# Escape Routing in Modern Area Array Packaging: An Analysis of Need, Trend, and Capability

Bhanu Jaiswal, *Member, IEEE*, Mihir K. Roy, and Albert H. Titus, *Member, IEEE*

***Abstract*—With the increasing complexity in the die and package designs and ever increasing cost pressure in today's microelectronic industry, the design for input/output (I/O) routing has assumed a vital role in the overall product design. This scenario is primarily driven by the increase in the I/O terminal counts in both die and package. Several authors have already described the possibility of using various escape routing models in order to maximize the number of I/Os in a given area. However, these models suffer from many drawbacks and fail to address the importance of processing factors and the actual manufacturing conditions. Therefore, a new design guideline for escape routing has been developed to achieve the maximum I/O density under the actual manufacturing, processing and cost related constraints. The correlation between the real world constraints and their impact on I/O routing has been explored and used as a foundation for developing design guidelines. This approach has been presented through a comprehensive case study that covers various design scenarios, provides the right set of real world trade-offs that need to be considered and simultaneously highlights the drawbacks in existing models.**

***Index Terms*—Area array package, die-size, escape routing design, input/output (I/O) terminals, layer-count, micro-via, package substrate, trace width and spacing, two-layer routing.**

## I. Introduction

THE ability to place extremely large circuits on a silicon chip with continuous advances in the fabrication technology has resulted in high input/output (I/O) counts both at the chip and the package level. The I/Os which consists of signal, power and ground terminals in electronic packages have significantly grown over the past few decades demanding advanced packaging technology. Over the past three decades, the complementary metal–oxide–semiconductor (CMOS) transistor density has been scaling according to the Moore's Law which states that the number of CMOS transistor in the same chip or die area will double every two years. With the current and most advanced CMOS fabrication process it is now possible to fabricate integrated circuits (ICs) with billions of transistors on a single silicon chip and this has enabled placing huge amounts of circuitry in a very small die area. As circuits on a single die have grown bigger, the number of I/Os required for these large integrated circuits have also grown in number considerably. Therefore, it is evident that these advances in the IC fabrication process as well as in the performance of electronic devices have resulted in high I/O counts both at the chip and the package level, which demands continuous development of new packaging technologies to accommodate these high I/O counts [1].

A large number of I/Os is an important consideration in the physical design of integrated circuit packages and board level routing, especially in area array packaging which mainly comprises of flip chip (FC) technology at the die level and ball grid array (BGA) packages at the electronic package level. While both BGA and FC technology provides a good solution for high I/O count and high-density component packages, it also presents new challenges for substrate or package level routing since area array routing is very complex and it directly impacts the cost and manufacturability [2].

A regular area array or a grid of I/O is inherently limited in its routability or the ability for traces to escape out from the array pattern. Escape routing complexity in area array packaging has been identified as one of the fundamental problems often leading to highly complex physical designs of package substrate. These new routing challenges did not exist until few years back with perimeter based packages, and package design was considered to be the relatively simple phase in the overall design process. The situation since then has changed considerably and escape routing design in area array packaging has started taking center stage in most design related problems.

As we will show in next few sections, the escape routing is not just a simplistic approach of geometric design, but it also needs to take into account several processing and manufacturing capabilities. In several circumstances although it is possible to geometrically incorporate a finer trace or traces of certain dimension in between the two bump-pads [3], in practice this is limited by manufacturing capabilities. Typically in a manufacturing environment several trade-offs need to be done to achieve the final product performance. A case study demonstrating the impact of die size and trace width and spacing on number of I/Os and routing layers has been provided while emphasizing the importance of manufacturing constraints. This case study also focuses on developing a sound understanding of these real industry-specific constraints which also serve as guidelines in the most routing design and analysis problems. As the industry is usually in a fluid situation, so instead of drawing a comprehensive conclusion an evolving scenario is provided along with some general conclusion from the case study analysis undertaken.

