and energy efficiency variation on Intel processors

E2SC'17, Nov. 12-17, 2017, Denver, CO, USA



(c) Quartz cluster - Broadwell Architecture

Figure 5: Performance of Intel processors running eight benchmarks on three clusters under five hardware-enforced power limits. Three rows show the results collected on Sandy Bridge, Ivy Bridge and Broadwell clusters. Results from the eight applications are organized column-wise. In each subfigure, one curve shows the normalized Instruction per Cycle (IPC) of one processor at different power limits. Dots on each curve represent the actual experimental measurements, and dots from the same processor are connected by line segments. Curves are colored according to the IPC of the processors when running Prime95 at 50W power limit, and the same coloring are applied to the same processor for other applications. For clarity, we only show data for same best 200 and worst 200 processors on each cluster.

4.1 Power-constrained performance variation

From top to bottom of Fig. 5, we demonstrate the performance variation of processors in terms of Instructions per Cycle (IPC) on Sandy Bridge, Ivy Bridge, and Broadwell clusters, respectively. The IPC-Power usage curves in each subfigure characterize the performance variation of the processors on that cluster when running a specific application under five different power limits. In each subfigure, the performance of an individual processor is characterized by an IPC-Power usage curve. The dots on the IPC-Power usage curves are measured IPC values averaged over the application run and also aggregated over all the cores on the same processor. The IPC values in each cluster and each application are normalized according to the *maximum* value in that case. The curves in Figure 5 are colored according to their performance (in terms of IPC) at 50W power limit when running the application Prime95, then the same coloring is applied to the same processor in the rest of the subfigures with other applications on the same cluster. In general, we observe increasingly higher inter-processor performance variations in all applications from Sandy Bridge to Broadwell.

We make the following interesting observations from Figure 5. First, within the same cluster, the sequence of the colors of the curves from top to bottom typically remains the same across applications. Specifically, the best to worst processors for Prime95 are typically also the best to worst processors for other applications. Second, due to the difference in application characteristics and their ability to draw power at different rate, the processors either meet the processor TDP first or meet the frequency limit first, resulting in different shapes at the top-right end of the curves. For Prime95 and Firestarter, the processors meet their TDP first, so there appears a limit at the end of the curves in terms of maximum power usage (we find that IPC is strongly correlated to frequency). For DGEMM, EP, MG, STREAM, and CG, a varying degree of processors meet the frequency limit first and draw varying amounts of power, so there appears a horizontal cut at the end indicating variation in processor power usage. In case of FT, some better-performing processors meet the frequency limit first as they cannot draw more power whereas the worse-performing processors meet the TDP first. We observe up to 20% variation in processor power usage at higher power limits for Broadwell compared to 10% on Sandy Bridge. Third, the shape and trend of the curves are generally consistent for the three clusters. However, the range and shape of performance variations among the processors are different in these three clusters. From older Sandy Bridge to the newer Broadwell, we observe a gradual expansion in the covered range of these curves. The worst-case power-constrained variation has increased from 30% on Sandy Bridge to up to 1.4x on Ivy Bridge and up to 4x on Broadwell. Also, the relationship between power limit and performance appears to become significantly more non-linear and diverse across applications from Sandy Bridge to Broadwell.

4.2 Variation in energy efficiency

Figure 6 compares STREAM and DGEMM benchmarks over metrics that describe the energy efficiency of applications. The X-axis shows five processor power limits including the TDP, which corresponds to the operational limit of individual processors. The Y-axis on each plot shows the measurements normalized per cluster for each application over all five power limits. Figures 6 (a) and (b) show that the variation in processor power efficiency (in terms of IPC per watt) depends on the application characteristics. For STREAM (Figure 6 (a)), Broadwell shows significantly higher variation in IPC per watt at 50W limit where the variation in IPC is the highest, and at TDP where the variation in power usage is high. The variation in IPC per watt reduces between 70W and 100W limit. The lower variation occurs due to the memory-boundedness of STREAM which spends much of its cycles waiting for memory accesses to finish. On the other hand, Figure 6 (b) shows that DGEMM suffers from higher variation in IPC per watt on Broadwell even between 70W and 100W limits. The worst-case variation on Broadwell is significantly higher (up to 4x at 50W) compared to lower variation in IPC per watt on Sandy Bridge (up to 30%) and Ivy Bridge (2x).

