Computer Modeling to Predict the Impact of Cold Air Stream on Multi-Chip Memory

Amal El Berry
Information Systems Department,
1.Faculty of Engineering &Computer Science ,
Salman Bin Abdulaziz University, Al-Kharj, KSA
2 .Mechanical Engineering Department, National Research Center, Cairo, Egypt

Kusum Yadav
Information Systems Department,
Faculty of Engineering &Computer Science ,
Salman Bin Abdulaziz University, Al-Kharj, KSA

Imtithal A. Saeed
Information Systems Department,
Faculty of Engineering &Computer Science ,
Salman Bin Abdulaziz University, Al-Kharj, KSA

ABSTRACT

RAM was a specialized electronic chip which multiple integrated circuits (ICs), semiconductor or other discrete components were packaged onto a unifying substrate. Facilitating their use as a single component (as though a larger IC). This paper presented model calculated the temperature distribution on four parallel memory modules and studied the effect of running a stream of cooling air . The density of the air domain depended on temperature, and introduced this influence with the ideal gas law. The inlet velocity of the air was given as a parabolic profile with a maximum velocities of 0.20, 0.21,0.22,0.23, 0.24 and 0.25 m/s, at this flow rates the flow will be laminar. This procedure yields average temperatures of 29.2, 32.4, 34.2 and 34.6 °C. The model defined the heat balance in the memory subdomains plus the air subdomain. The temperature field and flux were continued across these subdomains, which mean that there was no need specify any boundaryconditions between the memory and the air subdomain. The model used material properties of the IC-package material FR4 for the memory modules, and in the surrounding subdomain it used the properties for air.

Keywords: Memory, air velocity, Heat source

Introduction

In the case of memory chips, high temperature is expected to increase leakage current[1] which in turn leads to a higher likelihood of flipped bits in the memory array. R. Sabelka and S. Selberherr[2] were introduced the interconnect analysis program package SAP and demonstrated its usefulness for highly accurate interconnect simulation by a polysilicon resistor also they were found another advantage of their method was that the calculated properties are known at every point of the simulation domain, not only at the contacts, but it requires more memory and CPU time. Xiaowu Zhang [3] were studied a thermo mechanical analysis of an MCM using finite element modeling and they were performed to investigate the effect of package dimensions (the heat spreader thickness, the structural adhesive thickness and the substrate thickness) on the warpage, the maximum residual stress as well as the second level reliability[3]. The effect of temperature on memory induced by incomplete cycling in TiNi and TiNiCu ribbons, TiNiCu thin films, and TiNiCu wire with the two-way memory shape, Z.G. Wang .[4] were studied incomplete cycle, or a sequence of incomplete cycles with different temperatures which their results showed that the effect of temperature on memory was common phenomenon and its

alloys. The temperature memory effect (TME) phenomenon of Ni47Ti44Nb9 wide hysteresis shape memory alloy was studied by X.M. He [5] . The experimental data of average Nusselt number were correlated by using the module to channel height ratio within an error of ±18% as, Nu ¼ 0:84Re0:58 Pr0:25α0.47 $0.2 \leq \alpha \leq 0.5$; $1500 \leq Re \leq 7000$ (6). Sikka,.identified the thermal and mechanical challenges of a multi-chip module (MCM) used in a high-end computer system. The chip and thermal paste carrier for an IBM MCM package [7]. A futuristic microprocessor package uses micro channels and an embedded thermo electric device [8]. Natural convection heat transfer in rectangular enclosures is an active research area, due to its significance for both fundamental interest and engineering applications such as thermal management of electronic components which was studied by Mehmet Ilicak, Ali Ecder and Erhan Turan [9] also a numerical study had been performed to observe the effects of Rayleigh number, Prandtl number and boundary conditions on the solution of natural convection heat transfer from one or multiple heat sources mounted horizontally or vertically on substrates within enclosures. The temperature memory effect was found repeatable if the TiNi alloy was cold deformed larger than 12% and incompletely thermally cycled **[10]**. The feasibility and performance of a compact heat transfer model approach for a hot water-cooled heat sink utilized explore a new strategy to reduce the carbon footprint of data centers were evaluated by A. Kubilay**[11]**.

