



CHOOSING SYSTEM-ON-CHIP PROCESSES:

A TOUGH DECISION

IC-PROCESS SELECTION TODAY IS A COMPLEX, MULTIVARIABLE OPTIMIZATION PROBLEM WITH FINANCIAL, TECHNICAL, AND EMOTIONAL DIMENSIONS.

BY RON WILSON • EXECUTIVE EDITOR



An unwritten assumption of the chip-design profession is that it is always best to use the newest available process: best for your résumé, and best for the design. The most advanced process you can get will make the chip faster, lower power, and less expensive than that old “mature” process you used last year. Flaws in this reasoning have always existed, but the old rule is now breaking down on a grand scale. Far from assuming that they will use the latest and greatest, today’s design teams find that process selection has in itself become an important early step in the design flow.

The causes for this change are easy to find. Diminishing returns have set in, at least for some kinds of structures, on both performance and die area. A given block in a 65-nm process is no longer automatically smaller and faster than it was at 90 nm. Power no longer decreases monotonically with process geometry. Actual energy consumption today is a complex brew of process, library, and design choices. Seasoning this mix are inscrutable end-user behaviors and a growing list of process variations. The result is often not the pudding the designers had in mind.

So, how are design teams choosing their target processes? A number of designer managers and service providers provided some answers. Despite a variety of environments and viewpoints, some patterns have emerged.

WHEN YOU HAVE NO CHOICE

The easiest case to discuss might appear trivial: The chip has a technical requirement that dictates a particular process choice. Examples include integrated RF at frequencies greater than 10 GHz; high-sensitivity, precision analog circuits with very high dynamic range; and circuits that must operate at high voltage. “High interface-voltage levels sometimes dictate a solution,” says Bob Klosterboer, senior vice president for automotive and industrial applications at AMI Semiconductor. A specialist in mixed-signal and high-voltage ASICs, Klosterboer often sees these issues. “Sometimes, it is signal voltage, not interface level, that is the issue. Advanced processes use very low core voltages. But anything above a 10-bit dynamic range can be hard to achieve at even 1.8V.”

Designers can, in principle, address these issues and others, such as the need for large amounts of integrated nonvolatile memory, by adding modules to a standard logic process. But that approach can quickly become expensive in NRE (nonrecurring-engineering) costs, design complexity, and yield. “RF designers may ask for

metal-insulator-metal capacitors, thick upper metal layers, and triple-well RF transistors,” observes Ana Hunter, Samsung’s vice president of technology. “But management may tell them to work with what they get in the standard digital process.”

Another alternative is to use an SIP (system in package, **Figure 1**). Massively high-volume use of SIPs in cell-phone handsets has driven this technology to a level of maturity that makes it viable for even lower-volume applications, Klosterboer suggests. “Often, SIP is an alternative, but it sometimes doesn’t get the consideration it should because design teams don’t understand it well,” he says.

Yet, special cases sharply narrow the process choice, and designers can’t easily circumvent these scenarios. Two of these cases, Klosterboer points out, are high-temperature operation and extended product life—both facts of life in the automotive industry. “For example, chips that will go inside a transmission case may have to operate continuously at up to 150°C,” he says. “But, in advanced processes, the dice may be rated for only 50 to 70°C. You may find a foundry process characterized at 125°C, but it then turns out that the libraries you want to use have been characterized only to 85°C. It’s a problem.” Extended life can also be an issue, requiring the design team to ensure that its process

AT A GLANCE

▣ The most recent available process is no longer automatically the best choice.

▣ Specific needs constrain some process selections, but SIP (system-in-package) approaches can often address these needs.

▣ Finding the best process is a multidimensional optimization problem that involves lifetime costs, IP (intellectual-property) needs, design requirements, and analysis of risk.

▣ The whole team of design partners needs to be involved in the decision.

choice will still be around and running wafers in 10 years.

TOO MANY CHOICES

Beyond the cases in which process choice is mandatory lies the land of uncertainty: most chip designs in which many processes might be suitable. To settle on one, design teams go through a process of exploring, evaluating, and selecting. Hugh Durdan, vice president of marketing at eSilicon, lumps these considerations into about four categories. In rough order of priority, they are cost, IP (intellectual-property) maturity, technical requirements, and process maturity. “Each of these factors in isolation is straightforward,” Durdan says. “The difficulty is in balancing between them.”

