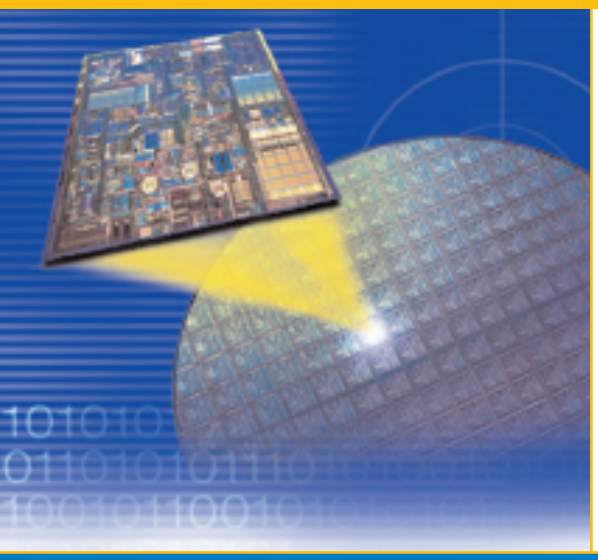
Volume 06 Issue 02 Published, May 16, 2002 ISSN 1535766X



Intel® Technology Journal

Semiconductor Technology and Manufacturing

Emerging Directions for Packaging Technologies

A compiled version of all papers from this issue of the Intel Technology Journal can be found at:
http://developer.intel.com/technology/itj/index.htm

Emerging Directions For Packaging Technologies

Ravi Mahajan, Technology and Manufacturing Group, Intel Corp.
Raj Nair, Technology and Manufacturing Group, Intel Corp.
Vijay Wakharkar, Technology and Manufacturing Group, Intel Corp.
Johanna Swan, Technology and Manufacturing Group, Intel Corp.
John Tang, Technology and Manufacturing Group, Intel Corp.
Gilroy Vandentop, Technology and Manufacturing Group, Intel Corp.

Index words: emerging directions, packaging, power delivery, vertical integration, thermal management, BBUL

ABSTRACT

The continual increasing performance of microelectronics products places a high demand on packaging technologies. Key drivers such as thermal management, power delivery, interconnect density, and integration require novel material development and new package architectures. In this paper, package technology migrations for microprocessors and communication products are described. Material needs for high thermal dissipation, high-speed signaling, and high-density interconnects are discussed.

Microprocessor scaling for increased performance and reduced cost places significant challenges on power delivery and power removal due to reducing dimensions, operating voltages, and increasing power. Meeting these challenges indicates a need for advanced packaging solutions, such as Bumpless Build-Up Layer Technology (BBUL); and power-delivery architectures such as On-Package Integrated Voltage Regulation (OPVR) that enhance the power-delivery capability of the packaging architecture. Similarly, solutions using advanced materials and heat management systems such as heat spreaders and high-capacity heat sinks are needed to facilitate power removal. Microprocessor scaling also requires improvements in package substrates and continues to drive major transitions in substrate materials and features while market constraints continue to exert significant cost pressures.

To support the ever-growing demand of cellular communication products for highly integrated, small form factor devices, new package architectures are described. Key research thrusts for the future are also highlighted.

INTRODUCTION

A review of the evolution of microprocessors in the past two decades and a projection into the near foreseeable future in the current decade shows that microprocessor performance continues to match the almost self-fulfilling prophecy of Moore's law [1]. This increase in performance places significant demands on packaging and assembly for performance and reliability. A paper published in the 3rd quarter of 2000 in the *Intel Technology Journal* [2] showed that in response to demand, microprocessor packaging has evolved from simple mechanical protection to a sophisticated electrical/thermal/mechanical platform that enables microprocessor performance. This paper elaborates further on the themes articulated in the earlier paper and provides additional details of some of the emerging trends in assembly and packaging. The key technical drivers for assembly and packaging in the areas of power delivery, power management, interconnect scaling, and integration are articulated. Future driver trends are discussed in order to explain some of the technical challenges these trends have created. Specific technical challenges in power delivery, thermal management, materials development, high-speed signaling, high-density interconnects, and integration are discussed, and the state of the art is reviewed. Opportunities for further work to continue to expand the cost-performance envelope of assembly and packaging technologies are identified; in particular the Bumpless Build-Up Layer (BBUL) technology is reviewed.

Attention is then shifted to the packaging of components used in portable and cellular devices. These applications demand low-cost, high-performance packaging in

compact form factors. The market segments present unique challenges in terms of cost, performance, and time to market. New package architectures developed to address these challenges are reviewed, and future developmental opportunities are highlighted.

POWER-DELIVERY CONSIDERATIONS

Microprocessor scaling has consistently adhered to Moore's law [1]. Increasing transistor density combined with the performance demanded from next-generation microprocessors result in increased processor power. Scaling also necessitates a reduction in the operating voltages both for reliability of the finer-dimension devices and for containing the power consumed. This reduction in the supply voltage further increases the supply currents drawn by the microprocessors while margins for noise in the power supply shrink in absolute terms.

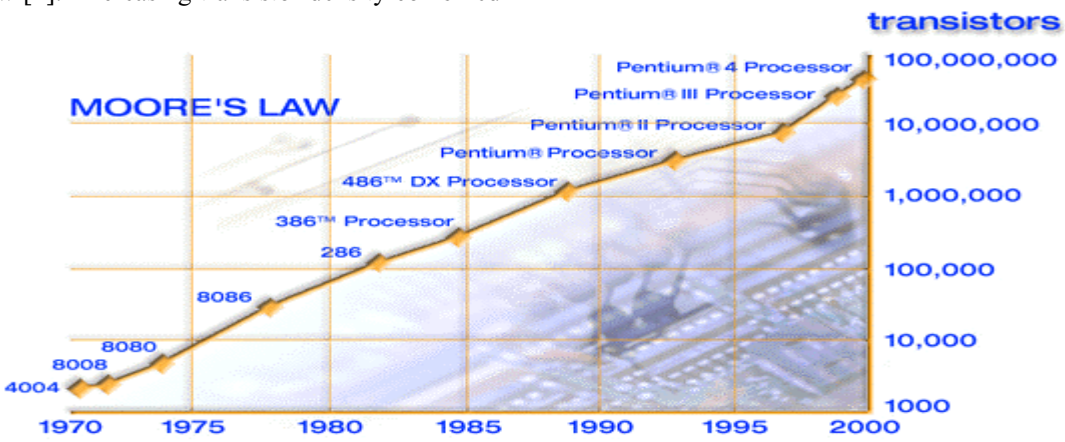


Figure 1: Intel CPU transistors double every ~18 months

The increasing supply currents and shrinking margins in supply noise place an enormous burden on the circuits that provide power to the chip. These circuits are collectively referred to as the power-delivery system for the processor, and they constitute the power-conversion devices or Voltage Regulator Modules (VRMs) that step a high voltage of 12 or 48V down to the processor operating voltage (~1.5V) as well as the hierarchy of capacitances located at the output of the VRM's extending all the way into the microprocessor package [3].

