MCM PRODUCTION: TESTING AND RELATED ASPECTS

R. Mariani, S. Motto, S. Giovannetti, CAEN Microelettronica, Viareggio, Italy
email: r.mariani@caen.it

Abstract

A careful strategy must be planned to bring complex multi-chip systems, as multi-chip modules (MCMs), into production. The different production and testing strategies must be characterized in term of parameters like production yield, Know-Good Dies (KGD) quality and Early Failure Rate. In any case, MCM yield depends strongly on the goodness of its individual components, i.e. on the Fault Coverage (FC). In this paper the production strategy and the testability activity adopted for the FERMI microsystem is presented. It is also shown as the intensive use of Design-for-Testability (DfT) methodologies, both at the chip and at the system level, is fundamental to achieve high production yield of dense multi-chip modules.

1. INTRODUCTION

The quest for higher integration levels in systems and competitive pressures to reduce system manufacturing costs make the Multi Chip Modules (MCMs) always more appealing in today's applications. A MCM is an electronic system or subsystem with two or more bare integrated circuits (bare die) or Chip Sized Packages (CSP) placed very close to each other and electrically connected to a common substrate with very dense lines (up to 10-25 µm spacing) [1]. The substrate provides mechanical support and interconnections: it is composed of multilayer conductors separated by suitable insulating dielectric material and vias connect different layers; wiring densities cover up to 90% of substrate, while in conventional boards it rarely exceeds 10%. The metalization technology can be either thick film (additive stacking on ceramic substrates of dielectric and conductive layers; the metalization is formed by deposition, drying and firing) either thin film (subtractive method; the metalization is formed by sputtering and selective photoetching) having a better stability and noise characteristics [1].

Three basic technologies are available for microwiring substrates:

1) MCM_L (Laminated). Essentially an advanced form of PCB technology with copper conductors on laminated base dielectric. The MCM_L is not always the best solution for every application. Especially with respect to long term reliability and wide temperature ranges MCM_L technology has a smaller application range than MCM_C and MCM_D technologies. MCM_L are less expensive than MCM_C and MCM_D.

2) MCM_C (Ceramic). There are two different processes in MCM_C categories: several conductive layers deposited on a ceramic substrate and embedded in glass layers or several conductive and ceramic layers cofired at high (HTCC) or low (LTCC) temperature. The minimum line to line pitch is about 150 µm, the minimum line width about 75 µm, the conductor thickness about 8 µm and the dielectric thickness about 40 µm.

3) MCM_D (Deposited). The technology is also known as Thin Film Technology: the interconnections are realized by thin film deposition of metals on deposited dielectrics, which may be organic polymers or inorganic dielectrics. The minimum line to line pitch is about 50 µm, the minimum line width about 10 µm, the dielectric thickness ranges between 3-15 µm and the via diameter ranges between 10-50 µm. The main advantages of the MCM_D are the very high wiring density, the improved thermal management (up to 10 W) and a higher mechanical stability with reference to MCM_L and MCM_C; on the other hand the cost of MCM_D is certainly higher than MCM_L and MCM_C.

The use of MCMs leads to performance enhancement and cost benefits at system level. The reduction of the average interconnection length between components causes a reduction in the line impedance and the shorter chip-to-chip interconnections and bare die allows increased operational frequencies, a reduction of the power need to drive line capacitances, a higher signal to noise ratio, a lower cross-talk and an overall reduction of decoupling capacitors, resistors and drivers [IBM source]. The cost saving at system level is mainly allowed by the increased assembly process yield for MCM (1 PPM vs 100 PPM for SMT version per solder joint measured at test [IBM source]), a thermal management enhancement and a reduced board complexity. On the other hand the development of MCMs is much more complex than the development of a PCB; the design has to start with a detailed specification including function, environmental and mechanical specifications, partitioning of the electric functions into standard circuits or ASICs and especially with a test strategy that has to be taken into account since the early stage of the project [2]. To produce high-quality and cost-effective MCMs, test and fault diagnosis has to be included as critical requirements early in the design cycle; treating test as an afterthought in this process may result in high costs. But if incorporating test and fault diagnosis as critical design requirements is necessary to achieve high-quality, high-reliability and cost-effective multichip systems, it takes considerably study to evaluate where and when to test and to decide upon the best test method and level; in fact it is necessary to determine a trade off between cheap test solutions with inferior quality and high quality with effective and highly expensive test solutions.

