WHITE PAPER

Custom SoCs Compete with Discrete IC Boards

New data shows that the break-even, return on investment (RoI) point continues to fall as custom SoCs compete directly with traditional discrete IC board designs.

John Blyler - Editorial Director at IoT Embedded Systems

Nov 2015

This research was sponsored by ARM Limited.

ARM

Introduction

Custom System-on-Chip (SoC) semiconductors have now reached a cost point to compete directly with traditional discrete IC board products. Many original electronic manufacturing (OEM) companies of printed circuit boards are reconsidering a SoC implementation as a way to reduce the bill-of-material (BoM) costs, shrink the product size and power consumption, and add new features such as wireless connectivity and sensor processing. Further, moving to a custom SoC has several inherent advantages, such as design intellectual property (IP) protection, supply chain end-of-life (EoL) security and mitigating the slowing of Moore's law.

What kind of company should consider developing a custom SoC? A typical enterprise might be a small to medium business (SMB) with an existing board-level product consisting of several discrete integrated circuits (ICs) that provide RF, analog, baseband, processor, memory and interface functions. All of these individual chips, along with associated capacitors and resistors, typically fit onto one large motherboard and at least one daughter board (often for the wireless connectivity circuitry). The BOM costs are large but could be far less once all functionality is captured onto a single application specific integrated circuit (ASIC) – by as much as 90%, according to the S3 Group. Similarly, an 85% savings in size reduction has been realized when moving from a board to chip implementation.

For most companies, cost will be the primary but not independent factor. "Cost can be broken down into unit cost, design cost and non-recurring engineering (NRE) charges," notes Merlyn Brunken, Director of Market Research at Mentor Graphics.

Richard Wawrzyniak Senior Market Analyst for ASIC & SoC Designs at Semico Research, looks at a variety of cost factors to produce a series of curves that plot out the profit and loss for a prospective custom design (see Figure 1). "These factors – from design to manufacturing (e.g., yield, die size, etc.) and production costs – are combined with assumptions on the average selling price (ASP), market size, expected product life and others to figure out the timeframe in which a company will earn back their investment. It's a simple goal but a very complex process to figure out."

Data Points and Parameters

Die size – sq. mm
Process geometry
Number of metallization layers
Yield – percent
Die shrinks over time
Wafer processing cost – K dollars
SoC design costs – M dollars
Target market for the silicon
Market size – units
Number of parts to reach breakeven point – at various process geometries, design costs, yields, ASPs and device margin points
Estimated device margin – in percent
Target market share – marketing driven
Labor – number of person months
Capital – M dollars
Design time – months
Product life cycle – months
Device ASP – both entry ASP and ASP reduction over time
Production volumes
Competition
Profit / loss – M dollars

Figure 1: Data points to consider in the cost analysis for a custom ASIC. (Courtesy of Semico Research)

One may wonder rather pursuing an ASIC design is worth the effort. Is a single SoC really a viable alternative to board-level implementations? "It costs a lot to build a custom SoC, especially if you don't have the volume," explains Jim Bruister, President at SoC Solutions. "But no matter the process node, there is always a (proportionally) hefty NRE. If you didn't have the appropriate volume, then the NRE would make it prohibitively expensive because you couldn't amortize the cost of the die over the up-front cost. You need enough product life to recoup the investment."

One alternative to building a custom SoC is to collect a number of existing dies in a System-in-Package (SiP). Here the problem shifts to the integration of various dies – typically one for digital and another for analog functionality. Still, at least one of the die might need to be custom, which brings us back to the question of when is a custom SoC be a cost effective alternative to a board.

SoCs for Small to Medium Businesses

Discrete board level integration – traditionally done by system houses – have often used custom ASICs to replace existing boards for reasons of both cost and IP protection. Let's consider the cost issue in greater detail.

"The real tipping point occurred around 2010 to 2011 when three things happened," explained Phillip Christie, Principal Scientist at imec IC-link. "First, the price of silicon reached a point where it became cost effective to replace board level integration with a custom ASIC. Secondly, application domains were popping up which were well suited for mature technologies at the higher process nodes, for example, wireless connectivity at 180nm. Thirdly, wafer volumes were so low in the very fragmented IoT market that there was no advantage in being a large company."

Christie prefaces his remarks by noting that the IC division at European-based Imec (known simply as "imec IC-link") is a value-added silicon broker to the foundries. As such, their business model is unique, based upon about 500 integrated circuit (IC) tape-outs per year from small to medium businesses (SMBs), which lack the wafer volumes (typically 10k wafer orders) to deal directly with large semiconductor foundries. Such SMBs work with value chain aggregators that provide technical support and, most importantly, sufficient aggregate order volumes from other SMBs to ensure the multi-project wafer (MPW) runs that large foundries will accept.