Increasing currents tax the power-delivery system of a CPU or network in two principal ways:

- Power-saving features in the CPU architecture mandate various operating conditions that lower power consumption to a minimum through 'sleep,' 'stand-by,' 'idle,' and 'power-down' states: when the CPU changes state to a fully operational mode, it demands a sudden surge in current in a short duration of time.
- The very large (>100A) currents flowing in the interconnect between the VRM's and the CPU cause power wastage and the associated self-heating in the board, socket, and CPU packages.

Power Path Loop Inductance Scaling and Reduction

The surges of current (often referred to in the literature as DI/DT events) as well as the sudden relaxation in the current demanded by the CPU, combined with the properties of the noise de-coupling capacitance hierarchy, result in a series of supply voltage variations referred to as supply droops and overshoots. The droops in the network are dependent upon the capacitance hierarchy, with the voltage variations being a consequence of damped resonant oscillations of the various de-coupling loops of capacitances and inductances in the network. It can be seen through simple analytical derivation that the magnitude of these droops can be calculated as follows:

$$\Delta V = \Delta I \sqrt{\frac{L_p}{C_d}} \qquad (1)$$

Where Lp and Cd in equation (1) refer to the package loop inductance and the die effective capacitance, respectively. The variation in supply voltage ΔV refers to the droop corresponding to the resonant response of the loop, including the die capacitance and the loop inductance, to the next level of de-coupling capacitance.

A key challenge in minimizing this droop is the scaling of the package loop inductance to meet the demands of CPU scaling. Exploring this further, we get the following:

$$\Delta V_1 = k_v V_1 = \Delta I_1 \sqrt{\frac{L_1}{C_1}} = k_i \frac{C_1 V_1^2 f_1}{V_1} \sqrt{\frac{L_1}{C_1}}$$

And

$$\Delta V_2 = k_v V_2 = \Delta I_2 \sqrt{\frac{L_2}{C_2}} = k_i \frac{C_2 V_2^2 f_2}{V_2} \sqrt{\frac{L_2}{C_2}}$$

Where X_i refers to a parameter X in the two processes. For example, L_1 is the package loop inductance in process-1 and L_2 is the same for the next-generation process-2. Defining S_l as (L_2 / L_1), and dividing the equations above, loop inductance scaling is given as

$$S_l = \frac{1}{S_c S_f^2} \qquad (2)$$

Where S_x refers to the scaling factor for parameter X. Interestingly, under the assumption that the die size and the architecture remain identical, the scaling factor for loop inductance is determined purely by the capacitance scaling factor (die capacitance per unit area) and the scaling factor for the operating frequency.

Power Progression for Intel CPUs

Historical data on the increase in power for Intel microprocessors is included in Figure 2. It is seen that the power doubles approximately every 36 months, which is approximately half the pace of the increase in the number of transistors as forecasted by Moore's Law. This difference could be attributed to greater amounts of memory content in microprocessors as they are scaled, leading to less overall capacitance contributing to power consumption than would be otherwise expected, or smaller die sizes. A capacitance scaling factor S_C = 1 reconciles the historical power trend with simple analytical predictions based on doubling frequency, transistor count, and less aggressive voltage reduction. These are described later.

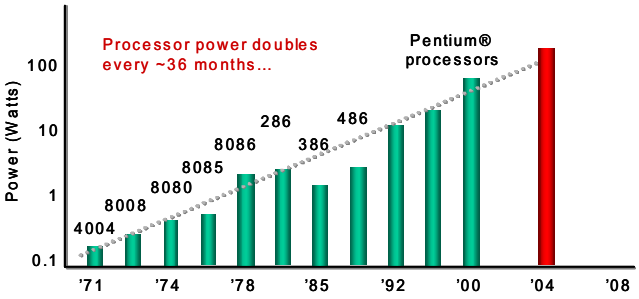


Figure 2: Historical power trend for Intel CPUs

Consider now a process scaling factor of $\frac{1}{\sqrt{2}}$ or 0.7 (as has been typical over a number of generations), and inspect power-related processor scaling scenarios:

- A *Quintuplet*[1] *'Q'* scenario, where C increases by (1/0.7) or S_c, the operating frequency increases by 2X or ~$(1/0.7)^4$, and the voltage reduces only by $\sqrt{0.7}$ or 85%, leading to a 2X increase in power and a $2^{\frac{5}{4}}$ or ~2.4X increase in the dynamic supply current.
- A *Realistic Scaling 'RS'* scenario where power P increases by a factor of $\sqrt{2}$ and the supply voltage V reduces by $\sqrt{\frac{1}{\sqrt{2}}}$ leading to an increase in the dynamic supply current of $2^{\frac{3}{4}}$ or ~1.68X.
- A *Triplet 'T'* scenario where the capacitance per unit area increases as before by S_c, but the frequency increases only by ~(1/0.7) or S_c, and the voltage reduces by a full 0.7 factor; the power in this scenario remains constant while the current increases by $2^{\frac{1}{2}}$ or ~1.43X.

It can be seen, by applying the simple scaling law (2) above that the *'Q'* scenario requires a scaling of the loop inductance by the 5th exponent of the process scaling factor S_c, while the *'T'* scenario also necessitates a scaling of the loop inductance by the 3rd exponent of the process scaling factor. The realistic *'RS'* scenario that reflects historical power scaling would require a loop-inductance scaling of the 4th exponent of process scaling.

[1] The motivation for the scaling scenario name will be evident in the consequence the scaling scenario imposes upon loop inductance scaling.

Figure 3: A Voltage Regulator Module (VRM)

As can be seen from equation (1), the knobs available to control the droops are limited from an assembly technology standpoint for the resonant loop of the highest frequency. Both ΔI and C_d are determined by the CPU architecture, circuit design, and layout as well as by process scaling. Components of assembly technology, particularly the package caps and the substrate, contribute to the effective loop inductance, L_p, that determines the droop in this damped resonant circuit. As supply currents increase and explicitly added die capacitances are removed (for die cost, leakage, and other die and circuit design reasons), reducing the power loop inductance to control droop ratios is a key driver that leads to the consideration of advanced packaging techniques such as the Bumpless Build-Up Layer (BBUL) technique, described in a later section.

Figure 4: Power pod power-delivery system

Power Path Series Resistance Scaling and Reduction

Power path resistance losses contribute significantly to system inefficiency as well as heat generation within the boards, sockets, and packages that support the microprocessor. The large current values anticipated in future microprocessors (in excess of 100A) are forcing greater pin counts in sockets, larger copper thicknesses in boards, and are arresting the thickness reduction (for aspect ratio control in fine trace widths) for metal layers within substrates. This is perhaps a more difficult challenge than the loop-L reduction.

