Integrated Systems

Dirk J. Bartelink*
Ultra Large Scale Integration Laboratory
HPL-95-130
November, 1995

bartelink@src.org

integrated circuits,
multi-chip assembly,
testing

Integrated Systems are defined as batch-fabricated interconnections of complex digital integrated circuits with analog interface circuits and transducers, such as sensors. By providing the cost, performance and reliability levels of monolithic integration, they offer potential advantages over multi-chip modules assembled with packaging technology. This paper studies the required process technology, as well as design, test and packaging issues, for integrating wide varieties of systems. The goal is to delineate the necessary steps in bringing Integrated Systems to market within a realistic period. With monolithic integration as the ultimate aim, a multi-chip entry point is identified that can start system technology on a learning curve of cost reduction using the same scaling principles that drive integrated circuits. Three challenges to be surmounted are identified in streamlining the I/O's and progressing along a learning curve, namely I/O scaling, I/O loading, and full-functional test. The 'composite IC' is the entry point. A large chip, containing only global interconnects and power distribution, acts as a silicon backplane. Subsystem-chips, such as digital microprocessors or sensors, are flip-chip mounted using the accuracy of MEMS processing to fabricate 'snap-together' physical and electrical interfaces with high reproducibility. While similar to conventional MCMs, this chip-to-chip connection has few compromises over on-chip connections. By keeping the fabrication responsibility within one organization, just as in monolithic chips, there is no need for incoming inspection. Added ESD protection and test-head loading are avoided on interior nodes by a new intrafactory method of testing.

Internal Accession Date Only

*Dirk J. Bartelink is with Hewlett-Packard Laboratories and is currently on Industrial Resident assignment to the Semiconductor Research Corporation, Research Triangle Park, NC 27709

I. INTRODUCTION

While the National Technology Roadmap for Semiconductors (NTRS) [1] predicts impressive advances for digital CMOS integrated circuits over the next decade and beyond, it does not specifically deal with their interface to the real world which works mainly with analog functions. This connection must be made through analog/digital conversion. For these applications, particularly in low-cost systems made possible by digital core logic and memory, the cost of connecting the core to transducers, such as sensors or actuators, must also be reduced.

With major advances in digital electronics having become commonplace, the lack of equivalent progress in integrating these application interfaces with equal cost-effectiveness is now very conspicuous. This paper explores technological alternatives and implementation requirements for such Integrated Systems, particularly the challenge of dealing with the wide variety of real-world systems, as well as the issue of their design and test.

Scope of Integrated Systems

The digital revolution and low-cost CMOS make it economical to build analog systems using a complex digital core surrounded by a minimal transducer layer. As the cost of the core is reduced by the learning curve effect of mainstream monolithic technology, optimal system design demands that equal progress be made for the transducer layer. The technology of this layer must likewise progress along a learning curve similar to monolithic integration. It is already evident for straight digital ICs, however, that no amount of hardware technology development by itself can result in products without the availability of designs through appropriate and timely design methodologies and cost practices. Thus, although analog system design presents the additional burden of integrating analog circuits and transducers, the true challenge of Integrated System design ultimately lies in meeting the complexity and variety of system products through a comprehensive design methodology.

An Integrated System, or effectively a 'system on a chip,' is defined as follows:

> *It is an intimate interconnection of mainstream high-complexity digital IC technology and analog system technologies, such as A/D converters and sensors, in a cost-effective scaleable batch-fabricated manufacturing approach, with the performance and reliability levels of monolithic integration. The design of its system functions is integrated to make effective use of the technological resources and is executed using efficient design tools. It uses, as needed, technologies that can be made competitive with best-of-breed technology of individual system components.*

Given the demanding requirements of this definition, there are today not many examples of true Integrated Systems that can be used in developing a general approach. An appropriate vision for the implementation of Integrated Systems, i.e., new concepts that can be used to achieve the end goal, is to emphasize innovative combinations of available technologies rather than relying solely on long-term high-risk new technologies with no proof of their manufacturing worthiness. A step-by-step approach is envisioned that allows market penetration to bear the cost of developing the ultimate solution.

As defined, an Integrated System may be a monolithic chip, but it is not necessarily one. Nor is it an MCM in the conventional sense. Advanced packaging solutions have their own avenue to system integration as extensions of current materials systems and manufacturing practices and conventionally work through the traditional I/O pads of separate chips. Integrated Systems instead directly address the underlying Si technology and are not likewise restricted. Many of the advantages of conventional MCMs, particularly their ability to combine widely divergent component technologies, are so compelling that they cannot be dismissed. Other assembly techniques with this same benefit form the centerpiece of the present approach to Integrated Systems, under the name "pseudo-monolithic integration."

Over time, passive components in ICs have been replaced with active ones, the last major step being TTL. Only the DRAM capacitor is actively developed. Integrated Systems, on the other hand, deal with all required components for a self-contained electronic circuit. In advanced systems, decoupling capacitors, terminating resistors or other components such as inductors cannot be avoided.

A trio of challenges in system assembly is identified as limiting MCMs or any other packaging initiative. These technological limitations must be surmounted in order to realize successful Integrated Systems technology.

1. *I/O Scaling* - Conventional flip-chip area-array packaging solutions improve on perimeter connections by providing more I/O's, but they almost always utilize rigid attachment between chip and substrate. With the commonly used, roughly spherical solder balls (C4 technology), this approach limits I/O density scaling to about 100 micron pitch because of the need to accommodate thermal expansion differentials for large chips.
2. *I/O Loading* - As currently practiced, chip-to-chip electrical interconnections internal to a system are designed to interface chips made by multiple suppliers. There is a business need to verify the functionality of incoming chips through tests applied to the pads. ESD protection circuits and the drive capability for interfacing a test head are then unavoidable. These internal I/O's are subjected, mainly by these factors, to significant loading as compared with a simple 'on-chip' wire, in speed, power and I/O circuit pitch.
3. *Full-Functional Test* - The prominent Known-Good-Die (KGD) problem must be resolved for assembled systems to succeed. At-speed testing to verify KGD specifications without package insertion, particularly at high speeds or for mixed signals, is not effectively handled by today's test and packaging strategies. The corresponding lack of feedback to the wafer fab related to speed-performance or low-signal-level failures also gives insufficient yield-learning information. This feedback loop includes the time delay in sending wafers to another location for packaging and test. Product time-to-market is intimately tied to this learning cycle.