## II. Substrate Routing Demand and Challenges

In typical computer architecture, the processing of data happens inside the processor or CPU (Central Processing Unit). The flow of data going in and coming out from CPU is managed

Manuscript received July 31, 2008; revised March 18, 2009; July 20, 2009. First published December 04, 2009; current version published February 26, 2010. This work was recommended for publication by Associate Editor M. Cases upon evaluation of the reviewers comments.

B. Jaiswal and M. K. Roy are with Intel Corporation, Chandler, AZ 85226 USA.

A. H. Titus is with the Department of Electrical Engineering, the State University of New York at Buffalo, Buffalo, NY 14260 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TADVP.2009.2035304

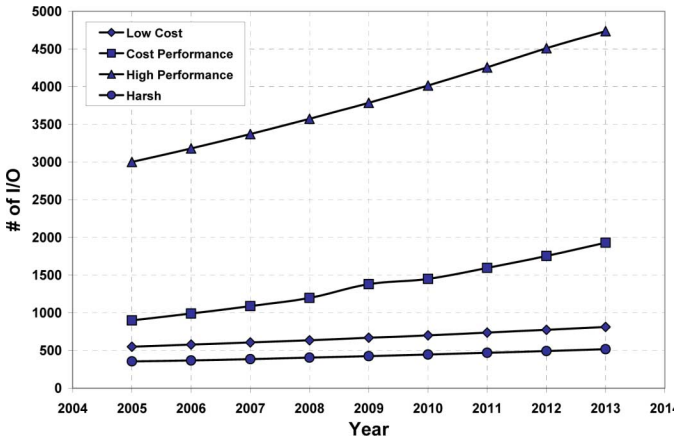


Fig. 1. I/O count for various packages over time of 2005 to 2013 (Source: ITRS Roadmap). Low Cost—Value segment where low cost is the main driver. Cost/Perf—Performance is an important criteria but determining factor is still the cost. High Perf—Performance is the main driver. Harsh—Typically used for high temperature and harsh environment like automobile, furnaces and military applications.

by a memory controller hub (MCH) which is also known in the computer industry as the North Bridge (NB). The MCH/NB typically has high I/O counts and as we move forward technologically, these high I/O counts of MCH are being integrated into the CPU chip for cost and performance reasons [4]. The most recent ITRS (International Technology Roadmap for Semiconductors) roadmap predicted an increase in I/O density count of more than two fold from 2005 to 2013 mainly in the performance CPU devices (see Fig. 1). In the past, the response of IC industry to an increase in I/O counts and associated IO routing requirement was to move away from peripheral based routing to area array routing. However, this area array based packaging is approaching a limitation in its ability to provide finer pitches and escape routes to IO bumps deep within the array. Therefore several designers are trying to route more and more I/Os by rearranging the patterns of the bumps, or by routing I/Os through multiple layers [5], [6].

To accommodate higher I/O counts, one of the potential solutions is to increase the die size; however this incurs heavy cost penalty on the silicon side. A deeper I/O pattern while maintaining smaller die size could be considered as an option, but this result in more package layers needed for escape routing often leading to an increase in cost. Therefore, increased die size versus substrate layer count cost trade-off analysis must be done before considering this option. In order to restrict package layer count increases, a deeper I/O pattern mandates very narrow widths and spacing of the routing traces also known in the industry as fine line and space (L/S) routing. The fine L/S can enable a reduced I/O bump pitch while giving an improved routability. This reduced bump pitch can also be achieved by reducing the pad size which requires tighter lithographic alignment capability in the substrate manufacturing process. Increased numbers of layers for routing is also dictated by stacked via capability. Higher numbers of layers would need micro via technology with higher numbers of via stacks. Unfortunately, this often comes with a penalty of reduced via reliability and via delamination problems. Stacked vias also require the micro-vias to be filled with metal which again adds to the manufacturing cost. To avoid these issues, the staggered via design is often used, which then shifts the challenge of accommodating increased I/Os back to the routing designs and designers [7].