1. Material and Method

RAM is important because it eliminates the need to "swap" programs in and out. When program is run such as a word processor or an Internet browser, the microprocessor in a computer pulls the executable file. [11] Large programs use large amounts of memory. The microprocessor also pulls in a number of shared dynamic link libraries (DLLs) shared pieces of code used by multiple applications. The DLLs take many more megabytes. Then the microprocessor loads in the data files at which want to look, which might total several megabytes if we are looking at more than one document or browsing a page with a lot of graphics. So a big application can easily take 100 megabytes of RAM or more, which can slow our system down significantly if there isn't enough memory. Random access memory (RAM) was the best known form of computer memory[12]. RAM was considered "random access" because the user can access any memory cell directly if we know the row and column that intersect at that cell. RAM data, on the other hand, can be accessed in any order. Similar to a microprocessor, a memory chip is an **integrated**

circuit(IC) made of millions of transistors and <u>capacitors</u>. In the most common form of computer memory, **dynamic random access memory** (DRAM), a transistor and a capacitor are paired to create a **memory cell**, which represents a single <u>bit</u> of data. The capacitor holds the bit of information a 0 or a 1. The transistor acts as a switch that lets the control circuitry on the memory chip read the capacitor or change its state. A capacitor is like a small bucket that is able to store electrons. To store a 1 in the memory cell, the bucket is filled with electrons. To store a 0, it is emptied. The problem with the capacitor's bucket is that it has a leak. In a matter of a few milliseconds a full bucket becomes empty. Therefore, for dynamic memory to work, either the CPU or the **memory controller** has to come along and recharge all of the capacitors holding a 1 before they discharge. To do this, the memory controller reads the memory and then writes it right back. This refresh operation happens automatically thousands of times per second.The capacitor in a dynamic RAM memory cell is like a leaky bucket. It needs to be refreshed periodically or it will discharge to 0. This **refresh operation** is where dynamic RAM gets its name. Dynamic RAM has to be dynamically refreshed all of the time or it forgets what it is holding. All previous process needs to do energy which converted into heat. The downside of all of this refreshing is that it takes time and slows down the memory. In this article, The model calculated the temperature distribution of four parallel memory modules when a stream of cooling air ran perpendicular to the array. The nature of air flow on extended cooling surfaces is to form a layer of stagnant, insulating air over any surface exposed to the air stream. Cooling area by acting as an insulation layer between the heated surface and the cooling air. The thickness, and its ability to transfer heat, of this insulating layer is dependent on many parameters, some are the velocity ,turbulent intensity , length of the cooling surface in the direction of air travel and roughness of the heat transfer surface. The model worked with Navier-Stokes and heat balance equations in the air domain using the non-isothermal flow application mode. The conditions which taken under consideration were parabolic velocity profile at the inlet, constant pressure at the outlet, no-slip conditions at all other boundaries , room temperature at the inlet, convection dominated transport at the outlet and insulation or symmetry at all other boundaries. The model could solve by using a special iterative segregated approach. A well-known approach to solving the Navier-Stokes equations is the pressure correction method [12], which differs significantly from built-in solvers in that it solves for one variable at a time in sequence. This type of method, called a segregated method, generally requires far less memory but considerably more time to converge to the same accuracy as the built-in solvers, but the difference in memory makes it an interesting alternative [13]. The algorithm described here was designed to solve time-dependent problems, but in this case it finds a stationary solution. The idea is to iterate with appropriate time steps until it reaches a stationary solution, which you evaluate by measuring the change of the variables at each time step. When the change is small enough you assume that the algorithm has reached a stationary solution.

1.1. Model description:

A typical finite element analysis consists of three stages, namely; preprocessing, processing, and post-processing. The preprocessing stage consists of geometry building, incorporating material properties, meshing (i.e., creating elements and nodes then assembling elements to represent the entire object), and finally applying boundary conditions and loads. The processing stage consists of solving a set of equations simultaneously to obtain nodal results, which are mainly the nodal displacements. The post-processing stage incorporates graphical representation of the solution and calculating the heat change in multi-chip module. A complete finite element analysis is the sequence of these three stages.

