

The Challenges of Creating a Successful Back-End Manufacturing Operations Flow for Fabless Companies in the Era of Deep Sub-Micron Silicon Technologies

Dr. Michael Orshansky

Prepared by eSilicon Corporation, Copyright November, 2002

1.0 ABSTRACT

It is becoming very clear that outsourcing of the back-end manufacturing operations part of the IC development flow is economically beneficial to fabless semiconductor companies of all sizes. By outsourcing post-GDSII operations, fabless companies can concentrate on their core competency, which is chip design, by reallocating their costs from non-core competency skills to differentiating skills. Economic efficiencies arise due to such factors as:

- Aggregation of wafer demand by multiple outsourced operations and the resulting reduced cost of a single wafer;
- Full loading of the operations team and the resulting improved efficiency;
- Access to a highly skilled team, which is hard to assemble otherwise;
- Reduced time to production through the use of a well-developed flow.

In this White Paper, we discuss in much greater detail what the manufacturing operations part of the flow consists of. We will show that what is traditionally encompassed by manufacturing operations, in reality, is a high-risk activity comprising a set of technically sophisticated and complicated tasks. Because the cost of a mistake in turning a design into silicon is so high, the failure to take back-end manufacturing operations seriously may ruin the company by compromising an otherwise good chip design. Thus, the goal of this White Paper is to paint a realistic picture of the complexities of manufacturing operations in the era of deep sub-micron CMOS, and dispel certain simplistic notions about the back-end manufacturing flow.

2.0 CHALLENGES OF DEEP SUB-MICRON DESIGN AND MANUFACTURING

Many fabless chip companies treat the back-end part of the IC design and manufacturing flow as something simple. This happens because, for these companies, their differentiation is in the system chip behavior, and all the technical attention is paid to high-level, front-end design. An additional reason is that typically the leaders of these companies come from the design community and are not

appreciative of the difficulties of back-end manufacturing operations. As a result, back-end manufacturing operations are dealt with as something that can be fixed at the last minute. Such an approach often leads to catastrophic results for the company, when an otherwise working chip experiences significant delay due to the neglect of certain operational aspects. In the remainder of this White Paper we will show that the back-end flow is a set of complicated steps that requires a lot of attention and a high degree of expertise. The discussion is organized around the traditional services that an established ASIC provider would offer to its customers.

2.1 Design Services

There was a time when, in order to ensure the correctness of an IC layout, one had only to run a generic design rule checking (DRC) program. The layout was regular and predictable, and discovering the violations was a simple matter.

This is no longer the case, because in order to improve the manufacturability of deep sub-micron circuits, a number of techniques are used to modify the layout after all the "traditional" design steps are completed. These techniques are usually combined under the umbrella of resolution enhancement techniques (RET). These techniques include optical proximity correction (OPC), phase-shift masking (PSM), scattering bars, and off-axis illumination. RET techniques modify the layout so that the final silicon, rather than the layout, looks as intended. We no longer live in the world of "what you see is what you get". Under these conditions, in order to effectively perform deep sub-micron verification, and promptly react to the detected violations, a deep understanding of the complex resolution enhancement techniques and their impact on the mask-making process is required.

2.2 Process Technology

Circuit design in deep sub-micron CMOS technologies also requires a sophisticated analysis of the interaction between the design and manufacturing of a chip. Selecting the optimal technology is difficult and requires a careful comparison of the multiple design objectives. A multitude of factors

goes into consideration. Foundries offer many different versions of supposedly identical technologies: TSMC's 0.18um CMOS technology may be significantly different from that offered by UMC. But even within the same foundry, there are multiple choices of the same nominal technology. Even if similar performance characteristics, such as ring oscillator delay, are quoted, the differences between technologies may be very significant. For example, leakage current, and as a result, static power consumption may vary by orders of magnitude. The proper choice can only be made through the comprehensive analysis of the design, process, and economic considerations.

In deep sub-micron CMOS, an important consideration in choosing the process technology is its manufacturability. Increasingly, the biggest yield loss is not due to random yield (e.g. particles), determined by the quality of the clean room, but is due to the systematic yield loss. Systematic yield loss results from the interaction of the features of design with the characteristics of the process technology. Therefore, once again, the choice of the optimal process technology inherently depends on the interplay between many objectives; for example: performance, power, cost, and area can be determined only by careful analysis that takes into account a multitude of factors.