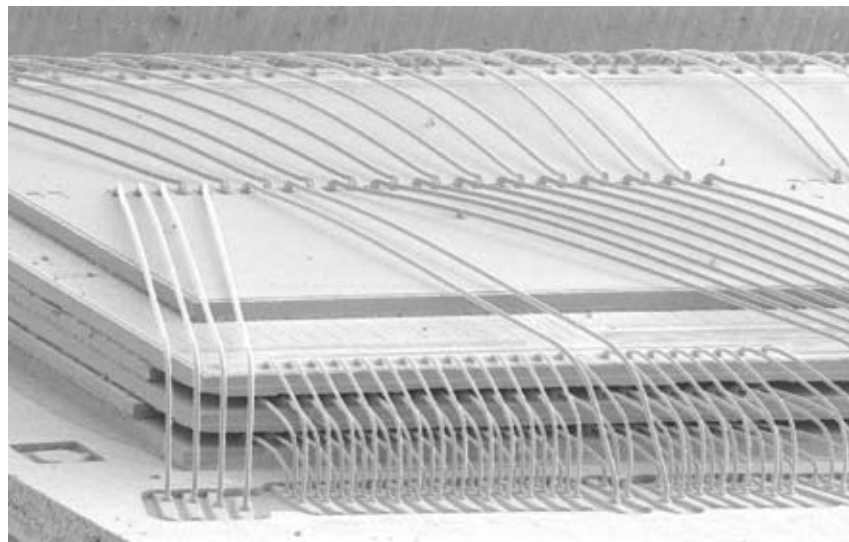


Figure 1 Complex, stacked-die systems in package have become business as usual for the packaging industry and represent an important approach to some process-selection problems for SOCs (courtesy Stats ChipPAC).

Perhaps the most obvious of these categories and, in some ways, the easiest to misunderstand, is cost. “The single factor for us is price; that pushes all the decisions,” says Jose Calero, chief technology officer of powerline-networking-chip vendor DS2 (Design of Systems on Silicon). “But we look at the cost of the full solution, not just of the silicon.” DS2 could be the poster child for consumer electronics. The company’s designs are not performance-constrained and do not have unusual technical requirements. But they are nonetheless complex SOCs (systems on chips) with formidable analog content and digital-signal processing, and they aim for high-volume markets. This fact makes unit cost more important than NRE.

Calero says that DS2 starts its cost-estimation process with a detailed knowledge of the digital and analog blocks that will go into the new design, because the new chip will usually be an incremental change from an older one. The analog-device designers can start early with a prospective process’ design kit and do their block designs through preliminary placement. The company then takes the gate counts for the digital blocks and the preliminary analog designs to prospective vendors for quotations. In general, the lowest quote wins.

COMPLEXITIES OF COST

Unit cost depends on die size. “In general, we are pretty good at estimating die size, given a good idea of the libraries the customer will use,” says Paul Rousseau, account manager on emerging accounts at TSMC (Taiwan Semiconductor Manufacturing Co). The estimation is more than a matter of counting standard cells, because I/O, power routing, decoupling capacitors for the power routing, and passive components for analog circuits can all be major factors in the final answer. But foundries or experienced design partners can often base their estimates on completed designs with similar characteristics. Experience is vital to this estimation. “We’ve been using a technology internally before we make it available to customers,” says Jonathan Stanley, senior account manager at Fujitsu. “So we often have prior experience with similar blocks, as well as internal estimation tools, to estimate die area.” And die size isn’t necessarily

the only component of unit cost: There are also yield, test costs, and packaging to factor in. Any one of these items has the potential to cost more than the die.

Another important unit-cost consideration is the opportunity to do a design reduction or a full redesign to bring down the silicon cost during a product's life. "Some people do their planning based on doing a simple die-shrink later," says Brad Paulsen, vice president of business development at TSMC. "They often use midstep processes that allow them to do a straight shrink and end up with a smaller die, rather than having to do a full redesign for a more advanced process node."

