A Thermal Management and Profiling Method for Reconfigurable Hardware Applications;
by Phillip H. Jones, John W. Lockwood, Young H. Cho;
16th Annual Conference on Field Programmable Logic and Applications (FPL);
Madrid, Spain, August 28-30, 2006; pp. 103-109.

A THERMAL MANAGEMENT AND PROFILING METHOD FOR RECONFIGURABLE HARDWARE APPLICATIONS *

Phillip H. Jones, John W. Lockwood, Young H. Cho

Applied Research Laboratory
Washington University
St. Louis, MO
email: phjones@arl.wustl.edu, lockwood@arl.wustl.edu, younghc@gmail.com

ABSTRACT

Given large circuit sizes, high clock frequencies, and possibly extreme operating environments, Field Programmable Gate Arrays (FPGAs) are capable of heating beyond their designed thermal limits. As new circuits are developed for FPGAs and deployed remotely, engineers are challenged to determine in advance if the device will operate within recommended thermal ranges. The amount of power consumed by the circuit depends on how an algorithm is compiled into hardware, how the circuit is placed and routed, and the patterns of data that pass through the system. The amount of heat that can be dissipated depends on the thermal transfer characteristics of the package, the air flow that passes over the package, and the ambient temperature of the remote systems. Rather than designing a system to handle unreasonable worst-case situations, we have implemented a thermal management system that continuously monitors the temperature of the FPGA and reprograms the device if the temperate approaches the outer limits of safe operating conditions. Our system measures the junction temperature of a Xilinx Virtex FPGA using a built-in thermal diode. Using the temperature monitoring mechanism, we have studied the steady-state and transient conditions of multiple benchmark circuits implemented in an FPGA logic on the Field-programmable Port Extender (FPX) development platform. We observed properties of these benchmark circuits that enable us to predict power and thermal characteristics for real applications. We propose a Dynamic Thermal Management (DTM) strategy for FPGAs based on temperature feedback.

1. INTRODUCTION

FPGAs provide the flexibility to deploy new circuits into remotely deployed systems. Predicting the power consumption and thermal behavior of circuits in different operating conditions, however, is difficult. Power consumption can be estimated using tools such as Xpower [1]. Xpower profiles power utilization using simulation. The actual power utilization depends on actual data patterns that pass through the system. Thermal considerations of the system can be estimated by analysis of the package and system hardware. Although such estimates may be valuable in characterizing typical situations, the actual thermal conditions heavily depend on external factors such as the ambient temperature and available airflow.

2. MOTIVATION

It is possible to design FPGA circuits that generate more heat than the package and platform can dissipate. This issue is of major concern for reconfigurable platforms like the Field Programmable Port Extender (FPX) developed at Washington University. The FPX platform contains two FPGAs: a small Xilinx Virtex FPGA called the Network Interface Device (NID) that is configured with a static bitfile, and a larger Xilinx Virtex FPGA called the Reconfigurable Application Device (RAD) that is reconfigured over the network [2]. New bitfiles are sent to the NID over the network as modules that reconfigure the RAD to implement new network processing functions [3].

Over sixty modules have been developed and deployed on the FPX platform [4]. Until recently, none of these circuits consumed more power than the FPX could dissipate. Recently, however, a RAD module was built and operated in a situation such that the device generated more heat than the platform could dissipate. The module was deployed onto a RAD that had no heat sink, and the FPX was mounted in a chassis without a cover which would have otherwise guided airflow across the device.

The result of this scenario was a chain of events that damaged an FPX platform. Retrospectively, we determined that the circuit loaded into the RAD overheated. Without a heatsink and sufficient air-flow, the FPGA overheated because it was not able to dissipate the power generated by the computationally-intensive circuit programmed into the de-

*SPONSORED BY NATIONAL SCIENCE FOUNDATION UNDER GRANT ITR 0313203.

vice. As the temperature increased beyond the maximum temperature range of the device (85 degrees C), the Silicon heated the package and the FPX printed circuit board to a point where thermal expansion caused warping of the printed circuit board. The warping became sufficiently intense so as to cause a power plane to short with another power plane that had a voltage with opposite polarity. The result was excessive current through a power connector that caused additional heating that eventually melted the plastic around the connector. The top view, and side view of the affected FPX board are shown in Figures 1 and 2.