Accordingly, the business trends seen by imec IC-link may be counter indicative to trends from major global integrated device manufactures (IDMs) like Intel, Samsung and others. This follows as imec IC-link customers are typically small to medium businesses in market segments that are growing in areas different from the industry as a whole.

Typically, a small company seeking to create a SoC starts with a multi-project wafer (MPW) run to get package samples, which is often a prelude for a second round of funding. If you do repeated wafer orders on an MPW run, each new wafer at 180nm costs around $2k and results in about 45 samples for a 25mm^2 die.

The cost versus volume gradient of these sets of curves for the mature nodes of 180nm, 90nm and such are very steep, notes Christie. "Although the initial non-recurring engineering (NRE) cost is around $16k at 180nm, the cost per die is around $50. So it is very expensive per unit price and a very low upfront cost."

Interestingly, small to medium businesses care much less about the unit cost than in minimizing their NRE cost. It is the later costs that they must overcome to really get their business going. Reducing the unit costs comes later once the company is more established and successful. At that time, they will transition from MPW runs to developing their own dedicated lithographic mask, which can result in a factor of 100 reduction in unit cost. The mask cost itself as another $150k, explains Christie.

The transition from the MPW segment to the essentially free silicon part of the curve for a dedicated mask are known as industrialization pathways (see Figure 2). The breakpoints occur where the foundries force small and medium customers to transition from a MPW to a dedicated mask.

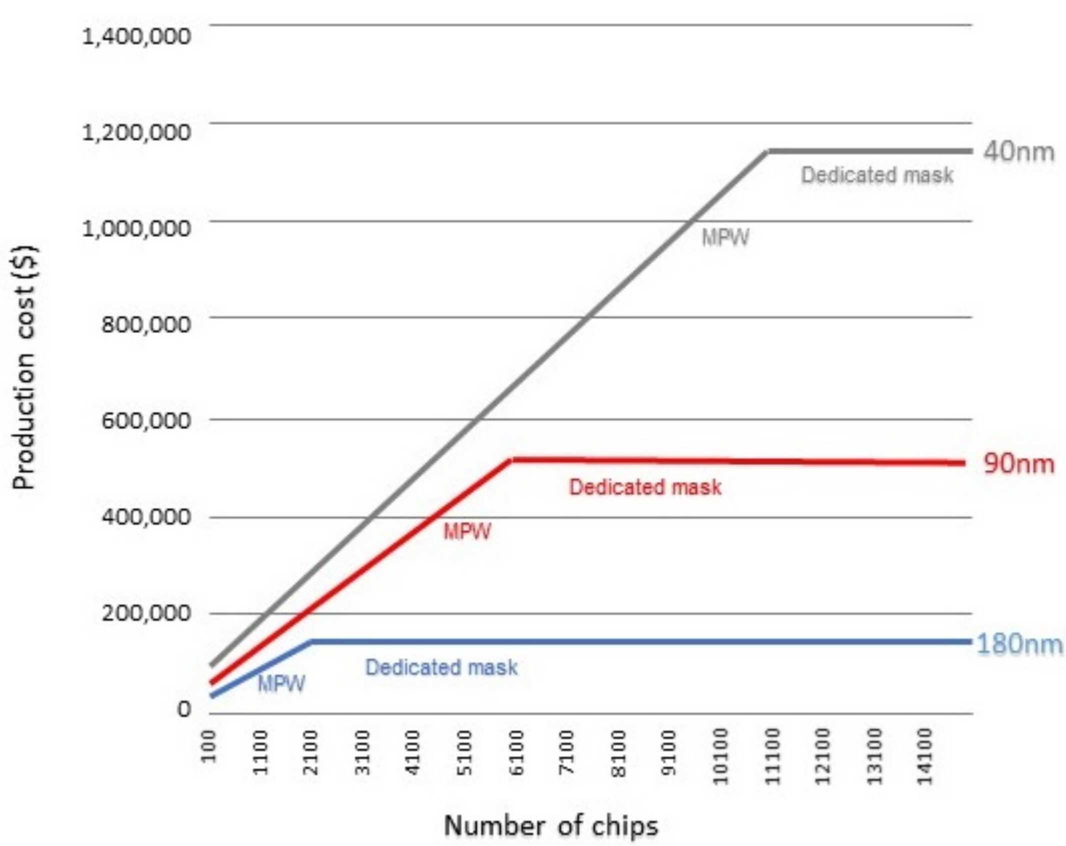

Figure 2: Industrialization pathway curves for 180nm, 90nm and (relatively lower node) 40nm, show the cost vs. volume and the breakpoints when small and medium companies must shift from MPW to a dedicated mask for limited runs. (Courtesy imec IC-link)

A dedicated mask reduces the unit cost, which has been the main driver for Moore's Law. Reducing the unit cost is why so many ASIC designs have gone from higher nodes like 180nm and 90nm to lower nodes like 28nm. But the down side of Moore's Law is that the NRE cost has been exploding. Small to medium companies can't afford custom ASIC designs at that cost.

