Circuit and Platform Design Challenges in Technologies beyond 90nm

Bill Grundmann, Rajesh Galivanche, Sandip Kundu
Design Technology, Intel Corporation
Contact: *Sandip.Kundu@intel.com*

Abstract

There are already a huge number of problems for silicon designers and it is likely to just get worse. Many of these problems are technical associated with shrinking geometries and increasing architecture complexities, but there are a significant number that seem to be caused by procedurally related mistakes and issues. Many of the technical problems are solved and re-solved on a piece-meal basis, focusing on local optimizations of small design-space problems. Unfortunately, many of these local solutions really create a less apparent but larger inefficiency in the whole design flow. The reason for this is that a few ever look at the whole design methodology, especially as it applies to large design teams. As a consequence, this lack of oversight for the whole methodology is causing project procedural problems and inefficiencies.

Introduction

In 1965, Dr. Gordon Moore predicted a doubling of transistors every year, but more importantly, he saw the shrinking size of the transistors as a means to making IC's that would become cheaper, more powerful and more plentiful. Someone took Dr. Moore's observations and called it Moore's Law, and hence this naming was established in the industry's terminology history. In 1975, he approved an adjustment to a doubling of transistors every 18 to 24 months, which for many reasons, has held quite true even to today (Figure 1). Since Moore's Law has many side benefits, it has even been extended, by others, to cover other business and commercial metrics beyond the domain of transistors. For example, it has been associated to the growth in microprocessor performance and clocking frequency expectations as well.

The impact of Moore's Law has been profound. The semiconductor industry uses Moore's Law as a reference to predict where the technology should be for a given year (Figure 2). To get to these smaller transistors and corresponding smaller interconnects on a recurring and compounding basis requires that many technology hurdles have to be overcome.

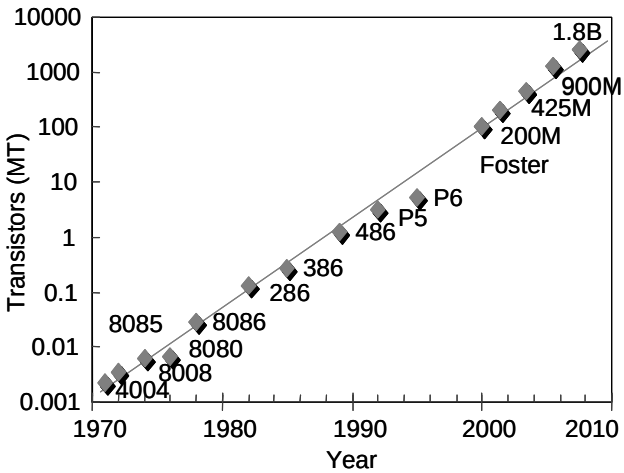

Figure 1 - Growth in number of transistors

There have been several times when experts declared that the industry would never be able to solve the technology barriers to maintain Moore's Law; one such example is when we reached 1µm. So far, these doomsayers have been proven wrong. Technology leaders are still predicting that Moore's Law can be extended through at least the next decade, and some others say even longer.

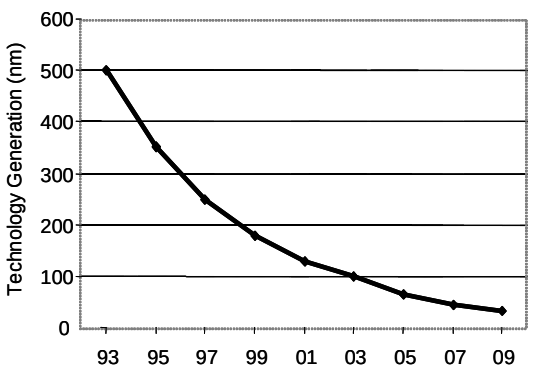

Figure 2 - Technology Generations

1530-1591/03 $17.00 © 2003 IEEE

Unfortunately, following any power-curve like Moore's Law doesn't come without a significant cost. Besides the huge manufacturing costs, there are the costs of creating the designs that take advantage of the abundant transistors.