2. MCM production flow and related aspects

In the figure 1 a complete production flow for MCM is shown: the component and substrate production is done in parallel (left and right branches) and then the MCM is assembled (central branch).
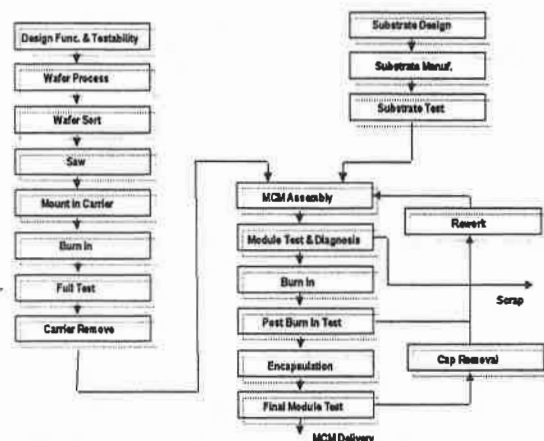

Fig. 1 : MCM production flow

Every step of the production flow has to be carefully planned to obtain high-quality and cost-effective MCMs, the cost and the resultant quality of an MCM depending mainly upon [3]:

– the yield of the chips;

– the number of chips in the module: a careful partition of the system should be planned at the very beginning of the project and the discussion should involve engineers expert in system design, circuit design, layout, manufacture, assembly, test and quality;

– the yield of the interconnection structure;

– the yield of the bonding and assembly processes;

– the effectiveness of the testing and rework process in detecting, isolating and repairing those defects: this aspect is also very closely related to the level of testability of the components assembled on the substrate.

If various chip types (1,2,...,n) are used within a module, then the first pass MCM yield can be expressed as it follows:

$$Y_{mcm} = (y_1)^A \cdot (y_2)^B \cdot ...(y_n)^N \cdot Y_S \cdot Y_I^Q \cdot Y_A$$

where: Y_{mcm}: first pass MCM yield; $y_{1,2,...,n}$: yield of chips 1,2,...,n (probability of chip 1,2,..,n being good); A,B,..,N: number of chips of each type 1,2,..,n respectively; Y_S: Known-Good probability of substrate; Y_I: Known-Good probability of die interconnects; Q: number of interconnects; Y_A: yield of bonding and assembly.

Chip yield plays a very important role: for example (figure 2) a 8-chip MCM with a 95% chips yield results in a 65% first pass MCM yield without considering any other source of yield loss. This means that 35% of the assembled MCMs must be diagnosed and repaired, a

costly and time-consuming task. The chip yield of bare chips must be pushed to nearly 100% to produce a module yield high enough to have cost-effective MCM process.
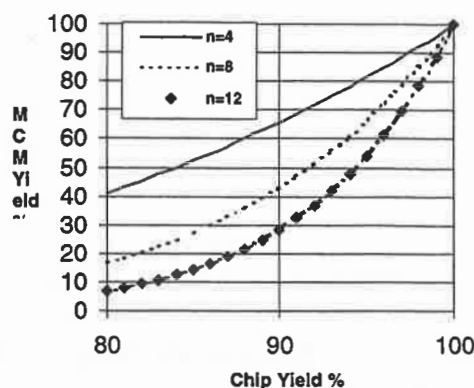

Fig. 2 : Chip Yield vs MCM Yield; n is chip number housed on the MCM

3. Chip yield: Fault Coverage (FC) and Design for Testability (DfT).

As underlined in the previous paragraph, high quality bare dies are needed to produce cost-effective MCMs; a high chip yield is achieved during wafer manufacturing, through process control-based approaches, and after manufacturing with bare die test and Burn-In to make the weak dies fail (infant mortality), increasing the confidence that each device is reliable and will continue to function for an extended period of time [1].