Figures 6 (c) and (d) compare the variation in execution time of STREAM and DGEMM, respectively. STREAM shows significant variation in execution time at 50W power limit on both Ivy Bridge and Broadwell, but the variation reduces significantly at higher power limits at different rates. While the execution time variation is consistently higher on Broadwell, the execution time of STREAM is largely unaffected by the effects of manufacturing variability due to its memory-boundedness. DGEMM (Figure 6 (d)), however, shows a significant variation in execution time on Broadwell even at high power limits due to its compute-boundedness. Also, the rate of change in variation over increasing power limits is different from STREAM on the three clusters. We observe that the execution time is strongly correlated with the effective processor frequency, which is severely affected on low-efficiency processor compared to high-efficiency processors on Broadwell.

Figures 6 (e) and (f) compare the variation in processor power usage for the two applications. For both applications, the variation in power usage on Broadwell only appears near TDP, whereas, the variation is consistently higher on Sandy Bridge and Ivy Bridge at lower power limits as well. This result on Broadwell is quite surprising as we expected the difference in effective frequencies to manifest itself into power usage variation even at lower power limits. We also observed that for compute-intensive applications such as DGEMM, the power usage on Broadwell was consistently lower than the applied power limit by 10% to 15% unlike on Sandy Bridge which showed the power usage within 1% to 5% of the applied power limits for all power limits except 50W. These observations suggest that for Broadwell cluster, power must be allocated differently for configurations with power limits set to TDP compared to configurations with lower power limits.

Finally, Figures 6 (g) and (h) compare variation in energy consumption of STREAM and DGEMM. For different architectures, both applications show different trade-offs in terms of median energy and variation in energy at various processor power limits. Ivy Bridge and Broadwell show higher variations in energy at extreme power limits primarily due to the variation in runtime and power usage at those power limits. On the other hand, Sandy Bridge shows a negligible change in observed energy variation over different power limits. These observations show that for efficient allocation of energy for Ivy Bridge and Broadwell, variation in energy must be taken into account.

Ramifications for modeling power and performance. Our analysis shows that simple linear models are only effective in capturing the relationship between processor power limit and the variation in its energy efficiency metrics on older Intel processors such as Sandy Bridge. Also, the same linear models are assumed to be effective across a variety of applications[12]. In contrast, on newer Intel architectures we observe that the relationship between the processor power limit and the observed variation in its energy efficciency metrics is often non-linear across the processors of the same architecture and is different across applications. This



Figure 6: Comparison of energy-efficiency metrics for two applications (DGEMM & STREAM) on three clusters with Sandy Bridge, Ivy Bridge, and Broadwell processors.

motivates future work to apply more accurate models to describe the variation in processor power and performance on modern HPC processors.

5 RELATED WORK

This section presents a brief overview of the existing literature on characterization and mitigation of manufacturing variation on HPC compute resources. Previous work on characterizing processor performance variation can be categorized into two distinct classes: 1. run-to-run processor variation, and 2. inter-processor variation. A large piece of work exists on the highly researched topic of characterizing system noise-induced run-to-run variation at various levels of a cluster. Recent work on this topic includes characterizing and mitigating noise induced by specific compute and non-compute components in the system. For example, Leon et al.[13] and Rosenthal et al. [16] characterized system noise on Sandy Bridge and showed that using Simultaneous Multi-Threading (SMT) to co-schedule system processes can reduce system noise. Bhatele et al.[5] studied the impact of network congestion on application performance variation. Since our work primarily focuses on processor-level performance variation due to manufacturing variability, our controlled experimental setup based on the feedback of previous work on run-to-run variation typically minimizes the impact on run-to-run performance variation. Even though the run-to-run variation on modern Intel processors has marginally increased, its characterization is beyond the scope of this work.

