Comparison of Finite-Difference and SPICE Tools for Thermal Modeling of the Effects of Nonuniform Power Generation in High-Power CPUs

Jeffrey L. Deeney

C. Michael Ramsey

This paper describes a thermal study of junction temperature variation across the surface of a large CPU resulting from nonuniform power generation. Results from Flotherm finite-difference thermal analysis software were compared to results from a SPICE simulation. Both simulations provided results close to measured values. Each tool offered strengths and benefits in different areas.

Because of unevenly distributed functionality in large integrated circuits, power dissipation can vary significantly across the surface of a device. In very high-power devices, these power variations can lead to significant variations in junction temperature across the die surface. If these temperature gradients are not fully understood or anticipated, product reliability may be degraded.

The HP PA 8000 microprocessor is a high-performance implementation of HP's PA-RISC computer architecture.[1] It is the central processor of the top-performing models of the HP 9000 C-class and J-class workstations and K-class servers. With a die size of 17.7 mm by 19.7 mm, it is the largest integrated circuit designed by HP to date. At the announced operating

Jeffrey L. Deeney

Jeff Deeney is a lead thermal design engineer for packaging development at HP's High-Performance Systems Division. He joined HP in 1981 after receiving his BSME degree from Washington State University. He is married, has three children, enjoys backpacking and climbing, and is active in public land use issues.

C. Michael Ramsey

Mike Ramsey is a member of the technical staff involved with presilicon verification of advanced processors at HP's VLSI Technology Center. With HP since 1980, he received his BSEE and MSEE degrees from Texas A&M University in 1979 and 1981. He is married, has a son, and is a student of Tae Kwon Do.

Figure 1

Fan/heat-sink forced air cooling solution.

frequencies of 160 MHz, 180 MHz, and 240 MHz, its power dissipation is also quite high. Removing this power presented challenges at both the chip and system levels and was anticipated early in the design process. However, results were still not without surprises. Specifically, the large size of the PA 8000 die and the nonuniform distribution of its heat-generating circuits led to unanticipated temperature differences across the die. The study described in this paper investigated how the thermal characteristics could be better modeled.

The PA 8000 is packaged in a 1089-pin ceramic land grid array (LGA) using solder bump bonding with epoxy underfill onto an alumina substrate. The back of the die is attached to a sintered copper-tungsten lid by means of silver-filled epoxy for enhanced thermal performance. The entire package is socketed onto its host printed circuit module. Different heat sink solutions are used in the various systems. In the C-class workstation, the PA 8000 is cooled by a fan/heat-sink forced air cooling solution (see **Figure 1**). In the J-class workstation and K-class server, heat is conducted through a three-pipe heat pipe assembly cooled by the cabinet's forced air flow (**Figure 2**).

Conventional IC design methodology partitions the functions of the PA 8000 into localized units, as shown in **Figure 3**. The floating-point processing unit is one such unit. The floating-point unit is a specialized coprocessor that delivers improved performance to applications using floating-point arithmetic operations. It is the largest source

Figure 2

Heat pipe mounted to printed circuit board.