Provided a test pattern for the bare die test, the Fault Coverage (FC) is defined as the ratio between the faults detected by the given test pattern and the possible faults of the device under test and it plays a very important role in determining the quality of the bare die test therefore the chip yield; defining the Defect Level (DL) as the percentage of chips of the same type shipped which passed the test, but may be faulty (DL=1-y_n with the notations used also in the previous paragraph), the following formula relates the Defect Level DL to the process yield Y_p and to the Fault Coverage FC [4]:

$$DL = 1 - Y_P^{1-FC}$$

The Fault Coverage FC depends on the goodness of the generated test patterns and on the use of DfT (Design for Testability) structures implemented on the device under test: a chip yield of 99% (a DL of 1%) and a process yield Y_p of 80% results in a FC of 98% (figure 3).

An approach to alleviating the need of sophisticated testers at all levels of integration is to incorporate the tester into the circuit under test itself; hence the notion of DfT and Built-In-Self-Test (BIST). These approaches (often called On-chip ATE, Automatic Test Equipment) eliminate the need for expensive testers an provide a mechanism for accessing and exercising internal design circuitry [3]. In the following sections, these approaches are briefly described and applied to the design of the MCM-V3.
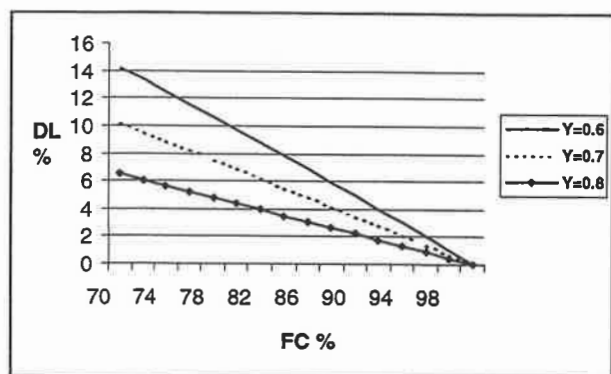
Fig. 3 : Defect Level vs Fault Coverage

4. Design for Testability in the design flow

As mentioned in the previous section, the yield of the components mounted on the MCM must be pushed to 100% in order to reach a cost effective MCM. Therefore, the standard design flow must be changed to insert some steps aimed to improve the component testability (what is usually called Design For Testability).

The figure 4 shows how the standard CAEN MicroElettronica design flow (left side) has been integrated with a typical testability flow (right side). Proper CAE tools (fault analyzers) implement the "Testability Analysis": a detailed map of the circuit in terms of Controllability and Observability (CO) values is computed, with different fault models (Single-Stuck At, Bridging Fault, IDDQ, etc...). Using these data and CAE tools (test synthesizers), scan logic (and Boundary Scan) can be inserted into the circuit to increase the CO values. Then Fault simulators and Automatic Test Pattern Generators (ATPG) are used to automatically generate the test patterns with the desired coverage. It is worth noticing that very often (especially with complex circuits), the tools are not able to reach high coverage; in these cases DFT rules must be used since the beginning of the design (at the HDL level) and "ad hoc" test structures must be inserted.

In very complex design using Deep-Sub-Micron (DSM) technologies, many parts of the chip are built using automatic tools (synthesizers) that often introduce redundancies and untestable nodes, therefore the use of testability CAE tools is more and more needed. On the other hand to use these tools in a smart and useful way means to educate design engineers on test related issues.

The use of TestGen (Synopsys), ex-Sunrise, one of the most powerful testability CAE tools currently available on the market, has allowed CAEN Microelettronica to considerably enhance the time-to-market and the quality of the shipped components.

