Adaptive Model-driven Facility-wide Management of Energy Efficiency and Reliability

Ananta Tiwari[†] Michael A. Laurenzano[†] Adam Jundt[†]
William A. Ward, Jr.[*] Roy Campbell[*] Laura Carrington[†]

[†]EP Analytics, Inc.
[*]DoD High Performance Computing Modernization Program

Abstract

We present the blueprint for the Energy Efficiency Management Platform (E2MP), a power-aware, green computing technology that can enact performance-neutral power and reliability management policies on high performance computing (HPC) centers. E2MP's design allows it to take a system-wide, holistic view of power and reliability management and dynamically make fine-grain power and thermal adaptations at the compute-node-level and at the facility-level in response to the behavior of the applications running in the facility.

E2MP continuously monitors a number of important metrics, including chip temperature, instantaneous per-component power draw and ambient room temperature, then relates those metrics via predictive models to specific application software behavior (e.g., quantity of main-memory traffic) and uses those relationships to steer systems towards better energy efficiency and reliability.

1. Background

Utility costs, constraints on power delivery and system reliability issues are limiting the expansion of large-scale HPC computing systems. Developing energy efficient and reliable modes of operation for these systems while meeting application performance targets has the potential to reduce the cost of operation, reduce carbon emissions, and enable the same science to be completed with substantially less energy. Towards that end, we present the design for the Energy Efficiency Management Platform (E2MP), a power-aware, green computing technology for managing energy consumption and mitigating reliability concerns on HPC systems while minimizing performance disturbance. E2MP is designed to take a facility-wide, holistic view and dynamically adapt the system's behavior by exercising fine-grained control over system-wide parameters affecting power consumption, reliability and performance. The control strategies are informed by predictive power, thermal and performance models as well as characterizations of applications running on the system.

The design and operation of E2MP is founded on our position that the operational settings that increase energy efficiency and reliability of a given HPC facility are not only governed by the capabilities and specifications of the computing and cooling hardware, but are also complex functions of the application software running in that facility. To take advantage of this and reduce energy consumption, E2MP will enact performance-neutral power adaptations to two main power consumers in HPC facilities – compute-nodes and facility's cooling infrastructure. Compute-node power adaptation will come in the form of processor and memory clock frequency adjustments and fan rotation speed throttling. Cooling power adaptation will come in the form of controlling compute-node temperature targets and adapting the set-point temperature of the Computer Room Air Conditioning (CRAC) units.

From the standpoint of reliability, over-heating of components is frequently cited as the cause for component failures [5, 11, 3]. By utilizing application characterization-based models to predict the level of thermal stress large-scale applications put on different components of the system, E2MP will be able to provide operational intelligence to avoid thermal hot-spots, and therefore improve reliability. This information can be used to drive reliability-aware management strategies such as scheduling "hot" applications on nodes that are closer to the rack/system fans.

2. Managing Energy Efficiency

Energy consumption in a large-scale computing facility is influenced by the complex interactions between a number of infrastructure components. Without an integrated, coordinated perspective on reducing energy consumption, local actions taken by one component with the intent to lower energy consumption can actually have the opposite effect on another component, thereby cancelling out the net effect. For example, increasing the temperature set-point of the CRAC to save cooling energy can lead to increased compute-node fan power and increased chip leakage power [9].

One of the major challenges in enacting energy management policies is meeting performance targets, such as time-to-solution or throughput goals, a top priority to HPC system designers and operators.

2.1. Compute-Node Power Optimization

E2MP will base its compute-node-level power management decisions on predictive models that provide information on performance, power and thermal profiles of large-scale applications. Certain subsystems within a node may be over-powered at any given point in time. For example, memory- or network-intensive workloads may leave CPUs starved for work, while compute-intensive workloads may not need to exercise the full capability of the memory subsystem or network. Isolating the properties of application software that manifest themselves as partially idle subsystems can be used by E2MP to control hardware "knobs" (e.g., CPU clock frequency and memory bus frequency) that reduce the capability (at a minimal performance loss) and the power consumed by those starved subsystems.