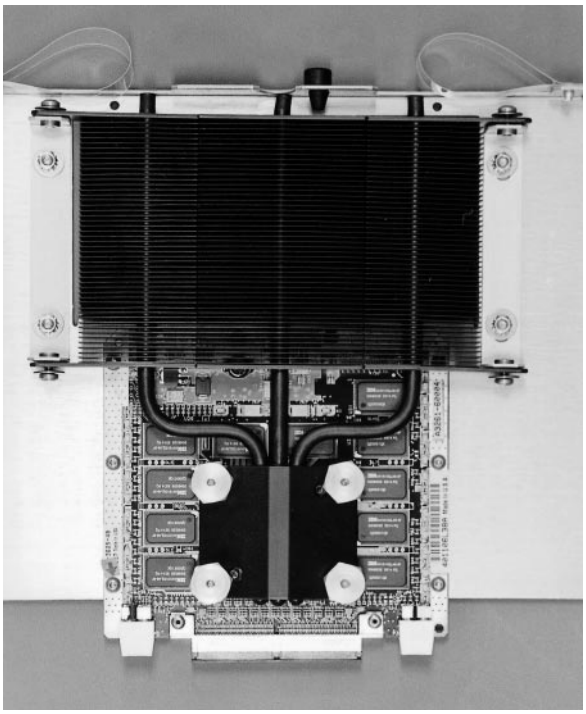
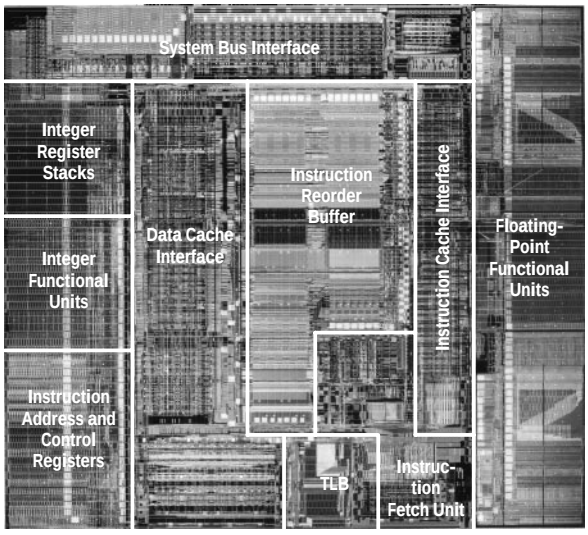


Figure 3

HP PA 8000 processor.



of heat in the PA 8000. The floating-point unit circuits normally dissipate power only when floating-point operations are being executed. Since this results in large power supply current transients, the floating-point unit can be configured to dissipate power even when inactive.

The PA 8000 includes two circuit cells whose purpose is to measure die temperature. Each of these cells contains a serpentine metal trace connected between two dedicated package pins. The resistance of these cells is temperature-dependent. These temperature monitor (TMON) cells must be calibrated for each die and package. The TMON cells were placed on the die in locations dictated by available pins. In retrospect, the placements were less than ideal.

During electrical characterization of the PA 8000, the effect of floating-point unit power-on die temperature was investigated. The TMON cells were instrumented and calibrated. Additional temperature measurements were made at the package/heat-sink interface and of the ambient air.

Electrical power to the die was supplied by a characterization fixture. This fixture is a system of power supplies, clock generators, a thermal chamber, and other instruments controlled by a computer workstation. It allows automated experiments and measurements to be performed on a circuit.[2]

Package-to-ambient-air thermal conductivity was consistent with previous experimental results as well as design objectives. However, TMON temperatures during peak floating-point unit power dissipation did not correlate with the known die input power and die-to-package thermal conductivity. The temperatures measured at the TMONs were lower than expected.

SPICE Simulation Using Electrical Analogs

It was suspected that the TMON temperatures were not representative of temperatures elsewhere on the die. The TMONs are both located 4 mm from one edge of the PA 8000 die. The floating-point unit, whose variable power was the object of the investigation, occupies a 4-mm-wide strip along the opposite side of the die. More than 10 mm of silicon lie between the floating-point unit and the TMONs, the only available means of measuring the floating-point unit's temperature. Thus, direct measurement of floating-point unit temperature was not possible. Another method of determining its temperature was needed.

Simulation of electrical circuits is used extensively in IC development. Drawing an analogy between electrical circuit theory and heat transfer made it possible to apply familiar IC design tools to model heat flow in the PA 8000. Heat flow can be modeled as electrical current flow, thermal resistivity as electrical resistance, thermal mass or capacity as electrical capacitance, and temperature as voltage. Using these analogies, a circuit model was simulated using HP Spice, a version of the public-domain SPICE program. The model was constructed using Piglet, a proprietary schematic capture and artwork design tool with a long history in IC design at HP.