Together, these constraints limit cost reduction opportunities in chip-to-chip integration compared with monolithic integration where scaling with successive generations is the norm (per the NTRS).

One answer to the question as to what is so desirable about monolithic integration may therefore be given as the continuous improvement along a learning curve by geometric shrinking of physical dimensions in an industry standard manner. As we shall see, it is not

necessarily contained in the integration of all functions in the same starting wafer. This paper studies this and other conditions for significant cost reduction in chip assembly as a means for realizing Integrated Systems.

II. Objectives of Initiative in Integrated Systems

Why are Integrated Systems not widely used? Their main objective of low-cost system interconnection was already desirable at the beginning of integrated circuits. The answer is that the benefits of scaling and the power of standardization that work for digital circuits do not work for systems because of the wide variety of system hardware configurations. The unique fit of system solutions to the application have made inventiveness and flexibility a priority, while it has been much more difficult to agree on standards.

To introduce Integrated Systems this late in the development of electronics requires that the market place decides on the acceptable technologies and that product sales bear the cost of development on a pay-as-you-go basis. While this is normal healthy business practice, the difference here lies in the emphasis on a well-developed starting point for entering the learning curve of scaled system technologies. Important also is a high profile and a high rate of observable progress to assure a general acceptance of this approach to systems. Only then can basic standards be established that drive the learning curve.

We can now ask what is new at this point in time that might encourage this aggressive approach and what is critical to its success? It is not our sudden recognition of cost reductions from monolithic integration. We must even concede that the newest components are becoming less compatible after many cycles of individual and separate refinement. Yet the pressure to integrate is stronger than ever, driven by cost, low-power and form-factor.

New technology

A possible answer to the question of what is new is the 'composite integrated circuit' (composite IC) which forms the basis of this paper. It is assembled from individual parts fabricated in separate wafers, but behaves like a system on a chip. The composite IC was envisioned in the 1994 NTRS [2] to address the challenges of scaled mainstream IC technology by limiting the necessary increases in stepper field size and wafer yield in manufacturing high-performance systems, while also answering the known-good-die problem. It may, however, find a more compelling application in merging divergent technologies for Integrated Systems.

Market development

Large capitalization demands pay as you go. The approach under consideration still needs much development and its timing is far from clear. The strategy then is to instigate a reasonable cost, limited risk entry point to interest potential customers, as well as providers of technology. The vision, as noted, is to use as much off-the-shelf technology as possible and enter in a modest way, but move decisively before the initial momentum for driving the learning curve dissipates.

Evolution from current practice attracts early customers. Technologists get trapped into developing the 'perfect' solution while ignoring that the market is shifting. In contrast, even an existing near-perfect solution may have difficulties, in spite of having wide acceptance and requiring 'only' capital investment. Entering the market is still

difficult, in view of the risk, because there are not enough developed customers nor distribution channels to roll out the product in quantities sufficient to justify the investment levels. New entries must, therefore, start at an investment level commensurate with the early customer base.

In sum, the entry point and learning curve are key. To succeed, the solution must be widely recognized as the system approach of choice. Still, the first examples should depend on a small set of new technologies and a test market must be chosen that is receptive and has suitable volumes. The automotive market is often cited as having these characteristics.

III. Step-Wise Approach to Monolithic Integration

An early entry point may mean that the initial solution is not the ideal final solution. The strategy is to accommodate both the transition from the entry point to the final state and the coexistence of low-volume, high-variety systems with high-volume commodity system products in a common long-term framework. The first step makes use of the composite IC concept to establish the entry point with pseudo-monolithic technology, while the second is structured around true monolithic integration for volume products.

A. Key factors influencing choice of a two-step strategy

The beauty of digital CMOS is its ability to handle a wide product variety through flexible design in a nearly standard process technology. Integrated Systems also have great need for variety but unfortunately also require flexible process technology to fabricate the interface to the real world.

Monolithic integration

The barriers to monolithic Integrated System are the following:

- Scaling has specialized logic and memory processes to optimize the different functions, while system components, such as sensors, have achieved greater capability through highly tuned individual fabrication techniques.
- Monolithic integration has typically meant significant compromises. Best-of-breed components are made in their own competitive environments. This has made system integration through packaging assembly more attractive.
- The benefit of a large digital core is so compelling that overall system cost demands cost-competitive implementation of this portion, as delineated in the NTRS. Adding any compromising steps will spoil the cost equation for conventional digital technology, which is always pushing its own limits.

Traditionally, in monolithic integration, combining the 'Digital' core with the surrounding 'Analog' interface has been done through one of two options. (Analog in this context is meant to represent all technologies except the mainstream digital logic and memory of the NTRS.) Neither option is ideal:

Monolithic Option 1: Add Analog (A/D, sensors, etc.) circuits to a complex Digital core

The cost effectiveness of microprocessors can be spoiled by processing variations that enable sensors, actuators, and linear circuits. This situation worsens as digital technology tries to maintain the forecast of the NTRS.

Even without this concern, the types of components that can be made by integrating them onto a digital starting point are very limited in number. Typically, at the niche market end, some specialized function provides the marketing edge and cannot be compromised too far. At the other end, low-performance high-volume products sometimes lead to mass markets and can support their own competitive standards. But these are limited to a lucky few winners, while a vast number of products will remain unsatisfied.

Most systems need several 'Analog' components to perform even the most elementary functions in the real world. If one combination of these would satisfy every need, a solution might be found, but the combinations are endless. The engineering cost of all these combinations alone is prohibitive. This problem is compounded when product shipments are small, or at least not predictable.

Option 1, therefore, does not look like a promising approach to Integrated Systems, nor does it have an appeal even in serving as an entry point for a major movement.

Monolithic Option 2: Add complex Digital circuits to Analog components

Adding digital circuits to sensors and actuators already results in substantial product shipments. The airbag accelerometer is a well-known example [3]. These digital circuits, however, cannot be classified as complex on the scale of a microprocessor or high-density DRAM. To keep the cost benefits of scaling, i.e., stay with the NTRS mainstream, it would be necessary to duplicate the technology of these complex parts in the many versions of processes suited to analog function. The economy of scale of standard parts would then be lost.