2.3 Packaging

The proper choice of a packaging solution is critical for achieving a good design. One of the factors that limits the maximum clock frequency achievable in ASIC designs is a lack of cost effective packaging solutions comparable in their performance to the packages used for the high-end custom designs. For example, high performance Intel processors - Mobile Celeron Processors - utilize the advanced socketable micro-FCPGA and surface mount micro-FCPGA packages. Because these packages have a very high thermal conductivity and are capable of effectively cooling the chip, they permit a significantly higher level of power dissipation than less expensive packages available to ASIC designers. The situation is further aggravated by the external cooling solutions available to high-end custom chips. These solutions prove to be too expensive for many ASICs, and further limit the choice of packaging solutions, and therefore maximum power levels. The capability of the board

manufacturer must also be taken into account when package choices are made. The package can affect PCB design and manufacturing costs.

The correct estimation of chip power dissipation is key. Because the choice of packaging solutions is constrained for ASICs, an inaccurate thermal and power analysis directly translates into the loss of performance. Despite the traditional practice and perception, thermal analysis must be a key part of the design process early on.

2.4 Design for Test and Test Engineering

The roadmap for semiconductors projects that in the very near future (if not now) the cost of testing a chip may exceed the cost of producing it. Testing the manufactured chip and ensuring its proper operation becomes a more challenging task than designing the chip in the first place. The first reason for this is the increase in the sheer size of chips to be tested, and the resulting inability of traditional testing approaches, such as scan, to scale up with increasing chip complexity. The second reason is the growing importance of novel deep sub-micron fault patterns; for example: internal transient errors due to delay faults, cross-talk, or ground bounce, as well as external disturbances, such as soft errors due to alpha particles.

Only a disciplined approach to testing that utilizes the latest testing methodologies can, in these circumstances, guarantee the proper operation of a complex chip. A full, conscious commitment to design-for-testability (DFT) is a must. DFT is a set of practices and techniques that inherently make a design testable, and allows achieving high fault coverage, which is a prerequisite of a high yielding chip. Advanced DFT techniques, such as memory and logic built-in self-test (BIST), and boundary scan, are required in addition to traditional ATPG.

A close interaction between the design team and the DFT group must be established early in the design process. DFT requires the designer to think about testing as an inherent part of the design process, rather than as an after-thought, a concept that is still foreign to many circuit design teams. Moreover, the complexity of new testing techniques is such that only highly experienced DFT experts

familiar with specialized test-insertion tools can perform the insertion of test circuitry. Design for testability does not come for free, since the test circuitry adds 10-25% of chip area in addition to the base design area. However, trying to save on DFT may turn out to be very costly. Tester selection is also an important consideration. The lowest cost production tester must be selected for the required pin count and speed in order to minimize production costs. Only experienced engineers with an intimate knowledge of the design can determine a proper compromise between cost and testing needs.

2.5 Product Engineering

One important trend brought about by deep sub-micron silicon technologies is the increased importance of design for manufacturing (DFM). The key aspect of design for manufacturing is ensuring that a high percentage of the manufactured chips will operate within the specifications in the presence of process, supply voltage, and temperature variations. Because manufacturing tolerances cannot be improved as fast as the nominal dimension is scaled, the relative magnitude and complexity of variation of many process parameters goes up. Just as an example, the gate length of the transistor, the driver behind Moore's Law, exhibits a rapidly growing variability between and within chips. This means that two transistors within the same chip that were supposed to behave identically may, in fact, exhibit quite different electrical and timing characteristics. Neglecting this important source of variation leads to a large parametric and circuit-limited yield loss.

The key to dealing with these complex issues is to ensure that the chip is operational under all possible physical and environmental conditions. This requires collection of the right characterization data. Process splits have to be performed to drive the devices to exhibit the limiting points of their electrical behavior. Process splits need to be able to uncover the complex patterns of variability that may be manifested in the real production cycle. On the basis of collected process data, worst-case corner simulations then can be run to ensure that all the critical circuit blocks operate properly. One also needs to address the supply voltage and temperature variation by validating the design under these corner conditions.

During production, one also has to constantly monitor manufacturing information in order to guarantee that all the reports from the foundry process floor are within the norm. The dynamic nature of modern IC manufacturing requires everyday attention. Step yields must be monitored to ensure that no particular step is drifting. In case a problem occurs, one has to alert the foundry so that process engineers can address the problem. The foundry publishes a multitude of yield-related reports and tracking them is a task that requires much experience: one has to read through the statistical noise and 'normal' fluctuations to discover trends that are truly worrisome. The amount of manufacturing data supplied by the foundry is so large that analysis can be performed only through the use of a sophisticated statistical database equipped with data-mining capabilities and powerful statistical analytics.