Every design generation has had its new issues and every generation creatively came up with usable solutions to those problems. For example, the early design teams worked only with pencil and paper to sketch their designs and manually created their manufacturing die masks using plastic sheets and pen-knifes. The next generation identified the need for basic tools to ease the design capture problem, which resulted in the creation of the first CAD tools. As time went on, problems like logic simulation, transistor equivalence, parasitic extraction, timing analysis, place & route, synthesis were all tackled and adequate solutions became available.

Until the mid-80's, everything was mostly focused on logic-centric implementation. What happened in the mid-80s was the sudden realization, by most, that interconnect was just an electrical element and it wasn't really immune to any of the laws of physics. They discovered interconnect was becoming worse in many ways and in many cases dominating how designs were created. Unfortunately, this realization hasn't caused the industry to rethink many of the previous logic-centric solutions, and we may be paying the price.

Different design limits

Technology scaling seemingly continues to follow Moore's Law, doubling the number of usable transistors every architectural generation. Along with this growing number of transistors is the seemingly insatiable market appetite for relatively higher product performance, which for many implementations requires the use of higher clock frequencies and fewer logic gates-per-cycle (Figure 3). The diminishing returns of either cranking up the clock frequency or reducing gates-per-cycle are become more visible by the existence of huge design teams and increased project schedules. Reduction of gates-per-cycle is also asymptotically limited to basically 1 gate + 1 latch/flip-flop per cycle, which has huge power dissipation, hard clock distribution and severe die-size area penalties. Also, traditional test methods that are based on full scan design-for-test become unaffordable due to the high latch/gate ratios.

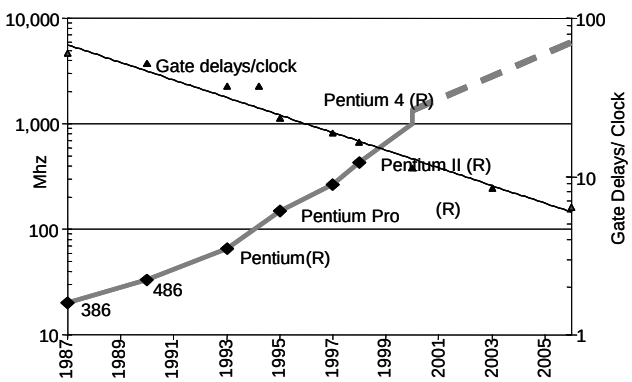

Figure 3 - Gate delays per cycle

There are many technical problems that are being solved, which probably means there are many problems that we haven't even thought about. Some of the more recent technical problems of interest involve power dissipation as a frequency limiter, leakage as a low-power inhibitor and reduction of design margin as number of gates-per-cycle continues to decrease.

The power dissipation problem has many interesting aspects. There are already products that throttle clock frequency when they detect over-temperature situations. Currently, in some cases, designers aren't even able to use the fastest parts as they exceed designed reliability limits (Figure 4). Because of this limit, these designers are starting to be creative with architectural and circuit tricks to conditionally move the power dissipation from one die area to another over time.