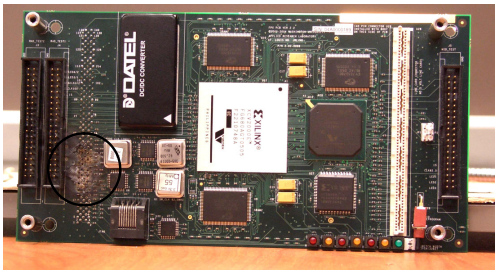

Fig. 1. Top view of the affected printed circuit board. The affected connector pins can be seen in the lower-left

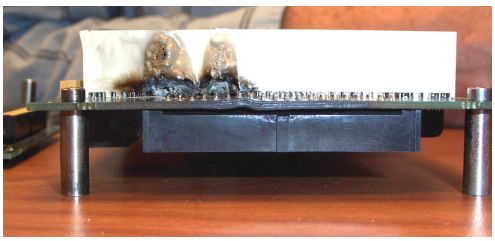

Fig. 2. Side view: Note how layers of the board were warped because of heat

The FPX platform had always included the ability to monitor and manage thermal conditions [1]. A Maximum 1618 temperature sampling device installed on the FPX converts the value of the current passing through the RAD's thermal diode to a digitized value that can be read over a two-wire serial bus [5]. The NID on the FPX could then read that value to determine the temperature of the RAD. In this work, we describe how a new circuit implemented on the NID allows monitoring of thermal conditions and enforces safe operation of the systems.

The rest of this paper is structured as follows. Section 3 gives a summary of the related work in the field of Thermal and Power management. Section 4 describes the setup for gathering thermal data and implementing an autonomous thermal shutdown safety circuit for the FPX development platform. Section 5 combines the temperature measurement mechanism with custom designed thermal/power benchmark circuits to profile the RAD FPGA. Section 6 summarizes the paper and suggests avenues of potential future work.

[1]Thanks are due to Bill Carter of Xilinx for the suggestion to include thermal measurement circuitry on the FPX platform

3. RELATED WORK

Microprocessors have been built that allow their voltage and frequency to be scaled to extend battery life of mobile computers. Companies like Intel and AMD extend this concept to manage heat dissipation on servers [6]. By introducing power management features, software running on the CPU can scale voltage and frequency to lower power usage before the device overheats. This technology is critical for servers located in large data centers that house hundreds and even thousands of compute nodes.

Low power embedded processors like Xscale [7] also have hooks that allow voltage and frequency scaling to increase power and thermal efficiency. Work presented by [8] makes use of these features to present a dynamic thermal management (DTM) system that would scale the processor frequency in response to temperature readings from an external thermal couple.

Xilinx Virtex FPGAs embed a sense diode for measuring the junction temperature (T_j) of the device. Thermal management issues have become a prominent issue for reconfigurable systems. The dynamic thermal management circuit described in this paper makes use of the sense diode to ensure safe operation of the FPX platform and to characterize applications programmed into the RAD on the FPX platform. Data read from the embedded sense diode can also be be used to regulate fan speed or allow the system to perform other mechanisms that would cool down the device.

3.1. Power Measurement

Xilinx integrates power estimation tools in their simulation suite. These tools enable comparison of the relative power used between circuit implementations. These tools generate results that differ greatly from physical measurements [9].

One method for physically measuring power consumed by circuits involves inserting small sense resistors in series with the circuit power supply. This is the method we use to make power measurements. Lee [9] presents a novel method for making cycle accurate power measurements that makes use of the concept of switching capacitance.

3.2. Thermal Measurement

Most tools simulate thermals at the board level. MIT researched tools that simulate thermals at the device level [10]. But this work was focused on reliability aspects such as electron migration, and not for estimating thermal dissipation of power from reconfigurable applications.

Lopez-Buedo [11] describe techniques to make physical temperature measurements. These techniques use external thermal couples, thermal imaging cameras, and sense diodes embedded into the die. A novel idea is presented for reconfigurable devices that configures ring oscillators in order to infer temperature measurement from changes in oscillator frequencies. This work is later extended in [12] using arrays of such oscillators to detect hot spots and thermal gradients in FPGAs.