Unit cost is not the only cost consideration for most design teams, although

it is for consumer-market products, such as DS2's. "When we hear people say, 'We can't afford to do a 65-nm design,' it raises flags," says TSMC's Paulsen.

"If you expect high volumes, the question is whether you can afford not to do 65 nm," TSMC's Rousseau adds.

If the expected unit volume is not that great, other factors intrude: NRE, IP-license and -royalty costs, personnel costs, and outsourced-contract costs are a few that Paulsen names. These factors can lead to trade-offs. "You can go for a smaller die at 90 nm and pay \$1 million in tooling costs," says AMI Semiconductor's Klosterboer. "Or you can accept a larger die size at 350 nm and pay \$30,000 in tooling. Volume is very important." Further, the design team has to

decide on licensing IP, contracting with someone to design critical blocks, or doing the job in-house. Here, questions of capability and risk become part of the cost equation. "Often, we'll see startups that assume they are going to do all the design themselves," Paulsen says. "Sometimes, we have to counsel them that it just doesn't happen that way."

There is also a matter of design teams biting off more than they can chew. Unquestionably, more advanced processes are more demanding on designers (Figure 2). More design steps, more expensive tool licenses, and greater risk of internal iterations or silicon re-spins all add to potential costs. Even large companies kill some designs when they run out of budget before tapeout. So, both design

KEEPING MEMORY OFF THE CRITICAL PATH ON PROCESSOR PIPELINES

By Vidya Rajagopalan, MIPS Technologies

Level 1-cache-memory access is typically a critical timing path in high-performance microprocessor designs. This situation is true for processors using custom-circuit techniques, and it is an even bigger challenge for synthesizable processor must not only meet its frequency target using SRAMs from a variety of memory vendors, but also scale well across process generations.

To achieve frequency targets, high-performance processors employ a variety of design techniques, the first and most straightforward of which is to budget enough pipeline stages for cache access. Depending on frequency targets, you can design custom processors with custom-cache subsystems such that the memory access completes in one cycle.

High-performance synthesizable processors using off-the-shelf SRAMs must, however, allow more than one cycle for memory access. The MIPS32 24K core

family, for example, uses this technique, in which the instruction- and data-cache accesses each span two stages of the eight-stage pipeline. Tag- and data-RAM access takes place in the first stage, and the tag-compare and data-RAM access complete in the second. Because associative caches have become almost de facto standards, data-RAM-way selection takes place during the second stage, based on the tag comparison. As frequencies push higher, this technique falls short, because the second cycle is under pressure to do a lot of work. As a result, in the next generation of high-performance synthesizable processors, cache access completes over three cycles. The MIPS32 74K core uses this technique, which allocates three pipeline stages for cache access. It uses the first stage for tag-RAM access, the second stage for tag comparison, and the final stage for data-RAM-way selection. This technique provides a significant amount of flexibility in the timing for the data-RAM access, which the core can access in either the first or the second stage.

The three-cycle memory access also enables the use of a technique

that modern EDA tools make possible. This technique involves using skew to solve the data-RAM-access bottleneck. The data RAM is larger and slower than the tag RAM. Because the critical timing path is, however, usually through tag RAM and then tag comparison, the data RAM has more time to complete its access. In the three-cycle access, data-RAM access can effectively straddle the first and second cycles, allocating more than one cycle for the array access. EDA tools can automatically determine the amount of skew necessary on the clock that drives the data RAM. Alternatively, you can manually specify the skew. The ability to use skew to move the clock edge makes it possible to scale effectively across vendor memories and process generations. These techniques represent some of the factors that allow the 74K core to achieve frequencies greater than 1 GHz in a standard 65-nm process, using generic standard cells and off-the-shelf SRAMs.

AUTHOR'S BIOGRAPHY

Vidya Rajagopalan is director of engineering at MIPS Technologies.