Programmability is needed to make the product applicable to a wider range of end objectives and to handle communications protocols and analog system functions such as self calibration. Supporting many limited-capability processors is impractical. Standard processors, for which all of the overhead functions such as compilers, software and documentation are carried by many users, are the accepted approach.

Thus, adding enough digital circuits in a specialized process has a limited potential. If small digital circuits are added anyway, the need remains to connect to complex circuitry by conventional packaging means and the product is not a true Integrated System, again leaving a large market segment unsatisfied with respect to cost reduction.

Option 2 also does not have the promise of starting or sustaining an Integrated System initiative.

Both approaches to monolithic integration of high-volume Integrated Systems have been tried extensively, but history has demonstrated that they do not work.

The lesson learned from mainstream CMOS is that product volume drives costs downward by underwriting solutions to difficult problems. Likewise, Integrated Systems

must address mass markets, but the flexibility of technology remains a challenge. Unlike the universal nature of digital technology, systems products are a specific fit to a need and are very much like ASICs. But ASICs do not drive IC *technology*, and therein lies the analogous problem of Integrated Systems: It is almost impossible to justify the investment in application-specific combinations of technologies as demanded by monolithic implementation without the assurance of market acceptance.

In sum, the principal negative impact of combining technologies through monolithic integration, particularly for low-volume products, comes from one of two sources:

1. a *compromise* in the 'best-of-breed' quality of the individual components,
2. the *cost of re-engineering* of (near) best-of-breed components into integrated processes for the wide variety of possible combinations.

If low product volumes are not also comprehended, the development of new technology is driven by a few winners, but these will so different from each other that no synergistic learning curve effect will be initiated and no infrastructure would be generated.

The conclusion we draw is that, while there are valuable attributes of monolithic integration that must be preserved for Integrated Systems, we must also maintain the native technology of the individual components. Similar arguments apply to other issues, including, the preservation of intellectual property of diverse components and the engineering costs associated with developing mixed-signal testing procedures for components embedded in an integrated environment of multiple technologies.

Beyond monolithic integration

What then is the solution to this problem of achieving functional integration without giving up the native environment of its component parts? What must be preserved from monolithic integration to receive its true benefits and what must be brought in from multi-chip module integration that satisfies the diversity of systems? What must be added that is unique to Integrated Systems?

Systems in general are composed of components with special fabrication and test requirements that are difficult and expensive to merge. The composite IC potentially answers this challenge by providing an interfacial layer between diverse components to allow them to be combined freely with a standardized assembly technology. This system technology is scalable and thus sidesteps the dependence on monolithic integration of diverse components, even in low volume applications. In response to the trio of challenges under consideration, it introduces the following:

1. Micro-assembly is performed using Micro Electro-Mechanical System (MEMS) concepts [4]. IC fabrication precision helps in self aligning 'subsystem-chips' to a 'backplane chip' and forming the (demountable) electrical spring contacts that accommodate mechanical variations such as thermal expansion or manufacturing tolerances and enable temporary assembly for test as in a probe card. Subsystem-chips are fully formed by conventional batch fabrication in wafers and 'snap together' with few if any additional processing steps.
2. Assembly is controlled within the IC factory to eliminate the need of ESD protection circuits and other electrical loading such as driving a tester for interior nodes [5].

Perimeter circuits leading to the outside still have conventional ESD and test capabilities.

3. KGD testing and yield learning are simplified through demountable snap-together assembly of the device under test (DUT) and known-good system components. This method brings the power of self test to a low-cost test head and enables a single insertion of the DUT directly from an untested wafer to be fully qualified at the wafer manufacturing site for all functionality, including burn-in [4].

The first two of these relate to I/O scaling and I/O loading which must improve with successive generations. In a monolithic implementation, this wiring is on-chip and there would be no question of such progress. As offered by this micro-assembly, Integrated Systems can gain these scaling benefits without compromise even when the components made separately.

The third challenge, full-functional test, has unique implications for Integrated Systems. If full system functionality is *monolithically* integrated, full system specifications need to be verified, both in initial debugging and in production testing. As more of the system is integrated, it may be expected that the external I/O count goes down. This makes testing of the increased component count even more difficult through these I/O's. In low-volume, test engineering may become prohibitively expensive for a variety of products because the interior nodes of the system are not accessible to diagnostics for mixed mode testing. The micro-assembly approach, by providing access to these nodes, opens up many new opportunities in system testing.

In addition to the benefit of the learning curve of scaling, there are valuable aspects of monolithic integration that have nothing to do with processing on a single starting wafer. It is impossible to design an IC without taking full responsibility for the whole design and modern design tools help assure design integrity. This integrity must also be assured for Integrated Systems, but without requiring true monolithic integration. This also applies to other specification, such as test. What these examples have in common is a single company environment and responsible engineering team attention to all aspects of the product at the initial stages and throughout production. This strict requirement may in time be relaxed to allow multiple teams or suppliers, provided suitable safeguards are in place.

B. Two steps to Integrated Systems

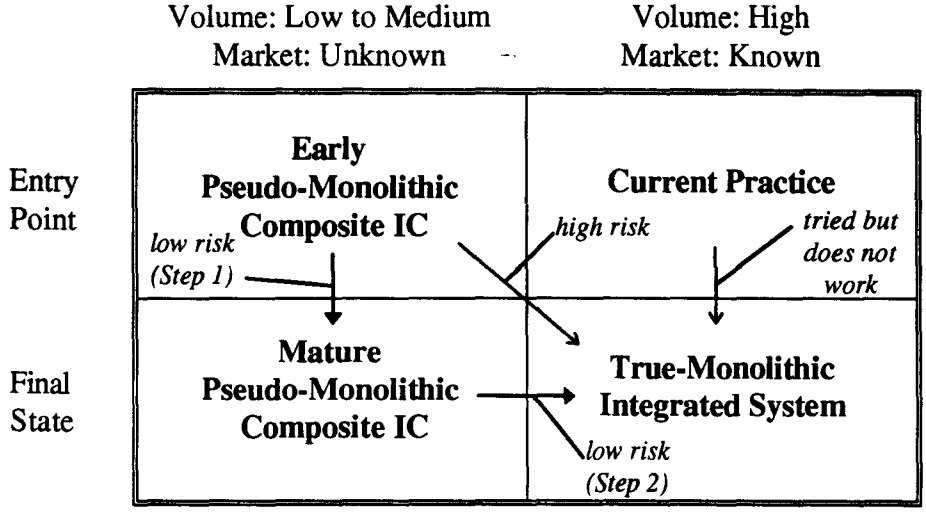
The late introduction of Integrated Systems with monolithic qualities into electronics requires significant rethinking of the approach, but limited exposure to risk. Key factors are market acceptance and the volume learning curve. What is proposed here, as Step 1, is a pseudo-monolithic composite IC for broad product coverage, controlled-risk, and a quick entry point. This is then followed, as Step 2, by true-monolithic products for the volume winners in established markets and with standards set in a solid infrastructure for Integrated Systems.