4. IMPLEMENTATION

4.1. Thermal Shutdown Circuit on the FPX

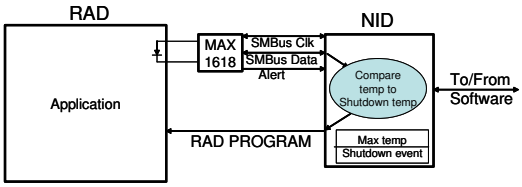

Fig. 3. Shutdown Circuit Architecture

Figure 3 shows the high-level diagram of the thermal shutdown circuit as it is implemented on the FPX platform. The RAD contains a built-in thermal diode in the silicon of the Xilinx Virtex 2000E FPGA. The anode and cathode of this diode are accessed through I/O pins on the 680-pin package. The current passing through these pins is then used to measured the junction temperature of the silicon.

The FPX platform includes an on board temperature measurement chip (Maxim 1618) to compute the T_j of the silicon from the current generated by the sense diode. The MAX1618 collects temperature samples at rate of 16 samples/second and communicates temperature readings using the SMBus serial protocol. These bus signals are routed to test pins on the FPX that are monitored by the NID (the statically-programmed control and configuration FPGA on the FPX platform). The temperature of the RAD can be monitored externally by sending a query message over the network to the NID. The NID responds with a status message that includes the temperature of the RAD.

The NID also compares the temperature of the RAD to a preset value to ensure that the RAD always operates in a safe mode. If the temperature approaches the threshold of safe operation, the NID can instantly and automatically issue a command through the SelectMAP interface of the RAD to reprogram the device. This command clears the RAD configuration memory thereby reinitializing the device.

Figure 4 shows a plot of temperature over time for a circuit that causes the RAD to surpass a thermal threshold of 70 C. As can be seen near time=400 seconds, the temperature exceeded the threshold and the thermal shutdown circuit reprogrammed the RAD. Thereafter, the RAD's temperature rapidly drops and eventually reaches it's idle (i.e. no configuration) temperature of 32 C.

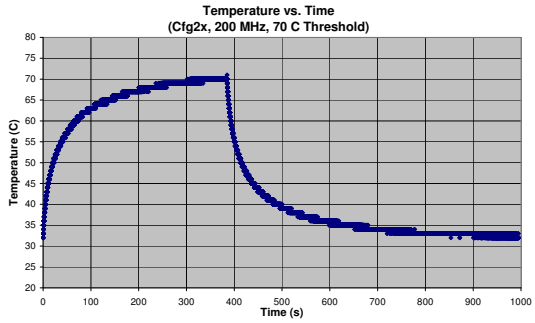

Fig. 4. FPGA being reprogrammed by the Shutdown Circuit upon reaching a thermal threshold of 70 C

The status cells returned by the NID allow an external PC to remotely query the temperature of the RAD. The remote PC can be located anywhere, just so long as there is a path through the network where query and status messages can flow. For the thermal profiling experiments ran for this paper the external software made temperature update requests approximately every 50 ms, and logged responses to a text file along with a timestamp. Section 5 makes use of these log files to help perform a thermal characterization of the RAD.

4.2. Accounting for Noise

Temperature measurements on sense wires are sensitive to crosstalk by signal transitions on nearby high speed data buses. Temperature measurement errors were observed on the FPX platform due to such crosstalk. The SRAM2 memory module (one of the four memory modules attached to the RAD) used address and data I/O pins that were located near the two wires connecting the thermal sense diode. Interference from data signal transitions caused the temperature reading to appear +30 C too high when the SRAM2 memory module was heavily utilized. In extreme cases, the temperature spiked by +60 C.

Two steps were taken to mitigate the effects of the crosstalk. First, a 2200 pF capacitor was added between the current sensing inputs of the MAX 1618 that connected to the Virtex FPGA. Second, the slew rate of the SRAM2 I/O pins were configured from fast to slow slew rate. The addition of the capacitor (as recommended by the MAX data sheet) removed the temperature spikes caused by crosstalk from the SRAM2 module and decreased the average temperature error from +30 C to +20 C. The slower slew rate on the SRAM2 I/O signals decreased the coupling current effect and lowered the temperature error to +9 C.