As pointed out above, the ability to manufacture extremely fine escape routing traces with narrow trace widths and fine spacing between them can help in routing out a deeper and denser I/O pattern. Using the fine L/S, a designer can allow more lines on the same package layer to escape between the bump patterns which significantly helps in high density I/O routing in minimal number of package layers. However, fine L/S from the manufacturing standpoint forces the package metal routing layer (typically copper) thickness to go down because the patterning of fine L/S necessitates the usage of thinner dry film resist for patterning the Cu in the build-up process of package substrate [8]. This in turn causes limitations on impedance control and high speed signaling. This happens because as the copper layer on substrate gets thinner, the thickness tolerance as a percentage of layer-thickness' increases dramatically and deteriorates the impedance control over the length of the routing trace. A significant impedance variation can increase the return loss for high speed signaling. A finer spacing between adjacent traces is another issue as it may lead to increased crosstalk for single ended signals. A higher return loss and crosstalk can lower the I/O channel performance at high frequencies thereby forcing the signal integrity engineers to lower the speeds of data busses. This in turn will limit the data transfer rate of I/O channel which can be highly unacceptable due to ever increasing demands for higher data bandwidths. Thus fine trace width and spacing may sound like an attractive option to enable high density I/O routing but it has its own penalties to the high speed signaling performance of the package.

## III. Analytical Models Available to Designer and Their Limitation

Numerous researchers have conducted routing studies using several I/O placement optimization methods and algorithms. Most of these methods have taken only the geometrical aspects into account. Among those methods, a novel method pioneered by Jaiswal and Titus, *et al.* [5], [6] studied the process of I/O pattern rearrangement to optimize the amount of area covered by the routing under the die shadow area for several patterns such as horizontal only routing, hexagonal array routing, and square grid array routing. This process takes advantage of minimization of the floor space through geometrical rearrangement of bumps, known as "Ball-shifted-as-needed" while creating additional space for escape routing. Fig. 2 depicts that for higher bump count, horizontal only routing only is most effective way of pattern rearrangement giving the maximum routing density. These geometrical factor based estimations are actually the theoretical highest number of I/O feasible, and is never seen by any packages in the entire industry. This argument therefore demands an investigation of the capability versus the theoretical maximum which is important from industrial stand point.

Typically the limiting factor for implementation of theoretically highest number of bumps comes from the processing ca-

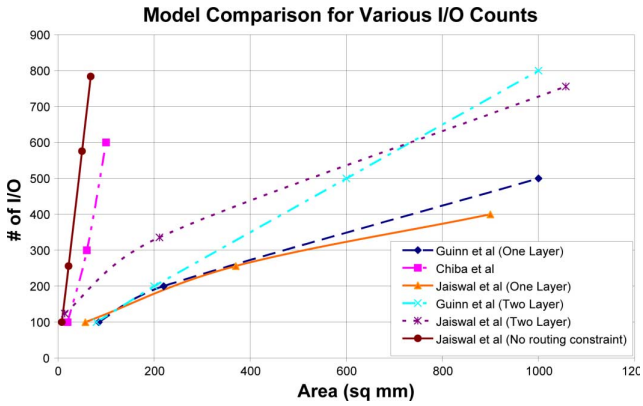


Fig. 2. Model based routing efficiency comparison for high I/O count packages [5], [6], [9], [10].

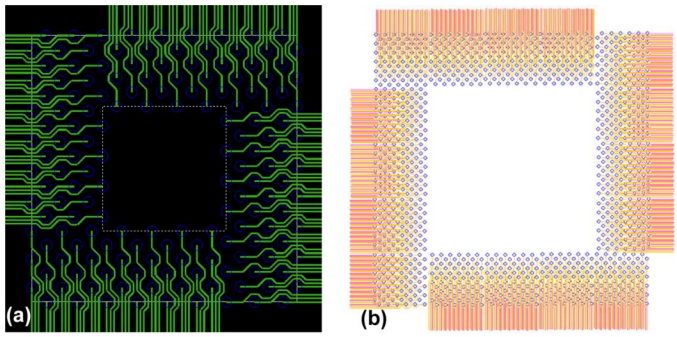


Fig. 3. (a) A single layer routing, and (b) double layer routing with two different color representing two different routing layers.

