Benchmark Standards for ASIC Technology Evaluation

Two benchmark circuits are used for objectively evaluating ASIC supplier performance claims. The method applies first-order equations relating capacitive discharge currents and transistor saturation current to arrive at a technology constant. The method has been used to survey 14 ASIC suppliers with over 76 different technologies. Results are shown for 48 CMOS technologies.

by Antonio A. Martinez, Aloke S. Bhandia, and Henry H.W. Lie

The issue of determining valid performance for ASIC suppliers is always paramount in the minds of designers. Choosing a supplier based on aggressive performance claims could lead to disastrous results. On the other hand, choosing a supplier with conservative performance claims results in higher costs than necessary because of suboptimal performance and area utilization.

One way to compare ASIC supplier technology performance is through benchmark circuits. The traditional simplistic 2-input NAND gate intrinsic delay or even fanout delay has given way to more complete benchmark circuits. But which benchmark circuit covers the range of design possibilities? Should it scale with technology parameters such as transistor drive, metal wire length, or capacitance? What about interconnect metal resistance? What are appropriate interfaces to the real world? TTL or CMOS? Beyond design considerations, what conditions were used to generate performance and delay numbers? What voltage, temperature, input slew time, and process conditions were used by the ASIC supplier? Our investigation addresses these issues.

After contacting several HP divisions, we ended up using two benchmark circuits that have been around in various forms since 1987 and were most recently published in 1993.[1] We surveyed ASIC technologies from various suppliers to compare benchmark circuit performance claims. Since the benchmark simulations give too much latitude to ASIC suppliers, we used specific device technology details to compare and evaluate the accuracy of supplier claims. We observed both aggressive and conservative performance claims given the underlying technology. The process we describe also allows HP designers to focus on areas where a supplier has suboptimal designs.

Description of the Benchmarks

The benchmark circuits are shown in Fig. 1. The first circuit, Benchmark 1 (Fig. 1a), shows a complete path from input pad to output pad. There is a substantial output pad load of 50 pF which more closely represents real-world applications. An internal path contains various typical logic gates with fanout and wire length specifications. The second circuit, Benchmark 2 (Fig. 1b), includes two D flip-flops and a 2-input NOR gate. Benchmark 2 does not include interfaces to the real world and is relatively simple in comparison to Benchmark 1. To a larger extent, Benchmark 2 gives a closer feeling for the intrinsic performance of a given technology. For both benchmarks, wire length and gate fanout capacitance scale with technology. For instance, if metal pitches become smaller when moving to the leading-edge technology, wire lengths should be shorter. Likewise, gate fanout capacitance should track when moving to smaller, stronger transistors in a leading-edge technology.

We obtained complete path delays for 48 CMOS technologies for 14 different ASIC suppliers, including internal and I/O path rising and falling delays for Benchmark 1 and maximum operating frequency for Benchmark 2. Generally, suppliers used models of their technology to estimate path delays rather than measurements of actual circuits. The data collected is shown in Table I.

Table I
ASIC Supplier Sample Data

Drawn transistor gate length (μm)
Effective transistor gate length (μm)
Transistor gate oxide thickness (Å)
Pitch for each metal layer (μm)
Power supply voltage used for benchmarks (V)
Junction temperature used for benchmarks (°C)
Process condition used for benchmarks (SLOW/NOM/FAST)
Input edge rate used for benchmarks (ns)
Benchmark 1 complete path T_{pLH} delay (ns)
Benchmark 1 complete path T_{pHL} delay (ns)
Benchmark 1 internal path T_{pLH} delay (ns)
Benchmark 1 internal path T_{pHL} delay (ns)
Benchmark 2 maximum operating frequency (MHz)

Suppliers were required to provide enough process details to evaluate the accuracy of performance claims. We wanted to verify performance using first-order figures of merit and comparisons between suppliers. To that end, we considered process and device parameters that strongly affect circuit performance, namely transistor effective gate length (L_{eff}),

transistor oxide thickness (T_{ox}), junction temperature, power supply voltage, and pitch for all metal layers. These parameters are shown in Table I.