Let's assume that the platform design requires that the loss in the power-delivery interconnect remain constant. As currents scale by a factor S_I, the resistance will need to be reduced by the square function of the current scaling factor to satisfy the constant interconnect power requirement, or

$$S_{ppr} = \frac{1}{S_I^2} \qquad (3)$$

where S_{ppr} is the power interconnect resistance scaling factor and S_I is the current scaling factor. It can be seen that the Q scenario for current scaling (where S_I = 2.4) will require a path resistance reduction by a factor of ~0.175 for the next process generation, the RS scenario requires a reduction factor of ~0.35, and the T scenario will need an S_{ppr} of ~0.5.

Figure 5 displays graphically the challenge in scaling assembly technology parameters in accordance with microprocessor scaling. Starting from hypothetical values for the loop-inductance and the series path resistance in the 0.18µm process generation, it is seen that the loop-L value required to meet the performance criteria three generations ahead is ~0.047pH for the Q scenario and 0.4pH for the T scenario. Similarly, the path resistance scaling beginning as before from a hypothetical value of 3.2 mΩ for the 0.18µm generation leads to a value of ~0.017 mΩ for the Q scenario and ~0.38 mΩ for the T scenario. The RS scenario lies somewhere in between these numbers. This illustrates the impracticality of scaling assembly technology parameters, as demanded by these scaling laws, in order that the performance requirements be met. It is becoming increasingly evident that increases in power or supply currents will require new architectures for power delivery to microprocessors in deep sub-micron processes.

While devices such as the VRMs shown in Figure 3 have advanced in their capability to deliver power, it is seen that solutions such as the Power-Pod (Figure 4), adopted for the anticipated high power and supply current numbers for the Intel Itanium™ family of processors will also be insufficient to meet the requirements of the 65nm node.

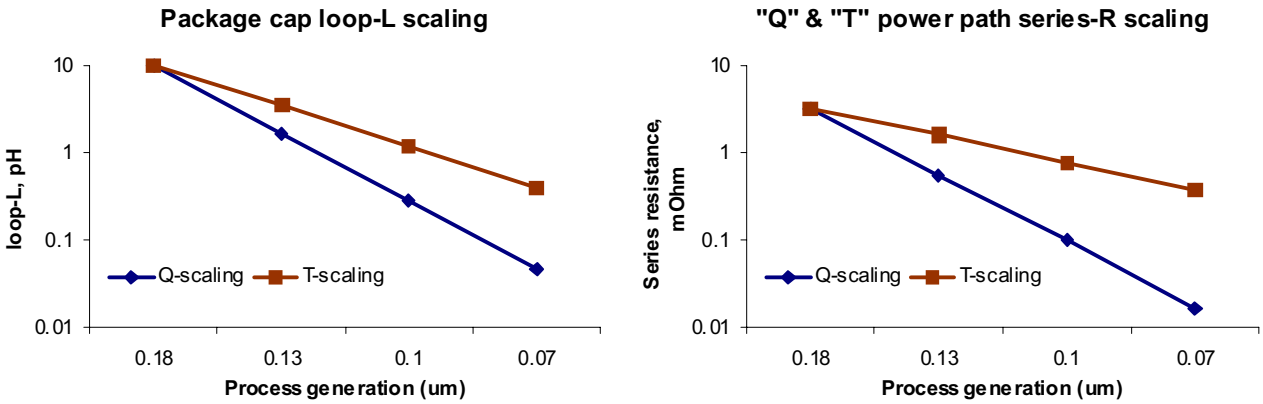


Figure 5: Scaling trends for package loop inductance and series path resistance

THERMAL MANAGEMENT CONSIDERATIONS

Technical challenges in the thermal management of microprocessors arise from two causes: (a) increasing power dissipation, which is concomitant with increasing performance; and (b) the need to cool regions of local power concentrations, often referred to as "hot spots." Typically, thermal management features are integrated in packages to spread heat while transporting heat from the die to the heat sink. The heat sink in turn dissipates heat to the local environment (see Figure 6 for a pictorial representation of this process).

The thermal management problem is one of ducting the Thermal Design Power (TDP) from the die surface at temperature Tj (commonly referred to as junction temperature) to the ambient at temperature Ta. In general terms, the temperature difference (Tj – Ta) is expected to slowly reduce over time since Tj can typically be forced lower by reliability and performance expectations, and Ta can be forced higher due to heating of the inside box air caused by increased integration and shrinking box sizes. Figure 2 shows TDP trends over time, and a simple scaling projection (as shown in the ITRS [4], for instance) indicates that the TDP will increase as a function of time, assuming no major design breakthroughs, which reduce microprocessor power, occur. Thus the thermal challenge arises from the fact that increasing values of TDP have to be ported between a diminishing temperature difference.

This challenge is exacerbated by another very important factor. On-die power distribution is typically not uniform. With increasing performance, the non-uniformity of on-die power distribution increases, and there are regions of the die dissipating high heat flux densities. These regions are commonly referred to as hot spots. Since the temperature of the hot spot can often affect performance and will always govern the overall reliability of the silicon, maintaining the hot spot temperature below certain limits is a requirement in thermal design. This leads to two undesirable consequences: (a) the focus on cooling the hot spot leads to a general over-design in the microprocessor cooling solution; and (b) the non-uniformity in the heat source limits the total amount of heat that can be managed by a thermal solution. The overall problem is graphically illustrated in Figure 7, which shows a plot of the overall cooling capability of different thermal solutions as a function of the Density Factor (DF), a factor defined in [5] as a measure of the impact of power non-uniformity.

Other significant constraints that must always be understood are the cost and integration constraints. While increasing power demands more sophisticated thermal management solutions, they have to stay within cost bounds dictated by the market segments in order to be economically viable. The solutions must also be capable of fitting within the chassis form factors and when assembled with the rest of the components, they should not reduce the reliability of the overall system.

Thermal Solution Strategies

An obvious solution is to mitigate the problem by design, i.e., design processors that are power efficient and have benign power topologies. This can be done by ensuring that thermal considerations are part of the design process rather than an afterthought.