pability (some of which are described in previous section). In the method described in [5] and [6], the most effective method, horizontal only routing, could only be utilized effectively if the floor space of silicon and substrate bumping matches one-to-one. This one-to-one condition is also true for two other methods (hexagonal array routing and square grid array routing) studied in the process [5] and [6]. Also, these method possesses several other processing challenges, in addition to one-to-one bump matching: 1) the impact of nonuniform bump pattern might affect epoxy underfill process due to effect of new bump positions on underfill (UF) flow properties, the process which is an integrated process of die attachment to package; 2) FLI (first-level-interconnect) formed by the C4 (controlled collapse chip connection) joint in FC technology [1]) reliability issue due to the modification of C4 bumping position; 3) challenges in deflux that might arise when the bumps come too close to each other locally thereby inhibiting the deflux chemicals to reach there which might cause reliability issues, etc.

These methods also describe routing that can be done using one-layer or two-layers. One-layer routing will mandate very thin L/S, increasing the cost significantly, whereas, two-layer routing requires cost of adding an extra layer and requires more vias required for routing. Fine L/S will also add some line resistance. So, there is a cost and performance trade-off for one-layer versus two-layer routing. Under the current industry scenario the cost trade-off appears to be the predominant factor in choosing the number of layers for routing.

Another approach for I/O maximization investigated by Guinn and Frye [9] considers the packing density to be limited by the maximum packing density of circular features of constant size. The assumption behind their study that all features may be regarded as circular is a logical, but an arbitrary decision. By dividing the total area available for traces and pads by the area occupied by the features, Guinn and Frye develop the equation for maximum I/O. This reciprocates strongly with Chiba's [10] approach that result in an equivalent equation dominated by two empirical constants: the effective channel utilization ratio and an empirical constant nearly equal to one which is a technology dependent constant based on an expression of how effectively the space available for routing tracks is utilized. This difference in results of both models is due to the way in which Guinn and Frye try to describe the packing density limitations that arise from geometric constraints. Their model is a function of the number of layers, the number of vias, the channel utilization, and feature spacing, router efficiency and via placement rules. These geometric considerations that essentially impact the packing density have not been considered in the Chiba's model. Chiba's model on the other hand assumes ideal and constraint less conditions for packaging density which compares very well with the specific case of Jaiswal and Titus model where the routing constraint has been assumed to be absent. However, the model by Guinn and Frye appears to over estimate the I/O counts in comparison to Jaiswal and Titus model for both one-layer and two-layer routing (see Fig. 2).

## IV. Analytic Investigation

To investigate the impact of various manufacturing components on escape routing, a case study was undertaken by changing some of the major components in the package that primarily impact escape routing. Some of those input factors are die size, trace, and L/S, routing layer count, etc. Since this approach essentially provides the limits for cost trade-off, this approach along with the processing costs is expected to provide the guide lines for future routing investigation. In addition, with this model the impact of one single variable's impact on overall result can be more easily comprehended than the two other methodologies investigated by the two groups of researchers. For this investigation the basic software used is APD (Advanced Package Designer software, version 2.0) from Cadence Design Systems, with some minor modification to it. The pad size was fixed at 120 $\mu$m diameter. The two cases of line width and spacing investigated here are 20/20 $\mu$m and 15/15 $\mu$m. The pad to line spacing follows the same rule as line to line spacing and the pitch has been adjusted to provide a deeper I/O pattern for a larger die size.