The meaning of this proposal and the risk containment brought about by it can be understood with the aid of Table 1. The table combines two market conditions (low and high volume) with the early and late stages of Integrated System development. For 'Current Practice,' which is aimed at high-volume markets but is still at the 'Entry Point,'

the direct transition to the monolithic high-volume 'Final State' has historically been tried, as noted, but does not work.

Table I

Risk Factors in Integrated-System Technology Introduction



The table highlights the fact that there are two transitions that need to be managed: between the rows as the maturing of the technology from entry to final state and between the columns as the maturing of the market or volume learning. Taking both transitions at once, on the diagonal, incurs high risk because of difficulty in forecasting the two separate transition events. The low risk path indicated calls for a maturation of a generic pseudo-monolithic technology in Step 1 within its own low/medium-volume market equilibrium. The transition to true-monolithics in Step 2 is taken only when for a particular product the market has been tested and the technology is predictable.

It can be expected that there is a steady-state role for the pseudo-monolithics, namely for those products that do not command a large volume but for which a high level of integration is an essential component, such as for products requiring high speed or high packing density. Unique technologies serving the pseudo-monolithic market can then be developed for long-term application. An infrastructure for this type of product then also serves as a launching point for products aimed at high-volume true-monolithic Integrated Systems.

Step 1: Pseudo-monolithics for broad product coverage, quick start

To snap best-of-breed components together in a flexible manner, the micro-assembly operation must be of a universal and standard specification. A large silicon chip, containing only global interconnect routing together with power and ground distribution, acts as a 'Silicon-Backplane' [6]. Several individual subsystem-chips are flip-chip mounted to this backplane with standard physical and electrical interfaces by using the accuracy and reproducibility of silicon processing. To understand the distinction between this interface and conventional flip-chip mounting, we note that, in advanced lithography, wafer flatness is controlled to better than a wavelength. Mating chips, therefore, is akin to placing optical flats together. MEMS technology can accurately self align their physical position and rotation, so that the issue of reaching a wire from one surface to the other is brought

into the micron-dimension regime of integrated circuits. We shall leave further details of the necessary MEMS steps to the next section, except to note that any added complexity associated with this chip-to-chip interface optimally resides on the backplane side because the subsystem-chips are then minimally modified from their native process.

Infrastructure

The composite IC has its own advantages over true-monolithics. Beyond integrating multiple technologies and improving test, it provides a larger backplane and allows independent optimization of the global wiring and power distribution metalization. A growing market for pseudo-monolithics could directly support the development of Integrated Systems.

The first pseudo-monolithics will come from a single manufacturer to ensure an integrated design. As the infrastructure for pseudo-monolithics builds up, however, multi-sourcing will become necessary. New standards for direct chip interfacing, paralleling but exceeding those of today's conventional chips, will arise.

Step 2: True-monolithics for volume winners, proven markets

For Step 2, individual technologies worthy and capable of integration must be identified. Development costs are high so that, to proceed, the needed confidence may come only from a successful Step-1 product. Success in Step 2 is enhanced if identical system design procedures were used for Step 1, but intellectual property rights and incompatible internal design styles of the components could still present problems. To justify Step 2, the cost of the more complex process and its lower yield, as well as the cost of technology engineering and test development, must beat the cost of multi-chip fabrication and assembly. The more compatible the processes are, the easier the transition will be. Higher volumes also favor true-monolithics, but the confidence of success must be correspondingly higher.

IV. Composite IC Technology

A. Visualizing entry-point Integrated Systems

To help visualize composite ICs, a simple example is presented. It is the least complex case and by no means forecasts the ultimate solutions that may be possible. A technically more sophisticated version of composite ICs has been proposed [7], but presenting it here would detract from the central issues.

Figure 1 shows a composite IC, as mounted in a conventional package, complete with subsystem-chips and perimeter wire bonds attached to the backplane chip. The first enlargement schematically illustrates a self-alignment structure with beveled top edges to guide the subsystem-chips. Once located, as shown, they have a small degree of position freedom to compensate for fabrication inaccuracies, thermal expansion, and system flex. The second enlargement shows MEMS-fabricated area-array springs on the backplane with a central probe tip and solderball contacts on the subsystem-chips. The springs allow for significant up and down travel and accommodate a small amount of in-plane motion.

The interconnect levels of this six-level metal composite IC are partitioned such that Metal 1 through Metal 3 (M1-M3) are on the subsystem-chips and M4-M6 are on the backplane.

The large-area backplane supplies the subsystem-chips with power and supports global communications, including both inter-chip and intra-chip signals. A fortuitous property of ICs aids this plan. For deep-submicron technology, the wiring pitch of the upper levels is relaxed because typically only power and global interconnects are designed into those levels. The larger feature sizes allow use of older style lithography tools, such as projection aligners, or even flat-panel display technologies, which impose almost no chip-size limitations compared with stepper-based lithography. This larger feature size, however, does not detract from the backplane's ability to serve as a high-speed, high-bandwidth communication path for uncompromised performance because one to two micron minimum dimensions are about optimum. They are, moreover, readily available in existing IC factories.

A characteristic of Integrated Systems is that they require many passive components. These consist of the widely used decoupling capacitors and terminating resistors that should be integrated into the backplane. Specialized passive components, such as inductors, are most easily fabricated as a dedicated subsystem-chip.