Inter-processor performance variation is increasingly becoming a hot topic of research due to its impact on system performance and efficiency. At the hardware manufacturing level, Borkar et al.[6] and Zhang et al.[23] present a detailed study of causes and characterization of chip-level variation in processor performance. Teodorescu et al.[21] and Herbert et al.[11] provide low-level solutions to mitigate effects of processor variation on early multi-processors with frequency scaling. Rountree et al. present one of the first surveys on performance variation under a hardware-enforced power constraint on Sandy Bridge [17]. Acun et al.[3] present a comprehensive analysis of inter-processor variation induced by dynamic overclocking in Intel processors including Sandy Bridge and Ivy Bridge. Schuchart et al.[18] show performance variation on Intel Haswell processors at different frequencies. We extend the previous work in two ways. First, we validate the findings in previous work on a similar set of benchmarks on our clusters with similar processor architectures at several processor power limits. Second, we extend previous work by including the results from our Broadwell cluster, which are quite surprising. Our results show that existing solutions[7, 9, 12] to mitigate processor performance variation are rather simplistic and cannot be practically applied to Broadwell.

6 RECOMMENDATIONS FOR BROADWELL

In this section, we propose a set of recommendations for the community towards designing and running applications on Broadwell.

- Characterizing node performance based on *averaged* performance distorts the true impact of manufacturing variation on processors on the node and therefore should be avoided.
- On large-scale clusters, co-locating high and low efficiency processors on the same node should be avoided because it complicates resource allocation under a power budget.
- Reporting just the average or best-case performance and power usage on individual processors is inadequate. HPC performance and power characterization must include extreme-case measurements along with average or median measurements on processors over several runs.
- Results reported for energy-efficient and power-efficient techniques must be collected on a range of pre-characterized processors to demonstrate their effectiveness.
- On hyper-threaded systems, leaving one hyper-thread on each core idle for the system processes shows lower variation.
- Previous approaches to model processor manufacturing variability [12] cannot be applied effectively due to non-linearity and non-repeatability in the relationship between processor performance and power usage across processors and applications. Sophisticated approaches towards modeling this relationship across several dimensions will be required.

7 CONCLUSION

Our experiments on three generations of Intel processors show that, while the computation capacity is increasing, the variation in processor performance and energy efficiency is becoming worse. With increase in the complexity of processor features and number of cores, we expect this variation to grow in the future. While the practice of filtering out the least-efficient, worst-performing processors may partially limit the variation at a practically high cost, a robust runtime-based solution will be required to effectively manage processor variation at the software level.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-CONF-738511). We thank Ayse Coskun and Stephanie Labasan for their invaluable suggestions.

