Quantifying IP Entitlement for 14/16nm Technologies

Meeting the pressures of increased cost per transistor below 28nm won't be easy, but memory compiler IP technology holds promise.

Executive summary

The scaling benefits of Moore's Law are being seriously tested at 28nm. It is no longer a given that the cost per gate will go down at leading edge process nodes below 28nm, e.g., 20nm though 14nm. Rising design and manufacturing costs are contributing factors to this trend.

Meanwhile, the competing trend of fewer but more complex system-on-chip (SoC) designs is reducing the knowledge base of many chip design teams. The reduction in knowledge means less IP availability at leading-edge process nodes. What can be done to mitigate these challenging trends?

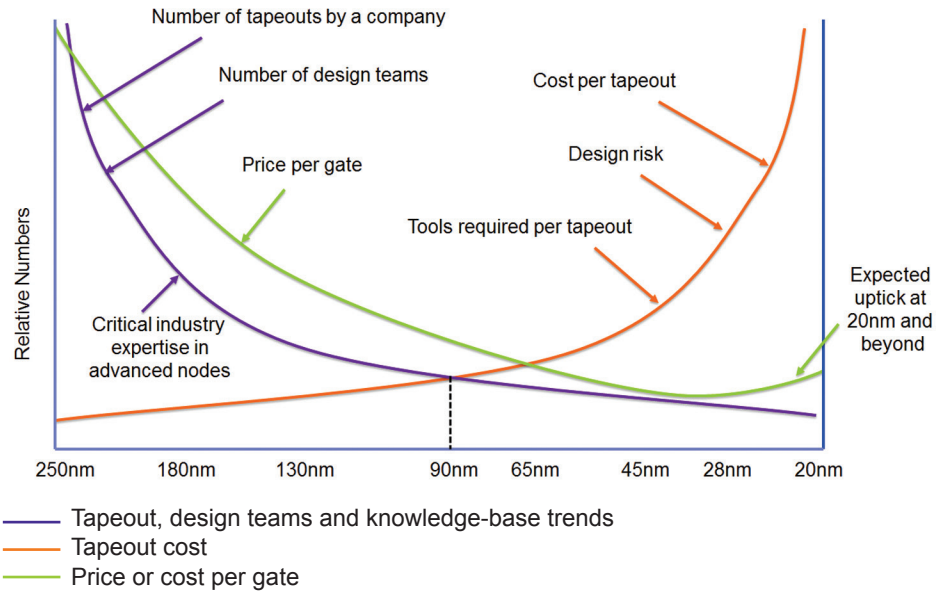
Embedded memory continues to dominate the die area of many chips regardless of the process node. This suggests that significant benefit can be achieved by customizing the memory architectures of an SoC early on in the design process as part of the overall SoC optimization for power, performance and die area.

This white paper explores these challenges and highlights custom IP memory optimization strategies as a solution at leading-edge nodes. These solutions have been implemented in a variety of market segments, including mobile tablets, multiprotocol switches, optical networks, carrier-grade Ethernet, gaming SoCs, digital TV (DTV) and medical devices.

I. Introduction

The design landscape at process nodes from 180nm to 14nm is changing in response to ongoing stresses to Moore's Law. These well-known stresses are impacting the design and manufacturing costs at each advanced node. For example, the cost per tapeout for an average chip goes up by a factor of 10[(1)]

Figure 1: Price per Gate at 20nm and Below



over the cost at the previous node. On the manufacturing side, more lithographic masks are needed at the lower nodes. On the design side, the cost of EDA tools and foundry process kits are also node dependent. Further, higher chip respin costs amplify the problem of any design or integration mistake.

The risk of design mistakes is caused in part by an ongoing decrease in design starts. Fewer tapeouts result in the need for fewer design teams, which results in a loss of the critical knowledge needed to ensure a successful tapeout at advanced technology nodes. This loss of knowledge coupled with an increase in the tool and other costs associated with a successful tapeout have led to the emergence of a new type of bathtub curve (see the purple line in Figure 1). This curve sets the stage for understanding why there is an uptick in the price per gate at 20nm and below.