Several methods exist to improve the circuit testability [5] and some of the most commonly used are listed below:

- Structured techniques, as Scan logic: A sequential circuit can be symbolized with a set of combinatorial circuits divided by flip-flops. This type of circuit is
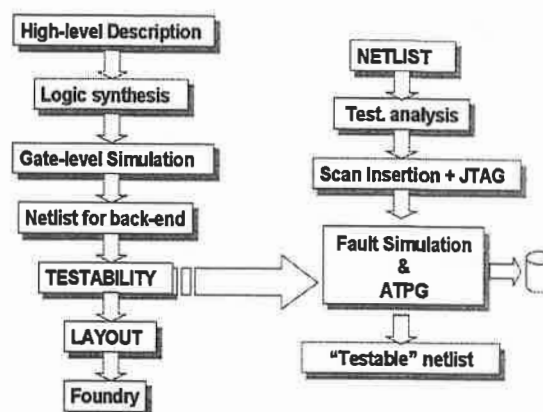

Fig. 4 : Testability design flow as adopted at CAEN Microelettronica

clearly difficult to be tested. For example it's difficult to insert a value in the last flip-flop or to observe a value of the first combinatorial circuit. With the insertion of the scan logic, it becomes easy for the ATPG program to insert the value in each flip-flop (FF) during an initial phase in which all the FF are serially loaded, and then is easy to read the result during a final phase in which all the FF are read by a serial shift. There are many kinds of scan techniques: Level Sensitive Scan Design (LSSD), Full Serial Scan, Partial Scan and Parallel Scan.

- Ad Hoc techniques, for example to reduce the number of Untestable faults. Untestable faults are faults that cannot be tested by the ATPG. This type of fault is found when a redundant part exists in the circuit. Another example of this fault is the case in which some nodes are tied to GND or VDD. They have to be removed redoing the design or using "ad hoc" testing structures: partitioning large sequential circuits, adding extra tests points, adding multiplexers.

- BIST (Built-In Self Test). Ad hoc blocks are inserted in the circuits to allow self-checking operations proving correct functionality, in general at-speed. There are many kinds of BIST: signature analysis and BILBO (Built-In Logic Block Observation), memory self-test.

- IDDQ testing, i.e. the monitoring of VDD supply current quiescent, mainly used to test bridging faults.

- Boundary Scan.

It is worth noticing that DfT and BIST are not free; they require an investment in chip area and in certain cases may themselves cause additional delays. These techniques will be briefly described in the following section applied to one of the MCM-V3 ASICs, and it will be shown as DfT structures can be employed without significant area and speed overhead if properly managed by the designer.

5. Application: MCM-V3

A MCM (MCM-V3) containing 14 bare dies flip-chipped on a MCM_D substrate (figure 5) was realized by FERMI collaboration [6]. Five complete channels are implemented: at the input of the MCM, non-linear data are applied to a linearising stage built around an adder for

offset correction and a multiplier for gain adjustment. The lineariser also includes a RAM that can be used as a Look-Up Table (LUT) and as a test pattern input for testing. The linearised data of the five channels are added and filtered in the level 1 filter ASIC, which will be described in the following paragraph. The lineariser data are stored in a pipeline during the level-1 trigger loop latency and is then, in case of a positive decision of the level-1 trigger, written into an event buffer. The readout filter (filter 2) contains three parallel Finite Impulse Response (FIR) filters and a non-linear order statistics operator; it can process one single channel at the time, and it offers a greater suppression of different artifacts, as timing jitter, in the input data stream, than a single FIR.
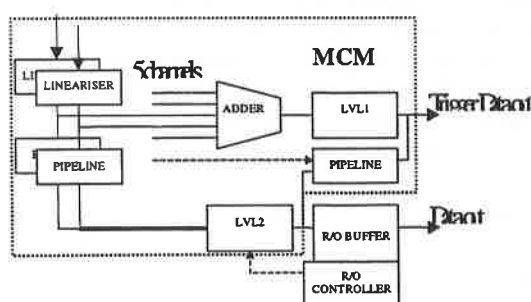

Fig. 5 : MCM-V3 structure

A well-detailed test strategy was planned before and during the design phase: a top-down strategy was adopted to specify the testability requirements at board level, MCM level and components level. The next section will introduce the activity and results carried out in the design of the LVL1 filter device, highlighting the concepts described in the sections 3 and 4.