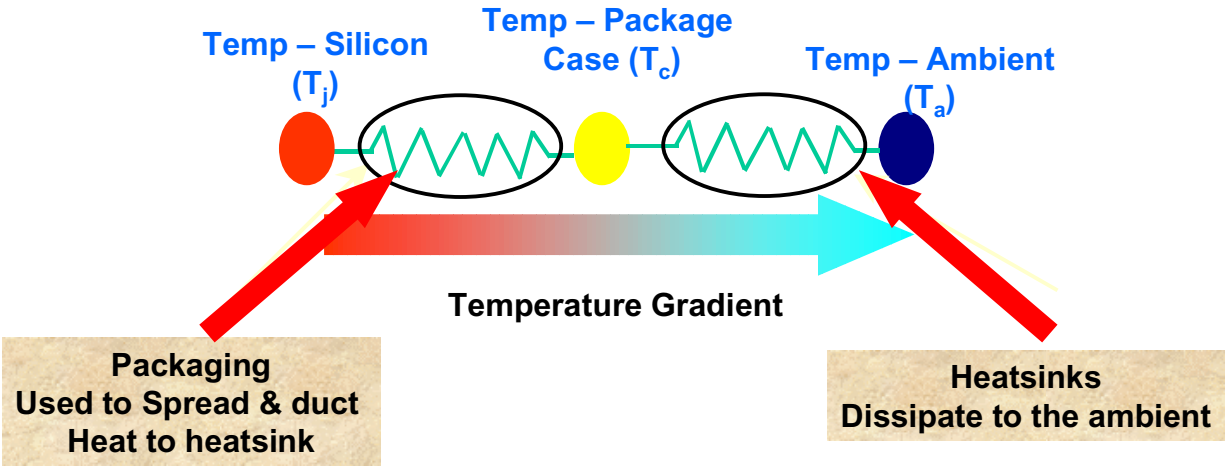


Figure 6: Temperatures and hierarchy in microprocessor cooling

Awareness of thermal issues and the need for thermal co-design has increased over the past few years as thermal management becomes one of the limiters of microprocessor performance. Thermally efficient designs could slow down the unconstrained increases predicted in Figure 2 and can help significantly in the development and deployment of cost-effective thermal solutions. However, while thermally benign designs are being developed, they cannot be depended on as the only strategy. Technology solution strategies need to be developed in parallel assuming that thermal demands will increase over time. These solution strategies can be broadly categorized as follows:

1. *Hot spot mitigation*: The goal of this would be to even out the temperature profiles due to non-uniform power distributions, as close to the source as possible, by spreading out the heat. The use of efficient and cost-effective methods to spread out the heat will increase cooling capabilities by reducing the Density Factor (DF) as seen in Figure 7. Spreading can be accomplished by optimal material and design development. One good example of this is shown in Figure 8. Figure 8 (a) shows an Integrated Heat Spreader (IHS) included in the packaging for Pentium® 4 processors, and Figure 8 (b) shows an integrated heat pipe lid developed for the Itanium™ processor.

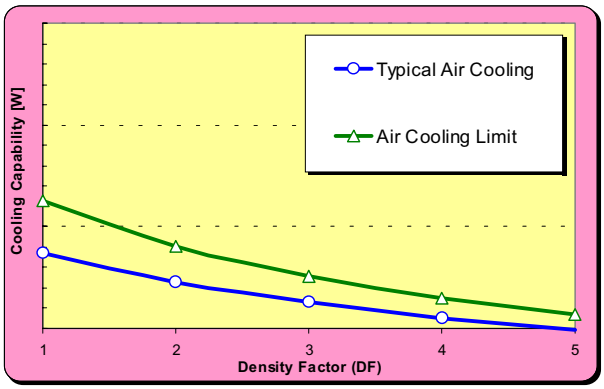


Figure 7: Impact of die power non-uniformity on cooling capability

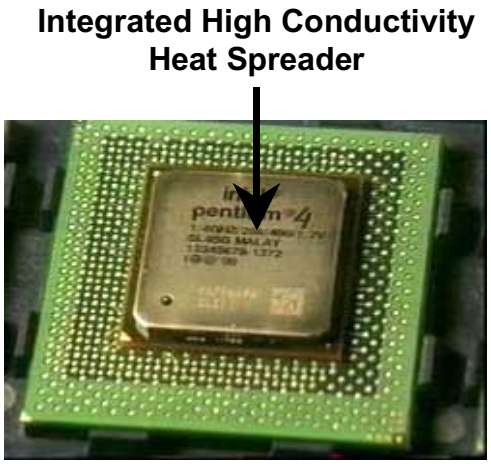


Figure 8(a): Use of IHS for Pentium® 4 processor

Pentium and Itanium are trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

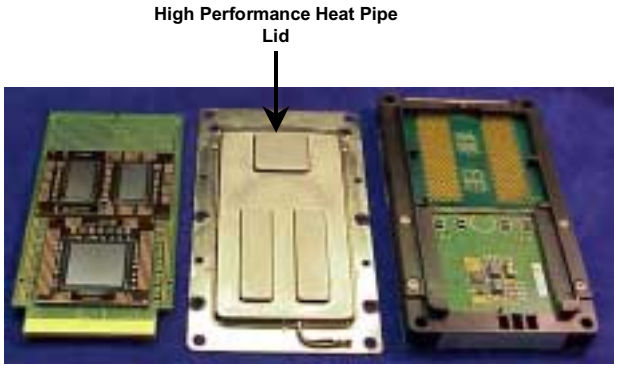


Figure 8(b): Heat pipe lid used for Itanium™ processor

2. *Increasing power-handling capability*: Currently air cooling techniques are used to cool microprocessors in most applications, i.e., a metal heat sink with air blowing over it is the primary cooling solution of choice. A representative analysis of the limits of air-cooling technology indicate that there are still some opportunities in air cooling that may be explored to increase cooling capability.

THERMAL MATERIALS TECHNOLOGIES

We can identify two basic architectures when describing heat removal for microprocessor packaging: (a) Architecture I (FC-XGA1), typically dealing with low-power (<30W) microprocessors or microprocessors in height-constrained applications, where the die is directly attached to the heat sink; and (b) Architecture II (FC-XGA2), typically dealing with medium- to high-power processors (>30W) where an Integrated Heat Spreader (HIS) is used to spread the heat. The term xGA applies to either PGA or BGA, and refers to the type of interconnect between the package and the motherboard. Figure 9 shows the basic implementation of these two architectures.

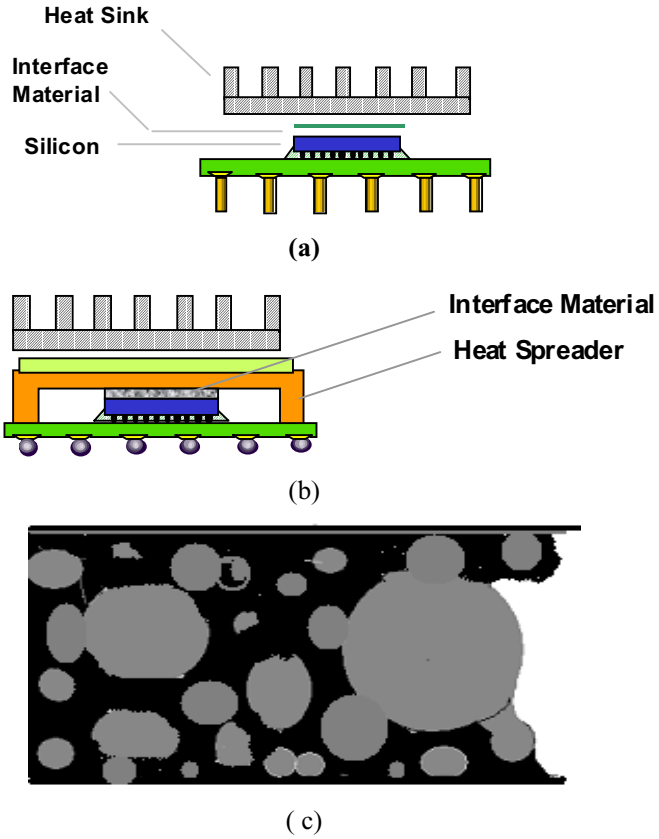


Figure 9: (a) Architecture I; (b) Architecture II; (c) schematic showing percolation in polymer TIM