### A. Impact of Die Size on Routing Layer Count

An analysis has been provided for the maximum number of bumps that could be routed using area array design as the die size changes. For this analysis, the substrate size is fixed at 40 mm $\times$ 40 mm and the die size is varied. 40 mm $\times$ 40 mm is chosen as it represents the substrate size for mainstream desktop and notebook processors available in market today. Fig. 3 shows a typical figure generated using the design software employed for this study. In Fig. 3(b), different colors are assigned to two

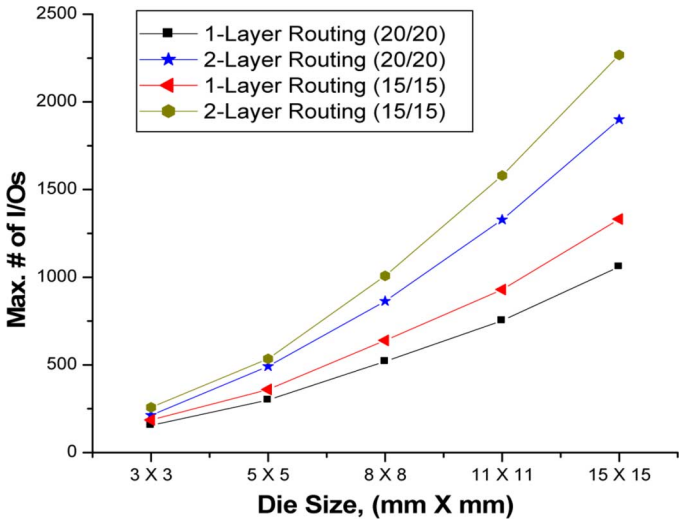


Fig. 4. Maximum number of I/O that could be routed using single and double layer routing for various die sizes.

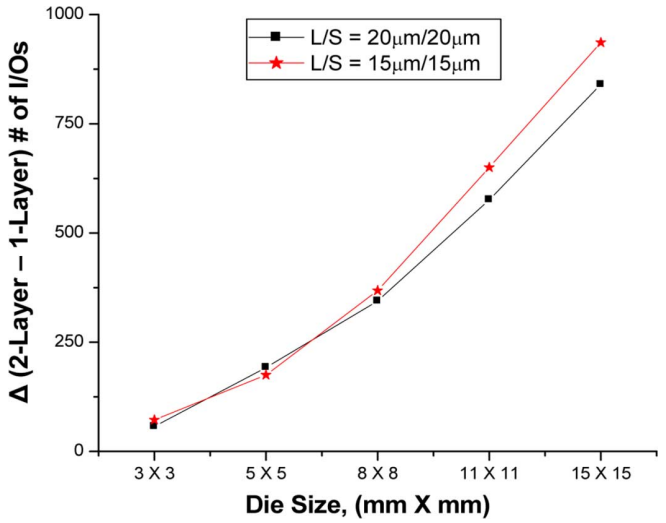


Fig. 5. Maximum I/O count difference between single layer and double layer routing for $L/S = 15/15 and 20/20 versus die size in mm × mm.

different layer of routing to depict the routing design much more clearly.

The plot for maximum number of bumps that could be routed as the die size is varied is shown in Fig. 4 for L/S of 20/20 $\mu$m and 15/15 $\mu$m with one and two layer routing. These two L/S dimensions are considered to be the leading edge technology in routing dimensions for organic packages. The Fig. 4 depicts that with increasing die size, the number of bumps that could be routed increases, which is an obvious result. However, for the one-layer routing, the limit of the maximum number of bumps that could be routed is much lower than that of the two-layer for both L/S dimensions considered. For both L/S values considered, the maximum number of bumps that could be routed using a one-layer for a die size of 15 mm × 15 mm is much lower than the estimated number of I/O provided by the ITRS roadmap [ref. ITRS 2005]. Therefore a two-layer routing approach is necessary in order to overcome the limits maximum created by the one-layer approach to routing. The number of I/O shown in Fig. 4 is estimated using the leading edge technology which far exceeds the I/O counts estimated by theoretical models. This trend is expected as the theoretical models in Fig. 2 are based on input parameters derived from older substrate technology generations.