Evaluating Performance Claims

We evaluate supplier performance claims by relating delay or frequency numbers to manufacturing process specifics. Roughly speaking, propagation delay t_d is proportional to L_{eff}, T_{ox}, and temperature, and inversely proportional to V_{DD}, as shown in equation 3 below. We arrived at equation 3, a first-order constant relationship, by equating the rate of capacitive discharge current (equation 1) to transistor saturation current I_{DSAT} (equation 2). Equation 1 models the capacitive discharge of a node in terms of the power supply voltage, current, and rate of discharge (roughly proportional to delay). Equation 2 relates transistor saturation current to transistor effective gate length L_{eff}, oxide thickness T_{ox} (inversely proportional to oxide capacitance $C_{ox)}$, mobility u_o (related to temperature in Kelvin), threshold voltage V_{to}, transistor width W_{eff}, and power supply voltage V_{DD}. For simplicity, V_{to} is assumed to be proportional to $V_{DD,}$ while temperature is inversely proportional to u_o.

$$I = C\frac{dV}{dt} \tag{1}$$

$$I_{DSAT} = \frac{u_o C_{ox} W_{eff}}{2\ L_{eff}}(V_{DD} - V_{to})^2 \tag{2}$$

$$\text{Technology Constant} \approx \frac{t_d V_{DD}}{L_{eff} T_{ox} \text{Temp}}. \tag{3}$$

Equation 3 is not very precise, and is valid only over narrow regions of operation. However, detailed SPICE simulations showed that there is a surprising amount of linearity, as shown in Fig. 2. The largest error was 6.4% for L_{eff} variations from 0.4 µm to 0.9 µm. The largest percent errors are shown in Table II. Simulations were done with HP SPICE for a circuit deck equivalent to Benchmark 2 using 0.5-µm CMOS SPICE models.

Table II
Maximum Percent Error for Linear Fit of Delay as a Function of Various Variables

Variable	Range	Maximum Error
L_{eff}	0.4 µm to 0.9 µm	–6.4%
T_{ox}	57 nm to 228 nm	–4.0%
Temperature	0°C to 135°C	–1.2%
V_{DD}	2.7V to 5.7V	5.8%

Supplier Performance Comparisons

We first charted all performance numbers for various suppliers as shown in Fig. 3. For each supplier, L_{eff} decreases as we move from left to right. Suppliers are generally quite competitive with their leading-edge technology, but some suppliers' leading-edge technologies appear slower than another supplier's mature technology. Some suppliers appear to do quite well in the total path delay for Benchmark 1 but do not perform as well in the maximum operating frequency for Benchmark 2.

We applied equation 3 to all delay and frequency numbers, computing a *technology constant* value for each supplier's technology performance claim. However, since the technology constant is only a constant under an idealized first-order approximation, we looked for patterns and trends. We computed the mean and standard deviation for all suppliers across all technologies for each Benchmark 1 and 2 constant. After plotting technology constants for each benchmark for all suppliers and their technologies, we placed guardbands one standard deviation from the mean as shown in Fig. 4. Technology constants falling above the upper guardband were labeled conservative and technology constants falling below the lower guardband were labeled