An often overlooked component of compute-node power is the power consumed by node fans, which can be as high as 14%[1] of the total power consumed by the node.

The compute-node fans at HPC centers most often are run on full-speed to make sure the processors do not reach the maximum thermal limit. Even when the control of the speed is given to the operating system, the control mechanisms are overly coarse-grained in terms of how many fan speed settings are explored. Various strategies have been proposed [16, 10] to determine the optimal fan speeds while also taking into account the exponential rise in static leakage power associated with rising chip temperature. However, prior work overlook the fact that the properties of computations running on the system also have a significant impact on the thermal state of the chips. We argue that a more sophisticated understanding of how much thermal stress an application puts on the chip (via predictive models based on application characterizations), together with a clear understanding of the temperature-induced increase in leakage power can be used as a guide to control the fan speed at much finer granularities.

2.2. Facility-wide Power Optimization

In large-scale computing facilities, nearly equaling the power consumption of the servers themselves is the power required to dissipate the heat generated by those servers. This power is transacted in the form of heat exchangers used to generate cool air that is pumped into the facility. Current practices in compute facility operation use an ambient or set-point temperature in the range of 64°F to 80°F [1]. Many in the industry, including Google [4], have suggested raising the set-point temperature to higher than 80°F to save on cooling power. Often lost in these debates is the associated rise in server fan power, which at higher set-points will increase because server fans now have to dissipate more heat. Furthermore, increased set-points will also lead to a rise in leakage power, which is projected to rise further with shrinking feature sizes in the next-generation processors. Therefore, optimal set-point determination will have to take these additional factors into consideration and, therefore, is a multi-objective optimization problem. E2MP will utilize models that inform the thermal impact of different applications on chip temperatures, the impact of temperature on the static leakage power and computation-aware optimal CPU, memory and fan speeds to navigate this complex optimization space.

If the processor and DRAM specifications allow, then specialized strategies that trade a small increase in leakage power (by letting the chips and memory modules operate at higher temperatures) for a larger decrease in fan and CRAC power can also be explored. However, this type of trade-off analysis has to carefully also consider the elevated failure rates associated with higher operating temperatures.

3. Managing Reliability

Research in computer reliability conducted at massively large datacenters such as Google and Microsoft show that hard disk drives (HDDs) and DIMMS are the two components that most commonly break down in large-scale computing facilities. Sankar and El-Sayed [11, 3] show that HDD reliability depends significantly on temperature, whereas temperature has a smaller impact on DIMM error rates in the field [12]. The need for the coordinated approach taken by E2MP to manage both energy efficiency and reliability is further illustrated by the fact that the node-level fan speed and its associated power, hardware temperature, leakage power, ambient temperature and computational characteristics of applications running on the system are all inter-connected (e.g., reducing the fan speed to reduce power will result in higher chip temperature and therefore, more leakage power and potentially higher failure rates).

Predictive models that utilize application characterizations to construct a thermal profile of the whole application can be used to make system adaptation decisions. For example, if the models inform E2MP that the upcoming phases[2] in an application are "hot", then E2MP can enact one of the various policies available to make sure that the chip (or any other system sub-component) temperature does not exceed a stipulated limit – e.g., preemptive CPU clock frequency throttling or application migration. Temperature limits can be set either based on the history of component failures or on temperature-based component failure probability models.

4. Conclusion

We presented a sketch of E2MP, a framework for enacting facility-wide power and reliability optimizations. E2MP's approach is novel in that it uses application-level behavior to guide its optimization strategies. Application characterization-based models that can predict power, performance and thermal profiles of large-scale applications are at the center of the proposed framework. Developing tools to automate the collection of data for application characterizations; using the characterizations to construct aforementioned models; using the models to inform procurements, application tuning for both performance and energy efficiency, and performance bottleneck detection and mitigation are key areas expertise of our team [7, 15, 8, 14, 13, 6, 2]. This expertise will be vital to the successful development and deployment of the E2MP framework.

[1] Based on in-house measurements on an Intel Sandy Bridge based dual-socket server node.

[2] We define application phases as contiguous regions of an application that exhibit similar behavior or put the same stress in a given system component.