4.3. Thermal Benchmark Circuit Description

Benchmark circuits were created to characterize the thermal behavior of the RAD. These circuits were designed to include four important characteristics. First, the benchmark circuits scale to operate over a wide frequency range. Second, the benchmark circuits scale in the number of on-chip resources utilized. Third, the circuits use a regular structure that can be readily analyzed. Fourth, the circuits are placed and routed evenly over the device to avoid creation of a large temperature gradient across the device.

The core building block of the benchmark circuits is shown in Figure 5. This circuit is made up of an 8x6 array of LUTs (Look Up Tables) each of which is followed by a DFF (D Flip-Flop). The LUTs are configured to be 4-input AND gates. Core blocks construct the smallest workload used in this paper call a Thermal Workload Unit. Figure 6 shows the structure of a Thermal Workload Unit. It is made up of two parts 1.) Input Generator 2.) Computation Row. The Input Generator is responsible for toggling the inputs of the Computation Row every clock cycle. The Computation Row is made up of an array of 18 Core Blocks. A single workload unit uses 876 LUTs and 876 DFFs (2.25%) of a the VirtexE2000(RAD) resources, approximately two rows of the FPGA. Xilinx specific VHDL directives allowed very controlled a precise layout. This allows for creation of high-speed Thermal Workload Units that can be evenly distributed over the FPGA. Figure 7 gives the FPGA resources used for each benchmark. Different configurations of the benchmark circuits use between 9% and 90% of the chip resources. The circuits placed and routed to operate at frequencies between 370-407 MHz. These circuits satisfy the desired properties of scalability with respect to circuit size and operating frequency. The variation in the maximum frequency was a result of variations in routing. Specifying that the routing tool try harder to maximize the frequency could allow all benchmark circuits to run at 400 MHz. The benchmark circuits had additional characteristics. First, each benchmark circuit has an equal number of LUTs and D Flip-flops. Second, each benchmark circuit uses a multiple of the Cfg1x benchmark resources. Third, the output of every LUT and DFF within a benchmark toggles every clock cycle (i.e. 100% activity rate), producing the maximum thermal heating for the chip resources used.

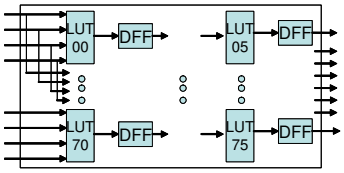

Fig. 5. Core Block: 8x6 array of pipelined LUTs configured as 4 input AND gates

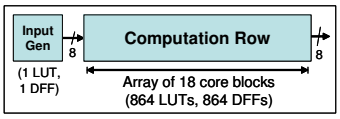

Fig. 6. Thermal Workload Unit

Cfg	1x	2x	3x	4x	5x	6x	7x	8x	9x	10x
LUTS	9%	18%	27%	36%	45%	54%	63%	72%	81%	90%
FF	9%	18%	27%	36%	45%	54%	63%	72%	81%	90%
Block RAM	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Fig. 7. Benchmark FPGA resource usage

5. BENCHMARK RESULTS AND ANALYSIS

5.1. Steady State Temperature Measurements

Figure 8 shows the data collected from running a subset of the benchmark circuits from 10 - 200 MHz. Each experiment was run approximately 20 minutes, this was empirically found to be the amount of time needed for a given circuit to reach its maximum steady state temperature. Figure 9 and 10 show plots of this data for temperature verses chip resources and temperature verses frequency. The data shows that temperature has a linear relation with respect to chip resources and frequency. This is to be expected because the first order approximation of power dissipated by a circuit is given as $Power \sim CV^2F$ which implies a linear relation between power with respect to both capacitance and frequency. The capacitance of the circuit is proportional to circuit size. Also, the relation $Power = Energy/second$, and $\Delta Energy \sim c_p \Delta Temperature$, where c_p is a constant that characterizes the amount of heat that a system can dissipate, shows that the change in temperature is proportional to power. Between 100 and 200 MHz there is a bend in the T_j vs. Frequency plot. This appears to be a side affect of needing to use an internal 4x clock multiplier to step up the base input frequency of 50 MHz to 200 MHz. All other experiments use the base clock as is.