The circuit model divides the PA 8000 die into 360 rectangular elements of 1 mm^2 each (**Figure 4**). Each element is connected to its neighbors through the thermal resistivity of the intervening volume of silicon. Heat flow out of the die is modeled by a resistive path from each element to a common node representing the heat sink.

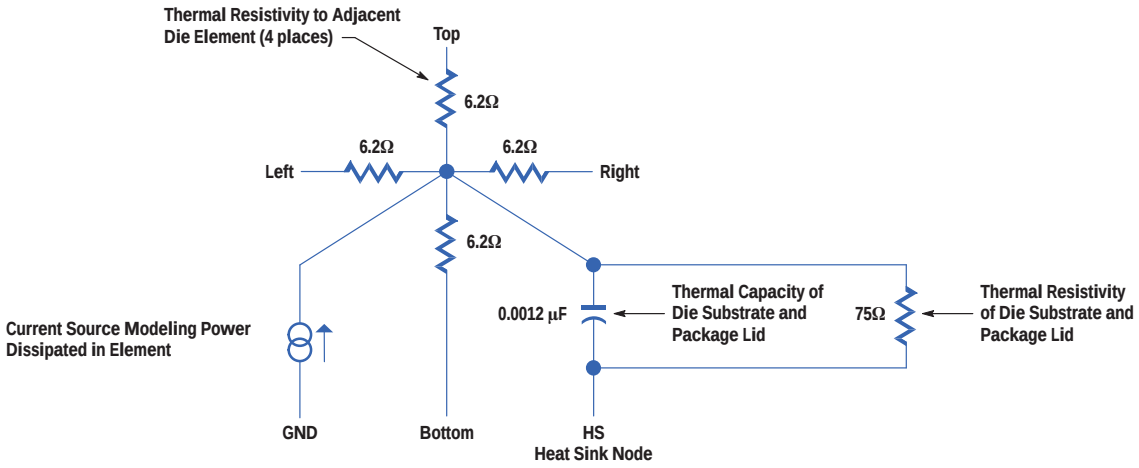
The thermal resistivity and thermal mass of the heat sink are modeled by a parallel resistor and capacitor to a node representing the ambient. A constant-voltage source sets the ambient temperature. A current source into each element of the die model models the power input into the circuits of that element. Power input was based on early simulations using Powermill software from Synopsys, Inc. This simulation summarized the power dissipated in each functional block at the top level of the design hierarchy. Power was assumed to be evenly distributed across the area of the block and a proportional current was sourced into the model elements corresponding to the physical location of the block.

Values for the components in the circuit model were derived by one of two methods. Components located between the package and ambient could be measured experimentally. Because the temperature measurements made using the TMON cells were suspect, components inside the package had to be calculated from the physical properties of silicon. The values used in the model are listed in **Table I**.

Steady-State Analysis

The first SPICE simulation confirmed that temperatures across the die varied greatly. As expected, the highest temperature was found in the floating-point unit. The corners on the side opposite the floating-point unit exhibited

Figure 4
SPICE cell.



the lowest temperatures. There was a surprising temperature difference of 21°C across the die.

Transient Analysis

A strength of SPICE is the analysis of transient phenomena. This is a very common application in IC design. A transient analysis of the PA 8000 thermal environment yielded insight into how the large temperature gradient across the chip was possible.

SPICE was modified to calculate an initial condition with the floating-point unit power turned completely off (see **Figure 5**). A short time into the simulation, the floating-point unit power was turned on to its maximum value.

Over the first 0.5 second after the floating-point unit power is applied, the floating-point unit temperature rises until the shape of the temperature profile across the die is the same as was found in the steady-state analysis. Subsequently, the differences in temperature between any two points remain constant. The absolute temperatures, however, are not at their final values. With a time constant of 70 seconds, the temperature of the entire die continues to rise towards the final temperature. This slow rise corresponds to the heating of the package lid and heat sink.