Looking at the relative tapeouts and volume from 180nm through 28nm, the success of these nodes has largely been driven by the availability of semiconductor intellectual property (IP).

Something significant happened at 28nm: the cost or price per gate stopped going down. There are several reasons why this appears to be the lowest point in the historically decreasing cost per gate curve (see green curve in Figure 1). During the transition to the 28nm node, several leading semiconductor companies struggled with available supply. They couldn't ship enough of their products. Part of the problem was lower-than-expected yield. This situation illustrates how traditional yield learning methods are running out of steam, largely because of the dramatic increase in the number and complexity of design-sensitive defects and longer failure analysis cycle times.

What we see is that few companies are willing to pay more per gate at or below 14/16nm and this puts more pressure on IP providers to offer more IP at the 28nm node, which is currently seen as the optimal price per gate[2,3] (see Figure 2).

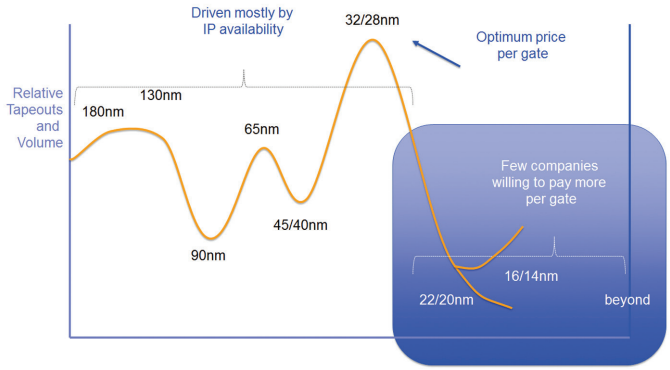


Figure 2: IP Availability vs. Technology

At advanced technologies, there are complex products that contain an ever-growing percentage of embedded memory as well as increased data and control bandwidth. There is increased complexity: IP integration with 500Mb+ of embedded memory in ASICs, gigahertz signals, package and board effects, signal integrity analysis, analog channel models, dynamic power analysis and advanced library characterization. Today's SoCs require a tight collaboration from IP through manufacturing and test.

From an IP perspective, it is ideal to work with SoC architects to define the best solution for the chip. The IP provider can develop IP that enables the chip to meet or exceed the driving market requirements. This could mean highly tailored data paths with specialized standard-cell library elements or large aggregated L3 caches to improve density, or very high-speed memory caches with speeds well above 2GHz. It could also define not-yet-available, off-the-shelf architectures such as CAMRAM (a combination of content-addressable memory and SRAM) and multi-port register files.

What about defining an extended memory IP portfolio that specifically addresses market requirements and attacks the new bathtub curve? Traditionally, we have seen that the first adopters have had the luxury of defining transistor characteristics on emerging technology nodes. They work very closely with the foundry and develop their own specialized IP to enable their market win. On the other hand, what about the second-wave customers who are also trying to come to market quickly with their new chips? Most of these companies are fabless semiconductor companies who are either developing their own ASICs or hiring an ASIC provider. Increasingly, these fabless semiconductor companies do not have their own IP development resources and are completely dependent on off-the-shelf IP providers. As the ASIC ecosystem continues to have fewer and fewer players, what happens when the stress of the implementation of the design and the new technology complexities collide, leaving a specification hole to fill?

II. Solution Strategies

A proven technique to address all of the challenges at advanced technology nodes is to utilize memory compilers, which create memory designs tailored to each instance needed. Application-driven memory compilers have been shown to increase performance or increase density while reducing die size. Further, these compilers offer the additional benefits of instance customization and Vt optimization

It is imperative to have early, silicon-proven memory architectures to use as a base for optimization of the rest of the SoC. This reduces the risk and enables a quick response for power, performance or area (PPA) optimization. The memory IP PPA window greatly expands to offer ASIC designers much more flexibility to meet or exceed their specifications with reduced design cycle time when expected technology entitlement is combined with the development of market-driven memory IP.