Subsystem-chips are supplied from starting wafers of varying diameter following conventional wafer sawing, as depicted in Figure 2. They may all be made by one manufacturer or, after suitable standards become established to control the chip interface, from several suppliers. The backplane chip is supplied from appropriately sized wafers, e.g., 300mm wafers or the flat-panel substrates shown in the illustration, but made with processes derived from standard IC and MEMS technology. To complete the backplane, separate alignment chips are attached.

The final step in manufacturing fully qualified composite ICs is test. Component testing to full-functional specifications is significantly simplified if the system consists of a small number of pieces that can be temporarily mated to each other with full electrical functionality. As depicted in Figure 3, the test station consists of a farm of test heads all driven from a single tester. The individual test heads are constructed identically to the composite IC of the product to be tested. Each backplane chip is populated with known-good subsystem-chips except for the DUT. A robotics arm inserts the DUTs. Test costs are minimized by sharing the tester over a large number of test heads. The resulting test cost per unit time is low enough that all testing operations can be performed during the same insertion, including temperature cycling and burn-in.

B. Self Alignment: system assembly and test insertion

An assembly of diverse subsystem-chips forms a self-contained electronic system. Such a system has many more interior nodes than exterior pads; its wide interior buses have speed and density demands not supported by the outside world. The backplane of Figure 1 provides the interconnection for these subsystem-chips using dedicated global wiring levels M4-M6. This creates inter-chip and intra-chip connections with equal facility, although the number of connections is thereby increased significantly. Alternative designs to the springs of Figure 1 can match the chip connection pitch to the desired bus pitch [7].

The fine pitch interior connections are in effect demountable vias connecting M3 and M4. They must, of course, be reliable. The MEMS technology for assembling the chips must be designed to accommodate these features and also allow scaling on a learning curve. Test insertion, as illustrated in Figure 3, demands easy and accurate assembly of the

subsystem-chips. Self alignment, therefore, is important in assuring that all the fine interior connections line up and make good contact during test.

Figure 1 shows a method of self alignment which is to guide the subsystem-chip to the correct location by use of an inclined plane on an alignment chip. Once within the tolerance range, the subsystem-chip falls over a vertical edge and into contact with the electrical connections. A certain amount of position freedom is designed in to accommodate thermal expansion and flexibility requirements, although these typically add up only to about a micron. The full manufacturing precision of ICs only applies to the front surface of the wafer and not to the saw-cut edges of the chips. To position the subsystem-chips or to guide them into place by these edges introduces an inaccuracy that may not be tolerable for the desired final assembly. A more comprehensive approach [7] uses a two-step course/fine method wherein the fine alignment is handled by IC technology.

The self-alignment chips are attached to the backplane chip and secured in place by snap-in action. This operation can be expected to have high yield because it is very simple and involves no electrical connections.

C. Docking Procedure

Subsystem-chip insertion as used here is a higher-accuracy extension of the conventional pick-and-place operation. To develop and maintain a robotics system that adjusts the needed degrees of freedom for each instance of subsystem-chip to backplane relationship would be very costly. Instead, the self-alignment chip guides the insertion for position, rotation and in-plane tilt by light contact to an inclined glide plane. The situation is similar to the docking maneuver between spacecraft. Knowledge of the original location of the subsystem-chip in the wafer can be preserved to make this simpler and cheaper. The speed of approach and the end of travel in seating the subsystem-chip have to be controlled to prevent chipping. The tooling used for placement should be universal, so that tool development can proceed from a common standard and evolve on a learning curve.

The insertion tool must have the proper degrees of freedom to allow the subsystem-chip to undergo the seating operation without releasing it prematurely. An example of a robotics pickup wand has been formulated with the necessary freedom [7]. It uses a low-mass, three-point suspension fixture for vacuum attachment to the back of the subsystem-chip.

D. Spring Contacts

Because membrane probe cards are not sufficiently flexible when several levels of metal are incorporated, front-surface springs are a preferred alternative. Batch fabrication by MEMS techniques provides individual spring force per contact. There are several degrees of freedom that must be accommodated: chip unevenness, thermal expansion differentials, overall system flexing, backing pressure for a heatsink, and test insertion. Ideally, all these requirements can be met with a single investment in spring fabrication sufficient to achieve a flexible flip-chip mating technology.

The largest amount of spring travel is in the vertical direction, especially if all subsystem-chips are backed up by a single flat heat sink. In that case, the error budget includes the

control tolerances of subsystem-chip and backplane surface flatness, heat sink flatness, subsystem-chip thickness after backgrind, the springs themselves, and the solderballs. The spring design illustrated in Figure 1 uses a long flat spring to allow this large travel at its midpoint.

When the spring and solder ball become fused together to make the final assembly, lateral forces come into play, although lesser amounts of travel are needed in the plane from thermal expansion and flexing. These forces, however, are very strong and not easily offset by stiffening the assembly with an underfill, as is a current practice in flip-chip packaging. With the scaled I/O's being considered, the approach here is to accommodate the travel rather than to resist these forces by means of many fine connections or compromise demountability. The spring of Figure 1 is supported by via connections that are free to tilt slightly. The two vias can tilt inward to allow the flat spring to flex. To accept lateral motion, the two springs tilt in the same direction. For test insertion there may be the need to provide some scrubbing action. Suitable design of the tilt forces of different springs in an array can provide this action [7].

E. Heat Sinking and Subsystem-chip Retention

Many future Integrated Systems will be operated near the thermal dissipation limit if device scaling density is pushed successfully. Even portable products, operated from a battery with a 4 W maximum [1], may encounter hot spots that need good heat distribution from a small volume. Since a universal solution is desired, a low-cost heat sink or heat spreader design as provided here could serve a wide market.

The need for separate retention of the subsystem-chips is brought about by the choice of flexible flip-chip mating technology. The self-alignment chips provide the appropriate location in the plane. All that is needed to complete the retention is a backing plate to keep the subsystem-chips in place. The integral heat sink performs this function, provided the expense of controlling the thickness and back surface of the subsystem-chips is justified. If not, for low power applications a compressible filler between a backing plate may be sufficient. A potting process would be even lower in cost.