Table I
Electrical Analogs of Thermal Elements

Component	Method	Value	Electrical Analog
Thermal Resistance of Silicon	Calculated	12.35 °C/W	12.35Ω modeled as one 6.18Ω resistor on each side of the element
Thermal Mass of Silicon	Calculated	0.0012 W/s·°C	0.0012 F
Thermal Resistance of the Die and Package per mm^2	Measured	64 °C/W	64Ω
Thermal Resistance of the Heat Sink	Measured	0.5 °C/W	0.5Ω
Thermal Mass of the Heat Sink	Measured	140 W/s·°C	140 F

Figure 5

Temperature gradient with floating-point unit power off (from SPICE simulation).

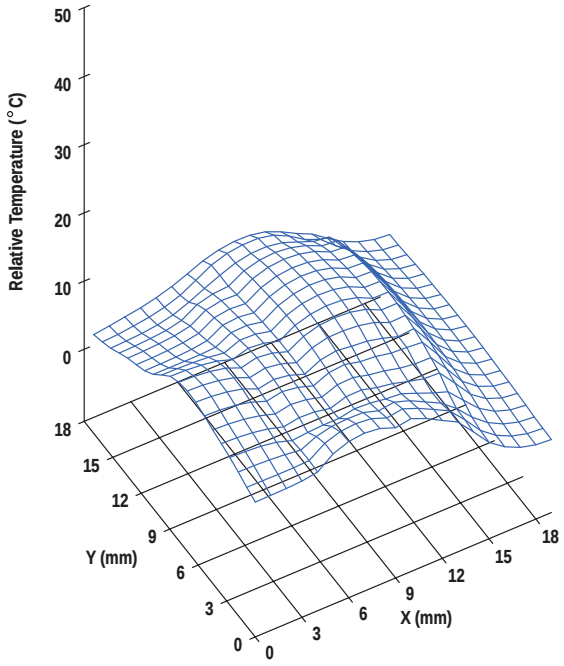
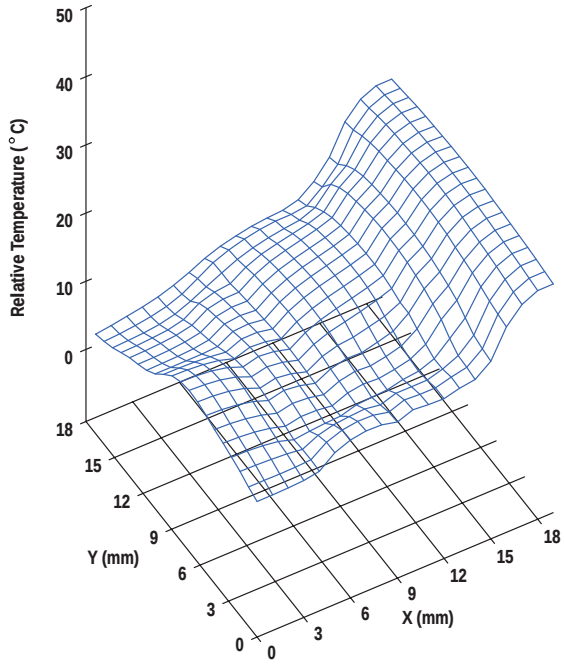


Figure 6

Temperature gradient profile established after floating-point unit power-on (from SPICE simulation).



The shape of the temperature profile as heat spreads out from the floating-point unit explains the magnitude of the gradient. **Figure 5** illustrates the initial state with the floating-point unit power off. Upon application of floating-point unit power, the heat from the floating-point unit can be seen to spread across the die. The advancing wave of heat is attenuated until it stops approximately halfway across the chip. **Figure 6** depicts this condition. The temperature profile of the remainder of the die is not affected directly by the floating-point unit. Instead, it rises only when the added power of the floating-point unit causes the temperature of the package and heat sink to rise. The final temperature gradient can be seen in **Figure 7**. This leads to the conclusion that the thermal resistance across the PA 8000 die is large enough relative to the thermal resistance to the package that locations more than 5 mm apart are effectively insulated from heat transfer.