For example, a 16nm chip that is 9.6mm x 9.6mm with 16 cores, each having 64Kb instruction and data caches, 49.9 percent memory content (749Mb) will operate at a maximum of 1.8GHz with generally available off-the-shelf IP. With market-optimized IP, the frequency can be extended up to 2.4GHz and the overall memory area can be reduced to save up to 7mm^2 or 10 percent die area.

Figure 3: Improving Power, Performance or Area

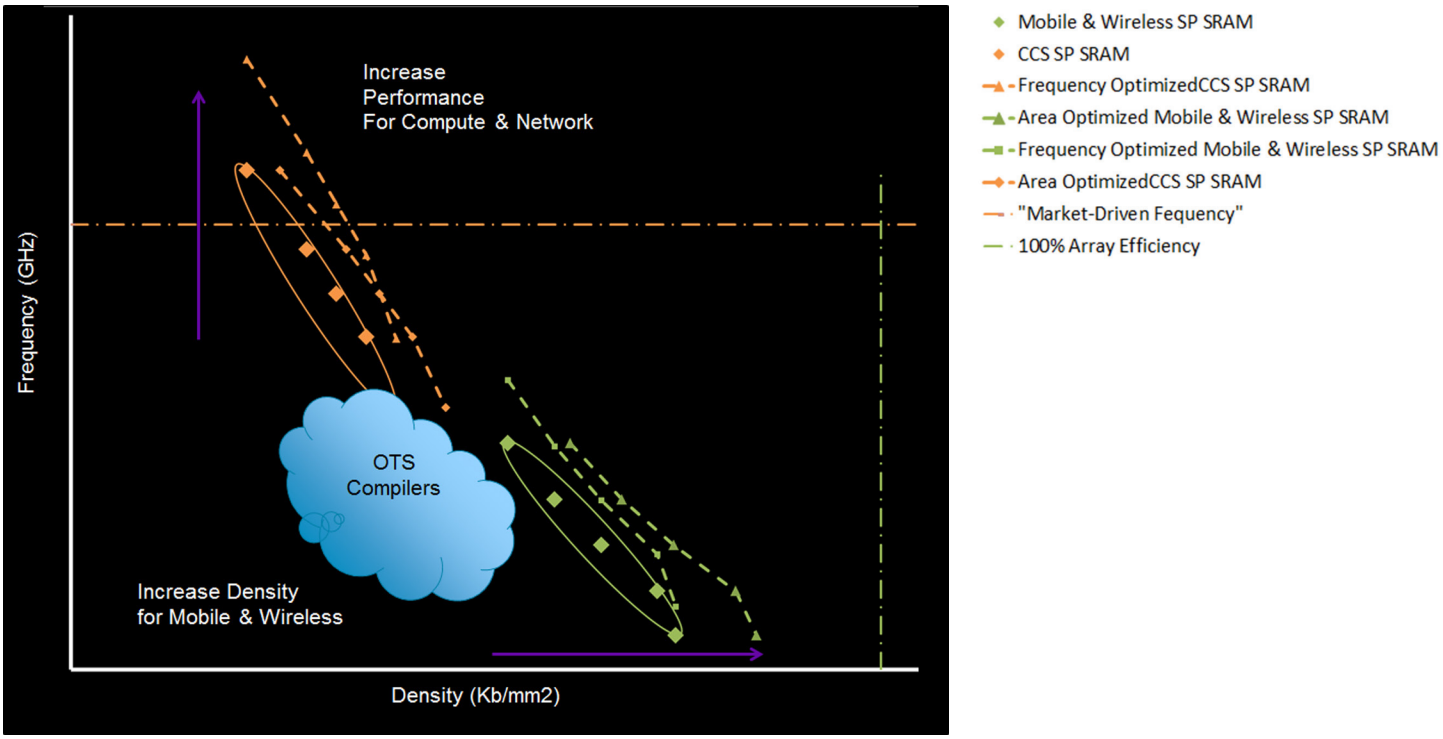


Table 1:
eSilicon Customer IP Optimization by Market Segment

Customer	Market	Tech	IP Differentiation
Customer A	Mobile Tablet	28nm	8X improvement in total power, 20X improvement idle power (custom SRAM instances)
Customer B	Multiprotocol Label Switching (MPLS) Core	28nm	25% chip power savings (CAM, UHD SRAM instances)
Customer C	Passive Optical Network (PON)	28nm	Memory subsystem: 45% improvement in memory access time and 19% area reduction (HD PDP SRAM)
Customer D	Carrier Ethernet	28nm	Memory subsystem: 24% area reduction, 65% power reduction, 62% leakage reduction using custom FIFO memories (CAMRAM, UHD SRAM, FIFO memories)
Customer E	SoC for Gaming Machines	28nm	11.6mm^2 area savings with custom SP SRAM compiler and custom large SP SRAM macro, LVDS I/O, custom cache memories for Cortex A7
Customer F	High-end Mobile Tablet	40nm	Memory subsystem: 12% area reduction
Customer G	DTV	40nm	7% die area reduction
Customer H	Medical	40nm	40% total power reduction (dynamic and static) with eSilicon-developed logic-rule bit cell

III. eSilicon Solutions

eSilicon has taped out several memory compilers on 14nm and 16nm FinFET and SOI technologies. eSilicon's 16nm ASICs are in development. We can look at 28nm and 40nm ASICs in production to extrapolate the future benefits of IP optimization. Below are examples of the benefits eSilicon customers have received by using customized memories to optimize their chips for market requirements.