### B. Impact of L/S

The difference between the maximum I/O feasible using single layer or double layer routing for L/S 15/15 and 20/20 over various die sizes is shown in Fig. 5. The increasing difference between one-layer and two-layer routing with increasing die size is quite notable. Also, it can be seen that for smaller L/S (i.e., 15/15) the difference in more pronounced for bigger die size, implying maximum I/O count using two-layers is higher than that feasible using one-layer count. This makes multilayer routing an attractive option for bigger die sizes.

One of the explanations for the increase in I/O counts that can be routed is due to the increasing I/O depth with double layer routing. I/O depth is defined as the distance between the centers of the innermost pad that is routed, to the edge of the die. Fig. 6 shows the plot for I/O depth versus the die size. For

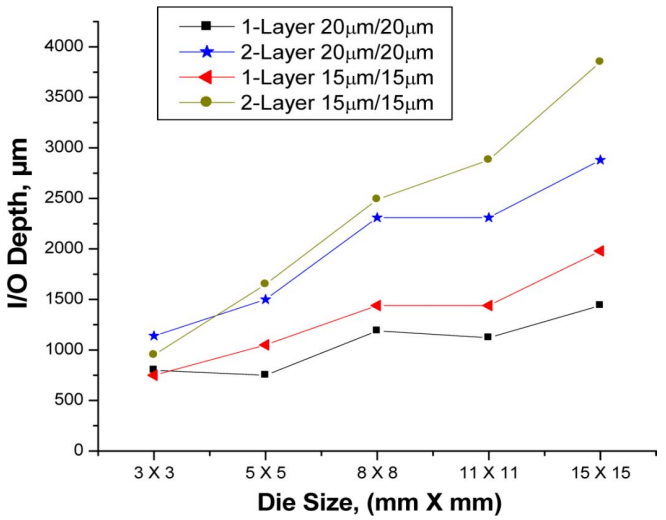


Fig. 6. Maximum I/O depth for L/S of 15/15 and 20/20 versus the die size.

smaller L/S (i.e., 15/15 in this case) the I/O depth is more than that of larger L/S (i.e., 20/20), as more traces can pass through the pads of same size and pitch. However, for dice of very small sizes (lower than 5 mm × 5 mm) this may not be entirely true, as two factors that come into play The first, the *Integer Rule*, can be explained by the following.

*The integral increase in number of traces that can be routed between two pads results in higher I/O counts, and any fractional increase in number of traces routed in between two bumps due to narrower L/S adds no additional value to I/O counts.*

Second, the traces that are routed near the corners are affected by the *Corner Rule*, which can be explained as.

*The I/O near the corner of the die can be fully routed using only one escape routing side and the presence of other escape routing side does not provide any benefit to further increase the I/O counts.*

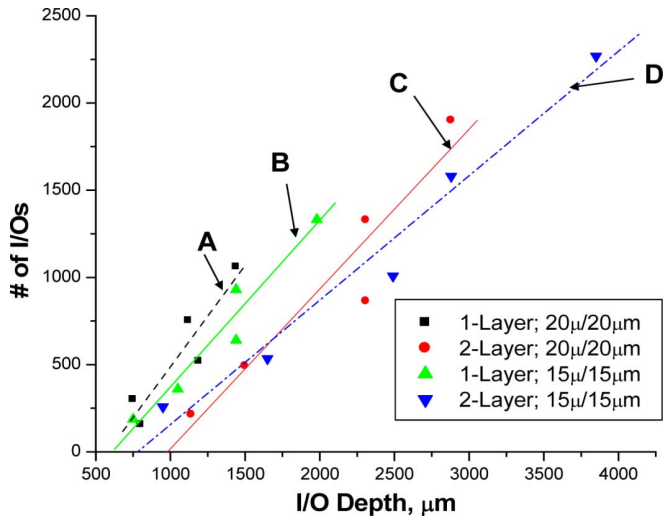


Fig. 7. Maximum number of I/Os feasible versus the I/O depth for 15/15 and 20/20 L/S using single and double layer routing. A = 1 layer, 20/20; B = two-layer, 20/20; C=one-layer, 15/15; D=two-layer, 15/15.