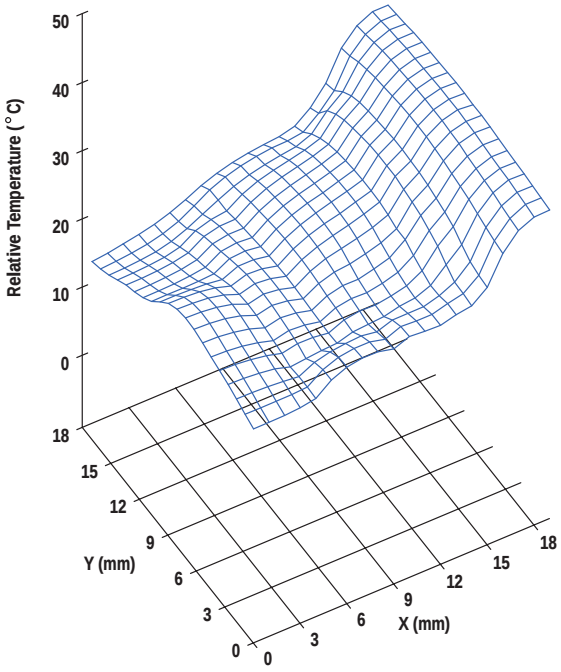
Parameter Sweep

It seemed intuitive that the high temperature gradient across the die was a result of the low thermal resistance to the package. While the low resistance was important in maintaining a low die temperature, it might also be causing a less-than-ideal temperature gradient by reducing heat flow across the die. How is the temperature gradient across the die affected by the package thermal resistance?

A typical application of SPICE is to execute multiple simulations while varying some characteristic of the circuit. This method was applied to the PA 8000 thermal model to determine the effect of changing the thermal resistance of the package. The resistors in the model that represented the package thermal resistance were swept across a range from one tenth of the typical value to 100 times the typical value. The gradient across the die increased with higher package resistance to a peak value. Beyond this peak, heat flow across the die became significant relative to the flow into the package and the gradient decreased with higher package resistance. The value at which the peak temperature difference occurred was 30 to 70 times the typical value, depending on which two points were compared. The die temperature reached unacceptable values long before this. Thus, for large values of package thermal resistance, the temperature gradient decreases with higher

Figure 7

Temperature gradient final state after heat sink heating (from SPICE simulation).



resistance, but at typical values, lower package thermal resistance is consistent with both lower die temperatures and a lower temperature gradient.

Finite-Difference Modeling

Finite-difference techniques involve solving a set of coupled, nonlinear, second-order partial differential equations. For solving thermal problems, the solution space is divided into a series of discrete cells and the differential equations for the conservation of momentum, energy, and mass are solved for each cell. For the purpose of this study, the Flotherm computational fluid dynamics software from Flomerics was used. Although the most popular application of Flotherm software is in the analysis of system-level air flow and convection cooling, it also contains features that allow accurate modeling of package-level details and conduction mechanisms. For the purpose of this study, a finite-element package such as Mechanica by Ransa, Inc. or ANSYS by ANSYS, Inc. would also have worked well.