Measured Temperatures

Cfg	1x	2x	4x	8x	10x
10 MHz	26	27	29.5	34	36
25 MHz	29	31.5	37.5	48	53
50 MHz	33	38.5	49.5	70	80
100 MHz	41	52.5	72.5	112.5	132.5
200 MHz	50.5	71	111	191	231

Extrapolated Temperatures

Fig. 8. Junction Temperature (T_j) measured for each benchmark (degrees C), FPGA package thermal rating = 85 C

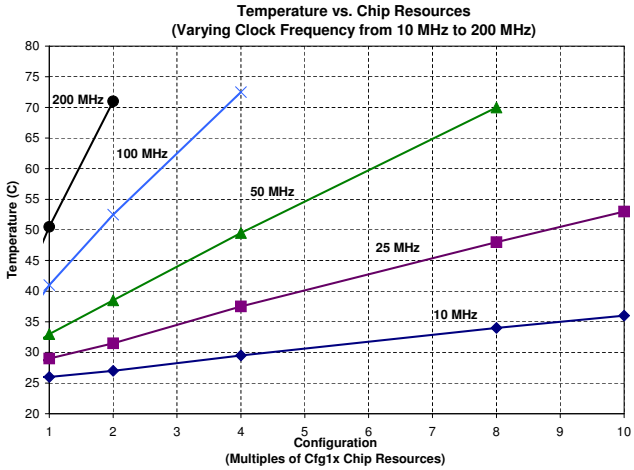

Fig. 9. T_j vs. FPGA resources for each benchmark

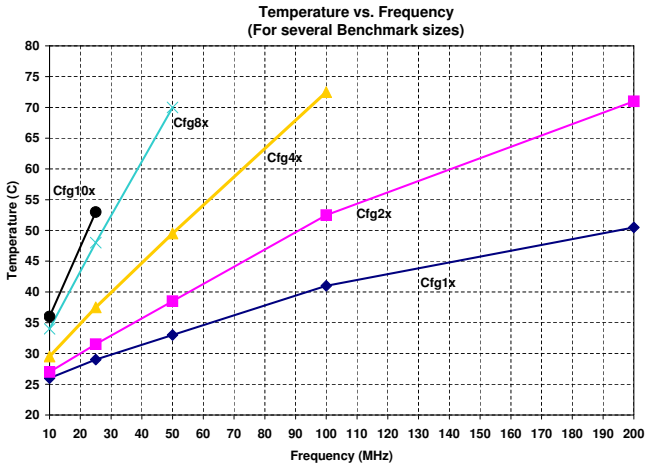

Fig. 10. T_j vs. Frequency for each benchmark

5.2. Transient Thermal Analysis

Since RAD temperature readings are updated quickly (50 millisecond intervals), the transient thermal behavior of the benchmark circuits can be examined. Figure 11 gives examples of the temperature profile of the RAD during the transition between idle operation and stead-state operation with two different benchmark circuits. The trajectory of the plots are characteristic of an exponential function in the general form of:

$$T(t) = Ae^{-t/\tau} + B \tag{1}$$

and the data from the upper measurement plot allows the the thermal time constant (τ) to be determined by solving the

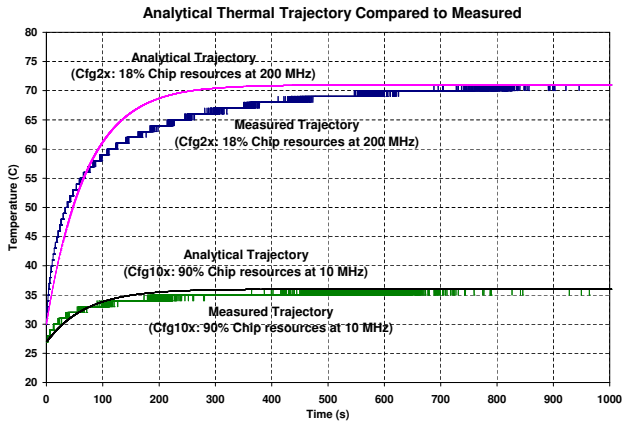

Fig. 11. Thermal Trajectory comparison: Measured vs. Analytical model

following equations:

$$T_j(\infty) = B \tag{2a}$$

$$T_j(0) = A + B; A = T_j(0) - B \tag{2b}$$

$$T_j(t = \tau) = Ae^{-1} + B = A * .386 + B \tag{2c}$$

The final steady-state temperature, $T_j(\infty)$ = B = 71. Thus A = 30 - 71 = -41. T_j(t=τ) is computed to be 55.8 C, which occurs at t=69.4 seconds, thus τ = 69.4 seconds.