2.6 Quality and Reliability

A number of quality and reliability checks and tests have to be performed before a chip is certified. The physical mechanisms leading to reliability problems of integrated circuits are very complex. As the performance and sophistication of circuits being designed and manufactured increases, the sensitivity of chips to all sorts of non-intended uses increases. A series of complex engineering experiments is required to ensure proper operation.

To ensure the reliability of the device in the first several months of its operation a burn-in test is applied. ESD (electrostatic discharge) testing has to be performed very carefully since chip failure due to electrostatic discharge is a big danger. For deep sub-micron chips the amount of electrostatic discharge that can kill the circuit is minute. Testing can be performed either under the human body model (HBM) or a charged device model (CDM). Latch-up testing is performed to eliminate the possibility of latch-up in the CMOS device. Highly accelerated biased humidity stress test (HAST) is

designed to detect the failure mechanisms related to die-to-package stress interactions, die passivation integrity, pad corrosion, ionic contamination (static stress), etc. A significant mechanical stress may build up in the chip due to the mismatch between the thermal expansion coefficients of

different materials. The temperature cycle test ensures the mechanical integrity of the chip by exposing it to alternating temperature extremes. Finally, life test is designed to rule out the susceptibility of the chip to latent process mechanisms, such as the possibility of time-dependent electro-migration and gate oxide breakdown.

2.7 Supply Chain

In order to establish a winning back-end manufacturing flow, a fabless company also has to successfully implement a host of tasks that are commonly combined under the term "supply chain management". In doing that, a new player in the fabless sector faces both business and technical challenges. Among the business challenges is a need to rapidly establish a number of mission-critical relations with the various partners and suppliers with the goal of guaranteeing a flawless execution to the customer schedule. The key business links include the foundry, the testing house, the package and assembly house, as well as the host of suppliers of materials for prototyping (e.g. probe cards, load boards, sockets, etc.)

To begin with, the company must secure a commitment from the foundry to manufacture silicon wafers, which, for a small company, may be quite taxing. New or small players in the industry will have to settle for longer turn-around times, higher wafer prices, and longer queues. This is especially true during the time of close-to full factory utilization. During such times, most foundries prefer to deal with larger players, significantly limiting the space of opportunities for smaller companies. Also, the small fabless companies have a harder time in getting detailed technical assistance from the foundry, for example in identifying yield problems, and implementing yield-enhancing process changes. At the time when chip yields are greatly influenced by the interplay between the design and process, the quality and ease of the technical interaction between the foundry and the design team is crucial to the final success of the product.

At the same time, the fabless company is faced with the task of implementing a set of highly specialized engineering procedures for the technical interactions with the foundry. For example, the

company has to ensure the qualified technical management of all the engineering lots being processed in the foundry, which requires a close familiarity with the transactions and nomenclature employed by the foundry, as well as day-to-day interactions with the foundry for each step in the prototype flow. Many other tasks, with a large number of logistical links, have to be handled in parallel, like ordering and properly managing the substrate and other materials, supporting sample requirements, and performing supplier and backlog management.

Finally, a fabless company should not underestimate the logistical difficulty of independently establishing the entire supply chain operation from scratch. Given the current trend towards the globalization and fragmentation of the semiconductor industry, a large number of worldwide, distributed suppliers have to be perfectly coordinated. All the legal, linguistic, and cultural specifics need to be considered to ensure a smooth synchronization of parties working in different time zones and different continents.

3.0 CONCLUSION

In this White Paper, we described in detail a series of technical and logistical steps that have to be established in order to implement a complete back-end manufacturing operations flow. The back-end flow is a high-risk activity comprising a set of technically sophisticated and complicated tasks. A high degree of skill and expertise is required of a team to handle these tasks properly and effectively. The price of a mistake in turning a design into silicon is high: the failure to take the back-end flow seriously may ruin the company by compromising an otherwise good chip.

Author:

Michael Orshansky is a Research Scientist at the University of California, Berkeley where he works on the development of techniques and algorithms for robust circuit design. His research interests include circuit design and analysis techniques for manufacturability of high-performance digital

circuit, statistical CAD algorithms for design-for-manufacturing and yield improvement, modeling and simulation of semiconductor devices. In 2000, Dr. Orshansky was an original member of eSilicon. He was a Fellow of Semiconductor Research Corporation and Advanced Micro Devices. He has published numerous technical papers in the area of statistical CAD algorithms, device physics, and modeling.