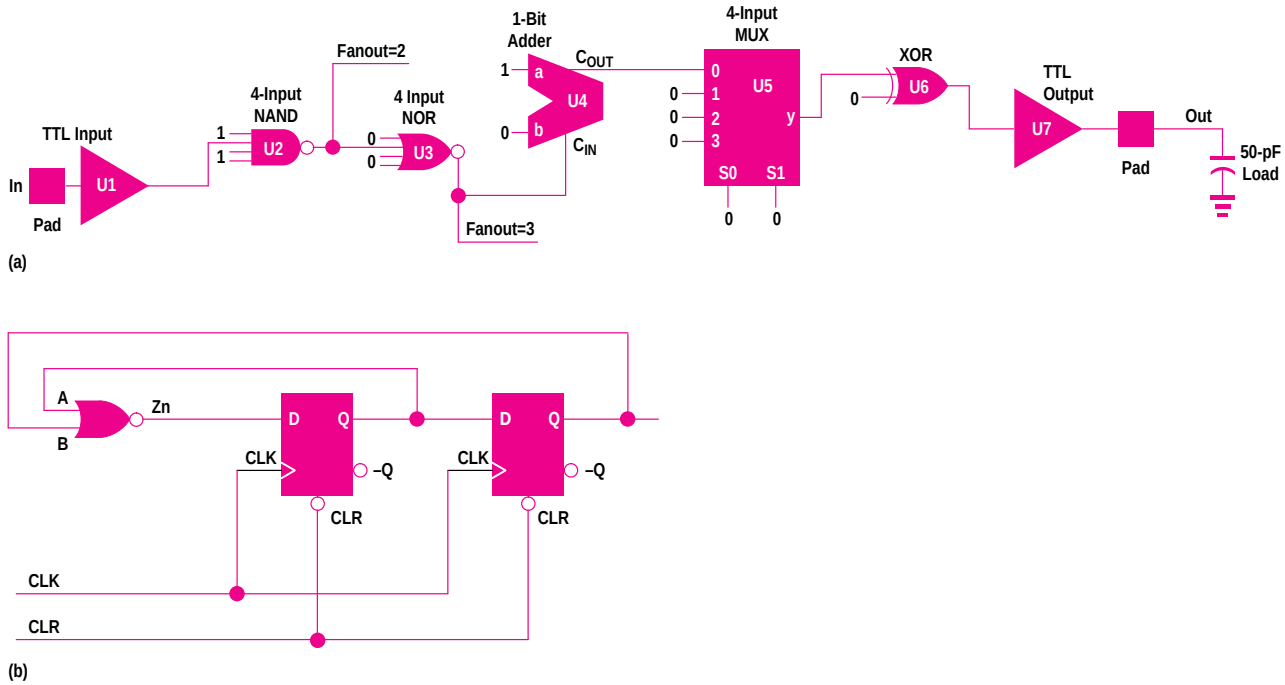

Fig. 1. (a) Benchmark 1 includes a complete path of internal and I/O delays. (b) Benchmark 2 includes D flip-flops with reset.