Model Construction

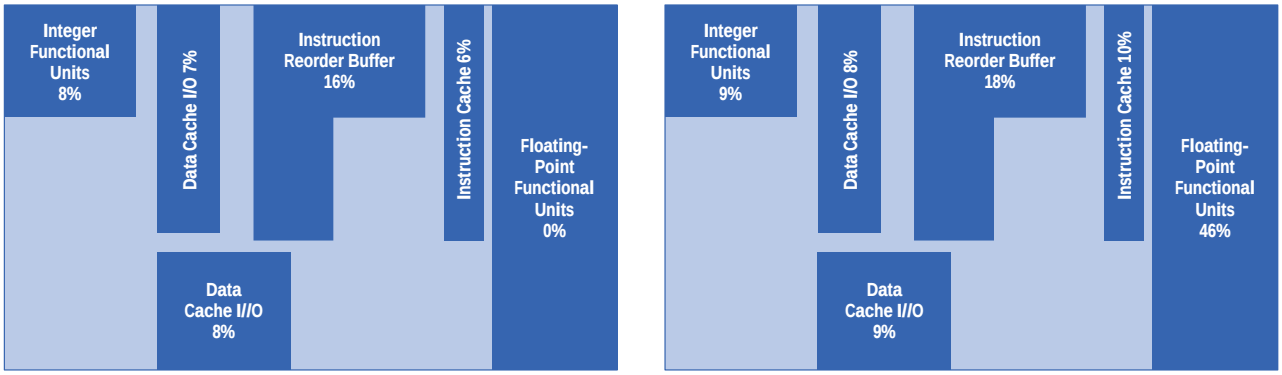
Power generation on the PA 8000 device is roughly symmetrical about a centerline. This allowed a reduction in complexity by creating a model representing only half of the device geometries. This is accomplished by placing the centerline of the half device against a surface that is considered to be a perfect insulator. The original PA 8000 thermal model used a single planar power source evenly distributed over the surface of the die. For our study, this power source was replaced with several distributed power sources. The top six highest-power-generating segments of the PA 8000 chip were isolated and mapped onto the half surface. The remaining elements contributed less than 20% of the power and were fairly evenly distributed. It was felt that their omission would not significantly alter the temperature gradient across the die. Simplified power distribution across the half surface of the chip in the high-power and low-power states can be seen in **Figure 8**. For the purpose of this study, the floating-point unit power in the low-power state was set to the minimum of zero. In the actual application the floating-point unit power is maintained in a 75% standby state to minimize sharp transient loads on the power supply. Maintaining a 75% power level also reduces microcycling stresses on the level 1 and level 2 interconnects.

All elements in Flotherm software must be created from cuboid blocks and plates. The major elements in the model consisted of the silicon die, ceramic substrate, printed circuit board, copper-tungsten package lid, and heat sink. All interface materials, solder bump junctions, and solder columns were modeled using internal plates. The reason for this change was twofold. First, Flotherm interprets internal plates as two-dimensional elements of fixed resistance. This simplifies the solution grid and eliminates many of the high-aspect-ratio cells. Secondly, internal plates allow quick changes to the model for sensitivity analysis. For example, rather than completely regridding the solution for a thicker epoxy, a simple resistivity change can be input and the new solution time shortened by using the previous results. While Flotherm internal plates only allow heat flow in a direction perpendicular to the plate, this was adequate for the elements that were chosen.

To decouple the effects of packaging and the final thermal path (i.e., heat sink or heat pipe), the heat sink was modeled by a large block of aluminum attached to the top of

Figure 8

Power distribution in the (a) low-power and (b) high-power states.



the package. The top surface of this block was defined as an isothermal (constant-temperature) surface. This is not unlike a heat pipe, in which the boiling point of the fluid creates an isothermal region. For the purposes of this study, the isothermal surface was set to 60°C. If a different ambient temperature is expected, the results from this study can be scaled up or down by an amount equal to the difference.

The ceramic package in this model is attached to a printed circuit board. For simplicity, the printed circuit board is modeled with a single lumped power plane to enhance heat spreading. A low air flow rate of 0.29 m/s is present, primarily to minimize the natural convection plumes. With the solution as modeled, more than 90% of the power flows to the isothermal surface. Convective heat transfer from the package and printed circuit board are almost negligible. Details of the model construction can be seen in **Figure 9**.