F. Testing.

The key assertion of the test approach is that a single insertion into the final product electrical environment is a realistic objective. The first benefit is the saving of engineering effort normally required to assure a match of the test operation to that of the product in normal use, basically a wasted step. This becomes particularly important in mixed mode operation and when product volumes are low. A second benefit is simply the savings in the number of operations, including queuing and inventorying. A truly important benefit is the reduction in cycle time in giving feedback to the wafer fabrication line about what functions are not meeting specs. This is most urgent when higher-level functions such as at-speed and analog properties set the performance limits, and thereby the selling price. These are the same issues that are now lumped under the name Known-Good Die. They are more challenging for Integrated System technology than for conventional digital CMOS.

The cost of the individual test head, being composed of conventional product components, is low enough that many test heads can be used in parallel. The tester is simplified to

consist only of electronics that generates accurate waveforms, 'edges', which are sent to each test head. On the test head, there are Built-In Self Test (BIST) sensing circuits consisting of the same CMOS type as the subsystem-chips and made in the same facility. These determine DUT faults using on-chip comparators gated by the accurate edges from the tester. As an option, these circuits can be contained in universal Built-Out Self Test (BOST) chips that are usable in different products.

In a variation of this testing approach, the test heads are actual product units undergoing test. If subsystem-chip sizes are small enough to give high yield, subsystem-chips can be combined in the product with little or no advance testing, as long as there is a good probability that the major test functions can be activated. The tester functions of Figure 3, together with BIST, can then determine which are the defective chips. The latter are replaced and the product is ready for shipment. The BOST chips may also be extracted to lower cost further.

G. Permanent Attachment and Reliability

After test and burn-in, there will be a time when the chip set and backplane are all functioning according to spec. At that point the manufacturing process is completed by giving the product long-term reliability. Although gold-to-gold contacts can meet some reliability standards, permanent attachment is preferred. The spring design of Figure 1 allows solder fusing of the solder ball and the center of the spring, as discussed.

One further consideration is rework. Normally, when a component fails it is necessary to remove it so that the good components can be saved. With conventional flip-chip solder ball methodology this normally occurs when the chips go through test and burn-in with a temporary solder connection. For the composite IC, desoldering is only a field failure necessity.

V. Summary and Conclusions

The new opportunities created by low-cost digital CMOS and the need for interfacing it to real-world analog functions in Integrated Systems combine to shape major market forces that can underwrite significant new technologies. Current monolithic and/or MCM approaches do not adequately address the broad market potential.

What has been proposed in this study is a two-step approach to monolithic integration of system products. An entry-point technology is formulated that starts technology development in harmony with the market, rather than relying on untried significantly new technologies or perpetuating the long-running battle of integrating incompatible system technologies in a piecemeal way. This entry point is a combination of the best features of monolithic ICs and multi-chip modules. It is is suggested that the key differentiator of monolithic chips is their development on an aggressive learning curve, together with their design in a single design environment and batch fabrication by a single responsible team and not the fact that they are made on a single starting wafer.

The composite IC embodies this entry point by combining subsystem-chips the way MCMs do but starting at the silicon technology level to furnish the necessary manufacturing capabilities. These are chip attachment and electrical contacting through mechanical and electrical features fabricated with MEMS technology. Self-alignment and individual contact springs per contact are similar to the kind of batch fabricated

technologies that have earned silicon technology its reputation for performance and reliability and that can be improved on a learning curve for chip-to-chip connection. Thus, the composite IC is assembled from batch fabricated components that can be snapped together to form the product. These components are connected with a standardized interface that uses the fabrication accuracy of silicon fabrication to mix and match parts. The major difficulty with multi-chip system products, namely the known-good-die problem, is solved by this assembly approach.

REFERENCES

[1] *National Technology Roadmap for Semiconductors*, San Jose, CA.: Semiconductor Industry Association, 1994.

[2] Ref. [1], p. 63.

[3] T. Core, R.S. Payne, D. Quinn, S. Sherman and W.K. Tsang, "Integrated, Complete, Affordable Accelerometer for Airbag Applications," Proc. Sensor Expo, Chicago, IL, 1991, pp 204B-1,4.

[4] D. J. Bartelink, "A Unified Approach to Chip, Test, and Assembly Technologies for MCMs," *Proc. 1995 IEEE Multi-Chip Module Conf.*, Jan. 1995, pp 221-228.

[5] Ref. [4], Figure 1 and discussion on p 222.

[6] M.Y. Lau, K.l. Tai, R.C. Frye, M. Saito and D.D. Bacon, "A versatile, IC Process Compatible MCM-D for High Performance and Low Cost Applications," *Proc. ISHM/IEPS Int. Conf. on Multichip Modules*, Denver, CO, 1993, pp 107 - 112.

[7] D. J. Bartelink, unpublished

FIGURE CAPTIONS

Figure 1 Schematic cross-section view of Composite IC showing component part and assembly details.

Figure 2 Sources of subsystem-chips and backplane chip for Composite IC.

Figure 3 Test Station for Composite IC testing showing portion of farm of test heads and robotics arm for DUT insertion.

[1] *National Technology Roadmap for Semiconductors*, San Jose, CA.: Semiconductor Industry Association, 1994.

[2] Ref. [1], p. 63.

[3] T. Core, R.S. Payne, D. Quinn, S. Sherman and W.K. Tsang, "Integrated, Complete, Affordable Accelerometer for Airbag Applications," Proc. Sensor Expo, Chicago, IL, 1991, pp 204B-1,4.

[4] D. J. Bartelink, "A Unified Approach to Chip, Test, and Assembly Technologies for MCMs," *Proc. 1995 IEEE Multi-Chip Module Conf.*, Santa Cruz, CA, Jan. 1995, pp 221-228.

[5] Ref. [4], Figure 1 and discussion on p 222.

[6] M.Y. Lau, K.I. Tai, R.C. Frye, M. Saito and D.D. Bacon, "A versatile, IC Process Compatible MCM-D for High Performance and Low Cost Applications," *Proc. ISHM/IEPS Int. Conf. on Multichip Modules*, Denver, CO, 1993, pp 107 - 112.