In either case, successful thermal management requires the development of a Thermal Interface Material (TIM) that comes in contact with the die and heat sink (Figure 9 (a)) or the die and the heat spreader (Figure 9 (b)). Typically the TIMs are made up of a polymer matrix in combination with highly thermally conductive fillers (metal or ceramic) and can be classified as Phase Change Materials (PCM) and Thermal Greases and Gels [6]. Heat dissipation through these materials occurs through the phenomenon of percolation, schematically illustrated in Figure 9 (c). One can also consider providing thermal solutions for heat dissipation entirely through conduction, by utilizing TIMs made up of metals such as lead, tin, or bismuth. However, to integrate these metallic TIM material technologies with the silicon and packaging technology of choice poses an entirely new set of challenges from a stress, cost, and infrastructure perspective. An ideal case would be to develop a composite TIM material that provides enhanced heat dissipation while balancing the mechanical properties to minimize package stress. Developing these composite TIMs can possibly be achieved in a variety of ways including utilizing recent developments in nano-material technologies. It is

expected that one can develop materials with bulk thermal conductivities 5-10 times of those obtained using conventional approaches.

PACKAGE SUBSTRATE TECHNOLOGY

Package substrate technology has undergone significant changes in the past two decades. The early X86 processors were packaged in ceramic substrate with tungsten (W) or molybdenum (Mo) interconnects. Ceramic substrates continued to be the substrates of choice up to the Pentium® Pro microprocessor generation. However, ceramic substrates suffered from the disadvantage of high dielectric constants, thick dielectric layers (leading to thick packages), poor conductor materials relative to Cu (W or Mo), and limitations on feature size just to name a few. In the mid 1990s, Intel pioneered the transition from ceramic to organic substrates. Organic substrates provide better performance at a lower cost, and have evolved to be the substrates of choice for microprocessor packaging. Figure 10 illustrates the substrate evolution for Intel microprocessor generations.

As future microprocessor and network processors run at increased clock speeds, significant challenges are imposed on the performance of the substrate technology. The key technology drivers are as follows:

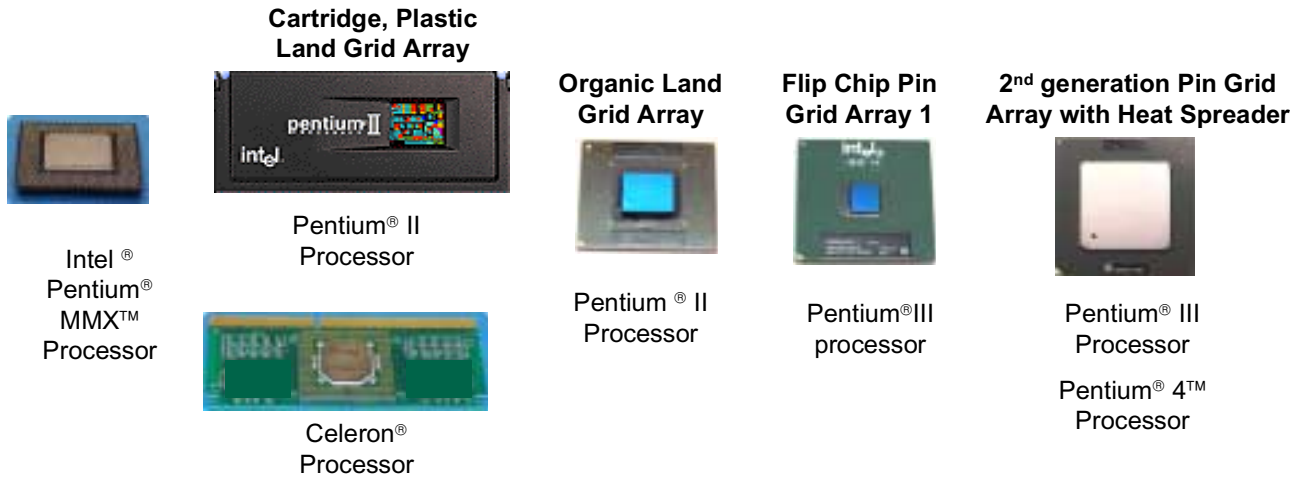


Figure 10: Substrate evolution

1) *Feature size reduction to increase routing density*. A reduction in feature size of the substrate includes a reduction in interconnect line width, line spacing, micro via diameter, capture pad diameter, plated through-hole diameter, flip-chip bump pad pitch, and flip-chip pad solder mask opening diameter. New processes and process materials are required. These include better photo resist and solder mask materials for higher resolution, better layer-to-layer alignment litho tooling and processes to reduce capture pad size, better mechanical drill bits to drill smaller plated through holes, better and higher through-put laser drill equipment for smaller micro vias, and better dielectric materials for improved metal adhesion. For today's state-of-the-art build-up processes, the line width and spacing is at about 25 microns. In about five years, the requirement will be substantially smaller.

2) *Increased performance*. In order to meet the requirements of future high-speed microprocessors, new dielectric materials and new process controls are needed. The new materials are expected to have low dielectric constants and low loss tangents. In addition, impedance matching and impedance control are critical for high-speed signals. Impedance control is a direct effect of dielectric thickness and line width tolerance; therefore, the control of line width and the tightening of tolerance are important parameters for future manufacturing processes. Improved processes and materials are also needed to reduce dielectric roughness while maintaining adhesion between the dielectric and copper lines in the substrate.

3) *Increase in mechanical loading*. Due to the increase in performance requirements, the power dissipation of future-generation microprocessors is rapidly increasing. As a result, larger and heavier heat sinks are employed to cool the microprocessor. In order to maintain intimate contact between the heat sink and package, a Thermal Interface Material

(TIM) is introduced between them. The two are then clamped so that the TIM is as thin as possible. This leads to fairly large static and dynamic loads on the package. In addition, the predicted eventual migration to Land Grid Array (LGA) sockets implies potentially higher static loading will be applied to the substrate. The substrate has to withstand all these mechanical loads through the life of the device. Another factor to consider is the implementation of Inner Layer Dielectric materials on the silicon with low dielectric constants (commonly referred to as low k ILD materials). These materials tend to become increasingly fragile with reduced dielectric constants. It is critical therefore to have a substrate material where the stress impact on silicon is mitigated while the thermo-mechanical reliability of the total system is maintained. This is another important requirement for the development of a new class of next-generation substrate materials.