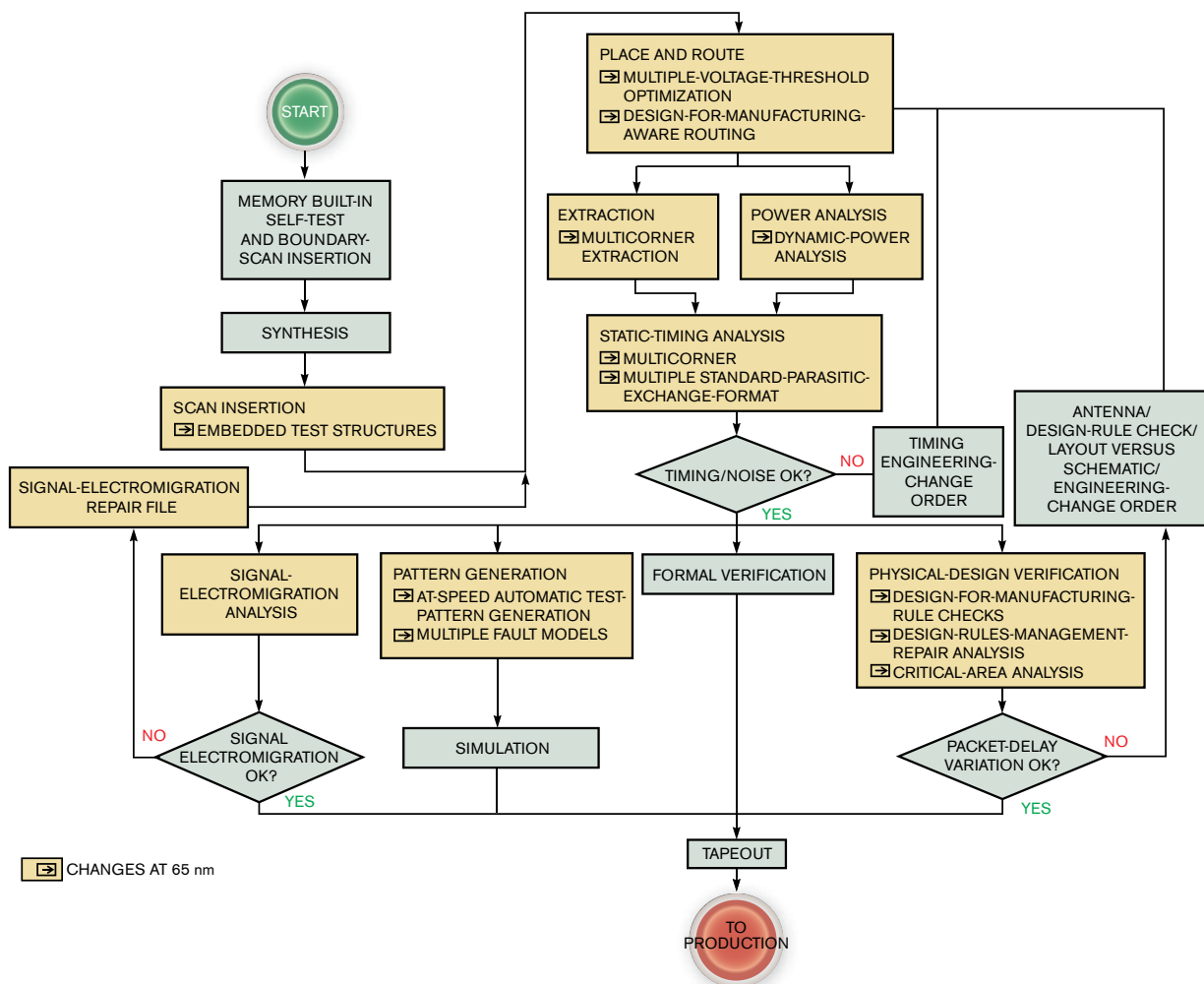


Figure 2 Moving from a 90-nm process to 65 nm adds a number of design steps (courtesy Open-Silicon).

partners and foundries try to ensure that design managers understand what they are getting into. “We train teams to do design at 65 nm,” says TSMC’s Paulsen. “We also recommend design partners, such as eSilicon, and we walk people through a tapeout sequence.”

“We have seen design managers change their minds on their process decision at this point,” Rousseau adds.