The Flotherm model was first run using the highest expected power densities, as outlined above. Modeling results indicate a 20°C temperature difference between the hottest and coolest portion of the chip. Two-dimensional thermal contour lines on the surface of the half die can be seen in **Figure 10**.

The Flotherm model was next run using the lowest expected power densities, as outlined above. Modeling results indicate a 10°C temperature difference between the hottest and coolest portions of the chip (see **Figure 10**). Compared to the full-power situation, the highest temperature on the chip decreased by 16°C, while the lowest temperature on the die decreased by only 5°C.

Sensitivity Analysis

Since the original physical thermal tests had been performed with uniform power distribution, one of the first questions to be asked was how much the maximum junction temperature in the nonuniform CPU differed from that measured in the uniform test die. To quantify the disparity between the test die and the PA 8000, a PA 8000 model with a uniform power source was created. The total power was equal to the power of the gradient model in the full-on condition. Under these conditions, the maximum junction temperature is 86°C (compared to a range

Figure 9

Model construction for finite-difference simulation.

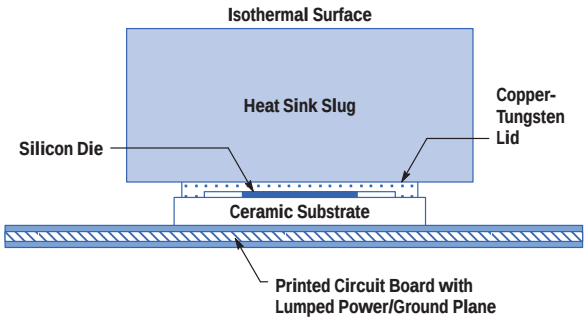
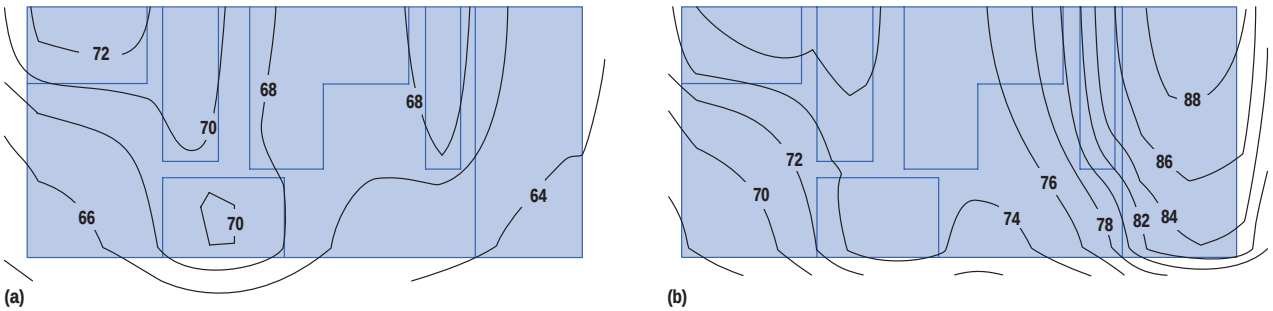


Figure 10

Temperature (°C) gradients during (a) floating-point unit low-power mode and (b) full-power operation (from finite-difference simulation).



of 90 to 68°C for the full-power condition with nonuniform distribution). This simulation showed that thermal testing with a uniform power source may have underpredicted the maximum junction temperature by 4°C. Depending on the amount of margin in the total thermal solution, this may or may not be an issue.

Once it was ascertained that a significant temperature gradient existed across the die surface, various sensitivity studies were undertaken to determine how the package construction might be used to minimize the temperature gradient. It was originally thought that the ceramic substrate might contribute enough heat spreading to equalize the temperatures across the die surface. A study was performed to determine the effect of either increasing or decreasing the thermal conductivity of the ceramic substrate. The nominal conductivity of alumina is 18 W/m·°C. The model was run with thermal conductivity an order of magnitude lower (1.8) and an order of magnitude higher (180). The lower thermal conductivity number is representative of organic laminate substrates. The upper thermal conductivity number is representative of an aluminum nitride substrate.