4) *Thinner substrate*. The increase in demand for thin, portable, laptop computers drives the requirement for thinner substrates. This challenge can be addressed on two fronts. Firstly, optimize design with reduced feature sizes to reduce layer count. Secondly, develop thinner materials for the substrate core. Flex substrates using pliant polyimide materials are also being used to reduce thickness; however, cost is a key issue. Lower-cost flex materials are needed. Meeting the thin substrate requirement will require the industry to invest in new process equipment, handling equipment, and carriers. This will negatively impact the cost of the substrate.

5) *Lower Cost*. If price were no object, then it would not be as difficult to custom tailor a substrate technology to meet all the needs stated above. However, in reality, market pressures require that the future substrate costs must be reduced to ensure competitiveness in the marketplace.

In summary, the requirements of continued performance improvement, higher reliability, smaller features, and lower cost are driving the development of break-through technologies.

CHIP-TO-PACKAGE-LEVEL INTERCONNECT MATERIALS

Organic flip-chip technology is today's cost effective technology of choice for meeting high pin count and high-performance requirements. The ITRS roadmap [4] predicts the I/O pitch for the die-to-package interconnect to approach 120 microns in the next five years and 80 microns in the next ten years. A combination of decreasing pitch; environmental concerns (Pb/Halogen free); mechanical stress concerns (e.g., for low k dielectric integrity); electrical requirements (e.g., current density); and cost constraints are driving the development of bump and underfill materials technology in an entirely new direction. Environmental concerns will eventually lead to Pb-free flip-chip technology through the development of new bumping materials. A variety of material choices can be pursued for Pb-free bumping depending on the process of choice, e.g., plating, printing, stud-bumping etc. However, as highlighted in the ITRS, the bumping and underfill technology of choice will have to ensure low k dielectric integrity. A Pb-free bumping material with low yield stress and low creep resistance will be a critical enabler in the future.

The choice of the bumping material guides one towards the mechanical properties required in the underfill material (modulus, Coefficient of Thermal Expansion (CTE), etc.) to ensure adequate fatigue life for the bumps. Decreasing bump pitch and chip height and increasing bump density will eventually push the limits of capillary flow underfill materials. In order to achieve a breakthrough, one has to look for new directions in both polymer and filler technologies. Polymer resin technologies that can provide fundamentally low CTE (<35 ppm) and low viscosity, and that can be 'integrated' with the bumping materials of tomorrow will emerge as the resins of choice. Filler technologies that can be combined with this resin of choice to manage the underfill CTE without impacting flow and/or the bump-to-package substrate interconnect would be ideal. Underfill materials using such technologies can provide further opportunities to develop new cost-effective processes.

The ideal underfill material/process technology would have to be cost effective, and it would have to be capable of being scaled in such a way as to be independent of bump pitch and die size. This would point towards wafer-level underfills, the best case being where the underfill material/process can be integrated with the back-end Fab process technology. In summary, the flip-chip technology of tomorrow will drive a tighter coupling of silicon processing technology and assembly packaging technology. An alternate approach to meeting the needs of reducing pitch in the die package interconnect scaling is discussed in the next section.

BUMPLESS BUILD-UP LAYER PACKAGING

Bumpless Build-Up Layer (BBUL) packaging is a novel technology developed to meet future packaging technology requirements. It is constructed by fabricating the package layers on top of the chip as opposed to attaching a chip and package. The BBUL package provides the advantages of small electrical loop inductance and reduced thermo-mechanical strain imposed on interconnects with low k ILD materials. Furthermore, it allows for high lead count, ready integration of multiple electronic and optical components (such as logic, memory, radio frequency, microelectromechanical systems (MEMS), among others), and inherent scalability. A cross section of a 3-layer BBUL package is shown in Figure 11. Intel initiated the BBUL project with the goal of addressing scalability, power delivery, and low-k ILD compatibility issues. Detailed analysis on the advantages of the BBUL technology may be found in the published literature [7-9]. Here, we present a summary of the technical advantages and manufacturing challenges.

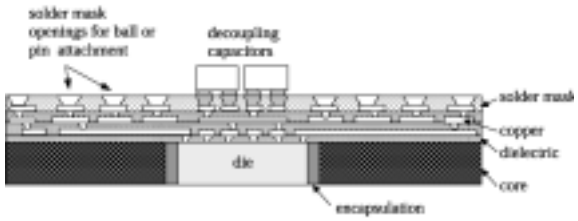


Figure 11: Schematic cross-section of a 3-layer BBUL package

The BBUL package shows a reduction in first droop, largely due to the decreased thickness of the package. BBUL package loop inductance is dominated by the inductance of the discrete decoupling capacitors; the package itself is a minor contributor to the loop inductance penalty. Another key advantage of BBUL packaging is in the area of mechanical reliability. The use of low-k ILD materials on the die is increasing the susceptibility of the die to mechanical failures caused by stresses imposed by the package. Figure 12 shows the relative out-of-plane stress for BBUL versus a standard package, as predicted using mechanical modeling [8] with the commercial finite element code ABAQUS (red is high stress, and blue is low stress on the rainbow scale) [10]. The figures show significantly reduced stress for the BBUL architecture. Equivalent comparisons for first principle stress and Von Mises stress are presented elsewhere [8]. We expect that the BBUL architecture will place lower stresses on the die, thus providing a mechanical advantage over a standard package.

BBUL also offers routability advantages over the standard package. Unlike many versions of flip-chip assembly (such as capillary underfill), die-package interconnections can be arbitrarily placed, as there are no restrictions imposed by an underfill process. This provides a significant advantage in the number of signals that can be routed out from the die on a single layer.

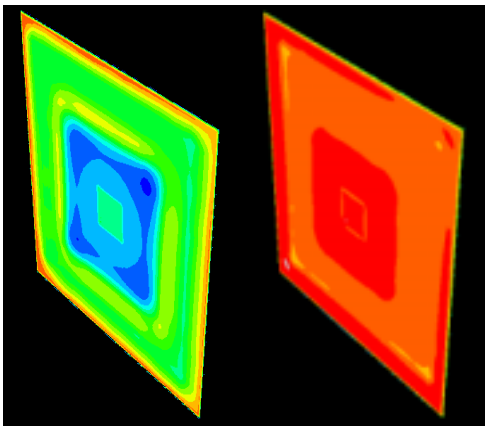


Figure 12: Out-of-plane die M6 to M7 stress for BBUL (left) and standard package (right)

In order to encapsulate the die inside the package, as is done in BBUL packaging, the process flow must deviate significantly from standard assembly. The process flow is shown in Figure 13. With standard assembly, the die and substrate are fabricated in parallel, tested independently, and finally assembled together to form the final package. With BBUL, substrate processing follows die processing, increasing the total throughput time and introducing known good die loss. These are significant manufacturing disadvantages and factor strongly in the final cost of the packaged die. Currently we are investigating, along with our package suppliers, methods of changing to the BBUL process flow and architecture in the hopes of retaining the performance advantages as shown above, while reducing the manufacturing penalty associated with the sequential process flow.

