ENERGY EFFICIENT COMPUTING USING CONTINUOUS TELEMETRY HARNESS

Ayse K. Coskun* Tajana Simunic Rosing* Kenny Gross[+]
*{acoskun, tajana}@cs.ucsd.edu – [+]kenny.gross@sun.com
*University of California, San Diego
[+]Sun Microsystems

Abstract

This work illustrates how Continuous System Telemetry Harness tracks thermal and energy dynamics inside a real-life server, enabling efficient thermal / energy management and intelligent fan control.

1. Introduction

Elevation in power consumption and temperature in computer chips has brought new challenges in reliability, performance, cooling costs and leakage power. Conventional workload scheduling and dynamic management techniques cannot always mitigate energy or temperature related problems at low cost. To increase system lifetime and reliability in a cost effective way, this project utilizes the Continuous System Telemetry Harness (CSTH) software innovation [1], which monitors temperatures, voltages, currents, and power dynamics versus load throughout the interior of computer servers. Using CSTH, we are able to reduce energy cost and extend system lifetime by a set of telemetry-enabled innovations, including enhancing the operating system scheduler for energy- and temperature-aware job scheduling [1], utilizing dynamically adaptive energy- and thermal-management methods [2], and applying Intelligent Fan Control [SMI patent pending].

The energy efficiency challenges are even more severe for data centers with massive size and computing capacity. This work discusses how to utilize the telemetry software for an intelligent load provisioning scheme to achieve energy- and thermal-aware scheduling, which result in more efficient computations, reduce the cooling burden, reduce the thermal gradients inside servers, and improve long term system reliability. In the demo, we utilize a 3D visualization utility that illustrates thermal and energy dynamics inside a real-life server during normal workload dynamics. We also show that Intelligent Fan Control is able to save a significant percent of energy by utilizing telemetry to monitor the temperature and then adjusting the fan speed to match thermal flux dynamics (regardless of the altitude of the data center, the local ambient temperature or the instantaneous load dynamics).

2. Continuous System Telemetry Harness

Integrating energy and thermal management into the OS requires fusing together data collected from temperature sensors in the system, in addition to performance metrics related to the workload and processors. For the OS to make correct dynamic management decisions, it must take these disparate telemetry streams and reconstruct a logically consistent snapshot of the system state. Misaligned telemetry streams, in which correlated events occur out of phase, may cause instabilities in the management and scheduling decisions, and lead to suboptimal results. Prior work in temperature-aware scheduling has focused exclusively on simulation results, with little attention paid to how the proposed solutions would be implemented in practice.

Continuous System Telemetry Harness (CSTH), developed by Sun Microsystems, systematically addresses the issues related to the collection, preprocessing, and analysis of time-series telemetry data [3]. It captures telemetry data from a variety of physical and logical sensors in the system. Physical sensors include distributed temperatures, voltages, currents, humidity, and vibration; and logical sensors include OS metrics, hardware performance counters, and quality-of-service metrics. CSTH is integrated into the existing software stack, and introduces negligible additional overhead.

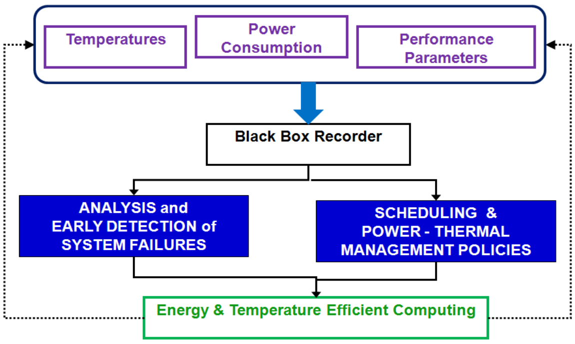

Figure 1: Energy and Temperature Efficient Computing Using Continuous System Telemetry Harness.

The CSTH mitigates the challenges of large scale remote monitoring schemes by providing a real-time telemetry architecture with a circular-file repository that acts as a system "black box" performance monitor. All monitored variables go to a novel, dual-stage circular file that has a very compact memory footprint. CSTH stores the variables for not only dynamic management, but also for analyzing and diagnosing system failures. During most of the time, the systems are performing without any problems and the new incoming signals simply overwrite the previously recorded signals. If an anomaly is detected in any monitored variable, then the dual-stage circular file is automatically compressed and frozen into persistent memory. Alternatively, if a systems analyst wants to access the data for validation of new techniques or for evaluating hypotheses about subtle phenomena that do not trip the anomaly thresholds, she can do a manual analysis of the