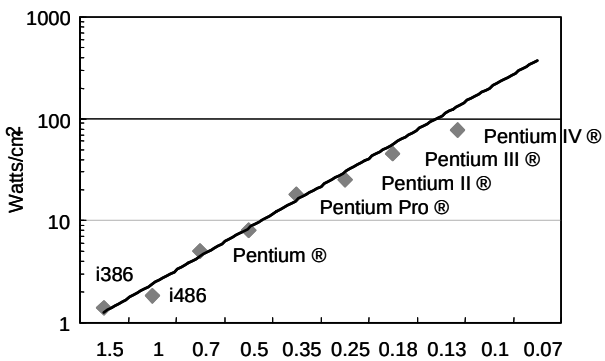

Figure 4 - Power Density

Power dissipation for most designs was just the summation of all of the dynamic CV^2F components with all other power dissipation components mostly ignored. A consequence of technology scaling seemingly is the steady, but steep, increase in various leakages: Source-drain, gate, diode, etc. Except for those worrying about quiescent current, this has been mostly ignored in the past. Unfortunately, these leakages are starting to equal many of the dynamic power sources. Calculating dynamic power has mostly been solved by counting signal transitions and knowing the signal's capacitance. Calculating leakage power requires the additional collection of the signal's state probability, or how many cycles it was high versus low.

Design margin is also being changed for the worse (Figure 5). With so few gates-per-cycle, the designers are finding it harder to get a timing convergence. There are many reasons for lack of timing convergence. One of the primary reasons relate to the reduction of the amount of design margin per gate. For instance, fixing a critical path that misses by one gate's delay is much easier in a

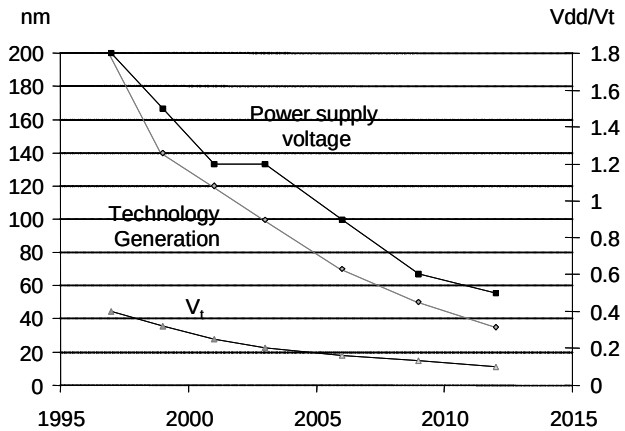

Figure 5 - Power Supply and Vt

20 gate path (5% problem and is often ignored) than it is in a 6 gate path (17%). Likewise, the manufacturing process variation can cause the same one gate's delay variation making the 20 gate design easier to hit the product's frequency goals.

Another problem with convergence is the traditional split between front-end and back-end design. No magic can be found that will guarantee back-end convergences from a poorly designed specification. You just can't build an aggressive product from a poorly written RTL. Designers have already been modifying their RTL to make design more build-able, but this is coming at an increasing cost to the design project. As the number of gates-per-cycle decrease, then more often an RTL has to be changed for retiming, restructuring, special circuitry and to meet the logic-equivalence requirements. The more the RTL changes, the less stable is the logic validation effort. Here's a good example of local optimization (RTL matching Netlist) hurting the success of the global project. Validation has to finish at least the same time as silicon implementation, preferably before implementation is done.

New kind of test problems

With the advent of designs that contain environmental sensors, we also are seeing the dawn of significantly new test problems. These on-silicon environmental sensors already observe temperature, frequency and voltage problems, allowing additional circuitry to somehow compensate some behavior of the die in hopefully a controlled fashion. Temperature sensors detect when the die is too hot or cool and respond by changing a combination of voltage, clock frequency and maybe even die functionally. Voltage sensors detect over or under-voltage situations that require change in clock frequency. Clock frequency detectors sense the out-of-range frequencies that a customer might impose on a die via over-clocking that may cause stress-related reliability problems. There may be a significant growth in the types of the sensors and, of course, designers will try their best to use them. Sensors to automatically detect critical-path requirements and cause an appropriate response in frequency or voltage probably aren't that far away.

Current designs have these sensors scattered across a die and generally group them together to perform a global optimization affecting the whole die. A typical sensor application might be to adjust to a lower clocking frequency on an over-temperature situation, keeping the part from getting over-heated and burnt. Voltage sensors will be used to compensate frequency during the first VCC droop problem and to compensate for abnormal voltage drops in critical areas of the design.