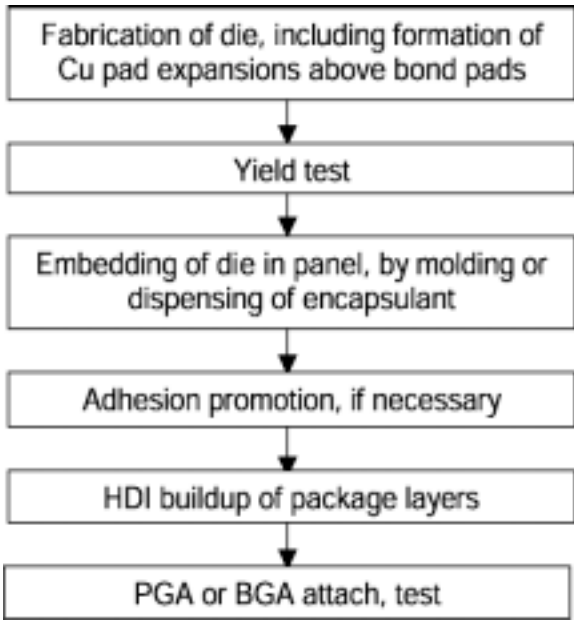


Figure 13: BBUL process flow

CELLULAR COMMUNICATION PRODUCTS

Another important direction for packaging is being driven by cellular communication products. Highly integrated, small form-factor packages are required to meet the demands of the emerging 3G application space. Integration in this environment goes beyond integration of flash with SRAM to integrating new memory (PSRAM, LPSDRAM) with a base band in addition to the traditional memory. In conjunction with the increase in the number of memory types, the available footprint on the board is shrinking, as seen in Figure 14. Vertical integration, i.e., the stacking of multiple die, is a typical approach to meeting the needs of this market segment [11]. While addressing in-plane constraints, vertical integration must also meet stringent height constraints. As a consequence of the increase in the types of die that are being stacked, an important package requirement is the flexibility to mix and change die within a package to meet demand and accommodate late changes. For example, the package architecture must allow for a memory upgrade without causing every layer of the package to be re-routed. These challenges are leading to new solutions beyond the current wire bond stacking found in present-day chip-scale packages.

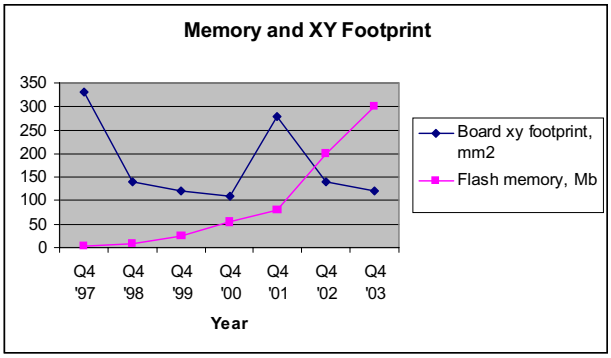


Figure 14: Memory increase and footprint on board

Current stacked-chip, scale-package architecture is shown in Figure 15. Technology efforts are underway to minimize the thickness of each component within the package, and the trend in substrate and die thickness is shown in Figure 16.

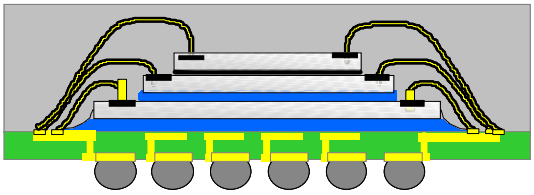


Figure 15: Current stacked CSP (1.4mm)

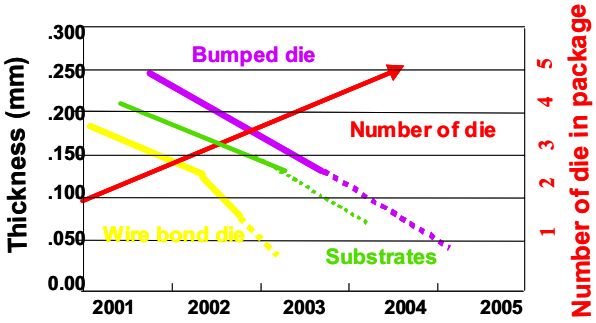


Figure 16: Trends in die/substrate thicknesses

As limits are approached on minimizing the die thickness, forming an interconnect through the silicon is another key direction [12]. The through silicon interconnect allows for a minimum in-plane and vertical footprint of the package. The vias are formed as part of the backend wafer processing, and typical dimensions are on the order of 10-50 microns in diameter with die thicknesses ranging from 25-150 microns.

There are two key issues that drive an alternate package architecture away from die-to-die stacking and towards package-to-package stacking. The first issue is the configurability of the different die within a stack. The

number of possible die combinations is growing exponentially. At the same time the need to respond to changes from the customer do not allow for long re-designs to upgrade memory or stack the different memory with the same base band. A stack-stack architecture can be designed that allows for minimizing the required changes by standardizing on-pin locations in the package-to-package interconnect. The top layer of the package can be quickly reconfigured without impacting the other two packages in the stack, as long as the interconnect terminals supply the same functionality. The second key issue is stacking a mix of Known Good Die (KGD) with non-KGD and the associated yield loss. Sensitivity to yield losses increases with the number of die in a stack. Memory can typically be completely tested by adding a third Sort step, while a base band cannot be as economically tested at the wafer level. Therefore, the ability to package and test the die prior to stacking allows for minimizing die yield losses. Assembly yield also may dictate the need to perform an open/short testing as the packages are stacked. The mix and number of die within a stack will eventually dictate the decision to use die/die stacking versus package/package stacking architectures.

SUMMARY

Packaging is one of the key enablers for microprocessor performance. As performance increases, the technical challenges in the areas of power delivery, interconnect scaling, interconnect performance, power removal, and mechanical reliability increase. This in turn requires the development of new materials and package architectures to enable microprocessor performance. A general overview of the emerging trends has been presented in this paper. An attempt has also been made to provide a context for some of the cost and integration constraints imposed by market conditions on the choice of materials, packaging architectures, and form factors.

Similarly, some of the unique requirements in the cellular market segment due to various form factor and integration requirements have been discussed to provide insight into some of the emergent packaging trends.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the encouragement and guidance in this work received from Bala Natarajan and also thank Mostafa Aghazadeh, Gerald Marcyk, and Nasser Grayeli for reviewing this document. Finally, we would like to thank the technical experts in the ATD Design, ATD Materials and Component Research groups for some of the analyses used to identify the emerging directions.

REFERENCES

[1] Moore, Gordon E., "Cramming More Components into Integrated Circuits," *Electronics*, Vol. 38, No. 8, April 1965.

[2] Ravi Mahajan, Ken Brown, and Vasu Atluri, "The Evolution of Microprocessor Packaging," *Intel Journal of Technology*, 3rd Quarter, 2000. http://developer.intel.com/technology/itj/q32000.htm

[3] Nair, Raj, "Assembly Technology Pathfinding Challenges," *Intel Assembly & Test Technology Journal*, 2001.