6. Application: Testability of the LVL1

The Level 1 trigger filter chip (LVL1) [7] operates at the sample rate (40 MHz) on the sum of five channels and performs three different operations: identification of the event time, measurement of the pulse amplitude and identification and flagging of overlapping pulses. LVL1 consists of two parallel FIR filters: the *energy* filter shapes the input data to measure the amplitude, and the *timing* filter shapes the data to identify the time of the pulse event. A three-point maximum finder produces the pulse flag and two pile-up flags (near and far) with programmable overlapping distances. The filter can be fully programmed by a serial interface.

The filter has been synthesized from a VHDL description and realized by using the AMS 0.8μm CMOS technology. A standard IEEE 1149.1 Boundary Scan path is included, with the I/O scan register and the bypass register. The whole filter is composed by about 46000 equivalent gates with a silicon area of about 58 mm^2 and 159 I/O pads. A few first prototypes have been fabricated and successfully tested by using functional test vectors.

For this ASIC, seven main activities concerning testability have been carried out:

a) *Analysis of the faults in the circuit.* The fault analysis with Single-Stuck-at-Fault (SSF) model showed that about 2% of the circuit nodes were untestable, mainly because of some VDD or GND tied nodes at the input of a block used in different part of the circuit. Other untestable nodes were due to tri-state bus logic.

b) *FC measurement using the functional test vectors.* This analysis showed that the FC (with SSF model) was 74.34% using a test-pattern of 180K functional test-vectors.

c) *Test-vectors generation and FC measurement using the ATPG.* Running the ATPG tool without any scan insertion was not possible due to a very slow convergence of the algorithms implemented in the tool. The main problems were due to the circuit initialization mainly concerning the serial interface decoder. In fact, running the ATPG on the circuit excluding the decoder removed the convergency problem and a test-pattern was generated allowing a FC of about 81%.

d) *Insertion of DFT structures.* To solve the problems shown in b) and c) (i.e. low coverage and ATPG convergence problems due to complex structures not easily testable), a Full-Scan chain (a Mux-scan methodology with a single scan chain) was implemented. The Full-scan methodology was preferred to the Partial-scan because it is easier to implement and it leads to a better ATPG efficiency (defined as # of detected faults plus # of untestable divided by # total nodes); on the other hand the Full-Scan methodology introduces an higher area overhead if compared with the Partial-Scan methodologies, but the use of DSM technologies makes this aspect less demanding. A Full-scan chain with a single scan allows a minimum I/O overhead: only 5 additional I/O are needed. Another possible approach was the use of a multiple scan chain, shortening the test-time but with a considerably increased I/O number.

e) *Test-vectors generation and FC measurement using the ATPG after insertion of DFT structures.* The results obtained are shown in the following table:

Statistics	Comb	Seq	2-Pass
Fault coverage	94.87%	97.58%	96.98%
Testable FC	99.19%	99.46%	98.77%
Efficiency	99.23%	99.47%	98.79%
ATPG Vectors	358	23308	940
Scan Chains op.	358	456	132

In the table above are shown the fault coverage FC, the Testable FC (obtained eliminating the untestable faults from the total number of faults), the ATPG Efficiency, the number of parallel ATPG vectors (i.e. in each vector the scan elements IO pins are considered as primary IO) and the number of scan chains serial operations (each operation consists of a full shift of the scan chain). These figures were calculated using a pure combinational ATPG, a pure sequential ATPG and a two-pass ATPG, where a first very fast combinational ATPG step is followed by a sequential ATPG step. In the last case, 940 parallel ATPG vectors are needed to reach a 96.98% FC, with a 96% reduction of the ATPG vectors needed in case of pure sequential ATPG.

f) *Untestable nodes reduction and insertion of "ad hoc" structures.* The untestable faults due to VDD and GND tied nodes were eliminated adding some "injector" structures (multiplexers and extra pins), increasing the controllability; the untestable faults due to the tri-stated bus logic were removed using injector and output multiplexers, increasing the observability and controllability. With these modifications, the number of untestable nodes were reduced by a factor equal to 81%.