With reduced substrate thermal conductivity, the temperature gradient was increased by two degrees. With increased thermal conductivity, the thermal gradient was reduced by six degrees. The conclusion is that varying substrate thermal conductivity has a small to moderate effect on the temperature gradient. This is partially because the primary heat path for the PA 8000 package is through the back of the die.

As mentioned in the SPICE analysis discussion above, the excellent thermal path out of the package is a major contributing factor to the high temperature gradients. A thicker copper-tungsten lid on the package would encourage better heat spreading, but would reduce junction-to-ambient thermal performance. Similarly, increasing the thermal resistance of the interface to the heat sink would encourage more uniform temperatures, but would increase the overall junction temperature.

Conclusion

The accuracy of both models was verified by performing various measurements of the package case temperature and using the temperature monitor features built into the PA 8000 chip. Measurements agreed well with the temperature predictions.

From the results of both the SPICE and the Flotherm simulations, it is clear that there is a significant temperature variation across the surface of the PA 8000 die. This variation is caused by the disproportionately high power density of the floating-point unit. This temperature variation may result in underestimation of the actual junction temperature.

Table II compares the SPICE and finite-difference predictions. Both models were surprisingly similar in predicting the temperature gradient across the surface of the die. The predicted areas of maximum and minimum die temperatures were approximately the same with both techniques. Because the SPICE model used higher ambient

Table II
Comparison of SPICE and Finite-Difference Predictions

Simulation Tool	Gradient across Die with Floating-Point Unit On	Gradient across Die with Floating-Point Unit Off	Difference in Maximum Chip Temperature between On and Off States
Finite-Difference	20°C	10°C	16°C
SPICE	21°C	9°C	18°C

temperatures and assumed a higher case-to-ambient resistance, the predicted temperatures were higher overall. Adjustment of the model would bring the nominal temperatures closer to the finite-difference predictions.

Both models predicted similar transient responses as the die was switched to the high-power state. The initial rise time in the SPICE model was somewhat shorter because the model did not include capacitive elements to represent the thermal mass of the ceramic substrate and printed circuit board.

Each model was found to have application advantages. Because of the two-dimensional nature of the SPICE model, the solution time was on the order of seconds, rather than hours, as required for the finite-difference model. At the same time, the three-dimensional nature of the finite-difference model offered a better visual exploration into the effects of various elements in the package construction. For an experienced user, the amount of time to construct each model was about the same. The SPICE model allowed the electrical designer to work with a familiar set of tools to perform thermal analysis, but required a solid understanding of the electrical analogs of thermal elements. Finite-difference tools required a good understanding of how to apply the elements available in the software package to the physical problem. The tools used to preprocess and postprocess the SPICE data allowed automated input of data directly from the chip power models. The Flotherm user interface required manual input of all geometry and power data.

The results of these simulations were discussed with the PA 8000 packaging vendor. They identified localized mechanical strain on the solder bump joints as a result of thermal cycling as a potential reliability issue. Because of concerns over step loads on the system power supplies, the minimum floating-point unit power in existing products is limited to 75% of the full power. Using updated nonuniform power distribution data, the packaging vendor performed finite-element stress simulations and determined that the mechanical integrity of the columns was not compromised at the current power cycling levels.

The knowledge gained from these simulations is also being applied to the development of future Hewlett-Packard microprocessors.

Acknowledgments

Many thanks to Rich Blanco for sharing his PA 8000 Flotherm model, which was modified for this study. Thanks also to the folks at Flomerics for their excellent customer support.

References

1. A.P. Scott, et al, "Four-Way Superscalar PA-RISC Processors," *Hewlett-Packard Journal*, Vol. 48, no. 4, August 1997, pp. 8-15.

2. J.W. Bockhaus, et al, "Electrical Verification of the HP PA 8000 Processor," *ibid*, pp. 32-39.