[4] ITRS 2001 Roadmap, www.sematech.org

[5] Chia-pin Chiu, Javier Torresola, Greg Chrysler, Dean Grannes, Ravi Mahajan, and Ravi Prasher, "Density Factor Approach to Representing Impact of Die Power Maps on Thermal Management" in preparation.

[6] Ram Viswanath, Vijay Wakharkar, Abhay Watwe, Vassou LeBonheur, "Thermal Performance from Silicon to Systems," *Intel Journal of Technology*, 3rd Quarter, 2000, http://developer.intel.com/technology/itj/q32000.htm

[7] Steve Towle, Henning Braunisch, Chuan Hu, Richard Emery, and Gilroy Vandentop, "Bumpless Build-up Layer Packaging," in *Proceedings ASME Int. Mech. Eng. Congress and Exposition* (IMECE), New York, Nov. 11-16, 2001, EPP-24703.

[8] R. Emery, S. Towle, H. Braunisch, C. Hu, G. Raiser, and G. J. Vandentop, "Novel microelectronic packaging method for reduced thermo mechanical stresses on low dielectric constant materials," in *Proceedings Adv. Metallization Conf.* (AMC), Montreal, Oct. 9-11, 2001, 7 pages, in press.

[9] H. Braunisch, S. Towle, R. Emery, C. Hu, and G. J. Vandentop, "Electrical performance of bumpless build-up layer packaging," in *Proceedings IEEE Electronic Components and Tech. Conference*, (ECTC), San Diego, May 28-31, 2002, in press.

[10] Hibbitt, Karlsson, and Sorensen, *ABAQUS* Version 6.2 User's Manual, Vol. 2, pp. 13.1.2-2, Pawtucket, RI.

[11] Smith, Lee and Zoba, David, "An Extremely-Thin Profile, Ball Grid Array Style Chip Scale Package," in *Proceedings of the SMTA International Conference*, Chicago, Illinois, September 2000.

[12] Yoshihiro Tomita, Tadahiro Morifuji, Tatsuya Ando, Masamoto Tago, Ryoichi Kajiwara, Yoshihiko Nemoto, Tomonori Fujii, Yoshifumi

Kitayama, and Kenji Takahashi, "Advanced Packaging Technologies on 3D Stacked LSI Utilizing the Micro Interconnections and the Layered Microthin Encapsulation," in *Proceedings of ECTC* 2001.

AUTHORS' BIOGRAPHIES

Ravi Mahajan received his B.S. degree from the University of Bombay in 1985, his M.S. degree from the University of Houston in 1987, and his Ph.D. degree from Lehigh University in 1992, all in Mechanical Engineering. He is currently in ATD Pathfinding and is primarily responsible for working with different groups within Intel Corp. to set strategic directions for thermal management of microprocessors. He is also a program manager responsible for establishing assembly and packaging strategies for the next-generation microprocessors. In addition, he is the Intel representative on the Technical Advisory Board for the Packaging and Interconnect thrust within the Semi-conductor Research Corporation. Ravi joined Intel in 1992 as a stress analyst in ATD. He then moved on to manage an analysis group and a TM Lab that developed a series of sophisticated analytical and experimental tools for design development and reliability assessments. He holds nine patents and is the author of several technical papers. His e-mail is ravi.v.mahajan@intel.com.

Raj Nair obtained his B.E. degree from the University of Mysore, India in 1986 and his M.S.E.E. degree from Louisiana State University in 1994. He joined Intel in 1995. He was the architect, designer, and implementer of Intel's first on-chip distributed regulation system. He then went on to be the architect and designer of an image sensor chip, and he researched CPU clocking, power delivery, packaging and signaling at Intel's microprocessor research labs before joining ATD-Pathfinding. He is currently responsible for project management in strategic initiatives enabling processor power delivery and IO. Prior to joining Intel, Raj spent about eight years developing and implementing machine automation systems, test and measurement instrumentation, and signal conditioning and data acquisition systems. Raj holds nine US patents and has numerous publications pertaining to voltage regulation, digital image sensing, digital clock distribution, and high-speed signaling. His e-mail is raj.nair@intel.com.

Vijay Wakharkar joined Intel Corporation in January 1991. Vijay received his B.S. degree in Metallurgy from The College of Engineering, Poona, India in 1982 and his Ph.D. degree in Materials Science and Engineering from SUNY at StonyBrook in 1989. He is currently managing the Materials group responsible for polymers and heat spreader materials and supplier development within the Assembly Technology Development Group. Vijay has worked at Intel for eleven years on materials development projects supporting the various package technology efforts within ATD ranging from TCP, PPGA, PLGA, Cartridge (SECc), and Flip-Chip Technology. Prior to working at Intel, Vijay spent two years as a Post Doctoral Associate at the IBM Alamden Research Center in San Jose. His e-mail is vijay.s.wakharkar@intel.com.

Johanna Swan received her B.S. degree in Mechanical Engineering in 1984 from Northern Arizona University. She joined Intel in 2000 as part of Assembly Technology Pathfinding. She works in the areas of stacked die and stacked packaging, as well as in the area of optical packaging to find solutions for emerging technology needs in the wireless communications arena. Before joining Intel she worked at Lawrence Livermore National Lab for sixteen years on projects ranging from X-ray diagnostics, laser beam delivery, magnetics and EUV lithography. Her e-mail is johanna.m.swan@intel.com.

John Tang received his B.S. degree in Chemical Engineering from the State University of New York at Buffalo in 1981 and his M.S. degree in Chemical Engineering from Northwestern University in 1983. He joined Intel in 1995 in the Assembly Test Materials Organization. He worked with substrate suppliers on the technology development and certification of the PLGA substrate. He also worked on the HVM ramp of PLGA in Intel factories worldwide. He then joined the Assembly Test Subcontract Group as a project manager, and was responsible for managing the assembly subcontractors for Intel's non-CPU products. He is now working at Assembly Technology Development pathfinding on substrate pathfinding and a long-range roadmap. Before Intel, John worked for IBM for eleven years. His experience included PCB board manufacturing, PCB board development, SMT development, Main Frame assembly development, direct chip attach development, and supplier management. His e-mail is john.tang@intel.com.

Gilroy Vandentop received his B.Sc. degree in Chemistry from the University of Alberta in 1986 and his Ph.D. degree in Physical Chemistry from U.C. Berkeley in 1990. He currently manages a packaging research group in Chandler, AZ, within Intel's Components Research organization. His group works in the areas of optical packaging and other novel packaging architectures to enable advancements in thermal, mechanical, and electrical package

performance. Before joining Components Research in packaging, Gilroy worked in Portland Technology Development for ten years, working on silicon process development in the areas of etch and photolithography. His e-mail is gilroy.vandentop@intel.com.

For further information visit:

developer.intel.com/technology/itj/index.htm