In the following subsections, we describe the application of each stage to the model:

1.1.1. Governing Equations

The finite element method divides the object to be modeled into shaped region, or what is called elements. Each element consists of a number of nodes, which varies according to the element type. An element can have the behavior of a structural element, fluid, heat, or other physical behavior according to the problem given. An approximate solution for the equation governing the behavior of that element can be developed for each of the elements forming the whole object. The total solution is then generated by linking together or assembling the individual solutions taking care to ensure continuity at the boundaries. The equation governing the behavior of the element must be developed; this equation represents a fit of the function to the solution of the underlying differential equation. Mathematically, the resultingnon-Isothermal Flow case, modeled with the Navier-Stokes equations, describes the fluid velocity, **u**, and the pressure, *p* as:

$$\rho u.\nabla u = \nabla.\left[-p\mathbf{I} + \eta(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathbf{T}}) - \left(\frac{2\eta}{3} - \mathbf{k}\right)(\nabla.\mathbf{u})\mathbf{I}\right] + (\rho - \rho_0)\mathbf{g} \quad (1)$$

$$\nabla.(\rho u) = 0 \quad (2)$$

The General Heat Transfer application mode is based on a general energy balance:

$$\nabla(-k\nabla T) = Q - \rho C_p u.\nabla \mathbf{T} \quad (3)$$

Due to heating of the fluid, deviations occur in the local density, ρ, compared to the inlet density, ρ_0. This results in a local buoyancy force expressed as $(\rho-\rho_0)\mathbf{g}$. The model also treats viscosity, η, as being temperature dependent. k represents thermal conductivity; C_p is the specific heat capacity; and Q is the heating power per unit volume, set to 0.0278 MW/m^3 component. The model treats properties for air as being temperature dependent according to the following equations:

$$\rho = (p_o M_\omega)/(\mathbf{R}T) \quad (4)$$

At the outlet the models use purely convective heat flux. We should also set the lateral boundaries periodic with respect to temperature, making the temperatures equal on both boundaries at every *y* value. Finally, the models apply continuity of temperature and heat flux at all interior boundaries.Then modified the first Non-Isothermal Flow application mode in the equation-system coefficients to follow the algorithm above. Finally, the implementation in Step 3 in the algorithm with the Init edits fields of the flow application modes. To achieve a rapid "convergence" to the stationary solution, you must determine the time step, Δt, in a suitable way. If you set Δt too small, the convergence is slow and takes considerable computational time; if you set Δttoo large, the accuracy or the convergence can become very poor. In this model you estimate the time step from the equation.

$$\Delta t = \min \left[\frac{\int_\Omega h dv}{\int_\Omega u dv}.\Delta t_{max}\right] \quad (5)$$

Where *h* is the local element size and *u* is the velocity norm. In addition to this algorithm for solving the Navier-Stokes equations, you also include a linear solution for the heat balance as a fourth step in the iteration loop because the temperature field affects the flow field through thermal expansion and buoyancy forces.

2.1.2 Geometry and Material Properties:

The used material properties and geometric parameters in finite element model for multi-chip module (MCM) were listed in Table 1. We used the geometric parameters to build up object's geometry in 3D; which is the (MCM). The material properties were incorporated in the object property. The model geometry has 4 memories every memory board which shown in figure1 used a 30-pin connector and was about 3.5 x .75 inches in size (about 9 x 2 cm). In most computers, it had to install SIMMs in pairs of equal capacity and speed. This is because the width of the bus is more than a single SIMM. Module was a separate subdomain, and the domain surrounding all of them represented by air.

Figure 1: Memory Board

2.1.3   Meshing

A finite element analysis of any physical problem requires that a mesh of finite elements be generated. Because the generation of finite element meshes is a fundamental step and can require significant human and computational effort, procedures have been developed and implemented that automatize the mesh generation without human intervention as much as possible. The accuracy of the finite element analysis results, measured on the exact solution of the mathematical model, highly depends on the use of an appropriate mesh, and this holds true in particular when coarse meshes need be used to reduce the computer time employed for complex analyses. Hence, effective mesh generation procedures are most important. Various mesh generation techniques are in use. Generally, these techniques can be classified into mapped meshing procedures, in which the user defines and controls the element spacing to obtain a relatively structured mesh, and free-form meshing procedures, in which the user defines the minimum and maximum sizes of elements in certain regions but mostly has little control as to what mesh is generated, and the user obtains an unstructured mesh. Of course, in each case the user also defines for what elements the mesh is to be generated. Mapped meshing techniques in general can be used only for rather regular structural and fluid domains and require some human effort to prepare the input but usually result in effective meshes, in the sense that the accuracy of solution is high for the number of elements used. The free-form meshing techniques in principle can mesh automatically any 3-D domain provided tetrahedral elements are used; however, a rather unstructured mesh that contains many elements may be reached. The challenge in the development of free-form meshing procedures has been to reach meshes that in general do not contain highly distorted elements (long, thin sliver elements must be avoided unless mesh anisotropy is needed), that do not contain too many elements, and that contain brick elements rather than tetrahedral elements. Two fundamental approaches have been pursued and refined, namely methods based on advancing front methodologies that generate elements from the boundary inwards and methods based on Delaunay triangularizations that directly mesh from coarse to fine over the complete domain. Although a large effort has already been expended on the development of effective mesh generation schemes, improvements still are much desired, for example to reach more general and effective procedures to mesh arbitrary three-dimensional geometries with elements[14].