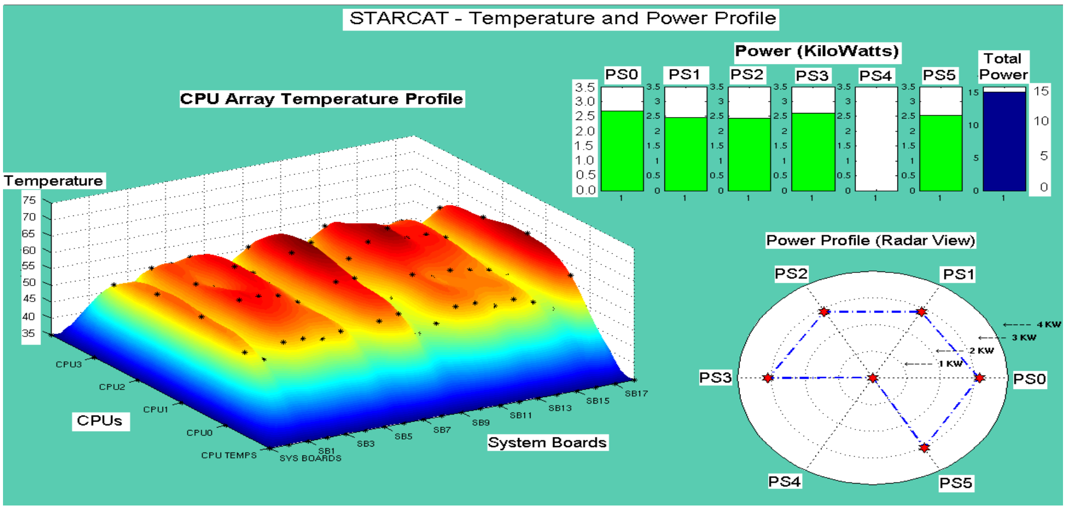

Figure 2. Temperature and power profile after the power supply failure (see PS4).

telemetry file at any time. Similarly, telemetry data can be easily passed to the OS at regular intervals for energy- or temperature-aware job scheduling. Figure 1 demonstrates the operational flow chart of a system with CSTH.

CSTH enables enhancements to availability, serviceability, performance, capacity planning, quality of service, and security; but without placing burdens on the monitoring infrastructure during the majority of the time when systems are behaving without problems.

Next, we show an example of failure detection with CSTH. Figure 2 shows the dynamic temperature and power consumption state of a SunFire 15K platform, collected by CSTH after a power supply failure. CSTH captured a power supply failure event (see PS 4 in the figure), in which the telemetry shows that the load is automatically shared across the five remaining functioning power supply units with no interruption in workload dynamics.

3. Intelligent Fan Control

In today's data centers, a significant portion of energy cost comes from cooling. In this section, we discuss how the temperature and power consumption data collected by CSTH can be utilized for Intelligent Fan Control (IFC). The goal of IFC is to reduce the energy cost associated with the fans. To accomplish this, IFC adjusts fan speeds to the level where the CPU temperatures (tracked by the telemetry harness) are at the upper edge of their "comfort zone". This way, while the temperature is guaranteed to remain below thresholds that would create reliability issues, we are able to save significant energy. In addition, reducing the fan speed makes the servers a lot quieter.

Figure 3 provides a screenshot of the 3D visual utility running on a Sun SPARC T5440 server, demonstrating the temperature profile and the total power consumption of the system before and after IFC is activated. By reducing the fan speed to 1500 rpm, IFC is able to reduce the overall power consumption by 200 to 300 Watts while keeping the temperature below 80^{o}C.

4. Conclusion

This work discusses how to utilize Continuous System Telemetry for reliable and energy efficient computing. In addition to discussing CSTH-enabled energy and thermal management, we have provided two examples of CSTH implementations on real-life systems: 1) Monitoring and detecting failures in servers; 2) Intelligent Fan Control for preventing over-cooling and saving energy in data centers.

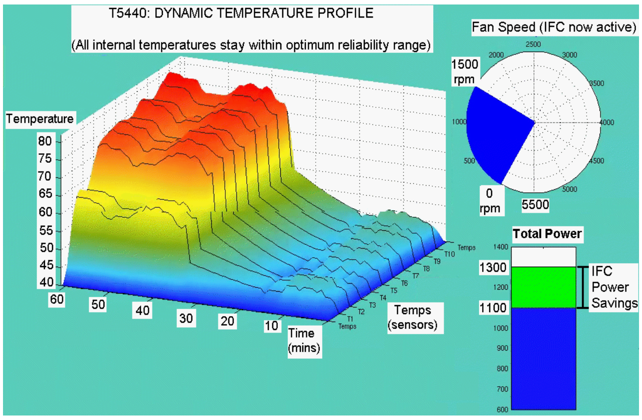

Figure 3. Energy savings achieved by intelligent fan control on a T5440 server.

5. References

[1] Ayse K. Coskun, Tajana Simunic Rosing, Keith Whisnant and Kenny Gross. "Static and Dynamic Temperature-Aware Scheduling for Multiprocessor SoCs". In *IEEE Transactions on VLSI, vol.16 no.9, pp. 1127-1140,* Sept. 2008.

[2] Ayse K. Coskun, Tajana Simunic Rosing and Kenny Gross. "Proactive Temperature Balancing for Low Cost Thermal Management in MPSoCs". In *ICCAD, pp. 250-257*, 2008.

[3] K. Gross, K. Whisnant, and A. Urmanov. "Electronic prognostics through continuous system telemetry". In *MFPT, pages 53-62,* April 2006.

Acknowledgement

The authors wish to thank Kalyan Vaidyanathan, Ram Dhanekula, and Aleksey Urmanov, lead innovators on the System Dynamics Characterization and Control team at Sun Microsystems, San Diego.