Referring again to Figure 11, the analytical model of temperature vs. time using the above equation fits the measured data well for a first order approximation. Now that τ is known for the RAD, we can use the measured values of $T_j(0)$ and $T_j(\infty)$ to predict transition behavior for other configurations. As an example the analytical trajectory of the lower plot of Figure 11 accurately matches actual measurements of the temperature trajectory of the RAD running a different benchmark configuration.

It is important to note that there is an enormous difference between the clock period of the system (t_c=20 nanoseconds at 50 MHz) and the time constant (τ=69.4 seconds) that characterizes the rate at which the system heats. The thermal mass of the FPX platform allows the FPGA that implements the RAD to operate at high power for a relatively long time before the maximum operating temperature would be exceeded. The highest temperature transition speed observed for the benchmark circuits was approximately 10 degrees/second, and this rate slows exponentially as steady state is approached. If a circuit has multiple thermal operating modes, then intensive computation can be performed over an interval until a threshold is reached, then less intensive computation can be performed to allow the FPGA to cool while operating in a lower thermal mode. This is useful for applications that process at bursty rates. An application can temporarily operate in a high power mode that would

exceed the thermal limits of the device if allowed to reach steady-state.

5.3. Temperature verses Power

An FPX platform was modified to enable power consumption measurements by placing a .008 ohm sense resistor in series with the 1.8 V FPGA core power supply. Measuring voltage across this sense resistor enables calculating power consumption at any instance of time.

Figure 12 shows the power measurements corresponding to each thermal experiment run. As with the temperature measurements, data points were obtained from running each experiment for 20 minutes. Unlike the transient analysis for

Cfg	1x	2x	4x	8x	10x
10 MHz	.61	.81	1.17	1.89	2.25
25 MHz	1.08	1.58	2.52	4.30	5.24
50 MHz	1.91	2.88	4.77	8.35	
100 MHz	3.56	5.42	9.04		
200 MHz	4.93	8.55			

Fig. 12. Power(W) consumed by each Benchmark Configuration

temperature which exhibits swings of up to 40 C, power consumption stays fairly constant over time for large thermal transitions. The largest change in power consumption from the initial time of the experiment to the point where the circuit reached a steady state temperature was only 3.5%.

The Xilinx Device Package User Guide [13] gives the following equation:

$$\begin{aligned} T_J &= Power * \theta_{JA} + T_A \\ \theta_{JA} &= (T_J - T_A)/Power \end{aligned} \tag{3}$$

for estimating T_J for a given power consumption. Where θ_{JA} is the thermal resistance from junction to ambient and T_A is the ambient temperature. This equation can be first applied to the data collected in Figure 8 and 12 to calculate the empirical value of θ_{JA} for the RAD (Figure 13). The User Guide specifies the package used for the RAD

Cfg	1x	2x	4x	8x	10x
10 MHz	4.94	4.94	5.55	5.82	5.78
25 MHz	5.55	5.40	5.75	5.82	5.72
50 MHz	5.23	5.38	5.55	5.63	
100 MHz	5.06	5.44	5.47		
200 MHz	5.58	5.61			

Fig. 13. Junction to Ambient Thermal Resistance (θ_{JA}) computed from empirical data (C/W), T_A = 26 C

(FG680) should have a θ_{JA} of 7.6 C/W for an air flow of 250 LFM (Linear Feet per Minute). The empirical value (5.4±.5 C/W) differs from this by about 27.6%. This difference may occur because the FPX platform sinks more heat than the Xilinx test board. The documentation about the Xilinx test board reports that their FPGA is mounted on a printed circuit board with only one power and one ground plane while the FPX has two ground and two power planes. These additional metal layers allow the FPX circuit board to effectively act as a more effective heatsink than the Xilinx test board.