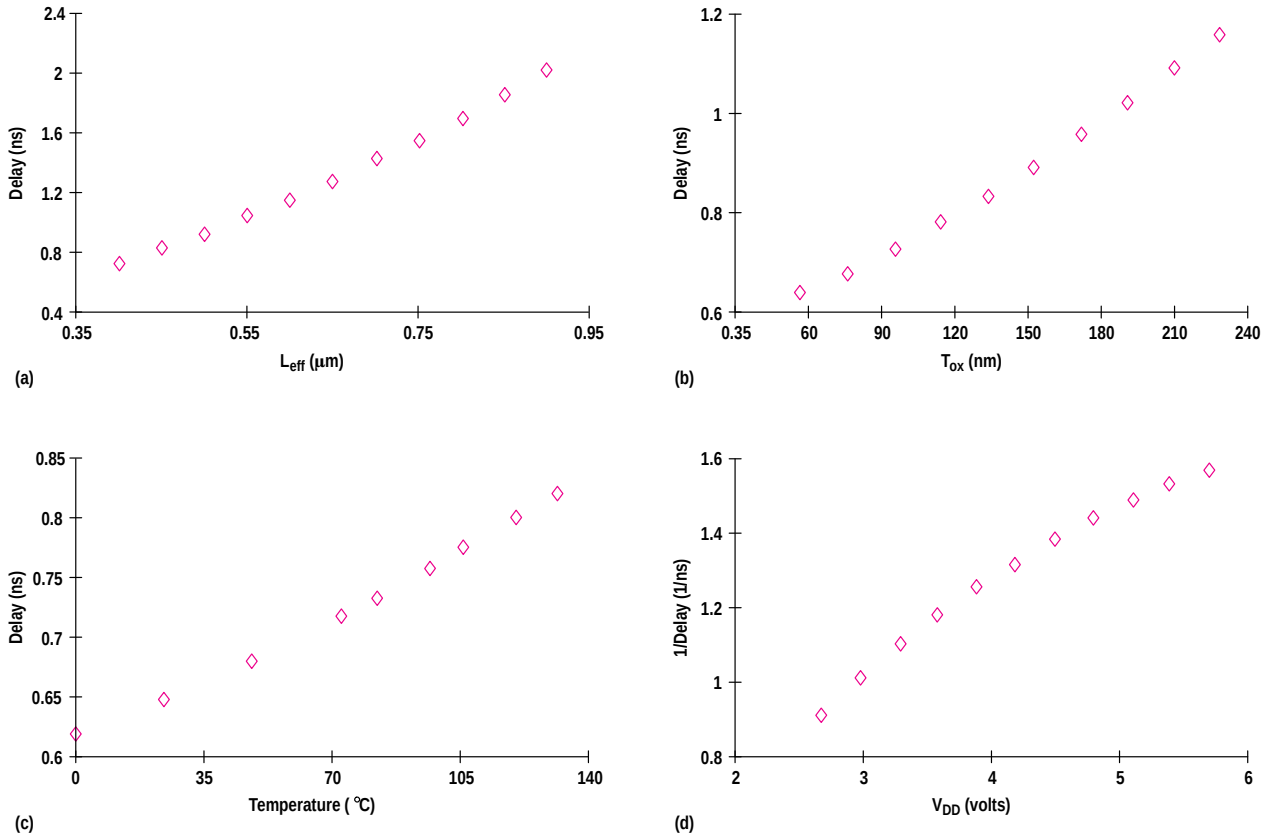

Fig. 2. SPICE plot of delay as a function of (a) L_{eff}, (b) T_{ox}, (c) temperature, and (d) V_{DD}.

optimistic. It should be noted that technology constant values are different for different benchmarks or portions of a benchmark.

From an examination of Fig. 4, some trends and patterns emerge. For instance, many suppliers tend to be more conservative for their leading-edge technology. This might result from inherent tendencies to be more conservative. It could also be that transistor performance, as measured through electron velocity saturation, will tend to level off for smaller L_{eff}.

Some suppliers clearly claimed performance that is simply unattainable given their technology description as seen in Fig. 4. As we investigated further, we found out why they appeared conservative or optimistic. For instance, supplier G used a 2-input NAND gate instead of a 4-input NAND gate, claiming that they were optimizing the critical path using a 2-input NAND gate for the critical path and a 3-input AND gate for the rest of the noncritical path. Supplier K used half the typical unit wire capacitance constant for computing internal net capacitances: 0.1 fF/µm instead of 0.2 fF/µm. On the other extreme, supplier F, while having competitive performance, clearly was capable of doing substantially better. After detailed conversations, it became apparent that their technology is immature and poorly defined, with high net