[7] D. J. Bartelink, unpublished

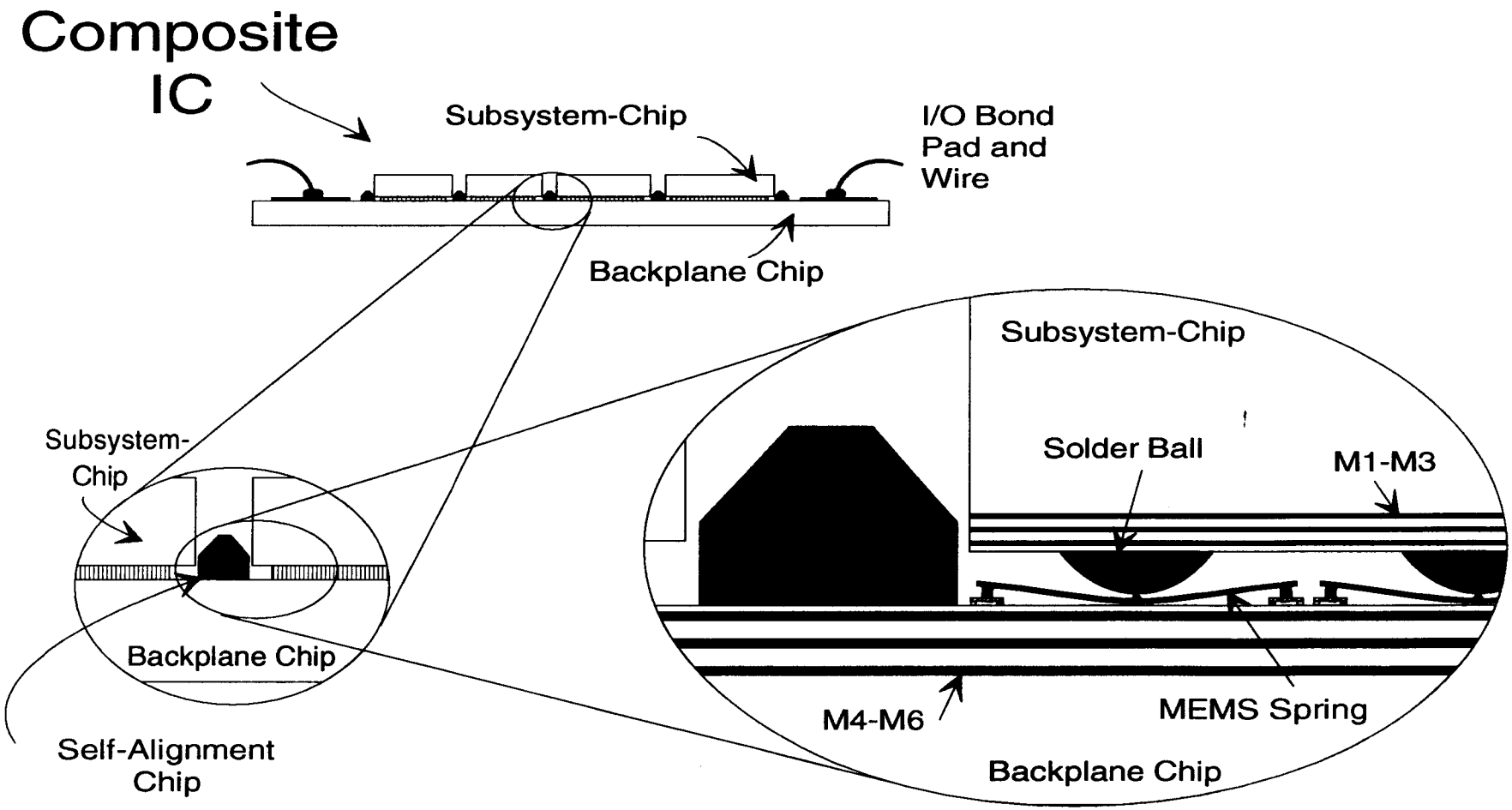

Composite IC
Subsystem-Chip
I/O Bond Pad and Wire
Backplane Chip
Subsystem-Chip
Subsystem-Chip
Solder Ball
M1-M3
Backplane Chip
M4-M6
MEMS Spring
Backplane Chip
Self-Alignment Chip