We should expect, in the future, that these sensors will start controlling more local die regions instead of controlling a single global aspect across the whole die. This will require more asynchronous design behaviors between each of the die regions. As we get more asynchronous behaviors we also get more uncertainty in test behavior. In addition, these sensors themselves have to be characterized, calibrated and tested to insure they perform correctly for their design purpose.

With all of these synchronous design problems, asynchronous design techniques are starting to look better than before, especially as a means to increase the effective clock frequency or product performance. As these large-scale asynchronous-like designs start to appear, so will the potential need for more asynchronous test technology versus today's dominant synchronous test technology.

Designers getting back to design

With the interconnect delay problems, low gates-per-cycle requirements, reduced design margins, high clock frequencies and power limited designs, designers need a break. They need a new overall design methodology.

A typical design methodology flow uses more CAD tools than its predecessors. With more tools, comes the need for more experts in understanding the individual tools. The ideal scenario is for every designer to be an expert, but the large number of tools and the difficultly of the

design itself makes it impossible for everyone to become an expert. So, what happens is that the tools are tailored to suit traditional methods that may be fine for the nominal design. The designer is getting more handicapped as they are not the tool experts and the design team structure usually requires the use of the tool per some design-flow. The designer using these scripted design-flow tools is in many cases just really be playing a game, like a video game, and trying different scripts and comparing the different results. This trial-and-error approach without really understanding how the tool does things has a problematic success. If none of the flows succeed, then most often these designers aren't left with enough information for why the tool failed. With the lack of this information, some designers perform a minor change to the design, hoping they are changing the right thing and repeat the whole process again.

Many of today's designers are hindered by this design-flow mentality that was established over the last 20 years. Although, these design-flows have helped the less experienced to be able to yield something, it really is a question of how efficient they are at what needs to be done. For instance, is it better for a designer to exhaustively work for 6 months on a piece of full-custom design and still not have a satisfactory solution? How about if they try all of the 100's of synthesis scripts and after 8 weeks none of them succeed? What would it take, for the designer to take a few days to assess the difficultly of design as just too hard or likely to be a design creation challenge? When a design challenge is found, what mechanisms are at the designer's disposal for solving the problem? For instance, how does that designer add additional architectural latency, which makes for an easier design, and at the same time quickly assess the impact to product performance or other product quality metrics? Most projects don't allow the implementation engineers to question the wisdom of the chip's architect. However, did the architect really understand and appreciate the impact or cost to design on his architectural decision? Small architectural performance improvements cause a significant amount of designer's time, and this might not be the best return-on-investment.

Summary and Conclusions

The ever decreasing circuit geometries are enabling us to pack more transistors and consequently more functionality in a given silicon area but also are bringing in additional tough challenges to overcome. Below is a list of most difficult ones that need to be addressed if we are to be successful in future silicon designs.

1. Leakage power is a significant portion of the total power and must be overcome to make the manufacturing of silicon design viable
2. Significantly increased susceptibility to design marginality some of which may have to be tested during manufacturing test flows
3. Increased use of adaptive design techniques with on-chip sensors making testing and validation more difficult
4. Need for design methods and tools that allow the designers to make global optimizations to get better return-on-investment of silicon area and engineering efforts.

References

[1] F. Pollack, "New Microarchitecture Challenges in the Coming Generations of CMOS Process Technologies," Micro32, 1999.

[2] G. Sery et al., "Life is CMOS: why chase the life after?," DAC 2002, 78-83.

[3] T. Karnik et al., "Sub-90nm Technologies - Challenges and Opportunities for CAD," ICCAD 2002

[4] S. Kundu et al., "Test Challenges in Nanometer Technologies," Journal of Electronic Testing: Theory and Applications, pp. 209-218, 2001

[5] S. Sengupta et al., "Defect-Based Test: A Key Enabler for Successful Migration to Structural Test", Intel Technical Journal, http://developer.intel.com/technology/itj/, 1999 1st quarterly issue