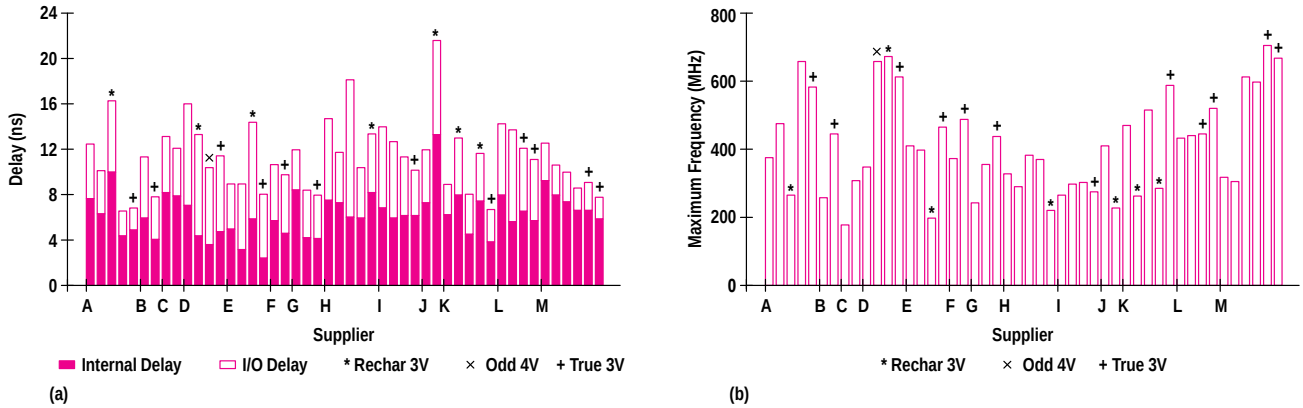

Fig. 3. Total path delay including (a) Benchmark 1 internal plus I/O delay and (b) Benchmark 2 maximum frequency for various technologies and suppliers. For each supplier, L_{eff} decreases from left to right for their respective technology offerings. Some technologies are simply recharacterizations from 5V to 3V (*) while others are true 3V (+).