Figure 1

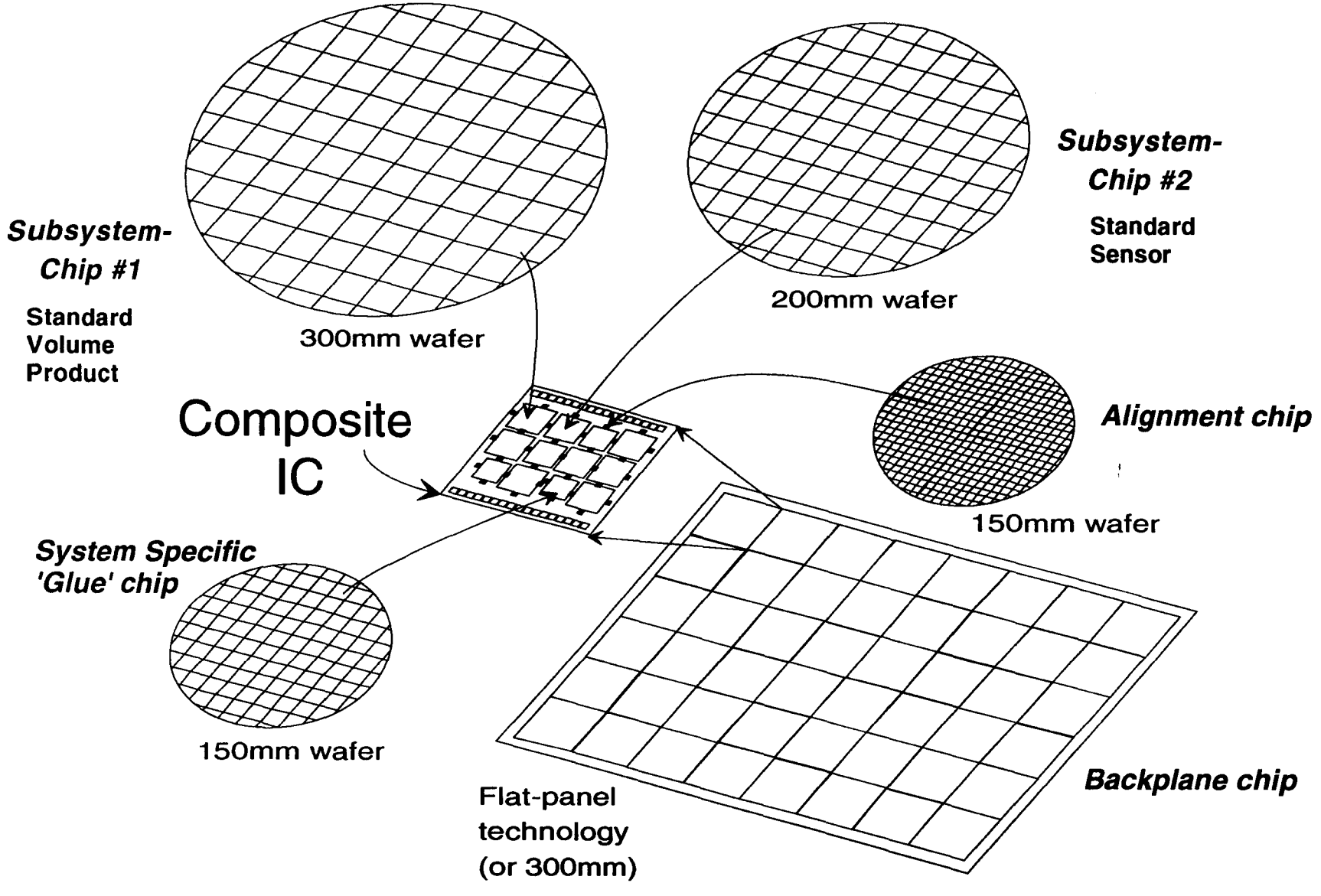

Subsystem-
Chip #1
Standard
Volume
Product
300mm wafer
200mm wafer
Subsystem-
Chip #2
Standard
Sensor
Composite
IC
Alignment chip
150mm wafer
System Specific
'Glue' chip
150mm wafer
Backplane chip
Flat-panel
technology
(or 300mm)

Figure 2

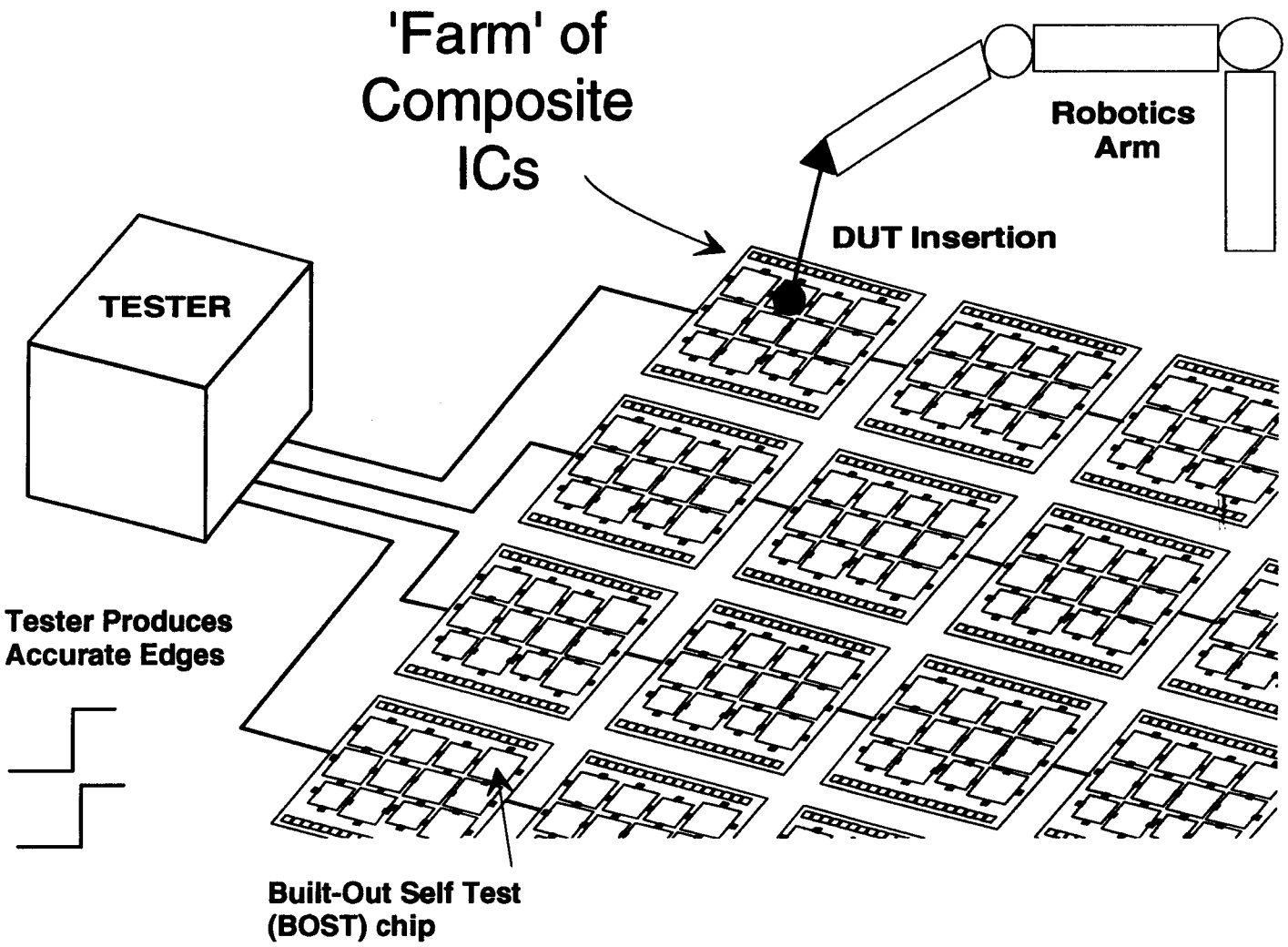

'Farm' of
Composite
ICs
Robotics
Arm
DUT Insertion
TESTER
Tester Produces
Accurate Edges
Built-Out Self Test
(BOST) chip

Figure 3