2.1.4 Boundary Conditions

The next step is to apply the boundary conditions on the meshed object, to make the elements know how to behave and deform upon those conditions. In the Boundary condition list specify Pressure, list specify Inflow/Outflow velocity, list select Convective flux and solve with any of the built-in solvers**.** for the flow inlet as boundary-normal flow with a known velocity field. For the natural-convection models, set the inlet velocity to zero. For the forced-convection cases, set up a parabolic velocity profile (similar to a fully developed laminar profile), u_y, at the inlet as a function of u_{max}, which equals 1 m/s. In terms of an equation it is:

$$u_y = s(1 - s)4u_{max} (6)$$

where *s* represents the normalized inlet width and at the outlet all the models use the normal flow/zero pressure boundary condition. In addition, they apply no-slip conditions at the surfaces of the board and the ICs. At the inlet boundary then fix the temperature to 300 K (room temperature).

2. 2 Steps Involved the Model

First, the modeling process was started by determining the physics of the problem, which was a structural problem, and then the (MCM) geometry and material properties was entered to the model. After that, a suitable element type was selected that was used in the meshing step. The quality of the elements in terms of size and fitness to the (MCM) was tested and checked. Following that, the boundary conditions were applied to the model. Thus, the solution phase was begun in which nodal displacements were calculated. Then, the graphical results were displayed and checked regarding nodal displacement profile to be logically acceptable and contact pair status to be fully closed. To optimize the cooling parameters.

3. Results and Discussion

The results of the 2D model shows that the temperature of the (MCM) increased considerably under a heating load of 1 W/component. Noted that the temperature increase of the sources varies from 30 K for the lowest chip up to almost 90 K at the top chip and result is that the (MCM) contributed a large amount of cooling power on its back side, although the thermal conductivity is quite small. This is apparent in the result plots as a temperature rise in the fluid at the right-hand boundary (that is, the back side of the next board in the stack). The fluid flow in the 3D problem is a bit more complex to solve because of the increased number of mesh nodes necessary to resolve the flow and heat transport fields. Theresult shows that the temperature increase at the hottest spot of each component is approximately three degrees higher for the 3D case than for the 2D case. In addition, the temperature difference among the various ICs is smaller in the 3D model, which predicts a more uniform temperature rise of the ICs. The ICs have an operating temperature between 70 and 80 K above ambient. This result is probably closer to reality compared to the 2D simulation because it also includes the horizontal gaps between the ICs. The difference in temperature rise along the board's height is

explained primarily by the fluid-flow pattern.This plots the fluid velocity for both the 2D and 3D models. The maximum fluid velocity is slightly higher in the 3D case than in the 2D case. More importantly, the flow field behaves differently in the two cases. When comparing (a) and (b), note that the velocity fields are rather similar along the center line of the heat sources. However, there is a channeling effect from thehorizontal gaps.

From the simulation results you can also determine the effective convection heat transfer film coefficient, *h*. Calculate it by integrating the heat flux across the fluid boundary of the source objects. Then divide that value with the temperature difference between that of the fluid at the surface and the inlet temperature. Put in terms of an equation, it is:

$$h_i = \left(\frac{1}{\Omega_1}\int q_i d\Omega_1\right)\left(\frac{1}{\Omega_1}\int T_{s,i}d\Omega_2 - T_{f,0}\right)^{-1} (7)$$

whereΩ_1 is the source surface, Ω_2 is the fluid cross section, q_i is the heat flux, while $T_{f,0}$ and $T_{s,i}$ represent the inlet fluid temperature and the surface temperature of source *i*, respectively. Thus, the value of *h* varies between the rows of the sources due to thermal footprints from upstream heat sources. In the case of forced convection, it is common to use the adiabatic film resistance, h_{ad}. Its definition is similar to *h* except it uses T_{cup} instead of $T_{f,0}$. T_{cup} is the cross-section average of the fluid temperature, T_f, upstream of each source, defined as:

$$T_{cup} = \left(\int((\rho n.u)T_f T\Omega_2)\right)\left(\int(\rho n.u)d\Omega_2\right)^{-1}(8)$$

The deviation from the experimental values for the natural convection might stem from differences in the geometry (which is not fully defined in [15]. In the forced-convection case you can compare the achieved results with experimental results by calculating the Nusselt number, *Nu*. It follows from:

$$Nu_L = h_{ad}\left(\frac{L}{k}\right) \qquad (9)$$

Where L in this case is the length of the heat source (20 mm). The calculated Nusselt numbers for the 2D model decrease from 16 to 11 between rows 1 and 4. These values agree well with the experimentally measured ones for similar geometries, being in the range of 15[15].