g) *Test-vectors generation and FC measurement using the ATPG after insertion of "ad hoc" structures".* The results obtained are shown in the following table:

fault coverage	98.51%
Testable FC	98.84%
Efficiency	98.85%
ATPG Vectors	1008
Scan Chains op.	145

It is important to notice that the untestable node reduction has leaded to a 1.5% enhancement of the FC on respect to the previous table.

h) *Area and speed overhead.* The results obtained are shown in the following table:

Area	+9%
IOs	+5
Speed	unaffected
Vectors	~200K

The area and I/O pads overhead does not significantly affect the design. Besides the delay on the critical paths was unaffected. Also the number of ATPG vectors needed to test the ASIC is relatively small, and it can be easily exercised by modern ATE. On the other hand the testability of the components was highly improved.

7. Future works

In the framework of the "Low cost Large Area Panel Processing of MCM-D substrates and packages" (LAP) ESPRIT Project [8], we are investigating a new partition for the system implementation on the MCM-V3, in order to reduce the number of components housed on the MCM and to obtain a higher general-purpose peculiarity of the MCM component leading to an increased opportunity to address different experiments. The main idea is to group together the Lineariser and the Pipeline in a single ASIC and to implement only the data acquisition (DAQ) path of the system. More investigation will be done for the ASIC FC, using IDDQ analysis (bridging fault), allowing further improvements for the chip yield. As concern the insertion of BIST structures, we are developing a modular solution for realizing RAM BIST, composed by a BIST controller to execute a test of the memory, a Test Pattern Generator and an Address Generator. We are going to implement the BIST controller as a programmable BIST processor, the Test Pattern Generator as a module generator, and the Address generator as an up_down counter. As concern the system level testability, different tests will be applied at each level (Die, MCM, Board, Crate and System) at different moments during the system life cycle: end of production, power-on, in-field and finally on-line. At each level of the hierarchy will be necessary to analyze the trade-off between constraints and goals and try to maximize the reuse.

8. Conclusions

In this paper we have introduced the MCM assembling technique and its advantages in term of system complexity and performance as compared to standard assembling techniques. However, we have also shown as the MCM approach needs a careful planning of the production flow in order to be cost effective. The topic role of chip yield in the production quality and therefore the need to push the component fault coverage to 100% has been discussed and it has been shown how the Design for Testability concepts must be included in the standard design flow at each level. Some results obtained in the case of the MCM-V3 design have been reported, and it has been shown how the use of testability CAE tools together with the design experience has allowed to reach high FC, without meaningful area overhead and performance degradation.

Acknowledgements

References

[1] Y. Zorian, "Multi-Chip Module Testing Techniques", *Proc. of International Advanced Course 'Test Technology for Digital and Mixed-Signal ASICs and MCMs'*, EUROPRACTICE Course Provider, 1997 Hannover, Germany.

[2] Authored by Members of the EUROPRACTICE MCM SERVICE, *Multichip Module Design Handbook*, Edition 1, EUROPRACTICE

[3] M.S.Abadir et al., Analyzing Multichip Module Testing Strategies, *IEEE Design & Test of Computers*, pp. 40-52, spring 1994.

[4] T.Williams, N.Brown, *IEEE Transactions on Computers*, 1981

[5] T. Williams, K.P.Parker, "Design for Testability – a survey", *Proceedings of IEEE*, vol. 7, no. 1, Jan. 1983, pp. 98-112.

[6] The FERMI Collaboration, "A Digital Readout System for High Resolution Calorimetry", *Proceeding of Third Workshop on Electronics for LHC Experiments*, CERN/LHCC/97-60, Sept. 1997, pp. 388-392.

[7] R. Mariani et al., "A Testable Fault-Tolerant VLSI Digital Filter for the FERMI Readout Microsystem", *Proc. of the 4° International On-line Testing Workshop*, IOLTW'98, July 1998, pp. 110-114.

[8] LAP Esprit WEB site: http://ife.ee.ethz.ch/mcm/lap