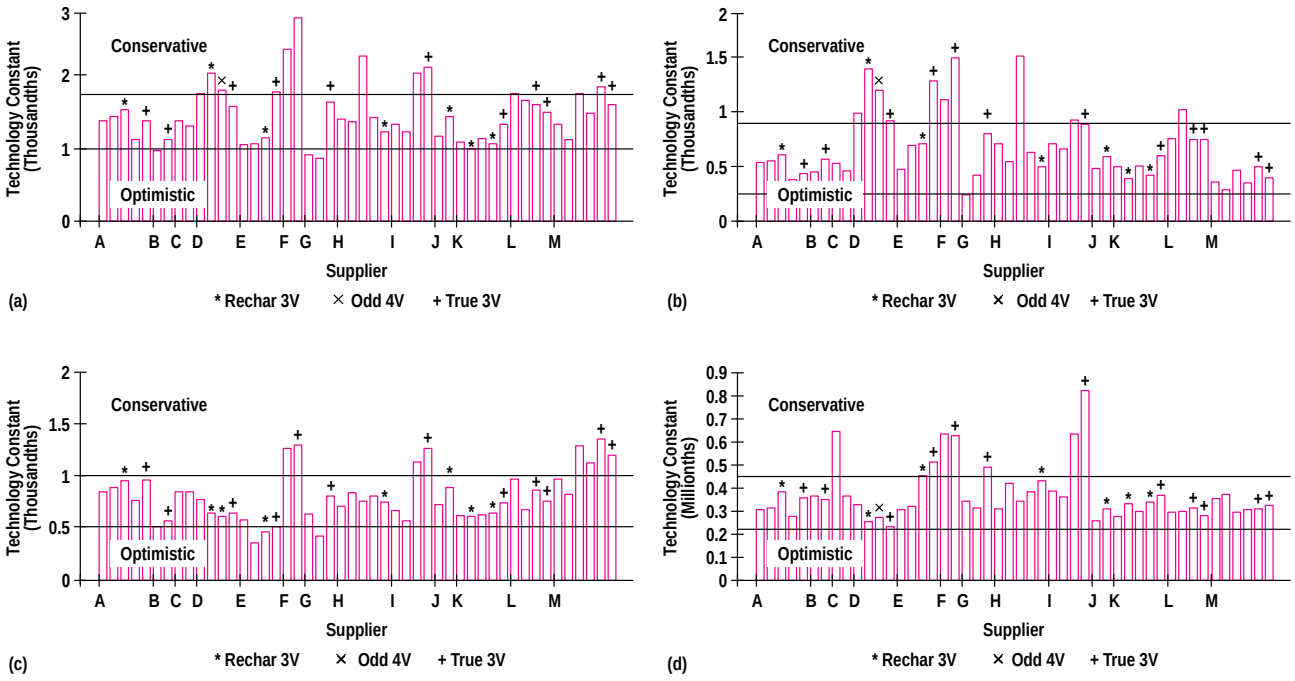

Fig. 4. Optimistic versus conservative technology constants for (a) Benchmark 1 total path delay, (b) Benchmark 1 I/O path delay, (c) Benchmark 1 internal path delay, and (d) Benchmark 2 maximum frequency for various technologies and suppliers. For each supplier, L_{eff} decreases from left to right for their respective technology offerings. Some technologies are simply recharacterizations from 5V to 3V (*) while others are true 3V (+).

capacitances computed based on previous technology offerings and less than compact layout. Comparisons through the technology constant method also allow us to determine if suppliers have potentially optimized I/O or internal cells as seen in Figs. 4b, 4c, and 4d, respectively.

Table III compares supplier K to another with very similar gate array technology, supplier H. Both suppliers offer gate array products. Supplier K used a wire capacitance constant of 0.1 fF/μm, which explains partly why they claimed higher performance. In fact, supplier K has larger metal pitches than supplier H, which should result in substantially longer net wire lengths after place and route, further increasing net wire capacitances and decreasing overall performance. Both suppliers use Cadence's Gate Ensemble for layout, so it is unlikely there would be substantial wire length differences even if they had the same metal pitches. As stated previously, supplier K was found to have overly optimistic performance claims. Comparisons through the technology constant method allowed us to home in quickly on the reasons why, chief among them being the unrealistically low wire capacitance constant of 0.1 fF/μm.

Table III
Comparison of Supplier Technologies

	Supplier	
	H	K
Effective transistor gate length(μm)	0.7	0.7
Transistor gate oxide thickness (Å)	150	150
Pitch for each metal layer (μm)	2.0,2.8,2.8	3.2,3.2,3.2
Power supply used for benchmarks (V)	4.5	4.75
Junction temperature used for benchmarks (°C)	85	85
Process condition used for benchmarks (SLOW/NOM/FAST)	SLOW	SLOW
Input edge rate used for benchmarks (ns)	1	1
Benchmark 1 complete path T_{pLH} delay (ns)	14.4	8.04
Benchmark 1 complete path T_{pHL} delay (ns)	12.2	8.72
Benchmark 1 internal path T_{pLH} delay (ns)	7.3	4.8
Benchmark 1 internal path T_{pHL} delay (ns)	7.1	4.4
Benchmark 2 maximum operating frequency (MHz)	317	460

Assumptions and Future Extensions

For the sake of expediting the process of ASIC technology benchmarking, we used a simplified approach that we can improve upon for future supplier analysis. Our detailed questions did not ask for wire unit capacitance; one supplier used wire capacitance values half those given by conventional wire capacitance statistical modeling. Neither did we include wire unit resistance effects. However, the benchmarks have relatively short wire lengths so it is unlikely that RC delays contribute significantly to total delay. In the future, we can specify gate array and standard cell height as

the unit of length. We can also include the effects of transistor mobility, in particular threshold drops (V_{to}), which become important as we scale down to lower power supply voltage levels.

We also should specify more precise input signal edge rates, power supply voltage, and junction temperature, such as 1-ns 10-to-90% rise and fall times, 10% off nominal power supply (4.5V for 5V or 3.0V for 3.3V), and 85°C junction temperature. The benchmarks left these numbers to the discretion of each supplier. Sometimes junction temperatures varied from 70°C to 125°C while power supply voltage varied from 5% to 20% off nominal. In the near future, we want to get supplier electrical and SPICE models, and ultimately verify performance through actual silicon.