4. Conclusions:

Rather than defining an absolute temperature limit as in hotspot-detection policies allows the policy to keep the temperature gradients controlled even at lower temperatures. In this model the Navier-Stokes equations in the air domain using the Non-Isothermal Flow application mode and Heat balance equations in the air domain for each of the modules, each with a heat source. The density of the air domain depended on temperature, and introduced this influence with the ideal gas law. The inlet velocity of the air was used as a parabolic profile with maximum velocities. From the results we can also find out the effective convection heat transfer. We can also calculate it by integrating the heat flux across the fluid boundary of the source objects. Then divide that value with the temperature difference between that of the fluid at the surface and the inlet temperature. A general conclusion we can draw from this example is that modeling can achieve accurate values of convective heat transfer film coefficients, although the values do differ somewhat between the 2D and 3D models. In addition, the good agreement between experimental and simulated values indicates the models' high accuracy.

5. References

[1] T. Hamamoto, S. Sugiura, and S. Sawada.(1998) ("On the retention time distribution of dynamic random access memory (dram). IEEE Transactions on Electron Devices, 45(6):pp.1300–1309.

[2] R. Sabelka*, S. Selberherr (2001) "A finite element simulator for three-dimensional analysis of interconnect structures" Microelectronics Journal 32 pp.163–171.

[3] Xiaowu Zhang, E.H. Wong, Charles Lee, Tai-Chong Chai, Yiyi Ma, Poi-SiongTeo, D. Pinjala and Srinivasamurthy Sampath(2004) "Thermo-mechanical finite element analysis in a multichip build up substrate based package design" Microelectronics Reliability 44 pp 611–619.

[4] Z.G. Wang, X.T. Zua, Y.Q. Fu and L.M. Wang (2005) " Temperature memory effect in TiNi-based shape memory alloys" Thermochimica Acta 428 pp 199–205.

[5] X.M. He, L.J. Rong, D.S. Yan, Y.Y. Li (2005)" Temperature memory effect of Ni47Ti44Nb9 wide hysteresis shape memory alloy" Scripta Materialia 53 pp.1411–1415.

[6] Mousa M. Mohamed (2006) "Air cooling characteristics of a uniform square modules array for electronic device heat sink"Applied Thermal Engineering 26 pp 486–493

[7] MEHMET ILICAK, ALI ECDER* and ERHAN TURAN(2007)"Operator splitting techniques for the numerical analysis of natural convection heat transfer" International Journal of Computer Mathematics Vol. 84, No. 6, June, pp.783–793.

[8] K. Sikka, "Multi-chip package thermal management of IBM z-Server systems",ITHERM, 2007.

[9] R. Yavatkar, M. Tirumala, "Platform wide innovations to overcome thermal challenges", THERMES, 2007.

[10] YanjunZheng, Juntao Li, Lishan Cui (2009)" Repeatable temperature memory effect of TiNi shape memory alloys" Materials Letters 63pp. 949–951.

[11] A. Kubilay, S. Zimmermann, I. Zinovik1, B. Michel, and D. Poulikakos (2011)"Compact Thermal Model for the Transient Temperature Prediction of a Water-Cooled Microchip Module in Low Carbon Emission Computing"Numerical Heat Transfer, Part A, 59: pp.815–835.

[12] A.J.Chorin, 1968, "On convergence of discrete approximations to the Navier-Stokesequations," *Math. Comp.*, vol. 22, pp. 745–762.

[13] Guermond, L. Quartapelle, 1998, "On the approximation of unsteady Navier-Stokes equations by finite element projection method ," *Numer. Math*, p. 217 in: J.-L. vol. 80, pp. 207–238.

[14] Klaus-Jürgen Bathe"Finite Element Method" textbook,13 JUN 2008.

[15] A. Ortega, "Air Cooling of Electronics: A Personal Perspective 1981-2001," presentation material, *IEEE SEMITHERM Symposium*, 2002