The exception to that rule has been the growth in the fragmented range of markets known collectively as the Internet-of-Things (IoT), explained Christie. Although analysts predict that the total volume of Silicon will be high for the IoT – around 100 Billion units, the individual wafer orders are relatively low because the markets are so fragmented. Thus, large companies have no particular advantage in this market.

Another advantage of the IoT market is connectivity. The distinguishing characteristic of most IoT applications are their wireless connections to the Internet, which requires the incorporation of antenna, RF-wireless and analog subsystems. For this technology, 180nm is a very good node.

"At 180nm, various analog metrics like the intrinsic voltage gain of the transistor are very high," says Christie. At 180nm, designers still have access to a supply voltage of 3.3v, which is good for dynamic range, noise margins and creating very efficient antennas. Since analog designs don't scale with lower voltage nodes, that means that 180nm is competitive with 28nm for many analog circuits.

Further, analog circuits don't scale with transistor gate width. Designers can make the gate channel thinner but it must then be longer to compensate in an analog circuit. So core analog building blocks are not scaling with technology and thus there is no real advantage in going to 28nm. The exception is a company that needs to reduce the unit cost. But as mentioned, for IoT applications, the challenge is the non-recurring cost, not the unit costs.

All of this has led to a phenomenon that Christie calls "sub-stream innovation," in which large corporate R&D divisions will out-source innovation in areas where they are not yet making money – like IoT. Traditional OEMs are not certain how to make money in an IOT world at levels that will satisfy the return-on-investment (RoI) demands of their shareholders. That is why large companies are not entering these markets but instead out-sourcing the innovation. However, large companies are watching the startup and small to medium business markets to see what happens. They are not yet in an acquiring mode. (See, "The New Hardware Hipsters")

Rule of 50

It's difficult to get precise data to determine the break-even point which justifies moving from a collection of discrete ICs on a board to a custom SoC. The break-even point is a function of the complexity of the design, the market segment, unit volumes of the product and more. Not surprisingly, such data is often proprietary.

Perhaps most importantly, the break-even point depends upon who is computing the RoI. Is it the semiconductor companies or original equipment manufacturers (OEMs) - that is, board and component system manufacturers? For semiconductor companies, the return comes from smaller sale prices spread over many customers. Conversely, OEMs compute their return from the gain in the total cost of the system and then internalize their supplier's margins. The RoI can be significantly different, even for the same product.

So how can OEM's determine this break-even point? First, they must get a reasonable estimate for the cost of a custom SoC. "We use the "**Rule of 50,"** explains Donnacha O'Riordan, Director of Services Strategy for the S3 Group. "If you have a product where $50 in BOM costs can be saved by integrating components off the board, and if you are shipping 50k units of that product per year, then we can justify a custom ASIC. Obviously, there are nuances, but a basic working number is $2.5 million."

This NRE cost is an approximation that covers going from a blank sheet of paper to a tested, package IC from which they can build prototype boards. While the NRE cost is still high – although an order of magnitude less than for a leading-edge node design – there is a simple financial justification that eases the sticker shock. The RoI break event time is typically between 12 to 18 months (see Figure 3). O'Riordan notes that, in general, most company chief financial officers (CFOs) will sign-off the expenditure with that break-even timeframe.

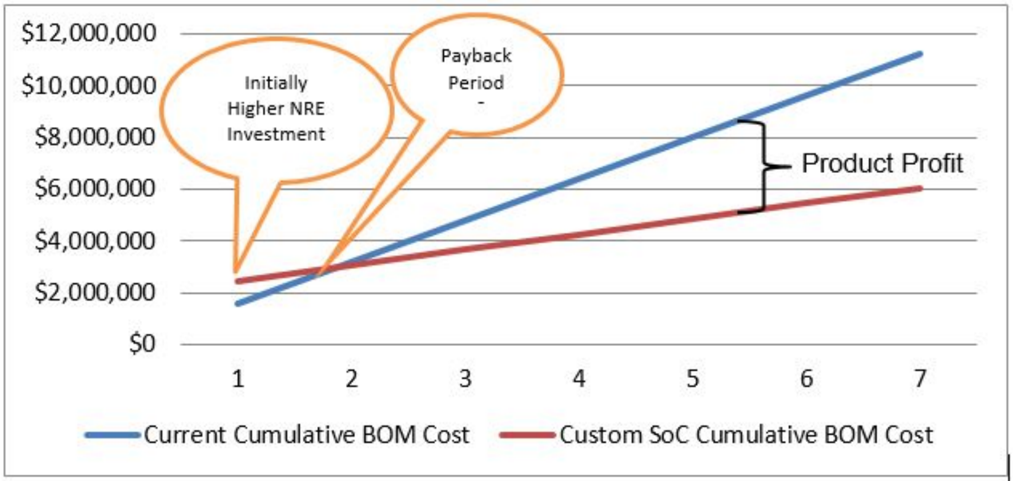

Figure 3: There is a cross-over point of 12 to 18 months for profit realization when moving from board to a custom ASIC. Note that the time-line axis is years. (Courtesy of S3)