IP DILEMMAS

IP looms as the second major decision in process selection. A design team must determine what third-party IP it will need and in what processes that IP is available. “Once you’ve gone through the IP requirements for your design, you’ve pretty well made your process decision,” declares Fujitsu’s Stanley. There is good reason for this strong statement.

“I don’t believe there is any such thing as nonsilicon-proven IP,” states AMI Semiconductor’s Klosterboer. “If it’s not running in the process variant

you intend to use, it’s just a data sheet. If the IP hasn’t at least been characterized from a shuttle run, the chances of a metal spin are better than 50%.”

Not surprisingly, IP vendors see this situation somewhat differently. “It’s true that each IP core has its own characteristics in each combination of process node, voltage, and libraries,” says Gideon Intrater, vice president of solutions architecture at MIPS. But synthesizable digital IP, such as a processor core, differs from a hard analog-IP block, such as the ones Chipidea makes. In that case, most people insist on at least test chips. The same situation was once true of critical synthesizable blocks. The variations

from interactions between constraints, synthesis switches, test insertion, libraries, and design rules put too much uncertainty into the design flow for most managers. They wanted to see a CPU core in a shuttle run, in their variant, with their parameters.

But Intrater says that some IP vendors are learning to go beyond this scenario. “We are learning to make processor microarchitecture very robust at the register-transfer level,” he says (see sidebar “Keeping memory off the critical path on processor pipelines”). This approach permits customers to hit their requirements with synthesis over a wide range of process and library choices. MIPS has synthesized its 4000 core, for instance, in processes of 65 to 250 nm. “There still are some issues,” says Intrater. “For instance, core cells are starting to get faster than memory cells. In highly pipelined machines like ours, we have occasionally had to redesign SRAM cells in an advanced process so that the memory could

MORE AT EDN.COM

Go to www.edn.com/080124cs for this article’s associated vendor box or to click on Feedback Loop to post a comment on this article.

keep up with the synthesized CPU core. Another question for the future is that it appears that, in the most advanced processes, flip-flop cells are becoming slower relative to the other cells. This [situation] could become an issue.”

PERFORMANCE AND POWER

It might seem strange that I have yet to mention performance and power as decision criteria. These are increasingly design, rather than process, issues. “If you are starting a fresh design, performance is more a matter of architecture than of raw gate speed,” says Paul Little, senior engineering manager at Fujitsu. “If the architecture is constrained by a previous design, then circuit speed can become a design requirement.” Power is an even more complex issue. Starting at the 90-nm node, designers had to consider leakage power because it had become sufficiently larger than switching power. By the 65-nm node, you must understand the chip’s application before you can begin power optimization. For example, an MP3 player that shuts off when it’s not fully active differs greatly from a cell phone that is almost always

in one or another standby mode.

In the advanced processes, design techniques, rather than intrinsic energy efficiency, make the difference in power consumption. “We see people come in the door with a power budget in mind,” says Fujitsu’s Stanley. “They start out with the process characteristics, choose the grid count of their libraries based on power/performance trade-offs, and then they start adding more and more aggressive power-management techniques as necessary to meet the budget.”

“Power has become a big issue for everything, not just for mobile devices,” says Samsung’s Hunter. “It gets much more analysis now, and it influences people’s choice of process variants and libraries. For instance, we see customers coming in planning to use the general variant of the process and switching to the low-power variant when they see that, with the right techniques and libraries, they can meet their performance goals that way at lower total power.”

This scenario may require designers to use different grid counts in different blocks of the design, along with voltage islands and adaptive voltage-frequency

scaling. And these approaches, in turn, may influence which third-party-IP blocks are compatible with the design. You also face the risk that the details of the voltage changes will hopelessly bog down the design verification. Managing the islands could also prove too difficult for the synthesis tools. Such considerations might lead designers to go for an older process with inherently lower leakage.

None of the major variables in process selection is complex in itself. But finding even a local optimum on this complex, sometimes-discontinuous, multi-dimensional surface is no easy matter, even with the best estimation tools and the best, most impartial advice. The only real advantage a design manager has in the battle is that all the design partners want to succeed. **EDN**

You can reach
Executive Editor
Ron Wilson at
1-408-345-4427 and
ronald.wilson@
reedbusiness.com.