For very small die sizes, the corner effect is more pronounced compared to a larger die. These two rules could possibly explain why there is a cross over for 20/20 two-layer routing and 15/15 two-layer routing. Another point to be noted is that the length of the traces used in industry are sometimes as large as 3–4 mm, which is comparable to that of the L/S dimension mentioned here. This suggests that the scenario depicted here is within the range of the manufacturing and processing capability and is realistic. The three earlier models proposed by three different group of researchers [5], [6], [8], [9] did not take into account the maximum length of L/S feasible through actual industrial capability.

The impact of the I/O depth on the maximum number of I/O feasible under current industrial capability scenario could be better explained using Fig. 7, which depicts the maximum number of I/O feasible for 15/15 and 20/20 L/S versus the I/O depth. The lines in the figure are the best regression fit to the data generated using APD software. The slope of these lines indicates the change in the number of I/O for a given change in I/O depth. Thus, a trend-line with a lower slope indicates that a greater change in I/O depth is needed to achieve the same increase in number of I/O as compared to a trend-line with a higher slope. By comparing the four lines, it can be seen that D (L/S of 15/15 with two-layer routing) has the lowest slope, followed by C (L/S 15/15 one-layer routing), which is also very close to B. Thus with 15/15 L/S two-layer routing the increase in I/O depth must be greater than other L/S and routing layer combinations to achieve the same increase in number of I/Os. The increase in I/O depth could be possible only if the packing density of I/O for 15/15 L/S with two-layer routing is much closer to its theoretical maximum value than that of their counterparts. Following same analysis, one-layer routing for 20/20 L/S is the farthest from the theoretical packing density. This shows that the efficiency of routing for smaller L/S and multiple-layer routing is higher than that for larger L/S and one-layer routing.

For I/O depth of 1700um there is a cross over between fitted lines of two-layer with L/S 15/15 and two-layer routing of 20/20. I/O depth smaller than 1700um is sufficient for routing die sizes less than 5mm $\times$ 5mm;for smaller dice, finer L/S does not provide any advantage. This is also one of the conclusions derived from Fig. 6.

## V. Emerging Picture

Since the addition of an extra layer for additional bump routing incurs more cost, a cost analysis of single versus double layer routing is essential. Typically, the cost for two-layer routing is $\sim$20% higher than for one-layer routing, even if the number of I/Os needed is at the maximum I/O feasible for single layer routing. This suggests that two-layer routing should be used only when the required number of I/Os could not be achieved by one-layer routing. However, in practice, it has been noted widely [1] that many PCB producers actually use two-layer routing even if the number of I/Os is below the one-layer threshold limit. This implies that as one approaches the threshold limit of maximum number of I/O, other reliability concerns may become the limiting factors. Such factors might be reliability of fine lines, resistance variation, dielectric delamination, etc., most of which has its origin in either material or processing factors.

There may be other concerns with one-layer routing with finer lines, such as yield issues that originate from stretching the limits of implementing finer L/S structures in substrate manufacturing. This will require much more tightly controlled processing and more sophisticated enabling equipment to implement finer line without a decrease in yield. Nevertheless, implementation of the finer lines leads to additional costs originating from either yield loss or equipment cost. Therefore, the manufacturers prefer to implement a two-layer routing which provides more room and can be implemented with similar functionality using coarser and, thus, more economical design rules.

## VI. Conclusion

From the routing case study provided some general conclusions can be derived. To support high I/O counts in future technologies while limiting the die size to a small size, the escape routing can be implemented by fine L/S and multiple layer routing. In case of large die packages, a coarser and single layer routing may be enough, but, if this is not the case, then a cost trade-off can provide a guideline regarding whether to go for single layer routing with finer L/S or coarser L/S and multiple layer routing. This is also true for very small die with small I/O counts and coarse bump pitch. Implementation of finer L/S has various processing and reliability concern which creates a threshold boundary for maximum I/O depth and maximum packing of traces.