Naturally, all of these numbers depend upon the project complexity, application and product type. Often, customers will want to add new features that may further affect the cost. The main point is that the cost and volume needed for a custom SoC at mature higher nodes is far less than for leading edge, high-volume lower nodes as found in the cell phone or similar markets.

IP Protection, Supply Chain Security, and Moore's Law Slowing

Once the cost barrier has been breached, other advantages inherent in ASIC implementation become available. These include intellectual property (IP) protection, supply security and safeguarding from the slowing of Moore's Law. Each of these advantages will be assigned different priorities, depending upon the company, application and market space.

IP protection is of particular concern for small to medium businesses (SMBs) and a significant advantage offered by even a low cost ASIC. Traditional IP protection methods like owning a patent are mostly irrelevant for small companies. While a patent gives one the right to sue for infringement, most SMBs don't have the financial resources needed to successfully conduct such a lawsuit.

The protection offered by ASICs result from the difficulty in reverse engineering the design. By comparison, a design build around a software programmable FPGA is relatively easy to extract or reverse engineer. Circuit boards with discrete ICs are even easier than an FPGA to clone.

The low cost of ASICs at mature nodes means that more companies can protect their IP in hardware. That's one reason why Jio Ito, the director of the media lab at MIT, calls hardware the new software. Today's hardware startups look like yesterday's software counterparts. That is why the spin-offs coming out of MIT are actually working in hardware rather than software - because of the falling costs of ASICs and the afforded IP protection.

As an example of this trend, consider a company that makes very advanced carpet making machines. This company was losing revenue to low cost knock-offs coming from across the globe. As soon as they released a new machine, it would be cloned at a much lower cost. For one of their recent product updates, this carpet machine company moved to a customer ASIC instead of the traditional discrete component board. To their satisfaction, that product has been the only machine that wasn't cloned.

The second advantage offered by ASICs is insulation from the obsolescence of discrete ICs. This security of the supply chain means that companies don't have to worry about end-of-life (EoL) scenarios for board components. When a supplier decides to stop supplying a particular component, the customer may have to redesign their product board. According to O'Riordan, the EoL scenario is a significant trigger for many companies when considering a custom SoC.

The magnitude of the EOL challenge is multiplied by the number of chips on the affected board, which makes it a major problem for small and medium businesses. Further, if such companies have completely subcontracted their production activity, then they may not receive advanced notification of upcoming EoL shortages from the subcontractor. The result can be production stops at the worst possible time, for example, when a new batch of products are needed because the company's inventory is low.

The capability to control the lifespan of vital components is often the number one criteria for companies that operate in markets with long product life cycles, such as rail transportation, aerospace, military and the like.

The third advantage comes from the slowing down of Moore's law. For past generations, semiconductor chip suppliers have enjoyed the 18-month cycle of benefits that came from Moore's Law – that is, chips became twice as powerful at half the power and cost. Today, OEM's building products with standard components from chip suppliers can no longer rely on the cyclic benefits of Moore's Law. Instead, the OEM's have to rethink the entire product architecture to achieve these benefits. During architectural rethinking is the perfect time for these companies to explore the possibilities of developing a custom ASIC.

Conclusion

There are many factors to consider when moving from a board with discrete ICs to a much smaller board that contains a single SoC. In general, many of the fastest growing markets like IoT can be built on the mature higher process nodes of 180nm, 90nm and even 65nm.

At 180nm, mask costs are below $200k. Multi-Project Wafer (MPW) runs for a 25mm^2 die are $16k and yield around 45 samples. The NRE costs for a designed, tested and fab-ready SoC are relatively high - typically at $2.5M – in comparison to unit costs. But those NRE costs reflect a BOM savings of only $50 and a unit volume of only 50k! Further, the break-even timeframe for profitability is relatively short – between 12 to 18 months.

The barrier for custom ASIC design is now within the reach of even startups, which is why several technology spin-offs from MIT are working in hardware instead of software. But the much lower cost of ASIC hardware isn't the only reason.

Custom ASICs afford design IP protection, safeguards against supply chain obsolesce and provide a way to deal with the power-performance-area effects of the slowing of Moore's law.