References

[1] ASHRAE Technical Committee. 2011 thermal guidelines for data processing environments - expanded data center classes and usage guidance. http://ecoinfo.cnrs.fr/IMG/pdf/ashrae_2011_thermal_guidelines_data_center.pdf.

[2] A. D. Breslow, A. Tiwari, M. Schulz, L. Carrington, L. Tang, and J. Mars. Enabling fair pricing on hpc systems with node sharing. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, SC '13, pages 37:1–37:12, New York, NY, USA, 2013. ACM.

[3] N. El-Sayed, I. Stefanovici, G. Amvrosiadis, A. A. Hwang, and B. Schroeder. Temperature management in datacenters: Cranking up the thermostat without feeling the heat. *USENIX*, 38(1), February 2013.

[4] Google. Efficiency: How we do it. https://www.google.com/about/datacenters/efficiency/internal/.

[5] C.-h. Hsu and W.-c. Feng. A power-aware run-time system for high-performance computing. In *Proceedings of the 2005 ACM/IEEE Conference on Supercomputing*, SC '05, pages 1–, Washington, DC, USA, 2005. IEEE Computer Society.

[6] M. A. Laurenzano, M. Meswani, L. Carrington, A. Snavely, M. M. Tikir, and S. Poole. Reducing energy usage with memory and computation-aware dynamic frequency scaling. In *Proceedings of the 17th international conference on Parallel processing*, Euro-Par'11, pages 79–90, 2011.

[7] M. A. Laurenzano, J. Peraza, A. Tiwari, L. Carrington, W. Ward, and R. Campbell. A static binary instrumentation threading model for fast memory trace collection. *International Workshop on Data-Intensive Scalable Computing Systems*, 2012.

[8] M. A. Laurenzano, A. Tiwari, A. Jundt, J. Peraza, W. Ward, R. Campbell, and L. Carrington. Characterizing the performance-energy tradeoff of small arm cores in hpc computation. In *Proceedings of the 20th international conference on Parallel Processing*, Euro-Par'14. Springer-Verlag, 2014.

[9] M. Patterson. The effect of data center temperature on energy efficiency. In *Thermal and Thermomechanical Phenomena in Electronic Systems, 2008. ITHERM 2008. 11th Intersociety Conference on*, pages 1167–1174, May 2008.

[10] B. Pradelle, N. Triquenaux, J. Beyler, and W. Jalby. Energy-centric dynamic fan control. *Computer Science - Research and Development*, pages 1–9, 2013.

[11] S. Sankar, M. Shaw, K. Vaid, and S. Gurumurthi. Datacenter scale evaluation of the impact of temperature on hard disk drive failures. *Trans. Storage*, 9(2):6:1–6:24, July 2013.

[12] B. Schroeder, E. Pinheiro, and W.-D. Weber. Dram errors in the wild: A large-scale field study. In *Proceedings of the Eleventh International Joint Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '09, pages 193–204, New York, NY, USA, 2009. ACM.

[13] A. Tiwari, M. Laurenzano, L. Carrington, and A. Snavely. Auto-tuning for energy usage in scientific applications. In *Proceedings of The Fifth Workshop Productivity and Performance 2011*, PROPER '11, 2011.

[14] A. Tiwari, M. Laurenzano, L. Carrington, and A. Snavely. Modeling power and energy usage of hpc kernels. In *Proceedings of the Eighth Workshop on High-Performance, Power-Aware Computing 2012*, HPPAC '12, 2012.

[15] A. Tiwari, M. Laurenzano, J. Peraza, L. Carrington, and A. Snavely. Green queue: Customized large-scale clock frequency scaling. In *Cloud and Green Computing (CGC), 2012 Second International Conference on*, pages 260–267, Nov 2012.

[16] M. Zapater, J. L. Ayala, J. M. Moya, K. Vaidyanathan, K. Gross, and A. K. Coskun. Leakage and temperature aware server control for improving energy efficiency in data centers. In *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '13, pages 266–269, San Jose, CA, USA, 2013. EDA Consortium.