5.4. Application of Thermal Information

The first order thermal trajectory of the RAD can be estimated from the measured instantaneous power (Power(t=0) ≈ Power(t=∞)). Power(t=0) can be used to estimate T_j(t=∞) via equation 3. T_j(t=∞) can in turn be used to complete equation 1, which gives the first order thermal trajectory. The thermal trajectory allows estimation of how long a circuit can operate in a given mode before reaching a thermal threshold. This information can be used by a dynamic thermal management system to help schedule what jobs to process during a thermal transition.

Scaling the activity rate appropriately could potentially allow the measurements made with these thermal benchmark circuits to provide accurate estimates of thermal behavior of real applications that use similar chip resources. A tool such as Xpower would be a practical method for extracting application specific activity rates. Also even though these benchmark circuits use equal numbers of LUTs and DFFs, they can be easily modified to implement various ratios. Further examination of how accurate an estimation of power and thermal behavior such a method would yield is worth exploring further.

6. CONCLUSION

A thermal management system was motivated and described for the FPX platform. The mechanism used for measuring FPGA temperature was then used to explore in depth the thermal characteristics of the RAD on the FPX platform. The thermal steady steady and transient behavior of the RAD, and thermal relationship to power suggests that the thermal delay between modes of a circuit could be used to implement new dynamic thermal management approaches. It was then proposed that estimation of thermal trajectories could help these management approaches better schedule work during thermal transitions. The benchmark circuits were also shown to have characteristics that may lead to an approach for more accurately predicting the thermal behavior of applications implemented on reconfigurable devices at design time. The thermal management methodology described in this paper can be applied to other platforms to enable modules to operate in short bursts at a higher power level than steady state thermal dissipation would allow.

7. REFERENCES

[1] S. Wenande and R. Chidester, "Xilinx takes power analysis to new levels with xpower," *Xcell Journal Online*, vol. 41, pp. 26–27, 2001.

[2] J. W. Lockwood, J. S. Turner, and D. E. Taylor, "Field programmable port extender (FPX) for distributed routing and queuing," in *ACM International Symposium on Field Programmable Gate Arrays (FPGA'2000)*, Monterey, CA, USA, Feb. 2000, pp. 137–144.

[3] J. W. Lockwood, N. Naufel, J. S. Turner, and D. E. Taylor, "Reprogrammable Network Packet Processing on the Field Programmable Port Extender (FPX)," in *ACM International Symposium on Field Programmable Gate Arrays (FPGA'2001)*, Monterey, CA, USA, Feb. 2001, pp. 87–93.

[4] J. W. Lockwood, C. Neely, C. Zuver, and D. Lim, "Automated tools to implement and test internet systems in reconfigurable hardware," *SIGCOMM Computer Communications Review*, vol. 33, no. 3, pp. 103–110, July 2003.

[5] *MAXIM Remote Temperature Sensor with SMBus Serial Interface*.

[6] I. Corporation, "Addressing power and thermal challenges in the datacenter," 2005.

[7] *Intel 80200 Processor based on Intel XScale Microarchitecture Developer's Manual*, 2003.

[8] E. Wirth, "Thermal management in embedded systems," Master's thesis, University of Virginia, 2004.

[9] Y. C. Hyung Gyu Lee, Kyungsoo Lee and N. Chang, "Cycle-accurate energy measurement and characterization of fpgas," *Analog Integrated Circuits and Signal Processing*, vol. 42, pp. 239–251, 2005.

[10] D. T. S.M. Alam and C. Thompson, "Circuit and system level tools for thermal-aware reliability assessments of ic designs," MIT, Cambridge, MA, Tech. Rep. RLE Progress Report 146, 2004.

[11] E. B. Sergio Lopez-Buedo, Javier Garrido, "Thermal testing on reconfigurable computers," *IEEE Design and Test of Computers*, vol. 17, pp. 84–91, 2000.

[12] S. Lopez-Buedo and E. I. Boemo, "Making visible the thermal behaviour of embedded microprocessors on fpgas: a progress report." in *FPGA*, 2004, pp. 79–86.

[13] *Device Package User Guide*, 2005.