eSilicon has silicon back and is currently testing its TSMC 16FF+ eFlex™ dual-port SRAM and two-port register file compilers. The eFlexCAM™ ternary CAM (TCAM) has taped out to a shuttle with silicon fabrication completion expected in November 2014. The 14nm FDSOI dual-port SRAM compiler has also taped out with silicon fabrication completion expected back in October 2014.

eSilicon also offers memory instance customization services. Our new web-based IP MarketPlace™ tool will be available for customer use in the fall of 2014. It will offer customers the ability to explore the existing PPA characteristics of our silicon-proven base compilers, as well as request custom memory instances to solve their unique problems. The IP MarketPlace tool will also enable customers to explore all eSilicon-developed IP (memory, standard cell libraries, I/Os, interface IP), as well as generate memory instances, explore memory PPA, and download views.

IV. Conclusion

The availability of IP continues to be a primary contributor to successful tapeouts at most modern process nodes. The increase of the cost per gate at 20nm through 14nm nodes is putting pressure on providers to offer more IP. Embedded memories represent an ever-growing percentage of leading-edge process chips. Chip designers, especially in fabless companies, can benefit greatly by using proven memory compiler technology to customize their application and lessen the effect of increasing costs beyond 28nm.

For more information on 14/16nm memories, custom memory or the IP MarketPlace tool, contact ipbu@esilicon.com.

V. References

1. GLOBALFOUNDRIES presentation at Common Platform 2012.

2. Solid State Technology, http://electroiq.com/blog/2014/03/moores-law-has-stopped-at-28nm, "Moore's Law has stopped at 28nm"

3. SemiWiki.com, https://www.semiwiki.com/forum/content/2768-28nm-transistors-cheapest-forever.html; "Are 28nm Transistors the Cheapest...Forever?" 9/17/2013

eSilicon Headquarters
2130 Gold Street, Suite 100
San Jose, CA 95002
www.esilicon.com
sales@esilicon.com

Phone: 1.408.635.6500
U.S. & Canada Toll Free:
1.877.7.MY-CHIP
(1.877.769.2447)