Equation 3 is only a first-order approximation. Curve fit analysis shows the best fit for the SPICE simulation data is not necessarily always linear. Equation 3 might be better evaluated as equation 4 below for some situations:

$$\text{Technology Constant} \approx \frac{t_d \ln(V_{DD})}{(L_{eff})^{1.25} e^{T_{ox} \times \text{Temp}}} \quad (4)$$

We are also currently evaluating other issues such as optimal metal pitch as a function of transistor L_{eff}, and library richness. Although having smaller metal pitches is advantageous in terms of routability and increased interconnection, there is a balance between transistor on-resistance and metal wire resistance, and between routing density and process cost and manufacturability. If wire pitch is too fine, wire resistance will dominate over transistor output drive. Furthermore, finer metal pitches increase process cost and reduce yield as well as long-term reliability.

Library richness is another area considered critical by many designers. In particular, designers want a rich and complete set of library functions to allow maximum flexibility in implementing chip designs. The library must be well-modeled and compatible with various CAD tools, especially mainstream synthesis and simulation tools. This includes optimizing cell drives and functionality for synthesis.

In the future, we need to review synthesis and simulation libraries for a number of ASIC technologies using small to medium-size benchmarks. Critical path timing and gate count should be evaluated after synthesis. We need to have commonly agreed-to benchmarks to evaluate and compare all major ASIC supplier libraries. These benchmarks should cover areas such as scan insertion, error correction and detection, RAM models, and others in addition to critical path timing and area optimization. The benchmarks should not include HP proprietary information so they can be freely used with external suppliers. This allows suppliers to run evaluations using their resources. HP divisions would simply have to corroborate ASIC supplier results. The benchmarks should be available in both VHDL and Verilog hardware description languages since both are being used within the HP community.

Part of the survey asked for detailed power dissipation for various portions of Benchmark 1 and Benchmark 2. Often, simulated power dissipation numbers were significantly out of line with common sense analysis, giving us an indication of the limitations of power estimation CAD tools for some suppliers. We are investigating the area of power estimation as it relates to mainstream CAD tools and supplier methodologies.

Conclusions

We have developed a simple, first-order method for quickly determining the degree of optimism or conservatism of ASIC technology performance claims from various suppliers. We found suppliers that are not capable of delivering performance as promised because of circuit tricks played with the benchmark or too-aggressive wire capacitances (0.1 fF/μm instead of 0.2 fF/μm). We also found suppliers that may have immature, poorly defined technology and design libraries.

Our method helps designers choose appropriately characterized ASIC technology while avoiding the disastrous consequences of choosing an ASIC supplier not capable of delivering the promised performance. On the other hand, it points out inefficient or immature suppliers that may be incurring extra costs because of suboptimal utilization of circuit performance and area. The technology constant method also allows us to identify suppliers with potentially superior or optimized I/O or internal design libraries.

The method is adaptable to any number of circuit benchmarks and ASIC suppliers. However, it is only a first step in the process of evaluating ASIC technologies. There are many other factors that should be considered when selecting ASIC technologies.

Acknowledgments

The authors would like to thank Hyoung Jo Youn for contributing his CMOS26 and CMOS14 SPICE decks. Also, thanks to Robin Purohit and Noori Din for discussions and enlightenment on the issues of library evaluation through synthesis. Our heartfelt gratitude goes to Sharon Fingold, wife of one of the team members, whose keen eye and clear mind kept our ideas and writing focused. Finally, we would like to mention the names of many other dedicated and creative people who contributed directly or indirectly through their discussions and otherwise, including but not limited to: Fritz Stafford, Harold Noyes, Theresa Prenn (Tideswell), Greg Marten, Bruce Schober, Rob Mankin, Scott Trosper, Gulu Advani, Atul Goel, John Jacobi, Donovan Nickel, Steve Chalmers, Margie Evashenk, Mike Stewart, Dave Fotland, Dave Goldberg, Bob Schuchard, Todd Spencer, Jon Riesberg, David Stewart, Paul Dembiczak, David Bond, Jeff Dawkins, Oren Rubinstein, and Alex Warshawsky.

Reference

1. L. Vereen and M. Jain, "Benchmarks Define ASIC Performance," *ASIC&EDA*, September 1993.