## References

[1] R. R. Tummala, *Fundamentals of Microsystems Packaging*. New York: McGraw Hill, 2001.

[2] T. C. Chung, F. Ghahghahi, and B. Oberlin, "Area array packaging technologies for high-performance computer workstations and multiprocessors," in *Proc. 46th Electron. Compon. Technol. Conf.*, May 28–31, 1996, pp. 902–910.
[3] C. S. Patel, K. P. Martin, and J. D. Meindl, "Optimal printed wiring board design for high I/O density chip size packages," *Circuit World*, vol. 25, no. 4, pp. 25–27, 1999.
[4] C. Natarajan, B. Christenson, and F. Briggs, "A study of performance impact of memory controller features in multi-processor server environment," in *Proc. 3rd Workshop Memory Performance Issues: Conjunction 31st Int. Symp. Comput. Architecture*, Munich, Germany, pp. 80–87.
[5] A. H. Titus and B. Jaiswal, "A visualization-based approach for bump-pad/IO-ball placement and routing in flip-chip/BGA technology," *IEEE Trans. Adv. Packag.*, vol. 29, 3, no. 3, pp. 576–586, Aug. 2006.
[6] A. H. Titus, B. Jaiswal, T. Dishongh, and A. Cartwright, "Innovative circuit board level routing designs for BGA packages," vol. 27, no. 4, pp. 630–639, Nov. 2004.
[7] J. H. Lau and S. W. Ricky-Lee, *Microvias for Low Cost, High Density Interconnects*. New York: McGraw-Hill, 2001.
[8] Y. P. Wang and T. D. Her, "Advanced substrate technology for ball grid array in electronic packaging," in *Int. Symp. Electron. Materials Packag. (EMAP 2000)*, 2000, pp. 335–339.
[9] K. V. Guinn and R. C. Frye, "Flip-chip and chip-scale I/O density requirements and printed wiring board capabilities," in *Proc. 47th Electron. Compon. Technol. Conf.*, May 1997, pp. 649–655.
[10] T. Chiba, M. Yamada, and F. Kobayashi, "Limitation of the signal pin density on wiring boards," *IEEE Trans. Compon., Packag., Manufacturing Technol.-Part B*, vol. 19, no. 2, pp. 391–396, May 1996.

**Bhanu Jaiswal** (M'06) received the B.S. degree from the Government Engineering College, Jabalpur, India, and the M.S. and Ph.D. degrees from the State University of New York (SUNY), Buffalo, in 2003 and 2006, respectively.

He works as a Senior Packaging Engineer with Intel Corporation. His current research interests involve high-speed signaling validation in package interconnects and package design for low-cost manufacturing. Other areas of research interests and expertise include substrate design rules, flip chip IO routing and electromagnetic modeling of packaging structures.

**Mihir Roy** received the Ph.D. degree from Rensselaer Polytechnic Institute (RPI), Troy, NY, in the field of nanodielectrics an interdisciplinary degree of electrical engineering and material science.

Currently, he is working in Intel Corporation as a Senior Packaging Engineer. His current interest includes undertaking basic research and pathfinding activates at Intel, manufacturing and design interactions, supplier management, and overseeing capacity expansion.

**Albert H. Titus** (S'86–M'97) received the B.S. and M.S. degrees from the University at Buffalo, Buffalo, NY, in 1989 and 1991, respectively, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, in 1997.

He is an Associate Professor in the Department of Electrical Engineering at the State University of New York at Buffalo. Prior to joining the faculty at Buffalo, he was an Assistant Professor at the Rochester Institute of Technology. His research interests include analog VLSI implementations of artificial vision, hardware and software artificial neural networks, hardware implementations of decision-making aids, optoelectronics and integrated sensor systems.

Dr. Titus has numerous research grants from federal and private sources, including an NSF CAREER award. He has been a reviewer for many journals and conferences. He is a member of INNS, ASEE, and SPIE.