REFERENCES

- [1] Great internet mersenne prime search. https://www.mersenne.org/download/.
- [2] Intel turbo technology 2.0. https://www.intel.com/content/www/us/en/architectureand-technology/turbo-boost/turbo-boost-technology.html.
- [3] B. Acun, P. Miller, and L. V. Kale. Variation among processors under turbo boost in hpc systems. In Proceedings of the 2016 International Conference on Supercomputing, page 6. ACM, 2016.
- [4] D. H. Bailey, E. Barszcz, J. T. Barton, D. S. Browning, R. L. Carter, L. Dagum, R. A. Fatoohi, P. O. Frederickson, T. A. Lasinski, R. S. Schreiber, et al. The NAS parallel benchmarks summary and preliminary results. In *Supercomputing* '91, Proceedings of the 1991 ACM/IEEE Conference on, pages 158–165. IEEE, 1991.
- [5] A. Bhatele, K. Mohror, S. H. Langer, and K. E. Isaacs. There goes the neighborhood: performance degradation due to nearby jobs. In *Supercomputing'13, Proceedings* of the 2013 ACM/IEEE Conference on, page 41. ACM, 2013.
- [6] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De. Parameter variations and impact on circuits and microarchitecture. In *Proceedings of the* 40th annual Design Automation Conference, pages 338–342. ACM, 2003.
- [7] D. Chasapis, M. Casas, M. Moretó, M. Schulz, E. Ayguadé, J. Labarta, and M. Valero. Runtime-guided mitigation of manufacturing variability in power-constrained multi-socket numa nodes. In *Proceedings of the 2016 International Conference on Supercomputing*, page 5. ACM, 2016.
- [8] J. J. Dongarra, J. Du Croz, S. Hammarling, and I. S. Duff. A set of level 3 basic linear algebra subprograms. ACM Transactions on Mathematical Software, pages 1–17, 1990.
- [9] N. Gholkar, F. Mueller, and B. Rountree. Power tuning HPC jobs on powerconstrained systems. In Proceedings of the 2016 International Conference on Parallel Architectures and Compilation, pages 179–191. ACM, 2016.
- [10] D. Hackenberg, R. Oldenburg, D. Molka, and R. Schöne. Introducing firestarter: A processor stress test utility. In *Green Computing Conference, 2013 International*, pages 1–9. IEEE, 2013.
- [11] S. Herbert and D. Marculescu. Variation-aware dynamic voltage/frequency scaling. In High Performance Computer Architecture, 2009. HPCA 2009. IEEE 15th International Symposium on, pages 301–312. IEEE, 2009.
- [12] Y. Inadomi, T. Patki, K. Inoue, M. Aoyagi, B. Rountree, M. Schulz, D. Lowenthal, Y. Wada, K. Fukazawa, M. Ueda, et al. Analyzing and mitigating the impact of manufacturing variability in power-constrained supercomputing. In *Supercomputing* 15, Proceedings of the ACM/IEEE Conference on, page 78. ACM, 2015.
- [13] E. A. Leon, I. Karlin, and A. T. Moody. System noise revisited: Enabling application scalability and reproducibility with smt. In 2016 IEEE International Parallel and Distributed Processing Symposium, pages 596–607, May 2016.
- [14] A. Marathe, H. Gahvari, J.-S. Yeom, and A. Bhatele. Libpowermon: A lightweight profiling framework to profile program context and system-level metrics. In *Parallel and Distributed Processing Symposium Workshops, 2016 IEEE International*, pages 1132–1141. IEEE, 2016.
- [15] J. D. McCalpin. Memory bandwidth and machine balance in current high performance computers. IEEE Computer Society Technical Committee on Computer Architecture Newsletter, pages 19–25, Dec. 1995.
- [16] E. Rosenthal, E. A. León, and A. T. Moody. Mitigating system noise with simultaneous multi-threading. *Proceedings of SC13 (poster)*, 2013.
- [17] B. Rountree, D. H. Ahn, B. R. De Supinski, D. K. Lowenthal, and M. Schulz. Beyond dvfs: A first look at performance under a hardware-enforced power bound. In Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012 IEEE 26th International, pages 947–953. IEEE, 2012.
- [18] J. Schuchart, D. Hackenberg, R. Schöne, T. Ilsche, R. Nagappan, and M. K. Patterson. The shift from processor power consumption to performance variations: fundamental implications at scale. *Computer Science-Research and Development*, 31(4):197–205, 2016.
- [19] K. Shoga, B. Rountree, M. Schulz, and J. Shafer. Whitelisting msrs with msr-safe.
- [20] H. Shoukourian, T. Wilde, H. Huber, and A. Bode. Analysis of the efficiency characteristics of the first high-temperature direct liquid cooled petascale supercomputer and its cooling infrastructure. *Journal of Parallel and Dist. Computing*, pages 87–100, 2017.
- [21] R. Teodorescu and J. Torrellas. Variation-aware application scheduling and power management for chip multiprocessors. In Computer Architecture, 2008. 35th International Symposium on, pages 363–374. IEEE.
- [22] O. Tuncer, E. Ates, Y. Zhang, A. Turk, J. Brandt, V. J. Leung, M. Egele, and A. K. Coskun. Diagnosing performance variations in HPC applications using machine learning. In *International Supercomputing Conference*, pages 355–373, 2017.
- [23] L. Zhang, L. S. Bai, R. P. Dick, L. Shang, and R. Joseph. Process variation characterization of chip-level multiprocessors. In *Proceedings of the 46th Annual Design Automation Conference*, pages 694–697. ACM, 2009.